PhD TESIS

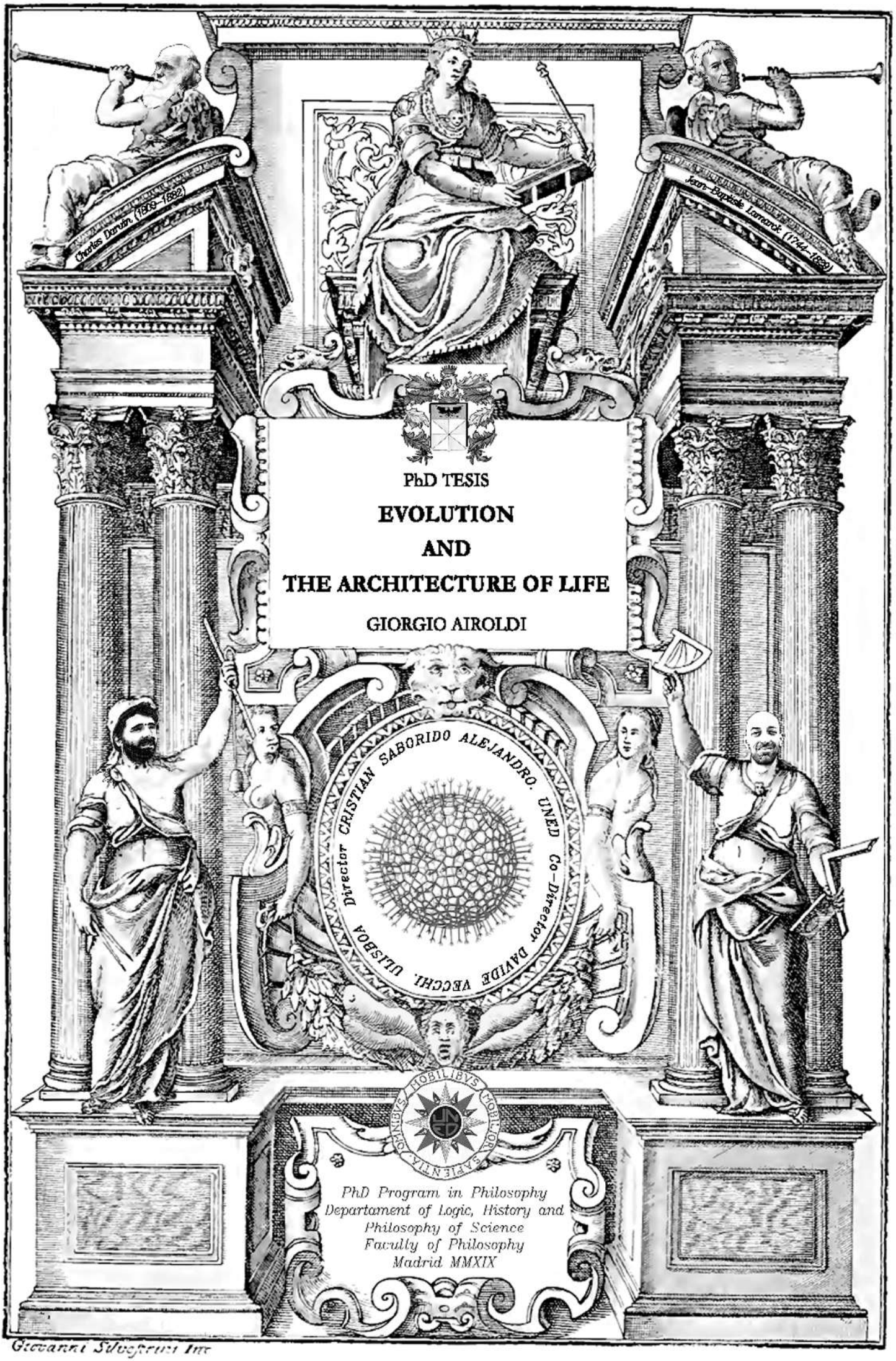
2019

EVOLUTION
AND
THE ARCHITECTURE OF LIFE

GIORGIO AIROLDI

Director: CRISTIAN SABORIDO ALEJANDRO, UNED
Co-Director: DAVIDE VECCHI, Universidade de Lisboa

PhD Program in Philosophy
Department of Logic, History and Philosophy of Science
Faculty of Philosophy
UNED



PhD TESIS

EVOLUTION AND THE ARCHITECTURE OF LIFE

GIORGIO AIROLDI

PhD Program in Philosophy
Departament of Logic, History and Philosophy of Science
Facully of Philosophy
Madrid MMXIX

Giovanni Silvestrini inc

Acknowledgements

I would like to thank Cristian Saborido for his support, both academic and moral!, during the years dedicated to my research and to the writing of this thesis; Davide Vecchi for accepting to be co-director and for providing countless advices and insightful suggestions; and David Teira, who was so kind to read the final draft of this work and to propose some brilliant improvements to its structure and contents.

I would also like to thank Víctor Luque, Javier Suarez and all the many students and professors I met during these years in congresses and seminars, and who shared with me their knowledge and researches, which contributed greatly to shaping the ideas behind this work.

I am grateful to all members and staff of the department of logic, history and philosophy of science of UNED in Madrid for providing a wonderful working and studying environment, and especially to Clara Bueno for revealing me unknown facts about Primates and Henry Potter.

A special thanks to Ana Peel for the reinterpretation of the original frontispiece from *I Quattro Libri dell'Architettura* of Andrea Palladio.

Contents

1. INTRODUCTION 3

2. NATURAL SELECTION AND THE VARIETY AND COMPLEXITY OF THE LIVING WORLD 15

2.1. DARWIN'S VIEW: SELECTION AS *EXPLANANS* OF BOTH VARIETY AND COMPLEXITY 16
2.2. NATURAL SELECTION AND VARIETY 19
2.2.1. Population genetics or: why are some phenotypic traits more frequent than others? 20
2.2.2. Foundational assumptions of population genetics 22
2.2.3. Population genetics models 24
2.3. NATURAL SELECTION AND PHENOTYPIC COMPLEXITY 29
2.3.1. What is the nature of the link between natural selection and biological complexity? 30
2.3.2. Two lines of research to explain phenotypic complexity: formal adaptationism and pluralism 37
2.4. A PROPOSAL TO FORMALISE BIOLOGICAL COMPLEXITY 39
2.4.1. What is complexity? Polysemy of the concept. 39
2.4.2. Biological complexity: organismal architecture as form and function. 46
2.4.3. A model to classify evolutionary complexity changes 53
2.5. BUT… DOES BIOLOGICAL COMPLEXITY INCREASE IN THE HISTORY OF LIFE? 60
2.6. CONCLUSIONS 65

3. ADAPTATIONISM 67

3.1. KINDS OF ADAPTATIONISM 67
3.2. THE FOUR PILLARS OF EMPIRICAL ADAPTATIONISM 69
3.2.1. Pillar I - Ubiquity of adaptation 69
3.2.2. Pillar II - Continuous, slow and incremental evolutionary change 70
3.2.3. Pillar III - Random and unconstrained variation 71
3.2.4. Pillar IV - Supremacy of selection 72
3.3. WAS DARWIN A DARWINIST? 74
3.3.1. The creative power of selection 74
3.3.2. Optimization ability of selection 78
3.3.3. Constraints and selection 80
3.4. FUNCTIONALISM AND EXTERNALISM 81
3.5. OPTIMIZATION PROGRAMS. 84
3.6. THE FORMAL DARWINISM PROJECT BY ALAN GRAFEN. 92
3.6.1. Two models, one equilibrium 93
3.6.2. A controversial project 100
3.7. THE FITNESS MATRIX 109
3.8. CONCLUSIONS 111

4. PLURALISM: BEYOND THE LIMITATION OF ADAPTATIONISM 115

4.1. CRITIQUES TO ADAPTATIONISM 116
4.1.1. What natural selection leaves uncovered 116
4.1.2. Pluralism and the pillars of adaptationism 117
4.1.3. An instance of a pluralist program: the Extended Evolutionary Synthesis 124
4.1.4. Structuralism and Internalism 126
4.1.5. The source of innovation as the basis of our classification of non-adaptationist evolutionary accounts 129

4.2. GENETIC ACCOUNTS 131
4.2.1. Classical population genetics' processes: mutation, recombination, and drift *131*
4.2.2. Wright's Shifting Balance theory and the metaphor of adaptive landscape *133*
4.2.3. Punctuated Equilibrium *139*
4.2.4. Kimura's neutral theory of molecular evolution *143*
4.2.5. Andreas Wagner's Genetic Networks *147*
4.2.6. Molecular Drive *156*
4.2.7. Final comments on the genetic sources of variation *156*
4.3. PHENOTYPIC ACCOUNTS 157
4.3.1. Exaptations *157*
4.3.2. Phenotypic plasticity and genetic assimilation *159*
4.3.3. Neo-Lamarckian accounts: induced mutations *167*
4.3.4. Epigenetics *174*
4.3.5. Final comments on the phenotypic sources of variations *178*
4.4. DEVELOPMENTAL ACCOUNTS 179
4.4.1. Evo-Devo *179*
4.4.2. Developmental constraints *182*
4.4.3. Developmental modularity and evolvability *185*
4.4.4. Final comments on developmental accounts of variation *187*
4.5. SYSTEMIC ACCOUNTS 188
4.5.1. Natural laws *188*
4.5.2. Evolution as a thermodynamic process *199*
4.5.3. The laws of complexity *203*
4.5.4. Zero Force Evolutionary Law *211*
4.5.5. Final comments on the systemic accounts of variations *217*
4.6. ARE PLURALISM AND ADAPTATIONISM COMPATIBLE? 217
4.7. CONCLUSIONS 221

5. EVOLUTION BEYOND SELECTION: A LOGICAL SPACE FOR ARCHITECTURAL CHANGES 225

5.1. ADAPTATIONISM, PLURALISM AND THE ROLE OF FITNESS 225
5.1.1. Evolutionary phenomena and architectural changes *225*
5.1.2. Stepping beyond the imperialism of fitness *228*
5.2. ROBUSTNESS AS SECOND DIMENSION OF ORGANIC DESIGN 235
5.2.1. Evolution as a two steps process *235*
5.2.2. Definition of robustness *241*
5.2.3. The two aspects of robustness *246*
5.2.4. Mechanisms behind robustness *249*
5.2.5. Types of robustness *252*
5.3. DESIGN SPACE 257
5.3.1. Dimensions and elements *257*
5.3.2. Dynamics of evolutionary movements in design space *261*
5.3.3. The design space and Fisher's fundamental theorem: the thermodynamic metaphor extended *278*
5.3.4. The design space and fitness landscapes: architectures and attractors. *282*
5.3.5. A classification of evolutionary phenomena *286*
5.4. CONCLUSIONS 293

6. FINAL THOUGHTS 295

7. REFERENCES 301

To A. A.

"et mutam nequiquam alloquerer cinerem"

Plate 3 from *Atlas der Natuurlijke Geschiedenis, voor scholen en huisgezinnen*, Amsterdam: J. C. Loman Jr, ca. 1865

1. Introduction

We humans show a natural tendency to knowledge, as Aristotle underlines in the opening sentence of the *Metaphysics*: "πάντες ἄνθρωποι τοῦ εἰδέναι ὀρέγονται φύσει."[1] (Book I, 980a-1). Our hunger for explanation finds its most immediate stimulus in the astonishing wonders of the natural world. Among them, we somehow instinctively distinguish between inanimate objects and living beings, and, fascinated by the latter, we strive to understand their origin. Our efforts, though, are always directed and limited by the cultural and social environment we belong to, as Hegel reminds us in the *Lectures on the Philosophy of History*: "no man can surpass his own time, for the spirit of his time [der Geist seiner Zeit] is also his own spirit" (Hegel 1986, 19:111). Let us see how the 'spirit of the time' has influenced through the centuries some of the mayor answers to the questions about life.

Aristotle was the first to study living beings in a systematic way. In his main works on biology (*Parts of Animals, Generation of Animals* and *History of Animals*), he analyses and compares species and their traits, making hypotheses on why organisms possess certain features, and applying his metaphysical system to the study of animals. Trapped within the concept of a circular, eternal time typical of Greek thought, the biological world is for Aristotle a fixed reality interpretable in terms of causes and degrees of perfection, and not something that came to be and could transform into something different in the future[2]. For example, he justifies the existence of two sexes by stating that each plays a different causal role (formal cause the male, and material cause the female), and that a way of reproduction that maintains the causes separated is more perfect[3]. This approach limits the development of his intuition that, although in many cases anticipate current biological themes such as functionalism[4], reductionism[5], and even adaptedness, stop on the verge of being two

[1] All men by nature desire to know.

[2] We refer here to the idea that biological species are fixed and cannot appear, change or disappear.

[3] "The nature of the first cause of change [*the male*] (to which belong the order and the form), being better and more divine than the matter [*the female*], *it is better that the superior one be a separate individual from the inferior one*. This is why in those in which it is possible and in respect to which it is possible, *the male has been separated from the female*." (Generation of Animals, II.1, 732a4-9, transl. Loeb, italics mine and comments in brackets)

[4] E.g. when he justifies parts of animals because of their functions, such as the functions of blood (Parts of Animals II.5, 651a10-18, transl. Loeb); or when he claims that male and female are to be

millennia ahead of his time, as it appears in the following passage, in which evolution by selection of advantageous variations is *almost* stated, but finally neglected:

> What applies to the blood applies as well to the other uniform parts and also to the non-uniform parts[6]; similar variations occur. And it must be supposed that *these variations* have some reference either to the activities of the creatures and to their essential nature, or else *bring them some advantage or disadvantage*. (Aristotle, Parts of Animals, II.2, 648a14-17, transl. Loeb, italics mine)

The 'spirit of the time' has always been an obstacle on the way to devise explanations around life's origin, variety and complexity. In the Christian world –and Darwin is of course included-, God has been the preferred *explanans*, although the nature of the answer changed over the centuries. If the Bible states:

> And God said, -Let the land produce living creatures according to their kinds: the livestock, the creatures that move along the ground, and the wild animals, each according to its kind.- And it was so. (Genesis 1: 24),

in the scholastic age, Saint Thomas strived to ground the biblical narrative on rational proofs. Instead of starting from the unquestionable existence of God to explain the existence of the natural world, he inverted the direction of the argument and started form the undeniable existence of the natural world to logically proof the existence of God:

> The fifth way is taken from the governance of the world. We see that things which lack intelligence [*including animals and plants*[7]], such as natural bodies, act for an end, and this is evident from their acting always, or nearly always, in the same way, so as to obtain the best result. Hence it is plain that not fortuitously, but designedly, do they achieve their end. Now whatever lacks intelligence cannot move towards an end, unless it be directed by some being

---

distinguished not for their parts, but for their functions (Generation of Animals I.1, 716a18ff; IV.1, 764636-765a1, transl. Loeb).

[5] E.g. when he denies the idea from Empedocles (Diels, *Fragmente*, 31B97) that many characters of animals are due to accidents in the process of their formation, claiming that any trait must already be in the 'seed' producing the animal (Parts of Animals I.1, 640a19-29, transl. Loeb).

[6] The concepts of uniform and non-uniform parts in Aristotle correspond roughly to tissues and organs.

[7] Saint Thomas included animals within this category, see e.g. the following passage: "Dumb animals and plants are devoid of the life of reason whereby to set themselves in motion; they are moved, as it were by another, by a kind of natural impulse, a sign of which is that they are naturally enslaved and accommodated to the uses of others." (Summa Theologiae, second part of the Second Part, Question 64, Article 1. Reply to objection 2).

> endowed with knowledge and intelligence; as the arrow is shot to its mark by the archer. Therefore some intelligent being exists by whom all natural things are directed to their end; and this being we call God. (Thomas Aquinas, Summa Theologiae, First Part, Question 2, Article 3, italics mine)

The view of the world radically changed as a consequence of the scientific revolution of the XVII century: the world is not anymore a creation in which all beings have a defined place and role, and are interlinked through harmonic relationships such as the elements of a building designed by God, the great architect[8] (Fig. 1.1), but a machine working according to precise laws.

Fig. 1.1 - Above: God as great architect of the universe (Royal MS 19 D III, f. 3r (1411), Harley MS 334, f. 34v (15th century), Royal MS 1 E VII, f. 1v (c. 1000)). Below left: man as a micro-cosmos linked to the four elements and to the planets[9] (BL MS Sloane 282, f. 18 (15th century). Below right: ascent of the Soul (Paris, Bibliothèque nationale MS 3236A, f. 90r)

[8] "God, Who is the first principle of all things, may be compared to things created as the architect is to things designed (ut artifex ad artificiata)." (Summa Theologiae, First Part, Question 27, Article 1, Reply to objection 3)

[9] As an example of the relationship between macro- and micro-cosmos, see Manilius, Book 4 of *Astronomicon*:

The universe becomes a great mechanism at all levels, from astronomical systems to animals, to man, and man is different in grade, but not in quality, from the other living beings, as La Mettrie, a bit contemptuously, reminds us in *L'Homme Machine* in 1747:

> […] however much those haughty, vain beings - who are more distinguished by their pride than by the name of men - may wish to exalt themselves, they are basically only animals and vertically crawling machines. (La Mettrie, in Thomson 1996: 35)

The flourishing of automata replicating human and animal functions testifies to this mechanistic enthusiasm (Fig. 1.2).

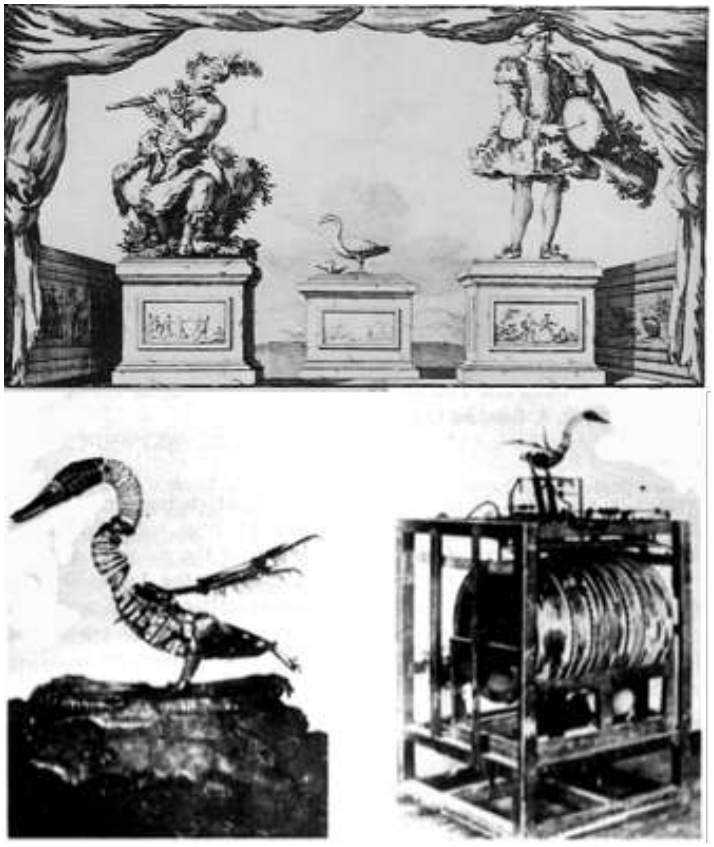


Fig. 1.2 – Above: drawing from a leaflet of Vaucanson's time of three of his most famous automata: the flute player, the mechanical duck and the tambourine player. Below: two XIX century pictures of the now lost mechanical duck, supposed to replicate all the living animal functions, including digestion (from Chapuis & Droz 1949).

Namque Aries capiti, Taurus cervicibus hæret;
Brachia sub Geminis censentur, pectora Cancro;
Te, scapulæ, Nemæe, vocant, teque ilia, Virgo;
Libra colit clunes, et Scorpius inguine regnat;
Et femur Arcitenens, genua et Capricornus amavit;
Cruraque defendit Juvenis, vestigia, Pisces.

This mechanist approach to organisms merged with the Thomistic attitude and became the ground for one of the most famous and influencing arguments in favour of creationism. At the beginning of the XIX century, William Paley drew his famous parallelism between the creative act of God and the production of a watchmaker:

> There cannot be design without a designer; contrivance without a contriver; order without choice; arrangement, without any thing capable of arranging; subserviency and relation to a purpose, without that which could intend a purpose; means suitable to an end, and executing their office in accomplishing that end, without the end ever having been contemplated, or the means accommodated to it. Arrangement, disposition of parts, subserviency of means to an end, relation of instruments to an use, imply the presence of intelligence and mind. (Paley 2008: 12)

Not even a brilliant creationist such as Paley could escape the spirit of his time: the world is a dynamic mechanism, not a static building, so God must be its mechanic, and not its architect.

Paley also wittily observed that the excellency of animals resides not much in their perfect constitution, but rather in the coherence of their parts towards the achievement of ends: something that chance cannot explain. This is what adaptedness consists of, and to this problem the theory of natural selection provides a revolutionary answer:

> Owing to this struggle [for life], variations, however slight and from whatever cause proceeding, if they be in any degree profitable to the individuals of a species, in their infinitely complex relations to other organic beings and to their physical conditions of life, will tend to the preservation of such individuals, and will generally be inherited by the offspring. The offspring, also, will thus have a better chance of surviving, for, of the many individuals of any species which are periodically born, but a small number can survive. I have called this principle, by which each slight variation, if useful, is preserved, by the term natural selection. (Darwin 1872: 76)

Darwin's idea changes the paradigm: he opens a third way between God and pure chance. Design is an optical illusion not to be explained, but to be dissolved. Instead of springing from the mind of the divinity through a voluntary act of creation, it derives from the unfolding of a potentiality already present in the environment, actualised through harsh competition among slightly different versions of the same basic individual. We are

astonished to learn that someone has guessed twenty times in a row the result of the flipping of a coin: we consider it so improbable, that we are led to believe that the subject possesses some prodigious power. If we observe, on the contrary, an experiment involving a thousand people trying to guess the result of twenty consecutive tossing, such that the persons that make an incorrect guess are eliminated, and the ones that guess correctly pass to the following round, the fact that one person out of a thousand guesses correctly twenty times in a row does not seem so prodigious anymore. In the same way, biological forms and functions appear to be a prodigious part of the natural world only if we do not analyse how they have arisen (by chance) and how they have been retained (through selection).

Darwin made the intervention of the divinity an 'unnecessary hypothesis', such as Laplace had done in the realm of astronomy at the beginning of the century[10]. He did not escape, though, the spirit of *his* time. In what sense? Because the environment takes the place of god, moreover as a bellicose deputy. A prominent member of the *intelligentsia* of the most extended empire of his time, that had reached the pinnacle of its power, Darwin believed that

> […] from the war of nature, from famine and death, the most exalted object which we are capable of conceiving, namely, the production of the higher animals, directly follows. (Darwin 1872: 554)

which, apart from a biological explanation, sounds like a political justification: not only the fittest wins, but its victory is good because it makes its environment, and the nature at large, a better and more perfect place . War is, in the end, the only way of progression:

> […] for some Victorians, at least, the philosophy of 'social Darwinism' legitimated the replacement of traditional values with the worship of brute force and success at any price. […] Evolutionism succeeded precisely because it could be adapted to many different social purposes. (Bowler 1992: 129-130)

Note that a very different view of the progression of life underlines the work, for example, of Kropotkin. In his book *The Mutual Aid*[11], he writes that

---

[10] Asked by Napoleon why he did not mention of God in his account of the solar system dynamics, Laplace is believed to have answered that he had not needed the hypothesis (firstly proposed by Newton) of God's intervention to counteract the secular perturbation that would end up destroying the system. There are many versions of this story, see for example (Rouse Ball 1960).

[11] *The Mutual Aid* is a collection of articles that Kropotkin wrote (in English) as an answer to Huxley's essay 'The Struggle for Existence in Human Society' (reprinted as an appendix in Kropotkin 2006: 742-819).

> There is an immense amount of warfare and extermination going on amidst various species; there is, at the same time, as much, or perhaps even more, of mutual support, mutual aid, and mutual defence […] Sociability is as much a law of nature as mutual struggle. (Kropotkin 2006: 58-59)

Kropotkin distinguishes the struggle among individuals of the same population for the access to limited resources, which fosters competition, from the struggle between a group of individual and the environment, which requires cooperation. In regions with harsh weather, like Siberia, animals cooperate, as birds do when migrating across long distances; but even where resources and food are abundant, as in the Amur and Usuri regions, "facts of real competition and struggle between higher animals of the same species came very seldom under my notice" (Íbid.: 60). Kropotkin does not deny the importance of selective competition, but claims that cooperation has been underestimated. Why? The spirit of the time, again: in this case, Russian anarchism. The cultural environment influences the interpretation of the physical world: "true believers tend to read their social preferences into nature" (Gould 1988: 18). And, in the XIX century, success was on the side of the British Empire, not of Kropotkin's anarchic ideal. So:

> If Kropotkin drew inappropriate hope for social reform from his concept of nature, other Darwinians had erred just as firmly (and for motives that most of us would now decry) in justifying imperial conquest, racism, and oppression of industrial workers as the harsh outcome of natural selection in the competitive mode. (Gould 1988: 21)

And today? How is our answer different from Darwin's? A century and a half after the publication of the *Origin*, our understanding of life has evolved too, especially through the integration of genetics into the theory of selection accomplished by the Modern Synthesis: an approach known as Neo-Darwinism. And genetic reductionism, glorified by its undeniable successes, together with Darwinian selection, seem to be the mark of *our* time: they are commonplace ideas in today's society, applied with enthusiasm well beyond their original scope. (Fig. 1.3).

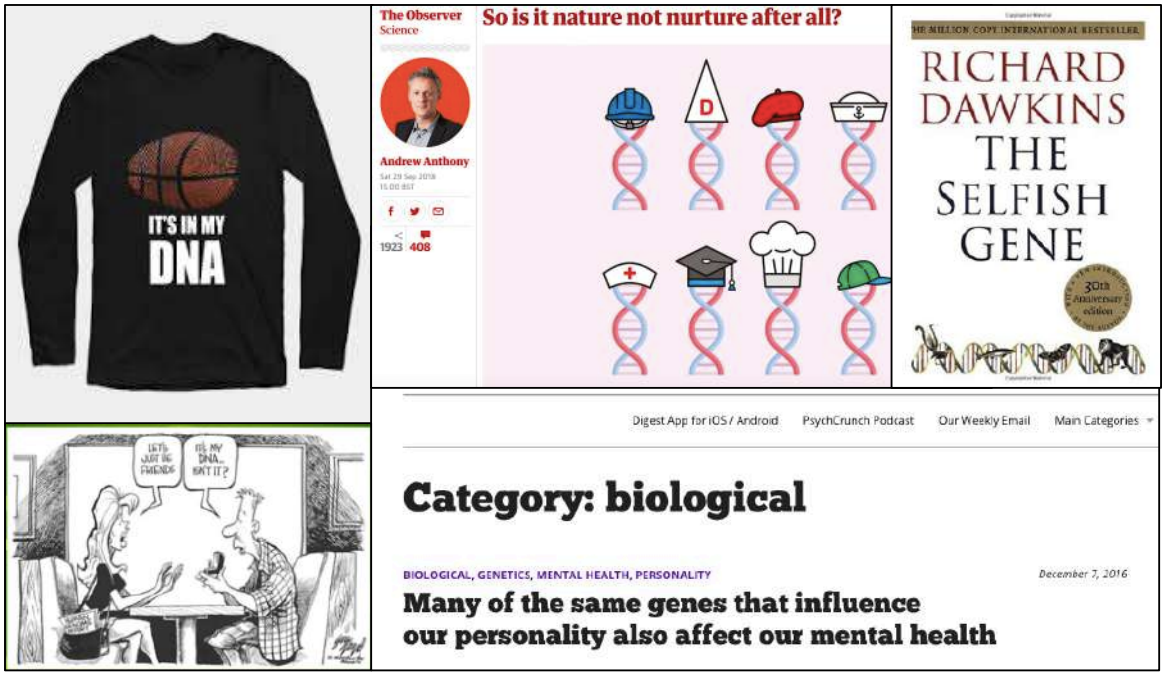


Fig. 1.3 – Genetic reductionism in academy and its interpretation in popular culture.

So, the suspicion arises: where are the limitations of this view, of the spirit of our time? I think the main one has to do with the question about complexity: selection on random mutations and recombination might be a part, but not all, of the answer. Let us see in what sense.

Darwin aimed, with his theory of descent with modification through natural selection, to account for three biological explananda: the variety of living forms, within populations, between species and across phylogeny; the complexity of organisms, apparently increasing through the history of life; and adaptedness, or the fit of organisms with their environment. The success of this original research project today, more than a century and a half after the publication of the *Origin*, is uneven. The Modern Synthesis formalizes, through population genetics, the claim about variety: traits in a population spread, get fixed, are lost and slowly change into new ones according to their fitness and to the strength of selection. On the other hand, population genetics does not explicitly address the issue of complexity, let alone its putative increase. Darwin's theory is also the only known scientific explanans for adaptedness. The evolutionary mindset known as adaptationism studies species' traits and provides ex-post explanations about their adaptive origin. A major limitation of adaptationist explanations is their narrative nature is that they are 'just-so stories', using S.J. Gould's famous dictum: they cannot be tested. Several models have been proposed to overcome this limitation, from optimization programs to the ambitious Formal Darwinism Project by Alan

Grafen. These proposals, however, focus on traits separately, and ignore the complexities of the architecture of organisms.

The issue of complexity thus remains either unaddressed (by population genetics) or taken for granted and explained informally and ex-post (by adaptationism). For the spirit of our time, that only sees the world through the glasses of natural selection and genetic reductionism, complexity originates from a Deus-ex-Machina hidden in the environment: we just have to explain why it appears (the task of adaptationism) and how it spreads (the task of population genetics). The formal explanation of how traits originate is left wanting in Neo-Darwinism.

This orthodox approach has been increasingly challenged in the last decades, through the analysis of non-purely selective processes and phenomena that affect evolution. Many of them have been collected under the name of Extended Evolutionary Synthesis (Laland et al. 2015). The range of phenomena targeted by these accounts spans from chemical-physical laws, to genetic (e.g. Cherniak & Rodriguez-Esteban 2013, Kimura 1983, Wagner 2015), developmental (e.g. Maynard Smith et al. 1985), systemic (e.g. Kauffman 2000) and neo-Lamarckian mechanisms (e.g. Koonin & Wolf 2009). None of these accounts denies the importance and even preponderance of selection in the history of life, and they rather aim at integrating non-selective phenomena into Neo-Darwinism (a view known as 'pluralism'). Although criticized by mainstream evolutionary biology, I believe that pluralistic views and classical Neo-Darwinism can be integrated into a unified vision of evolution that formally accounts for organismal complexity: this is the fundamental hypothesis of this dissertation. We need to go beyond the 'monotheism' of the spirit of our time, in which selection is the only recognised creative power in the shaping of life, to a 'polytheism' in which there is place for other such powers.

This transformation of paradigm requires, in my opinion, two changes in the theoretical framework of evolutionary biology: in the first place, a new definition of organismal architecture going beyond the adaptationist – which equates of organisms to the sum of their optimised traits - is required; and, in the second place, a new measure of the impact of non-selective mechanisms on phenotypes beyond fitness - an intrinsically selective measure – is required. With these two tools, it is possible to build a formal model of evolutionary changes that tracks the effect of selection and of non-selective forces on populations and organisms. Complexity appears in this model to be the result of several processes acting at the same time, among which selection plays an important but not preponderant role.

The structure of this work is as follows.

Chapter 2 introduces the three main biological *explananda* that the theory of natural selection aims to address: the variety and complexity of life, and the adaptedness of organisms to their environment (§2.1). It briefly shows how Neo-Darwinism successfully explains variety (that is, why and how some existing phenotypes are more frequent than others) through population genetics' equations (§2.2), while complexity (that is, why and how exiting phenotypes transform into different ones) remains an open issue lacking a correspondent formalization, and still relies on informal narratives (§2.3). It also presents the two main research program trying to overcome this limitation: formal adaptationism and pluralism, that are the subject of chapters 3 and 4 respectively. The chapter finally introduces the issue of complexity in biology, proposing a definition of biological architecture as sum of form and function, and a model to track complexity changes in organisms (§2.4). It also addresses the problem of whether and in what sense complexity increases in the history of life (§2.5). The rest of the work is devoted to the second dimension of life -complexity- and how the two mentioned approaches explain it.

Chapter 3 develops the view that explains complexity uniquely through selection, an approach known as adaptationism. After introducing its different versions (§3.1), it focuses on the extreme version of Neo-Darwinism known as empirical adaptationism and it illustrates its main tenets (§3.2) and its attempts of formalization, including the already mentioned *Formal Darwinism Project* of Alan Grafen (§3.5, §3.6). After underling some problems of this latter approach, the chapter suggests a formalization of fitness that could help overcoming them (§3.7). The chapter also compares Darwin's original writings to extreme adaptationism to show how they sometimes diverge (§3.3), and illustrates in what sense adaptationism is a functionalist and externalist position (§3.4).

Chapter 4 is dedicated to the universe of non-adaptationist explanations of complexity. First of all, it suggests in what sense adaptationism falls short of explaining what is going on in evolving species, and illustrates some of the main phenomena that adaptationism ignores (§4.1). The main section of the chapter introduces some of the non-selective theories of the evolution of complexity, classified according to where they consider that variation appears: the genotype (§4.2), the phenotype (§4.3), the developmental system (§4.4) or the organism as a whole (§4.5).

Chapter 5 is the core of this work: it introduces an original model that aims at unifying selection and non-selective proposals within a novel framework. Our underlying working assumption (based on the criticisms to adaptationism and the description of pluralism

provided in the previous chapters) is that the mechanism of natural selection on variations caused by point mutations and re-combinations can explain local optima in the shaping of organismic phenotypes, but not substantial novel architectures (as defined in chapter 2 as the sum of forms and functions). To account for these novelties, other processes must also be considered (§5.1.1). The proposal defended here is that fitness changes, by definition, can only track the action of natural selection: some other measure should be used to detect non-selective processes (§5.1.2). Interpreting evolution as a two-stage process of variation-creation followed by variation-selection, I suggest that 'reproduction', measured by fitness, is the key issue in the second stage, while 'survival' is the key issue in the first: I propose to measure this first stage through robustness, a concept I thoroughly analyse (§5.2). By distinguishing the effects of natural selection from the effects of other non-selective processes, and by linking fitness changes to the former and robustness changes to the latter, a clearer picture of evolutionary phenomena can be sketched on a bi-dimensional design-space, in which complex evolutionary phenomena can be decomposed into simpler ones, and the forces fuelling them can be identified (§5.3). The action of selection and of the main non-selective forces can be represented in this bi-dimensional design-space, as the examples illustrated will show.

Chapter 6 summarises the findings of the thesis, suggests some possible further developments and concludes.

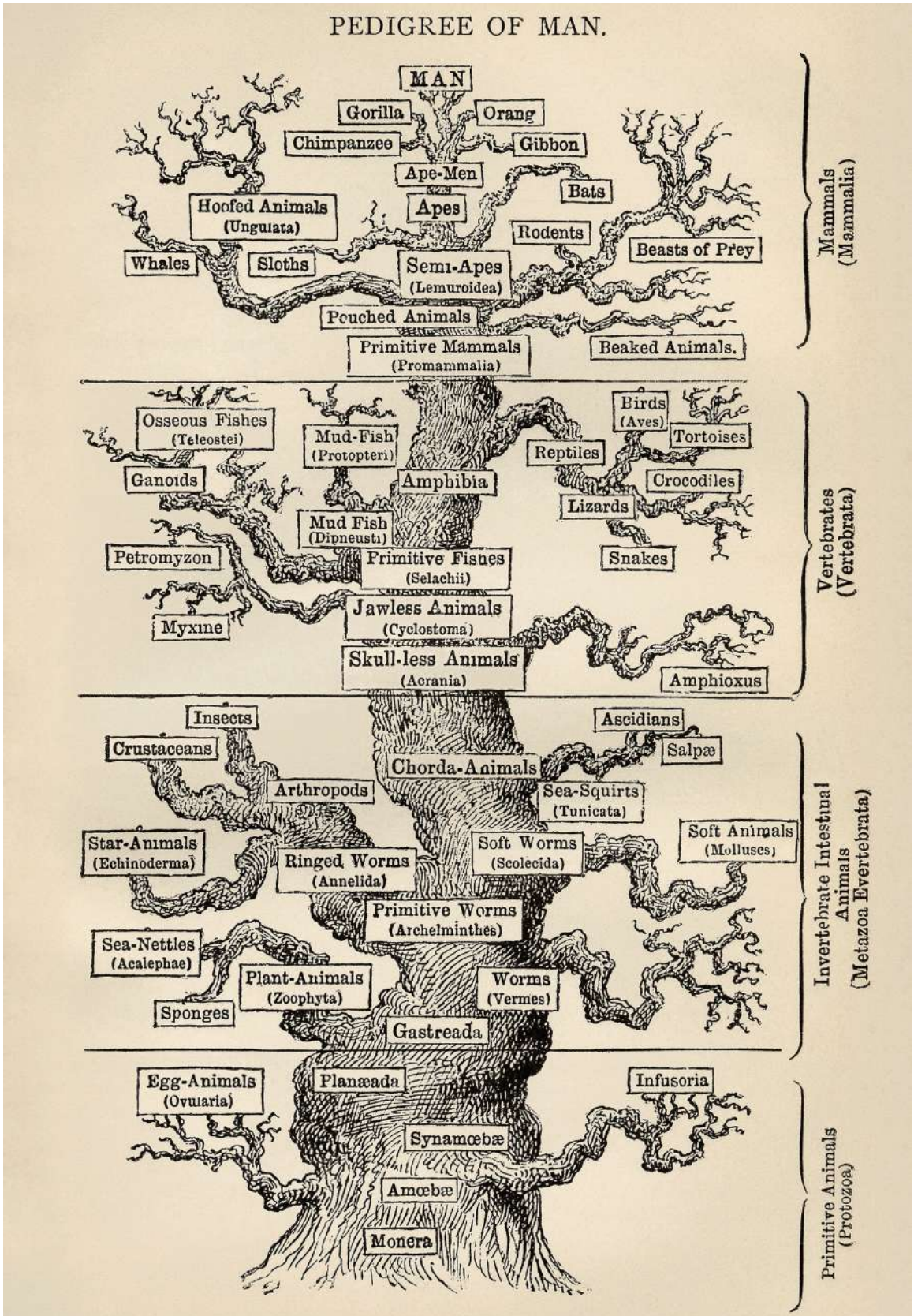


Plate XII – The pedigree of man -from Haeckel, E., *Anthropogenie oder Entwickelungsgeschichte des Menschen. Gemeinverständliche wissenschaftliche Vorträge über die Grundzüge der menschlichen Keimes- und Stammes-Geschichte*. Leipzig: Engelmann, 1874

2. Natural selection and the variety and complexity of the living world

This chapter discusses two main issues that are fundamental for the rest of this thesis. Firstly, it introduces in §2.1 Darwin's thought about the relationship between selection and the three main evolutionary *explananda*: adaptedness, diversity and complexity: natural selection, by fostering adaptness, can explain the two dimensions of the tree of life: its increasing diversity and the (apparent) progressive rise in complexity of biological organisms (Brown 1991). Grafen (2014) has underlined that, while population genetics has formalised the first of these claims, showing how selection configures traits' frequencies in populations based on their relative fitness contributions, it remains agnostic regarding the second, namely how complexity appears: §2.2, by providing some high-level notions of population genetics models, illustrates these points. Adaptationism, on the other hand, aims at supplying narratives to justify the role of selection not only in the distribution of traits, but also in their appearance (and thus explanations of the complexity of organisms): its attempts, however, are controversial, let aside that the creative role of selection is not universally accepted (§2.3.1). As a consequence, two main approaches in current biology strive with the problem of the evolution of complexity, a problem that will be comprehensively illustrated in the next chapters: formal adaptationism and pluralism. The former applies optimization programs to population genetics models, relying on a reductionist definition of complexity (usually identified with some scalar parameter, typically linked to fitness); the latter includes non-selective processes and forces into the picture (§2.3.1). Chapter 3 and 4 are dedicated to these approaches and their proposals.

The second main issue of this chapter is the definition of complexity. After briefly reviewing the main definitions of complexity in the literature (§2.4.1), we focus the analysis on biological complexity and propose a definition in terms of form and function (§2.4.2), that can be applied to classify evolutionary phenomena in a novel way (§2.4.3). Our definition will be used in chapter 5 within the model that we propose in order to unify adaptationist and pluralistic accounts.

We finally review the main positions about the putative increase of complexity in the history of life, to clarify that our interest in the present chapter is not in supporting either of the options, but only in understanding how complexity changes happen (§2.5).

2.1. Darwin's view: selection as *explanans* of both variety and complexity

Among the several ideas introduced by Darwin in the *Origin*, there is the hypothesis that natural selection is behind not only the branching of the tree of life (the variety of organisms), but also the increasing 'organization' of the living (the complexity of organisms[12]):

> Natural selection acts exclusively by the preservation and accumulation of variations, which are beneficial under the organic and inorganic conditions to which each creature is exposed at all periods of life. The ultimate result is that *each creature tends to become more and more improved* in relation to its conditions. This improvement inevitably leads to the *gradual advancement of the organisation* of the greater number of living beings throughout the world. (Darwin 1872: 141, italics mine)

Darwin's theory encompasses much more than the theory of natural selection: Mayr (2004) claims that it includes five hypotheses, each in some way independent from the others:

1. The fact of evolution, that Darwin defines "descent with modification";
2. The branching theory of common descent;
3. The multiplication of species;
4. Gradual evolution (as opposed to saltationism);
5. The theory of natural selection as main force (not the only one, though) behind evolution.

The claims are independent because, for example, the theory of natural selection could be true even if the gradualist hypothesis turned out to be false; or if life had originated several times on earth and different species or phyla did not descended from the same

[12] As we will see in §2.3.1, complexity is a polysemous concept. At this introductory stage of the discussion, we must content ourselves with a 'common sense' interpretation of complexity, one that considers a multicellular organism as more complex than a prokaryote cell, and a human brain more complex than a reptile one.

ancestor. Evolution, on the other hand, could be driven by a mechanism different than selection; and the same independence can be shown among all the five hypothesis.

Each of the five hypotheses above was not particularly original on its own, and Darwin just took sides in disputes that had been around for long: for example, the gradualist vs. saltationism quarrel (see e.g. Gould 2002: 396-466). As for the concept of 'evolution', it was already accepted as characteristic of the inorganic universe and human culture: Darwin just extended it to the realm of life (Lewontin, 1968). Darwin's novel contribution consists in putting together existing ideas in a new way, that allows him to propose a naturalistic solution to some apparently puzzling questions that were, and are, trending topics in biology:

> […] the essential nature of the Darwinian revolution was neither the introduction of evolutionism as a world view (since historically that is not the case) nor the emphasis on natural selection as the main motive force in evolution (since empirically that may not be the case), but rather the replacement of a metaphysical view of variation among organisms by a materialistic view. (Lewontin, 1974: 4).

For Darwin, variation among individuals, and the fixation of such variation in populations, is at the base of evolution: for the pre-Darwinian thought, variations of actual organisms were just deviations with respect to an 'ideal' type. Additionally, previous evolutionary theories were *transformational*: the individual, like Lamarck's giraffe, was the subject of change. Darwin's formulation of the process of evolution was an epistemological break because it was based on a *variational* principle: the system evolves by changes in the proportions of the different types of individuals, now the *object* of changes (Lewontin 1983).

Sober likewise underlines that the novelty of Darwin's theory was in how he combined pre-existing ideas:

> I mentioned before that Darwin was not the first biologist to think that current species were descended from ancestors different from themselves. The same point can be made about the second ingredient in Darwin's theory: the idea that natural selection can modify the composition of a population was not original with Darwin. But if the idea of evolution wasn't new and the idea of natural selection wasn't new, what was new in Darwin's theory? Darwin's innovation was to combine these ideas - to propose that natural selection is the principal explanation of *why evolution has produced the diversity of life forms we observe*. (Sober 2000: 10, italics mine)

The theory of natural selection addresses three main biological *explananda* behind the five hypothesis listed by Mayr (Brown 1991):

A. The fit of Form and Function: why are organisms so well-fitted to survive in their ecological niche[13]?
B. The variety of life: why are there so many different forms of life?
C. The progression of life: why do organisms (seem to) become more and more complex along the tree of life?

While the next chapter (§3) focuses on the first of these questions, we now analyse Darwin's thoughts about the remaining two: that natural selection, by improving the fit between the individual and the environment, explains at the same time the variety and the complexity of living organisms. A way to see this is by considering the two dimensions of the Tree of Life (Fig. 2.1).

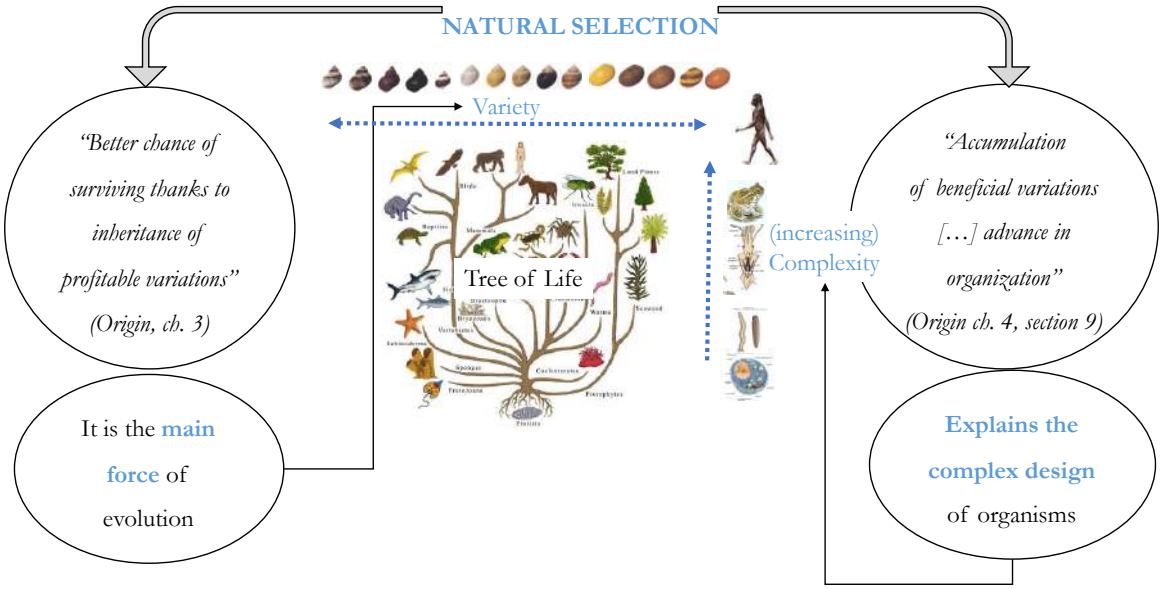


Fig. 2.1 – Darwin's 'Tree of Life' illustrates how both variety and complexity tend to increase thanks to selection.

Individual variation is at the heart of Darwin's theory, contrary to the previous metaphysical thought, dating back to Plato and Aristotle, that the individual is always an imperfect approximation of an ideal 'type', and variation a mark of nature's imperfection (Mayr 1963, Lewontin 1974). The width of the tree widens as generations pass, giving rise to a greater

[13] Some authors claim that this question is based on a fallacy, as the notions of 'ecology' and 'phenotype' are inter-defined (Fodor & Piattelli-Palmarini 2010: 143). See §4.1.4 for a more detailed discussion on this subject.

variety of living forms, both intra-species (among individuals of the same species) and inter-species (among individuals of different species), as

> [A]ny being, if it vary [sic] however slightly in any manner profitable to itself, under the complex and sometimes varying conditions of life, will have a better chance of surviving, and thus be NATURALLY SELECTED. From the strong principle of inheritance, any selected variety will tend to propagate its new and modified form. (Darwin 1872: 17, capitals in the original).

Natural Selection is thus the process behind the evolution of variety. But there is another idea underlying the sketch: generation after generation, the "preservation and accumulation of variations, which are beneficial" leads to "the gradual advancement of the organisation of the greater number of living beings throughout the world" (Darwin 1872: 141). And also:

> The inhabitants of the world at each successive period in its history have beaten their predecessors in the race for life, and are, in so far, higher in the scale, and their structure has generally become more specialised; and this may account for the common belief held by so many palaeontologists, that organisation on the whole has progressed. (Darwin 1872: 414).

Selection is also the process that explains the increasing complex design of organisms: "And as natural selection works solely by and for the good of each being, all corporeal and mental endowments will tend to progress towards perfection" (Darwin 1872: 554). In contemporary biological vocabulary, 'evolution for'[14] leads, gradually and uniformly, to a directional and progressive trend that increases the fit of the organism with the environment (Rasskin-Gutman & Esteve-Altava 2008). Cain's *The perfection of animals* (1979) is a paradigmatic defence of Darwin's ideas: Cain claims that all organismic traits are adaptations, and even criticizes Darwin when talking about 'trivial characters': no trait is functionless.

2.2. Natural selection and variety

The Darwinian view of evolution has been defined as the conversion of inter-individual variations into intra-population and intra-species ones, in space and time (Lewontin 1974).

[14] For a definition of 'evolution for' as opposed to 'evolution of', see (Sober 2000: 82)

Population genetics studies the first part of this conversion: the origin and dynamics of genetic changes within populations.

2.2.1. Population genetics or: why are some phenotypic traits more frequent than others?

Let us analyse the claim about the ability of natural selection to explain the variety of life (point B of Brown's list – see §2.1). From domains to species, to the differences among individuals of the same population, the collection of extinct and living organisms is countless:

> Biologists use the vernacular term diversity in several different technical senses. They may talk about 'diversity' as number of distinct species in a group: among mammals, rodent diversity is high, more than 1,500 separate species; horse diversify is low, since zebras, donkeys, and true horses come in fewer than ten species. But biologists also speak of 'diversity' as difference in body plans. Three blind mice of differing species do not make a diverse fauna, but an elephant, a tree, and an ant do -even though each assemblage contains just three species-. (Gould 1989: 49)

However different, these organisms can be organised, according to Darwin, in continuous series nested into one another, to form a whole whose elements are just a little 'adaptive' bit apart. Species that are at first sight very much separated from each other can be linked through a long but finite series of very small changes, that can be qualitative (a new trait unknown before, like the ability to give birth -viviparity- together with the already existing of laying eggs -oviparity- of the lizard *Saiphos equalis*, Smith & Shine 1997), quantitative (an existing trait with new dimensions or proportions, like the thickening of the shell of the marine mussel *Mytilus edulis* in response to the introduction of a new predator, the crab *Hemigrapsus sanguineus*, Freeman & Byers 2006), relational (the same traits arranged in a new fashion, like the loss of ears' symmetry that allows some species of owls to identify the altitude of a sound -what changes is not one or the other trait, but their spatial relationship-, Konishi & Volman 1990, 1994; or the morphological changes linked to changes in the diet of the lizard *Podarcis sicula* when moved to a new environment, Herrel et al. 2008), or a mix of these (like the increasing ability of the invasive cane toad *Chaunus [Bufo] marinus* to invade the Australian territory, Phillips et al. 2010).

The theoretical efforts to unify the work of Mendel[15] and Darwin by Fisher, Haldane and Wright in the 1930s, known as the 'modern synthesis' (Huxley 1942) or Neo-Darwinism, have successfully formalized the nature of the link between variety and natural selection and the corresponding inter-generational dynamics, through the testable mathematical models of Population Genetics. Population genetics studies "the extent of genetic variation within populations, why it exists, and how it changes over the course of many generations" and predicts "how the gene pool will change in response to fluctuation in size, migration, and new environments." (Brooker 1999: 704-706). The cost of population genetics success is the simplification of both Darwin's theory and the reality of biology. In order to become a testable theory and to show predictive power, the formalization of population genetics ignores some caveats of the original Darwinian ideas, and transforms some others into new concepts: the idea that the sources of variation can be multiple (as the pluralistic view suggests) is an example of an element of Darwin's thought that the modern synthesis do not take int account (see for example Gould 2002: 505-508); the transformation of the Darwinian idea of 'fit between individual and environment' into Fisher scalar number 'fitness of an individual' is an example of a radical change in a concept's scope (§5.1.2). Population genetics also simplifies the nature of biological reality: the focus moves to genes, while the phenotype is considered as its deterministic result, built through a 1-to-1 gene-trait mapping[16] (at least as far as relevant traits are concerned); the organism is reduced to its fitness value; and novelties appear thanks to the *deus ex-machina* represented by mutations. By substituting the organism with its genes, and assigning to each allele a fitness value, population genetics models can thus explain how and why some traits increase their frequencies in a population while others tend to disappear: evolution is explained by postulating that it corresponds to changes in alleles' frequencies. Under such simplifications and assumptions, it is not surprising that the population genetics models work best for simple organisms like the asexually reproducing bacteria *E. coli*, as in Lenski's long-term evolution experiment (Lenski 2003, Dawkins 2009). Starting from 12 identical populations in

[15] The idea that characters' transmission is Boolean rescues Darwinism from the problem of how to avoid homogeneity in a population after a few generations. If the offspring's characters are just a continuous mix of the parents', variation is quickly lost. The Mendelian bits of information about character are transmitted without blending, so that they are preserved even if unexpressed in one generation. This is the real and novel contribution of Mendel to Neo-Darwinism.

[16] Phenomena like epistasis (expression of a gene modified by another gene) and pleiotropism (a gene affecting more than one trait), that fall outside the 1-to-1 paradigm, are also accounted for in advanced models of population genetics (Hamilton 2009, Burger 2011); this, however, does not affect the idea that the phenotype is a deterministic function of the genotype, where the relevant 'stuff' happens.

1988, Lenksi has tracked the genetic changes appeared during 60,000 generations, such as initial rapid improvement in fitness that later became slower, appearance of defects in DNA repair, and, in one of the population, a novelty: the ability to survive on citrate, unusual in *E. coli*. Once the mutation allowing the new metabolic capacity appeared, it spread across the population according to population genetics forecasts.

The major limitations of population genetics are thus its inability to explain how new phenotypes appear in the first place:

> The real guts of evolution -which is, how do you come to have tigers and horses, and things [*that is, phenotypic novelties*]- is outside the mathematical theory [*population genetics model*]. (Waddington 1967, quoted in Gould 2002: 584, italics mine)

and its inability to explain how speciation and extinction occur:

> While population genetics has a great deal to say about changes or stability of the frequencies of genes in populations and about the rate of divergence of gene frequencies in populations partly or wholly isolated from each other, it has contributed little to our understanding of speciation and nothing to our understanding of extinction. (Lewontin 1974: 12)

2.2.2. Foundational assumptions of population genetics

Population genetics assumptions are foundational in the sense that they are not empirical, but define the scope and limits of the discipline. That whatever lies outside of the scope of population genetics is ignored or denied, is not a fault of the models but of their interpretation. The model is not the reality, the same way as "a map is not the territory" (Korzybski, 1933): we cannot refuse to see an actual tree only because it is not marked in our map; likewise, we should not deny the existence and relevancy of evolutionary impact of processes and forces not included in population genetic models only because they are not included!

Evolution as change in genic (alleles) frequencies. Population genetics defines evolution as any change in alleles frequencies within a population: "Since evolution is a change in the genetic composition of a population, the mechanisms of evolution constitute problems of population genetics" (Dobzhansky 1937). Its focus is thus at the genetic level. It is peculiar, one might notice, that one of the ideas at the core of Darwin's theory – i.e., the fact of

evolution in terms of phenotypic differentiation - is reduced to a concept unknown to Darwin. The phenotype, on which selection directly acts, is ignored[17], as it is assumed that information about the phenotype is not incremental with respect to what the genotype[18] already reveals, and that the next generation's gene pool only depends on the current gene pool and its alleles' fitness[19].

Phenotype as function of the genotype. Evolution by natural selection entails variety, fitness differences and heritability. Population genetics models, however, also considers evolution without fitness differences, that is, without selection, as in the case of drift, or migration: variety in the gene pool, in these cases, is not linked to differential survival and reproduction, but to random phenomena. What the models do not consider is evolution *without genetic heritability* (e.g. epigenetic phenomena, see e.g. Jablonka & Lamb 2005). This is the fundamental justification for ignoring the phenotype: as long as the process of the formation of the phenotype in terms of expression of the genotype is deterministic, information concerning the phenotype becomes irrelevant; and as long as information concerning the phenotype can supply additional information concerning phenogenesis or development, it is irrelevant as it cannot be inherited, i.e., transmitted to the gene pool of the following generation.

Genetic fitness as comprehensive description of an organism. Given that evolution equals changes in alleles frequencies and that the phenotype is reducible to the genotype, the only information required to forecast how the population will appear in the following generation, in addition to the configuration of the genetic pool in the current generation, are the alleles' fitness, or why some individuals have more offspring than others (Bethel 1976). The organism is its

---

[17] Population genetics models do not usually explicitly include information about the phenotype. They do include such information implicitly in the alleles' fitness values, but the phenotypic mechanisms that give rise to fitness remain unveiled.

[18] The genotype can be considered as a systemic property whose components are alleles (thanks to Davide Vecchi for this definition).

[19] There is an interesting caveat of this approach: the phenotype depends not on alleles directly, but on the genotype. And genotypes frequencies can vary even when alleles frequencies remain constant. Sober proposes the example of a population whose members only mate with organisms having the same genotype at a given locus (assortive mating): after one generation, genic frequencies are the same, while allelic frequencies are not (homozygote frequencies have increased and heterozygote frequency has decreased): "Is the process generated by assortative mating an evolutionary one? It is standard fare in evolution texts and journals. To exclude it from the subject matter of evolutionary theory would be a groundless stipulation. I conclude that evolution does not require change in gene frequency." (Sober 1999: 2).

genotype and its destiny is measured by its genetic fitness. Only the reproduction side of Darwin's argument is considered relevant, while the survival is reduced to a pre-requisite for reproduction. How fitness is reached, by which mechanisms and causes, is of no interest to population genetics: the aim is not to explain *why* some genotypes are more successful than others, but, given that they are, just to track their frequency increase or decrease. It's interesting to note that, although the adaptationist attitude seems quite the opposite (see Fig. 2.4) - as it analyses the phenotype, ignoring the genotype, building narrative explanations of why a given trait exists and provides a competitive advantage - it is likewise based on the reduction of the organism to its fitness: the difference between population genetics models and adaptationism is that the latter considers *phenotypic fitness* instead.

2.2.3. Population genetics models

Once it has reduced biological reality to fitness and alleles, population genetics builds models able to forecast the evolutionary trajectory of populations (Fig. 2.2).

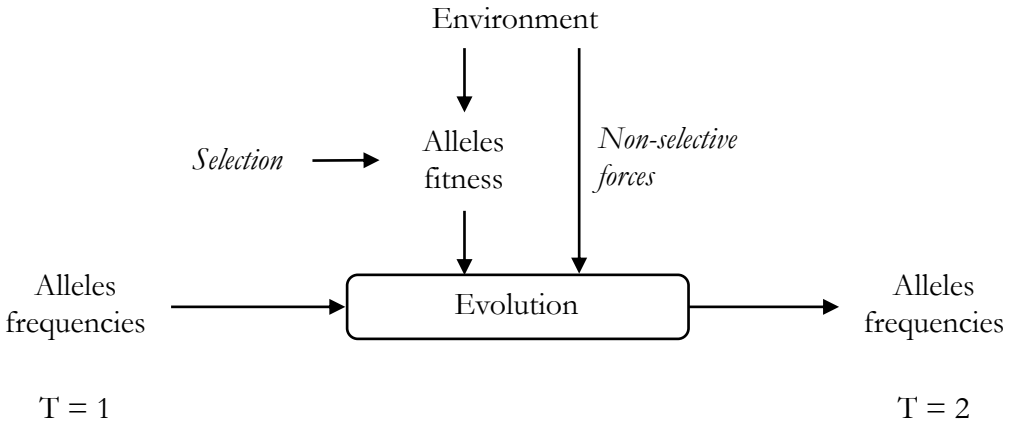


Fig. 2.2 – Concept of evolution in population genetics as change in alleles frequencies. The change is driven by the environment, either directly thorough non-selective forces (mutations, migrations, etc.) or indirectly through fitness differences defined by selection.

As we have seen, population genetics' focus is thus not on the individual and its phenotype, but on the collection of genes of a population, or its *gene pool*. Two variables are considered: alleles' frequencies (the number of copies of an allele vs. the total number of alleles in the population) and genotype frequencies (the number of individuals with a genotype vs. the total number of the population). These variables are linked through the Hardy-Weinberg equation (

Table 2.1) provided that the following conditions are met (Brooker 1999: 710):

- Populations are large enough to avoid any drift effect;
- There is no migration;
- There are no selective pressures;
- There are no mutations;
- Mating is random.

Table 2.1

Allele	Frecuency
A	p
a	q

⇔ Hardy-Weinberg eq. $\mathbf{p^2 + 2pq + q^2 = 1}$ ⇔

Genotype	Frequency
AA	p^2
Aa	2pq
aa	q^2

The Hardy-Weinberg equation and how it links alleles and genotypes frequencies. In a diploid species with two alleles (A and a) for each gene, the equation forecasts the frequencies at equilibrium (that is, after some generations have passed) of the three possible genotypes AA, Aa and aa, based on the alleles frequencies p and q[20]. Provided the equation conditions of applicability are met, alleles' and genotypes' frequencies remain constant.

The Hardy-Weinberg equation is thus considered as describing a population in which 'nothing happens' in evolutionary terms: it is the equivalent of the inertia principle in classical mechanics, the null hypothesis against which hypotheses around evolutionary phenomena are tested[21]. As such, it is widely used: in the first semester of 2003, the Science Citation Index counted 57 papers testing the Hardy-Weinberg law (Pigliucci & Kaplan 2006: 244).

If any of the listed conditions fails, the equation does not hold: mutations can introduce new alleles, and migration, drift or selection can spread it through the population. In case of selective pressure, for example, it can be mathematically demonstrated that the three terms of the equation (p^2, q^2 and 2pq) do not add up to 1 (as in the H-W case, when selection is null), but to the mean fitness of the population (Brooker 1999: 721). Population genetics has developed mathematical formalisms to account for these cases, and to forecast the evolution of gene frequencies (Hamilton 2009).

[20] The equation can be generalized for more than two alleles and for more than two sets of homologous chromosomes (polyploidy).

[21] The status of the H-W law as inertial principle has been lately questioned: stasis (Luque 2016) and drift (McShea & Brandon 2010) have been proposed as alternative principles to describe a population under no evolutionary forces.

All population genetics models can be described in terms of Fisher's Fundamental Theorem of Natural Selection (Grafen 2002), which links *changes in average fitness* between generations with the *fitness variance* of the original population. The theorem states:

> The rate of increase in fitness of any organism at any time is equal to its genetic variance[22] in fitness at that time. (Fisher 1930).

The theorem is based on certain simplifying assumptions (Wright 1982):

> [F]irst, that the likelihood that a mutation would be favourable falls off so rapidly with the magnitude of its effect that only those with minor effects need be considered[23]; second, that the local effects of accidents of sampling are negligible because they are overwhelmed by dispersion from neighbouring localities and third that one-to-one relationships between gene and character are the norm. (Wright 1982: 3)

The first assumption is clear: mutations with large effect are usually deleterious, so only small ones provide feasible phenotypic effects. The second simply means that drift can be ignored. The third assumption is the strongest one, because it entails that all genes have additive phenotypic effects[24], and that the organism is "a mosaic of unit characters" (Wright 1982: 4), of independent traits coded by the corresponding genes.

There is another tacit assumption behind the theorem: that fitnesses remain constant. This equals to assuming a constant environment (both in terms of physical environment -e.g. no temperature changes- and of other species sharing the same environment -e.g. no new predators-), and the absence of frequency-dependent selection (traits whose fitness depend on their relative frequency in the population)[25].

Common sense might help to understand this latter principle: if all individuals in a population have the same fitness, the population variance in fitness is null and we don't expect average fitness to change from one generation to the next. If, on the other hand,

---

[22] In probability theory and statistics, variance is the expectation of the squared deviation of a random variable from its mean. Informally, it measures how far a set of (random) numbers are spread out from their average value. (In Wikipedia, The Free Encyclopedia. Retrieved 09:26, September 13, 2019, from https://en.wikipedia.org/w/index.php?title=Variance&oldid=911844733)

[23] The concept of 'magnitude of effect' of a mutation is open to interpretation: if acquiring a new metabolic capacity is considered a 'large effect', Lenski's experiments (Lenski 2003) would invalidate Fisher's statement (thanks to Davide Vecchi for this comment).

[24] Genes' effects are additive if each gene contributes to the phenotype independently. If genes interfere with each other, their effect is non-additive: we have dominance (if the interaction is between genes at the same locus) or epistasis (if the interaction is between genes at different loci).

[25] For an analysis of these factors see for example (Okasha 2018: 60-61).

individuals have very different fitnesses, the population fitness variance is high and we expect the average fitness of the population to grow from one generation to the next, approaching the higher values, given that natural selection pushes the fittest organisms to have more descendants[26]. Fisher compared his theorem to the second principle of thermodynamics, according to which the entropy in a closed system never decreases: in the same way, fitness in an isolated population never decreases, given that variance is, by definition, never negative (Fisher 1930). As there is empirical evidence that fitness does indeed decrease sometimes, the validity of the theorem was highly controversial until Price (1972), through a purely mathematical deduction, clarified the conditions under which it is applicable. Fitness does not decrease if natural selection is the only force acting on the population: that is, if the population is isolated and not affected by migration, drift and other non-selective forces. If this is not the case, Fisher's theorem should take into account changes in expectation (i.e., changes in average fitness due to migration, mutations, etc. and not to simple re-assortment of existing alleles) that in fact can be negative and reduce average fitness.

Fisher's theorem turns out to represent a particular case of the Price equation, which describes the evolution between generations of the frequency of some characters of the population as the sum of two terms:

- The covariance (Cov) between fitness and the relevant character, representing the contribution of the latter to survival and reproduction;
- The expected change (E) of the value of the character regardless of its fitness contribution, representing phenomena like mutation, recombination and so on. This term is also known as 'transmission bias' (Okasha 2006).

The Price equation can be represented as follows:

$$W \Delta Z = \text{Cov}(w, z) + E(w\Delta z)$$

where:

- z is the value of a character and Z its average value in the population;
- ΔZ is the change in average value of the character between generations;

[26] Fisher's theorem holds when social behaviour and frequency-dependence selection are ignored. Otherwise, selection can cause fitness to decrease (Grafen 2015). Altruism is a good example of frequency-dependence: the higher the frequency of selfish individuals, the lower the altruist fitness. If selfishness' fitness is higher than altruistic fitness for any percentage of selfish individuals, a population of 99% altruistic individuals will end up with 100% selfish individuals because of selection: the finale average fitness will be lower than the initial one (see Sober 1998b: 416 for details).

- w is the absolute fitness of an individual (number of descendants) and W its average value in the population;

If we consider natural selection only, the second term disappears; and if the trait z is fitness itself, the covariance term becomes the variance in fitness: we thus obtain Fisher's theorem:

$$W\,\Delta W = \mathrm{Var}\,(w), \text{ or}$$

$$\Delta W = \mathrm{Var}\,(w_{rel})$$

where w_{rel} is the individual relative fitness with respect to the population average.

Apart from Fisher's theorem, many other population genetics formalizations can be derived from the Price equation (Queller 2017).

	Price Equation z = any measurable phenotypic trait				
Equation	$w\bar{\Delta}\bar{z}$	=	$Cov(w, z)$	+	$E\,(w\Delta z)$
Interpretation	Changes in average trait between generations		Covariance bewteen fitness and the trait in original population		Expectation of trait changes
	Fisher Theorem z = fitness w (considered as target of selection)				**Price extension of Fisher T.**
Equation	$w\bar{\Delta}w$	=	$Var(w)$*	+	$E\,(w\Delta w)$
Interpretation	Changes in average fitness between generations		Variance in fitness in original population		Expectation of fitness changes
Value	always ≥ 0		always ≥ 0**		can be > or < 0
Forces	N/A		Natural Selection		Non-Selective forces (mutation, recombination,…)

** Cov(w, w) = Var (w); ** a variance is never negative by definition*

Fig. 2.3 – Connection between Fisher's Theorem and the Price equation. The Price equation is a generalization of Fisher's theorem in two senses: it extend its application to any trait (and not just fitness) and it also considers the effects on the trait of non-selective forces (mutation, recombination, etc.). Fisher's theorem box. The theorem links changes in average fitness between generations to the variance in fitness in the original population, while ignoring any effect from other non-selective forces. Price extension of Fisher theorem box. Price clarifies that the theorem holds true if selection is the only acting force; if other forces act, the second term E, that can be negative, comes into play and can outweigh the positive contribution of selection. Price equation box. Generalization of Fisher theorem: instead of fitness w, the equation considers (1) the values of any measurable trait z and (2) the impact E of non-selective forces.

In the metaphor of evolution as a theory of forces (Sober 1984), while the Hardy-Weinberg law plays the role of inertia, the Price equation acts as Newton's third law. The metaphor is highly debated. Lewontin (1983), for example, criticises the view of the organism as a passive object subject to external (e.g. selection) and internal (e.g. mutations) forces; given that the organism is merely the result of this clash, the implication is that it is evolutionarily irrelevant, making "the gene the only real unit of selection" and the study of evolution "nothing but a combination of molecular biology and geology". *The Selfish Gene* by

Dawkins (2006) is a classic example of such a view. By ignoring the phenotype, population genetics does neither share with adaptationism the problem of justifying how traits appeared in the first place through some narrative about their contribution to fitness, nor the need of proving that they tend to optimality. Consequently, population genetics models do not make any claim about fitness optimization (Edwards 2007, Birch 2018, Grodwohl 2016[27]) an issue that, as it will be detailed in chapter 3, is at the core of the efforts of formalization of adaptationism and its claims around complexity.

In this section, we have presented the answer provided by population genetics to the question about the variety of life: why are individuals of the same population different from each other? And why do these differences not soften after few generations, as it should be the case if the offspring's characters were just an average of their parents' traits? Why populations, once separated, tend to diverge and transform I new species? Population genetics explains not only individual differences in traits (through the study of phenotypic effect of alleles), but also why these differences are maintained in a population (according to the Hardy-Weinberg law), and why some traits tends to become more frequent (following Fisher theorem), or to be substituted by new, more efficient, ones (as forecasted by the Price equations). By unifying the discoveries of genetics and the theory of Darwin, population genetics provides a powerful and unsurpassed tool for the understanding of life's variety. This success comes, however, at a cost: giving up the task of addressing phenotypic complexity. The remaining of this chapter (and of this whole work) is devoted to show how this gap can be addressed.

2.3. Natural selection and phenotypic complexity

Let's now turn to the question of complexity (point C in Brown's list, see §2.1): how do new types arise and why are they more complex than the existing ones (if they are)? The question can be decomposed into two sub-themes: the definition of complex architecture and the putative existence of a tendency towards higher complexity in the history of life. Different lines of research give different answers to these themes. As we have seen, population genetics describes the organism simply through alleles' fitness, and does not directly engage with complexity at the level of the phenotype. Besides, it remains agnostic regarding fitness

[27] Grodwohl traces the history of the idea of fitness optimization in population genetics, until its fall in the 1980s.

optimization, thus abandoning any ambition to link phenotypic complexity to selection[28]. This strategy can have a two-folded reading: either phenotypic complexity is reduced to its genic and genotypic side or, if it does have additional elements, these are irrelevant for forecasting evolution. On the other hand, phenotypic complexity and its relationship with natural selection is at the core of the adaptationist line of research, which proposes unstructured definitions of complexity through narrative justifications of phenotypic traits. To bring some order into this debate, we illustrate in the next sections the main ideas in favour and against the existence of a causal link between selection and complexity (§2.3.1), and describe the two resulting approaches (§2.3.2): the one that defends the link and tries to formalize it (formal adaptationism), and the one that rejects the link and suggests other sources of complexity (pluralism).

2.3.1. What is the nature of the link between natural selection and biological complexity?

Given that life originated from inorganic matter, it is clear that there has been (at least in some cases) an increase in phenotypic complexity (however defined) over the past 3.5 billion years, although long-term stasis has been the predominant pattern in most lineages. The question is whether natural selection is a necessary or sufficient force to explain the emergence of the genomic and cellular features central to the building of complex organisms (Lynch 2007), or whether other forces, mechanisms and constraints contribute and how. Even among the authors claiming that a trend towards higher complexity exists, there is no agreement about its causes. Gregory (2008b) lists several mechanisms behind evolutionary trends in general, among which selection, either because a directional change in some trait grants a fitness advantage or because stabilizing selection creates a lower boundary to a trait value (e.g. minimum complexity), so that any change would be in one direction only[29]; constraints, such as the higher probability of some mutations, or physical limits; directional speciation, if new species tend to differ from their ancestors in biased manners; irreversibility of some changes, that makes further changes possible only in one direction; differential speciation rate, if species with higher value of some trait tend to speciate more often; species hitchhiking, meaning that a trait correlated with another one showing a trend, will also show a trend.

---

[28] Adaptationism identifies fitness optimization with improvement of design, e.g. in (Grafen 2004): organism are designed through selection for fitness optimization.

[29] See note 95 for the definition of different kinds of selections.

Among evolutionary biologists, especially those endorsing an adaptationist view, the idea that natural selection fosters increasing complexity is common (e.g. King 2004, Jacob 1977), even if there is disagreement around the relative importance of micro-evolution (at population level) and macro-evolution (at species-level) in pushing the trend (Gregory 2008b). Bonner, in a book whose title *The Evolution of Complexity by Means of Natural Selection* anticipates all of its content, provides a good example of this extreme adaptationist view:

> Why has there been an evolution from the primitive bacteria of billions of years ago to the large and complex organisms of today? The answer to the question of 'why' is straightforward: it is natural selection. (Bonner 1988: ix)

However, the acceptance of the preeminent role of natural selection in improving phenotypic design is not so universal (see e.g. Gould & Lewontin 1979). Maynard-Smith and Szathmáry (1999), as we have seen, believe that complexity increase is a fact, but claim that there is no reason that evolution by natural selection alone should be behind it. Its real cause, they argue, is linked to major changes in the way information is stored, transmitted and translated. They identify eight major transitions in complexity along the history of life:

- From replicating molecules to populations of molecules in protocells. Simple replication is not informational, evolution requires different kinds of replicating molecules to co-operate and, for this to happen, they have to be enclosed within a compartment.
- Form independent replicators to chromosomes. The co-ordinated replication of genes belonging to the same chromosome prevents competition among genes of the same compartment.
- From the world of RNA – a molecule that performed both the activity of storing and transmitting information as well as that of catalysing reactions and form the organism's structure - to the world of DNA and protein, with each taking care of one of the activities. This transition requires the evolution of a genetic code, whose sequences determine proteins' structures.
- From prokaryotes, lacking nucleus and having a singular circular chromosome, to eukaryotes. Among the many events involved in this change, the spatial separation of transcription and translation, the emergence of rod-shaped chromosomes with no limitation on genome size and with multiple origins of replication, and the assimilation of cell organelles with own genetic information, all represent major transitions in the authors' sense.

- From asexual clones to sexual populations. Sexual reproduction radically changes the way genetic information is transmitted to offspring. Among its advantages, it allows the production of offspring that are not all alike (as in parthenogenesis), the unification in the same individual of two advantageous mutations, or the elimination of harmful ones (thanks to recombination).
- From protists, organisms of one or few kinds of cells, to animals, plants, and fungi, composed by many different kinds. Given that all cell types carry the same genetic information, this transition requires cells differentiation at some stage of development, and thus a new way of using genetic information.
- From solitary individuals to colonies, similar to superorganisms, and in which only a few individuals reproduce.
- From primate societies to human societies, based on human language, sharing with the genetic code the ability of unlimited heredity based on a limited set of basic rules.

Each transaction entails the loss of the ability of the initial entity to replicate independently: sexual organisms, for example, cannot reproduce on their own as asexual organisms can. This leads McShea (1996) to claim that the authors talk about a specific complexity, namely hierarchical process complexity. Even applying this specific definition, he criticises their analysis as lacking a unified and measurable definition of 'change in the way the information is transmitted', and of a common set of features that could define each transition.

Lynch (2007) underlines that, if organismal complexity does grant adaptive advantage, the massive dominance of unicellular species over multicellular remains unexplained (multicellularity only arose in animals, fungi and in vascular plants, see Bonner 1998). The author also questions that the independent appearance of multicellularity can be due to natural selection, given that "multicellular species experience reduced population sizes and reduced recombination rates, and increased deleterious mutation rates, all of which diminish the efficiency of selection" (Lunch 2007: 8600). Lynch not only denies the existence of any evidence supporting the 'myth' that selection promotes complexity, but he also claims that genomic complexity is fostered by a less efficient selection process. It seems that "non-adaptive processes have played a central role in driving evolutionary patterns" and that "the origins of biological complexity should no longer be viewed as extraordinarily low-probability outcomes of unobservable adaptive challenges, but expected derivatives of the special population-genetic features of DNA-based genomes" (*idem*: 8603-4). A similar point has been made previously by Kauffman (2000), although his conclusions were derived from

models far removed from mainstream population genetics. Koonin also denies any trend in complexity fuelled by natural selection: "[…] there is no trend for complexification in the history of life" (Koonin 2009: 1024). Any increase in complexity is linked to weak purifying selection, thus to *a failure of the mechanism of selection*.

McShea & Brandon (2010) defend a similar conclusion. The authors claim that a trend towards complexity exists and that it is a consequence of the Zero Force Evolutionary law, as several known phenomena confirm, for example:

- Nucleotide positions that are not under selection diversify spontaneously;
- Tissues and organs that are not under selection are more variable among individuals;
- Homologous characters can be maintained by stabilizing selection in multiple lineages and yet developmental and genetic mechanisms that underline them might diverge;
- Laboratory populations spontaneously drift.

However, less complexity than forecasted by their theory is found in the history of life: they deduce that selection may work not for but *against* complexity. McShea suggests the same conclusion when, talking about 'internal variance' of the organism (or horizontal complexity) he claims that " it is unclear whether selection will enhance this vector [i.e. tendency to higher complexity], act neutrally, or oppose it[30]" (McShea 2005: 146).

The criticisms to the role of natural selection in shaping complex design are partly justified by the lack of universal formalization of the accounts supporting this hypothesis: their advocates usually recur to arguments that explain phenotypic traits either through informal narratives impossible to falsify (Gould & Lewontin 1979), or through *ad-hoc* mathematical models presupposing the action of some fitness optimization mechanism (see e.g. Cain 1989 for examples of this kind of adaptationism; and Ch. 3). Population genetics models supply solid and widely accepted universal formalizations, but they are usually agnostic regarding the optimization of fitness, which they even deny in some cases (e.g. in the case of overdominance[31]):

[30] Stabilizing selection, e.g. tends to destroy variation (see note 95 for a definition of stabilizing selection).

[31] That is, of better fitness of heterozygote individuals vs homozygotes. In a population of 100% heterozygotes, average fitness decreases from a generation to the following, as the offspring includes homozygotes individuals.

> The idea of optimization was an acceptable part of the operation of population genetics in the 1950s and 1960s, but it became increasingly frowned on because the arguments based on optimization were dynamically insufficient. (Grafen 2003)

Provided that, for both population genetics and adaptationism, the link between selection and phenotypic design is identified with some measure of fitness (phenotypic or genotypic, of a trait or of the individual), which is implicitly considered as a complete and sufficient description of the organism, we face a puzzle, firstly identified by Grafen (2014): on the one side, population genetics formally supports the Darwinian claim that selection explains life's variety, but it is agnostic about design and denies the optimization of fitness, which is usually considered as design optimization by adaptationist accounts (e.g. Grafen 2014)[32]; on the other side, adaptationist accounts develop possible narratives to back up the Darwinian claim that links selection to design, but recur to unformalized considerations about fitness optimization and, unable to produce a universal formal model, propose *ad-hoc* ones through optimization programs. This situation was, in some way, already underlined by Lewontin:

> Apparently, then, we have two parallel systems of evolutionary dynamics, one operating in the space of genotypes and bypassing the phenotypic space, and another operating entirely in the phenotypic domain. (Lewontin 1974: 15)

Lewontin, however, considers that the incommunicability between what he calls the 'biometric' corpus (framed in phenotypic terms) and the 'mendelian' corpus (framed in genetic terms) is illusory, as genotype and phenotype are inter-linked: genetic fitness depends on phenotypic traits, and phenotypic traits are in great part the output of genetic information. State variables in population genetics models are not phenotype-independent, and state variables in phenotypic equations (e.g. the breeder's equation, see Kelly 2011) are not genotype-independent. The illusion rests on a poor interpretation of state variables: by measuring a quantitative trait, we are indirectly including genetic information; and by measuring fitness, we are indirectly taking into account the population genetic pool.

Lewontin's analysis is very clarifying but ignores, in our opinion, two assumptions that aim at deactivating the genotype-phenotype inter-dependency:

- Population genetics ignores non-heritable phenotypic traits (see above): phenotypic fitness might add information to the genotypic one, but this information is irrelevant;

[32] See chapter 3 for the role of fitness as proxy for complex design in adaptationist models.

- Adaptationism considers that selection will solve any problem regardless of genetic details (Mayr 1963) and endorses what Grafen defines 'phenotypic gambit' (Grafen 2014): the idea that phenotypic traits can be analysed regardless of their genotypic encoding (a typical functionalist approach). Slatkin, for example, proposes a model that works "directly at the phenotypic level where selection acts" and avoids "the necessity of a direct statement about the underlying genetic description of the population." (Slatkin 1970: 88). As Stoltzfus puts it: "once we understand the "gene pool" and the power of "shifting gene frequencies", the genetic details cease to matter" (2017: 2).

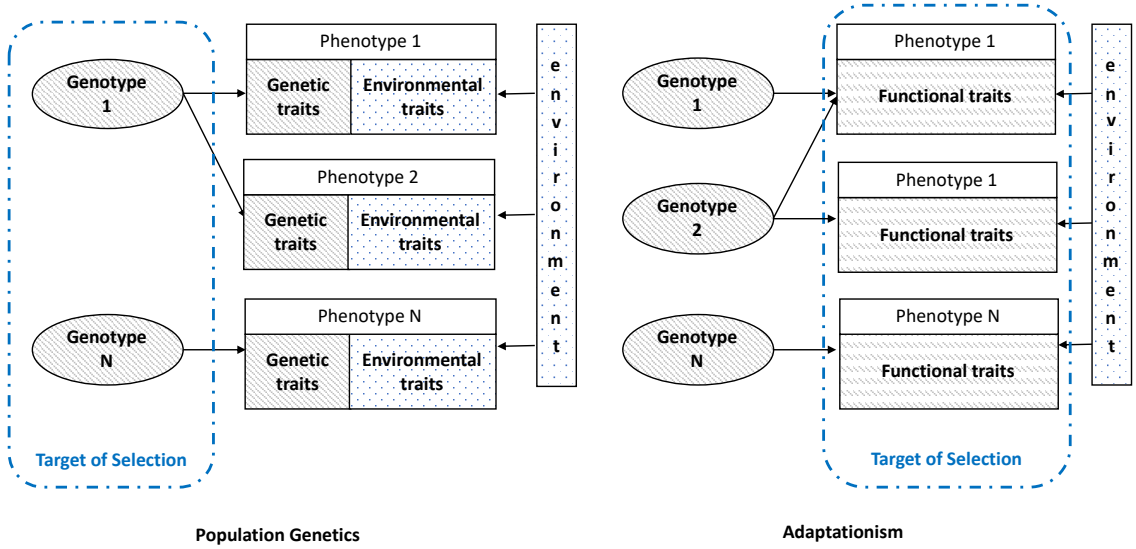


Fig. 2.4 - The different focuses of population genetics and adaptationism. Population genetics believes that only heritable traits are evolutionary significant, and that they are a deterministic output of the genotype, where selection acts (e.g. Dawkins 2006). Environmental or other traits exist and have an impact on the individual fate, but are invisible to selection, whose target is the genotype. Adaptationism considers that the target of selection lies at the phenotypic level, where functional traits are relevant as long as they contribute to fitness and regardless of their genetic encoding. As the same functional solution can be based on different genotypes, the relevant information to understand evolution cannot lie in genes.

The illusion of independency mentioned by Lewontin is thus theoretically real, but practically ignored: and it is towards this practice that Grafen draws our attention. As a matter of fact, both views, the 'biometric' and the 'mendelian', ignore one of the sides of the genotype-phenotype relationship, based on opposite arguments (Fig. 2.4).

Population Genetics mathematical models focus on the *genotype*, where the action of selection is tracked through differential *genic fitness*. These models capture the *mechanism of selection* and forecast the evolution of alleles' frequencies in a given population and under given environmental conditions. Existing alleles, and new alleles appeared thanks to mutations, migrations and the like, become more or less frequent thanks to their fitness: but explaining

the reasons behind these fitnesses is not at the core of population genetics. In particular, population genetics models do not rely on ideas of fitness optimization, nor can fitness optimization be deduced from them: "mathematical population geneticists today recognize no general principle of fitness maximization" (Grafen 2015: 10). This makes sense, as such reasons are to be found at the phenotypic level, as "the whole real guts of evolution -which is, how do you come to have horses and tigers, and things- is outside the mathematical theory" (Waddington 1967).

Adaptationist accounts, on the contrary, focus on the *phenotype* as both the target and the output of selection, and justify phenotypic traits by explaining how they contribute to the organism fitness. As each trait is considered as a separable unit, the global configuration of the organism is justified as the optimal mix of qualitative traits and/or of the optimal values of quantitative traits. Unfortunately, no universal formal model exists of such informal narrative explanations.

We thus face an unsatisfactory situation with regards to organismal complexity. From one side, population genetics explains and forecasts variation between individuals in a population or between populations through sound mathematical models, but it is silent about how and why novelties arise:

> Population genetics sets a much more modest goal than general evolutionary theory. […] While population genetics has a great deal to say about changes or stability of the frequencies of genes in populations and about the rate of divergence of gene frequencies in populations partly or wholly isolated from each other, it has contributed little to our understanding of speciation and nothing to our understanding of extinction. Yet speciation and extinction are as much aspects of evolution as is the phyletic evolution that is the subject of evolutionary genetics, strictly speaking. (Lewontin 1974: 12)

From the other side, adaptationism suggests how and why novelties arise and result in speciation, but it does so mainly through narratives, and is unable to provide a comprehensive and unified formal model. Fig. 2.5 summarises the peculiarities of both accounts.

In the next section, we introduce two approaches that aim at overcoming this situation: formal adaptationism and pluralism.

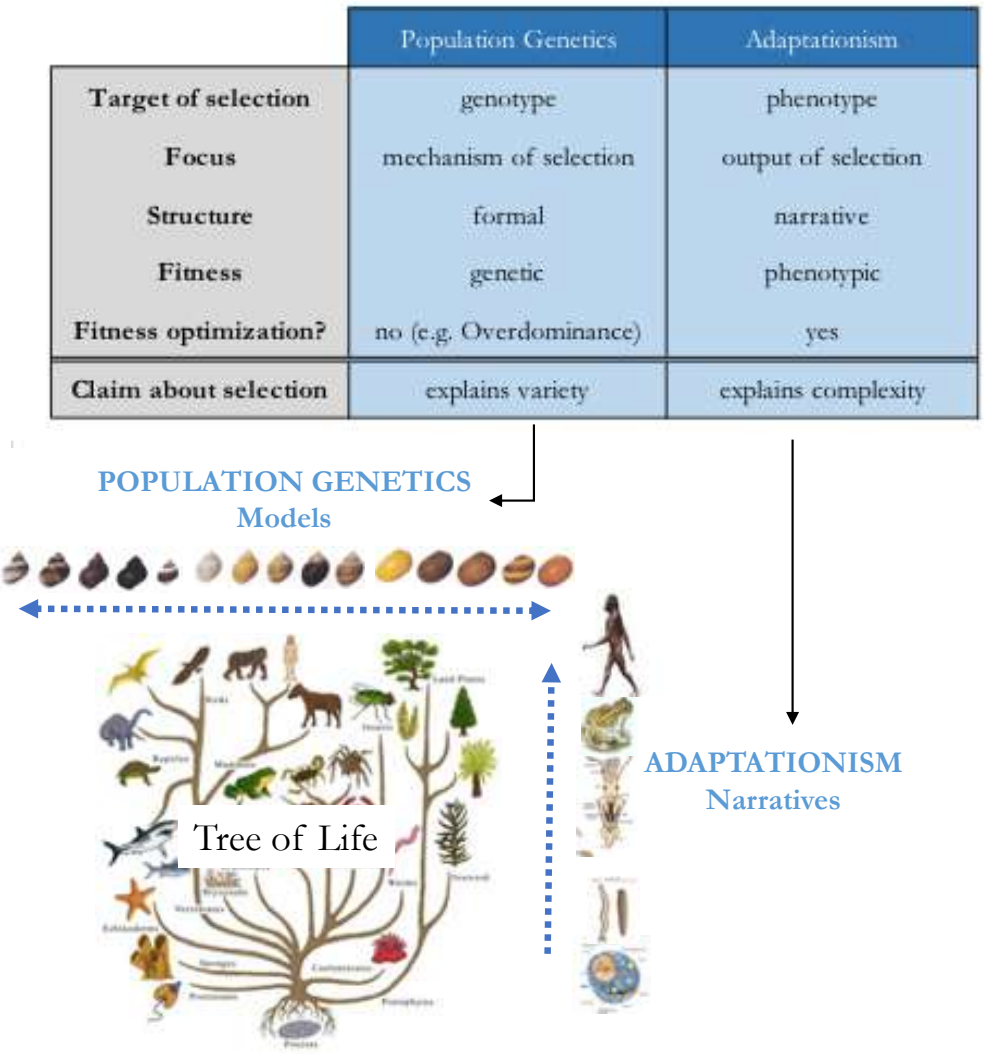


Fig. 2.5 – How population genetics and adaptationism explain the Tree of Life. Population genetics formalises the evolution of variation within populations, but does not address issues of phenotypic complexity; what is done by adaptationism, but with narrative explanations not relaying on formal models.

2.3.2. Two lines of research to explain phenotypic complexity: formal adaptationism and pluralism

To overcome the tension between the ambition of adaptationism to explain complexity and the weakness of its narratives to do so, there are two possible strategy: defending the idea that selection alone is sufficient for the explanation of complexity and trying to formalise the claim, or giving up the idea and including also other forces to explain it. The first approach can be defined as *formal adaptationism*, the latter as *pluralism*. Let us briefly illustrate them.

Some authors have tried to translate the adaptationist narratives into formal models. *Optimization programs*, for example, -an interesting approach borrowed from economics-, reduce complexity to the value of a phenotypic variable (Parker & Maynard Smith 1990). In

optimal foraging models, where such programs have been successfully applied, this variable measures the average foraging time in a place before the individual moves to a new one, and its value is deduced from considerations around the maximization of energy gain per unit of time (Charnov 1976). As a general trend, the narrative that explains a trait is formalized based on the assumption of fitness maximization (to which the trait contributes), an assumption that population genetics often denies. These models, therefore, although predictive and thus falsifiable, lack a solid link with genetics. Grafen's *Formal Darwinism Project* (Grafen 2007, 2014) aims at solving this limitation and represents the most ambitious among adaptationist formalization attempts. Applying an optimization approach to population genetics equations, Grafen shows that, at equilibrium, genetic frequencies as forecasted by these equations lead, as a general tendency, to the maximization of fitness (even if such maximization is not reached because of genetic constraints). We introduce formal adaptationism in chapter 3.

Pluralism is the idea that the evolution of complexity is caused, together with natural selection, by other laws and processes, as well as by historical contingencies (Gould 1997). An extensive literature denies that selection alone can explain all traits (e.g. Maturana & Varela 1980, Pigliucci 2008a, Leland et al. 2015) and that it shows unlimited capacity to produce new traits (e.g. Wagner 2015, Moczek 2008, Eldredge & Gould 1972). Based on these critics, pluralism looks for alternative mechanisms, non-linear and non-progressive, to explain the appearance of new traits and novel architectures in organisms. In many of these proposals, selection too plays an important role: the difference with pure adaptationism lies in its relative importance. Pluralistic accounts do not reduce complexity to fitness, either genic or phenotypic, and try to provide more holistic definitions. We list some of the evolutionary accounts that can be grouped under pluralism in chapter 4.

Table 2.2 summarises the main tenets of the two approaches with regards to the problem of complexity.

Table 2.2

Approach	What is complexity?	Does selection explain complexity?	Proposals	Chapter
Formal Adaptationism	Fitness of a trait, or mix of traits	Main *explicans*	Optimization Programs Formal Darwinism Project	3
Pluralism	Comprehensive, holistic definitions	Together with other other, possibly more relevant mechanisms	Genetic mechanisms Phenotypic mechanisms Environmental mechanisms Systemic mechanisms	4

How formal adaptationism and pluralistic accounts interpret and explain complexity.

2.4. A proposal to formalise biological complexity

We have been talking a great deal about complexity in an informal way: in this section, we amend our fault and we will illustrate some alternative definitions of complexity found in the literature (§2.4.1), just to show that they are of little practical use for our aim of formalising complexity changes due to evolutionary phenomena. We then propose a description of organismal complexity based on for and function (§2.4.2) and apply it to the classification of evolutionary changes (§2.4.3).

2.4.1. What is complexity? Polysemy of the concept.

Complexity is a common sense concept not easy to formalise, and biological complexity is no exception. Is a mammal more complex than a reptile? Or than a plant? And in what sense? What are we comparing when making such claims? Even when the answer seems obvious, a deeper analysis shows that appearances deceive: is the *E. coli* metabolism really less complex than the structure of an eye?

Charles Darwin, in the *Origin*, admits that "[…] naturalists have not defined to each other's satisfaction what is meant by an advance in organisation." (Darwin 1872: 141) There are many proposals for defining complexity in the literature. Classical population genetics' accounts reduce phenotypic complexity to the complexity of the genotype: "[We assume that] an organism's complexity is a reflection of the physical complexity of its genome" (Adami and al. 2000: 4463). An adaptationist like Grafen (2007, 2014) does not provide a formal definition, but claims that "adaptation is design, and maximizing fitness is what

organisms are designed for." (Grafen 2007: 1248) and thus seems to identify complexity as the remote cause of increment of phenotypic fitness, whose immediate cause is natural selection. The 'science of complexity' approach, on the contrary, identifies complexity as an emergent phenomenon that arises from the interaction among homogeneous and relatively simple elements that constitute open, non-hierarchical and far-from-equilibrium systems, at the edge between order and disorder (Kauffman 2000). For Mitchell (2009), complex systems exhibit non-trivial, emergent and self-organizing behaviour. For Johnson (2010), a complex system is a subtle mix between order and disorder, moving between the two. The HOT *Highly Optimized Tolerance* (HOT) approach, proposed by Carlson & Doyle (2002), on the contrary, characterizes complex systems as composed by heterogeneous elements organized in hierarchical structures that constitute organizations robust against expected turbulences, yet fragile against unexpected ones.

Mitchell (2009) summarises several alternative ways to define and measure complexity:

1. As 'Shannon entropy', or the average amount of information (or 'surprise for the receiver') carried by a message. An ordered message composed only of the letter A has zero entropy; a message composed of random letters has maximum entropy. The problem, of course, just translates into the new one of how to describe an organism in form of a message. Even extreme reductionists would probably agree that the strings of genetic information themselves are not a complete and sufficient description of the organism. Even if they were, a higher Shannon entropy would not necessarily mean more genomic complexity: a non-functional, totally random genome is formally more complex than a functional, internally-ordered one. The number of genes of an organism is likewise of little significance to measure complexity: the mustard plant *Arabidopsis* has about the same number of genes as humans.
2. As 'algorithmic information' content, or the shortest computer program that could generate the description of an object. Again, a string of n times the letter A can be described as 'print A, n times'; while a random string of all possible letters does not have a compressed description. The problem is that, under this view, the more random an object, the more complex it results. To avoid this pitfall, only 'regularities' within the string are considered, and defined as 'effective complexity'. Both very ordered and random strings lack regularities, so their 'effective complexity' is low. Genes could be an example of regularities among the 'junk DNA' within the genome.
3. As 'logical depth' of an object. It measures how difficult it is to construct an object, for example, as a viable genetic string. Once the object is translated into a sequence of 0s

and 1s, its logical depth is the number of steps the shortest Turing machine would need to produce the string. Although theoretically interesting, this definition is of little practical interest, as it would be complicate to translate, for example, a phenotype into a Boolean string, and to calculate the corresponding Turing machine.

4. As 'thermodynamic depth', or amount of thermodynamic and informational resources required by the physical construction process of the object. In the case of a type of organism, this could mean, in an evolutionary approach, describing all events that led to its appearance; alternatively, in a developmental approach, it would mean accounting for all developmental events needed to construct the organism. A human has therefore a much higher thermodynamic depth that an amoeba. The difficulties of identifying all such events, as well as to define what an event is, make this method of no use for our practical aim of describing complexity changes following an evolutionary phenomenon. It is clear that,, apart from the simplest cases (e.g. a geometrical solid), it would be practically impossible to come up with a value of 'thermodynamic depth' for any real entity (e.g. a virus).
5. As 'statistical complexity', or the minimum amount of information about the past behaviour of a system to forecast its future statistical behaviour. Contrary to Shannon entropy, a random string of letters has a low 'statistical complexity', as its description does not correspond to the whole string, but to a simple rule: 'choose a random letter'. Very ordered and very disordered object have thus low complexity, as common sense suggests (Fig. 2.6). Translation of the object into a message source is not always a trivial task, therefore statistical complexity remains an interesting theoretical concept with hard implementation in the realm of biology.

String complete description	**String shortest description**	**Statistical complexity**
H B F T H C I N D L A R T O P	15 random letters	low
A A A A A A A A A A A A A A	13 times the letter A	
A B D H P	1) The first letter of the string is A 2) The n+1 letter of the string is the letter whose position is n positions after the position of the n letter of the string 3) The last letter of the string is P	high

Fig. 2.6 – Statistical complexity of different alphabetical strings. The shortest descriptions of completely random and completely orders string show low complexity, while strings built following some rule have higher complexity.

6. As 'fractal dimension'. Fractals are geometric objects with some peculiar characteristics; in particular, they show fine organization at different scales and self-similarity properties among scales. To illustrate this, consider the ideal snow-flake in Fig. 2.7: its shape has a pattern repeated also at smaller scales. To build it, one starts with a triangle; than builds, in the middle of each of its sides, three triangles with side sizes equal to 1/3 of the original; than one builds similar triangles on the sides of these three triangles; and so on indefinitely.

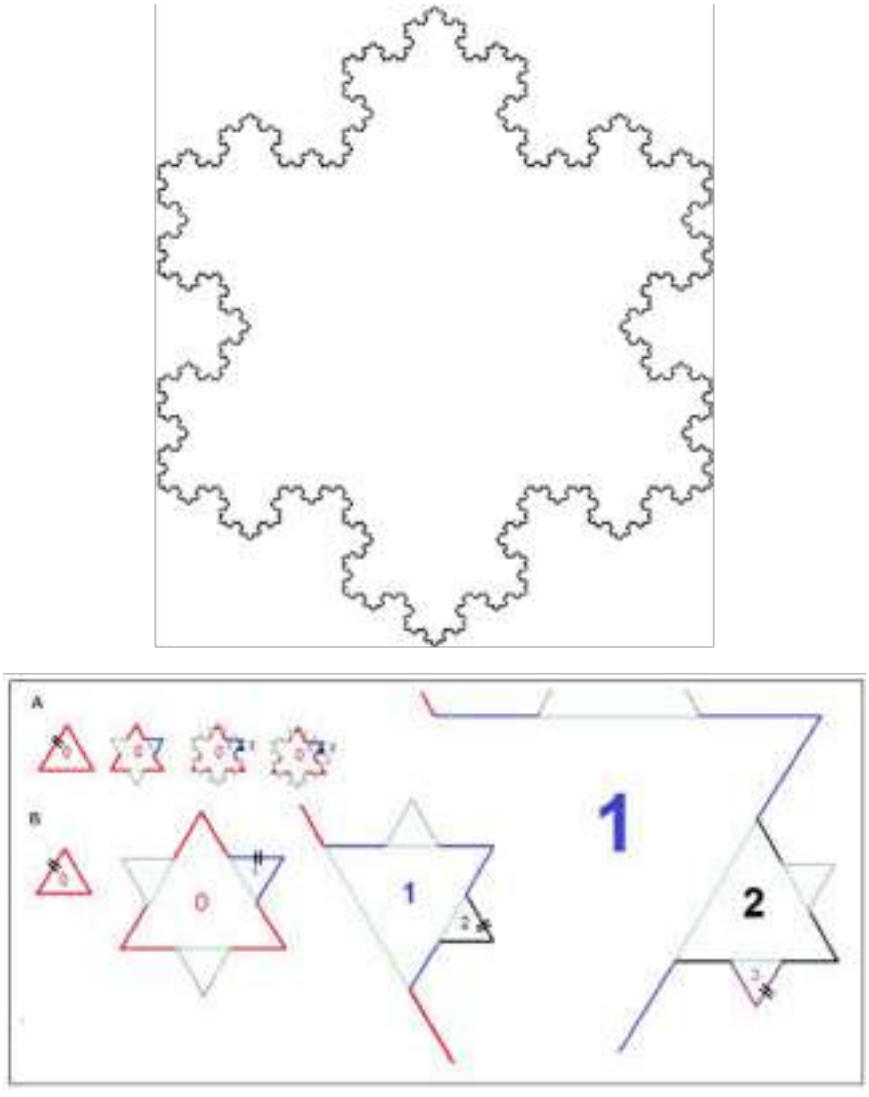


Fig. 2.7 – Fractal structure of an ideal snow-flake. The dimension of a fractal is defined according to the following rule: "Create a geometric structure from an original object by repeatedly dividing the length of its sides by a number x. Then each level is made up of $x^{dimension}$ copies of the previous level.[33]"

[33] This definition is coherent with the common-sense idea of dimension. If you divide a line by x in order to get equal sub-lines, you get x of them: the dimension of a line is thus x=1. If you divide a

Applying the rule to the snow-flake example, you have x=3 (you divide the previous level line into 3 parts), and you obtain 4 new structures (the new level lines). The dimension is thus 1.26, given that $3^{1.26}=4$. Fractal dimensions are not integers, and they measure how interesting the details discovered by looking at the object at higher scale can be. A circle, for example, shows the same information no matter what the scale is, so it has a low fractal dimension.

Fractal dimension captures the geometrical complexity of a structure; however, the complexity of an organism can be considered to depend not only on form, but also on function, and this is something that fractal dimension cannot account for.

7. As 'degree of hierarchy'. Simon (1962) defines a complex system as a large set of parts interacting among themselves in non-linear ways and leading to chaotic regions of the phase space, and whose behaviour is different than the simple sum of its parts (see also Callebaut & Rasskin-Gutman 2005: ix-xiii). Such systems have two distinctive features:
    - Their parts are organised hierarchically, until some elementary subsystem is reached. For examples, organisms are organized in organs, organs in tissues, tissues in cells, cells in cell's parts like nucleus, membrane, etc.
    - Interactions *among* parts are of several orders of magnitude lower than interactions *within* parts. For example, metabolism within a cell consists of many more interactions among molecular components than the number of interactions existing among cells. The system is thus 'nearly decomposable': the short-run behaviour of each subsystem is approximately independent of the behaviour of others; in the long run, it depends on the aggregate behaviour of the other subsystems. This means that the high-frequency dynamics of the system (linked to the internal structure of subsystems) is decoupled from the low-frequency dynamics (linked to interactions among subsystems).

    Simon claims that biological systems are usually hierarchical, as their 'building-blocks' structures speed evolution up. Selection judges the whole organism, not each component trait, given that improvement in one trait might worsen the performance of others: but these cases are made less probable by the interdependency granted by modularity, so that the design of each organ can be improved independently (Simon & Ando 1961).

---

square by x in order to get equal sub-squares, you get x^2 of them: for example, if x=2, you get 4 of them. The square dimension is thus x=2. With a cube, you get a dimension x=3.

McShea (2001) suggests an operational way to measure hierarchy in biological organisms in terms of 'level of nesting' of lower level entities within higher-level individuals. His measure includes two components:

- *Nestedness*. It refers to physical inclusion: higher-level entities are composed by lower-level entities, which in turn are spatially arranged and interact in some specific way. Lower level entities must be homologous with some (extant or extinct) organism in a free-living state (not part of a bigger entity). McShea suggests several level of nestedness:
  - Level1: prokaryotic cells (e.g. bacteria). There is no lower level than cells, because their components (proteins, etc.) are not free-living organisms;
  - Level 2: eukaryotic cells, as aggregates of prokaryotic ones;
  - Level 3: multicellular organisms, as aggregates of eukaryotic cells;
  - Level 4: colonial individuals, as aggregates of cells (e.g. biofilms) or of multicellular organisms (e.g. insect colonies).
- *Individuation*. It refers to the requirement for the higher-level entity to be identified as a new whole, i.e., as 'an individual'. Although it is difficult to make a list of such requirements without recurring to functional concepts, McShea (2001) suggests three criteria: connectedness of the lower level entities among them; differentiation into different types of the lower level entities; and presence of intermediate-level parts larger than the lower-level entities.

We can draw some conclusions form from this short outline of definitions. In the first place, complexity cannot be reduced to a scalar measure.

> It is hard to imagine how a useful notion of overall complexity could be devised. (McShea 1996: 480)

This should not come as a surprise, if we consider the etymology of the term from its Latin origin. Complex comes from *cum-plexus*, and indicates the structure of fabric, whose fibres are closely interwoven and cannot be simplified without destroying it. On the other hand, something complicated (*cum-plicum*) has folds or wrinkles (*plicum*) that can be made simple (*sine-plicum*) through a process of ex-plication, or unrolling (*ex-plicare*) by eliminating them (Fig. 2.8). Complexity is a holistic phenomenon involving several structures and processes interacting at a time, and it is maybe the serial functioning of our understanding that makes it hard to grasp, while generating the frustrated need for a reduction impossible to obtain.

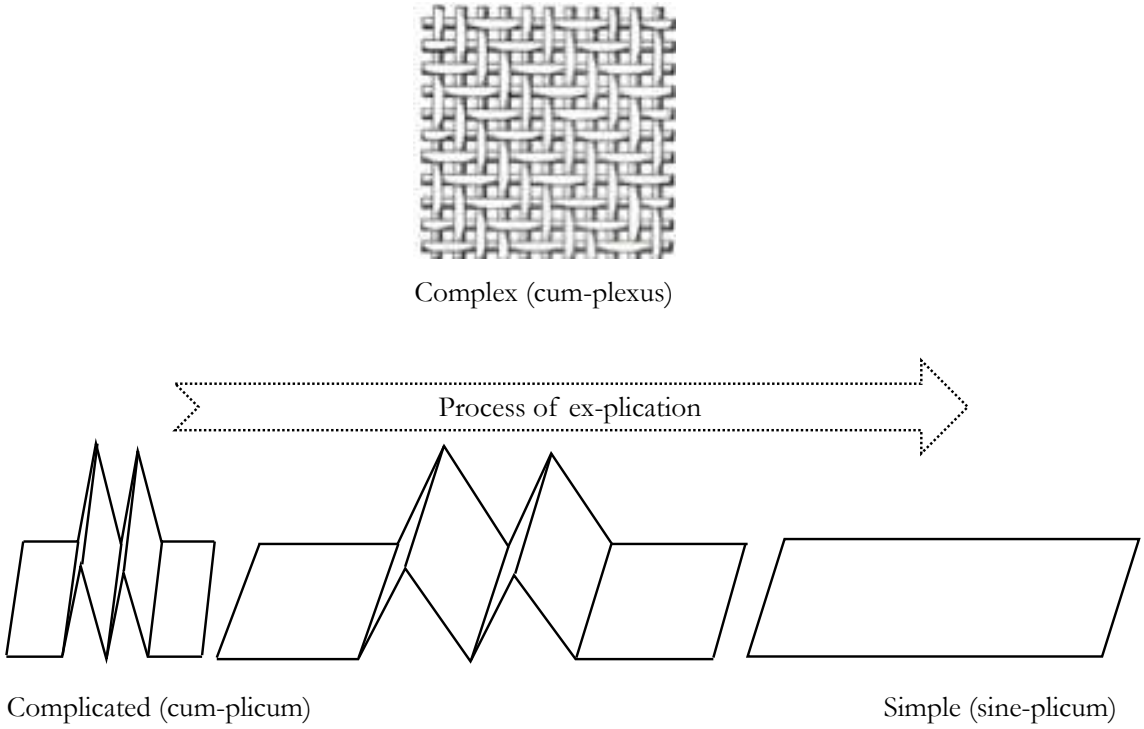


Fig. 2.8 – Difference between (irreducible) complex and (explicable) complicated structures.

Secondly, this brief list of definitions reveals that an entity complexity can be addressed from several points of view, depending on whether the focus is on how complex is to build the entity (as in 3. 'logical depth' and in 4. 'thermodynamic depth'), to describe its structure (as in 6. 'fractal dimension', 7. 'degree of hierarchy', and in 2. 'algorithmic information'), or to describe its relationship with the external world (as in 1. 'Shannon entropy' and in 5. 'statistical complexity'). These three issues correspond in biology to the questions around how an organism comes to be (the main evolutionary question), around why it shows certain traits (the basic structuralist question), and around why it performs certain functions (the typical adaptionist question). The three issues are of course interwoven through the forces, processes and mechanisms that link form, function and evolution: an organism evolves into a different one because its forms and functions change generation after generation.

In the next sections, we propose a way to analyse biological complexity as sum of structural (or morphological) and functional complexity, while in chapter 3 and 4 we describe how adaptationism and other approaches explain the evolution of form and function. It is useful to keep morphology and function separated: a system may be organised in order to perform some function, even with a simple structure. On the other hand, a morphologically complex system may or not have a function: a heap of junk may function as a work of art or just as a random collection of parts (McShea 1991: 305). Prum, for example, claims that a

theory of the origin of feathers is required "based on possible evidence and that is independent of hypotheses about their ancestral function" (Prum 1999: 292). This separation also helps avoiding the trap of looking for a uni-dimensional, scalar definition of complexity.

We do not propose any absolute way to measure morphological and functional complexities, but we suggest a model to track their changes in organisms, and we show how evolutionary phenomena can be described as complexity changes either in form or in function.

2.4.2. Biological complexity: organismal architecture as form and function

In our discussion, we propose to consider the evolution of biological complexity in terms of changes in the architecture of an organism. The idea of comparing living beings and artefacts in terms of their architectures is not new (Davies 2016). We use the term 'architecture' instead of 'organization' because it entails the idea of a set of entities interlinked in a complex structure, while 'organization' just refers to the their arrangement in a design that can, or not, be complex[34]. As the University of Edinburgh's Biological Architecture Group underlines:

> [B]iological architecture includes, but goes beyond, consideration of physical structures. It encompasses the structures of systems, of networks, of connections; it involves structures in time and space, and it is concerned with the fundamental principles of organization that separate the living from the non-living world. (Davies et al. 2017)

How do we define and characterize a biological architecture? We propose to summarise it as the sum of form and functions of the organism.

Many evolutionary phenomena can be interpreted either as a qualitative or quantitative change in *morphology*, or in the *functions* performed by a trait (e.g. Mayr 1960): thus, in its architecture as defined here. We focus our model on morphological phenotypes. Phenomena linked to behavioural or physiological phenotypes, however, can also be accounted for if the reasonable hypothesis is accepted that behaviour and physiology must be implemented in some physical structure and can be interpreted as a set of functions: e.g. in physiology, the

[34] The Oxford English Dictionary defines organization as "the way in which the elements of a whole are arranged", and architecture as "the complex or carefully designed structure of something".

functions of cells, tissues and organs. The reduction of behaviour to functions and traits is supported by Tinbergen 'four questions' model (Tinbergen 1963). The model, based on Aristotle's four causes, explains behaviour through two pairs of complementary categories: diachronic vs. synchronic perspective, and how vs. why questions. Development answers the diachronic/how question (how does the individual come to be); mechanisms answer the synchronic/how question (how do the individual's traits do what they do); evolution answers the diachronic/why question (why did the species come to be); and, finally, function answers the synchronic/why question (why does the species have the traits it has). Reference to evolutionary changes in terms of form and functions are common in the literature. Erwin (2015), for example, defines the changes in morphology as phenotypic novelty and as innovation the changes in functions, and underlines that they are independent: a phenotypic novelty in not necessarily a functional innovation, and an innovation need not involve a novelty (e.g. an exaptation[35] is based on an existing trait). An example of functional change without changes in morphology is found in diet of proboscideans (Erwin 2015): the amount of grass consumed (function) increased before the acquisition of high-crowned teeth resistant to the abrasion of grass (morphology). Likewise, McShea (2001) considers the cytoskeleton as a functional innovation, but not a relevant morphological change.

The form-function dichotomy was already at the base of the Cuvier–Geoffroy debate in 1830[36] (Appel 1987, Asma 1996), and reflects the functionalism/structuralism clash: while the former considers that function comes before the form supporting it (through the fundamental question: what is the function of this form?), the latter believes that form must be explained regardless of its function (e.g. because of geometrical, physical or developmental constraints). Homologies, for example, are explained as output of the same selective pressure by functionalists, while are linked to a reality deeper than adaptation for structuralists (Amundson 2001).

Although the idea that form and function should be kept separated in the analysis of organismic evolution is quite extended, the related terminology is still confused (Brigandt & Love 2012). Erwin (2015), as we have seen, uses the term 'novelty' when referring to morphology, and 'innovation' when talking about functions. Müller (1990, 2007) too talks

[35] An exaptation is a trait contributing to fitness without having been selected for its contribution to fitness. See §4.3.1 in Ch. 4 for an exhaustive definition and some examples.

[36] Georges Cuvier and Étienne Geoffroy Saint-Hilaire debated in the French Academy of Science for almost two months in 1830 around animal structure: Cuvier claimed that it was shaped by the animal needs, while Geoffroy believed that the structure of any animal could be interpreted as a modified version of a same unified plan.

about new structural elements (thus about morphology) with discontinuous origin as 'novelties'. Together with Wagner, he also claims that they must be neither homologous to any ancestral structure nor serially homologous to any other structure in the same organism (Müller & Wagner 1991, see also Moczek 2008). McShea (2001), on the contrary, uses 'novelty' for function and 'innovation' for form. Mayr (1960) defines novelty as a new structure allowing a new functions (thus mixing form and function). Banzhaf et al. (2016) uses novelty indifferently for form and function, although distinguishing among three types of novelties according to the extension and hierarchical level of the change[37]. Wagner (2014), coherently with his theory of gene networks, defines novelty as a part of the body that acquires individuality and quasi-independence. We will apply Erwin's vocabulary, although with an important *caveat*: Erwin considers that a new function measures the reproductive success of a novelty through its contribution to fitness, and that the distinction between form and function reflects the distinction between origin and diffusion of inventions (Erwin 2015). We think that such a view reintroduces a link between the two independent dimensions of organismic architecture. A morphological novelty can have no function and still perpetuate; or an existing morphological trait can suddenly gain a new function. Likewise, a new function can spring from a novel combination of existing, separated traits; or can be maintained even after its contribution to fitness has disappeared.

To better shaper our concept of complexity as sum of forma and function, we develop hereafter each of these two constituting elements.

Morphology

McShea's analysis of biological complexity (2001) focuses only on its morphological aspect, that has two elements: a hierarchical, or 'vertical', one, and an 'horizontal' one. Functional complexity remains outside of the scope of his definition. His approach supplies a way to measure complexity while judgements about function are suspended:

> […] our ignorance of function is profound in biology, and the attempt to separate the functional from the 'random' would be frequently frustrated. We do not know, for example, whether the placement of the human heart on the left or the number of fingers in the hand is functionally significant. (McShea 1996: 481)

He suggests for example that:

[37] See §5.3.5 for more details about this classification.

> [t]he cytoskeleton […] may have represented a key innovation in functional terms, but in structural-hierarchical terms it constituted only one element among many in the elaboration of the host. (McShea 2001: 412)

The indifference to function is a virtue, as it is impossible to understand the link between complexity and any kind of functionality (including fitness) if the concept of complexity has function built into it: question such as 'are organisms with more part types more functional, more favoured, more fit?' make sense only with non-overlapping notions of complexity and function (McShea & Brandon 2010: 50-51). The distinction between morphology and function is especially relevant when analysing phenotypic changes (Erwin 2015).

The second aspect of morphological complexity is 'horizontal complexity' (McShea and Brandon 2010), or the number of different parts of an organism, also referred to as 'internal variance' (McShea 2005). As it happens with 'vertical' complexity, parts are considered regardless of their ability to do something: even functionless ones count. The question to be answered is thus 'how many?', and not 'what for?'. Under this approach, a human with appendix is more complex than one without it. Complexity as number of part types is level relative. A fish, for example, has 120 part types at the level of cells, 90 at the level of tissues/organs, a huge number at the level of molecules and about a dozen at the level of atoms (McShea & Hordijk, 2013). Similarly to McShea and Brandon, Kauffman (1993: 74), in his NK model of biological entities, defines complexity in terms of numerousness, as the number N of elements of a complex system, e.g. the number of genes. Bonner (1998) manages a similar concept of complexity when he measures it in terms of number of cell types (what he calls 'internal' complexity) and number of species (what he calls 'external' complexity).

McShea (1996) extends his analysis to physiological and developmental 'processes', whose elements (metabolic steps and morphogenetic events) have vertical complexity (number of levels in a casual hierarchical chain) or horizontal complexity (number of interactions at a single hierarchical level). The double classification 'object/process' and 'vertical/horizontal' generates four different senses in which a system is complex, each of them measured in a different way (Table 2.3).

Table 2.3

	Hierarchical (vertical complexity)	Non-hierarchical (horizontal complexity)
Objects	*Hierarchical object complexity* Number of levels of nestedness of parts within wholes	*Non-Hierarchical object complexity* Number of parts at the same hierarchical level
Processes	*Hierarchical process complexity* Number of levels in the causal chain	*Non-Hierarchical process complexity* Number of interactions at the same hierarchical level

Four senses of complexity and their possible metrics (based on McShea 1996)

Function

It is now necessary to give a brief outline of the treatment of functions in biology. Biological functions are usually linked to a phenotypic trait (what bears the function) and analysed in terms of teleology (what is their end for the organism) and normativity (what is their correct operation to reach that end). The teleological issue allows to distinguish between relevant and accidental functions, while the normative issue has to do with the contribution of functions to the system. Biological function is a polysemous word, and, similarly to its use for artefacts, machines and their parts, it can have three possible meanings, that should be expressed by three different terms (Canfield 1990):

- As function of the whole artefact, or function as *proper operation*. For example, the function of a car is to move from one point to another; the function of an animal is to live. The normativity considerations concern this sense of function, as they underline the correct operation of the artefact in order to fulfil its end.
- As function of a part, or *function in the strict sense*. For example, the function of the cooling system of a car is to maintain the temperature of the engine below some temperature; the function of the digestive system of an animal is to provide energy, supply nutrients, eliminate waste, etc.
- As function of its function, or function as *utilization*. For example, the utilization of a car can be to run a race, or to transport people. A biological example is the microorganism that rots an apple and makes at the same time appear it repellent:

what is the function of making it so? Because no other organism is thus led to eating the apple. The teleological considerations concern this sense of function.

In order to treat these issues, there are several ways to consider functions in biology: the traditional etiological approach, the dispositional approach, and the more recent organizational approach. The *etiological approach* (from *etiology*, i.e. causal history) is basically an adaptationist interpretation, according to which (1) the function of a trait contributes to the organism's ability to survive and reproduce and (2) this ability causally explains the trait's appearance and maintenance in a population through an historical process of selection (Godfrey-Smith 1994). This view has several drawbacks, for example the fact that function and trait (morphology) are interlinked in a hierarchical structure, where the former is ontologically previous to the trait 'bearing' it (as it is to be expected from a functionalist view as adaptationism): it therefore falls in the 'hierarchical process complexity' category of McShea classification (Table 2.3). A main weakness of the etiological approach is its 'epiphenomenalism': by explaining functions through their past (adaptive) history, it does not provide information about the current function (the 'phenomenon') itself (Christensen & Bickhard 2002). On the other hand, it supplies an answer to the teleological and the normative issue: the telos of a function is its contribution to the existence of the individual, and it is malfunctioning if it fails to do what it was selected for. The normativity side of this account, though, is deeply linked to an adaptationist view: as it justifies the current existence of the function by its history, this approach is 'backward' looking, contrary to the second functional account that we present, the dispositional approach, that is 'forward' looking.

The *dispositional approach* does not recur to any selective advantage to justify the existence of a trait's function, and just supplies some criteria to identify a function. There are several version of this view. In the 'systemic approach' (Craver 2001), for example, the criterion is the ability of a process to contribute to some higher-level capacity of the system: that such capacity is or not useful to the system is irrelevant to the identification of the function. The view is criticised, among other reasons, for being unable to distinguish dysfunctional or irrelevant effects: by avoiding the teleological dimension in the definition of functions, the systemic approach is unable to ground their normative side (Mossio et al. 2009).

A third account has been recently proposed: the *organizational account* of functions (Mossio et al. 2009). Based on the consideration of organisms as complex, self-maintaining systems, this approach justifies the existence of a function based on its contributions to the system's self-maintenance, and grounds its normativity on the actual efficacy of such contribution. According to these authors, this approach solves both the teleological issue

(why a function exists in a system) and the normativity issue (how a functions contribute to the system) of functions, and also allows to distinguish between real functions and useful contributions or simple side-effects.

In this paper and in the model we will propose, we consider function under the dispositional approach. The inability, typical of this view, to ground the teleological dimension of function is a virtue in our model, because the existence of a function and its relevance with regard to the attaining of some end are two different logical steps in our proposal. Baldwin already noted that, in the early evolution of a trait, structure and function are only partially correlated (Crispo 2007). Traits and functions may arise in a population for several reasons (as it will be shown in Ch. 4) that sometimes have nothing to do with their contribution to survival and reproduction. For example, phenotypic changes could arise "through the alterations in developmental mechanisms (the physical interactions among genes, cells and tissues), whether they are adaptive or not" (Müller 2007). The strength of selective forces will determine whether their fate will get fixed or disappear, or whether they remain in the population as a residual character waiting to be taken advantage of as soon as the environment allows it. A function may contribute to survival and reproduction, but this empirical observation cannot ground the identification of any biological function, only of adaptive ones. A similar critique could be advanced against the organizational approach. As McShea and Venit suggest: "[…] functionality and partness are independent notions" and "[…] an essential organ might be a part, but nonfunctional structures, or even deleterious ones, can also be parts" (McShea & Venit 2005). Moreover, the identification of adaptive functions does not explain their existence based on their contribution to survival and reproduction. If two species developed alternative ways to perform the same function, and the survival of one over the other was determined purely by chance, adaptation by selection would be no explanation for the existence of the trait (Cain 1990). The 'why' of a trait can explain its maintenance, but not its existence in the first place: in some sense, all adaptations are really ex-aptations! Adaptationism has so deeply invaded biology that all areas of research are trapped within its paradigm: fitness and function are just two examples. If these concepts are not purified first, any alternative proposal for explaining the appearance of characters is doomed to fail.

The choice of the dispositional interpretation of function is instrumental to its use in our model. It does not entail an ontological defence of this approach, just a semantic clarification and a pragmatic use.

2.4.3. A model to classify evolutionary complexity changes

Once defined any evolutionary phenomenon as a change in architecture, we can map and track it on a bi-dimensional space like the one in Fig. 2.13 - E, whose details are explained in this section.

Morphology axis

The element of the morphology axis are inspired by Müller's classification: form can be defined operatively by focusing on the constituents of phenotypic changes, such as "the generation of new structural elements (novelties), the establishment of standardized building units (modularity, homology), the arrangement of such units in combinations (body plan)." (Müller 2008: 19). Analysing an organism as the sum of elements always entails some bias in the fragmentation of the whole into parts; however, it can be justified by the general acknowledgment that "evolution proceeds in a mosaic fashion, modifying some characters but not others in a certain lineage" (Schlosser 2005: 143). McShea and Venit (2005) suggest a method to identify such parts, by checking if the object is (i) enclosed in a definable region, (ii) different in composition from its surroundings. Fig. 2.9 lists the identified traits of the marine invertebrate *C. elegans* identified applying their method.

Ciscipora elegans

candidate part name	status (larger part)	reasons for classification
annular muscle	part	enclosed, contiguous, different comp.
anus	nonpart	empty space
anal sphincter muscle	part	enclosed, contiguous, different comp.
attachment ligaments	nonpart* (attachment organ)	enclosed, contiguous, not different comp.
attachment organ	part	enclosed, contiguous, different comp.
cardia	nonpart* (digestive tract)	enclosed, contiguous, not different comp.
caecum	nonpart* (digestive tract)	enclosed, contiguous, not different comp.
communication pore	nonpart	empty space
digestive tract (excluding pharynx)	part	enclosed, contiguous, different comp.
epidermis	part	enclosed, contiguous, different comp.
funiculus	part	enclosed, contiguous, different comp.
laminar skeletal layer	nonpart*(skeleton)	outer structure, contiguous, not different comp.
lophophore	part	outer structure, contiguous, different comp.
loph. retractor muscle	part	enclosed, contiguous, different comp.
membranous sac	part	enclosed, contiguous, different comp.
mouth	nonpart	empty space
orifice	nonpart	empty space
perisome	nonpart* (skeleton)	outer structure, contiguous, not different comp.
pharynx	part	enclosed, contiguous, different comp.
pylorus	nonpart* (digestive tract)	enclosed, contiguous, not different comp.
rectum	nonpart* (digestive tract)	enclosed, contiguous, not different comp.
skeleton	part	outer structure, contiguous, different comp.
tentacle	nonpart* (lophophore)	outer structure, contiguous, not different comp.
tentacle sheath	nonpart* (lophophore)	outer structure, contiguous, not different comp.
transparent skeletal layer	nonpart* (skeleton)	outer structure, contiguous, not different comp.
vestibular wall	nonpart* (lophophore)	outer structure, contiguous, different comp.

Fig. 2.9 – Parts of C. elegans (from McShea & Venit 2005)

An architectural change can, with regards to morphology:

1. leave existing traits unaffected, with only quantitative (e.g. size), qualitative (e.g. colour) or relational changes among them (Fig. 2.10). This is by far the most common type of change: "[m]uch of phenotypic evolution consists of changes of shapes and proportions." (Müller 2008: 8). A trait (or character) is a physical structure constituting the organism (e.g. an eye), including processes such as metabolic cycles or behaviours, and it is different from the attributes of the trait (e.g. the eye shape or colour), that describe or measure it (Colless 1985).
2. entail a new trait (Fig. 2.11a): a novelty, in our meaning of the term (Erwin 2015);
3. create a new module (Fig. 2.11b): an arrangement of traits forming a standard building unit, which is part of a system and whose elements have ties among themselves stronger than the ties they have with other parts of the system (Rasskin-Gutman 2005). Modules are units in the development of an organism (Wagner & Altenberg 1996);
4. require a new body plan (Fig. 2.11c): a new arrangement of modules, "shared by a group of phylogenetically related animals at some point during their development." (Willmore 2012). The idea of body plan is associated with the concept of phyla (Valentine 2004). As there are just about 35 phyla (disagreements exists about the exact number), such as chordates, echinoderms, molluscs, etc., changes in body plan are major macro-evolutionary changes.

One can think of evolutionary changes involving modules and body plans as events of macro-evolution[38], as these morphological concepts are linked to the level of species' differentiation; and of changes in traits as events of micro-evolution, or fine-tuning. This is illustrated by the red and blue rectangles in Fig. 2.13.

[38] "The patterns and processes pertaining to the birth, death, and persistence of species" (Lieberman & Eldredge 2014: 185)

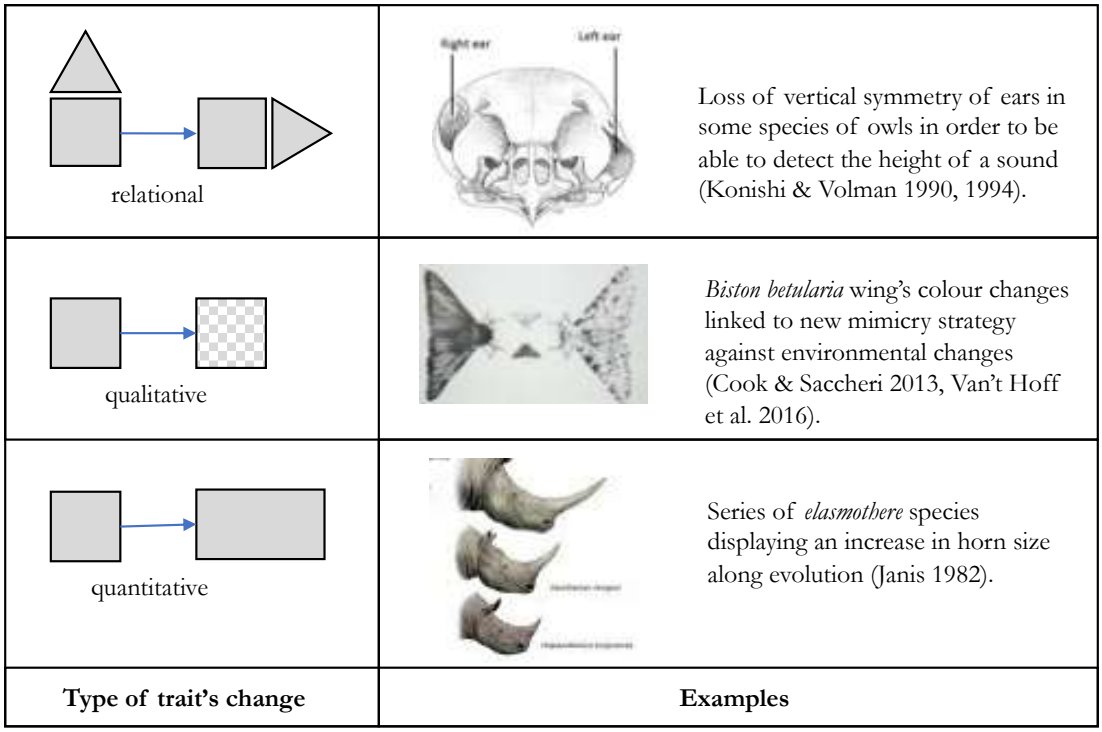


Fig. 2.10 – Types of current trait's change in evolutionary phenomena. Relational: unchanged single traits relate to each other in a new way. Qualitative: the morphology of the trait is unchanged, but its qualitative characters do change. Quantitative: some measurable characteristic of the trait changes its value.

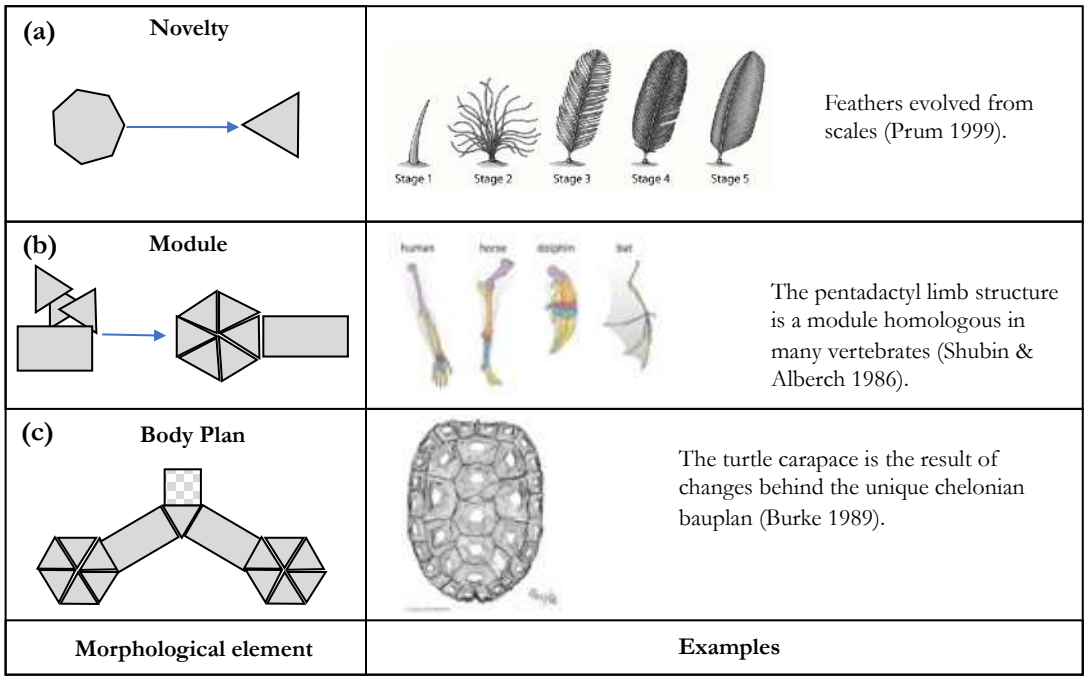


Fig. 2.11 –Types of morphological changes in evolutionary phenomena. (a) A trait is modified and transformed in a new one: a novelty. (b) Traits (triangles and rectangles) are arranged into a module through ties among themselves stronger than ties with other parts. (c) Modules are arranged into new body plans.

Functional axis

The functional axis of our model considers that an evolutionary phenomenon might:

- leave an existing function unaffected at time t[39]. A change in the qualitative value of a trait (e.g. wing's colour of *B. betularia*) can just insure the maintenance of the function of mimicry, without introducing any new one. The same can be said of redundancy for duplication of an organ. It's a '*do the same in the same way*' situation. A novelty might also have no function, as non-functional structures can also be traits (McShea & Venit 2005): the human chin is an example. We classify such cases in this category because, in some way, the lack of function is maintained before and after the architectural change.
- incrementally improve some ability already owned by the organism. We can refer to this situation as '*do the same in a new way*'. New metabolic routes allowing the exploitation of new energy sources are an example of such improvements.
- entail a completely new function: the motto is here: '*do something new*', as in the thermoregulation provided by feathers or the ability of flying.

We can now apply our model to some known evolutionary changes and place them in the map:

- The polymorphism of the *B. betularia* does not involve a change neither in form nor in function: wing's colour, either white or black, insures mimicry against predators. We can hardly affirm that an architectural change has happened at all.
- Exaptations are a change of function, but not of form: the new function is performed by an existing trait40. Under the view of complexity as part types (McShea 1991, 1996), or part numbers (McShea & Brandon 2010), an exaptation is therefore not an increase in complexity. According to our proposal, on the other hand, it does represent a change in architecture.
- The evolution of the cytoskeleton is another example of functional innovation based on a previously existing phenotypic trait (proteins in this case), without change in form (McShea 2001), as illustrated by the following excerpt:

[39] We specify 'at time t' because any change, even though initially functionally irrelevant, might become functionally relevant at a later stage 't+1' (e.g. a duplicated organ might later give rise to new functions). Such later development is considered as a different evolutionary phenomenon in our model.

[40] See note 235

> Although they are *constructed from homologous proteins*, the functions of prokaryotic and eukaryotic filaments are not broadly homologous. In eukaryotes, DNA segregation is ubiquitously performed by the tubulin-based cytoskeleton, whereas cytokinesis involves actin–myosin. In contrast, most prokaryotic cytokinesis is based on the tubulin homologue FtsZ, whereas actin-like, tubulin-like, or WACA proteins may be used for DNA segregation. This suggests that *a switch must have occurred in cytoskeleton function* in the proto-eukaryote. (Wickstead & Gull 2011, italics mine)

- The evolution of forelimbs in mammals from a common tetrapod ancestor represents a change in architecture based on new functions and new modules. The salient elements of this phenomena are (Brandon 2005: 53):
    - The earliest mammal was a tetrapod with not much differentiated forelimbs and hind limbs (functionally and developmentally).
    - The forelimb has evolved into flipper (whales), wings (bats) and hands (humans), apt for the innovative function of swimming, flying and manipulating objects.
    - The forelimb has evolved as a new module, without changing the whole body-plan, as whales, bats and humans has similar circulatory systems. This is clear also because forelimbs have evolved independently of hind limbs (bats do not have two sets of wings, nor human two sets of arms).
- Redundancy attained through the duplication of an organ (e.g. two kidneys) represents a change in morphology (basically a new module, although it could be claimed that the body plan is modified as well), without functional innovation.
- Redundancy attained through the creation of alternative ways to perform the same or improved function (e.g. alternative metabolic routes) does entail a modification of form (new trait).
- Stereopsis (stereoscopic vision) represents an example of redundancy by duplication that, at the same time, improves the function of vision and modifies morphology through a new body plan. The same is true regarding the redundancy represented by having two symmetric ears, that supports the new functions of identifying directionality of sounds and balancing oneself.
- Loss of symmetry in the ears of many species of owls allows to identify the elevation of sounds with respect of eye level: noises originating below or above eye level sound louder in the right or left ear, respectively. The asymmetry is reached in three different ways in different species, pointing to an independent appearance of the trait, due to an

important adaptive advantage: locating prays in total darkness (Konishi & Volman 1990, 1994; Payne 1971). Like the case of stereopsis, this is a new function, but linked to a loss of symmetry, although not of redundancy. It is clearly an improved function, and not a completely new one, as the ability to identity the horizontal direction of a sound already existed; and it is the same trait.

- The ability to fly has been reached through substantially different paths by insects and dinosaurs.
    - In insects, wings appeared as thermoregulatory mechanisms, and the transition from aerodynamic function involved only isometric changes in body size, while "changes in body form were not a prerequisite for this major evolutionary change in function" (Kingsolver and Koel 1985). A radical new function appeared without any relevant change in traits[41].
    - In dinosaurs, the ability to fly is linked to the appearance of feathers as thermoregulatory tool, and their exaptation to fulfil an aerodynamic function. The evolutionary phenomenon was more complex than in case of insects, and involved several intermediate steps entailing a deep modification of the whole body plan (Prim 1999).
- The turtle carapace involves a new body plan and a new function (Burke 1989).
- Non-functional traits have their place too in the map: the human chin is a new trait with no apparent function.
- Some evolutionary phenomena might involve more than a change at a time. The evolution of toes in horses involved at the same time quantitative changes in the measures of the limb's bones and their re-arrangement in a new relational structure (Fig. 2.12), while the already existing function (locomotion) incrementally improved to support higher bone stress linked to increased body mass (McHorse et al. 2017).

[41] In fruit flies, wings also underwent a secondary exaptation, and their vibration are used to produce a 'courtship song' (Jablonka & Lamb 2005: 321).

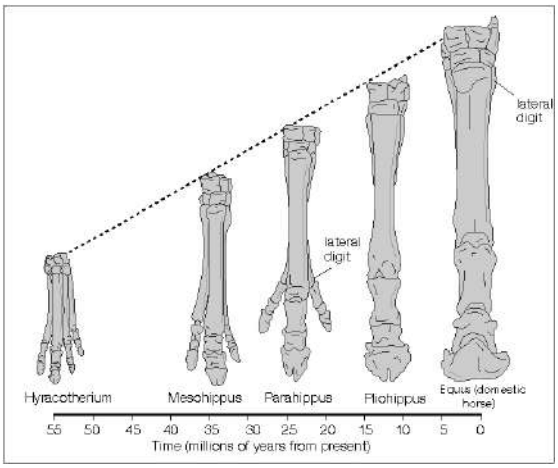


Fig. 2.12 – Toes evolution in horses as an example of simultaneous quantitative and relational changes in morphological traits.

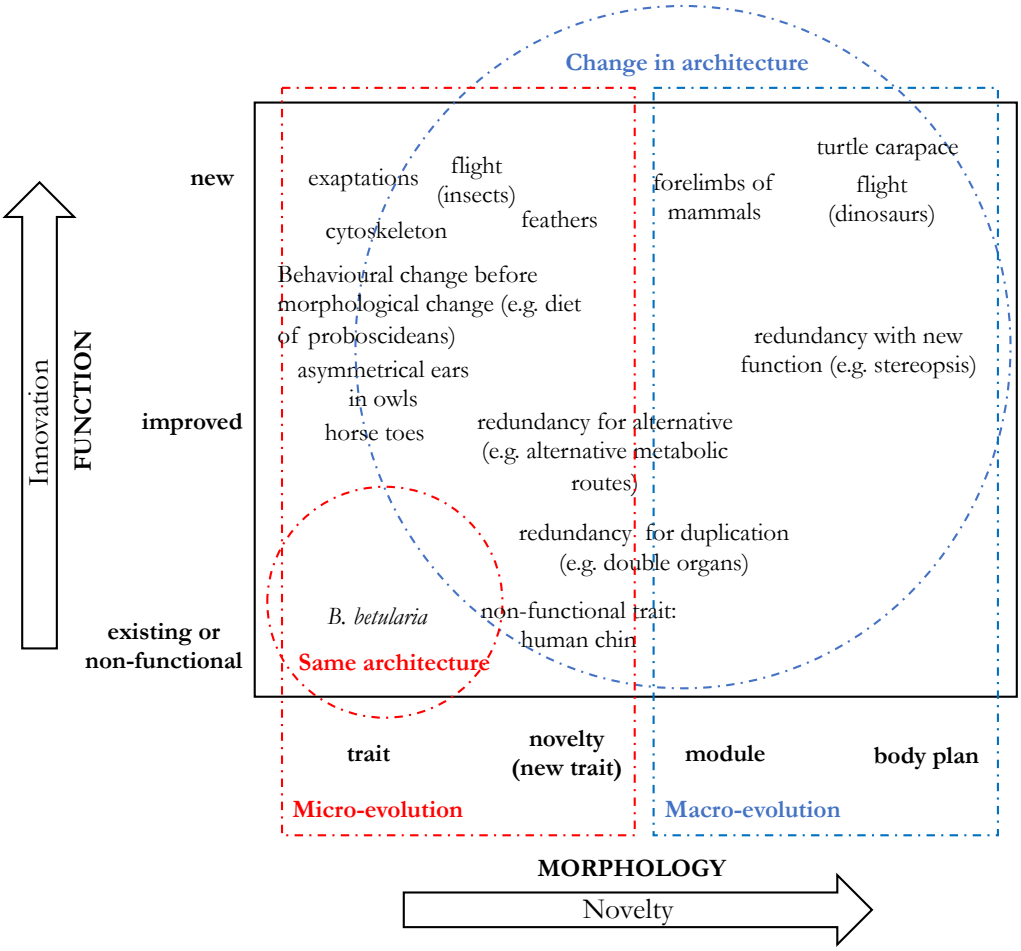


Fig. 2.13 - Evolutionary phenomena in organismic architecture. Changes along the morphology axis are 'novelty' and changes along the functional axis are 'innovations' (Erwin 2015). Null or small changes in both morphology and function keep architecture unaffected (small red circle), and are mainly linked to fine-tuning of the organism through selection, without speciation. On the other hand, relevant changes in any of the two dimensions entail changes in architecture (big blue circle), and might entail macro-evolutionary changes and new species. Note that micro- and macro-evolution classification is based only on the morphological axis, while the architectural classification involve both axis: the two classifications do not coincide (e.g. micro-evolution might or not entail architectural change).

Note that no hypotheses are made as to what caused the architectural modification: the model is agnostic with regards to the role of selection or other forces in driving change. This distinguishes our proposal from other classifications found in the literature. Schlosser (2005: 143), for examples, lists several evolutionary changes that can alter the morphology of an organism by affecting some of its modules: (1) loss of a module; (2) shift in timing of development of one module with respect to others; (3) shift in localization of one module with respect to others; (4) redeployment of a module in additional contexts; (5) other correlated trait changes (size, shape,…) among modules. These kinds of classifications mix morphological changes with their causes, as it happens with point (2), which can be a cause of any of the other points, e.g. heterochrony can cause point (5) -changes in size of a module with respect to others-. Such mixtures of causes and effects are confusing, and we prefer to separate the classification of changes from the evolutionary forces behind them. Like in the case of heterochrony, modifications could spring from events affecting development (see Linde-Medina & Newman 2014, fig. 1, for qualitative/quantitative trait changes during development that can be reflected in the adult phenotype). Applying a physical metaphor, ours is a 'kinematic' analysis (the study of movement without inquiring for its causes) of evolutionary phenomena, while their 'dynamic' side (the study of the causes of movement) is introduced in the next two chapters, and a model to unify them is proposed in the last one.

2.5. But… does biological complexity increase in the history of life?

Darwin's idea about 'progress in organization' contains two independent claims: the first being that 'organization' progress is a general trend (as opposed to punctual events of complexity increase) in the biological world; the second being that natural selection is the main cause of this trend. The first claim is shared by several evolutionists like Lamarck, Cope, Spencer, but also increasingly questioned by many others, like Lewontin and Gould (McShea 1996). The second claim is defended mainly by adaptationism, and increasingly questioned by several alternative theories (e.g. complex systems' theory or Evo-Devo[42]). Adaptationist arguments supporting the idea of a trend towards complexity appeal to the higher efficiency of more complex organizations, the need for more complex organisms to

[42] See Ch. 4 for an articulated list of such accounts.

face more and more complex niches, and the consideration that it is less disruptive to add parts than to delete them (McShea 1991, 1996[43]). The two claims are independent because it is possible to defend one of these two ideas while criticising the other. For example, both adaptationist accounts and population genetics models defend the primacy of natural selection as a source of phenotypic novelties; however, the former believes in a trend towards better designs, while the latter is to some degree agnostic regarding it, at least in its mathematical formalizations. On the other hand, both adaptationist accounts and alternative theories usually defend complexity increase, but they grant natural selection a different weight in explaining it. In this section, we treat the issue of complexity's putative increase; in the next one, the issue of the relationship between complexity and selection.

A closer look at the claim of complexity increase reveals that it is not trivial to formalise. Additionally to the already mentioned problem of defining complexity, there is the issue of how to measure its eventual increase: is it a general trend affecting the majority of species? If just a few lineages become more and more complex, while the remaining ones do not, can we still claim that complexity increases in the history of life? And what if half of the known lineages show increasing complexity, while the other half shows decreasing complexity? Increases in complexity in some lineages, however this is defined, could be counterbalanced by decreases in other lineages (McCoy 1977). Body size in horse evolution is an example of such counterbalanced trends: although it increased in some clades during the last 20 million years, it remained constant or even decreased in others (MacFadden 2005). Fig. 2.14 shows some of these cases: in (a), a general trend toward higher complexity; in (b), an increase in maximum complexity ($Cmax_1$ to $Cmax_2$) counterbalanced by a decrease in lower complexity ($Cmin_1$ to $Cmin_2$), so that average is unaffected. What to measure (Y axis) is also an issue, even if complexity is clearly defined: should we count species? Or individuals of each species?

[43] McShea also suggests that evolution can progress by subtraction (McShea & Hordijk 2013): genome reduction is such a case (Morris et al. 2012). Thus, there is certainly a trade-off between accumulation of parts and their deletion (thanks to Davide Vecchi for this note).

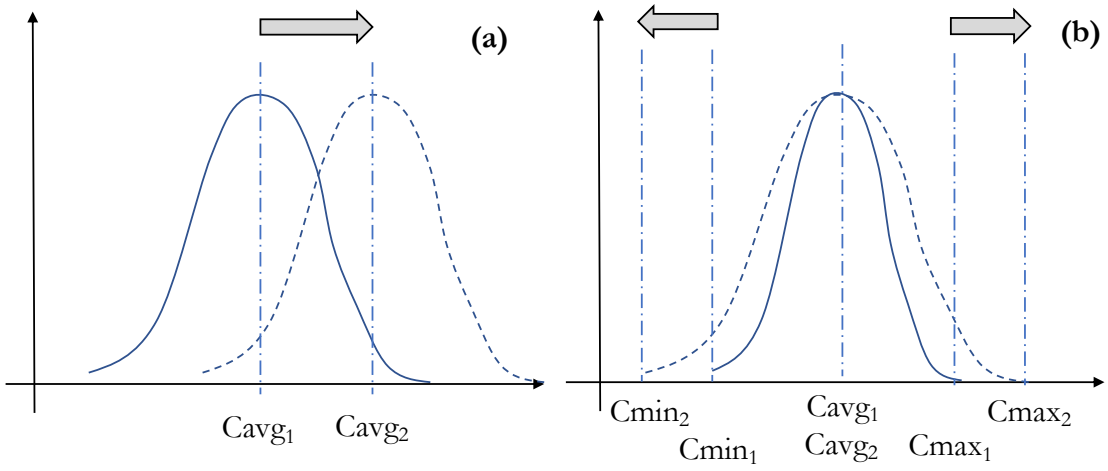


Fig. 2.14 – Different interpretations of complexity increase, as increase in average (a) or in maximum value with constant average (b). The Y axis can be interpreted as number of species or number of individuals for each level of complexity, the X axis measures complexity.

A similar issue has to do with hierarchical levels, as McShea et al. (2018) underlines: "An upward trend in cell types does not mean that there was an upward trend in complexity at the level of organs, tissues, or molceules." (p. 12)

There are several definitions of evolutionary trends in the literature. Gregory (2008b) suggests that it consists of "identifiable patterns in which the overall evolution of a trait occurs in a given direction within a group for a prolonged period of time". He admits that his definition is less technical than others like McShea's, i.e., "a long-term directional change in a summary statistic for a clade, such as the mean" (McShea 2005), Gould's "directional character gradient through time in a well-defined monophyletic clade" (Gould 1990), or McKinney's "persistent statistical tendencies in some state variable(s) in an evolutionary time series. Such variables may be point estimates (e.g. mean, maximum) of a group (e.g. cladogenetic, concerning a number of species) or a single lineage (anagenetic, concerning a number of individuals in a species)." (McKinney 1990).

Some population genetics accounts defend that complexity increases: "Darwinian evolution […] has gracefully produced organisms of vast complexity" (Adami and al. 2000). To do so, as we have seen, they reduce phenotypic complexity to the complexity of the genotype: "[We assume that] an organism's complexity is a reflection of the physical complexity of its genome" (Adami and al. 2000: 4463). Computational simulation shows that genomic complexity, measured with regards to the information stored in the genome, must increase if faced with a challenging environment and a series of mutations (Adami and al. 2000); but horizontal genomic complexity (number of genes) is not always a good proxy of

organismal complexity (see comments to Shannon entropy at §2.4.1). Snaders and Ho (1976) believes that organisms performing more functions need more complex structure to survive. Bonner (1998), an extreme adaptationist, links complexity increase to the evolutionary trend toward bigger sizes, that he attributes to selection, as if to be bigger mean more success both as a predator and as a prey. Bigger size fosters cell differentiation (due to division of labour), so he concludes that evolution usually progresses by increases in complexity caused by natural selection.

Complexity increase is not always explained as output of adaptations. Maynard Smith and Szathmary (1999) defend the same idea without linking it with selective forces, and measure such increase as the length of the shortest list of instructions able to generate a complex structure (a concept similar to algorithmic complexity and logical depth as discussed in §2.4.1), list that they identify with the number of genes of an organism (Table 2.4).

Table 2.4

Species	Type	Approximate gene number
Prokaryotes		
Escherichia coli	Bacterium	4000
Eukaryotes (except vertebrates)		
Oxytrochis similis	Ciliated protozoan	12 000–15 000
Saccharomyces cerevisiae	Yeast	7000
Dictoyostelium discoideum	Slime mould	12 500
Caenorhabditis elegans	Nematode	17 800
Drosophila melanogaster	Insect	12 000–16 000
Strongylocentrotus purpuratus	Echinoderm	<25 000
Vertebrates		
Fugu rubripes	Fish	50 000–100 000
Mus musculus	Mammal	80 000
Homo sapiens	Mammal	60 000–80 000

Number of genes in various organism (from Maynard-Smith & Szathmáry 1999: 16)

Doubts about the effective trend towards higher complexity are widespread. Gould (1980) notices that the hypothesis that changes are directed towards higher level of organizations is not a clear statement of Darwin, who seemed to have different ideas on the subject at different times. In a letter to Alpheus Hyatt late in his life, Darwin stated that "after long reflection I cannot avoid the conviction that no innate tendency to progressive

development exists" (Gould 1977). Haldane (1958) believes that, for any case of increase in complexity in evolution, there have been ten cases of regressive evolution. McShea denies that there is clear empirical evidence of a general complexity increase, although there is an abundance of proposed mechanisms to explain it, and calls for a change of focus from theoretical explanation to empirical studies, based on a formal and measurable definition of complexity (McShea 1991). The only aspect of complexity that seems to show an increasing trend is hierarchical object complexity (see) (MsShea 2001).

Finally, we clash with the issue of explaining how 'progress' or 'increase' should be interpreted. Williams (1966) lists several possible categories of progress:

- Accumulation of genetic information;
- Increasing morphological complexity. Although probably the most intuitive interpretation of 'complexity increase', it is not clear in what it should consist of, considered that a human skull is simpler than a fish skull (for example, in terms of number of parts);
- Increasing histological differentiation;
- Improvement in effectiveness of adaptation. The lack of an objective criterion to measure such effectiveness makes it of little practical use; moreover, ancient species continue to survive, so its scope cannot be universal.

In conclusion, there is no agreement among biologists and philosophers of biology about whether a tendency towards more complex organisms exists and whether it involves all lineages if it does, nor about how it should be measured and what forces would drive it. From one side, this is no surprise, given the equivalent disagreement around the definition of complexity itself. From the other side, our approach to architectural change does not rely on any hypothesis of complexity trend, as it just describes evolutionary phenomena regardless of whether the final result is a progression or a regression with respect to the initial organism. We can thus afford to maintain an agnostic position in this area. Surely in some cases, and in some sense, some organisms at least seems to be more complex than others: but whether this acknowledgment has any philosophical or biological value is a reflection that falls outside of the scope of this work.

2.6. Conclusions

In this chapter, we have discussed two of the biological *explananda* that the Darwinian theory addresses and putatively solves: the variety of living beings at all levels (among individuals of the same population, among populations of the same species, and among species); and their complexity (however defined), that seems to increase along the history of life[44]. We have briefly discussed how population genetics, by subtly changing some of the concepts used by Darwin, successfully formalises the first of these issues, and is able to explain the distribution of traits in a population and to forecast its evolution given the presence of some evolutionary forces. On the other hand, population genetics does not directly address the issue of complexity, and especially of phenotypic complexity: something that adaptationism does, by justifying phenotypic traits in terms of their contribution to environmental challenges, but at the cost of renouncing to the same level of formalism than population genetics. One of the reason of this uneven formalization of the two Darwinin claims is that complexity, more than variation, is a polysemous term, and hard to formalise itself.

We have introduced two approaches that aim at overcoming this situation and at explaining the issue of complexity. Formal adaptationism sticks to the idea that selection is a sufficient explanandum, and extends population genetics models to include phenotypic variables to demonstrate it. Pluralism believes that the source of complexity has to be found in other evolutionary forces, not mainly in selection. The following two chapters illustrate with some detail these two approaches.

In the last part of this chapter, we have proposed a definition of organismal complexity and we have shown how it can be used to classify evolutionary phenomena, in order to clarify what exactly is meant when one talks about 'change in complexity' in the history of life. We have also discussed the issue of whether complexity increases in the history of life, and we have clarified that our focus is on complexity changes and on the evolutionary forces that affect them, and not on the direction of the change. Our definition will be applied in chapter 5 to build a model that unifies the formal adaptationist and the pluralist approaches.

[44] The third *explanandum* being the fit of form and function (Brown 1991)

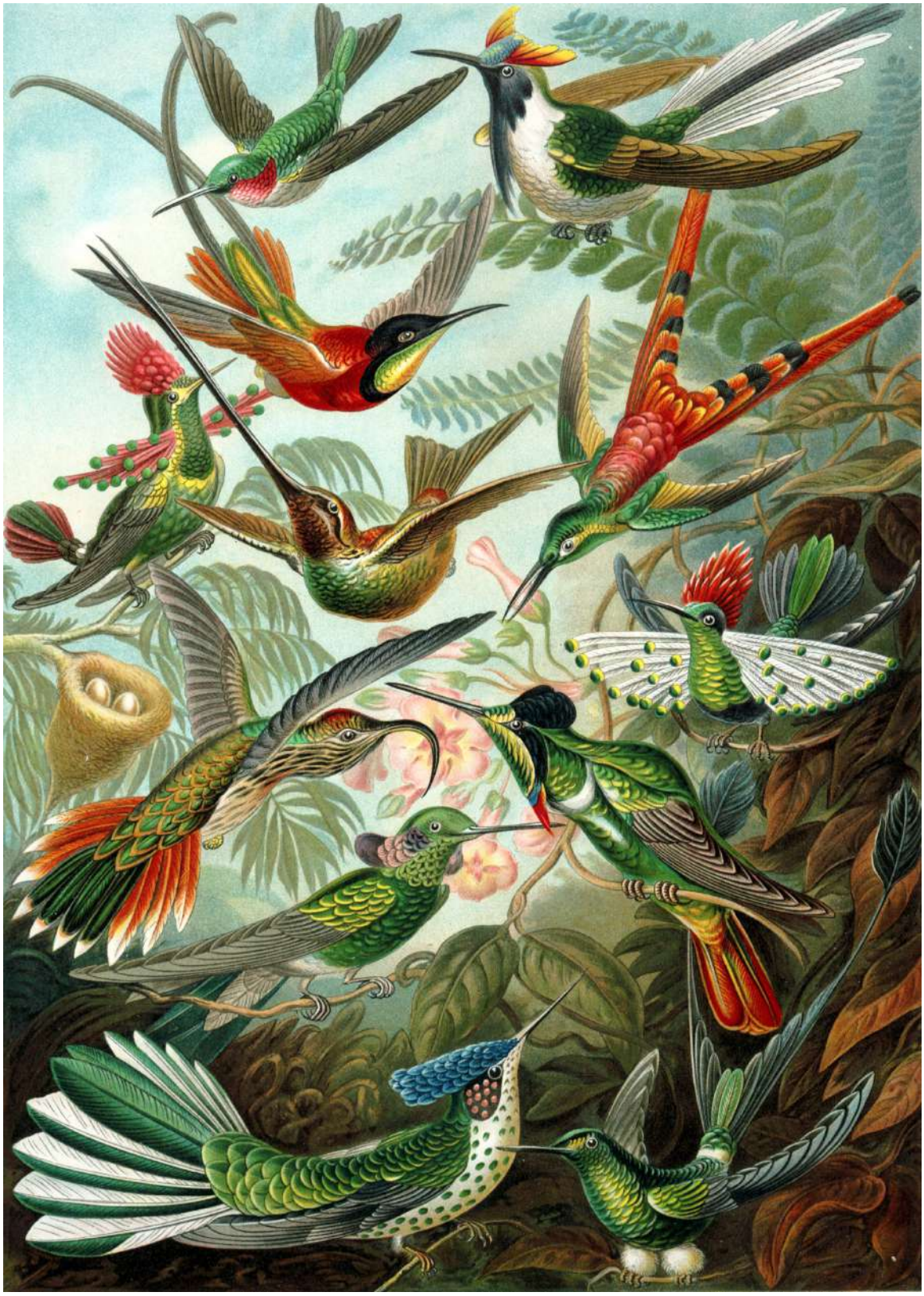


Plate 99 – Trochilidae, from Haeckel, E., *Kunstformen der Natur*. Leipzig und Wien: Verlag des Bibliographischen Instituts,1899-1904.

3. Adaptationism

This section focuses on the family of evolutionary explanations that consider selection as the preeminent or sole force behind evolution: a view known as adaptationism. Adaptationism accounts for phenotypic traits by invoking the optimization of some kind of fitness property (individual, inclusive, etc.), via the optimization of a trait or of a group of traits. Although purely narrative at the beginning (starting with Darwin's beautifully written accounts in chapter III of *The Origin*), various formalizations have been proposed in order to provide a sound mathematical model of adaptationist explanations and to make them practically testable.

Adaptationism can be interpreted in different ways: we start this chapter with a brief description of its main flavours (§3.1), and a more detailed analysis of its most extreme version, empirical adaptationism (§3.2). After discussing how the original Darwinian thought has been interpreted and fitted within current adaptationisms (§3.3), we illustrate in what sense adaptatonism is a functionalist and externalist attitude (§3.4). Finally, we briefly introduce two of its formalizations: optimization programs (§3.5) and the Formal Darwinism Project by Alan Grafen (§3.6). Based on our critiques to the use of fitness in these models, we suggest a way to unify different concepts of fitness (§3.7), and we conclude by underlining some limitations that make formal adaptationism models blind to a relevant range of evolutionary phenomena (§3.8).

3.1. Kinds of adaptationism

Broadly speaking, adaptationism is a collection of schools of thought that explain the existence of phenotypic traits through narratives based on the evolutionary history of the species, illustrating how each trait was selected thanks to its contribution to the fit between the individual and the environment: it is "the myth that all of evolution can be explained by adaptation" (Lynch 2007). Sober (1998a) identifies adaptationism with the claim that "natural selection has been the only important cause of most phenotypic traits found in most

species", while constraints are irrelevant and "*nonselective processes are ignored*"[45] (Sober 2000: 123, italics mine). These ideas are in line with Gould's claim (Gould 2002: 157) that, in Darwin's theory, adaptation is *the* central phenomenon of evolution, and natural selection is by far the most predominant and frequent cause of adaptation.

As it is the case with many concepts in philosophy of biology, however, adaptation is a polysemous concept, and different authors assign it a different meaning and apply it in different ways to tackle different problems. Although the versions of adaptationism could in principle be countless, a classification of three main logically-independent types has been proposed (Godfrey Smith 2001, Orzack & Forber 2010):

1. *Empirical adaptationism*. It claims that natural selection is a sufficient explanation for all traits, and that no constraints limit neither the generation nor the variation it acts upon, at least in the long run. According to the three ideas discussed in §3.1, this position treats selection as the only relevant creative force explaining traits, shaping them towards optimality, while constraints have, if any, just a negative, temporary role. The pillars of this view are discussed later in §3.2 and criticised in §4.1.2.
2. *Explanatory adaptationism*. It focuses on selection as *the fundamental answer* to adaptation, which is in turn *the fundamental question* of biology. It can be considered as an aesthetic claim about the world, more than a scientific one, and its importance is, for example, to counter the creationist arguments from design (Dawkins 2006, Dennett 1995). Alternative *explananda* in evolutionary biology could be, for example, the diversifying pattern of the history of life, or the complexity of living beings[46], and not adaptation and the forces influencing it.
3. *Methodological adaptationism*. It is a practical attitude that suggests resorting first to an adaptationist approach when trying to explain a trait: by supposing that the trait is optimized by selection, one can either confirm the guess, or look for other intervening forces by detecting deviations from the theoretical optima. It is a 'policy recommendation' (Godfrey-Smith 1999). According to Sarkar (2005), it is the a priori methodological principle implied by the hypothesis that natural selection is always the most important process of evolution and that it always leads to an optimal fit with the environment.

[45] Chapter 4 is devoted to illustrate some of these processes.

[46] Brown (1991) suggests that Darwin aimed at explaining no only the fit of form and function, but also the variety and the progression of the tree of life.

It is easy to see that none of these positions entails the others (Orzack & Forber 2010, Godfrey-Smith 1999). Empirical adaptationism does not make any claim about what questions are the most important in biology. Explanatory adaptationism does not claim that natural selection is the most powerful evolutionary force, only that it is the only relevant one to explain design and adaptation. Methodological adaptationism does not suggest anything about fundamental questions nor about fundamental forces, it only suggests starting any investigation with an adaptationist hypothesis.

In the remaining of this chapter, we refer to adaptationism in its strong (empirical) version.

3.2. The four pillars of empirical adaptationism

Empirical adaptationism relies on the following four general assumptions[47] (Godfrey-Smith 2001, Orzack & Forber 2017):

3.2.1. Pillar I - Ubiquity of adaptation

The action of natural selection is ubiquitous, i.e., it acts on all phenotypic traits (Gould 1976): all traits of an organism are (or have been, or will be) adaptations. Fisher underlines the marvellous level of adaptation of all aspect of organisms (1930:41). Huxley considers organisms as 'a bundle of adaptations' (1942: 420). Mayr (1983) believes that every trait can eventually become a target of selection, either actual or potential, and claims that "one can never assert with confidence that a given structure does not have selective significance" (Mayr 1963: 190). Selection acts on heritable variation because it is blind to non-varying traits (Amundson 2001). Lethi et al. (2009) represents a nice example of the 'ubiquity' approach, when they take for granted that a trait with no apparent ecological role must have had one in the past, and when they claim that uselessness should be declared with caution, as "new and previously secondary functions are not always easily identified" (2009: 492).

[47] The triadic classifications of 'adaptationisms' is of course a simplification, as each of the three views come in degrees too. Our account of empirical adaptationism reflects a theoretically extreme position.

3.2.2. Pillar II - Continuous, slow and incremental evolutionary change

Organismic changes between generations are continuous, slow and incremental, without major jumps (a view known as 'gradualism' , Gould 1980, 2002: 146-155):

- "Continuous" means that there is an "unbroken historical connectedness between putative ancestor and descendant", otherwise the fact of evolution would be challenged, as "[i]f new species originate as creations ex nihilo by a divine power, then connectivity fails." (Gould 2002: 149). Gregory (2008), for example, assumes that "evolutionary explanations for the origin of biological features […] are based on the assumption of continuity" or "an unbroken chain of ancestry and descent" (p. 359);
- "Slow" means that change occurs on (geologically) long time-scales. Cain (1989) claims that "if we allow a time-lag of twenty-million years […] we still have plenty of time from the Cambrian to the present day to completely remodel the older groups if it were necessary." To account for phenomena like the similar rate of change at the molecular level of genes and pseudo-genes, or of introns and exons (Kimura 1983)[48], seemingly undermining a faster evolution of molecules under strong selection, adaptationism would claim that these differences will tend to even out over very long geological time frames (Gould 1993);
- "Incremental" means that any novelty is built in a long series of small, cumulative steps. Speciation is a gradual process, while saltations (e.g. macro-mutations) are very rare: macromutations are usually lethal, while small mutations, given the additive and linear relationship of genotype to phenotype, translate in small incremental changes in the latter (Fisher 1930, Fodor & Piattelli-Palmarini 2010). Selection is 'creative' because it allows such cumulative selection of tiny variations, accumulation that would otherwise be very improbable (Razeto-Barry & Frick 2011). Darwin, as we have already mentioned, considered that his theory would collapse if any complex organ existed "which could not possibly have been formed by numerous, successive, slight modifications" (1859:189). Gould (2002: 144-146) specifies that the existence of phenotypic macro-changes would not invalidate the theory of selection, but

[48] Pseudo-genes and introns are genetic entities that are not involved in phenotypic expression and therefore cannot be target of selection. Their rate of inter-generational change due to random mutations should be higher than in the case of genes under selection (whose mutations rarely pass through to the next generations because they usually have lethal phenotypic effects), but it turns out that the rates are very similar. See §4.2.4 for a detailed account of this phenomenon and of Kimura's theory.

would relegate it to the secondary role of spreading the new phenotype, while the source of novelty would be the force of variation itself.

3.2.3. Pillar III - Random and unconstrained variation

The source of evolutionary variation is "random" with respect to adaptation (in the sense of not being biased towards beneficial variants production) and unconstrained (Gould 1976), and its source are genetic mutations, recombinations, etc.: "[…] studies of both mutation and recombination show that, although selection may lead to the adaptive modulation of the amount of variation, *there is no bias towards the production of beneficial variants*." (Charlesworth et al. 2017: 9, italics mine).

Gould summarises the requirement of being *random* as fundamentally meaning that there is no relationship between appearance of variants and direction of evolution (Gould 2002: 144), so that, for example, the environment cannot influence the appearance of beneficial mutations, but can only select them if they appear. Razeto-Barry and Vecchi (2016) highlight that the term random can have several interpretations based on the supposed absence of correlations between any two of the three relevant variables: environment, fitness and occurrence of mutation. The authors suggest that none of these 'dyadic' interpretations makes evolutionary sense: for example, occurrence of mutation and fitness do show a correlation within the neo-Darwinian paradigm, given that deleterious mutations are more probable than beneficial ones, as previous selection already 'quasi-optimized' phenotypes. The authors propose to apply the concept of 'independence' to all three variables at a time, instead of 'absence of correlation' between any two of them: a mutational mechanism is random "[…] if the occurrence of mutations is conditionally independent of the fitness effect of the mutations given the environmental variable." (Razeto-Barry and Vecchi 2016: 679).

The concept of *unconstrained* means that there are no unsurmountable physical or developmental limitations to the emergence of new variants. Adaptationism can admit the existence of constraints on variations in the short-term, but these are assumed to be eliminable by mutations in the long-term (Parker & Maynard Smith 1990): natural selection will solve any problem regardless of genetic constitution (Mayr 1963, Grafen 2014). There exists no bound to possible phenotypes, and this ability to produce new variants is iterative: evolved phenotypes are themselves susceptible of further evolution. This permits explanations *in media res*, that is, of the phenotypes we have to deal with at any time: for a

phenotype to evolve, the theory "presupposes some presumably very simple unevolved self-replicators with phenotypic traits to which the laws of evolution apply in the first instance" (Fodor & Piattelli-Palmarini 2010: 6). Gould (2002: 141-143) adds that variation must be "copious": only abundancy of variants insures that all potential evolutionary paths are explored and that new paths are always available. The requirements of being copious and incremental together ensure that "[r]andom variation supplies the indispensable 'fuel' for natural selection but does not set the rate, timing, or pattern of change." (Gould 1993: 542): random variations are therefore a raw material that does not show any *a priori* evolutionary significant structure or function - the task of shaping structures and functions being left to natural selection, as the following and last pillar states-

3.2.4. Pillar IV - Supremacy of selection

The adaptedness and complexity of organisms are shaped solely by natural selection[49]. Variation, independently of how it is generated, supplies the raw material only. One of the most prominent evolutionary biologists, Maynard Smith, claims that selection is the most important evolutionary mechanism and that it leads to optimal adaptation in any environment (Sarkar 2005). Brandon underlines that (1990: 175): "It is worth noting that presumably no serious biologist thinks that other [non-selective] evolutionary mechanisms, such as drift or pleiotropy, can produce complex and intricate traits that appear to be adaptations". Mayr claims that "all substantial change is adaptive" (Gould 2002: 521, note): natural selection is so important, and nonselective forces (e.g., saltationism and orthogenesis, Gould 2002: 507) so unimportant, that accurate explanations and predictions of the phenotypes of organisms can be obtained by simplified models in which only selection is represented and nonselective forces are ignored (Orzack & Sober 2001). Developmental mechanisms and processes, for example, (everything included in the black-box leading from genotype to phenotype, Hall 2003b), are considered irrelevant to the process of evolution. As Kaufmann summarizes it: "If we take selection as the *sole source of order*, it is because we have come to suppose that without selection there could only be chaos" (1993: 11, original italics). Even Kimura, while claiming that natural selection has little impact at the molecular level, affirms that "progressive evolution is almost always brought about as a result of

[49] Note that this is a different concept from the first pillar: *ubiquity* means that selection acts on *all traits*, but does not exclude that other mechanisms could also contribute to shape them. *Supremacy* means that *no other mechanisms* apart from selection shape evolutionary traits, but does not imply that all traits are adaptations.

organisms' response to environmental challenge", and that "evolution at the level of form and function is largely determined by Darwinian natural selection that brings about adaptation of organisms to their environment" (Kimura 1983:61-62). The phenotype is shaped by causal interactions with its actual environment while endogenous factors have a negligible role (Fodor & Piattelli-Palmarini 2010).

The pillars are logically independent: each can be true regardless of the truth of the others[50]. Although independent, we can organize the four pillars into two sets (Fig. 3.1):

- The first two (about ubiquity of adaptation and pace of evolutionary changes) are claims about where in the phenotype and how the theory of evolution by natural selection acts. They talk about the *internal mechanism* of natural selection, without committing in any way with what happens (if anything) beyond selection. For example, they do not deny nor affirm anything about other potential sources of evolution: that selection acts on all traits does not exclude the action of other processes.
- The last two (about random and unconstrained source of variations and supremacy of selection) are claims about where variations fuelling selection come from and what is final impact of selection in transforming such variations into the evolved phenotype. They define the *external conditions* in which natural selection acts. They do exclude other evolutionary processes, for example that developmental laws might pre-determine some evolutionary paths and preclude others. The fact they are external to the theory is clear if we apply them to other potential evolutionary forces: we could claim that drift is the only source of evolution, that is acts on random and unconstrained variations and that it wipes out all effects from other forces, including selection.

[50] See note 49

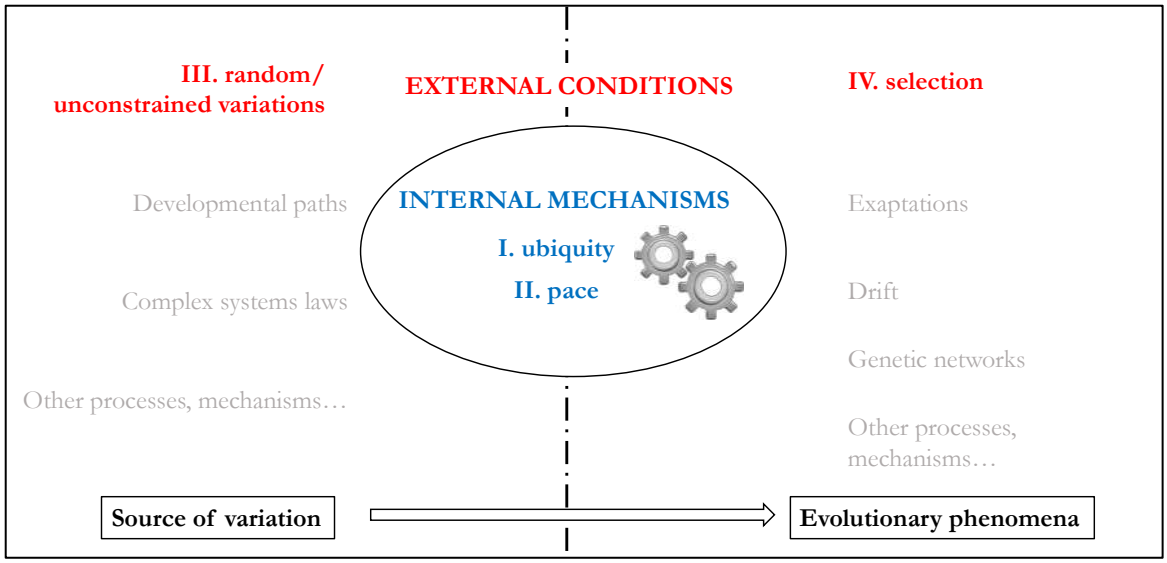


Fig. 3.1 – Conceptual classification of the adaptationist pillars. Pillars representing the internal mechanisms of the theory are in blue. Pillars representing external conditions to the theory are in red. Shaded text represents potential evolutionary forces and processes considered irrelevant by empirical adaptationism.

3.3. Was Darwin a Darwinist?

One of the reasons behind the polysemy of the concept of adaptationism can be traced back to the historical interpretations of Darwin's ideas. As it happens with many '-isms' (consider the difference between Marx's thought and Marxism, or Christ's teachings in the gospels and Christianism), there is a gap between what Darwin's followers, defenders and interpreters have been claiming during the last 150 years, and what Darwin actually wrote (see also Razeto-Barry & Frick 2011). To show this, it is useful to compare Darwin's thought with later interpretations of it, especially around three implicit claims hidden behind the umbrella-idea that all traits are adaptations, and behind the adaptationist pillars, namely that: (1) selection is responsible for traits to appear in the first place, (2) traits are optimally fine-tuned thanks to selection, and (3) selection is able to dodge any possible obstacles to (1) and (2).

3.3.1. The creative power of selection

Darwin believed that:

> If it could be demonstrated that any complex organ existed, which could not possibly have been formed by numerous, successive, slight modifications, my

> theory would absolutely break down. But I can find out no such case. (Darwin 1872: 204)

Natural selection explains not only 'phenotypic prevalence' (i.e., why certain structures have become established), but also 'organic form' (i.e., why structures look the way they do) (Linde-Medina 2011): thus, both why traits get fixed or eliminated as well as how they appear in a population. The two phenomena are quite different, as exaptations[51] show. While the idea that selection preserves and propagates traits has predecessors, Darwin was the first to claim that it also creates them, and that it is the main source of complex trait.

The Darwinian idea that natural selection explains not only the distribution of traits within a population, but also their origin, is known as the 'positive view' (Neander 1995). Genetic mutation and, above all, recombination, just supply variation, but

> [E]volution is not primarily a genetic event. Mutation merely supplies the gene pool with genetic variation; it is selection that induces evolutionary change. (Mayr 1963: 613)

Novelties, as a consequence, are not due to mutations or other genetic changes, but

> [S]election pressure for it is generated by the appearance of novel challenges presented by the environment and by the ability of certain populations to meet such challenges. (Stebbins 1982: 160)

The validity of this claim has been defended more or less strongly. Dawkins (1986) claims that selection is like a 'blind watchmaker', and that it succeeds in solving problems by improving organisms through the testing of casual incremental changes and the accumulation of the ones that work. He accepts few limits to the creative power of selection and lists several examples of selection's creative power. Forber (2005) circumscribes this power only to traits affected by several factors, typically genes. His reasoning goes as follows: if a trait is coded by more than one gene, it is improbable that all these genes suffer a mutation in the same individual. If this happens, the new version of the trait is not really due to selection, but to the casual sum of all mutations. On the other hand, if there is a selection gradient such that even a mutation of a single gene results in a fitter trait, selection helps fixing the single mutations and accumulating them: once all genes have undergone mutation, the new version of the trait can be said to be the result of selection. Birch (2012) claims that

[51] Exaptation is a term created by Gould and Vrba (1982) to refer to traits that currently increase the chance of surviving of the organism possessing them, but were not shaped by selection to that end. See §4.3.1 for a more detailed treatment of exaptations.

denying creative power to selection makes adaptedness inexplicable: there would be no evolutionary explanation of why an individual has the traits it has. The standard neo-Darwinian claim is that there is no fundamental difference between micro- and macro-evolution, the latter being just an extension over time of the former (e.g. Linde-Medina & Newman 2014); this view requires that selection not only selects, but must also create the fittest through accumulation of small mutations (Gould 1980): other sources of novel phenotypes (for example, macro-mutations) are not considered feasible or relevant for evolution (Orr 2005). The adaptationist paradigm, at least in its empirical version, endorses the positive view, and it is so widespread among biologists (Sober 1998a, Orzack & Sober 2001, Stoltzfus 2006) that even Kimura, the founder of the neutral theory (see §4.2.4), while claiming that natural selection has little impact at the molecular level, affirms that "*evolution* at the level of form and function *is largely determined by Darwinian natural selection*" (Kimura 1983:61-62, our italics).

Adaptationism under the positive view thus considers that all traits are adaptations[52]. Lewens (2007) gives an overview of different definitions of adaptations, classifying them between historically and non-historically based ones. The former ones appeal to the evolutionary history of the population, as in the following case:

> A is an adaptation for task T in population P if and only if A became prevalent in P because there was selection for A, where the selective advantage of A was due to the fact that A helped perform task T. (Sober 1984: 208)

Non-historical definitions focus on the current role of a trait, independently of its evolutionary history:

> An adaptation is a phenotypic variant that results in the highest fitness among a specified set of variants in a given environment. (Reeve & Sherman1993: 9)

The concept of adaptation generates a new set of problems though. West-Eberhard (1992) summarises the ways in which the term is used in evolutionary biology, but there seems to be no unique criteria to define a structure as an adaptation (Linde-Medina 2011). Many definitions includes concepts that are themselves controversial. Dobzhansky (1968)

[52] This is of course a theoretical claim fully valid only on very long time-scales, in an ideal world with strong selection, constants environment and no genetic and developmental constraints. Adaptationists do not believe that such is the case of the real world, but they do believe that any element that distance the current form the ideal world cannot have any positive role in shaping the phenotype.

claims that "adaptation is the evolutionary process whereby an organism becomes better able to live in its habitat or habitats": but it has been argued that an ecological niche can only be identified in terms of the activities of the organism living in it (Lewontin 1983), making this definition viciously circular. An alternative view defines adaptations as "genetically programmed features and behaviours that promote survival and reproduction." (Maynard Smith et al. 1985); by recurring to fitness considerations, this definition seems to avoid the niche definition problem. McShea and Brandon (2010), however, point out that improvements in fitness are not necessarily linked to adaptations. To show this, they identify three necessary but not sufficient conditions for selection: (a) variation in some traits, (b) heredity of this variation, and (c) differential reproduction of each variant. The additional condition to induce the action of selection is that (c) depends on (a): if not, drift could also explain the evolution of the population, and the trait, although fulfilling Maynard Smith's definition, would not be an adaptation. This dependence is hard to establish without recurring to niche considerations, such as: What are the problems that the niches pose? What functions are required to cope with them? Which traits supply these functions and how? Williams (1966) identifies the same issue in Fisher's definition (1930) of adaptation as a relationship between the organism and its environment such that any slight change in the former or in the latter results in a less adapted organism: this view entails that we can define the environment of a species independently of the species, as an ontological entity in itself that we can fill with organisms more or less adapted (see also Fodor & Piattelli-Palmarini 2010). Williams also underlines the lack of a general criterion to distinguish adaptations from fortuitous effects: he claims that if selection seems unable to explain a trait, this should not be considered as an adaptation. He suggests using similarity with human design as a discrimination criterion, and mention the use of paws by a fox to create a path through the snow as an example of a behavioural trait which is not an adaptation, as paws were not selected to that end.

According to the opposite idea, or 'negative view' (Sober 1985), selection explains the distribution of traits in a population, but not the possession of traits by individuals: it is a claim about which traits get to fixation, not about how these traits appear (Sober 1998a). The negative view limits the creative power of selection:

> It suggests that strong selection for, say, long necks in giraffes over many generations does nothing whatsoever to explain why a particular giraffe has a long neck, despite its manifest causal impact in the past. (Stegmann 2010).

This leads to the scission of the two roles of selection (in the creation and in the fixation of traits):

> It follows that adaptation and adaptive are not interchangeable concepts. A trait is adaptive now if it currently confers some advantage. A trait is an adaptation now if it currently exists because a certain selection process took place in the past. The two concepts describe different temporal stages in the trait's career--how it got here and what it means for organisms who now have it. A trait can be an adaptation now without currently being adaptive. And it can be adaptive now, although it is not now an adaptation (*for example, if it arose yesterday by mutation*). (Sober 2000: 84, italics mine).

The reference by Sober to mutations is relevant because it points to the existence of alternative ways for traits to appear which are logically and historically prior to selection. This issue is treated in Ch. 4.

3.3.2. Optimization ability of selection

The creationist tradition believed in the perfect adaptation of creatures to their environment. In Paley's *Natural Theology* there are several mentions about organism's perfection (e.g. Paley 2008: 149, 156, 216). Here is an example regarding feathers:

> This admirable structure of the feather, which it is easy to see with the microscope, *succeeds perfectly* for the use to which nature has designed it. (Paley 2008: 118, italics mine).

On the contrary, Darwin argues that organisms are fit for survival in their environment, but he does not consider that they are *'perfectly' fit*:

> Natural selection tends only to make each organic being as perfect as, or slightly more perfect than the other inhabitants of the same country with which it comes into competition. […] *Natural selection will not produce absolute perfection*. (Darwin 1872: 223-224, italics mine)

Sober follows Darwin on this point: "Adaptationism, as I construe it, does not demand that the process of natural selection maximizes the fitness of the organisms (or the genes) in a population" (Sober 2000: 123-124). Sober (1984) considers that adaptation says something about the past history of a trait, while fitness predicts its future. Williams (1966) seems to express the same idea when he claims that selection is just a statistical bias in the relative

rates of survival of alternative individual organisms, and as such it is not an explication of evolutionary change (where traits come from), but of the maintenance of adaptation (why traits get fixed) (see also Razeto-Barry & Frick 2011). Gregory (2008) underlines the 'bricolage and collage' type of action by selection, that works with what is already present and available, something that "has been often overlooked when authors characterize natural selection as an optimization process" (p. 364).

Sober's position is not common within neo-Darwinian thought: adaptationism defends the view that natural selection pushes traits to (at least) local optimality (Orzack and Sober 2001). Williams (1966) interprets Fisher's definition of adaptation (a relationship between the organism and its environment such that any slight change in the former or in the latter results in a less adapted organism) as a declaration that all adaptation is optimization[53]. Lewontin captures this idea when he defines adaptationism as the "approach to evolutionary studies which assumes without further proof that all aspects of the morphology, physiology and behaviour of organisms are adaptive *optimal solutions* to problems" (Lewontin 1979, italics mine). Dennet underlines this view of natural selection as a creative force acting on random changes when he writes: "Darwin's central claim is that when the force of natural selection is imposed on this random meandering, in addition to drifting [random changes in the distribution of traits] there is lifting [improvement of the individual phenotype]" (Dennet 1995, comments in brackets are mine), a process that, we can imagine, only stops when an optimum is reached. The interpretation of 'optimum' is another original contribution of the adaptationist paradigm that sees phenotypic design as the sum of the traits of an organism, each contributing to the total fitness of the individual. The slow and continuous action of natural selection configures the best value for each trait and of their mix, through a process of fitness optimization. Grafen gives an explicit definition in this sense: "Adaptation is design, and maximizing fitness is what organisms are designed for." (Grafen 2007). Some formal models have been produced to support such claim, mainly based on optimization programs (see §3.5). The most ambitious, and controversial, attempt in this sense is Grafen's Formal Darwinism Project (see §3.6). By applying population genetic formalism to optimization programs, Grafen aims at demonstrating that fitness maximization under natural selection is a general trend.

[53] The optimum can be a local one, though.

3.3.3. Constraints and selection

Darwin believed that the variability of the forms of life upon which selection acts is almost *limitless*:

> I can see no limit to the amount of change, to the beauty and complexity of the coadaptations between all organic beings, one with another and with their physical conditions of life, which may have been effected in the long course of time through nature's power of selection, that is by the survival of the fittest. (Darwin 1872: 123-124)

Today, no advocate of adaptationism denies the existence of genetic, developmental and physical constraints, at least in the short run[54]. Constraints are considered, though, as simple limitations that might prevent the attainment of theoretical optimal solutions, either because they keep a population stuck at a local optimum ('constraints on adaptation'), or because they prevent the evolution of a developmental pathway leading to such optima ('constraints on form') (Amudson 1994). While Amudson's point is ontological (as it talks about how nature is), a specular epistemological point could be raised regarding the necessity of constraints on the possible values of variables in order to build significant optimality models (see §3.5).

Other authors consider constraints as genuine opportunities to explain traits, not just as limitations to the action of selection, but as causes of evolutionary novelty *instead of* selection (see e.g. §4.4.2). For instance, Sterelny (2008) argues that the emergence of multicellularity is due to a bias that pushed lineages toward certain evolutionary paths. While Gould and Lewontin (1979) defend the central role played in evolution by constraints (on organisms' development, for example). Boundaries, a special type of constraints that limit the upper or lower value of some variable, might also have a positive role in shaping traits: Gould (2002), for example, believes that the lower boundary of complexity is fixed by biochemistry, while there is no theoretical upper one; while Maynard Smith and Szathmary (1999) consider that there are upper thresholds that, until overcome with a 'sudden jump', limit the evolution of complexity; for example, the structure of the prokaryotic cell has remained stable in evolution until the eukaryotic model suddenly emerged.

We can now conclude this section by summarizing how the fundamental tenets of empirical adaptationism spring from the three claims discussed. Selection is creative because,

[54] In the long run, some authors argue that selection is always able to find roundabout ways to avoid any constraint.

by inspection of the organism's traits, we can draw conclusions about the environment for which they are adaptations, but *not about their source*. How a trait appears is irrelevant, as it is not settled until it undergoes the sieve of natural selection: the environment, through selection, puts its final seal on the phenotype. Potential constraints (physio-chemical laws, developmental processes, etc.) are thus irrelevant: whatever mark they might initially leave on the phenotype, selection will wipe it away if they are adaptively irrelevant. Finally, traits are the optimal solution to environmental challenges.

In the remaining of the chapter, we show how these tenets are formalised in some adaptationist formal models; but first we have to introduce some other peculiarities of adaptationism, what we do in the next section.

3.4. Functionalism and Externalism

Regardless of how far it borrows from Darwin's ideas, all types of adaptationism embrace a functionalist view. Functionalism is the view stating that function is causally primary and explains the emergence of structure; for instance, functionalists would say that giraffes have a long neck *because* it helps them feeding themselves. The opposite view, structuralism, states that structure is causally primary and explains the eventual, possibly serendipitous, emergence of function[55]. Structuralism, for example, looks for developmental laws that define and limit the range of possible morphologies, and predicts forms by applying such laws (Kaufmann 1993: 15). Both views can be articulated as either historical positions, if they address the issue of the origin of biological systems, or as non-historical ones, if they just consider the synchronic relationship between form and function (Brooks and Wiley 1986). Adaptationism is clearly an historical version of functionalism, as it states that, given an environment, there exists an optimal way to survive in it, way that can be expressed as a set of functions. Natural selection is the mechanism that allows such functions to be implemented through the traits' configuration that is optimal to achieve survival. The link between adaptation and function is already pointed out by Aristotle, who believes that animals are designed to survive and reproduce[56] (Leigh 2001). Adaptations are, under this view, the transformation of environmental information into internal changes of form,

[55] See §4.1.4 for an introduction to the structuralist view in evolutionary biology.

[56] Of course, Aristotle, in his view of the universe as eternal and unchangeable, did not explain how organisms able to perform these functions came to exist: he just acknowledge that they do, and that they, the same as artefacts, cannot be understood without understanding their functions.

physiology and behaviour (Gould 2002: 157). The environment becomes the *deus-ex-machina* of evolution: organisms are, for the great part[57], a repository of past (historical) environmental changes. The process of change must possess the right pace: if the environment is too stable, no new adaptations are needed and evolution comes to a halt; if it changes too abruptly, natural selection does not have the time to shape organisms, and 'catastrophes' decides what survives and what perishes (Gould 2002: 161-163).

According to functionalism, homologies are a consequence of similar selective pressures[58] having unique optimal solutions, which selection, sooner or later, will find and 'reinvent' as many times as they are needed. The talk about algorithms is a good signal of a functionalist approach:

> The diversity of animals known to have stereo vision allows us to begin to investigate ideas about its evolution and the underlying selective pressures in different animals. It also further prompts the question of whether *all animals have evolved essentially the same algorithms* to implement stereopsis. (Nityananda & Read 2017: 2502, italics mine)

Organisms are mechanisms implementing algorithms, which are in turn the real target of selection:

> Form is arbitrary (subject to functional adequacy) and the cause of the existence of a feature is its functional significance. (Linde-Medina 2011: 582)

Selection will always find the best solution to problems:

> That is the one point which I think all evolutionists are agreed upon, that it is virtually impossible to do a better job than organism is doing in its own environment. (Lewontin, quoted in Dawkins 2006: 164).

Not only selection does a far better job than any human artefact[59]: engineers should copy selection's inventions, as the quoted paper about the evolution of stereopsis suggests:

> If so, this [i.e. the algorithm common to all animals] *must be the best way to do stereo vision*, and *should be implemented by engineers* in machine stereopsis.

---

[57] This does not mean that all environmental changes imply internal changes, but the reverse is true: all (significant) internal changes are developed as effect of external changes.

[58] Structuralism states that they are due, for example, to common developmental or physical constraints.

[59] However, recent biotechnologies go a long way to demolish this argument, it could be argued (thanks to Davide Vecchi for this comment).

> Conversely, if animals have evolved a range of stereo algorithms in response to different pressures, that could *inspire novel forms of machine stereopsis* appropriate for distinct environments, tasks or constraints." (Nityananda & Read 2017: 2502)

The movement on wheels, being unknown in the natural world, must be a very inefficient technology, one would conclude!

The last quotation suggests that, if constraints exist, alternative algorithms are applied. Nesse (1994) mentions the differences between the cephalopods eye, where the optic nerve does not interrupt the view, and vertebrates eye, where the 'blind spot' is caused by the optic nerve fibres running through the retina; in the latter case, selection has come up with an elaborate work-around of the eyes wobbling back-and-forth to correct for this "engineering defect". Another example of alternative engineering solution concerns the cranial nerves in tetrapods. In early vertebrates (sharks, skates and rays, collectively named Chondrichthyes), the cranial nerves run from the part of the brain that interprets sensory information and radiate out towards the organs producing those sensations. In tetrapods, however, and mammals in particular, the nerves take an elaborate winding path through the cranium around structures that evolved after the common ancestor with sharks.

Adaptationism might also be characterised as an externalist attitude. In pre-Darwinian theories like Lamarckism, the individual organism was the subject of evolution (its efforts changed, fostered and improved its abilities), pushed by internal forces (e.g. Lewontin 1983). Adaptationism, on the contrary, considers the organism as a passive object whose changes are triggered and driven by the external process of selection, while internal dynamics are irrelevant (Gould 1977):

> If a feature exists, it is because it has been forged by natural selection to meet external functional demands. (Linde-Medina 2011)

Kimura, the founder of the neutral theory (see §4.2.4), while claiming that natural selection has little impact at the molecular level, affirms that "progressive evolution is almost always brought about as a result of organisms' response to *environmental challenge*" (Kimura 1983:61-62, our italics). Macro-evolution, or speciation, is also explained by the same external causes (Linde-Medina & Newman 2014), typically the rise of new ecological opportunities: the Galapagos finches are a text-book example of adaptive radiation linked to colonization of a region offering not-yet-used resources (Erwin 2015).

We conclude this section with two final comments. Firstly, the 'functionalization' of phenotypes is what allows adaptationism to formalize evolution as a fitness optimization process, as the next section will illustrate. Secondly, the approach to evolution as an issue of gain or loss of functions pushed by the external environment places the action of selection on the vertical axis of our architectural model (§2.4.2): in chapter 4, we will analyse alternative non-functionalist views of evolution that, focusing on structure, suggest alternative forces that might explain the horizontal movements in our matrix.

3.5. Optimization Programs.

Population genetics, as we have seen (§2.2), interprets evolution as changes in alleles frequencies, without relying on the hypothesis that these changes are directed by a process of fitness optimization (Ewans 2004, 2014). Evolutionary biologists, on the other hand, claim that the phenotypes of organisms are combinations of traits individually optimised by natural selection:

> […] with constant phenotypic fitness and infinite population size the evolution of the average phenotype in response to selection is always in the direction which increases the mean fitness in the population. (Lande 1976: 317)

> Individuals should be designed by natural selection to maximize their fitness. (Davies et al. 2012: 81)

> Adaptationists often assume that natural selection has fine-tuned organisms to their environment and that the phenotypes observed in nature are optimal or nearly so. (Halama and Reznick 2001: 246)

According to this view, traits are justified thanks to their contribution to the organism's fit with the environment: if a trait is advantageous, a genetic combination encoding for it will sooner or later appear (Grafen 2014, Huneman 2014). Unfortunately, such adaptationist accounts lacks formalizations comparable to the mathematical models of population genetics, and usually recurs to narrative descriptions of the phenomena that presumably led to the appearance and evolution of a trait (see Fig. 2.5 for the comparison between population genetics and adaptationist views). This limitation opens the door to the charge of un-falsifiability (Gould & Lewontin 1979), making adaptationism fall on the wrong side of Popper's demarcation criteria for science (Popper 2002).

Adaptationism is thus in a somehow difficult position with respect to the formal arguments of population genetics, whose conclusions often seem to contradict the hypothesis that selection is an optimizing force. Nevertheless, including within the definition of adaptationism the hypothesis that natural selection pushes traits to local optimality *is what allows its formalization by making adaptationist narratives practically testable* (Sarkar 2005): because one thing is to say that the observed 'guarding' time after copulation of the *Scatophaga stercoraria* is optimal (a purely narrative, thus untestable, hypothesis), and another very different thing is to show that any other guarding time would result in a lower number of total fertilised eggs (a formal, testable claim)[60]. With this later claim, adaptationism moves from being an a priori true or false axiom to becoming a falsifiable, thus scientific, theory (Orzack and Sober 2001a):

> Optimality models, or a dynamical variant of such models, represent the only reasonable way to determine which traits would be adaptive under hypothetical circumstances when verbal reasoning does not suffice. (Abrams 2001: 293).

In the last decades, optimization models based on the adaptationist claim that natural selection leads quantitative traits to values maximising fitness have been widely used to generate testable interpretations of adaptationist claims (Abrams 2011, Parker & Maynard Smith 1990 for an introduction to the different kinds of models found in the literature).

Common in mathematics, informatics and economics, optimization programs aim at finding the best solution to a problem among all feasible ones. In its simplest form, the optimisation problem consists in selecting the value of a variable that maximises (or minimises) the value of a function. The fundamental elements of an optimization program are (Snyman 2005):

- the function to be maximised (called the *maximand*);
- the variable whose value is to be chosen (called the *strategy*);
- the *constraints* limiting the variable.

For example, the maximand could be the distance flown by an airplane with a given amount of fuel; the variable to choose could be the 'average speed'; and the constraints could be the

[60] Females of *S. stercoraria* copulate with several males, and it turns out that the second male fertilizes more eggs than the first. A male is thus interested in 'guarding' the female after copulation. The time spent guarding her, however, is not used to copulate with other females. There is thus a trade-off between flying away immediately after copulation and increasing the number of mates, and staying and guarding, increasing the chance to fertilize eggs in the already copulated mate (Sober 2004).

maximum and minimum speeds of the airplane, determined by its technical features and by the laws of fluid dynamics. The trade-off behind the problem is the following: if the pilot flights at a very high speed, the airplane will cover a long way in a short time, but it will soon run out of fuel; if she flights at a minimum speed, the airplane will flight much longer, but the covered distance will be less. By writing the equation connecting speed and fuel consumption, and by applying calculus techniques, it is possible to choose a speed resulting in the longest distance covered.

Optimization programs can also be used to minimize a function (the *minimand*). For example, to optimize the cost of aluminium used to produce a beverage can of a given volume, the elements of the program would be:

- variable: radius r of the basis of the can;
- constraint: volume V;
- minimand: total can surface S;
- it is easy to write the equation S = S(r) linking S to r, and then to find the optimal r* by imposing dS/dr=0. If the actual values of commercial 0,33 litre cans are used, it turns out that the aluminium consumption is not optimised, possibly because of other constraints not considered in the model (e.g. filling technology, transportation requirements, or marketing considerations).

In the realm of biology, optimization models have been widely applied to the theory of optimal foraging (Schoener 1971, Charnov 1976, Krebs and McCleery 1984, Stepehn & Krebs 1986, Pulliam 1989).

Fig. 3.2 – Optimal foraging strategy… (image: Swarthmore 2010)

Firstly proposed by MacArthur and Pianka (1966), the theory claims that selection favours animals that maximize their energy intake per unit of total time T spent in foraging,

sum of the time T1 dedicated to looking for food and T2 dedicated to actually getting and eating it once found. The strategy consists in deciding how much time to commit to each of these two tasks (Fig. 3.2). The authors clearly assume that animal behaviour can be predicted and explained by applying optimization considerations, as "[t]here is a close parallel between the development of theories in economics and population biology" and "natural selection will often achieve such optimal allocation of time and energy expenditures" (MacArthur and Pianka 1966: 603). Sinervo (1997: 114-115) proposes an especially clear optimal foraging model. The author considers total foraging time T as the strategy variable whose value is to be decided. As simplifying assumption, the time T1 (travel time) is considered fixed: any two foraging areas are supposed to lie at the same distance. The function to be maximised is the energy intake per unit of time, or E/T. The energy intake during travel (T1) is null; during T2 (actual feeding), it is positive but marginally decreasing[61] (Fig. 3.3).

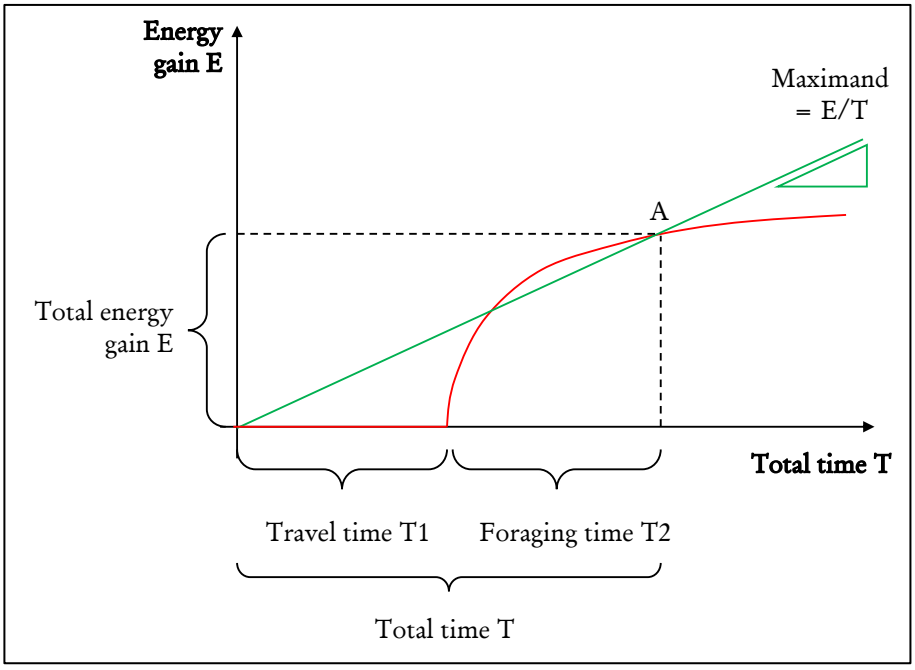


Fig. 3.3 – The red line shows total energy gain E as function of total time T. Energy gain for T<T1 is null. For T>T1, the energy gain is positive but marginally decreasing with T: for the same time interval ΔT, the energy gain from T' to T'+ΔT is smaller than the energy gain from T" to T"+ΔT, if T">T'. The *maximand* is the average energy intake during total time T, or the slope of the segment connecting point A to the origin of the axes.

[61] To see why it is decreasing, think of the task of plugging apples from a tree: at the beginning, when the branches are full of fruits, a basket is quickly filled with apples; at the end, once the majority of apples have already been plugged, it is harder and harder to find the remaining, scattered ones. It takes much longer to fill the last basket then the first one.

The relationship between time and energy gain is thus not linear: there comes a time when remaining in the same area causes a decrease in average energy gain. The individual should thus leave before this threshold is passed (Fig. 3.4).

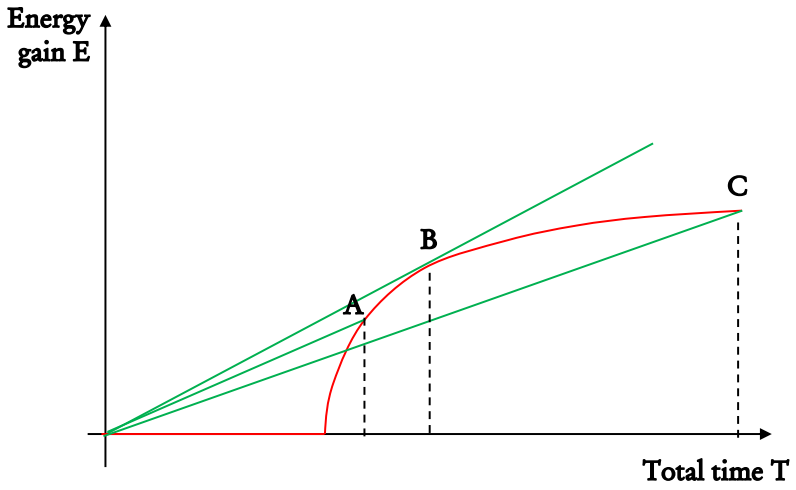


Fig. 3.4 – Different foraging strategies: in A, the individual chooses to leave the area too soon; in C, too late; in B, the individual leaves the area at perfect timing, hence its energy gain per unit of time is maximized.

The hypothesis underlying foraging time optimization models is that organisms with a more efficient energy assumption show a higher fitness, so that their strategy spreads throughout the population and gets finally fixed. Obviously, the function that links the trait value (in this case, foraging time) with the individual fitness, although fundamental for the model to be significant, is difficult to estimate; let alone its genetic encoding:

> Optimizing selection has been particularly difficult to model because of the gap in our knowledge between the well-founded theory of population genetics based on elementary Mendelian mechanics and empirical observations on the inheritance properties of polygenic characters. At this time we are unable to describe the effects of individual genes on a continuously varying character in order to develop an accurate model of the effect of selection. (Slatkin 1970: 87)

Apart from foraging theory, optimality considerations have been used to formalize adaptationist claims about very different phenotypic traits. Roff (1981) builds a model linking fitness to body size in *Drosophila*, calculates that the optimal size is 95 mm., and shows that individual from actual populations with the calculated size have a fitness maximum, while fitness of individuals with body size even slighter higher or lower than the

calculated optimum decreases abruptly. The author's approach also represents a vivid example of the adaptationist stance: if "the optimum does not correspond with the observed size we may conclude either that one or more of the components has been incorrectly specified or that density-dependent or interspecific interactions are important" (Ross 1981: 406); the option of the trait not being maximised by selection is not considered. Bidder (1923, 1937) studies the relationship between a sponge form and its current, calculates the optimal osculum diameter that maximizes the water jet, measures the actual dimension in organisms and concludes that sponges' design is perfect. Parkhurst and Loucks (1972) apply an optimization model to photosynthesis. They deduce leaf size by assuming that the function to be maximized is the amount of photosynthesis per unit of water transpired. Even Kimura, who advocates for the preponderant importance of drift at molecular level, believes in the power of selection for optimizing mutation rates (Kimura 1960, Gillespie 1981).

Birch (2018) distinguishes four types of fitness optimization principles, logically independent in the sense that none entails any other, and based on two distinctions:

- whether the principle aims at describing what happens at *equilibrium* (E[62]) or rather the *direction of change* (D) when the population is not at equilibrium;
- whether the principle concerns the *population* (P) mean fitness or the *individual* (I) behavioural strategy.

Sinervo's model, for example, considers the equilibrium condition for individual strategy (E-I model). Fisher's fundamental theorem (§2.2.2), on the other hand, can be interpreted as an optimization principle (Grafen 2015) regarding the direction of change of average population fitness (D-P). The logical independence of Birch's principles nicely appears in Lehman's model (Lehman et al. 2015) formalising strategy optimization in a patch structured population: the authors demonstrate that lineage fitness maximization does not involve individual strategy optimization.

Simple optimization models like the one described are based on the assumption that the value of a strategy does not depend on how many other individuals make the same choice. If this is not the case, and the optimal strategy for an individual is frequency-dependent (it depends on what others do, especially on how many others do the same), then more refined, game-theory based models taking into consideration the dynamic of the social environment can be applied. Instead of an optimal strategy, an evolutionary stable strategy (ESS) arises

[62] The labelling of the options is ours and does not appear in Birch.

(Maynard Smith & Price 1973): a strategy that, once fixed in the population, cannot be displaced by any mutant behaviour. It is a Nash equilibrium[63] secured by the action of natural selection, thus evolutionary stable. Contrary to the simpler case of optimization models, an ESS does not necessarily optimize the population average fitness, but just maximises it, in the sense that any other strategy played by mutants does worse. Several ESS might co-exist for a population in a given environment, so that the actual one depends on the initial conditions (Maynard Smith 1982, Parker and Maynard Smith 1990). Brown (1991) claims that game theory optimization models applied to adaptationism can explain both the variety and the increasing complexity of organisms.

Apart from frequency-dependent cases, that result in multiple, not necessarily optimal equilibria, the optimal solution to a phenotypic problem can be out of reach because of many other reasons, for example a too quick environmental change (Abrams 2001), or constraints such as a genetic system composed of few loci, thus with limited potential variation (Hines and Bishop 1983; Eshel and Feldman 1984). The action of other forces like drift, mutation and migration, might also invalidate the results of optimization models and prevent the optimum to be reached (Abrams 2001). All these disturbing factors do not discourage the adaptationist, who rebuts that natural selection just *tends* to bring traits to their optimal relationship with other traits, or with the environment, not that it actually does (Orzack and Sober 2001b).

Even if considerations about frequency-dependence, constraints and non-selective forces are included in the optimizations models in order to make their predictions more accurate, these models still rest on a questionable functionalist assumption: that ecological niches can be identified as objective entities and that the relationship of the organisms with the environment can be modelled through functions. The following quote represents an example of such view:

> The enormous diversity of organisms may be envisaged as correlated with the immense variety of environments and of ecological niches which exists on earth. (Dobzhansky 1958: 9).

Dobzhansky seems to claim that niches are there and can be identified for their intrinsic properties rather than in terms of their relation with respect to organisms.

---

[63] The Nash equilibrium is a solution of a non-cooperative game in which each player knows the strategies of the other players, and in which no player gains by changing her own strategy (Osborne & Rubinstein 1994).

Species and clades just occupy an (onto)logical, pre-existing space, captured by the image of a fitness landscape (Gould 2002:527-528), and their success in surviving and reproducing depends on their ability to solve an equation: organisms are computing machines fuelled by selection. This is not, however, what Darwin claimed: he saw natural selection as a *mechanism* leading to local adaptation, and not as a perfecting *principle* (Gould 1976). Sarkar (2005) points out that fitness can and do decrease between generations, due to genetic mechanisms[64]: natural selection and optimization are thus not equivalent.

The idea of selection as an optimizing engine acting on a function linked to fitness bears the implicit limitation of ignoring phenomena that could radically change the definition of a problem, as Prigogine underlines:

> […] the way in which biological […] evolution has traditionally been interpreted represents a particularly unfortunate use of the concepts and methods borrowed from physics […]. The foremost example of this is the paradigm optimization. It is obvious that […] the action of selective pressures tends to optimize some aspects of behaviour […] actions or modes of connection but to consider optimization as the key to understanding how populations and individuals survive is to risk confusing causes with effects. *Optimization models thus ignore both the possibility of radical transformations* -that is, transformations that change the definition of a problem and thus the kind of solution sought- and the inertial constraints that may eventually force a system into a disastrous way of functioning. (Prigogine 1984: 207, italics mine)

Wagner's B-Matrix model (Wagner 1989) is an interesting version of optimization program based on the mapping of genetic variables into phenotypic variable, and in turn phenotypic variable into fitness. The model summarises the limitations of the extreme functionalist assumptions upon which adaptationist formalizations rest: that each phenotypic variable is *a priori* supposed to show an optimal value (in terms of its contribution to fitness); that the organism is reducible to the sum of its traits, considerations around architecture being irrelevant (Altenberg 2005); that no evolutionary mechanism other than selection can shape a trait (Sarkar 2005). Lande claims explicitely that:

---

[64] E.g. overdominance, or heterozygote advantage. Even in a 100% heterozygote population, some homozygote will appear in the next generation, thus reducing average population fitness. See (Grafen 2014) for other examples of genetic mechanisms limiting fitness optimization by selection.

> [T]he actual phenotype of an organism is composed of many inter-related characters. If these characters follow or can be transformed to a multivariate normal phenotype distribution, a rotation of axes can be performed to find linear combinations of the characters which have no genetic correlations and the above analysis [*an optimization program*] can, in principle, be applied to these genetically independent characters. (Lande 1976: 317, note in italics mine)

Lewontin (1974) also identifies operational issues that, if not satisfactorily clarified, make evolutionary models of little practical and theoretical use: they are not dynamically sufficient, not allowing to forecast the future state of a population based on their laws and state variables; the tolerance limits of their description is usually ignored; and their state variables are hard to measure. The author claims that the last requirement is especially difficult to fulfil, for example when trying to measure genetic fitness, and that evolutionary models, by not recognizing this "epistemological paradox", often become

> […] a vacuous exercise in formal logic that has no points of contact with the contingent world. The theory explains nothing because it explains everything."(Lewontin 1974: 11-12)

3.6. The Formal Darwinism Project by Alan Grafen.

Even if all these limitations of adaptationist optimization models were overcome, there would still remain an undeniable tension between their results and those of population genetics formalizations: the assumption of fitness optimization is in fact defended by the former and denied by the latter. The Formal Darwinism Project proposed by Grafen (1999, 2002, 2006a, 2007, 2008, 2009, 2014a, 2014b) aims, among other things[65], at dissolving this tension, by unifying in a single model the evolutionary explanations of phenotypic traits provided by optimization programs, and population genetics formalizations of the evolution of genotypic frequencies. Grafen especially believes that selection, and selection alone, accounts for the rise of organismal design[66], and that it can thus explain not only the

[65] For example, Grafen believes that his model can account for "modern extensions of evolutionary theory including ESS theory and inclusive fitness, and Dawkins' synthesis of them into a single structure." (Grafen 2014a: 155)

[66] Grafen talks about 'design' without giving a formal definition of the term. We will keep his terminology, by interpreting 'design' as 'phenotypic architecture', as defined in §2.3.2

diversity of life, but also the increasing complexity of the organisms represented by the tree of life (see Fig. 2.1 in the previous chapter): he does so by defending the thesis that design *is* fitness optimization, and that fitness optimization is the inevitable output of evolution by selection. The core of the project consists in describing the evolution of a population both through an optimization program and through population genetics equations and in showing that, under appropriate assumptions, one model can be translated into the other, so that the equilibrium state of both descriptions is the same. Grafen claims that this demonstrates that evolution by natural selection optimizes traits' fitness (an adaptationist claim) by modifying alleles' frequencies (a population genetics description):

> The Formal Darwinism Project aims to provide a mathematical framework within which important fundamental ideas in large parts of biology can be articulated, including Darwin's central argument in *The Origin* (that mechanical processes of inheritance and reproduction [i.e. selection] can give rise to the appearance of design [i.e. optimization of traits]) […] [and] […] the centrality of fitness maximisation ideas to many areas of biology (Grafen 2014a: 155)

Grafen claims that his project fills the gap between biology and physics: contrary to what Newton did with the latter, no formal account of the Darwinian evolutionary theory was yet available before his proposal.

In this section, we firstly present the main aspects of Grafen's proposal, and then underline three implicit logical premises at the base of the model that, we believe, limit its field of applicability.

3.6.1. Two models, one equilibrium

1. Introduction

To carry out his project, Grafen firstly builds a model of the evolution of a *population* based on the *genetic approach*, applying the Price equation. He calls this formalization 'mathematic of motion'. He then builds a second model of *individual* evolution based on the *adaptationist approach*, using an optimization program that he refers to as 'mathematic of optimization'. This terminology reflects the fact that population genetics models account for the dynamics of the evolution of the population between generations, without suggesting that it points to any equilibrium; while optimization programs focus on explaining equilibria, without suggesting how the system has reached them. Grafen demonstrates that each of the two models can be interpreted and translated in terms of the other so that, at equilibrium, both

give the same result. The translation is performed by showing that there are some necessary links among the results of the two approaches.

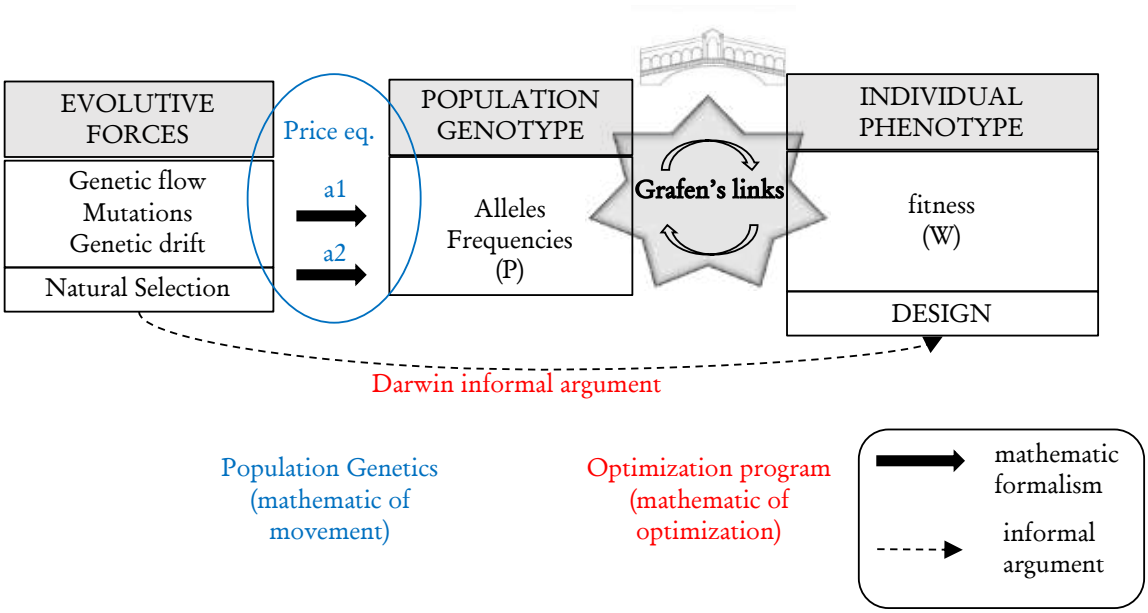


Fig. 3.5 – Grafen's project – Evolutive forces affect the genotype according to the Price equation (that accounts for selection, arrow a2, and other forces, arrow a1). The design of the phenotype is explained by selection through the informal Darwinian argument (dotted arrow). Grafen aims at creating a bridge between the two views, by building an optimization program for a population that shows some necessary links at equilibrium with the description of the same population in terms of genetic dynamics.

2. Assumptions

The project relies on a set of hypotheses that Grafen's attributes to Darwin's original argument in *The Origin*:

1. *Selection maximises fitness*, to be interpreted as individual fitness, and not average fitness like in population genetics models[67] (Grafen 2002: 86);
2. *Genetic details are not important* to understand phenotypic functions (Grafen refers to this hypothesis as '*phenotypic gambit*', Grafen 2002: 77). Grafen aims at formalizing Darwin's argument, so he justifies the gambit based on Darwin's ignorance of genetics. The separation between genes and functions is also justified because (see also §2.3.2):
    - Functions spring from a process of fitness maximization, independently from which traits perform them and how these traits are genetically encoded. The

[67] Grafen applies inclusive fitness, a concept articulated by Hamilton (1964) with the aim of explaining the evolution of altruism through natural selection. Although selection maximises inclusive fitness, this remains a trait of the individual. Selection changes the target of maximization, not the maximization agent, which is still the individual, not the group (Gardner 2014b: 105).

evolution of the phenotype (form and function) is led by natural selection, but "positive selection cares little about how such form and function are brought about genetically" (Kimura 1983: 62);

- The same function can be performed by traits with different genetic encoding. For instance, bacterial and eukaryote flagella share some of their functions but have very different molecular structures (see e.g. Blair & Dutcher 1992);
- The same phenotypic trait can have several functions. Marine turtles' flippers work for swimming as well as for digging in sand when laying eggs.

3. The *unit of selection is the individual organism*, neither the group nor the gene: only the individual organism has a phenotype, is subject to selection and shows some correspondent fitness (Grafen 2014a: 211-212).

3. Population genetics model

Grafen believes that all population genetics models can be described in terms of the Price equation (§2.2.3), as the equation summarises their virtues and limitations: it identifies fitness as the target of evolution, and it clearly separates the evolutionary adaptive 'engine' that increases it (selection) form other causes that can reduce it (mutation, drift, etc.). Nevertheless, the equation does not entail fitness optimization, it considers average and not individual values and, above all, it focuses on the genotype and not on the phenotypes, where design is to be found (Grafen 2002: 77). Grafen, coherently with his adaptationist view concerning the supremacy of selection, ignores the second term of the Price equation[68], thus assuming unbiased transmission (Grafen 2008: 427) and discharging the influence of the environment - interpreted in a broad sense, e.g. chromosome segregation or mate pairing system (Grafen 2002: 90) -, as it did not contribute to traits' optimization. The Price equation is reduced to:

$$\Delta P = \mathrm{Cov}\,(Pi, Wi/W)$$

[68] The reader should remember that the Price equation describes evolutionary changes (changes in alleles' frequencies) as the sum of two terms, one depending on selection and the other on any other evolutionary force (see §2.2.2). By ignoring the 'transmission' element of the Price equation that, the reader might recall, describes changes in alleles' frequencies due to mutations and other non-selective phenomena, Grafen does neither deny that they exist nor does he minimize their importance *tout-court*: he rather contends that their contribution to phenotypic design is neither continuous nor inevitable: they represent an evolutionary 'noise' whose possible success/causal role in affecting the phenotypic design is just due to chance.

where Pi and Wi are the values, respectively, of the quantitative character P and of fitness W in individual i, and P and W the respective averages over the population. Grafen additionally postulates the existence of two functions: one (ω) linking fitness and the phenotype: ϕi: Wi = ω(ϕi), thus assuming that *only the phenotype affects fitness*; the other (ν) linking the phenotype to the genotype: Γi: ϕi = ν(Γi), thus assuming that *no variable other than the genotype* (e.g., environmental elements) *affecting the phenotype is evolutionary relevant* (Leheman & Rousset 2014). By substituting these functions in the Price equation, we get the dynamical model used by Grafen in his demonstration:

$$\Delta P = \mathrm{Cov}\,(Pi, \omega(\nu(\Gamma i))$$

4. *Optimization model*

The corresponding optimization program is the following:

- The *instrument* (or optimization variable) is *the individual phenotype ϕi*. The individual solves the optimization program if its phenotype corresponds to the maximum possible value of the maximand. The phenotype can be identified with any trait (physical or behavioral) that, at the same time, is target of selection and has some impact on the reproductive outcome: "The instrument can be anything that natural selection acts on that influences reproductive success. It can be sex ratio of offspring, foraging time in a patch, body mass and so on." (Grafen 2007: 1244-1245).
- The *constraints* to the set Φ of all possible values of the phenotype reflects *physiological, physical, and informational limitations*, linked to issues such as: "what can an animal do? What can a mutation produce? What events can an animal's behaviour be conditioned on?" (Grafen 2008: 424). These are constraints linking form and function.
- The *maximand f* measures the success of the phenotype (Grafen 2007: 1247). Grafen postulates that *f is the same function*:
    - For *all traits*: this is coherent with the choice of considering individual instead of trait fitness[69] (Grafen 2002: 80).
    - For *all individuals* (what he calls '*strategic equivalence*'): this means that individuals can be distinguished for their genotypes and phenotypes, but not for the strategy they follow in the 'selection' game (Grafen 2008: 428).

[69] See §3.6.2, point 2, for a detailed analysis of trait and individual fitness.

The optimization program is thus[70]:

$$\phi i: \max f(\phi i), \text{ where } \phi i \in \Phi$$

5. Translation

The relationship between the genetic and the adaptationist models is not obvious at all: population genetics has to do with *population* and *genotypes*, while optimization programs have to do with *individuals* and *phenotypes*. The project aims to build a bridge between the two approaches, by identifying, thanks to a proper translation, which elements of the dynamic model equate to which elements of the optimization program.

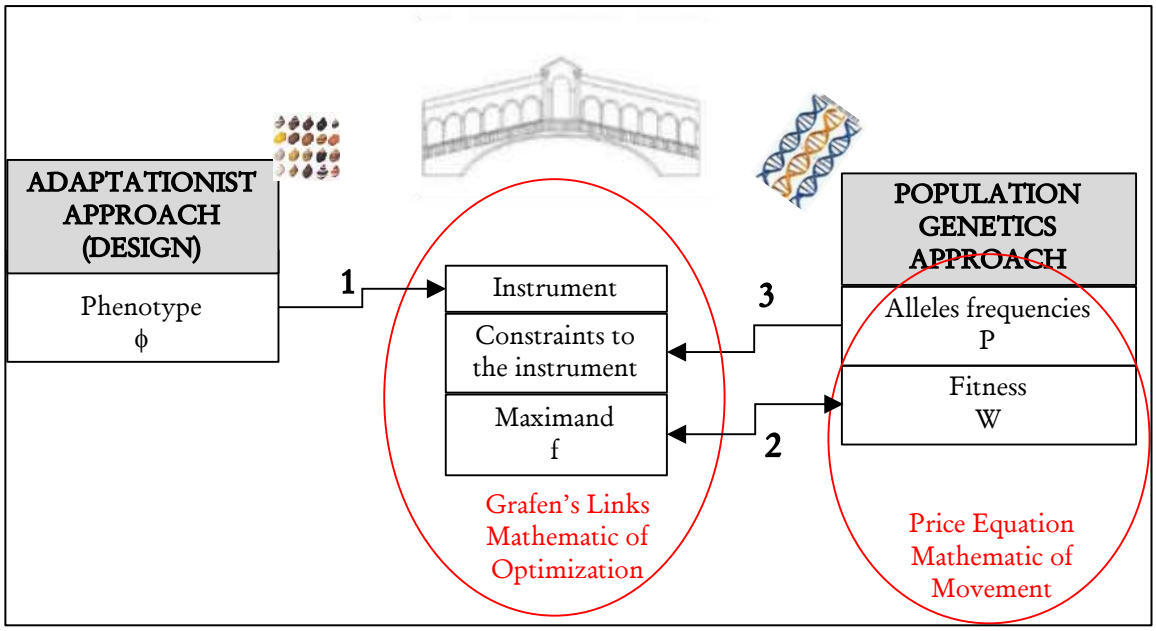


Fig. 3.6 – Grafen's translation. The adaptationist approach focuses on phenotype's design, justified through informal narratives. The population genetics approach focuses on alleles' frequencies and their evolution, based on differential fitness, formalised through the Price equation (the 'mathematic of movement'). Grafen builds a bridge between the two approaches through an optimization program whose elements are interpreted in terms of both: the instrument is the phenotype; the constraints are set by genetics; and the maximand is fitness.

Fig. 3.6 summarises the main elements of the two models and the relationships stipulated among them, allowing the translation of a model into the other:

1. The *instrument* of the optimization program is the *individual phenotype* ϕi[71].

[70] In later versions, Grafen refines his model to contemplate elements such as uncertainty (Grafen 2008), social interactions (Grafen 2006a) and classes of individuals (Grafen 2006b). The fundamental hypothesis and the mechanics of the model, as well as the conclusions, remain, nevertheless, unchanged, and we therefore limit the present discussion to the basic version of the model.

[71] This view considers the phenotype as exogenous to the model, and not as its product. Grafen's model does not consider that the environment could prevent the optimal phenotype from appearing and get fixed in the population (Huneman 2014: 276).

2. The *maximand f* is the *function* ω that links the fitness to the phenotype in the Price equation (Grafen 2007: 247):

$$f = \omega$$

According to Grafen, this is the only way to give a formal definition of fitness (Grafen 2008: 424-425). This equivalence is allowed thanks to the two above mentioned hypotheses: each individual and each trait share the same maximand, the same fitness function, so we can move from individual to average values.

3. The *constraints* to the set of possible phenotypic values are determined by genetics (alleles frequencies).

Thanks to this translation, the states of the two models, be it of equilibrium or not, can be compared. Grafen identifies four necessary links behind this relationship (Grafen 2002: 87-88):

I. *If each individual solves the optimization program* (adaptationist approach), then the expected variation of alleles frequency is null, and *no new phenotype* will spread throughout the population (population genetics approach). All individuals of the population show the same phenotype, which is a perfect adaptation to the environment;

II. *If each individual gets to the same value* of the optimization function, but not to its maximum value (adaptationist approach), then the genetic variation is null but *a mutant type could appear* and spread (population genetics approach). As in the previous case, all individuals of the population show the same phenotype, but this is not the best possible adaptation to the environment: changes in the genetic pool could result in improved adaptation, and these changes would spread;

III. If *each individual gets to a different value* of the optimization program (adaptationist approach), *variations in alleles' frequencies equal covariation among genes and fitness* (population genetics approach). There are different phenotypes showing different degrees of adaptation: selection acts and eliminates the less adapted phenotypes, spreading and fixing the more adapted ones. This is a reformulation of Fisher's theorem: it describes the evolution of a population climbing toward a fitness peak;

IV. If the *expected change in each allele's frequency is null*, and no new phenotype can spread (population genetics approach), then *each individual solves the optimization program* (adaptationist approach). This link mirrors the first one.

The first three links stipulate a state for the optimization program and allow deducing some conclusions about the correspondent changes in alleles' frequencies. The fourth link works in the opposite direction, from changes in alleles' frequencies to the deduction about the state of the optimization program: the model captures the idea that the mechanisms of reproduction and inheritance allows the individual to act as a rational decisions maker[72] (Grafen 2007: 1248). The 4th link is thus just the 1st one interpreted from a population genetics view: it is not logically independent.

6. Conclusions

Grafen believes that his model demonstrates that alleles' fixation according to the Price equation automatically leads to the maximization of the optimization function, that is, of phenotypic individual fitness. Once the optimization function is maximised for all individuals, no fitness variation exists in the population: the inter-generational variation in the Price equation is null. There are isomorphisms between natural selection and optimization (Huneman 2014: 274). The links show that *optimization by natural sel*ection is not a particular case, but *a general tendency*. Such unification of the adaptationist and genetic approaches formally demonstrates, according to Grafen, the Darwinian idea that natural selection explains the complexity of life.

Grafen argues that the models of population genetics necessarily lead to an equilibrium state described by an optimization program; and that, conversely, the equilibrium state is necessarily reached through a path described by genetic models. It is thus possible to establish a parallelism between Grafen's project and the semantic and syntactic rules of an axiomatic system: the former, like the optimization program, tell us which states are meaningful within the system, and thus allowed; the latter, like population genetics models, tell us which deductions are possible to move between allowed states. The admitted deductions (population genetic formulas) necessarily lead to meaningful states (optimality equilibria), and the meaningful states (optimality equilibria) are reachable through a finite series of admitted deductions (population genetics formulas). It turns out, however, that there are admitted states in an axiomatic system that are not reachable just through the deductions allowed by its syntactic rules[73]. The same situation could be reflected in this

[72] Note that this is of course a metaphor: Grafen does not impute any intentionality to the individuals. (See Okasha 2018)

[73] This is what Gödel incompleteness theorem affirms: in an axiomatic system, there are true statement that cannot be proved without recurring to some new axiom not included in the original set of axioms. The new axiomatic system that includes the new axioms, though, shows the same

metaphor, as the next chapter shows: there might be phenotypic equilibrium states not reachable through the movement allowed by selection alone.

3.6.2. A controversial project

Grafen's project represents an important contribution to the task of formalizing the adaptationist claims about the role of selection in maximizing fitness and in shaping organismal design: a good sign of the significance of Grafen's project is the intense discussions that it has triggered among biologists and philosophers alike. The main controversies are well summarized in a series of articles appeared in the monographic issue of the journal *Biology and Philosophy* (Volume 29, Issue 2, March 2014) dedicated to Formal Darwinism. Many criticisms question the explicative power of the links. Birch (2014: 178-180) believes they hold in too weak a form insufficient to demonstrate that selection leads to design, as they allow suboptimal equilibria. For example, even if alleles' frequencies do not change, genotypes' frequency could, and the population would not be evolutionary stable even if Grafen's link I holds (see also Sober 1999: 2). Link III affirms that selection causes fitness-improving traits to get fixed within the population, but as a matter of fact selection could be unable to do so, for example in case of polymorphic equilibria. Osaka and Paternotte (2014) argue that Grafen's links form a tautology:

> […] in every case, two out of three of links 1, 2 and 3 must hold trivially […]. The reason is simple: the antecedents of these three links[74] constitute a partition of logical space, so exactly one of them is true in any case. (Osaka & Paternotte 2014: 227)

Fig. 3.7 details the tautology in the logical space. Individuals can get to the same or to different values of the strategic variable; and this can be or not a maximum. The three links are trivially true in each of the three possible areas of the matrix. Note that, if each individual gets to a different value, these values cannot be maxima, as Grafen supposes that the fitness function does not have multiple peaks: this is why the cell 'different values/maximum value' is marked as Not Applicable.

---

problem: it includes true statements not provable without adding new axioms. And so on. See (Smith 2007) for an introduction to the Gödel theorem.

[74] Recall that the 4th link is logically equivalent to the 1st one.

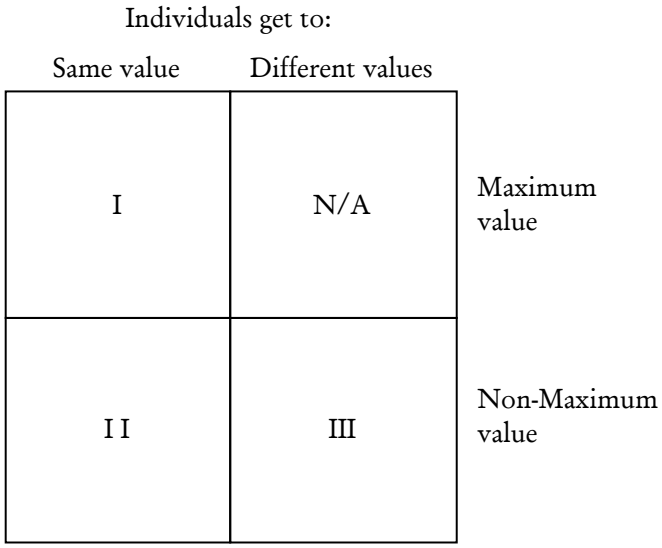


Fig. 3.7 – Partition of the logical space by means of Grafen's first three links (see text for explication).

Even if all these criticisms could be satisfactorily answered, however, Grafen's model would still provide an incomplete account of how organismic architectures arise and evolve. In order to show why and how its limitations could be overcome, let us first consider the main logical hypotheses behind the project (Fig. 3.8):

1. Natural selection is the only evolutionary force contributing in a consistent way to shape the phenotype. Positive contributions by other forces are purely accidental (e.g., a particularly positive mutation, or a fortunate drift episode). Genetic details just act as constraints that might delay the appearance of the optimal phenotype.
2. Selection fine-tunes and combines the organism's traits in such a way as to maximize individual fitness.
3. The improvements in the phenotypic design of the organism consist in the movement towards the optimal mix of phenotypic traits.

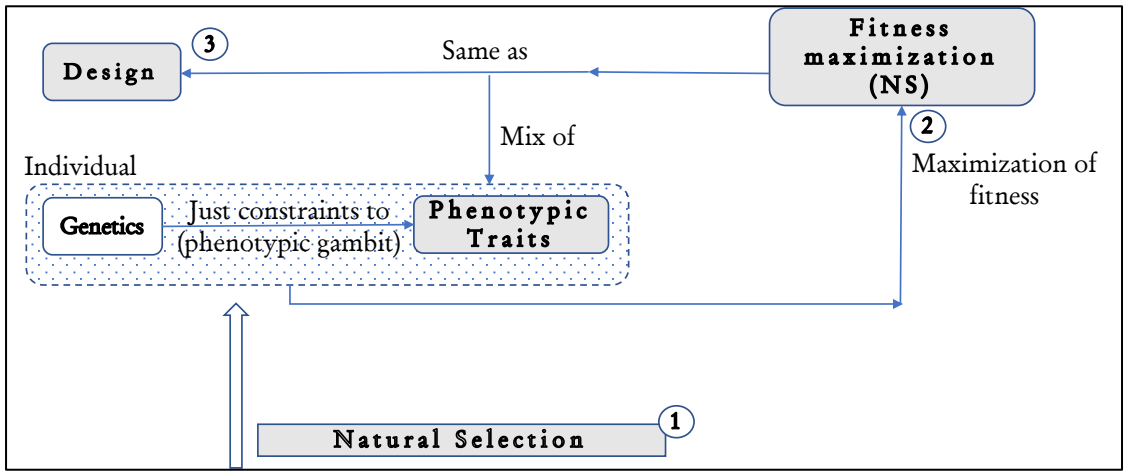


Fig. 3.8 – Logical structure of Grafen's project.

1. The prominent role of natural selection

Grafen's focus on selection as the sole, all-mighty and optimizing evolutionary force is widespread among biologists: Erwin (2015) believes that it dates back to the founders of the modern synthesis, who claimed that all evolutionary innovation appears through adaptive radiations driven by ecological opportunity rather than morphological novelties. Although Grafen reckons that it is a central idea in the *Origin*, arguably the original Darwinian theory and the initial neo-Darwinian synthesis did not grant natural selection an exclusive role in the shaping of phenotypes. Gould analyses in detail, through the comparison of the early and later works of Wright, Dobzhansky, Simpson and Mayr (Gould 2002:518-541), the phases of what he defines 'the hardening' of the synthesis: the progressive central role granted to natural selection and the reduction of any alternative process to irrelevant noise in evolution. Consequently, adaptation became an *a priori* assumption for the evolution of any trait and phenomena studied at a few generations timespans were extrapolated to explain evolutionary changes covering millions of years (Gould 2002: 521). Gould, via a meta-theoretical use of the concept, explains this hardening by means of a 'founder effect' hypothesis: the community of evolutionary biologists was small and stratified in the 1940s, facilitating the general acceptance of the new focus on natural selection (Gould 2002: 543). An example of this widespread bias is the treatment of phenomena like the parallel evolutions by adaptive radiation[75] of marsupials in Australia and of placental mammals in America, which produced similar morphological forms as responses to similar ecological opportunities, as evidence for the universal power of natural selection (Kimura 1983: 61-62): internalist mechanisms (like developmental laws) could provide an equally valid explanation. Another example from the literature is in Brooks and McLennan (1991). When trying to explain why one- and two-horned rhinos species coexist, with the one-horned species seemingly the better adapted, they suggest that "[t]he 'whys' of the two-horns must be examined within the context of the ancestors and the environment in which the trait originated", through "detailed knowledge of past environments" (1991: 146): an adaptationist explanation is taken for granted, it's up to us to find it by collecting more details about the history of the species' environment.

In this context, Lynch affirms that "The literature is permeated with dogmatic statements that natural selection is the only guiding force of evolution" (2007: 8598). Other authors likewise criticize Grafen's claim: Orzack (2014: 262) underlines that evidence in favour of selection is not evidence against the existence of other relevant evolutionary forces

[75] Adaptive radiation is the evolution of ecological and phenotypic diversity within a rapidly multiplying lineage (Schluter 2000: 10)

impinging on traits' evolution. Drift, for example, is ignored in the project[76] by affirming that it cannot improve adaptation, a dogmatic claim that should be tested, for example, through comparison of the predictive capacity of a model including drift and another excluding it. Orzack (2014) suggests the same methodology to test the primacy of selection.

The idea - formalised in Grafen's model - that selection is able not only to sort out the optimal mix of traits, but that it can also limitlessly generate new traits and new traits' versions, implicitly reduces the logical space of applicability of the Formal Darwinism Project. We clarify this point by resorting to Pigliucci's classification of evolvability (2008), which also includes a description of scales of evolution, based on the evolutionary forces predominating in each (Fig. 3.9).

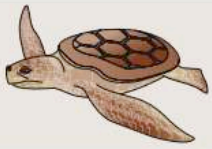
Suggested term	Scale	Description	Effects	Example
Heritability (*sensu* Houle)	Within populations	Standing pool of genetic variation and covariation	Determines the response to natural selection within populations	
Evolvability (*sensu* Wagner & Altenberg)	Within species	Includes variability (*sensu* Wagner & Altenberg), depends on genetic architecture and developmental constraints	Affects long-term adaptation, channels evolution along non-random trajectories, allows mid-term exploration of phenotypic space	
Innovation (*sensu* Maynard-Smith & Szathmary)	Within clades	As for within species, but includes the capacity to overcome standing genetic and developmental constraints, opening new areas of phenotypic space for further evolution	Generates major phenotypic (morphological, behavioural or physiological) breakthroughs (novelties)	

Fig. 3.9 – Classification of evolutionary changes (from Pigliucci 2008a)

At the lower level (*heritability*), evolution consists in changes in alleles' frequencies between generations of a given population, solely or mainly fostered by selection and described by population genetics models. Traits are known before and after the evolutionary phenomenon, and their optimal mix is shaped by selection: Grafen's Formal Darwinism has its most fruitful application at this level. The evolution of the *B. betularia* wings' color is a classic example of such a phenomenon: the trait and its (qualitative) variants are well defined (white and black wings), and their relative frequencies are the output of a process of selection. Still, it is at the higher levels, labelled by Pigliucci as *evolvability* and *innovation*, that new traits and functions appear. Processes linked to genetic and phenotypic architecture, or to developmental dynamics, seem to play a relevant role at these levels, while selection's role

[76] Grafen considers drift in at least two of the papers related to the Formal Darwinism Project (Grafen 2002, Gardner & Grafen 2009), only to demonstrate that its statistical effect has no importance when averaged and that, as a consequence, it can be ignored.

is possibly secondary (Albrech 1991, Houle 1992, Wagner & Altenberg 1996). It is not evident that Grafen's model can be applied to this juncture without substantial modifications, given that its hypotheses do not straightforwardly apply: selection does not act alone nor it is the dominant force, and the phenotype cannot be summarised in terms of a the list of constituent traits and their values. Phenomena occurring at these levels and reference to non-selective processes possibly causing them will be introduced in chapter 4, while a model encompassing them and extending Grafen's model will be proposed in chapter 5.

2. Fitness maximization.

The second fundamental idea behind the Formal Darwinism project is that selection maximises individual fitness. However, we believe that a bigger conceptual problem is the unclear definition of fitness used by Grafen. Let us first quickly mention how maximization is questioned in the literature. Abrams (2001) underlines that the majority of species are far from optimality. Ewans (2014: 202-204) does not believe that any genetic and mating system (to quote the main elements considered in Grafen project) leads by definition to fitness maximization, an issue that theory alone, without analysis of real data, cannot settle. Moreover (198), this author claims that Fisher's theorem allows for fitness' decrease when fitness depends on more than one gene and locus; interestingly, note that Fisher never explicitly talks about maximization. Okasha and Paternotte (2014: 255-256) underline that Fisher only claims that fitness increases under selection if environmental conditions, including other genes' effects, remain constant. They also observe that average fitness (Fisher's fitness measure) and individual fitness (Grafen's fitness measure) can as a matter of fact take divergent paths: the former can be maximised even if the individual ones are not, and vice versa (as in evolutionary theory of games). Gilchrist & Kingsolver (2011) likewise underline that optimality is local and referred to individual values in adaptationist models, while in in population genetics models it refers to average values. Adaptive peaks in the two approaches are in general different, especially when fitness landscapes are rough[77], and maximization corresponds to different phenotypes. Sarkar (2014: 255-256) suggests reformulating the project leaving aside fitness optimization, an idea that does not appear in Darwin[78]. In this weaker version of the model, selection improves the adaptedness of the phenotype but only until a certain point: the final theoretical optimal fitness could be out of

[77] That is, they have many peaks connected by steep flanks.

[78] See §4.6.2 for an account of the differences between Darwin's original idea of 'fit' and the current biological concept of fitness.

reach given the initial conditions of the population, as it happens in cases of frequency dependent selection. A similar idea is defended by Lewontin and Goss (2005): due to the impact of environmental conditions on the relationship between genotype and phenotype, the selected genotype will be the one that, considering the probability distribution of environmental variables and developmental noise[79], maximizes the *chance* of producing the optimal phenotype, although the actual phenotype might finally be less-than-optimal.

A more subtle problem is linked to the polysemy of the concept of 'fitness': one can talk, for example, about genotypic or phenotypic fitness (Lehmann & Rousset 2014), or about the fitness of an individual or the average fitness of a population (Ewens 2014), in which case maximization of the former and of the latter do not necessarily coincide. Grafen swings in his writings between ecological and reproductive fitnesses (Sober 2009). Let us now see what they are, why this 'swinging' is a problem in the Formal Darwinism model, and how it could be solved.

Ecological (or trait) fitness refers to a single phenotypic trait and measures its contribution to the individual organism's reproductive success, through the a priori analysis of its 'good design'. As Gould (1976) puts it, certain traits (morphological, physiological and behavioural) grant an a priori advantage for surviving in a specific environment, and can be judged as superior by engineering criteria. Adaptationist accounts apply this concept, for example, when justifying the giraffe's neck for its contribution to survival and, eventually, reproduction. Considering that it represents the expected future performance of a trait, ecological fitness is a dispositional concept and, as such, elusive to direct measuring. Moreover, traits never appear isolated (Fig. 3.10), and fragmentation of an organism in single traits is a subjective task (is the length of a limb a trait? Or is it the length of each limb's element?). Organisms are not the simple sum of modular phenotypic traits, but a whole ensemble (Dupré 2006: 66). Altenberg underlines this aspect: "[…] the process by which traits are distinguished from one another is a human measurement process" (Altenberg p. 100 in Callebaut & Rasskin-Gutman 2005). Attempts to define a trait as "a phenotypic variant that results in the highest fitness among a specified set of variants in a given environment" (Reeve and Sherman 1993) are clearly biased towards an adaptive view.

[79] Any variable that can affect development in unpredictable ways, so that different phenotypes can result from the same genotype in the same environment: e.g. human finger-prints.

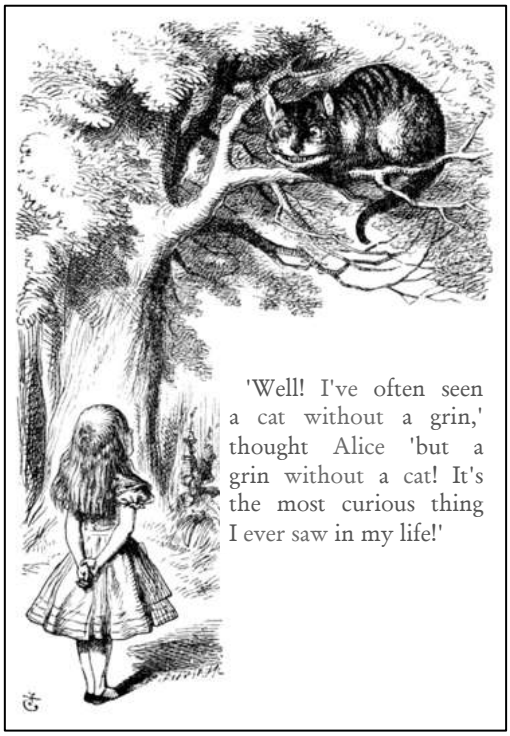


Fig. 3.10 – Alice and the Cheshire cat: individuation of a trait (the grin) without the individual (the cat) is a curious thing indeed!

Reproductive (or individual) fitness has peculiarities mirroring the ecological concept. It is a concept linked to the whole individual and to its actual reproductive success, measured a posteriori. By measuring its past performance through the count of its offspring, its measurement is objective and does not present the theoretical complications typical of the ecological fitness concept (although the practicalities of the measurement are far from obvious)[80]. These characteristics of the concept allow for an easy integration of individual fitness in formal models: indeed, this is the concept of fitness used in population genetics. Reproductive fitness has nevertheless a couple of problems too. In the first place, without a criterion of fitness independent of reproduction, it turns the Darwinian argument into a tautology: the fittest individual has more offspring, and it is the fittest because it has more offspring[81] (Bethel 1976). Secondly, by considering the individual as a 'black box' that transforms genetic frequencies into fitness, it seems inadequate to explain the organism's

[80] Individual fitness is here interpreted as the actual number of offspring of an individual, that is, neither as a probability nor as an estimation of probability based on actual frequency, as it happens, for example, in Sober (2013).

[81] Ecological fitness, by recurring to a priori engineering criteria, avoids the tautology. See §4.6.2

architecture, given that the details of the architectural structure are in some sense to be found *within* the 'black box'[82].

The following table summarizes the peculiarities of these two kinds of fitness.

Table 3.1

	unit	focus	performance	used in
Ecological (or Trait) Fitness	trait	a priori	expected, future	adaptationist accounts
Reproductive (or Individual) Fitness	individual	a posteriori	measured, past	population genetics

Peculiarities of ecological and reproductive fitness concepts.

Confusion between the two concepts is not a novelty. In the diatribe between Bethel and Gould (Bethel 1976, Gould 1976), the two authors are really talking about different fitnesses: the former considers evolution by natural selection a truism because he considers the problem of individual fitness; while the defence of Darwinism by the latter is based on the virtue of trait fitness. Both seem to forget half of the story.

Grafen too appears to be swinging between the two views of fitness[83]: on the one hand, he claims that natural selection leads to the optimization of the individual organism's (or reproductive) fitness[84]; on the other hand, he seems to interpret design as the mix of traits that results in the highest adaptive performance, recurring to trait (or ecological) fitness. This latter interpretation is, in fact, quite logical, because optimality models are about traits, not about token individuals (Sober 2013).

In the last section of this chapter (§3.7), we propose a way to unify the two concepts of fitness.

3. Design as fitness optimization.

Even if selection were to optimize fitness, and the concept of fitness were well defined, the causal relationship between optimization and design would still not be evident (Birch 2014).

---

[82] Fitness, its interpretation and its explanatory role in evolution are highly debated (see e.g. (Matthen & Ariew 2002. Abrams 2012).

[83] Ewens (2014) suggests a similar criticism when he shows that Grafen interchanges the concepts of individual and average fitness. We believe that the unifying framework proposed in this section answers that criticism too.

[84] Grafen (2014) claims: "We are dealing with the maximization of individual fitness. Correspondingly, the evaluation of a phenotype is the number of offspring an individual with that phenotype has."

An optimised fitness does not necessarily reflect an improved design or an increase in its complexity (however defined), in the same way as a molecule that spontaneously reorganizes from a less to a more stable configuration following the optimization of some thermodynamic potential does not either. Grafen's implicit definition of design rests on the hypothesis that selection improves adaptation, and that adaptation might lead to maximization of fitness, through the configuration of the best value of each trait: the resulting phenotype is 'well designed'. This definition reduces design to a 'fine-tuning' of pre-defined variables (the traits), organised in ways that do not change in a relevant way and can be thus considered to configure fixed architectures. While coherent with the idea of optimization by selection, can this view encompass phenomena like the appearance of new traits, or a substantial change in the relationships among traits? The idea of 'improved design' is also poorly treated in Grafen's model: the change in wings' colour in *B. betularia* represents a new fitness optimum and, as such, an improvement in adaptation; but in what sense is it also an improvement in design?

If optimization and design are linked, the formalization of such a link deos not seem to be a trivial task.

Fig. 3.11 extends Fig. 3.8 by adding some of the main issues discussed above, that the project leaves open.

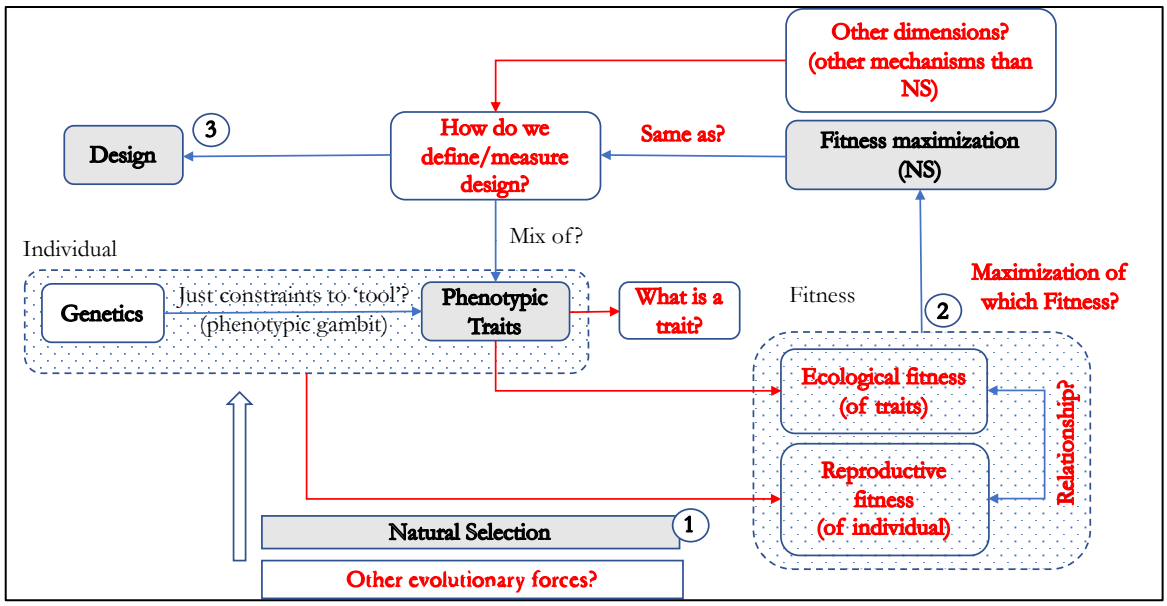


Fig. 3.11 – Open issues in Grafen's project (in red). What other evolutionary forces, apart from selection, shape the phenotype? Are the genetic details just constraints to the optimization of phenotypic traits or do they have a positive role? Is development relevant in evolution? Also: what is, and how do we identify, a trait? Which fitness is maximised, and how do individual and trait fitnesses relate to each other? Is fitness maximisation the same as design, or do we need other dimensions to measure it? Is design just a collection of traits?

3.7. The fitness matrix

In the discussion about fitness in Grafen's model, we have highlight a potential confusion between individual and trait fitness. We now propose a way to dissolve the confusion, so that the use of individual or of trait fitness in formal models becomes interchangeable. The two concepts of fitness are linked, and it is possible to make this link explicit through a matrix, built upon the idea that *individual fitness is a function of the ecological fitness of the individual's traits*. This function, however, contrary to what adaptationist and optimization models presume, will in general be:

- *Non-monotonous* for a given trait: incremental differences in the value of a trait's fitness might have incremental or decremental effect on the individual's fitness, depending on their initial values. For example, longer legs might increase the ability to run (away from a predator or after a prey), but legs too long might reduce fitness as they are prone to break (Sober 2009);
- *Non-linear* among traits (non-additive): there is in general a trade-off between traits, so that the positive contribution of a trait to individual fitness might cause the reduction of the contribution of another trait (see Sober 2000: 77-82 for a detailed example). The non-linear relationship is, furthermore, linked to phenomena of frequency-dependent fitness (Sarkar 2014).

The steps to estimate this function are (Fig. 3.12.a):

- the measurement of the reproductive or individual fitness (RFj) of each member j of the population;
- the somehow subjective fragmentation of the organism into traits i considered relevant for survival and reproduction (e.g., the length of a limb), the definition of their unit of measure (e.g., centimetre), and the actual measurement of their value for each individual (Wij);
- the estimation, through statistical regression, of the contribution of each trait to individual fitness (see e.g. Endler 1986): the corresponding value represents the trait's fitness (TFi).

As expected, *individual fitness is measured*[85]*, while trait fitness is estimated* (Fig. 3.12.b). This estimation is, however, performed ex-post by means of the following proposed methodology: the contribution of traits is not estimated by engineering design alone, although the identification of relevant traits implies some degree of engineering considerations.

a

	Individual 1	Individual 2	…	Individual j	…	
Trait 1	W11	W12		W1j		Ecological Fitness 1
Trait 2	W21	W22		W2j		Ecological Fitness 2
…						…
Trait i	Wi1	Wi2		Wij		Ecological Fitness i
…						…
	Reproductive Fitness 1	Reproductive Fitness 2	…	Reproductive Fitness j	…	

b

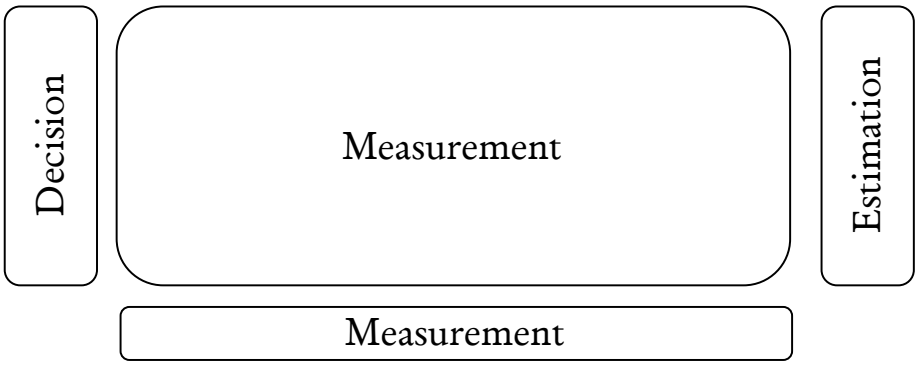


Fig. 3.12 – (a) Matrix linking individual and trait fitnesses. Each column represents an individual and each line represents a trait. The coefficients Wij measure to which extent an individual j possesses trait I (e.g. the length of its legs). Based on Wij and the individual fitness of each individual, traits' fitness can be estimated.
(b) Epistemological analysis of the matrix: each of its areas is the output of a different epistemological step: the division of the phenotype into traits is a somehow subjective *decision*; the coefficients of the matrix and the reproductive fitness are *measurements*; ecological fitness are statistical *estimations*, not measurable directly.

The advantage of this approach is that the value of each trait is measured 'in the field', interacting with all others: something that engineering considerations alone could hardly

[85] For a very different approach, see (Sober 2013), where the author claims that individual fitness is practically impossible to measure and theoretically useless to predict evolution, claims challenged by Pence and Ramsey (2013).

accomplish. The so-built function thus represents a proxy of the internal architecture of the organism: of its traits *and* of their interactions. It is therefore a tool that:

- Helps reducing the gap between adaptationist arguments (that appeal to trait fitness) and population genetics models (that use individual fitness);
- Eliminates the confusion in Grafen's model, by allowing the double interpretation of fitness as individual or as output of the best mix of traits.

In chapter 5, we will apply this unified interpretation of fitness in the model we propose to unify adaptationist and pluralistic evolutionary explanations (illustrated in chapter 4).

3.8. Conclusions

In this chapter, we have analysed how adaptationism explains the appearance and evolution of phenotypic traits solely through the process of selection. We have mentioned the main internal[86] problem of adaptationism: the purely narrative nature of its explanations, that makes them hard to test, and therefore not 'scientific'. Aware of this pitfall, adaptationism builds formal models by borrowing the concept of fitness from population genetics and reformulating it. We have introduced some of these models: optimization programs and the Formal Darwinism Project, the most ambitious attempt to make adaptationism a theoretical discipline. We believe that all these models show external limitations that make them applicable only to a narrow set of evolutionary phenomena, mainly:

- The a priori hypothesis that selection is the only relevant force in evolution; or, in other words, that all phenotypic traits are adaptation.
- The identification of phenotypic design changes with fitness changes; or the reduction of design to a mix of fine-tuned, pre-defined traits.

In the next chapter, we will introduce other potential, non-selective, sources of phenotypic change, and we will show how they might contribute to explain the evolution of phenotypic traits. In the final chapter, we will propose an innovative way to formalise these other sources without recurring to fitness optimization considerations; in this sense, we will map selective and non-selective processes in our model of phenotypic architecture (see

[86] 'Internal' because it is not linked to any adaptationist assumption, but to their formalization. Adaptationism acknowledges such internal problems. When we talk about 'external' issues, we refer to limitations caused by adaptationist assumptions, e.g., supremacy of selection, that adaptationism does not consider.

§2.4.2 and 2.4.3) and propose a unified model encompassing both adaptationist and non-adaptationist explanations.

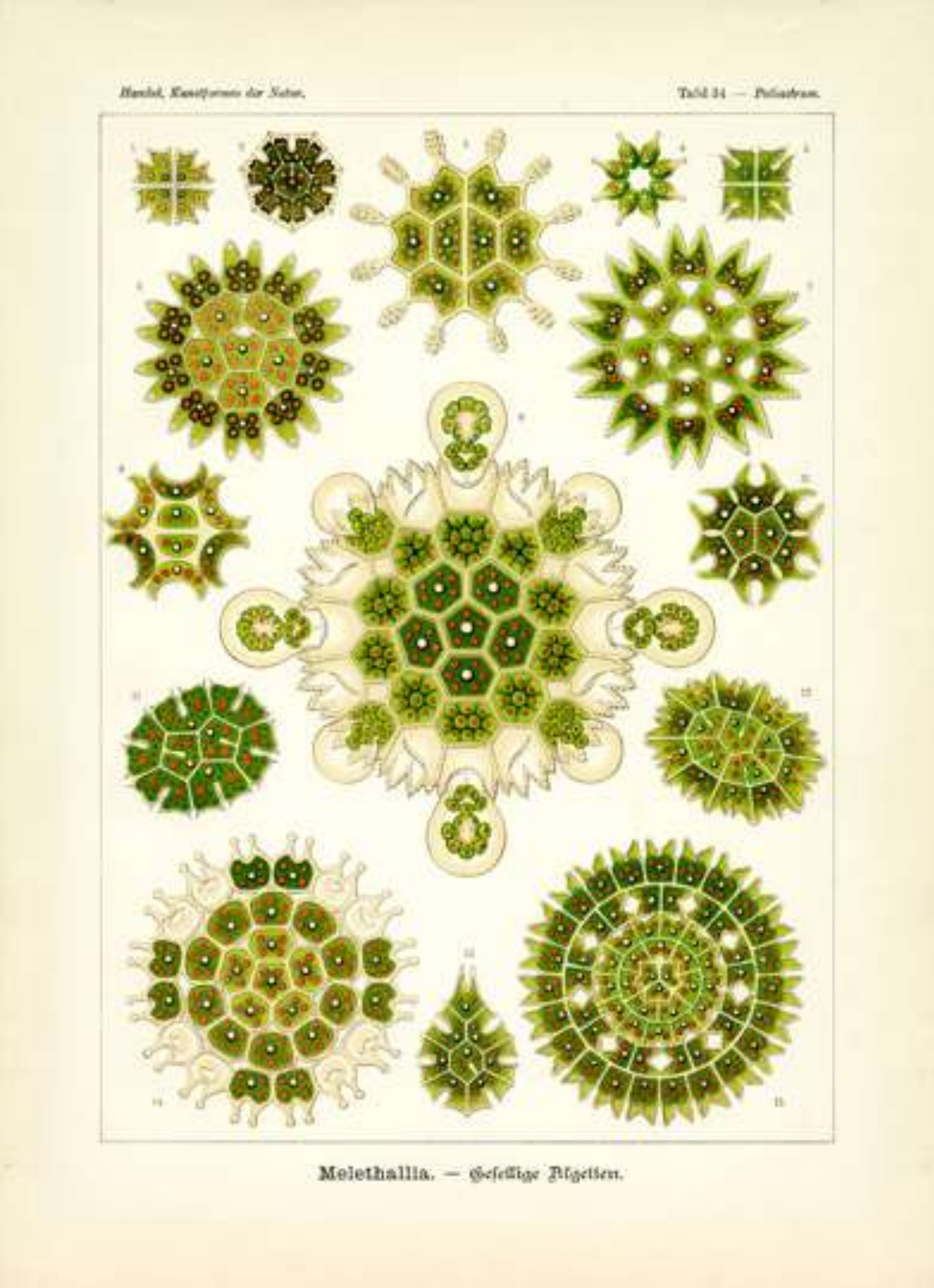


Plate 34 – Melethallia, from *Report on the scientific results of the voyage of H.M.S. Challenger during the years 1873-76 under the command of Captain George S. Nares ... and the late Captain Frank Tourle Thomson*, R.N. Edinburgh: Neill,1880-1895.

4. Pluralism: beyond the limitation of adaptationism

Evolution is a fact, but what originates and drives the changes behind it? One thing is to accept the hypothesis that the process of natural selection is part of the explanation of evolutionary change and a different thing is to claim that, through the accumulation of slow, incremental and random fitness-increasing variations, selection is a *necessary and sufficient* condition for the appearance of novelties in evolution, as empirical adaptationism[87] defend (Orzack & Forber 2010). One can soften this latter, extreme position without denying the importance, and even the preponderance, of adaptation in nature. Dupré warns that "we must not assume that any feature of an organism that attracts our attention is an appropriate subject for evolutionary explanation." (2003: 28), especially when the feature's function is not clear beyond any doubt. The giraffe's neck is the result of the "full history of the lineage complete with varying ecological and climatic circumstances, not to mention brute chance." (2003: 38). Even neo-Darwinian biologists admit that: "[…] there is more to evolution than an increase in adaptation" (Maynard Smith et al. 1985).

In this chapter, we perform a constructive critique of adaptationism (§4.1), that starts by showing the limitations of its scope and hypothesis to introduce some of the main accounts proposed to overcome them (§4.2, 4.3, 4.4, 4.5). These proposals, broadly unified under the umbrella concept of 'pluralism' (the idea that evolution is caused, together with natural selection, by other laws and processes, as well as by historical contingencies, Gould 1997), are quite diverse, although all of them explain drawing on factors other than fitness differences or optimization. Finally, we analyse the reaction of adaptationism to pluralism and the putative incompatibility between them, only to suggest that it is not inevitable (§4.6). We conclude (§4.7) introducing our proposal of unification, that is developed in the next chapter.

[87] In this chapter, we will refer to empirical adaptationism every time that the term 'adaptationism' is used. For a definition of the different types of adaptationism, see §3.1.

4.1. Critiques to adaptationism

We believe that adaptationism ignores many phenomena, processes and laws that affect organisms (§4.1.1), and that this attitude rests mainly on the uncritical acceptance of the four pillars introduced in chapter 3, and scrutinised here (§4.1.2). The result of this scrutiny is a broader view of evolutionary forces, as considered in the evolutionary extended synthesis research program (EES), a pluralistic approach that aims at integrating some of these non-adaptationist accounts of evolution proposals within a renewed synthesis (§4.1.3). We then underline the differences in attitude of many of these proposals with regards to adaptationism: namely, their structuralist and/or internalist view (§4.1.4). We propose a taxonomy to classify some of the main non-adaptationist proposals (§4.1.5), introduced in §4.2, §4.3, and §4.4.

4.1.1. What natural selection leaves uncovered

Empirical adaptationism seems to substitute one unlimited creative power for another: "[Darwin] killed off God [...] but Mother Nature and other pseudo-agents got away scot-free" (Fodor & Piattelli-Palmarini 2010: 163). Some mystical claims of adaptationist biologists are good example of it (Fig. 4.1).

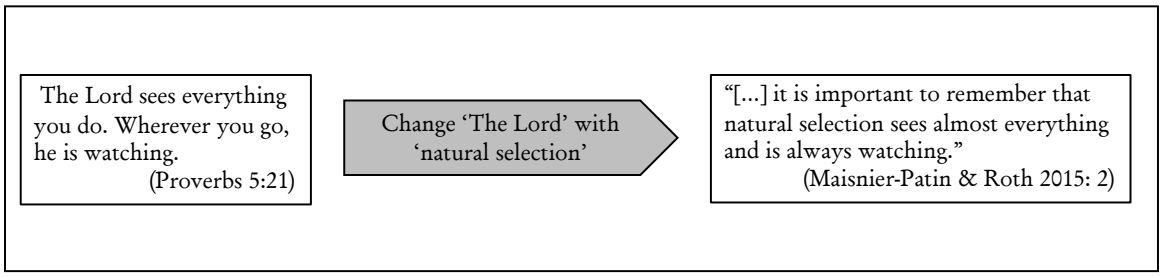


Fig. 4.1 – Biblical and adaptationist mysticisms.

Natural selection can solve almost any environmental problem encountered by the organism and ignore almost any interference and constraint. This attitude (there is always an optimal solution, and selection will sooner or later provide it) is especially evident when other possible sources of evolutionary change are considered. Brooks and Wiley (1986: 3), for example, believe that the adaptationist focus on natural selection as the primary force behind biological organizations and evolution leaves four areas uncovered:

1. The chemical and physical basis of natural selection. Organisms are first of all material entities, and natural selection cannot overcome physical, chemical and

thermodynamic laws binding them. Brooks and Wiley as well as many others interpret evolution as an entropic process (see section §4.5.2 below). D'Arcy Thompson (§4.5.1) underlines some recurrent relationships between physical magnitudes in organisms. The 'zero force evolutionary law' claims to describe a natural tendency of any system possessing certain features (§4.5.4).

2. The integration of developmental biology within population genetics. The evo-devo research program (§4.4.1) underlines the impact of developmental pathways and constraints on evolution.
3. The explanation of the appearance of higher taxa. The Punctuated Equilibrium hypothesis (§4.2.3) challenges the idea that the sum of micro-adaptations can result in the gradual creation of new species. Accounts focusing on organisms as systems (§4.5.3) or networks (§4.2.5) suggest that these have features that naturally lead to increasing complexity.
4. The relationship between form and function in evolution. D'Arcy Thompson (§4.5.1) underlines the recurrence of forms in distant taxa and species. Some neo-Lamarckian (§4.3.3) and epigenetic accounts (§4.3.4) reverse the adaptationist dogma that function always precedes form. Exaptations (§4.3.1) and phenomena of genetic assimilation (§4.3.2) point in the same direction.

The processes and mechanisms captured by the non-adaptationist accounts are therefore at the root of many aspects of biology that are superficially explained by natural selection (Lynch 2007).

In the next section, we go back to the adaptationist pillars to see what pluralism has to tell about them.

4.1.2. Pluralism and the pillars of adaptationism

Non-adaptationist accounts question one or more of the four pillars of adaptationism (presented in §3.1): ubiquity, continuous, slow and incremental change, random and unconstrained variation, and supremacy of selection.

Ubiquity. Regarding the ubiquity, or pervasiveness[88], of natural selection, Darwin himself (1883) admitted that trivial characters which are "really of no considerable importance in the struggle for life, […] could not be modified or formed through natural selection". He believed this was especially clear in case of polymorphic traits with no apparent adaptive role: "I am inclined to suspect that we see in these polymorphic genera variations in points of structure which are of no service or disservice to the species, and which consequently have not been seized on and rendered definite by natural selection […]" (Darwin 1859: 46). Although phenotypic polymorphism might hide an adaptation not immediately evident[89] (Halama & Reznick 2001), Mayr (1942) included it, together with most geographic variation, in the category of non-adaptive traits[90]. Crampton (1916, 1932), for example, claims that the different phenotypes of the land snail *Partula* are based on non-adaptive traits (see Fig. 4.2).

Fig. 4.2 - Possible phenotypes of the land snail *Partula*, whose different traits versions have no adaptive meaning, according to Crampton (196, 1932).

[88] Recall that ubiquity and supremacy of selections are different claims; the former means that selection acts on ALL traits, but does not exclude that other mechanisms could also act. Supremacy means that there are no other mechanisms than selection acting on evolutionary traits, but does not claim that all traits are evolutionary.

[89] The different phenotypes co-exists in what seems to be the same environment, but strict adaptationism would claim that it is not: there are not yet identified sub-niches in which each phenotype is optimally adapted.

[90] Mayr will later consider all traits as adaptive (Mayr 1963), after joining what Gould calls the 'hardening' of the synthesis (Gould 2002:536-538).

Others authors highlight the use by taxonomists of non-adaptive traits to classify species (Haldane 1932:113-114, Robson & Richards 1936). The neutralist theory (Kimura 1983) defends that many traits at the molecular level are selectively neutral. The most known critique of the thesis of the ubiquity of adaptation attacks its supposed 'panglossianism'91 (Gould y Lewontin 1979): it is always possible to build ex-post a narrative explaining why a species or a population shows a trait, but these narratives are usually impossible to falsify (Fig. 4.3).

In Lee et al. (2018) we can find an example of the paradigmatic attitude of the adaptationist approach to explaining traits.

After showing that a significant proportion of observed phenotypic variation of the coat pattern of the wild Masai giraffes is shared between mother and offspring, and it is thus an heritable trait, the authors analyse "whether variation in complex coat pattern traits was related to a measure of fitness (survival) and thereby infer the effect of natural selection (viability selection) on giraffe coat patterns" (Lee et al. 2018: 4). They claim that there seems to exist a statistical relationship between the trait and fitness, and postulate that the function of coat pattern is camouflage, only to conclude that "[w]hether or not spot traits affect juvenile survival via anti-predation camouflage, spot traits may serve other adaptive functions such as thermoregulation[…], or social communication […], and thus may demonstrate associations with other components of fitness, such as survivorship in older age classes or fecundity. Individual recognition, kin recognition, and inbreeding avoidance also could play a role in the evolution of spot patterns in giraffes and other species with complex coat patterns […]" (Ibid., 16). In a typical 'panglossian strategy', if one selective explanation does not work, they look for another.

Fig. 4.3 – Coat pattern in Masai giraffes: an example of 'panglossian' attitude in adaptationist explication of traits.

Mayr's statement that "one can never assert with confidence that a given structure does not have selective significance" (1963:190) is a paradigmatic example of such an attitude. The critique is based on Popper's denial that existential statements of the form '∃xQx' (like 'there

[91] Professor Pangloss, in Voltaire's *Candide*, following Leibnitz's claim that we live in the best of all possible worlds, finds a justification for any misfortune. Gould and Lewontin apply the term "panglossianism" to adaptationist narratives because these, as Pangloss's reasoning, justify *a priori* any existing trait as the best possible adaptation to the environment.

are unicorns'[92]) are scientific (Popper 2002: 47-50), as it is impossible to find a counter-example: one can always go on searching indefinitely for a unicorn, as one can always search indefinitely for an adaptive explanation for any trait. Moreover, if a selective explanation turns out to be wrong, the reason can be ascribed, apart from the inappropriateness of the theory, to unreliable data[93] (Leigh 2001); and, given that a theory usually resorts to auxiliary assumptions, these can always be blamed for its failure[94].

Continuous, slow and incremental change. The idea that natural selection acts progressively on long timescales is challenged, according to Eldredge and Gould (1971), by the gaps we find in the fossil record, showing how "gradual, progressive evolutionary change is a rare phenomenon". These gaps are not to be interpreted as lost information about the gradual transformation of phenotypes, but as evidence that evolutionary changes happen 'quickly' between long periods of stability. Contrary to the idea of slow rates of evolution, Thompson (1998) lists several examples of evolutionary phenomena whose timespan is less than hundred years, in some cases less than one year: for example, changes in beak morphology of the *Geospiza* finches following change in seed availability took less than 10 years (Grant & Grant 1995). Other recent studies show very fast rates of phenotypic changes, suggesting that ecological and evolutionary times are not always incommensurably different, and thus that evolutionary change is not always slow and incremental (Hairston et al. 2005). Cases of loss of traits following relaxation of selective pressure also seems to "provide the most significant exception to Darwin's generalization that evolution proceeds by small steps" (Lethi et al. 2009:494). Kimura (1983) suggests that the directional[95] selection step of

---

[92] 'There are unicorns' is logically equivalent to 'There is an object x such that x is a unicorn', of which '$\exists$ xQx' is the formal translation, where Q = 'is a unicorn'.

[93] The Copernican mathematical model is an historical example of a valid theory questioned because of unreliable data. The little accurate astronomical data available in the XVI century matched the prediction of the Ptolemaic model not worse than the heliocentric ones, and this fact represented a mayor argument against the new vision of the universe: "[F]or the planets, Ptolemy's predictions were as good as Copernicus' […] In fact, Copernicus' theory was not more accurate than Ptolemy's […] Until [*the improved data collected by*] Kepler, the Copernican theory scarcely improved upon the predictions of planetary position made by Ptolemy." (Kuhn 1962: 68, 154, 156; note in italics mine).

[94] Thanks to Davide Vecchi for this comment.

[95] Evolutionary biologists distinguishes different kinds of selection:

- *Directional* selection happens when an extreme phenotype is favoured and the corresponding advantageous alleles' frequency increases within the population (e.g. wings' colour of the B. betularia moth tends to become black when predators identify more easily white wings) (Martin et al. 1996). It entails some phenotypic property which alone can confer a fitness advantage if it changes in the right direction (Altenberg 2005: 110). Selection pushes the phenotypic trait towards the value with optimum fitness. Although Darwin thought mainly

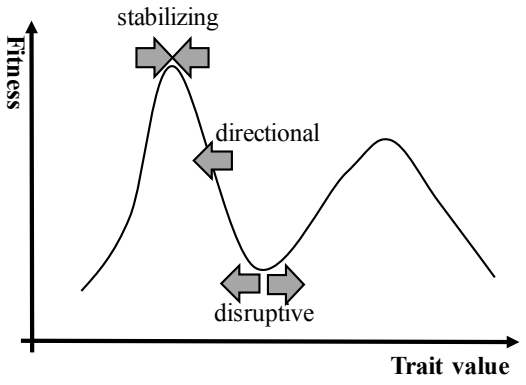
evolution occurs quite quickly, triggered by very relevant but rare environmental changes, and takes advantage of genetic variety already accumulated at the molecular level during the previous long period of environmental stability and stabilizing[96] selection. According to Fodor and Piattelli-Palmarini (2010: 52-54), gradualism makes a tacit assumption regarding the transitivity of fitness: if a phenotypic variant A has a fitness advantage over variant B, and variant B over variant C, then variant A also has a higher fitness than variant C. While fundamental for the fitness-landscape metaphor and the corresponding continuous 'hill-climbing' process[97], the authors claim that the assumption is false in most cases: when A, B and C are all present at the same time, it might happen that C wins over A. The *Achillea* plant, for example, shows variants of different height at different altitudes, without a clear relationship between trait and environment.

McCandish and Stoltzfus (2014) criticises the 'origin + fixation' model that underlines much of Neo-Darwinian theoretical work: a mutation introduces a small, incremental variant, and selection fix it if selectively advantageous. They claim that it captures only a

---

about this kind of selection, it is rarer than stabilizing selection and limited to situations in which the evolutionary phenomenon is triggered by a change in the environmental conditions (Kimura 1983: 120);

- *Stabilizing* selection, a mode of selection opposite to directional selection, maintains the genetic pool of a population and eliminates extremes phenotypes (e.g. weight at birth is under stabilizing selection, as babies above or below the average weight have greater mortality Karn & Penrose 1951) (Martin et al. 1996). It entails some phenotypic property which alone can impose a fitness disadvantage if it changes in any direction away from its current value (Altenberg 2005: 110). Individuals possessing the mean value of the trait have the highest fitness; individuals with higher or lower value are selected against. When it entails the elimination of a deleterious allele, it is also called *purifying, or negative* selection.
- *Disruptive* selection occurs when there are two or more optimal values of a phenotypic trait: the population segregates into two groups. Disruptive selection can produce polymorphism and eventually sympatric speciation (Mather 1955).



[96] See previous note.

[97] See §4.2.2

small part of evolutionary dynamics, such as "the evolution of quantitative traits based on abundant standing variation, overdominance (which
can result in prolonged polymorphism) […] interference between multiple beneficial mutations" at either a single locus or several partially linked loci (Íbid., 238).

Gradualism is also denied by Koonin (2009): single gene duplications, horizontal gene transfers, deletion and acquisition of larger regions, genome rearrangements, whole-genome duplication, etc. are not infinitesimally small! "Gradualism" he concludes "is not the principal regime of evolution." (Íbid., 1027).

Random, unconstrained variation and the supremacy of selection[98]. The idea that random, unconstrained variation[99] (genetic mutations, recombination, etc.) can supply all the material upon which natural selection alone creates any new trait and function has been questioned since the dawn of Darwinism. However, the idea that variation can occur in any direction is "too useful a simplification to give up" (Kauffman 1993: 11). Cope, in *The Origin of the Fittest* (1887), underlines the difference between the laws that restrict, direct and destroy a structure and the laws that originate it: natural selection might account for the former, but cannot explain the latter. This is the general tenure of the critiques against empirical adaptationism, or the idea that without selection there could only be chaos: their critiques do not deny that natural selection affects the frequency of traits in a population, but doubt that it is the (main) source of novel traits. Lynch, one of the most renowned experts in population genetics considers mutation, recombination and drift as forces in the same category as selection in shaping an evolutionary pathway, and even claims that some genetic and developmental details cannot have an adaptationist origin: adaptationism is described as "permeated with dogmatic statements" (Lynch 2007: 8598) and "adherence to the adaptationist paradigm" is "religious" (ibid., 8599). Some authors go one step further and doubt that natural selection can be the source of any real innovation at all: like artificial selection, it can shape existing traits and maintain them in a population, but cannot create new ones. The denial of any creative power of natural selection – a position that could be labelled Negative View -, has a long tradition (a thorough list of authors that challenge the creative view of natural selection is presented in Razeto-Barry & Frick 2011). De Vries, one of the precursors of the modern

[98] We treat the critiques to the last two pillars together because they are more closely linked than the others: see (Gould 2002) and put discussion at the end of §4.6.

[99] Recall that 'random' and 'unconstrained' are two different, independent concepts. See §3.2.

synthesis and the postulator of genes as information carrier for traits[100], believed that natural selection can explain the survival, but not the arrival of the fittest. Waddington (1967) underlines that the mathematical theory behind Neo-Darwinism does not explain how "horses and tigers and things" appear. Kauffman (1993: 173, italics mine) believes that "much of the order we see in organisms may be the direct result *not of natural selection but of the natural order* selection was privileged to act on". Neander (1995) lists several arguments in favour of the Negative View. Here we mention just the 'counterfactuals' one: any organism exists because of the line of its ancestors, regardless of which other organisms existed besides, and of their destiny[101]: a world with unlimited resources (thus with no selection) would be populated by all kind of creatures, even by the ones that, in the current, competitive world, were eliminated by selection. A similar idea is defended by Morgan (1925: 127), who writes that selection does not produce anything new, only more of something that already existed. He also claims that, in a world without natural selection, existing traits would co-exist with traits ruled out by natural selection (Morgan 1932). Macroevolutionary lags between the origin of a novelty and its adaptive success represent another argument against the pervasive action of selection in the formation of novelties: for example, grassland spread tens of millions of years after its appearance, showing that the origin of novelties is decoupled from their ecological success (Erwin 2015). According to Moczek (2008), the last century has seen great advancements in the understanding of how phenotypic traits diversify, not in how they appear. After decades of investigations around genes and the genotype, we know much about how traits and innovations spread through populations, but how innovations appear is still, to a great extent, a mystery (Lynch 2007). And Wagner (2015) believes that the real mystery of evolution is not the selection, but the creation of new phenotypes. In the same line, Rasskin-Gutman & Esteve-Altava (2009) claim that natural selection just explains 'what is successful among the possible', while other processes and mechanisms explain 'what is possible among the conceivable' (Rasskin-Gutman & Esteve-Altava 2009). Denton (1985) challenges the idea that the random search in a limited time postulated by adaptationists can give rise to the design of life that we find in nature: not even the probability of discovering by chance a single protein molecule has been calculated, let

[100] In his book *Intracellular Pangenesis* published in 1889, De Vries proposes that the carriers of hereditary traits are particles, that he called 'pangenes', from which the current term 'gene' comes from. See Heimans 1962 for an historical overview.

[101] "This premise is implicitly based on the existence of just vertical inheritance; but with horizontal gene transfer it crumbles. In fact, not all organisms exist because of the gen-phenotypic legacy of their ancestors. Only multicellular and sexually reproducing organisms might solely rely on vertical inheritance." (Davide Vecchi, personal communication).

alone the probability of complex organisms. Even though Fisher sees natural selection as "a mechanism for generating an exceedingly high degree of improbability" (Huxley 1936), a probabilistic account of the putative creative power of natural selection is still lacking (Razeto-Barry & Flick 2011). Gould (1993), commenting on Kimura's neutral theory (1983), believes that the importance of selection and adaptation is magnified by our vision at the organism's scale: at the scale of genes, "randomness might predominate". Darwin himself explicitly claims that "natural selection has been the main but not exclusive means of modification" (1859: 5-6)[102] and that "our ignorance of the laws of variation is profound" (1859: 151). Optimality, or near-optimality, does not exclude the action of non-selective forces: the trait can be the net result of competing mechanisms (Abrams 2001). Muller (1949) believes that, even in the absence of natural selection, "higher forms would nevertheless have arisen". Pigliucci et al. (2006) distinguish between the ultimate cause of adaptation during evolution (natural selection) and the proximate cause of change (e.g. developmental plasticity): selection acts on variation due to plasticity. And to conclude, Lewontin claims that

> […] it is by no means certain, even now, what proportion of all evolutionary change arises from natural selection. (1974: 3)

4.1.3. An instance of a pluralist program: the Extended Evolutionary Synthesis

The expression 'Extended Evolutionary Synthesis' (EES) identifies a conceptual framework for evolution more comprehensive than the Modern Synthesis (MS) (Pigliucci 2007, Pigliucci & Müller 2010, Laland et al. 2015). The EES aims at including, rather than discharging as irrelevant, all mechanisms, processes and constraints that directly or indirectly affect the 'descent with modification', that is, among other elements, the origin of variation, and the direction and rate of evolution. Pigliucci (Pigliucci & Muller 2010) claims that biology proceeds by continuous expansion of its empirical and conceptual structure: the EES is just the last of a series of steps, within which many of the approaches analysed in this chapter fit

[102] The misinterpretation of Darwin's though must trace back to his own time, if he felt compelled to write, in the final edition of the Origin (1872):

> [As] it has been stated that I attribute the modification of species exclusively to natural selection, I may be permitted to remark that in the first edition of this work, and subsequently, I placed in a most conspicuous position—namely, at the close of the Introduction—the following words: "I am convinced that natural selection has been the main but not the exclusive means of modification." This has been of no avail. Great is the power of steady misrepresentation. (p. 421)

(Table 4.1). Others (Craig 2010) defend a more extreme positions and claim that Evo-Devo research highlights concepts that create insurmountable problems to the Modern Synthesis: it is not a question of extending, but rather of getting beyond it.

Table 4.1

Step	Novelty
Darwinism	common descent and natural selection
Neo-Darwinism	no Lamarckism
Modern Synthesis	population genetics and its extension to palaeontology
Extended Synthesis	epigenetics, plasticity, evolvability, multilevel selection, niche construction, punctuated equilibria, developmental bias

Empirical and conceptual evolution of biology (Pigliucci & Müller 2010).

Laland et al. (2015) list six areas of assumptions where the EES overcomes the limitation of the MS:

- Instead of considering natural selection as the only *explanans* for the direction and rate of evolution, and for the fit between organisms and environment, developmental processes are given an equally important role in the reciprocal shaping of organisms and environment.
- Inheritance is not limited to genes, but encompasses epigenetic, physiological, ecological, social and cultural transmission. Acquired characters can play an evolutionary role by biasing selectable traits.
- The source of variation is not only random and not only genetic. Phenotypic variation can be biased due to developmental processes and constraints that trigger well-integrated responses to mutation and environmental changes.
- The rate of evolution is not always and only slow and gradual. Radically new viable variants appear not only through unlikely macro-mutations with direct phenotypic effect, but also thanks to mutations within gene networks, and to coordinated response of development systems to environmental challenges.
- Evolution is not just 'change in genes frequencies within a population' due to mutations, drift, gene flow and natural selection. The EES moves the focus from the gene to the entire organism, thus considering transgenerational changes that might not be immediately genetically encoded.

- Macro-evolution is not just the sum of many micro-evolutionary phenomena, and is driven by different processes other than just selection, drift, mutation and gene flow: developmental bias and ecological inheritance also intervene to configure macro-evolutionary patterns.

4.1.4. Structuralism and Internalism

Adaptationism, as illustrated in the previous chapter, can be seen as a functionalist and externalist approach: *functions* drive forms, and the *external environment* defines relevant functions (see e.g. Dwyer 1984). One of the problems with adaptationism under this view is that it provides *ex-post* explanations and definitions. The definition of ecological niche (what elements of the environment are relevant) and the identification of traits within that niche (how to split the organism in parts) depend on the adaptation whose existence is to be demonstrated. Instead of deducing a trait from some objective environmental description (e.g. what functions are needed to survive and reproduce in it), adaptationism infers the niche's constituents from the supposedly adapted trait, and then explains its functions and why they increase adaptedness. The following analogy neatly explains this critique:

> The order of metaphysical dependence is that keys solve the problem of finding something to open locks, not that locks solve the problem of finding something for keys to open. In adaptationist theory, by contrast, there's a sort of topsy-turvy: whether a feature of the environment constitutes an evolutionary problem for a creature depends on whether the creature's phenotype was selected for solving it. (Fodor & Piattelli-Palmarini 2010: 140)

To see this more clearly, it is useful to briefly review the concept of 'niche'. The term was first used by Grinnel (1904) to identify whatever biotic (e.g. predators) and abiotic (e.g. temperature) factors impacts the existence of a species in a given geographical area. Hutchinson (1957) redefines the concept as the set of environmental states that allow the survival of a given species, making the 'niche' an attribute of a species and not of the environment (Pocheville 2015). Mayr (1963) summarises the two positions when he claims that a 'niche' is definable either as a property of the species living in the environment, or as a kind of objective ontological entity incrusted in the environment. The objectivity of the niche in this second option, however, is illusory, because he admits that, to identify the relevant elements of the environment, on must recur to its niche-specific components, thus falling in a vicious circle. In both visions, the niche is finally defined in term of the species

adapted to it. Fodor and Piattelli-Palmarini (2010) believe that the concept of ecological niche is intentional: it is not an objective entity waiting to be filled by the first organism able to adapt to its challenges. Rather, niches (and the evolutionary problems they impose) are identified *ex-post* by reference to the organisms they host: that the creatures' phenotypes are extremely well fit for them is a tautology, because, if not, we would not identify that environment as their niche! According to adaptationism, if the fit between niche and organism is not 'perfect', "all that follows is that that isn't exactly the niche that the species occupies" (Fodor & Piattelli-Palmarini 2010: 147). The authors suggest that one thing is how a creature 'makes a living' now, and a different thing is how it came to make a living in that way: natural selection cannot answer the first question, and the adaptive solution to the second one is just one possible answer, not 'the' answer (ibid., 151) -developing specific physiological (e.g. immune system) or behavioural phenotypes are, for example, alternative developmental explanations for the fit-. Adaptationism is fundamentally a theory about the preponderant importance of exogenous factors in the shaping of phenotypes. Gould christens this attitude as 'billiard ball' model, because the organism is a passive entity pushed by external forces, and he asserts that "environmentalism will have to be tempered with a strong dose of internalism to explain these major events [e.g. the Cambrian explosion]" (1977: 22). A growing number of biologists and philosophers of biology advocate for such a broader view of the processes and forces behind evolutionary phenomena. Many of the processes postulated in the following sections can be seen as such internalist proposals: they see the organisms as systems under the control of internal laws and constraints that define and limit their possible morphologies. Natural selection can assign functions to these morphologies and eliminate unfit ones, but has a limited role in creating and shaping them: it just 'fine-tunes' the organism (Fodor & Piattelli-Palmarini 2010: 86).

Challenging one or more of the adaptationist assumptions can lead (but not necessarily does) to a structuralist critique of Darwinism (Gould 2002: 159). Although almost wiped out by the functionalist wave, structuralism[103] has a tradition dating at least to the XVII century, when Rational Morphologists (Cuvier, Geoffroy St. Hilaire, and later Owen) strived to find the underlying, universal laws of form that could account for the similarities among apparently distant organisms. Linnean taxonomy was also meant to group living beings on the basis of similarities that could reveal such underlying laws:

[103] For a brief definition of functionalism and structuralism, see §3.4 and 4.1.4.

> […] the intellectual aims of a systematic taxonomy were not merely to categorize organisms into nested clusters, but to discover the natural orderings and groupings among all living things. If similarities in morphology were to reveal what the proper laws of form might be, then proper clustering of organisms into a system was critical to seeing the proper similarities, hence finding the laws. (Kauffman 1993: 4)

Organisms were not considered just as the accumulation of a series of 'lucky historical accidents', but the result of a small group of common mechanisms that explain homologies within and among species (Kauffman 1993: 4-5). In the beginning of the XX century, in a book considered the result of the eccentricity of a brilliant mind, D'Arcy Thompson (see §4.5.1) claimed that physical laws, and not heredity and selection, mould the forms of life. Sixty years later, in what has been defined as "the most influential structuralist manifesto" (Wagner 2014: 7), Gould and Lewontin (1979) sanctioned the same idea and introduced the concept of exaptation (see §4.3.1), or phenotypic structures that contributes to the feasibility of an organism without being built by selection for their current role. Through the famous metaphor of spandrels, they suggest that function can be parasitic on form, and that form can be just a by-product of pre-existing structures. Evo-Devo research (see §4.4.1) similarly focuses on how developmental constraints configure phenotypes and their potential evolution: Wagner (2014) claims that evolutionary novelties emerge from such constraints and not from the cumulative selection of random variational changes. Selection adapts form to some function only after form is already available. Likewise, novelties are not a consequence of selection but of the structural features of complex systems, able to self-organize, according to Kauffman (see §4.5.3); or of physical and thermodynamic laws (see §4.5.1 and 4.5.2). Cherniak (2009) applies network engineering considerations to brain structure, deduces that the optimization of neural networks (minimization of connection costs) comes "for free, directly from physical processes", without DNA involvement, and claims that "the neural optimization paradigm is a structuralist position". Form anticipates function also at the genomic level, according to Wagner (see §4.2.5), whose investigations reveal how the underlying structures of genetic networks allow the exploration of novel functions and their subsequent judgement by the court of selection.

As mentioned, not all critiques to adaptationist accounts are structuralist. Neo-Lamarckian (see §4.3.3) and epigenetics (see §4.3.4) evolutionary accounts suggest that novelties are somehow externally induced by the environment: although challenging the idea of random and blind variations, they still place the source of change outside of the

organism, and still consider function as driving form. The same can be said of Wright's theory (§4.2.2) or of the so-called Baldwin effect (§4.3.2).

4.1.5. The source of innovation as the basis of our classification of non-adaptationist evolutionary accounts

Adaptationism claims that selection is a sufficient *explanans* for all three *explananda* of life: adaptation, variability and complexity (Brown 1991). The accounts of this chapter do not question the first claim: natural selection explains why organisms are so well-adapted to their niches. In this sense, they are not *anti*-adaptationist: all of them explain drawing on factors other than fitness differences or optimization., but they are generally not in contradiction with neo-Darwinism and empirical adaptationism as long as these are not considered to be exhaustive theories of evolution: rather, they aim at complementing them by trying to identify where the variability comes from, upon which selection acts. Brooks and Wiley, for example, claim that their entropic theory "extends rather than conflicts with neo-Darwinism" (1986: 201) and that "the empirical core of neo-Darwinism, namely population biology, can be accommodated within our theory" (1986: 203).

Non-adaptationist proposals suggest that other processes might act before, together with, or at a different hierarchical level than selection. Some are compatible with the idea of natural selection as the main directional agent of change, but propose sources of variation others than random mutations, thus suggesting that variations already come with some intrinsic degree of complexity. As an example of such a process acting before selection, Kauffman (1991,1993, 2000, see §4.5.3) identifies in the process of self-organization of complex systems the source of order in organisms (including phenotypic novelties), and relegates selection to the role of picking up the most evolvable among such spontaneously generated systems. Likewise, Wiley and Brooks (1986: x, see §4.5.2) accept the theory of evolution by natural selection, but as an auxiliary hypothesis and, as such, they claim that it just sorts a subset from all possible phenotypic outcomes generated by another, more fundamental process: evolution as an entropic process. Others authors advocate the action of non-adaptationist mechanisms not only in the generation of variations, but also in their fixation in the population. The shifting-balance mechanism (Wright 1931, 1932, see §4.2.2) acts together with selection, allowing to overcome some of the genetic constraints limiting its optimization power. Some authors limit the main role of natural selection to the phenotypic level of adaptive evolution: Kimura (1983: xi) proposes his neutral theory as an

alternative to Darwinism at the different hierarchical level of molecular evolution, without questioning its supremacy at the phenotypic level.

In conclusion, they do not pretend to demolish the adaptationist building by eliminating all four pillars: they just reinforce some of the pillars and add new ones, to make the building more stable. As Gregory claims: "[…] an important point to bear in mind about complex organs is this: not everything about them should be viewed in the light of adaptation." (2008a: 366).

The classification of non-adaptationist accounts applied in the rest of this chapter is based on where variations and novelties are supposed to come from. It is an heuristic choice and it lacks any presumption of exhaustiveness, or of ontological foundation. It is based on the identification of the main logical blocks that constitute organisms: their genotype, their phenotype, and the developmental system creating the latter from the former. All components of these blocks are connected in inextricable and complicated ways, thus configuring a natural system interacting with the environment (as a sum of ecological niche and physical laws). Fig. 4.4 shows these elements of the organism and the evolutionary mechanisms assumed to be working under the adaptationist view.

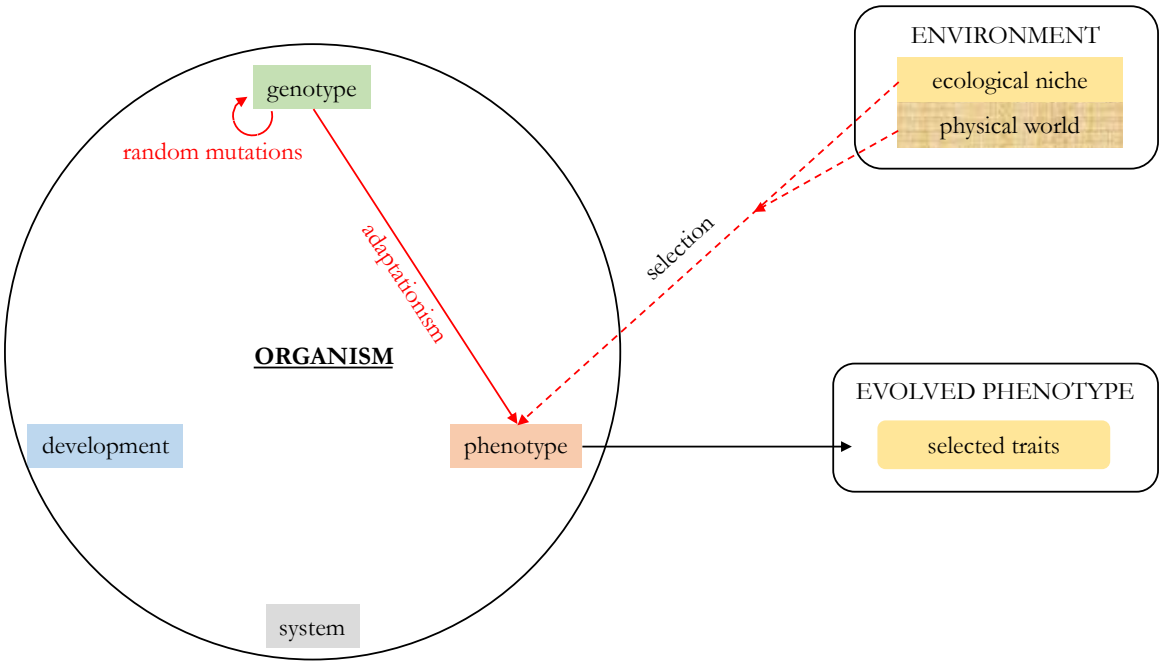


Fig. 4.4 – Organism as a complex system constituted by three logical blocks: genotype, developmental system and phenotype. According to adaptationism, genotype creates the phenotype. Developmental changes, which are not genetically inheritable, are ignored. Colours in the 'Evolved Phenotype' box show the final source of novelty that, according to adaptationism, is always the environment (Compare with Fig. 4.41, that includes accounts from the pluralistic view).

Fig. 4.41 at the end of the chapter shows the same representation integrated within pluralistic accounts. Some accounts consider the genotype as the primary source of phenotypic variability: they loosen some of the hypotheses of the modern synthesis (e.g. big populations, random mutations, etc.) and introduce the mediation of mechanisms other than random mutations. The second group of accounts abandon the gene-centred view and look for variability at the phenotypic level, where selection really takes place. Developmental systems are also considered as a potential source of variation, fostering quick evolutionary changes thanks to their plasticity. The last group of accounts focuses on constraints and the evolutionary consequences of physical, chemical and thermodynamic laws as sources of variation, as well as on some putative universal principles governing complex systems, suggesting that the living world is no exception to their applicability.

4.2. Genetic accounts

4.2.1. Classical population genetics' processes: mutation, recombination, and drift

We consider population genetic as non-adaptionist because it distances itself from adaptationism in at least two respects: it does not claim that fitness is optimised under selection, and it does not grant a special role to selection compared to other forces like mutation or drift, considered by adaptationism, at best, as evolutionary noise. Additionally, recent population genetics theoretical analyses seem to suggest a non-adaptive origin of genome architecture, at least in multicellular eukaryotes (Sarkar 2015): the 'baroque' structure (size, diversity, etc.) of most eukaryotic genomes persists regardless of selection acting against it.

Koonin claims that the point mutations (that represent the Darwinian concept of infinitesimal changes) are not the primary mechanism of new variations in populations: "gene and whole genome duplications, large deletions including loss of genes or groups of genes, horizontal transfer of genes and entire genomic regions, various types of genome rearrangements, and interaction between genomes of cellular life forms and diverse selfish genetic elements" have a far more important impact on genome evolution (Koonin 2009: 1027).

Lynch (2007) especially underlines the idea that natural selection is just one in the foursome of evolutionary forces that also includes mutations, recombination and genetic

drift. Many genetic and developmental aspects (e.g. genomic architecture, gene structure, developmental pathways,…) cannot be accounted for without recurring to these non-adaptationist forces that, even if stochastic in nature, do have impact on the *direction* of evolution[104]. Mutation is not just the ultimate source of variation upon which natural selection acts on the basis of alleles' phenotypic effects, but it represents a weak selective force "by differentially eliminating alleles with structural features that magnify mutational target sizes" (Lynch 2007: 8598): complex traits are coded by several *loci*, so, the more complex a trait, the bigger is the size of the set of genetic information that codes it, and the higher the probability that some mutation affects part of it. Given that most mutation are deleterious, alleles that foster such complex genic architectures are more prone to be eliminated. Traits are influenced by many loci, and hence affected by mutational damage in a relatively large proportion of the genome. Recombination generates new assortments of genes within and among chromosomes. Genetic drift does not just introduce noise in the evolutionary path led by natural selection, but influences its direction by "increasing the likelihood of fixation of deleterious mutations and decreasing that of beneficial mutations" (Lynch 2007: 8598). Higher-level features like evolvability and robustness might just be by-products of such forces and not the direct result of selection, among other reasons because there is no evidence that they are always advantageous. On the other hand, Lynch denies that there is any other mechanism of evolution: "These four broad classes encompass all of the fundamental forces of evolution" (Lynch 2007: 8597). Evolutionary change is just a change in genotype frequencies, and "population genetics provides an essential framework for understanding how evolution occurs". Other proposals, if not well-founded on work done by evolutionary biologists, are either wrong or obscure (e.g. Kauffman's account for complexity, see §4.5.3). Lynch is critical against adaptationism in claiming that evolution is not just natural selection; but, against the spirit of the extended synthesis, believes that micro-evolutionary theory (including all four mentioned forces) is sufficient to explain all the complexity of the biological world and that no novel mechanism of evolution has been revealed by accounts like e.g. Evo-Devo (§4.4.1).

[104] Lynch makes the interesting claim that a stochastic phenomenon can have non-random consequences: "It is tempting to think that stochastic processes have no implications for the direction of evolution. However, the effects of mutation and recombination [*stochastic phenomena*] are nonrandom, and by magnifying the role of chance, genetic drift indirectly imposes directionality on evolution by encouraging the fixation of mildly deleterious mutations and discouraging the promotion of beneficial mutations." (Lynch 2007: 8598, italics mine)

4.2.2. Wright's Shifting Balance theory and the metaphor of adaptive landscape

When Wright formulated his Shifting-Balance Theory, the problem he tried to solve was the apparent confinement of a population in a small portion of all possible gene combinations due to natural selection. New re-combinations can oppose deleterious mutations and at the same time introduce advantageous novelties, but evolution cannot access such distant areas of the genotype space under the action of selection alone. Wright proposes therefore a 'trial-and-error' method, based on casual loss and fixation of alleles pushed by drift (Pigliucci 2008b). The process of drift is thus behind Wright's well known Shifting-Balance Theory and his successful metaphor of fitness landscapes.

Genetic drift refers to changes in alleles' frequencies in a population linked to random events (e.g. accidental deaths) instead of selection of fitter phenotypes. As a consequence of chance, some alleles can disappear from the population while others can get to fixation. McShea and Brandon (2010: 90-92) define the principle of drift in terms of sampling error: it is any outcome that deviates from the expected one. They also claim that any population at equilibrium/on an evolutionary trajectory will tend to drift from it unless acted on by some evolutionary force.

According to Gould (Gould 2002: 522-524), Wright's idea about the role of drift changed along the years, moving from being considered as a non-adaptationist agent of change (strong negative view of selection) to becoming a source of 'raw material' upon which natural selection acts (weak negative view):

> Wright therefore (and accurately) depicted his later shifting balance theory as adaptationist, and as invoking drift only for a source of variation among demes. (Gould 2002: 523).

In both cases, however, Wright's theory proposes a mechanism complementary to selection, as illustrated in the following paragraphs.

Phase I: drift alternative to selection. In its initial formulation of the theory (Wright 1931, 1932), Wright talks about multi-dimensional surfaces mapping each possible gene combinations to its selective value (e.g. its fitness) (Fig. 4.5). Contrary to Fisher's hypothesis[105], he considers

[105] Fisher based his Fundamental Theorem on several assumptions, among which, that the relationship between gene and character is one-to-one, or additive (see §2.2.3).

gene's effects as non-additive[106]: evolution is selection of systems of genes, not of single genes. The mapping corresponding to such 'epistatic' picture has many peaks, as there are many possible combinations of gene with the same selective value[107]. On the other hand, populations under some kind of selective pressure are typically confined to an infinitesimal part of all possible gene combinations, and thus accumulate in a small area around a peak. This means that, once pushed on a local peak by selection, a population will stick there unless some unusual phenomena take place. Wright identifies several such phenomena. Increased mutation, combined with reduced selection, can expand the extension of the population on the peak and decrease average fitness. The landscape shape can be modified by environmental changes, so that the peak disappears and selection can move the population to a new peak. Inbreeding in small population can lead to extinction, while in medium-sized populations it usually moves the population away from the optimum without reaching a new one. All these mechanisms cannot but relocate the population to new peaks that are close to the original one. Real innovation within a population can only occur when a different and more radical phenomenon occurs: new traits arise, in relatively short evolutionary time spans, from genetic drift acting on populations of small dimension and geographically isolated, through the casual fixation and loss of alleles. Given that the relationship between alleles and phenotypic traits is complex and non-additive, novel phenotypes can arise even without the contribution of mutations. What differentiates the new population from the original one, thus, is typically a non-adaptive trait: and the fact that systematics often uses such traits to classify subspecies (e.g. Haldane 1932:113-114) is considered as evidence that speciation is not always adaptive. The mechanism of 'mutation plus selection' cannot explain the most significant phenotypic novelties because it needs too long timescales, because mutations are usually lethal and because it tends to destroy variety. The polymorphism exhibited by many species is the final equilibrium of a selective process rather than the source for future adaptations. Wright puts special emphasis on the size of

[106] Genes' effects are additive if each gene contributes to the phenotype independently. If genes interfere with each other, their effect is non-additive: we have dominance (if the interaction is between genes ate the same locus) or epistasis (if the interaction is between genes at different loci). See also note 16.

[107] Wright's intuition is confirmed by Kauffman`s N-K model (1993), that shows how the shape of the fitness landscape depends upon the number N of gene loci affecting fitness and upon the average number K of other loci affecting each locus (a measure of epistasis). If there is no epistasis (all loci provide a purely additive contribution to fitness, K=0), there is just one peak. In case of limited epistasis (K<=2), multiple peaks appear, but close to each other. If K tends to (N-1) (all loci affect all other loci), there are many small, dispersed peaks.

evolving populations (Wright 1931). If the population is too small, all genes are fixed but there is no equilibrium: the population moves toward extinction, as frequencies drift slowly and selection cannot counteract this slow, random drift. If the population is too big, genes are fixed at equilibrium and evolutionary change, fuelled by mutations or environmental pressure, is slow and reversible: variability is lost and further evolution is improbable. Populations of an intermediate size show no fixation of genes, whose frequencies drift randomly around equilibrium even in the absence of environmental changes or selection: they can go through continuous, irreversible evolution which is random (non-adaptive) in the short run.

Drift is, under this view, a non-adaptationist agent of change, as Provine (1986) claims: "[…] Wright believed non-adaptive random drift was a primary mechanism in the origin of races, subspecies, species and perhaps genera". Kauffman (1993: 10) affirms that Wright's vision of drift is the most successful attack to the neo-Darwinian belief that all evolutionary change is pushed by selection.

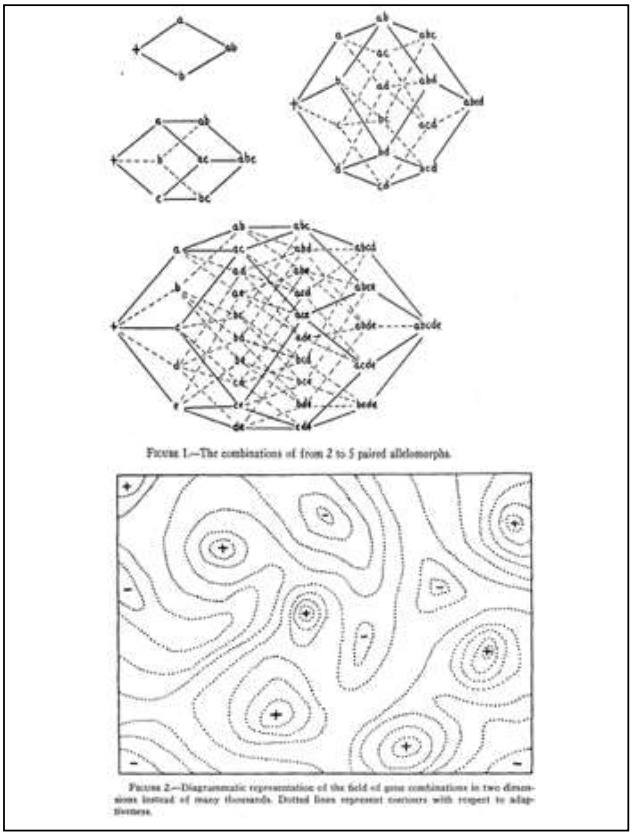


Fig. 4.5 - Mapping of all possible combination of alleles (above) and altimetric landscape of selective value (below) (Wright 1932)

Phase II: drift as step previous to selection. In later versions of the theory, Wright (1982) moves to a different idea of drift, whose role is just that of supplying additional variation for the

action of natural selection, the only relevant evolutionary force. Drift ceases to be a non-adaptive force and becomes thus compatible with empirical adaptationism. To illustrate his theory, Wright introduces the metaphor of the adaptive landscape, a function that maps the fitness of phenotypes into a continuous surface having peaks and valleys (Fig. 4.7). Given that, according to Fisher's fundamental theorem (Fisher 1930), average fitness never decreases when natural selection is the only acting agent of change, populations tend to move, generation after generation, upward along the function, till they reach a peak, which represents a local fitness maximum[108]. The population ends up stuck in what is probably not an absolute fitness maximum and, if natural selection is the only force, it cannot reach any higher peak, from which it is separated by an unbridgeable fitness valley. Drift can overcome this problem and help the population cross the valley. Drift is the random sorting of some individuals from a group (e.g. because the group was suddenly separated from the rest of the population). The resulting sub-group might have a lower average fitness than the original population, e.g. if the sorted individuals are not among fittest ones (in a population, not all individuals are equally fit, as the environment usually allows for a range of phenotypes around the optimal one). The sub-group, or deme, might thus 'jump' into a point of the fitness landscape far away from the original peak, and probably located in a valley. Under selective pressure, the deme, in the absence of genetic, functional, and demographic constraints, will then evolve in the direction of the steepest slope in case of asymmetric landscapes, and could thus climb to a new, higher peak (Lande 1976).

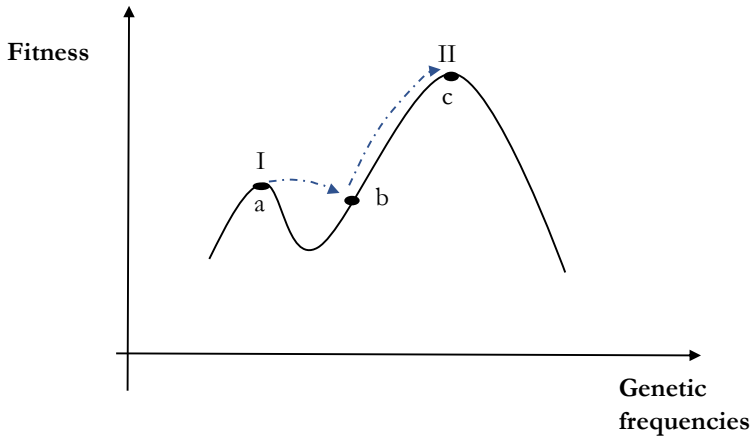


Fig. 4.6 – The 'shifting balance' process. The original population a lies at fitness peak I, from where it cannot be displaced by selection only. A random change in genetic frequencies (drift) might push the population away from the peak to point b, form where selection move it till the new, higher peak II.

[108] Fisher (1941) did not agree with the metaphor of population climbing the adaptive landscape based on his theorem, as he claimed that it mixed individual with group advantage (see Okasha 2018: 80).

This is a random process of 'trial-and-error' that Wright names 'shifting-balance' (Fig. 4.6). Under this interpretation, drift just supplies raw material for the action of natural selection, which in turn configures landscapes and drives populations' movements along them. Drift pushes a deme away from the peak into a valley, and selection afterwards decides its fate: it either takes it back to the same peak, makes it extinct, or pushes it to a new peak The advantage of drift is that it is quicker than mutations[109], and can cause farther jumps (big mutations are usually lethal). It is thanks to the cooperation of selection and drift within the shifting-balance mechanism that macro-evolution takes place: selection generates reproductive isolation by eliminating unfit hybrids, while drift establishes novel characters without recurring to improbable mutations.

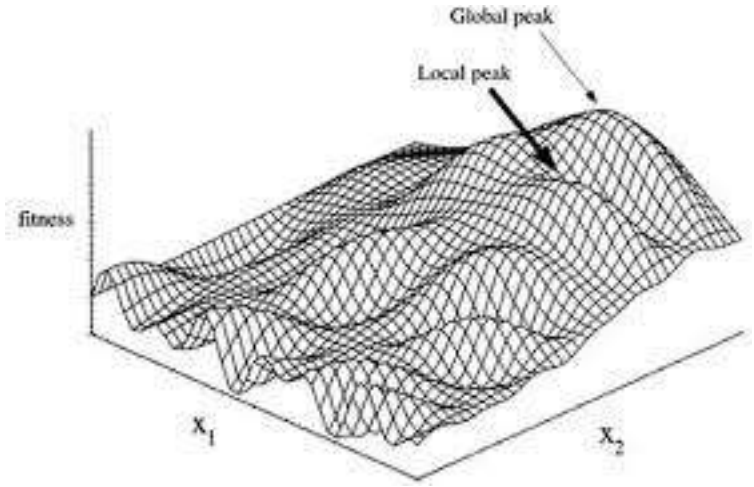


Fig. 4.7 – Fitness landscape

Wright's landscape (in the literature, the term *fitness* landscape is used when considering individual values[110], and the term *adaptive* landscape when considering average values) has three different versions (Gilchrist and Kingsolver 2001):

- Individual fitness of multi-locus genotypes on the z axis, and genotypes on x/y axis. It is inapplicable in real cases, as it is impossible to map the complexity of such genotypes on a pair of axis (Wrigth 1932).

[109] Small mutations move the phenotype away from the peak, and selection eliminate them. Big mutations could allow a jump to another peak or to a valley, but are rare. Drift can allow jumps of demes thanks to new recombination that mutation plus selection preclude.

[110] Other terminologies are also used in the literatura, e.g. "'individual fitness surface' for the plot of individual fitness against individual genotype or phenotype." (Okasha 2018: 75).

- Average fitness of genes frequencies on z axis, and genes frequencies on x/y axis. Fitness is averaged because frequencies are a property of populations, not of individuals (Wright 1935).
- Individual or average fitness on z axis, and phenotypic values on x/y axes (Simpson 1944, Lande 1976). Only this version, when considering averages, can be used with models of evolutionary dynamics.

The adaptive landscape metaphor has been very successful, but it has also been the target of heavy critiques. Plutynski (2008) summarises the main ones:

- It is unclear what the unit of measure of the independent axis of the landscape represents.
- The environment does not remain fixed once and for all, so the landscape will change as well.
- Frequency-dependency can produce situations not reproducible in any landscape, because the independent variables are not independent anymore.
- Under a wide variety of circumstances, selection will not drive populations up a hill of the adaptive landscape, but down a valley.
- The fitness of the whole population is not a direct function of the frequencies of genes in the population.

Kaplan (2008) claims that adaptive landscapes are void of explanatory content and hide a circular description: the population moves along the landscape, which in turn is a representation of how the population moves. Moreover, Kaplan criticises the assumption that there is a fixed and known genotype-fitness mapping, where this is generally unknown. Finally, in multidimensional landscapes, the topography changes radically and, above a 'percolation threshold', the vast majority of peaks are connected through a one-step move (e.g. simple mutation) without passing through a valley. Calcott (2008) rejects this last critique, as he believes that the vanishing of valleys is not linked to multi-dimensionality, but to the hypothesis that close genotypes have similar fitness.

Wright's idea of drift clearly stands out, at least in part, as a non-adaptive model, possibly the first one within the paradigm of the modern synthesis, in the sense that it does not resort to fitness in order to explain phenotypic novelties. Even in the 'weak' version of his negative view, Wright introduces two important elements behind the appearance of new phenotypes:

- The *current genetic pool* of a population can potentially express *much more phenotypic variety than the actual one*. Selection keeps the population within a very limited portion of the whole possible genetic space. Mutations are not an efficient way to escape this limitation because they have either relevant phenotypic effects -and these effects are probably deleterious or even lethal-, or very small phenotypic effect -and thus moves the population only slightly, just insuring slow and gradual optimization to local conditions-. A lot of potential phenotypic variety (linked to combinations of *existing genes* having complex, non-additive effects) remains unexpressed, because it is eliminated by selection when competing with fitter phenotypes in the peak. The 'shifting-balance' mechanism, pushing the population away from the peak and, without need of mutations, foster distant jumps in the genetic space, in areas where the potential phenotypic variety can appear, and be submitted to a 'trial-and-error' strategy able to open paths to new peaks and phenotypic novelties.
- *Selection among individual* of the same population can *only lead to micro-evolution*, a slow, self-terminating process of fine tuning of genes to the local environment. Macro-evolution is based on selection among groups, triggered by the opening of a new niche (e.g. a new territory, a catastrophe modifying the current niche, a new sub-niche created by the great efficiency of the current specie) and fuelled by the quick, continuous, not self-terminating process of 'shifting-balance'.

Wright's shifting-balance mechanism anticipated in some way the Punctuated Equilibrium theory by Gould and Eldredge, introduced in the next section, which appeals to the similar idea that phenotypic changes are rapid and limited to marginal groups, and that, if successful, they later spread to the rest of the population. This pattern of allopatric speciation, followed by geographical expansion, is proposed to explain the gaps in the fossil record.

4.2.3. Punctuated Equilibrium

Darwinism (as an extrapolation unjustified by Darwin's ideas themselves) claims that selection slowly converts *variation within* population into *differences among* populations (Gould 2002: 748). Speciation is thus just adaptation on large timescales, and macro-evolution just the sum of many micro-evolutionary events: both are individual-centred phenomena. The logical evidence of such gradualism should be found in the fossil register, with slightly

different fossils to be found in successive bedding planes: but this is not the case. Gould and Eldredge's starting point is this widely recognised and at the same time ignored fact: "The long-term stasis, following a geologically abrupt origin, of most fossil morphospecies" (Gould 2002: 750). Morphogenetic explosions during the history of life have taken place over (geologically) short periods, as evident from the Ediacara fossils (575 to 542 million years ago), the White Sea assemblage (560 to 550 mya), the Nama assemblage (550-542 mya) and the Cambrian explosion (545 mya), when in about 5 to 10 million years most of complex life forms appeared (Fodor & Piattelli-Palmarini 2010). Darwin was aware of the problem for his theory of the absence of intermediate varieties in the fossil record:

> Geology assuredly does not reveal any such finely graduated organic chain; and this, perhaps, is the most obvious and serious objection which can be urged against my theory. (Darwin 1872: 351)

To discuss this objection, he dedicates a whole chapter of the *Origin*, the 10th, whose title is revealing: 'On the Imperfection of Geological Record'. The graduated organic chain left few evidences because fossilization is a rare phenomenon:

> The explanation lies, as I believe, in the extreme imperfection of the geological record. (Darwin 1872: 351)

However, practical palaeontology does not recur to theoretical 'stages of evolution' to date rocks and determine their stratigraphic sequence: it uses fossil species actually discovered instead, as fixed entity within a given stratigraphic range. If the fossil record were so incomplete, such approximate method would bear catastrophic results. The morphological stability of fossils is especially evident in traits used to identify species-level taxa (Eldredge and Gould 1977). Long periods of stasis are predominant and well documented in the literature. For example, in the horse lineage, transition from *Mesohippus* to *Miohippus* shows stasis and rapid branching, overlap of genera, well-defined and static species over millions of years, against the idea of gradual phyletic transformation and closely related species (Prothero & Shubin 1989). Gould (2002: 752-755) lists several other examples. Available data point at stasis of species: it is thus at least odd that Darwinism defends gradualism on the basis of supposed lost evidences and ignores the abundant evidence we do have against it.

The theory of Punctuated Equilibrium gives a new interpretation of this deficiency of fossil record that shows 'instantaneous' origination and stability of 'paleontological morphospecies'. (Gould and Eldregde 1993). It is not, thus, anti-selectionist *tout court*: it proposes a pattern of speciation alternative to phyletic gradualism (Eldredge & Gould 1972,

Gould and Eldregde 1977), and suggests that micro-evolutionary mechanisms automatically do not translate into macro-evolutionary events like speciation. The theory asserts that "the great majority of species […] originate in geological moments (punctuations) and then persist in stasis throughout their durations" (Gould 2002: 766).

Gould identifies and defines operationally three basic concepts of the theory (Gould 2002: 67-774):

- *Stasis*. It is not a claim about the absence of any change in the morphology of a species through time; rather, it is the denial that changes accumulate over time, in the sense that final samples do not differ from the initial ones, or at least that their statistical fluctuations do not lie outside of the range observed during the life of the species. Stasis is the normal status of species, while speciation is a (relatively) 'quick' punctual event, not the result of phyletic gradualism (Lieberman & Eldredge 2014). 'Living fossils', under this view, do not reveal optimal adaptation, but just very low speciation rate, that is, the lack of mechanisms triggering the geographical and genetic isolation of a sub-population. Some authors interpret long evolutionary stasis as the result of strong stabilising selection (Charlesworth et al. 1982). Stasis in a population can be due to an optimal phenotype or to the impossibility to get to the optimal phenotype due to some genetic constraints linked to the current genotype. Genetic networks allow to move away from the current genotype and explore the genotypic space through neutral mutations, until constraints are released and the optimum reached (Wagner 2011, Ch. 12).
- *Punctuation*. It is defined, with a somehow arbitrary analogy with human gestation[111], as a period of about 1-2% of the duration of the subsequent stasis. The average duration of a species being 4 million years, punctuation covers around 40.000 years. Given that such a timespan allows for micro-evolutionary gradualist phenomena, the theory cannot be accused of 'saltationism', or being anti-Darwinian (Gould & Eldredge 1993). The authors make no special claim about the mechanism behind speciation: it can be adaptive or include non-adaptive factors like drift, neutral evolution, etc. (Lieberman & Eldredge 2014). Russkin-Gutman and Esteve-Altava (2008) suggest several mechanisms behind punctuated equilibria, e.g. developmental canalization, genomic drift, the relationship between body and population size.

[111] The 1% rule applies in most mammals (from cats and dogs, to pigs, to whales), with relevant exception in cows (3,5%) and Giraffes (6%).

- *Dominant relative frequency* of punctuated equilibrium vs. phyletic gradualism. Gould's claim is not limited to assert the fact of punctuated equilibrium phenomena; but it defends its dominant role in macro-evolution.

The fossil record can thus be interpreted without formulating any ad-hoc hypothesis about its incompleteness. A species' morphology remains fixed along its whole life, thus creating a homogeneous bedding plane. From this centrally located population, a daughter sub-population, peripherally isolated, forms a new species. Old and new species coexist for at least some time (Lieberman & Eldredge 2014). The new species, if successful, later invades the territory of the original ancestral population, creating a new homogeneous bedding plane of different fossils. The lack of intermediate forms is thus accounted for, as the new species did not originated *in situ*. Ancestral species occupies the central territory, while new species appear in peripherally isolated areas.

The basic differences between the neo-Darwinian view and the Punctuated Equilibrium theories concern the concept of species, the tempo and mechanism of their birth (Gould 2002: 776):

- *Concept*. Contrary to the gradualist view of speciation by anagenesis[112], that entails the problem of clearly identifying species out of a continuum, under the Punctuated Equilibrium hypothesis species are well-defined Darwinian individuals that appear during a (geological) short timespan by cladogenesis. Although the issue of defining species and their ontological status might be considered immaterial, the ability of the Punctuated Equilibrium theory to (dis)solve the species-problem in palaeontology shows that the theory solves more problems than the ones it was initially conceived for, it makes new predictions and is more falsifiable and thus more scientific (Popper 2002).
- *Tempo*. Species appear abruptly (in geological terms) and rarely, and do not change significantly during their whole existence vs. species transform into new ones through slow, gradual and continuous modifications (phyletic gradualism).

[112] "During anagenesis […] a single species is transformed into a different species over the course of many generations. […] cladogenesis […] involves the division of a species into two or more species." (Brooker 1999: 738)

- *Mechanism*. Species rise through speciation (splitting, e.g. through geographical isolation of a daughter population, followed by genetic differentiation) vs. through anagenesis (transformation of the whole population).

Adding this 'species-view' level to evolutionary analysis decouples macro- from micro-evolution (Gould and Eldregde 1993): not all evolution follows phyletic gradualism. It might be so at the level of micro-evolution (within species), but macro-evolution (among species) involves a higher-level sorting of stable species, which can or not involve 'species selection' (Lieberman & Eldredge 2014). There are thus two different research programs in evolutionary studies (Gould and Eldregde 1993):

- Micro, whose main question is: how did natural selection within a lineage lead to adaptation over time?
- Macro, whose main question is: what species did better or what bias prevailed among species within a clade?

Speciation by branching is acknowledged by Darwinian gradualism, but interpreted either as (Gould 2002: 777) just an 'engine' for spreading a favourable trait into more than one taxon, a real but rare phenomenon; or as the simple sum of two anagenetic processes proceeding at a very slow tempo. Other authors accepts the theory but relegate its validity to special cases: punctuated equilibrium would happen in benthic environments, while gradualism in pelagic environments (Johnson 1975, 1982); or, respectively, in unstable and stable environments (Sheldon 1990).

4.2.4. Kimura's neutral theory of molecular evolution

Kimura neutral theory of evolution (1977, 1983) claims that, at the molecular level, drift, and not selection, is the fundamental process in evolution. The theory claims that most DNA changes during evolution are invisible at the phenotypic level (where natural selection acts) and are thus not adaptive (and neither deleterious). Variability and polymorphism at the molecular level are therefore neutral with regard to the survival and reproductive success of the organism, and are explained by mutation and drift and not by positive selection[113].

[113] Positive selection is the selective pressure that leads to the fixation of advantageous variants in a population. In negative selection, deleterious variants are removed from a population. See also note 95 in this chapter.

Kimura underlines that "the great preponderance of nucleotide changes over time and of nucleotide variability in populations at any time are selectively neutral" (Kimura 1983: xiii). This neutrality does not mean that the involved genes are functionless, but only that the mutant version is selectively equivalent, in a particular environment, to the original one (Kimura 1983:50).

Molecular changes should be expected to be very rapid when neutral, that is, free from selective pressure: and this is exactly what happens. The rate of mutant substitution at molecular level is constant per year and independent from living conditions or population size: for example, haemoglobins and other molecules of 'living fossils' (e.g. the Port Jackson shark) have undergone as many DNA base substitutions as corresponding genes in more rapidly evolving species (e.g. humans) (Kimura 1983: 81). This does not mean that no selective 'tuning' happened in the genes of the latter: but a much larger number of neutral substitutions have superimposed on it (Kimura 1983: 89). So, regardless of the evolution rate of genes with adaptive phenotypic effect (low in 'living fossils', high in rapidly evolving species), the rate of substitution of neutral gene is the same. Gould (1993, 2002: 688) offers an interesting example of how neutral substitution is faster if the corresponding phenotypic effect is unaffected by selection: the major lens protein of *Spalax ehrenbergi*, a subterranean blind mole rat, evolves much faster than in other murine rodents with normal vision. Gould (1993) also underlines the use of the theory to understand if the eye still has some adaptive value: the higher substitution rate of the gene encoding the lens protein is still only 20% of the typical rate for pseudo-genes[114], thus pointing to some selective relevance.

Evidence for neutral substitution comes from the maximum tempo of molecular substitutions in DNA forms that do not affect the selective value of the organism (Gold 1993):

- Rates of change at the redundant third position of DNA amino acid coding sequences are five or more times more rapid than in the relevant first and second position.
- Introns change at much higher rate than exons.
- In pseudo-genes, rates of change are the same at all three positions.

The third position in functional genes is redundant, introns and pseudo-genes have no detectable phenotypic effect: in all these cases, high change rates correspond to invisibility to selection, one of the theory's hypotheses.

[114] Genes with no known phenotypic effect.

Kimura lists five features peculiar of molecular evolution (interpreted as mutant substitutions) that distinguishes it from phenotypic evolution:

1. The rate of molecular evolution for each protein (in terms of amino acid substitution) per year and site for various lines is approximately constant;
2. Molecules that are functionally less important evolve faster that more important ones (as in the example above of the major lens protein of *Spalax ehrenbergi*);
3. Less disruptive substitutions (with respect to form and function) are more frequent than more disruptive ones;
4. Gene duplication always precedes the emergence of a gene with a new function. While one of the two copies continues performing the original functions, the other one, liberated by stabilizing selection, can accumulate mutations that are neutral or slightly deleterious in the current environment, and experiment new functions that could become adaptive in a new one;
5. Elimination of selectively deleterious mutants and random fixation of selectively neutral mutants are much more frequent than selection of advantageous mutants. This means that a mutation is not neutral because 'almost adaptive, but not enough', but because 'not enough deleterious'. Evolutionary change (in Kimura's sense: any molecular change) happens as soon as the negative-selection barrier is lifted: an interpretation already captured in the idea that Darwin's theory does not talk about the survival of the fittest, but of the fit enough (Maturana & Varela 1980), of 'the non-survival of the not-enough fit' (den Boer 1999).

These features make Kimura's theory, which is not incompatible with Darwinism at the phenotypic level, quite incompatible with neo-Darwinism (Kimura 1983: 25). Kimura accepts that evolution at the phenotypic level (Darwin's level) is fuelled by natural selection (mainly stabilizing selection). Given that selection 'doesn't care' how a phenotype is genetically encoded (the 'phenotypic gambit' of Grafen 2014), he claims that evolution at the genetic level is driven by a different process: something that neo-Darwinism cannot accept, as

> A constant and random process, such as the accumulation of point mutations without selection, is not compatible with the traditional concept of directional evolution. (Rassking-Gutman & Esteve-Altava 2009)

Each phenotypic trait is encoded by several genes: this complex relationship between genotype and phenotype makes the accumulation of slightly deleterious or neutral mutations possible, given that each one of them has only a tiny fitness effect. This abundant variability

at the molecular level represents the potential raw material for future selection under new environmental conditions (Kimura 1983: xiii): if the environment changes, some of the initially neutral molecules can become adaptive. Moreover, by the accumulation of neutral mutations, the genome can hide cryptic changes and, if some thresholds are met, give rise to phenotypic novelties without intermediate steps (Kimura 1977). Evolution is a two-step process (Kimura 1983: 325-327):

- During long periods of time, environmental conditions are stable and species are under stabilizing selection. Random fixation of neutral mutant alleles occurs extensively, generating abundant raw material for future evolutionary changes;
- From time to time, an environmental change occurs and shifts the fitness landscape. Directional selection occurs, but the genetic variants upon which it acts are not caused by selection: they are already present thanks to the accumulation of neutral mutations during the previous phase. In a parallelism with phenotypic exaptations, Wagner talks about 'molecular exaptations' (Wagner 2011, Ch. 7) to refer to mutations that are selectively neutral not in absolute terms, but because the problem they can solve has not been yet identified.

This account is compatible with the 'Punctuated Equilibrium' thesis (Gould and Eldredge 1977).

The theory of 'constructive[115] neutral evolution' (Stolzfus 1999) suggests that repetition of neutral steps can result in the emergence of complex traits that *appear* to be adaptations, but are really 'aptations'[116]: they contribute to fitness, but they were not built through incremental selective steps. They did not appear because of their contribution to fitness, but they were retained because, following a change in a related trait, their loss would decrease fitness. The author considers duplicate genes, and claims that the classical neutral theory cannot account for the high percentage of retention of duplicates over long evolutionary periods. He identifies instead a mechanism to explain how the retention is possible based on the interactions among three phenomena:

---

[115] "The term 'constructive' is not meant in the vernacular sense of 'sympathetic' or 'positive' but is intended merely as a descriptive term, useful in conjunction with 'reductive' and 'conservative' to refer to increase, decrease, and lack of significant change (respectively) in the complexity of features that contribute operationally to organismal fitness". (Stolzfus 1999: 170)

[116] Meaning: 'fit for a role', independently from how they appeared: because of selection (ad-aptation) or because of its effect (ex-aptation) (Gould & Vrba 1982).

i. The existence of excess capacity in the genome in the form of redundant duplicate genes that perform the same functions;
ii. The action of selective constraints that prevent the loss of the duplicate gene when the activity of the original gene decreases (e.g. epistatic link of a selectively relevant site with a neutral evolving site);
iii. The bias in favour of activity-reducing variations vs. activity increasing ones.

A change in a gene that reduces its capacity (but is a step in the creation of a new function) can cause the maintenance of the (initially gratuitous) same capacity linked to a duplicate gene. The inclusion of activity-reducing mutations differentiates this model from the classical neutral account that considers only null mutations.

The separation between regulatory and coding function in eukaryotic genomes could be an example of this phenomenon: not a functional necessity, it separates two dimensions of genetic variation and thus allows the exploration of genetic spaces of each function independently, without affecting the other. Sub-functionalization is a similar process: a gene coding for different functions can duplicate, and each copy can silence one of the functions, thus allowing independent explorations of the space of functions independently (Altenberg 2005).

4.2.5. Andreas Wagner's Genetic Networks

Wagner believes that there is something special in the 'architecture of life' that makes it amenable to improvement through random change (2011:2). The source of innovation is not external to the organism (environment), and the mechanism of improvement is not natural selection. Evolution is more than raw material (random mutations) submitted to trial in the court of selection. He identifies the peculiarity of life's architecture in *genotypic networks* within which potential genotypes, although differing from one another for some elements (e.g. proteins that differ in one amino acid), share the same primary functions: thanks to the accumulation of cryptic mutations[117], the genotype can explore this network and acquire new functions without losing the original ones. This explains the sudden appearance of novel traits.

In his book *The Arrival of the Fittest* (2015), Wagner clearly states the focus of his research: "The real mystery of evolution is not selection, but the creation of new

[117] "Cryptic mutations are undetected [with no or very slight phenotypic effect] base changes in genetic DNA (or hereditary RNA)." (Lamb 1975: 305: mine note).

phenotypes". According to Wagner, the modern synthesis, very successful in accounting for the changes in allele frequencies in population, is unable to answer the question about innovative phenotypes. The reason behind its success is also the cause of its limitation: the oblivion of the phenotype, reduced to a fitness value, based on the assumption of a straightforward genotype to phenotype relationship. Two classes of empirical evidence support Wagner's view: the "complex beyond imagination" relationship between genotype and phenotype (p. 30) and the "enormous amount of genetic variation, everywhere, even in organisms that have not changed for many millennia" (p. 27), that points to the frequency and importance of silent genetic mutations. Selection cannot act on such variation and, above all, cannot generate it.

Innovation in life is widespread since its appearance on earth 3,8 billions years ago, and evolution has advanced at a high pace already a few hundred millions years afterwards, based on the emergence of new molecules and the reactions that create them. This restless process of trial-and-error producing phenotypic novelties seems to proceed in the backyard of life, as the incredible amount of cryptic genetic variation suggests. Wagner proposes a mechanism behind innovation that is based on genetic networks and genetic robustness.

First, he underlines that living organisms use around 5.000 different chemical reactions[118] (for example, sucrose and water react to produce glucose and fructose, see Fig. 4.8) to produce the 'building blocks' of life, or biomass molecules (nucleotides for DNA/RNA, amino acids for proteins, lipids, energy storage molecules, etc.), although each organism only catalyses some of them in its metabolism.

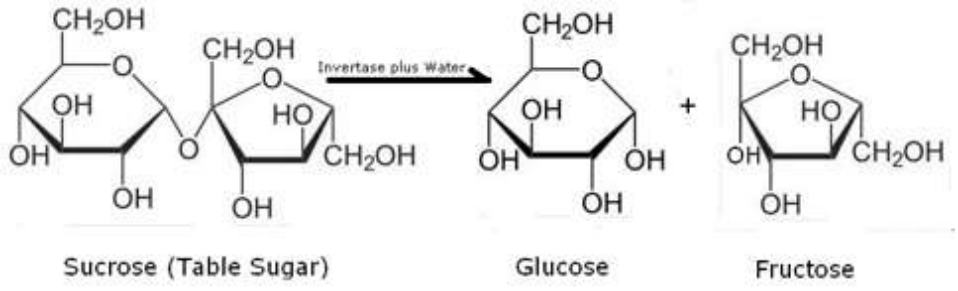


Fig. 4.8 - Metabolic reaction from sucrose and water to glucose and fructose (source https://www.decodedscience.org/wp-content/uploads/2014/03/Reaction-Sequence.jpg)

The bacterium *Sphingobium chlorophenolicum*, for example, extracts energy and carbon from a molecule (pentachlorophenol) that kills life anywhere else, and does so by mixing some new combinations of these reactions, without adding any new one:

[118] A complete list of reactions for thousands of organisms is available at http://biocyc.org/

> […] its genome encodes four enzyme-catalysed reactions that convert pentachlorophenol into molecules that are digestible as glucose […] The combination of these reactions is unique to *S. chlorophenolicum*, but the reactions themselves are not. (p. 72-73)

E. coli can evolve to survive on glucose through many alternative metabolisms, each one of these different only for some reactions. The evolved *E. coli* might share only 20% of their metabolic reaction with the original ancestor, and still survive on glucose (p.97).

How can such new combinations be explored? The classical population genetic answer, by sexual reproduction and spontaneous mutations, is of little effect on large, multicellular animals, whose individuals share 99,9% of the coding part of their genome and that produce a new generation only each few decades. Between mutations and selection, Wagner claims, there is an additional step, a logical structure that allows at the same time to maintain the original functions of the mutant gene and to explore new potential ones: genetic networks. This structure is possible because most biological problems have many alternative solutions (the 'neutral space' in Wagner's terminology), and some of them harbour the seeds of innovation to solve other problems.

Wagner (p. 82) considers the list of the 5.000 reactions: different organisms can produce some of them and not others, and this information is coded in their genotypes. Thanks to the known relationship between genes and enzymes, and between enzymes and chemical reactions, it is possible to translate genomic sequences into '*metabolic genotypes*', described through Boolean strings of 1s and 0s, each digit corresponding to one of the reactions. For example, the metabolic genotype of a genomic sequence unable to produce any reaction would be a string of 0s (Fig. 4.9).

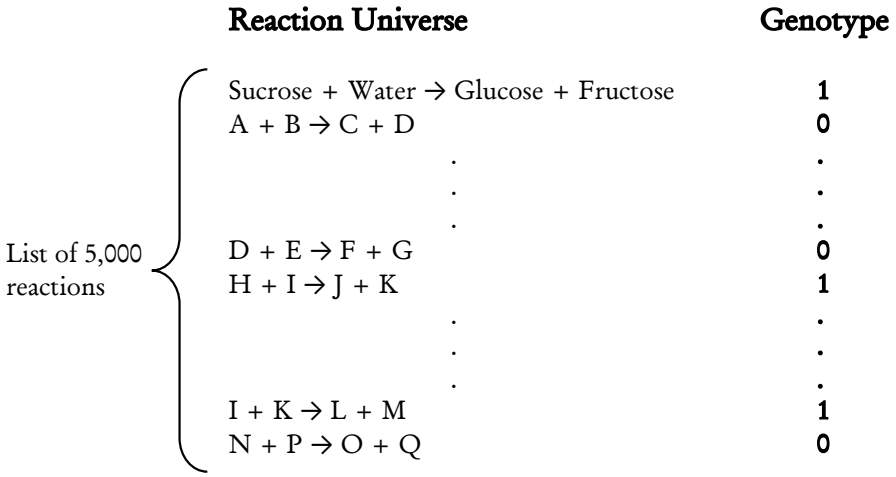


Fig. 4.9 – A metabolic genotype (from Wagner 2015: 70)

Changes in the metabolic genotype (gene deletion or gene transfer) add or eliminate some reactions from the available ones, switching 0s and 1s. Given any genotype, there are thousands of others that differ from it in just one digit: Wagner calls them 'neighbours' of the original one, and their set its 'neighbourhood' (Fig. 4.10).

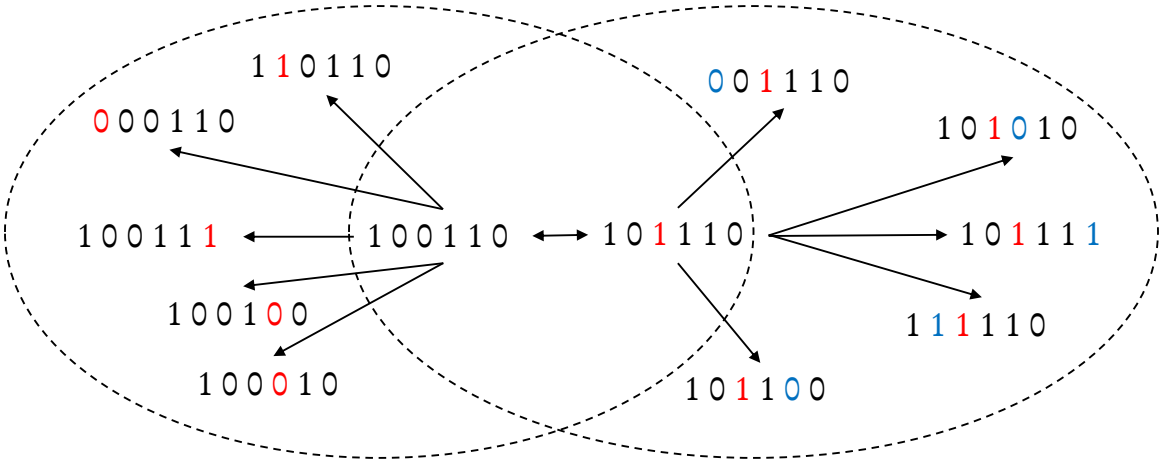


Fig. 4.10 – Neighbours and neighbourhood of metabolic genotypes. The set on the left includes all neighbours of the genotype '100110', each different for one digit. The set on the right is the neighbourhood of one of these neighbours, the genotype '101110'

It is possible to define the 'genotype distance' D between two organisms: if A catalyses a reactions and B catalyses b (b < a) reactions, the distance D = (a – b) / a. This distance is 1 if the two organisms do not share any reaction, 0 if they share all of them (Fig. 4.11).

Genotype 1 0110111100
Genotype 2 0100111100

Distance: $D = 1/6$

Fig. 4.11 – Metabolic genotype distance. The first organism can produce six reactions, the second only five: their distance D is (6-5)/6 = 1/6 (From Wagner 2015: 82)

Wagner then defines a '*metabolic phenotype*' as the list of all fuel molecules an organism can survive on, by synthesizing from them all biomass molecules. There are hundreds of such fuels (glucose, ethanol, sucrose, acetate, etc.), but not all organisms can survive on each: some can use many different ones, while others, more specialised, only very few. Similarly to the representation of the metabolic genotype, the metabolic phenotype is represented by a list of 1s and 0s, corresponding to a given list of fuels that the organism can or can't use (Fig. 4.12).

Fuel Molecules	Phenotype
Glucose	1
Ethanol	0
.	.
.	.
.	.
Sucrose	0
Fructose	1
.	.
.	.
.	.
Citrate	1
Acetate	0

Fig. 4.12 – Metabolic phenotype (from Wagner 2015: 86)

The set of all possible metabolic phenotypes is the '*metabolic library*'. By switching one of the digit of a metabolic phenotype to the opposite value, a new phenotype is created: for examples, by switching the digit corresponding to glucose from 1 to 0, we move from an organism able to survive (among other possible fuels) on glucose to one unable to do it. The concept of 'neighbour' and 'neighbourhood' introduced for genotypes are also applicable to phenotypes. The phenotype is encoded by the genotype, so these switches corresponds to changes in the metabolic genotype (gene deletion or gene transfer), caused by any evolutionary phenomena: "evolving organisms are like visitors to the metabolic library" (p. 92) and, after 3,8 billion years of life on earth, only a small fraction has been actually explored. This library shows a peculiar feature: all metabolisms solving the same problem (e.g. surviving on glucose) constitute a network whose neighbours members have metabolic genotypes that differs for just one digit from each other. To see this, consider an initial metabolic genotype G and its list of reactions. Let's assume that the corresponding metabolic phenotype P can survive on many fuels and, among them, on glucose. Let's analyse each of the thousands of neighbours G^n of G, each differing from G by just one digit (one reaction more or less) and their corresponding phenotypes P^n. Some phenotypes among the P^n (but not all) will still have the ability to survive on glucose: let's call them $P^{n'}$, and $G^{n'}$ their genotypes. Let's perform the same analysis on each of the neighbours of each of the $G^{n'}$, and retain the ones that can survive on glucose, $G^{n''}$. These new genotypes will have a higher genotype distance D from G than the $G^{n'}$. By reiterating the analysis till no modified genotypes gives a phenotype able to survive on glucose, the network is fully explored (Fig. 4.13).

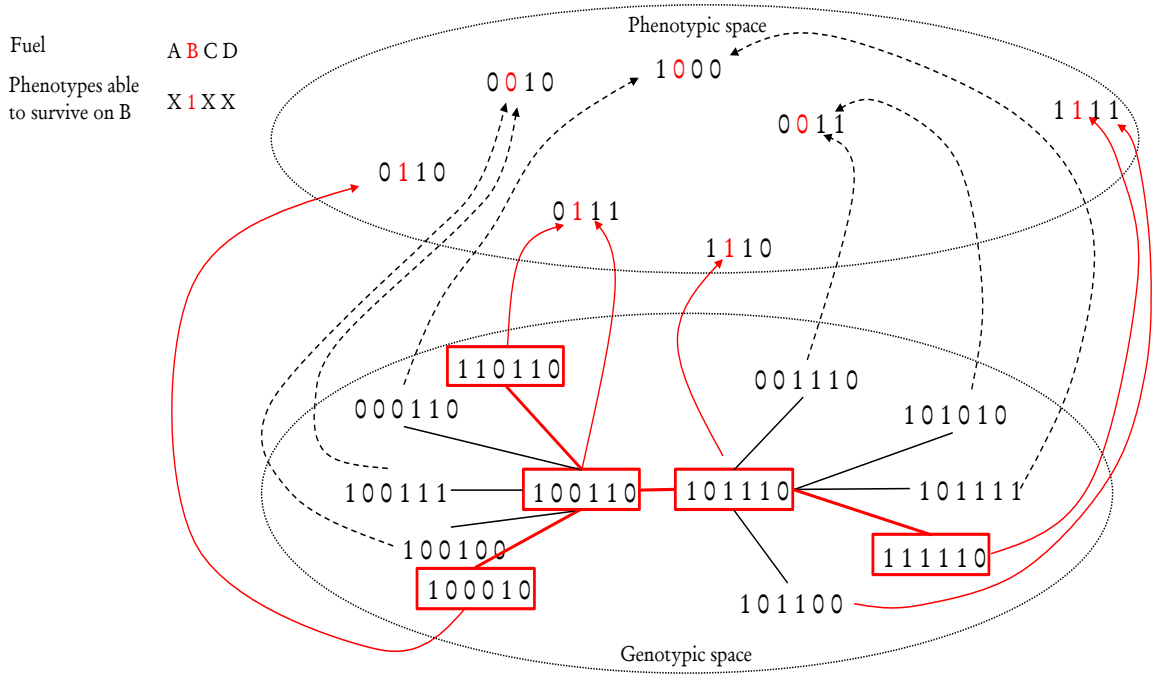


Fig. 4.13 – Genotype network for 'B' fuel. Lower set represents genotypic space, with neighbours genotypes linked by a segment. Upper set represent phenotypic space, in which the second digit shows the ability to survive on 'B' fuel. Arrows represent the genotype-phenotype relationship (note that different genotypes can result in the same phenotype). Genotypes highlighted in red have phenotypes able to survive on 'B' fuel and configure the correspondent genotype network.

Elements of the network can differ among them in as much as 80% of their reactions, and still solve the same problem. This feature allows exploring the network without the deadly consequence that losing the initial function would carry when changing any of the initial reactions. Paraphrasing Popper: testing bad areas of the network allows the organism and its descendants to remain alive to test again[119]. This result is true not only of glucose-based metabolisms, but of all metabolisms based on any 'fuel'. Wagner discovered a second important feature of these networks: different networks, solving different problems, share the same members, and are thus inter-connected. This means that a modified genotype can gain new functions (enter a new network) while exploring its initial one (Fig. 4.14). If the exploration just preserved the initial function without adding any new one, exploring would be pointless, (p. 102).

[119] The quote usually attributed to Popper is: "Good tests kill flawed theories; we remain alive to guess again.", but I could not find the original source.

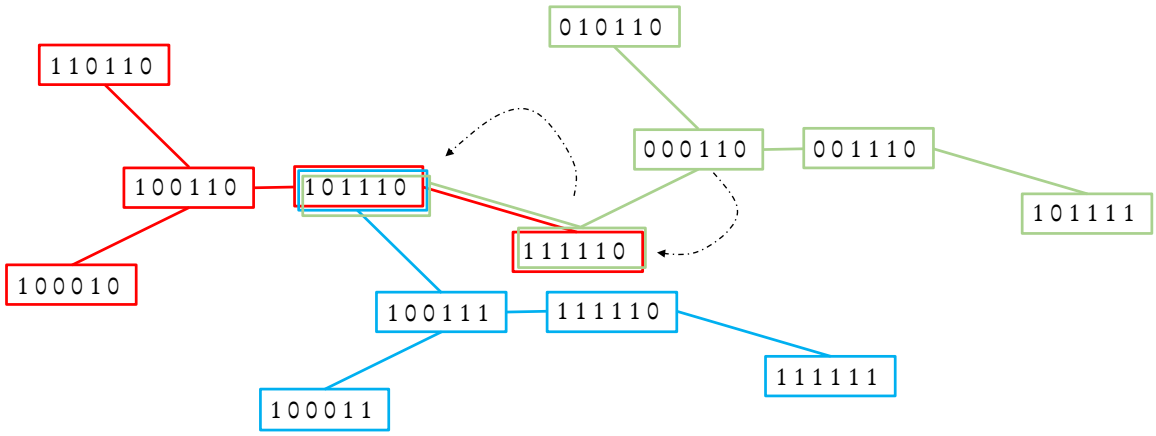


Fig. 4.14 – Inter-connection of different genotype networks. Genotype '111110' belongs to the red and the green networks. Genotype '101110' belongs to all the three networks. Subsequent modifications of the genotype '000110' (green ability) for some evolutionary phenomena could lead it to become '111110' and gain the red ability (without losing the green one), and then '101110', adding the new blue ability to the previous ones.

The same network structure with the same two features is found when analysing the library of twenty amino acids and the corresponding protein space. Many proteins can support the same functions: e.g. anti-freezing function is performed by different proteins in the Arctic and Antarctic fishes (p. 112); different organisms use different proteins (enzymes) as crystalline refracts light (p.177).

Regulation of gene expression too can be analysed under the paradigm of library/network. Regulation consists in the tuning of the activity of molecules, typically genes (e.g. the beta-galactosidase gene controlling the enzyme production in *E. coli* in the presence of lactose, Jacob & Monod 1960), and it is a fundamental mechanism in embryonic development, that allows cell differentiation and, hence, the development of a body from a single cell (p. 142). The fruit fly deposits RNA transcriptions of several genes in different areas of the egg before laying it. Once fertilized, the egg divides into differentiated cells thanks to the proteins (regulators) synthetized by these RNA molecules that switch on or off the cells' genes according to their position. Regulators are connected in chains, cascades and circuits, and they activate and repress each other according to complex rules. Hox genes are one of such chains of regulators that control the body plan along the organism's main axis (e.g. backbone in vertebrates). Once body segments are formed during development, Hox proteins define the correspondent appendage (e.g. arms or legs in humans; legs, antennae or wings in fruit fly), although they do not intervene in the subsequent part development. The sequence of Hox genes in chromosomes often follows the same order than the animal axis (see Fig. 4.15).

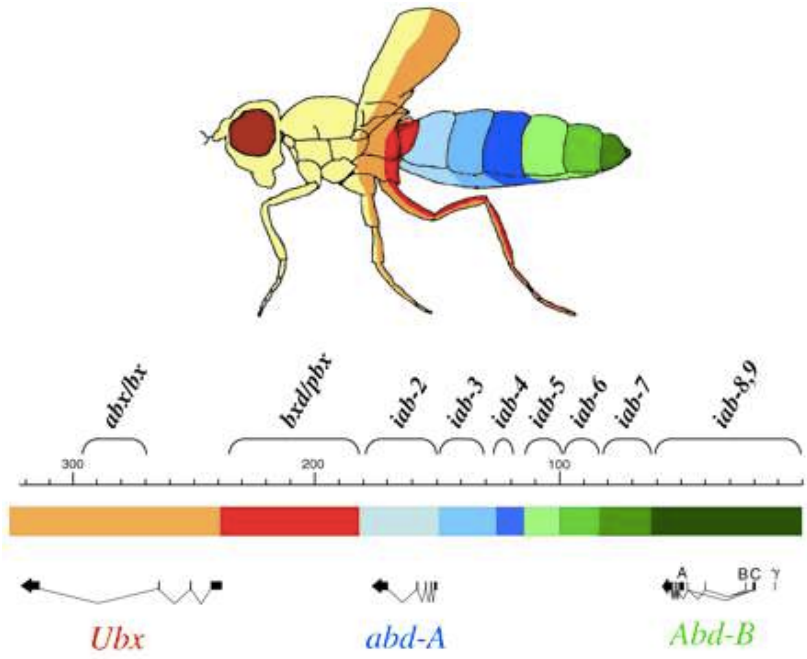


Fig. 4.15 – Collinearity between sequence of Hox genes in chromosome and body plan (from Maeda & Karch, 2006)

Hox genes circuits are fundamental for the appearance of innovations, both as new body parts and as new body plans (p. 154). Over millions of years, evolution used combinations of Hox genes to create fins of fishes and limbs, and to design the limb's different parts; and to define body plans for walking, swimming and flying animals.

Wagner builds a library of circuits based on all possible regulators and their interconnections that can be either on or off. Again, the resulting circuits turn out to be integrated into networks having the same features as in the case of metabolism and protein networks. An organism can explore the network through silent mutations, moving to neighbour circuits that express the same phenotype.

Wagner concludes that the three kinds of genetic networks show common features (2011, ch. 5):

- Different genotypes connected in a network express the same phenotype;
- Genotype networks traverse a large part of the genotype space;
- Therefore, a network has generally many other networks in its neighbourhood, and different areas of such neighbourhood contain different phenotypes;
- Innovation arising from these networks is thus combinatorial.

The peculiarity of the analysed networks resides in the robustness of their architectures, meaning the "persistence of life's features in the face of change" (p. 170). There are two fundamental ways to insure robustness (Wagner 2005, 2011): the first is to have duplicate

genes, so that, if one is lost because of a mutation, the other can still perform the needed function in the same original way (*robustness of a part*). It is the kind of robustness of an airplane with two engines: if one fails, the other can keep the airplane flying; or of a ship with multiple sails, which can also lead to sub-functionalization (Fig. 4.16).

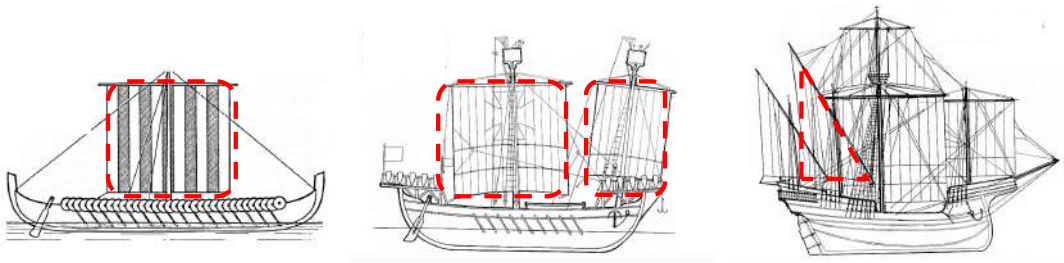


Fig. 4.16 – Evolution of sails from single square sail in Viking boat (900 aC), to double square sails (duplication) in Norse boat (1100 aC), to multiple square and triangular sails (sub-functionalization: triangle sails allows to sail against the wind) in Venetian boat (1500 aC).

The second way to make a system robust is to have different *alternative* ways to perform the same function (*distributed robustness*). If a part of a system fails, there isn't a duplicate unit, but the system switches to an alternative way to perform the same function. When your navigator recalculates your route in order to avoid a traffic jam, it is recurring to this kind of robustness. A great percentage of genes in the genotype of any organism apparently has no function at all; and the majority of genes does not have duplicates: the loss of one of them is like a traffic jam in, e.g., a metabolic route: other genes, till that moment silent, guarantee an alternative way to get to the desired destination.

Robustness allows the organism to survive to otherwise lethal mutations; and, at the same time, it allows the exploration of the genotypic space through the accumulation of changes as a cryptic stock not expressed in the phenotype[120] (defined as the sum of selectively-relevant traits: traits irrelevant to selection can of course appear or be modified by the changes in the genotype). The exploration of new traits allowed by robustness does not, however, fix them across the population: that is a task performed by natural selection, once and if the traits become selectively relevant. Most mutations are neutral, and they are fundamental to browse the networks. They stay neutral until a change in the environment makes them relevant and selection eventually fixes them.

A last, fundamental feature of networks is their close interconnection. Starting from any point of the network, it takes an average of only fifteen steps away to find another network, and acquire new functions: this is equivalent to exploring just 10 exp (-100) of the library.

[120] Kimura (1983) suggests that most genetic variation seems neutral, Wagner suggests the underlying reason: the peculiar architecture of genotype networks.

This is why, according to Wagner, life is so immensely diverse: there are many solutions to the problem of surviving: "the haystack contains more than one needle" (p.186).

Robustness makes very complex organisms: this is the cost of the ability to survive in many different environments. *Buchnera aphidicola*, a bacterium inhabiting the body of aphids, has a very simple metabolic network of just 263 reactions but it can only survive in its current environment. The virtuous circle is in brief made up of environmental change requiring complexity, begetting robustness, begetting genotype networks, enabling innovations. Allowing life to cope with environmental change is "an ascending spiral of ever-increasing innovability" (p. 194). And innovability is allowed by the self-organized architecture of genotype networks.

4.2.6. Molecular Drive

Dover (2002) proposes a new process ('molecular drive') that acts at the genetic level but is fundamentally different from the classical population genetics' accounts because it affects evolution by changing a population genetic pool independently of natural selection and drift. Molecular drive rests on "ubiquitous mechanisms of DNA turnover (MOT), such as gene conversion, unequal crossing over, slippage, transposition, retrotransposition and so on [which are] essentially non-mendelian, in that the initial mutant sequence can increase or decrease in copy number within the lifetime of an individual" (Dover 2002: 588). Such mechanisms affect DNA behaviour, while chromosome behaviour is affected by meiosis: such separation is at the basis of molecular drive. Although the rate of MOT mechanisms is too low to affect evolution of a small number of progeny or of a few generations, it can gain relevance over thousands of progenies and/or generations. The Theory of Molecular Drive has not gained a wide acceptance.

4.2.7. Final comments on the genetic sources of variation

The importance of genetic footprints for the configuration and evolution of phenotypes is undeniable: but so is the complexity of the genes' expression and mutual relationships. Concepts like genes, alleles and functions are theoretical fictions that sit on a complex molecular richness: as such, they are useful to formalise the evolution of populations as long as such hidden complexity does not play a primary role. The proposals illustrated in this section try to improve our understanding when it does play such a role, by integrating the simplistic account of the phenotype as deterministic output of the genotype through a linear

mapping, that causes small changes in the phenotype following small changes in the genotypes.

Wright's 'shifting balance' points at the creative potential of new mixes of existing alleles: shuffling the cards can change the rules of the game. Drift is not just a sorting process, but also a source of novelties that can open new, potentially more efficient, routes to selection. Phenomena of 'punctuated equilibria' show that phenotypic changes are not necessarily incremental, and that evolution is not always a strategy of continuous 'trial and error'. Wagner's genetic networks highlight that 'sudden' appearance of novelties can occur without the need for macro-mutations, thanks to the subjacent architecture of the chemistry of life.

All these proposals could in some way, at least in theory, be integrated within population genetics formalizations, something that could be harder for the accounts of the next section.

4.3. Phenotypic accounts

4.3.1. Exaptations

Exaptation is a concept introduced by Gould and Vrba (1982) to identify traits whose contribution to fitness is not the original reason for their fixation on the population. Darwin already considered in the *Origin* that an organ might have originated for a completely different purpose than the current one: sutures in the skull of mammals are an advantage in parturition, but cannot be an adaptation to this end because they are also present in reptiles, and are thus probably due to the law of growth (Darwin 1859: 197). In another passage, the idea is even more explicit:

> The illustration of the swim-bladder in fishes is a good one, because it shows us clearly the highly important fact that an organ originally constructed for one purpose, namely flotation, may be converted into one for a widely different purpose, namely respiration. (Darwin 1872: 205)

Gould and Vrba (1982) develop the same idea by underlining two common criteria to recognize a trait as 'apt':

- Historical genesis: if the trait has been built by selection because of its current *function*. The trait is an *adaptation*.

- Current utility: if the trait has a positive *effect* on current fitness, regardless of its historical genesis (thus, not necessarily a process of selection). The trait is 'apt' but not 'adapt'. Gould and Vrba baptize it an *exaptation*.

An exaptation can either be a trait appeared as an adaptation for some original function that is later leveraged to perform a new one; or a trait due to some kind of constraints and lacking any function, until an environmental change grants it one.[121] Examples of the first category are feathers, a thermoregulation tool later exploited for flight (Kingsolver and Koel 1985); bones, selected as a phosphate storage before becoming a structural supporting tool; and the gas bladder of fishes, an exaptation of the lungs of basal fish (Colleen 1997). Exaptations of the second type are the mentioned sutures in the skull of mammals, due to developmental constraints; the shape of sponges and corals due to marine currents (Gould & Lewontin 1979); crystallines, proteins originated from metabolic enzymes that, thanks to their ability to retain transparency at high concentrations, originated lenses; and repetitive DNA, neutral at phenotypic level and thus invisible to selection, until co-opted for some new function.

The concept of exaptation can help solving the puzzle of 'pre-adaptation', or the problem of justifying incipient stages of an adaptation (e.g. the difficult-to-identify selective advantage of, say, the 5% of a wing). An incipient structure can be at the same time an adaptation for a function *and* a potential exaptation for another one. Pre-adaptation becomes a meaningless concept, a backward neo-Darwinian interpretation of an exaptation 'before the fact'. There is no such thing as a 5% of a wing: there is the 100% of a thermoregulation structure that, at a certain point, allows (rudimentary) flight and 'exapts' to the 100% of a (rudimentary) wing.

Once appeared, an exaptation will usually be shaped and optimized by selection: any complex structure is probably a mix of exaptations and (secondary) adaptations. Some authors (Barve & Wagner 2013) claim that the architecture of genetic networks shows how exaptations exceed adaptations several-fold, and that this could mean that selection is not the primary source of evolutionary innovations.

[121] Gregory (2008a: 361-363) lists an exhaustive classification of evolutionary changes that can be defined as adaptations.

4.3.2. Phenotypic plasticity and genetic assimilation

Some authors claim that novelties can first appear in the phenotype, and only afterwards are encoded in the genotype. Three of such proposals are summarised here: the Baldwin-effect, Waddington canalization and West-Eberhard genetic accommodation.

Baldwin-effect. The so-called Baldwin effect (Baldwin 1896) consists in a process "by which non-heritable developmental accommodation of novel inputs, which makes an organism fit in its current environment, can become internalized in a lineage and affect the course of evolution." (Badyaev 2009: 1125). The Baldwin effect entails that phenotypic plasticity can affect the evolution of a trait: "If individuals vary genetically in their capacity to learn, or to adapt developmentally, then *those most able to adapt* [the ones showing higher level of phenotypic plasticity] will leave most descendants, and the genes responsible will increase in frequency. In a fixed environment, when the best thing to learn remains constant, this can lead to the genetic determination of a character that, in earlier generations, had to be acquired afresh in each generation" (Maynard Smith 1987, italics mine and note). Phenotypic plasticity is the ability of individuals to adapt some trait of their phenotype to their environment during their lifetime, or "the genetically influenced capacity of an individual to develop into one among a range of phenotypes" (Ancel & Fontana 2000: 243; see also Schneider 1993, Pigliucci et al. 2006). A phenotype showing no variation to a varying environment has null plasticity (see Fig. 4.17).

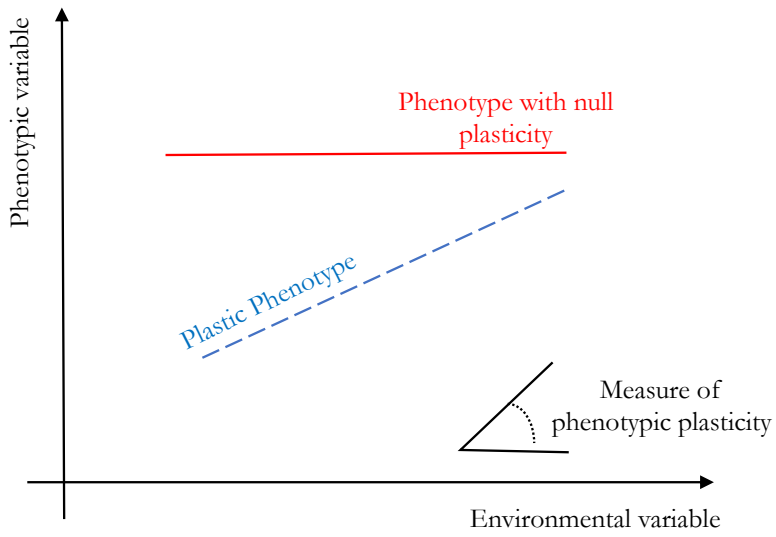


Fig. 4.17 – Phenotypic plasticity of a trait to a given environmental variable.

The Baldwin effect rests upon the following two hypotheses about phenotypic plasticity:

A. It is adaptive in that it allows an individual to improve its fitness during its lifetime by acquiring a trait through 'learning'. Note that 'learnt' trait here means just 'acquired during the individual's lifetime': it refers to phenotypic accommodation[122] and not to any cognitive ability (see the two-legged goat effect quoted in Crispo 2007 and West-Eberhard 2005b).
B. It is costly. *Ceteris paribus*, this means that a genetic encoded trait grants a higher fitness than the same trait when learned, as learning a trait is a costly activity, thus less efficient than the option of a hard-wired alternative (Price et al. 2003). Mayr (1963) refuses validity to the Baldwin effect because it does not believe that phenotypic rigidity is always selectively superior to phenotypic flexibility.

The steps of the Baldwin effect are:

- A change in the environment, or the colonization of a new environment, makes a new trait, not yet genetically encoded, relevant for the individual's fitness;
- Some individuals possess the phenotypic plasticity needed to 'learn' (acquire) the new trait (hypothesis A);
- Selection acts on the population and grants a fitness advantage to the plastic individuals displaying the learned trait (Baldwin calls this phenomenon *organic selection*);
- After some generations, the majority or all the individuals of the population show a plastic phenotype and the learned trait;
- At some point, an individual appears that genetically encodes the new trait (a 'talented individual') thanks to e.g. mutation or drift (Pigliucci et al. 2006, fig. 2);
- The fitness of the mutant is higher than the fitness of the individuals having to learn the new trait (hypothesis B);
- Finally, the mutant phenotype takes over the plastic one and the trait is integrated within the genetic pool of the population (Baldwin calls this process *orthoplasy*).

At the end of the process, a phenotype induced by the environment is thus genetically encoded and inherited. The key here is that plasticity ensures the survival in the novel environment, so that selection can act on the population without destroying it (Sznajder et al. 2012, Crispo 2007). Learning is clearly a quicker way to explore phenotypic space than genetic experimenting, although potentially less efficient for the individual. Although the

[122] "Phenotypic accommodation is adaptive adjustment, without genetic change, of variable aspects of the phenotype following a novel input during development." (West-Eberhard 2005b: 610)

Baldwin effect does not entail the inheritance of acquired phenotypes ("The ontogenetic adaptations are really new, not performed; and they are really reproduced in succeeding generations, although *not physically inherited*", Baldwin 1896: 451, italics mine), it does identify in the phenotype and its plasticity a source of evolutionary change, at least in the sense that the plasticity allows and accelerates evolution of certain traits that, for example in a new environment, become relevant for survival (Hinton and Nowlan 1987). On the other hand, the lack of plasticity could lead to the extinction of the population (Price et al. 2003). Opinions around the existence and relevance of the Baldwin effect are discordant (Depew 2003, Sznajder et al. 2012, Santos et al. 2015).

Badyaev (2009) gives a thorough analysis of a possible Baldwin-effect that could explain the origin of novel adaptations during the (on-going) colonization of North America by the house finch. New environmental conditions (e.g. temperatures exceeding egg-viability limits) induce new behaviours (e.g. modification of incubation onset), that are selected and accommodated, and might later be genetically retained.

Crispo (2007) mentions two examples of putative cases of Baldwin-effect at different stages of progress:

- A population of the bird *Junco hyemalis*, after colonising a coastal environment with milder weather than the original mountainous one, shows a longer breeding season, a clear adaptive modification. The plasticity of the trait 'breeding time' clearly facilitated the colonization of the new area. The species is undergoing organic selection. This is an example of the first step of the Baldwin effect, in which genetic accommodation has not yet occurred;
- The salmon *Salvelinuns alpinus* shows three sympatric head-morphologies, reflecting preferred prey type. The various morphologies also appear in laboratory-raised individuals, suggesting a genetic encoding. Phenotypic plasticity followed by genetic accommodation could be at the basis of such niche-diversification. The species has undergone both organic selection and orthoplasy.

The Baldwin effect has something in common with exaptations: the ability to plastically adapt following a change in the environment is an already existing neutral trait that becomes relevant to fitness in the new situation without having been selected for that purpose. We can consider the Baldwin effect as a secondary adaptation that, based on cost minimization to reach the same fitness, moves an exaptation from the domain of phenotypic plasticity to the

genetic domain. Note that such an optimization reduces the flexibility of the phenotype: in the original environment, plastic individual would be better off than genetically-encoded ones.

Canalization. Phenomena similar to the Baldwin effect have been repeatedly noticed. Artificial selection for tameness in silver fox also led to unexpected physiological and phenotypical changes. Interpretation pointed to a great reservoir of 'dormant' genes that stressful situations (such as domestication) can activate: an epigenetic event can increase heritable variability (Jablonka & Lamb 2005: 259-260). Waddington (1956) drew the same conclusion by observing that the 'wild-type'[123] phenotype is usually very homogeneous in a population, while the phenotypic effects of the same mutation can vary widely. He coined the concept of 'canalization' to account for the phenomenon: normal development is constrained through certain paths (firmly established by selection during a long evolutionary history) that protect it from a range of environmental or genetic shocks, and that make it a robust process. New mutant types do not have the same stability, so their phenotypes reflect any small change in genotype. Canalization allows the accumulation of cryptic genetic changes that can be expressed and become available to selection when the triggering perturbation reaches some threshold: a new developmental path can be opened and, if selected, can create a new stable pattern even in the absence of the original perturbation. Waddington defines as 'genetic assimilation' the process through which an induced epigenetic event[124] creates a new phenotype that, if selected, will be genetically encoded. It is a quasi-Lamarckian mechanism, but not a pure one, because the phenotype induced by the environment is not necessarily adaptive: genetic assimilation is "a process whereby environmentally induced phenotypic variation becomes constitutively produced (i.e. no longer requires the environmental signal for expression)" (Pigliucci et al. 2006).

Waddington proposed his hypothesis based on experiments with *Drosophila*. He induced a *bithorax* (a second thorax) abnormal phenotype in some individuals of the fruit-fly by spraying eggs with vapour of ether in non-lethal doses. The abnormal individuals were selected as parents of the following generation, and their eggs submitted to the same treatment. Generation after generation, the frequency of the *bithorax* type steadily increased

[123] The most common phenotype in a population: all other types are mutant. The tiger wild-type, for example, has orange fur and black stripes; albino tigers are a mutant-type.

[124] Epigenetic events causes heritable phenotypic changes not due to DNA alterations (e.g. mutations), but to genes' activity or expression (Dupont et al. 2009). See also § 4.3.4 later in this chapter.

and, after 30 generations, the type appeared *even in the absence of treatment*. What was initially an environmental-induced phenotype allowed by (developmental) plasticity became later genetically encoded.

The perturbation need not be environmental: a mutation can lead to the same effect. The protein Hsp90 protects embryo development in *Drosophila* from heat shock. Phenotypic variants appear if the protein is damaged, either because of environmental causes (e.g. drugs) or genetic causes (e.g. mutation). If a variant is selected, it gets fixed in the population *even after the mutation that originally caused it disappears* (Rutherford & Lindquist 1998): Hsp90 returned to its original (non-mutant) form in the new variant phenotype. This example underlines the difference between novelty due to genetic assimilation, where mutation is just a kind of 'catalyser' or trigger for change with a temporary role, and might even disappear in the new phenotype, and the classical neo-Darwinian account of novelty as mutation plus selection, where the mutation becomes part of the new genetic pool.

Phenotypic Plasticity. West-Eberhard (2003, 2005a) proposes a similar mechanism, based on phenotypic plasticity, followed by genetic accommodation, to account for novel phenotypes.

Novelties can be triggered by environmental factors, and not only by genetic mutations. As a matter of fact, adaptive evolution requires genetic variation, but selection (differential reproductive fitness) just requires phenotypic variation, regardless of whether its bases are genetic or not: any phenotypic trait can be the target of selection. This means that *there can be selection of phenotypes during generations without evolution* (in the population genetics sense of changes in genes' frequencies) nor genetic variation. Cases of great morphological changes with little genetic changes are easily accommodated under this view. A morphological change can appear and persist thanks to developmental plasticity linked to environmental factors (Fig. 4.18). As Lewontin points out:

> The consequence of the interaction of gene, environment, and developmental noise is a many-to-many relationship between gene and organisms. The same genotype can give rise to different organisms, and the same organism can correspond to different genotypes. Ontogeny is not a linear array of stages, one leading always to a particular next stage, but a branched set of pathways. (Lewontin 1983: 94)

If a mutation backing the phenotypic novelty appears, the novelty gets fixed in the population: an evolutionary phenomenon occurs.

The steps leading to the fixation of a phenotypic novelty are:

1. An ancestral population of plastic organisms, i.e. organisms potentially able to produce phenotypic variants under the effect of new inputs either from the genome (e.g. mutations) or the environment (e.g. colonization of an island is a typical environment change favourable for a process of genetic accommodation of changes triggered by phenotypic plasticity (West-Eberard 2005), like in the Galapagos' finches), actually produces such a novelty. The source of the novelty is the plasticity of the developmental system. Developmental recombination allows traits to be expressed in new combinations or to undergo quantitative changes. Recombination can be linked to specific decision points during development, switches that can direct development along one or another direction. Such decision points can also help defining what a trait is. For example, digit variation of vertebrate limbs might be due to experimental, mutational, pathological, evolutionary modifications to the limb development system. The recurrence of certain phenotypic traits in discontinuous phylogenetic distributions can be explained by these switch mechanisms.
2. The novelty has a fitness effect and becomes target of selection. This just means that the individual exhibiting the new trait will e.g. have more descendants, not that the descendant will show the same trait. However, if the environmental trigger (e.g. food, temperature, parental behaviours) continues to act across generations, and the descendants have the same plasticity, they will also express the new trait. The trait will *spread through the population even without genetic base*. While mutational-based novelties can be rare, an environmental-based one has two peculiar features easing its propagation:
    a. It affects several individuals at the same time;
    b. It persists across generations even if it has a null or even negative fitness effect.
3. The novelty finally gets a genetic base (because of genetic accommodation) and leads to adaptive evolution. Genetic recombination of the existing genetic pool is usually sufficient to support genetic accommodation, without the need of mutations. Under this view, *the genotype is a follower and not a leader in the evolutionary process*.

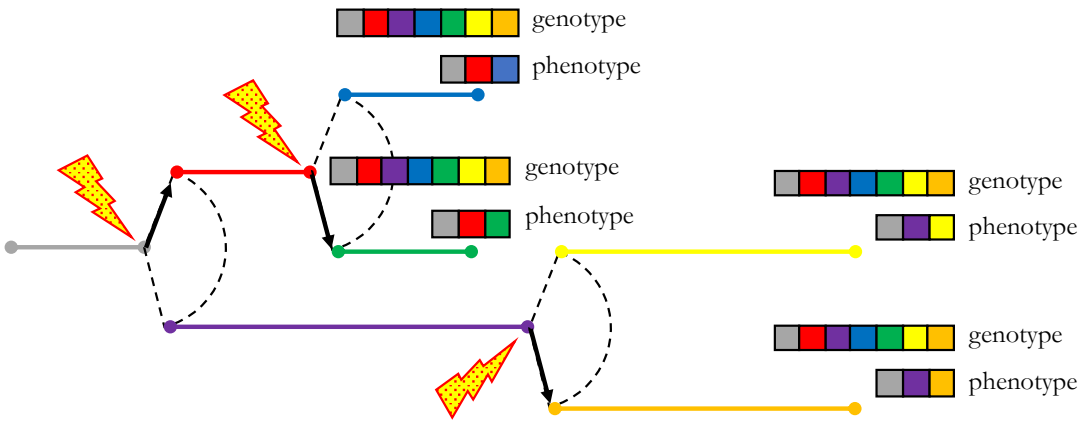


Fig. 4.18 – Logical developmental switches leading from genotype to phenotype: different phenotypes can result from the same genotype.

The proposed mechanism might account for several peculiar properties of species:

- Permanence of the same genes in distant related species and over long period of times;
- Small genetic differences between species with very different phenotypes;
- Frequent homoplasy (a shared trait between two or more species that did not arise from a common ancestor) and parallelism (an adaptation developed independently by two or more species);
- Appearance of phenotypic differences within species before reproductive isolation.

The main points of genetic assimilation can be summarised as follows (Jablonka and Lamb 2005: 272-273)

- Natural populations have huge hidden genetic variation;
- Canalization, through buffering systems like the Hsp90 protein, insures that environmental shocks and new mutations do not affect development and, at the same time, allows the accumulation of additional hidden genetic variation;
- Buffering capacity is not unlimited: extreme shocks above certain thresholds can overcome canalization and lead to the sudden expression of the hidden genetic material in new phenotypes;
- If advantageous, the new phenotype is selected and a new developmental canalization is created, so that it will be maintained even in the absence of the inducing agent.

The discussion around phenotypic plasticity, genetic accommodation and their relationship with natural selection helps clarifying the proper domain of each concept (Pigliucci et al. 2006):

- Phenotypic plasticity is (in part) a developmental process, a target of selection;
- Natural selection is an evolutionary mechanism (one among others);
- Genetic accommodation is (under certain conditions) the outcome of selection.
- Therefore, plasticity generates variation, and selection fixes it fostering accommodation.

Fig. 4.19 compares, among them and with Neo-Darwinism, the three proposals analysed in this section.

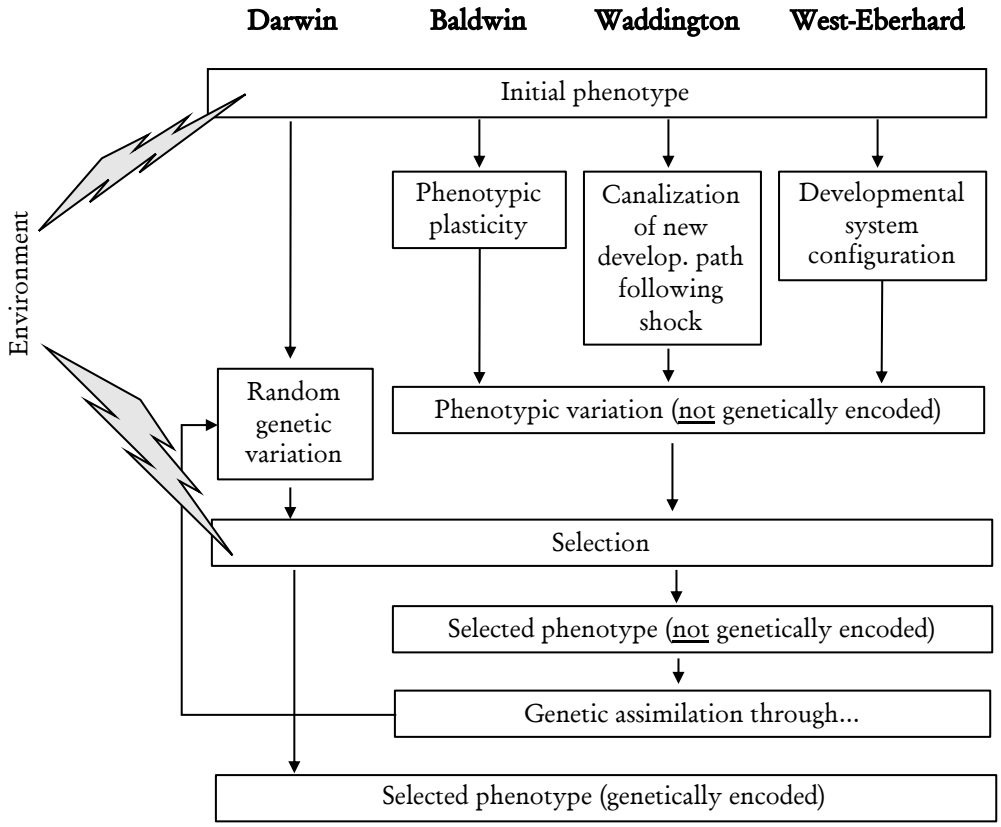


Fig. 4.19 – Comparison among different variation-generating hypothesis (see text for details).

According to classical neo-Darwinian though, variation, upon which selection acts, is random and independent from environmental pressures. Baldwin believes that variation is linked to the response of a plastic phenotype to a new environment. Waddington focus on the ability of the developmental system to integrate some changes initially due to external shocks, even after the shock disappears. West-Eberhard focuses likewise on the developmental system, that can be configured in many different ways to better cope with the

environment, without genic modification. All three proposals agree that the corresponding phenotypic variation, although submitted to selection, is initially not genetically encoded, and would disappear if the environment went back to its original state. Genetic accommodation is needed at certain point, through random genetic changes, to fix the new variants in the population.

4.3.3. Neo-Lamarckian accounts: induced mutations

Random mutations occurs spontaneously even in healthy cells, and are undeniably a genetic source of new phenotypes: the mutation is at the same time the origin of the novelty and the mechanism by which it is expressed in the phenotype. Mutations induced by environmental causes (including mutagens), on the other side, are classified here as an environmental source of novelty because, while their mechanism is genetic, the event that triggers them is external to the genome.

Lamarck's theory of evolution is based on the belief in the innate tendency of organisms toward increasing complexity. The mechanism behind the accumulation of improvements is the inheritance of the characters acquired through use and the loss of those fallen into disuse. Habit is thus a fundamental intermediate step between environment and form: environment creates needs that are fulfilled by habits that entail use and disuse of organs, whose modifications are inherited by the following generations (Koonin and Wolf 2009). Although denied in current biology as incompatible with the central dogma of molecular biology[125], Koonin and Wolf believe that some Lamarckian and quasi-Lamarckian phenomena do exist and impact the evolutionary path of species. The mechanism behind these phenomena is compatible with the Central Dogma because it does not cause the information to be in some way transferred back to nucleic acid; rather, it causes directed adaptive change in genomes. Lamarckism under this interpretation is adaptive, because only advantageous traits are inherited. However, unlike in Darwinism, adaptation happens in the generation of traits and not in the selection stage: variability is not the result of chance (Fig. 4.20).

[125] The Central Dogma states that once 'information' has passed into protein it cannot get out again. In more detail, the transfer of information from nucleic acid (DNA or RNA) to nucleic acid, or from nucleic acid to protein may be possible, but transfer from protein to protein, or from protein to nucleic acid is impossible. Information means here the precise determination of sequence, either of bases in the nucleic acid or of amino acid residues in the protein (Crick 1958). Koonin (2012a) suggests that prion diseases could be a problem for the central dogma.

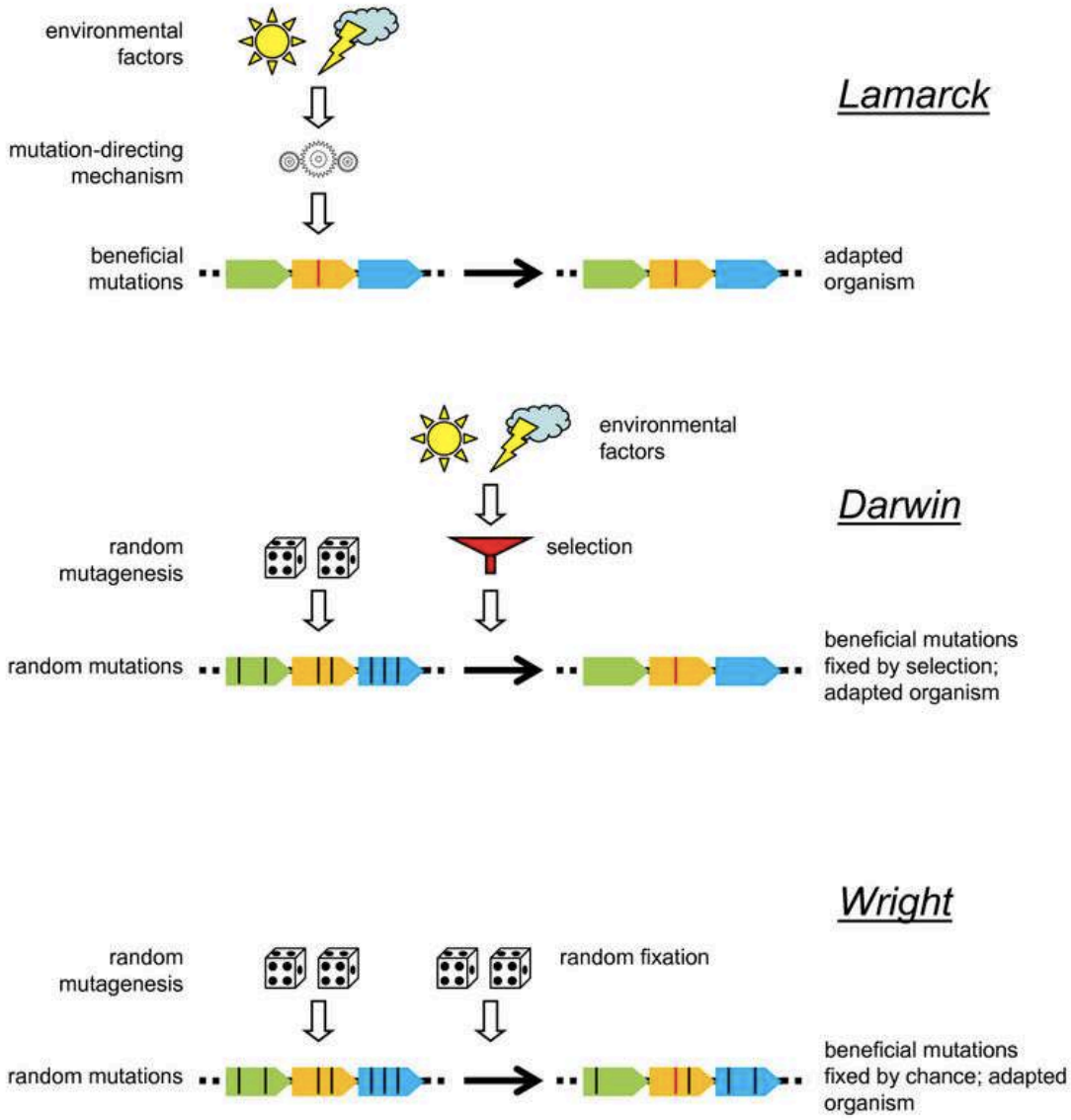


Fig. 4.20 - Similarity and differences between the evolutionary ideas of Lamarck, Darwin and Wright. For Wright, both generation and fixation of characters are random processes (linked to genetic drift). For Darwin, generation is random but fixation is selective. For Lamarck, generation is directed and pre-selected by previous adaptation. (Koonin & Wolf 2009)

Koonin and Wolf translate the three basic hypothesis of Lamarckism into modern genetics vocabulary:

1. Environmental factors cause genomic heritable changes (mutations);
2. These changes (mutations) are targeted to a specific gene;
3. These changes provide adaptation *to the initial environmental factors*.

While the first two points are compatible with Darwinism, the last one is where neo-Lamarckism departs from strict Darwinism, for which mutations are always random with respect to environmental factors (see §3.2). Razeto-Barry and Vecchi (2016) claim that all definitions of randomness in evolutionary theories are not formal, which casts a shadow also

on the concept of non-random, induced or Lamarckian mutational processes, usually defined as the negative of random ones. Induced mutation definitions end up being either too wide (like in Jablonka and Lamb 2015, see below in this section), or too strict (like in Koonin and Wolf 2009 - the one just presented), thus including phenomena that are not genuinely Lamarckian or excluding others that are. Some of the 'too strict' definitions require, for example, that adaptive mutations outnumber non-adaptive ones. The authors suggest that the total net fitness effect of all mutations (beneficial and deleterious) should be considered instead, calculated as the expected value of fitness impact of each mutation averaged over its probability of occurrence. A mutational phenomenon is Lamarckian if "[…] (i) the probability that the correlation between the occurrence of M [mutation] and E [environment] is positive, given that the correlation between S [net fitness effect] and E is positive, is higher than if the correlation between S and E is not considered, and (ii) the net fitness effect of the mechanism in some particular environment *e* from E is positive." In unformal and less precise words, if a positive net fitness effect in an environment exists, it helps explaining the relationship between mutation and the environment.

According to the Koonin and Wolf (2009), several evolutionarily relevant phenomena can be interpreted as fulfilling some or all of these hypotheses (see Fig. 4.21):

Table 1: Lamarckian and quasi-Lamarckian phenomena

Phenomenon	Biological role/ function	Phyletic spread	Lamarckian criteria		
			Genomic changes caused by environmental factor	Changes are specific to relevant genomic loci	Changes provide adaptation to the causative factor
Bona fide Lamarckian					
CRISPR-Cas	Defense against viruses and other mobile elements	Archaea and bacteria (present in ~1/3 sequenced genomes)	Yes	Yes	Yes
piRNA	Defense against transposable elements in germline	Animals (apparently, all)	Yes	Yes	Yes
HGT (specific cases)	Adaptation to new environment, stress response, resistance	Archaea, bacteria, unicellular eukaryotes	Yes	Yes	Yes
Quasi-Lamarckian					
HGT (general phenomenon)	Diverse innovations	Archaea, bacteria, unicellular eukaryotes	Yes	No	Yes/no
Stress-induced mutagenesis	Stress response/ resistance/ adaptation to new conditions	Ubiquitous	Yes	No or partially	Yes (but general evolvability/ Adaptability enhanced as well)

Fig. 4.21 – (Koonin and Wolf 2009)

- *CRISP-Cas*. It is an antiphage defence system in archaea and bacteria that seems to act via a straightforward Lamarckian mechanism. Parts of the genome of archaea and bacteria that are essential for resistance turn out to be identical to fragments of the attacking bacteriophage and plasmid genes. The heritable immunity is acquired by archaea and bacteria thanks to the modification of their genome due to an environmental cue (virus) and directly affects the same cue that caused it.
- *Horizontal* (or *Lateral*) *Gene Transfer*. Prokaryotes and unicellular eukaryotes can obtain and integrate DNA from the environment (via phages, plasmids and other organisms) in their genetic material. The new genetic material is eventually fixed in the population if selectively advantageous. Evolution of antibiotic resistance is an example of such mechanisms. The movement of genetic material between genomes within microbial communities can have significant effects on the evolutionary trajectory of a microbial lineage, such as the loss or gain of biological functions on short timescales (Eva Boon 2017). Once inserted into the genome of an organism, the new genetic material can represent a source of novelty and is inherited through the 'classical' vertical process. Almost 45% of human genes could derive from horizontal transfer: in microorganisms, it is a much more common mechanism than vertical transfer (Fodor & Piattelli-Palmarini 2010: 68-69)
- *Stress-induced mutagenesis*. Bacteria are subject to a wide range of stresses, including nutritional deprivation, DNA damage, temperature shifts, and exposure to antibiotics. Responses include functions increasing genetic variability (Foster 2007), e.g. increased mutation rates, some of which could help opposing the stress[126] (Jablonka & lamb 2005: 93). Activation of 'jumping genes', DNA sequences that can change their position in the genome, is an example of such response. McClintock (1950) was the first to show this mechanism in plants under stress and its importance in creating resistant phenotypes thanks to the mobility of controlling elements in the genotype. Van't Hoff et al. (2016) describe how the mechanism of 'jumping genes' allowed for the appearance of the *carbonaria* type of *B. betularia*. Industrial melanism was due to a mutation consisting in the insertion of a transposable element into the

[126] The majority of mutations will be deleterious or lethal, but by randomly trying many possible alternative solutions to the problem that created the stress, the probability increases of *at least one individual* finding a solution.

first intron of the gene called cortex, around 1819. The authors consequently defend the importance of 'jumping genes' as a source of major phenotypic novelty.

- Koonin and Wolf define stress-induced mutation as a quasi-Lamarckian mechanism because: "i) mutations are triggered by environmental conditions; ii) the induced mutations can [*but not always*] lead to adaptation to the stress factor(s) that triggered mutagenesis; iii) mutagenic repair is subject to elaborate regulation which leaves no reasonable doubt regarding the adaptive nature of this process." (2009, italics mine).

Jablonka et al. (1998) defend likewise the existence of quasi-Lamarckian mechanism: "adaptive mutation, epigenetic inheritance, behavioural inheritance through social learning, and language-based information transmission have properties that allow *the inheritance of induced or learnt characters*" (italics mine).

The classical neo-Darwinian account of evolution points at random mutation as the chief source of variation, where random is to be interpreted as 'blind to its selective consequences'. Jablonka & Lamb (2005) describe four kinds of environmental-induced mutations that do not seem to be completely random in this sense: "evolution by natural selection has led to the construction of mechanisms that alter DNA in response to the signals that cells receive from other cells or from the environment." (2005: 88). They call these 'half-random/half-directed' variations or 'interpretive mutations' and identify four types:

1. *Induced Global Mutations*. It is a mechanism similar to Koonin and Wolf 'induced stress mutagenesis' (2009). Mutation rate increases under stress (a well-studied phenomena in, but not limited to, bacteria). The result is that, among the many new mutants, the majority will not be viable, but a few might develop a novel solution to survive the stressing environment. Increase mutational rate can be an adaptation or just a consequence of the failure of the DNA maintenance system due to the stress. Whatever their origin, some of these mutations are non-random because they appear *when possibly useful*. *Escherichia coli* seems to increase its rate of beneficial mutations under conditions that prevent growth: without chromosomes replication nor cells division, directed mutagenesis and not selection should explain the phenomena (Cairns & Foster 1991, but see critiques in Maisnier-Patin & Roth 2015).
2. *Local Hypermutations*. They are mutations appearing in particular places in the genome at a rate hundreds or thousands of times higher than elsewhere. They are non-random because they appear *where possibly useful*, not when: mutations appear

constantly at these places, not when they are triggered by external events. The genes determining the surface structure of the bacterium *Haemophilus influenza* present local hypermutations in so called 'contingency genes', which allow the bacterial population to continuously try different surface structures that make them survive in new environments and to puzzle the host's immune system. The authors claim that selection is behind the local hypermutation mechanism.

3. *Induced Local Mutations*. A kind of mix between the first two types, these mutations occur *where and when* they can be useful. They are environmentally-induced and located in specific locations in the genome: thus, they are "both required and acquired" (2005: 97). Their mutation rate lies in between the first and the second kind of mutations: five to ten times higher than normal mutations. Experiments in *E. coli* show that local mutations in a non-functional gene producing a given amino acid can be induced by the shortage of that amino acid together with a stress signal: some of the mutations can make the gene functional (Wrigh et al. 1999, Wright 2000).
4. *Induced Regional Increased Mutation*. Interesting because found in multicellular organisms, they are triggered by a change in external conditions and involve a specific set of genes, whose mutation rate is increased by several orders of magnitude.

These mechanisms suggest that phenomena happening to individual organisms in a generation can impact the genetic variation in the following generation. The Central Dogma is not broken, however, because "backtranslation is not necessary for acquired characters to be inherited" (Jablonka & Lamb2005: 104): genetic change simulating the acquired change affects the regulatory, not the coding sequences.

The difficulty of finding examples of these four kinds of mutations shows that they are rare. Causes for their low frequency, despite the potential benefits they can provide to the organism, are twofold: environmental problems do not usually repeat themselves, so we do not expect fine-tuned induced solutions; and complex organisms have many intra-genic interactions and an intricate genotype-to-phenotype map, so induced mutations favourable in one sense are probably deleterious in many others. Complex organisms inherited probably acquire characters through supra-genetic mechanisms (epigenetic, behavioural, symbolic) (2005: 105-107).

Fig. 4.22 summarizes the characteristics of these four kinds and of the classical, neo-Darwinian blind mutations.

Type of genetic change	Targeted to a specific gene or region?	Induced or regulated?	Adaptiveness of the type of change	Type of DNA alteration
Classic blind mutation	No	No	None	Changes in bases, mistakes in repair and replication, movements of mobile elements, breakage and rejoining, etc.
Induced global increased mutation	No	Yes, by extreme stress	None, but elevation of the general mutation rate may be adaptive	Elevation of overall blind mutation rate
Local hypermutation	Yes	No	Yes	DNA sequence organization leads to high mutability in specific regions
Induced local increased mutation	Yes	Yes, by nonextreme stress	Yes	Mutation targeted to specific active genes
Induced regional increased mutation	Yes	Yes, by changed environment	None (as far as is known)	Mutations targeted to particular DNA repeated sequences
Developmental	Yes	Yes, regulated by developmental signals	Yes	Precise genomic changes and mutations in well-defined regions

Fig. 4.22 - Characteristics of randomness and directedness of different kinds of mutations (from Jablonka & Lamb 2005: 100)

The table clearly shows that the dual categorization 'blind vs. directed' variation only works at the extreme of the range. The distinction between generating and selecting the variation is not so clear-cut in many cases. Local mutations, for example, are randomly generated, but they rise in specific spots of the genome: selection might have generated an 'informed-guessing' method to increase efficiency in the exploring of genotypic space. The conclusion is that

> […] Darwinian evolution can include Lamarckian processes, because the heritable variation on which selection acts is not entirely blind to function; some of it is induced or 'acquired' in response to the conditions of life. (2005: 102)

If as Neo-Darwinism claims, the process of generating new variations were completely random, it would be the only aspect of the living world not produced by selection: a strange fact indeed.

4.3.4. Epigenetics

Epigenetics is "the study of *changes in gene function* that are mitotically and/or meiotically heritable and that *do not entail a change in DNA sequence*" (Dupont et al. 2009, italics mine). Failure of cloning based only on genetic information clearly shows that the "popular notion that the genome contains 'all the information needed to make a worm' is simply false." (Cavalier-Smith, quoted in Jablonka & Lamb 2005: 123). The genetic information of specialized cells (liver, skin, kidney, etc.) in the same individual is identical: what changes among them is which genes are active or silent, and how their products (inter)act. This configuration, peculiar to each cell type, is determined during development and later transmitted to the daughter cells (from liver cells, only liver cells are obtained). This information transmission, independent of DNA, is called *Epigenetic Inheritance System* (EIS) (Jablonka & Lamb 2005: 113). Fundamental in ontogeny, EIS can also have great evolutionary potential in phylogeny. Müller defines epigenetic causation as "the idea that developmental systems do not merely transform genetic change into phenotypic change but also represent a generative component in phenotypic evolution." (Müller 2008: 10).

There are four types of EIS, each allowing the inter-generational maintenance ('memory') of past events (2005: 119 et seq.):

- *Self-Sustaining Loops, or the memory of gene activity*. A system constituted of two elements is self-sustaining if A causes B that causes A. A gene that produces a protein that keeps the gene active is an example of such a system: the original activation of the gene might be due to an external cause but, once activated, the gene remains active even if the initial cause disappears. After cell division, the daughter cells might inherit enough level of the protein and, consequently, the active state of the gene. Two genetically identical cells can therefore show different patterns of gene activity,

depending on the their ancestors' histories, and this pattern might be transmitted generation after generation.

The unit of heritable variation is the *state of the loop* and not, like in genic inheritance, just its components. Considering that a cell can have many different loops, and that each loop can have at least two states (active/inactive), the number of functional variants that are epigenetically inheritable, and upon which selection can act, increases exponentially.

- *Structural Inheritance, or structural memory*. Cells can have alternative versions of their part's structures, versions that the daughter cells inherit, because existing structures guide the creation of new ones, acting as a 'template'. For example, changes in the direction of cilia through microsurgery in ciliate *Paramecium* is transmitted to offspring (experiment by Tracy Sonneborn in the 1960s, discussion in Preer 1997): they showed the new cilia direction, although genetically identical to descendants generated before the surgery. The ability of prions of transferring their abnormal secondary structure[127] to normal proteins of a subsequent generation of cells is another example of structural inheritance.

  What is inherited in this type of EIS is the *organization* of a system, not its state like in self-sustaining loops. Structure replication does not relay on any mechanism, as it happens with DNA replication: "The ability of a structure to be reconstructed in daughter cells is inherent in its organization." (Jablonka & Lamb 2005: 126).

- *Chromatin-Marking Systems, or chromosomal memory*. Chromatin is the complex of DNA, RNA and proteins (e.g. histones), found on eukaryotic cells, whose functions are to package DNA, to reinforce it to allow mitosis, to prevent DNA damage and to control gene expression. There can be a variety of ways to package the same DNA sequence in different cells or at different times in the lifecycle of the same cell, and each way activates or silences different genes. Some features of chromatin ('chromatin marks') are heritable. One of the most known, DNA methylation, consists of the attachment of a methyl group (CH_3) to some of the DNA bases: during replication, a likewise methylated pattern is maintained. When density of such groups is high, protein transcription is less likely to occur.

[127] A protein secondary structure is the three dimensional form of local segments of proteins, whose primary structure consists of the sequence of amino acids in the polypeptide chain. Prions are misfolded proteins which characterize several fatal neurodegenerative diseases in animals and humans (Prusiner 1998).

'Protein marks' and 'histone marks' are other types of chromatin marks, although their details and workings are still largely unknown. They all allow inter-generational modular[128] transmission of information about gene activity.

Epigenetic marks influence the probability that the affected DNA region will be target of genetic changes (mutations, recombinations, movement of 'jumping' genes) (Jablonka & Lamb 2005: 247).

- *RNA Interference, or the silencing of the genes.* In an RNA interference process, RNA molecules inhibit gene expression or translation in a stable and heritable way. When abnormal RNA is produced, an enzyme called 'dicer' chops it into smaller pieces called siRNA (small interfering RNA) that can cause the destruction of the original RNA and silence the gene expressing it. By introducing the correct siRNA into a cell, it is virtually possible to silence any gene.

Epigenetic inheritance systems can be found in single-cell eukaryotes and in asexual multi-cellular organisms, e.g. in plants reproducing by fragmentation, where each fragment can have different epigenetic modifications linked to different environments or initial conditions. In sexually reproducing organisms, on the other hand, the fertilized egg must start development from an unbiased epigenetic state to allow cells to differentiate correctly. Paternal and maternal genetic contribution seem however to be sometimes different: the phenotypic differences of genetically identical mules and hinnies depend on being the father a horse and the mother a donkey, or the other way round. Some epigenetic marks are probably left on egg and sperm during their production.

Epigenetic marks can be very stable across generations. A good example is the *Peloria* variant of the *Linaria vulgaris,* neither a new species nor a mutation, but an 'epi-mutation' (Fig. 4.23).

[128] Like in DNA, information is divided in independent modules, so that "[o]ften (though not always) it is possible to change the methylation state of one cytosine without affecting any others." (Jablonka & Lamb 2005: 131)

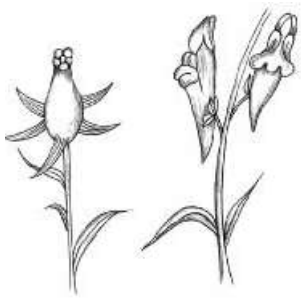


Fig. 4.23 - *Peloria* and normal variant of *Linaria vulgaris* (Jablonka and Lamb, 2005: 141)

Fodor and Piattelli-Palmarini (2010: 66) mention a case of human epigenetic inheritance. During the famine in Netherland in the winter of 1944—1945, women had babies of less than usual height and weight. Several decades later, their granddaughters' babies, despite normal nutrition, show the same unusual characters.

Epigenetic variation can thus occur even under constant genetic conditions: selection can act on such variation and lead to evolutionary phenomena. It usually occurs at higher rates than genetic variation because it often depends on environmental conditions and it *might not be blind* to function: resulting adaptation might be quite rapid compared to adaptation due to genetic variability (Jablonka and Lamb 2005: 144-145).

Induced epigenetic variation has several interesting characteristics (Jablonka and Lamb 2005: 275):

- It is especially relevant for providing non-genetic variations to small, genetically homogeneous populations;
- It arises when environmental conditions change, insuring new alternatives to selection when most needed;
- It is reversible if the change reverses;
- It can anticipate genetic changes through canalization, and give rise to genetic accommodation, Baldwin effect, etc.

Fig. 4.24 illustrates the interactions between genetic and epigenetic systems.

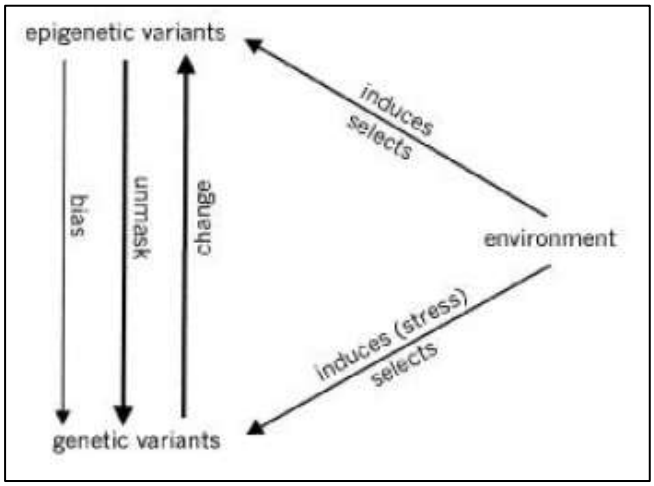


Fig. 4.24 - Interactions between genetic and epigenetic systems. Both genetic and epigenetic variants are subject to selection. Genetic variants affects epigenetic ones (e.g. by affecting the marks the genes carry) and the other way around (chromatin marks bias when and where mutations occur). Epigenetic variants, triggered by environmental events, might unmask cryptic genetic variants already present in the population, and make them available to selection. (Jablonka and Lamb 2005: 276)

The importance of epigenetic inheritance in evolution is not universally accepted. The controversial issues are mainly the persistence of epigenetic marks (provided that they appear, for how many generations they persist), the relevance of epigenetic phenomena for 'core' DNA sequences (these affecting the phenotype), the possibility that some phenotypic variation springs from epigenetic phenomena and that it is adaptive (Charles worth et al. 2007).

4.3.5. Final comments on the phenotypic sources of variations

The proposals of this section challenge the neo-Darwinian view in several ways. Exaptations reject pure functionalism, as they suggest that function can be parasitic on pre-existing forms[129]. Genetic assimilation and induced mutations can be considered as externalist accounts, as they trace back to the environment the cause of phenotypic changes: but they extend this cause well beyond what adaptationism claims. In Aristotelian terms, the environment is not just the efficient cause of the phenotype (by sorting it), but it also supplies its material and formal cause (by directing its genetic configuration *before* the sorting stage). Epigenetics contradicts the modern synthesis dogma: traits can be inherited outside of genetic mechanisms.

[129] Function is always parasitic on form in the sense that there must a morphological trait performing the function. What Neo-Darwinism claims is that form appears incrementally from a casual mutation, and function directs the increment.

To include these accounts in population genetics models is harder than in the case of the genetic accounts, but not impossible: for example, one could consider epigenetic factors as 'invisible' genes, and induced mutations as non-stochastic functions of environmental variables. The task is however close to impossible for the developmental accounts discussed in the next sections, that attack externalism.

4.4. Developmental accounts

4.4.1. Evo-Devo

The Modern Synthesis presupposes a deterministic relationship between genes and phenotypes[130] (Pigliucci 2010, Buchanan 2009) and limits the source of phenotypic variation to gene mutation and recombination: "In recent decades, the phenotype of an organism (i.e. its traits and behaviour) has been studied as the outcome of a developmental 'programme' coded in its genotype. This deterministic view is implicit in the Modern Synthesis approach to adaptive evolution as a sorting process among genetic variants" (Sultan 2017).

The developmental step that transforms genetic information into a mature individual was not included in the original synthesis, as it was considered "an incidental black box with no direct causal relevance to the evolutionary process" (Pigliucci 2010: 557): this drives "a wedge between developmental biology and the population-genetic understanding of adaptation and evolution" (Amundson 2001), although developmental constraints "undoubtedly play a significant role in evolution" (Maynard Smith et al. 1985). The human thumb, for example, cannot be explained "outside the context of its generative mechanism (the mechanisms of cartilage morphogenesis and differentiation, under particular local tissue conditions)" (Linde-Medina 2011: 582). According to Mitchell (2009: 277 et seq.), three puzzling phenomena are not accountable for within the Modern Synthesis' simplified view of development as a 'black box' where nothing relevant happens. In the first place, humans have 25.000 protein-coding genes, while other, simpler species can have substantively many more (for example, common rice has more than 46.000 genes): it is hard to see a linear relationship between information stored in the genome and organismal complexity. Moreover, humans share 90% of their genes with mice and 95% with chimpanzees, while the similarities between the respective phenotypes, however measured, are several-fold smaller. Finally, sudden (in evolutionary times) morphological changes, such

[130] See §2.2

as the ones considered in the Punctuated Equilibrium theory, are difficult to explain if based just on slow accumulation of genes' mutations[131].

Evolutionary Developmental biology (or Evo-Devo) aims at understanding how development, through its impact on phenotypes and on phenotypic innovation, can account for these phenomena and at integrating the results into the modern synthesis: its core question is thus the inter-relationship between development and evolution. The main questions that Evo-Devo tries to answer have to do with (Müller 2008: 6-11) the origin of development in multicellular organisms; how developmental processes arise and evolve; how they contribute, through developmental constraints, to phenotypic variation and novelty; how development affect the phenotypic organization; and how it is affected by the environment. Four main research programs lead the empirical research in Evo-Devo (Müller 2008: 11-16):

1. The comparative embryology programme, that studies the processes behind the evolution of developmental mechanisms of existing and extinct organisms;
2. The epigenetic and experimental programme, that analyses the ability of developmental systems to produce evolutionary relevant phenotypes, through the study of the impact on development of physical and geometrical properties, of constraints and threshold, etc. It also suggests how development can transform qualitative selective changes in qualitative changes;
3. The theoretical biology programme, that tries to formulate general rules, theorems and formal models of evo-devo events (e.g. three-dimensionals modelling of gene expressions in embryos);
4. The evolutionary developmental genetic program, that studies the genetic machinery behind embryological development.

This last program focuses on mechanisms of genes' switches: through complicated genetic networks, *regulatory genes* turn different *functional genes* on and off, so that the same functional genetic information can result in very different phenotypic traits. Cell differentiation is a typical case: skin and lever cells have identical DNA, but different phenotypes, because different functional genes have been activated during their development. Very similar genetic material can express a human, a mouse or a chimpanzee thanks to different uses of the available genetic stock. Similarly, changes in the control network can cause 'sudden' morphological evolutionary novelties without slow accumulation

[131] This final can be explained by accumulation of cryptic genetic variation (see 4.2.4 4.2.5), a phenomenon not integrated in the Modern Synthesis, as it entails, among other things, the conservation of neutral mutations.

of small mutations: evolution is based on modifications of switches more than of genes. Darwin's finches might be a striking example of this. A study by Abzhanov et al. (2004) suggests that beak differentiation among finches species is not the result of successive incremental changes selected for because advantageous in different ecological niches, but has a developmental basis. Expression of two genes, Bmp4 and calmodulin, shape beak form so that, "when mis-expressed in chicken embryos, Bmp4 caused morphological transformations paralleling the beak morphology of the large ground finch *G. magnirostris*". Chickens got a new beak just by changing the expression of a gene, without any new mutation[132]! The role of the gene Pax-6 in the development of the eye is another classic example of how control genes can shape morphology in absence of mutations. By expression of the gene in abnormal sites, eyes can be induced in legs, wings and antennae of Drosophila (Gehring 1996).

Under this view, macroevolution is fuelled by mutations in developmental genes, while microevolution by mutations in other kinds of genes or re-arrangements of genes pushed by fitness optimization processes; although, in both cases, natural selection decides of the destiny of the new phenotypes.

Integrating development in a more encompassing theory of evolution allows increasing its explanatory power, by:

1. adding degrees of freedom to the genotype-phenotype relationship (so that the same genotype can result in different phenotypes if different developmental processes act).
2. limiting the range of theoretically possible morphologies, due to developmental and morphological constraints such as modularity.
3. making the limited range more easily accessible even in the absence of genetic modifications (for example when novelties spring from inner dynamics, e.g. modularity, of the developmental system).

Note that points 1 and 2 are not in contradiction, although the first seems to suggest an increase in the phenotypic range of outcomes and the second a decrease. A new theory should explain all phenomena that the old theory explained, plus new ones: and this is exactly what integrating development into Neo-Darwinism does. Firstly, it explains why some phenotypes based on the same genotype are evolutionary significant: because developmental

[132] This does not entail that previous mutations are not required: gene expression does not happen in a vacuum; between species there are loads of accumulated small genetic variational changes; any regulatory gene produces a phenotype in such a context; past accumulated mutations are still required for the regulatory gene to have the same phenotypic effect. New mutations are not needed, but old ones are essential. (Davide Vecchi, personal communication).

systems and their modifications are heritable, something overlooked by adaptationism, that considers such differences evolutionarily immaterial. Secondly, it explains why some transitions between phenotypes are not possible and why some others are highly probable: because of constraints. Adaptationism believes all constraints are just difficulties on the way to optimization, and that can be overcome in the long run, thus playing no active role in evolution.

Evo-Devo is therefore an internalist account of evolutionary phenomena, in that it does not identify the external environment as the main drive of phenotypic change, but points out to internal development processes such as (Müller 2007):

- Heterochrony: it consists in the change in speed of a development event (onset, offset or tempo) that can lead to difference in anatomical parts in related species;
- Heterotopy: it describes changes in the location of developmental events (Russkin-Gutman & Esteve-Altava 2008);
- Phenogenetic drift, or drift of the developmental system: it produces homologous characters through different developmental processes and gene regulatory networks (True and Haag 2001, Weiss and Fullerton 2000);
- Homeotic transformation: it refers to the change of a body part into another due to a genetic/epigenetic change in development;
- Polyphenism: it refers to the appearance of alternative phenotypes based on the same genotype under different environmental conditions (Müller 2007).

4.4.2. Developmental constraints.

Developmental constraints are "biases on the production of variant phenotypes or limitations on phenotypic variability caused by the structure, character, composition, or dynamics of the developmental system" (Maynard Smith et al. 1985). They influence the direction of evolutionary change because any developmental step will tend to be integrated into development pathways and hard to remove (Rasskin-Gutman & Esteve-Altava 2008). Absence of photosynthesis in higher animals[133] and of birds giving birth to live young, or the peculiar relationship between metabolic rate and mass (rate proportional to mass exp. 0,75) are examples of probable influence of developmental constraints on evolution (Wagner 2011). Developmental constraints can also affect morphological evolution: in the radiation

[133] The definition of higher animal is of course discretional: the sea-slugs *Elysia chlorotica* do photosynthesize (thanks to Davide Vecchi for the example).

of the Anolis lizards of the Caribbean, most island were colonised by a single species, that afterwards evolved towards very similar suites of species in terms of size, colour and feeding strategies. The parallel repetition of the same morphological patterns in different islands could be explained by the existence of constraints that limit evolution along pre-determined paths (Losos 2009).

Constraints can arise from different sources, from general physical law to taxa-linked features. The accessibility or inaccessibility of some phenotypes can be linked to (Maynard Smith et al. 1985):

- the peculiar developmental mechanisms arisen in a certain taxon. Given that any taxon must have some developmental mechanism, and that these are not necessarily linked to adaptive reasons, these constraints are somehow 'accidental'. For example, palms do not branch and have trunks uniform in diameter because the process of secondary thickening has not evolved in monocotyledons (taxa to which palms belong). Wagner (2011) suggests that the fifth digit in salamander related to *Ambystoma mexicanum* (axolotl) does not develop because it is the last to develop and there is no space for it;
- physiochemical processes or qualities of complex systems. For example, banding patterns on gastropod shells can be interpreted as the solution of a wave equation, a result of standing or travelling waves of pigment in the mantle. Wagner (2011) mentions the limited size of terrestrial organisms, due to mechanical constraints), or the kind of chemical reactions that can occur in water and limits the possible metabolisms of aquatic organisms;
- the hill-climbing nature of evolutionary processes. Selection can limit the morphological solutions adequate for a given population in a given environment. For example, it is not probable that Kangaroos will evolve towards a bipedal running locomotion from the current locomotion by leaps, because the change would initially be maladaptive: populations are usually committed to a 'way of life'. Selection can also shape the developmental processes, and not only the phenotype: some constraints can themselves be the result of adaptation. Canalizing selection is an example: uniform morphology of wild-type compared with high variability of mutants might be due to constraints developed via stabilizing selections across many generations (Waddington 1956, 1957). Stabilizing selection tends to move all individuals towards the same phenotype, and selection can eliminate some

phenotypes (e.g. absence of cyclopia, or one eyed-organism, in nature) (Wagner 2011);

- the genetic system. An ideal phenotype might require mutations that are not accessible to the lineage, e.g. for epistasis.

Not all constraints are *effective* or *constitutive*, though. A constraint out of reach is real but not effective, like the maximum combinatorial possibility of the human genome (4 nucleotides in L genes give 4^L different combinatorial possibilities, a value way bigger than the possible number of individuals). A constitutive constraint like the mechanism of reproduction and heredity is real; but an imposed one can be removed at some point, like the number of nucleotides (a 5^{th} one would increase genomic possible combinations) (McShea and Brandon 2010)

Constraints contribute to determine morphologies: Rasskin-Gutman postulates the existence of a 'logic of development' that defines the possible morphologies among all conceivable ones[134]. Within the theoretical morphospace, defined by the mathematical and generative possibilities of form, only some regions are filled by actual organisms (Rasskin-Gutman 2005). Instead of on the equilibria points themselves, Evo-Devo thus focuses on the *path* from one equilibrium state (attractor) to another. The intermediate steps of the classical neo-Darwinian account might be excluded. The change to a new phenotypic architecture might therefore be not gradual, slow and driven only by the environment, but relatively quick and along a limited set of evolutionary paths (Alberch 1982, 1991). In *The Material Basis of Evolution*, Goldschmidt (1940) already underlined the importance of developmental mechanisms in the shaping of the phenotype: even minor mutations in developmental genes could result in large phenotypic changes, given that "a single mutational step affecting the right process at the right moment can accomplish everything". A new phenotype can appear in just one or very few evolutionary 'jumps'. He names these new phenotypes "hopeful monsters". As an example, he proposes to interpret the different wings arrangement in the homeotic *podoptera* mutants as result of differences in concentration of a leg inducer during development (Dietrich 2003, Box 2). Gould denies that certain phenotypic novelties, such as the split maxillary bone of the upper-jaw of two boid snakes in the island of Mauritius, could have appeared by small incremental steps, given the lack of plausible adaptive accounts for such intermediate forms (Gould 1980). Tooth and beaks represent

[134] D'Arcy Thompson anticipated this 'logic of form' idea, see §4.5.1

good examples of how development plays a main role in shaping the phenotype and in producing novelties:

> In the case of the tooth, a single-primordium organ with morphodynamic shaping in which the strengths and relationships between signalling centres is influenced by the cell and tissue movements they induce, and the beak, in which the final form is influenced by the collision and rearrangement of multiple tissue primordia, *abrupt appearance of qualitatively different forms (i.e. morphological novelties) can occur* with small changes in system parameters induced by a genetic change, or by an environmental factor whose effects can be subsequently canalized genetically. (Linde-Medina &Newmann 2014: 211, line italics)

Lenski's experiment with *E. coli*'s populations shows that the evolving bacteria's relative fitness increases in 'jumps', and that these 'jumps' become smaller in later populations (Chouard 2010). It is worth recalling that Kauffman's model makes the same prediction (see end of §4.5.3).

To make adaptationist accounts of complex traits compatible with developmental processes, developmental constraints must both allow *incremental* changes in one trait and avoid that changes in one trait affect the fitness contribution of other traits (Maynard Smith et al. 1985).

Adaptationism gives a different interpretation of constraints like 'phylogenetic inertia', or the influence of ancestors on descendants (Harvey and Pagel 1991). Vertebrates have four limbs because their ancestors (fishes) also had the same morphological configuration, even though this seems to be an optimal adaptation for terrestrial locomotion (Lewin 1980). Although the presence of inertia is undeniable, adaptationists consider that the maintenance of the configuration is due to stabilizing selection and not to inertia (Orzack and Sober 2001b): inertia is simply a constraint that, if not advantageous, is going to be overcome sooner or later by the creative power of natural selection.

4.4.3. Developmental modularity and evolvability

Rasskin-Gutman (2005) believes that evolution tends to give an advantage to modular architectures. Species will preferably occupy areas of modular design in the morphospace, leaving gaps in areas of non-modular design. Such regions are thus 'islands' of possibility in

a sea of impossible forms, and lineage would 'jump' from one region to another in successive evolutionary phenomena. While the exploration of the genetic space is random, only the genetic combinations giving rise to a modular morphology will be successful. A module can be seen as a set of elements of the genome (DNA sequences) organized in such a way as to constitute supra-elementary structures, with changeable composition but same structural properties and same functions[135]. Modules are "highly integrated developmentally and/or functionally but […] develop and function relatively independently from other [modules]" (Schlosser 2005: 145). The evolutionary advantage of modularity is based on the principle that modules are semi-independent from the rest of the organism, and can change without altering the integrity of the whole. Modules are thus the 'row material' for complexity. Evolution acts by creating, replicating and modifying modules. The modification can consist of changes in the proportions of the parts of a module, in the orientation of a module, in the connections within the parts of a module, in the articulations among parts or in the loss of some parts of a module. Only genetic changes generating any of these modifications will give feasible morphologies that will then be screened by natural selection.

Altenberg (2005) analyses the modularity of the genotype-phenotype map to arrive to similar conclusions. Higher modularity improves evolvability, roughly definable[136] as the potentiality of a genotype to produce heritable phenotypic variation, which depends not only on the current standing pool of genetic variation and covariation of the population, but also on the structure of the genotype-phenotype mapping function (Wagner & Altenberg 1996). Although apparently the phenotypic space can include many dimensions, genetic (e.g. pleiotropy) and other kinds of constraints (of the type: $f(x, y, \ldots.) = 0$ & $g(x, y,\ldots) = 0$) limit the possible variations and reduce it to a lower dimensionality. Therefore, two phenotypes apparently very far from each other in the phenotypic n-space can be really close to each other in the genotypic m-space, depending on few variables in the genotypic-phenotypic mapping.

Modularity allows defining traits without recurring to the arbitrary slicing of the phenotype, but based on functions that rests on sets of genes. Modules (sets of genes) encoding functions have variable potential to enhance evolvability: little under stabilizing selection, but high under directional selection. Suppose a module codes for two functions, one under stabilizing and the other under directional selection: the result will be low evolvability, because selective forces collide. If the modules break into two sub-modules,

[135] Thanks to Davide Vecchi for this definition of genetic module.

[136] See Pigliucci 2008a for different definitions of evolvability in the literature.

evolvability can increase dramatically. For example, the separation between regulatory and coding function in eukaryotic genomes is not a functional necessity: it could arise from such a splitting that, separating two dimensions of genetic variation, allowed the exploration of genetic spaces of each function independently. Sub-functionalization is a similar process: a gene coding different functions can duplicate, and each copy can silence some of the functions, thus allowing exploration of the space of alternative functions independently.

Brandon (2005) gives some examples of evolution potentially based on modular architecture. The earliest mammal was a tetrapod with not much differentiated forelimbs and hind limbs (functionally and developmentally). The forelimb has evolved into flippers (whales), wings (bats) and hands (humans). The forelimb has evolved without changing the whole body-plan, as whales, bats and humans have similar circulatory systems. Moreover, forelimbs have evolved independently of hind limbs (bats do not have two sets of wings, nor humans have two sets of arms). Schlosser (2005) proposes other examples of limbs as modules or units of evolution, linked to amphibian variations, and cites classical experiments of extirpation and transplantation of limb buds that do not perturb the original trunk development nor affect normal limb development in the ectopic site.

Although modularity can contribute to the evolution of organisms, especially under directional selection, it does not seem to be the result of selection. Ancel and Fontana (2000, 2005) suggest that modularity is a side-product of selection. The authors build a simple model of evolution of RNA secondary structures, in which RNA folding patterns are the possible phenotypes, each possessing a given thermodynamic stability. They show, through simulations, that selection for the RNA configuration with the lowest free-energy locks the RNA 'population' in a limited set of shapes, showing high modularity, but at the same time showing very limited variation. Evolution thus shifts from the creation of new modules to the polishing and novel combinations of existing ones.

4.4.4. Final comments on developmental accounts of variation

Empirical adaptationism claim that, from an evolutionary pint of view, the phenotype is a deterministic output of the genotype. It does not deny that many non-genetic elements can affect the process of development: it denies that these elements, and their effects, have any evolutionary significance, as they are not heritable.

The accounts of this section underline that genetic information alone is not enough to create a phenotype: development also plays an equally fundamental role. This role has several facets: developmental paths can suffer modifications and these modifications can be inherited without being coded in the genotype; developmental constraints can contribute positively to shape the phenotype; and the organization of the organism can increase its success and its ability to further evolve.

4.5. Systemic accounts

4.5.1. Natural laws

Organisms are physical entities, and it is thus obvious that they are subjected, as all other inorganic objects, to universal natural laws that apply "to all physical systems (and hence to all organisms), to all things built out of the materials in question (including organisms), and to all physical systems of the requisite complexity (including organisms)" (Maynard Smtih et al. 1985). It is a matter of debate, though, if and how these laws are also causally responsible for the development and evolution of biological beings, either directly or acting as constraints. Basic physical and chemical mechanisms, whose action in not limited to living matter, contribute to shape the organization of organisms at the level of cell aggregates and lower. Biochemical entities such as DNA sequences and proteins, and processes, such as protein biosynthesis and folding, are affected by physical, chemical and thermodynamic processes and mechanisms.[137] The cell membrane, for example, is a well-known example of biological structure that appears spontaneously when hydrophilic-hydrophobic lipids are placed in water (Fig. 4.25).

[137] Thanks to Davide Vecchi for these comments.

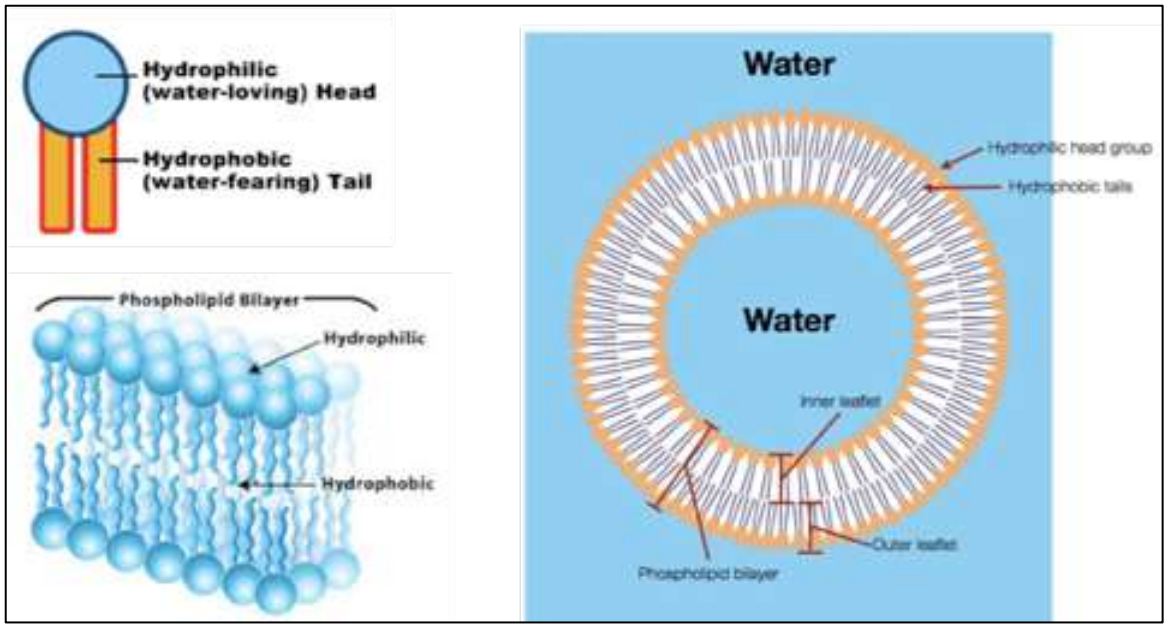


Fig. 4.25 – Cell membrane structure is the result of purely chemical and physical properties. Phospholipids constituting the membrane have a hydrophilic and a hydrophobic end. In water, they spontaneously organize themselves in bilayers and spheres in order to maximize/minimize water contact of the former/latter.

The cell's morphology is thus not determined by genetic information alone, but also by the chemical and physical environment (Newman et al. 2006, Müller 2007, see Fig. 4.26 below).

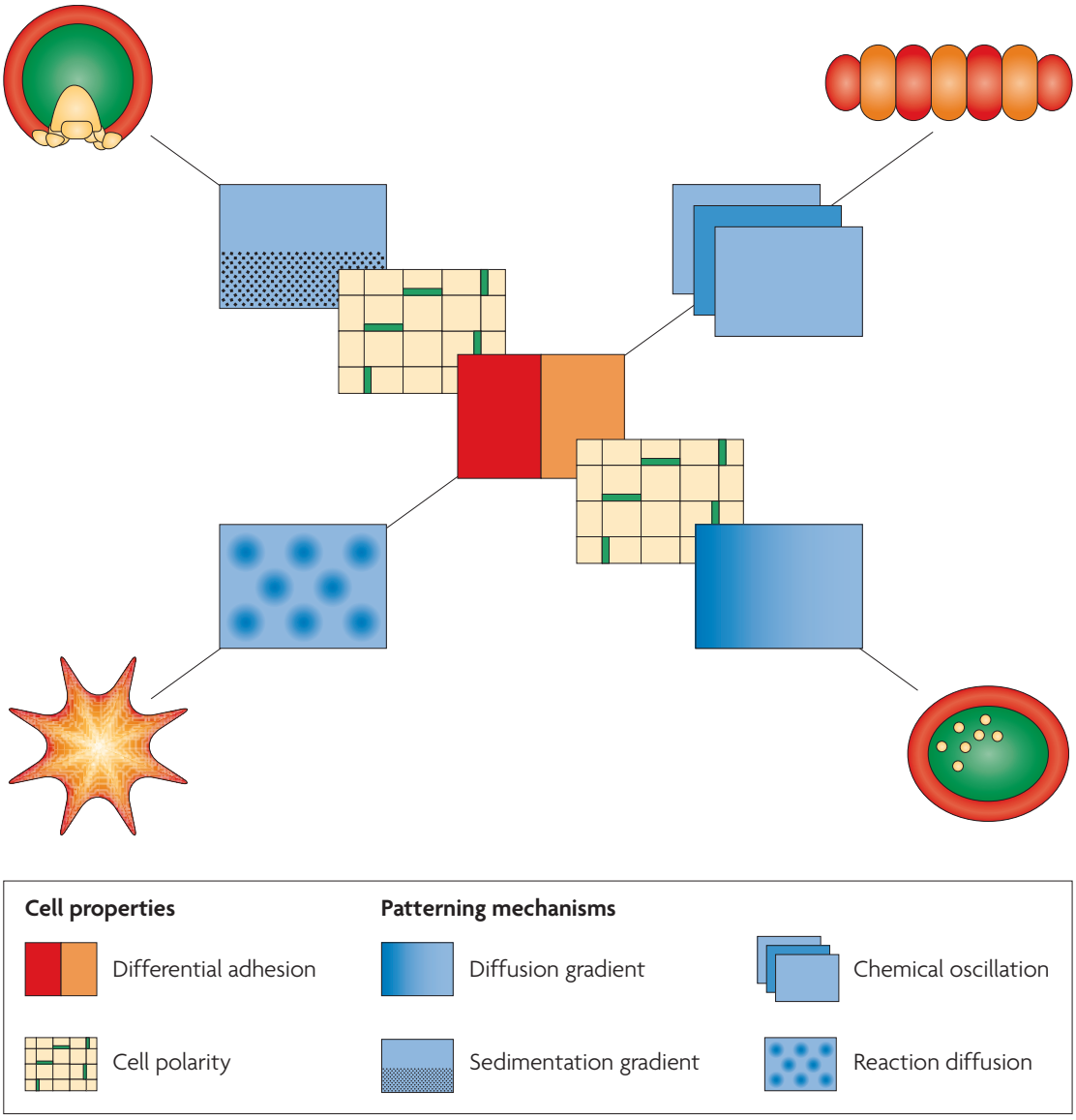


Figure 2 | **Generic forms that result from the interaction of basic cell properties with different pattern-forming mechanisms.** Differential adhesion and cell polarity (centre boxes), when modulated by different kinds of physical and chemical patterning mechanisms (blue boxes), lead to standard organizational motifs. On the upper left to lower right axis, differential adhesion properties and their polar distribution on cell surfaces lead to hollow spheres when combined with a diffusion gradient, and to invaginated spheres when combined with a sedimentation gradient. On the lower left to upper right axis, the combination of differential adhesion with a reaction-diffusion mechanism generates radially periodic structures, whereas a combination with chemical oscillation results in serially periodic structures. Early metazoan body plans represent an exploitation of such generic patterning repertoires Modified with permission from REF. 82 © (2006) UBC Press.

Fig. 4.26 – Cell properties and pattering mechanisms depend on natural laws (Newman et al. 2006)

The paucity of open coiling in logarithmically coiled shell of molluscs is another example of biomechanical laws impacting the strength of the structure and the quantity of material needed to build it (Maynard Smith et al. 1985, see Fig. 4.27).

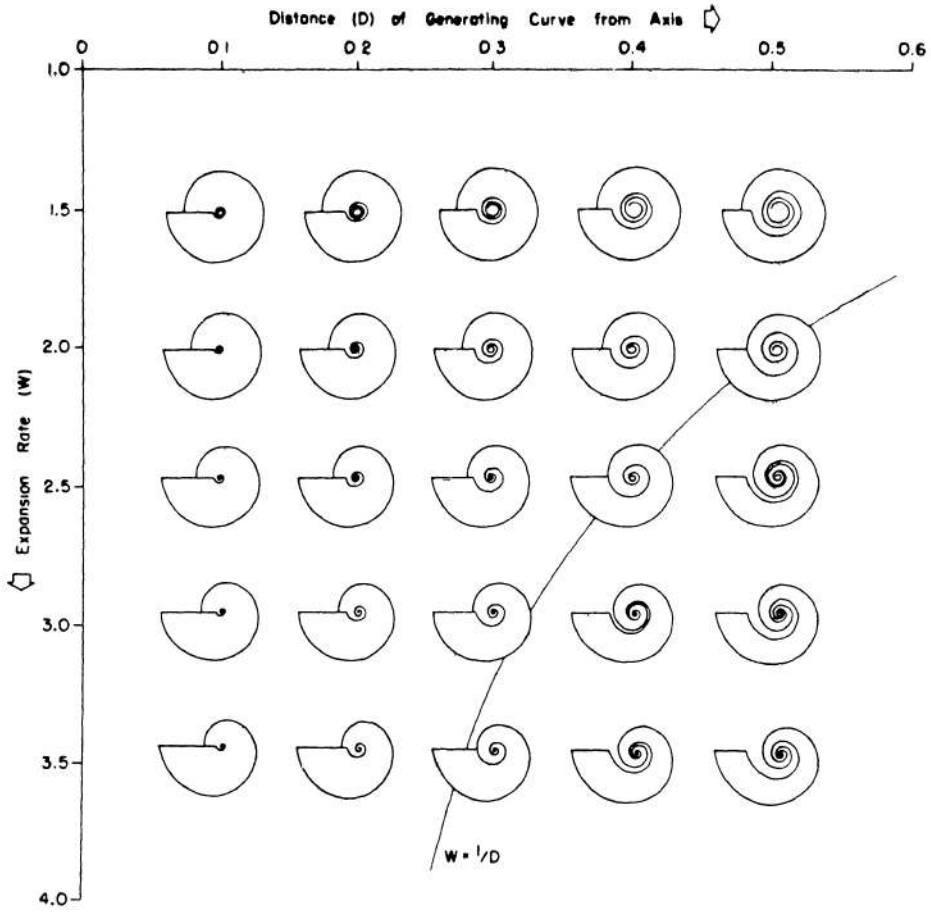


Fig. 4.27 - Distribution of shell forms: open-coil types, under the hyperbole curve, are very rare or do not exist as wild-types (from Maynard Smith et al. 1985)

D'Arcy Thompson, in his famous book *On Growth and Form* (1942), presents an extensive study of several morphological traits (shells, horns, leaf-arrangements, eggs among others) to show the importance of physical laws and constraints for explaining morphology. While not refusing selection and adaptation, he claims that their importance is small compared with the action of natural laws and constraints in the generation of phenotypes, a view clearly committed with structuralism: "Our own study of organic form […] is but a portion of that wider Science of Form which deals with the forms assumed by matter under all aspects and conditions, and […] with forms which are theoretically imaginable" (Thompson 1942: 1026). Here are some of the geometrical characteristics found in organisms that D'Arcy Thompson explains in chapter I ('On Magnitude') recurring to physical laws:

- An animal surface is proportional to the square of its linear dimension, and its volume/weight to the cube of its linear dimension. These geometrical properties have an effect on the organism because some forces act proportionally to its volume (e.g. Archimedes force), others to its mass (e.g. gravity), and others to its surface (e.g.

fluid dynamics forces). As a consequence, in order to maintain the same equilibrium, it is impossible to have a magnified or reduced version of an organism without changing the relative dimensions of its parts. For example, given that the weight of a fruit increases as the cube of its linear dimension, while the strength of the stalk only as the square, the relative dimensions of fruit and stalk cannot be maintained: because of this, and not because of an adaptive response to an environmental challenge, melons grow on the ground and apples on trees. The same principle, rather than an adaptationist account, explains why quadrupeds with large heads have necks either thick and strong (e.g. bulls) or very short (e.g. elephants) (1942: 26-27).

- Aquatic animals can have bigger dimensions than terrestrial ones, as their weight is counterpoised by Archimedes law. Given that available energy depends on the mass of muscles, and friction depends on the square of their linear dimension, the bigger the animal, the quicker it is and its speed is proportional to the square root of the linear dimension (1942: 31).
- In animals, energy loss (by radiation) increases with surface, while energy gain (by oxidation) with volume: the smaller the animal, the higher the ratio of loss to gain. This explains why smaller animals need more food compared with their weight: the metabolic activity of a whale (150.000 kg) is about 1/7 calories per kilo; that of a guinea pig (0/7 kg) of 223 calories per kilo. Therefore, a warm-blooded animal much smaller than a mouse is impossible, as it would need a quantity of food impossible for it to obtain and digest. The loss of energy is increased by low temperatures or water convection: big birds are found in the Arctic winter, but not small ones; and no small mammals exist in the sea (1942: 34-35).
- The minimal speed for a flying object in order to have a stable trajectory depends on the square of its linear dimension. This explains while large birds have difficulties in taking off from the ground, while humming-birds are capable of almost stationary flight (1942: 46).

Based on these and other considerations, D'Arcy Thompson claims that "a great deal of evolution is involved in keeping due balance between surface and mass as growth goes on". Based on the classification of living beings according to their dimension (see Fig. 4.16), he adds: "The things which fly are smaller than the things which walk and run; the flying birds are never as large as the largest mammals […] insects come down a step in the scale and more […]. Fishes, by evading gravity, increase their range of magnitude both above and

below of that of terrestrial animals […] in a certain narrow range of magnitude where gravity and surface tension become comparable forces […] a population of small plants and animals not only dwell in the surface water but are bound to the surface film itself" (1942: 66).

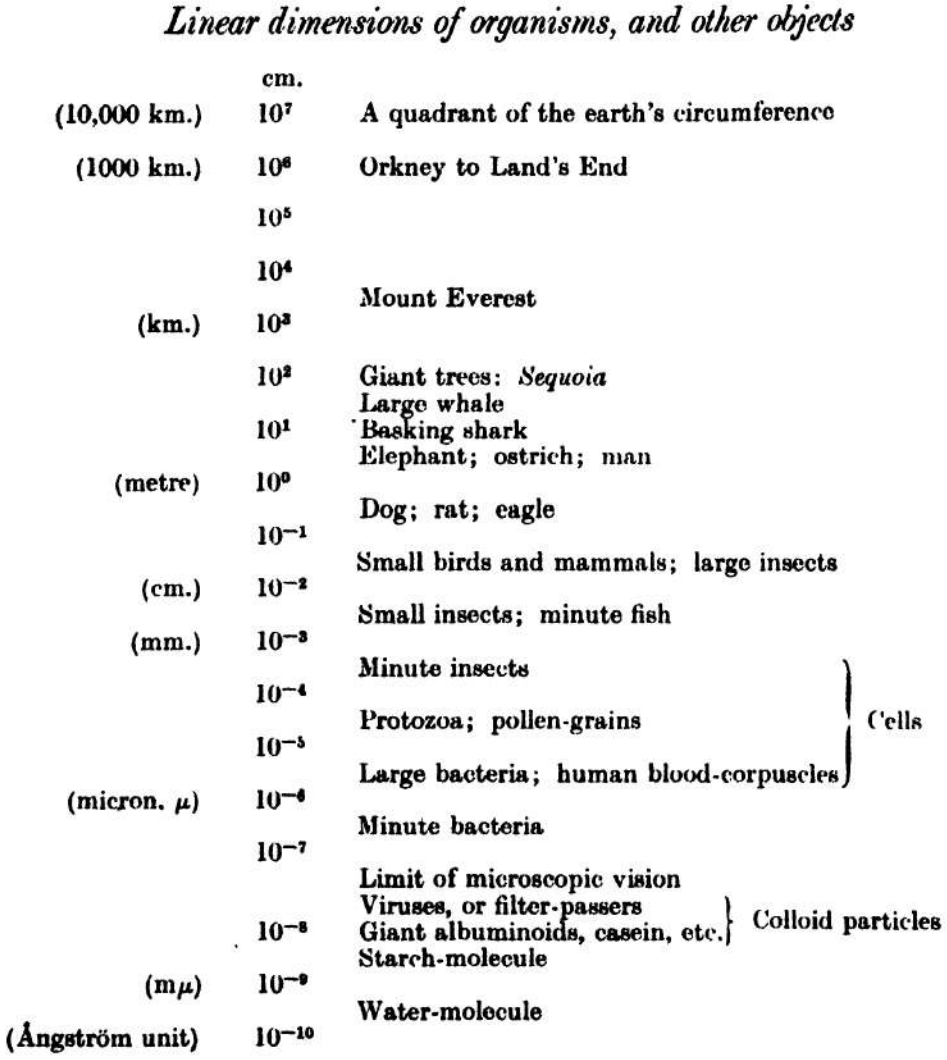


Fig. 4.28 - Living beings organized according to their dimensions (from D'Arcy Thompson 1942: 66)

In the last chapter of the book, D'Arcy Thompson proposes a method to study organic forms by first describing them through bi-dimensional Cartesian coordinates, and then transforming the resulting geometrical figures into deformed ones, corresponding to other organisms, by applying some kind of transformation. With this method, he shows that the form of an organism can be transformed into the form of another, sometimes assigned to distant genera or species. Again, his aim is to replace adaptation with natural structures and laws. He lists a series of possible transformations (see Fig. 4.29): (i) linear change in relative dimensions (e.g. to transform a circle in an ellipsis); (ii) non-linear change (e.g. to transform a

circle into an egg-shaped figure); (iii) modification of the angle between the X and Y axis (e.g. to transform a square into a rectangle); (iv) substitution of XY coordinates with radial (or polar, etc.) ones (e.g. to transform a square into a section of annulus).

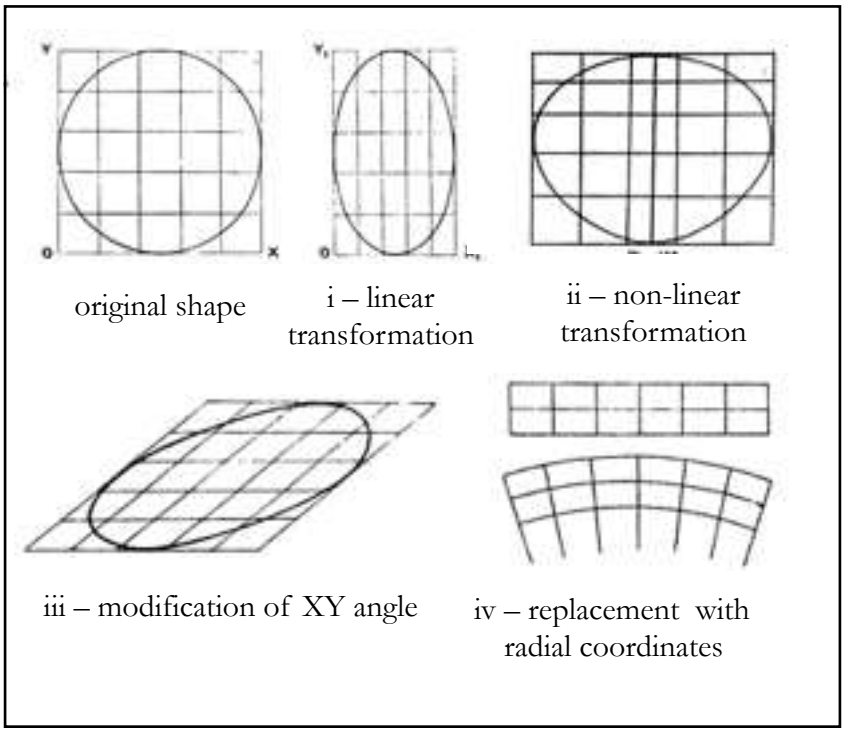


Fig. 4.29 - Possible morphological transformations (from D'Arcy Thompson 1942: 1038-1039)

He applies his method to several biological cases with startling results[138] (1942: 1052-1065). He shows (Fig. 4.30a) that the relative dimensions of the foot's bones of ox, sheep and giraffe are linked by the same proportions; that the carapace of different crabs can be reduced to a single form (Fig. 4.30b), and, in one of his most famous examples, that the morphologies of different kind of fishes can be transformed into each other (Fig. 4.30c).

[138] It is worth noting that D'Arcy Thompson does not hide cases against his own theory. He underlines, for example, the failure of his method to account for the transformation of hominids skull into modern human skull, imputing it to a lack of continuity due to divergent, rather than continuous, variation (1942: 1085). The efficacy of his method cannot be accused of resting on manipulation of data through selection of only favourable examples.

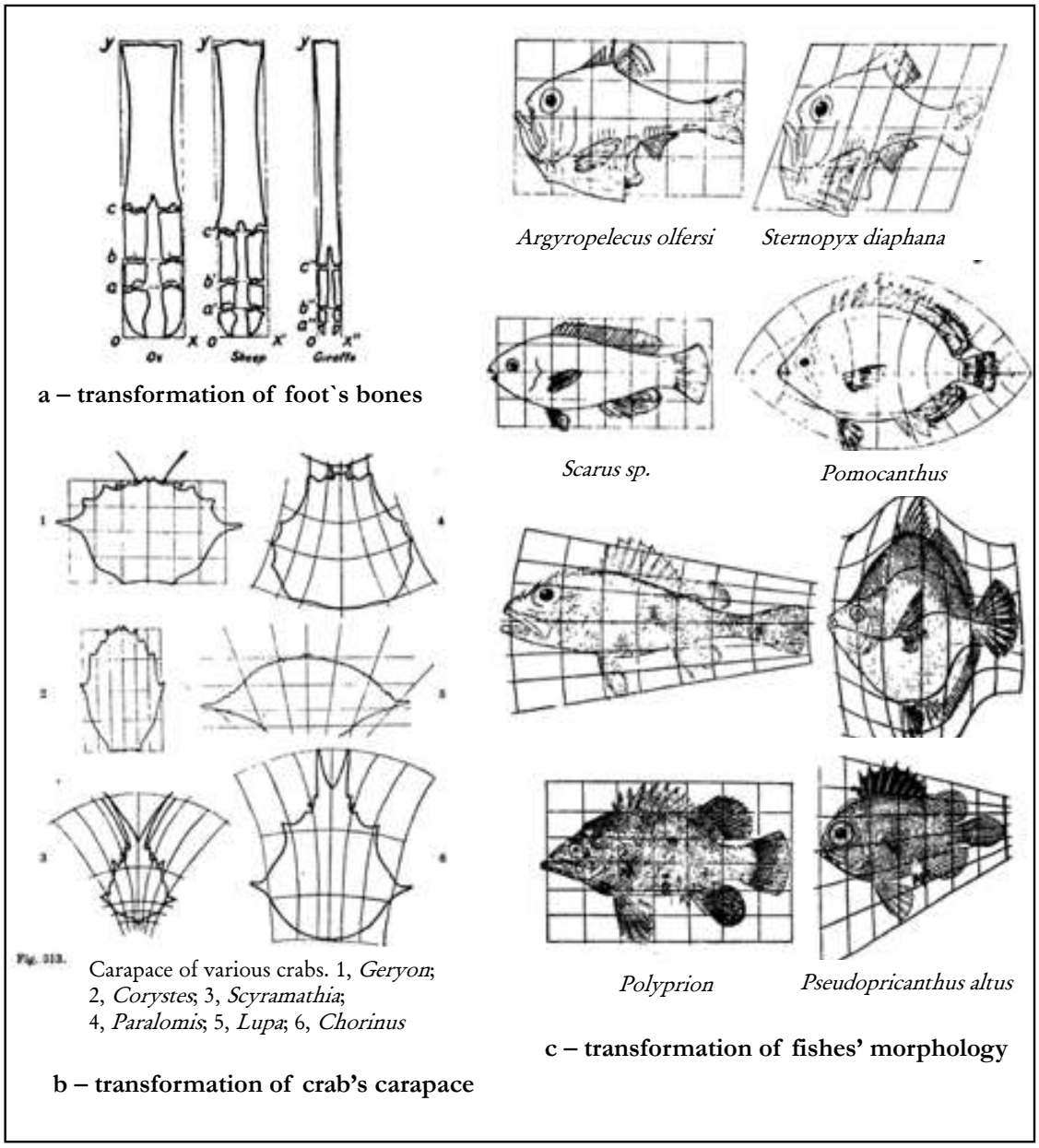


Fig. 4.30 - Example of morphological transformations (from D'Arcy Thompson 1942: 1052-1065)

The chapter closes with the claim that these geometrical analogies "weigh heavily against Darwin's conception of endless small continuous variations; they help to show that discontinuous variations are a natural thing" (1942: 1094): a thing driven by geometrical and physical laws.

D'Arcy Thompson is not alone in his search for non-selective laws of form. Lewontin claims that the slower size increase in tooth than in body in adult primates is non-adaptive (Lewontin 1980:248). Empirical trends deduced by fossil register and physiological statistics

show peculiar relationships between organic and physical variables, for example (Rasskin-Gutman & Esteve-Altava 2008):

- Ecomorphological rules of body mass to latitude relation. Bergman's rule, for example, states that individuals of the same species tend to be bigger in colder climates. Sub-species of warm-blooded vertebrates from cooler climates, for example, tend to be larger than the ones living in warmer climates (Mayr 1963). Even when the rules is inapplicable, there still exists a relationship between latitude and body mass, as in different species of *Sorex* shrews, whose body mass is negatively correlated with latitude (Ochocinska & Taylor 2003);
- Symmetry rules in the body plans of bilateral animals and serial repetition. An example of repetition is segmentation, or "the repetition along the anterior–posterior axis of a structural unit that comprises a suite of characters involving the entire body" (Davis & Patel 1999), such as in earthworms, whose segments have the same organs and muscle tissue.

Power laws of allometry relate the value of a quantitative trait with another through a non-linear relationship. For example, mass and metabolic rates are universally related by a 3/4 exponent, what is known as Kleiber's law (Wagner 2011) (Fig. 4.31).

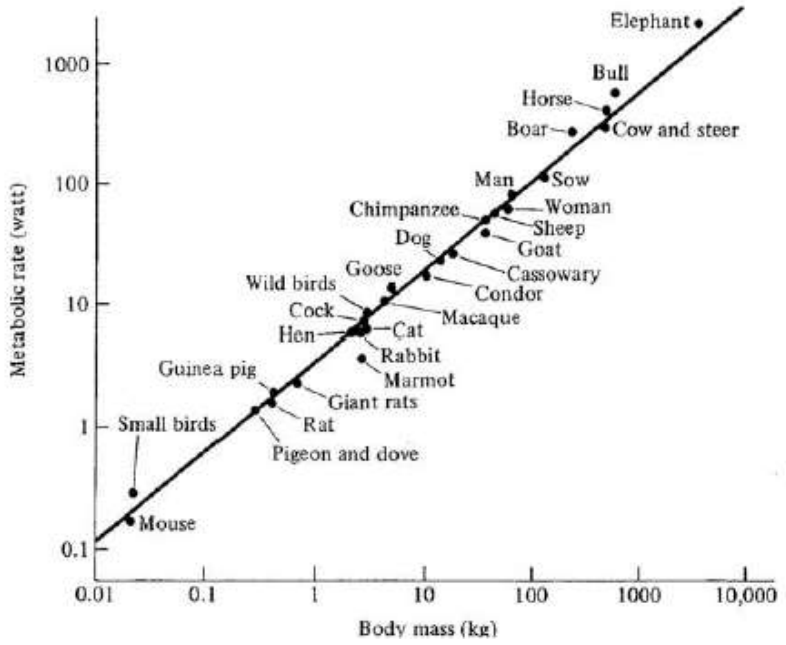


Fig. 4.31 - Metabolic rates vs. body mass – note the logarithmic scales of the axes (from Mitchell 2009: 261)

The same power-law with a 4 in the denominator appears in many natural general equations of the form $Y = Y_0(\text{body mass})^b$, for example (Fodor & Piattelli-Palmarini 2010: 204, note 16; Dawson 2014):

- Rates of body mass and heartbeat: b=-1/4 (Fig. 4.32)
- Rates of body mass and capillary radius: b=1/(4 • 3) (Fig. 4.33)
- Rates of body mass and cardiac output: b=-1/4 (Fig. 4.34)

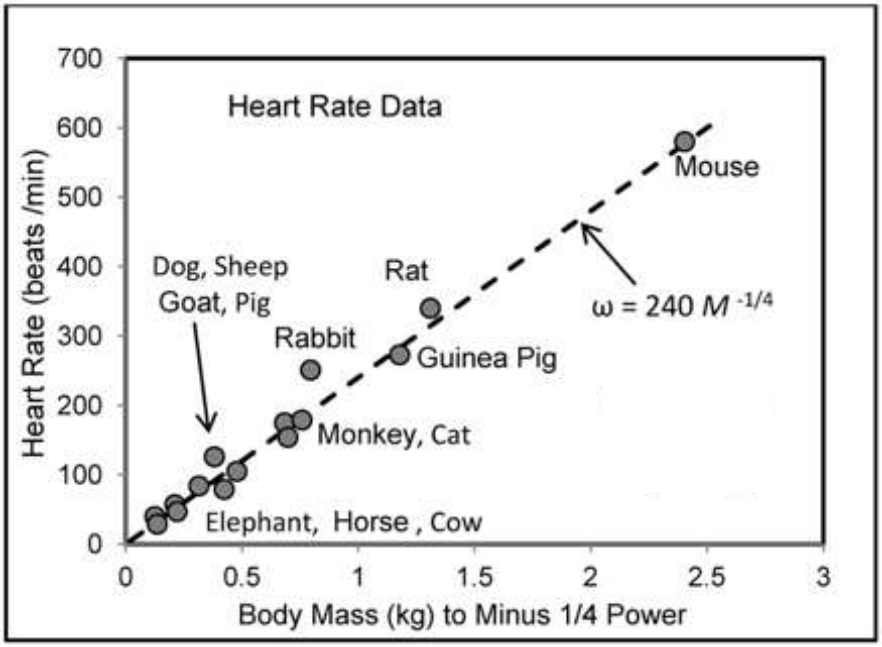


Fig. 4.32 – Relationship between body mass and heart rate (Dawson 2014).

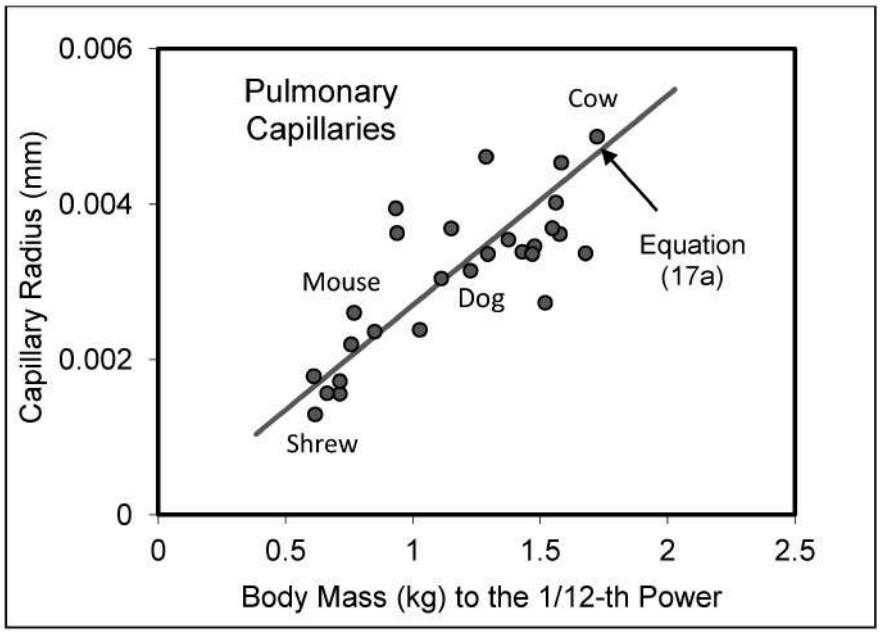


Fig. 4.33– Relationship between body mass and capillary radius (Dawson 2014).

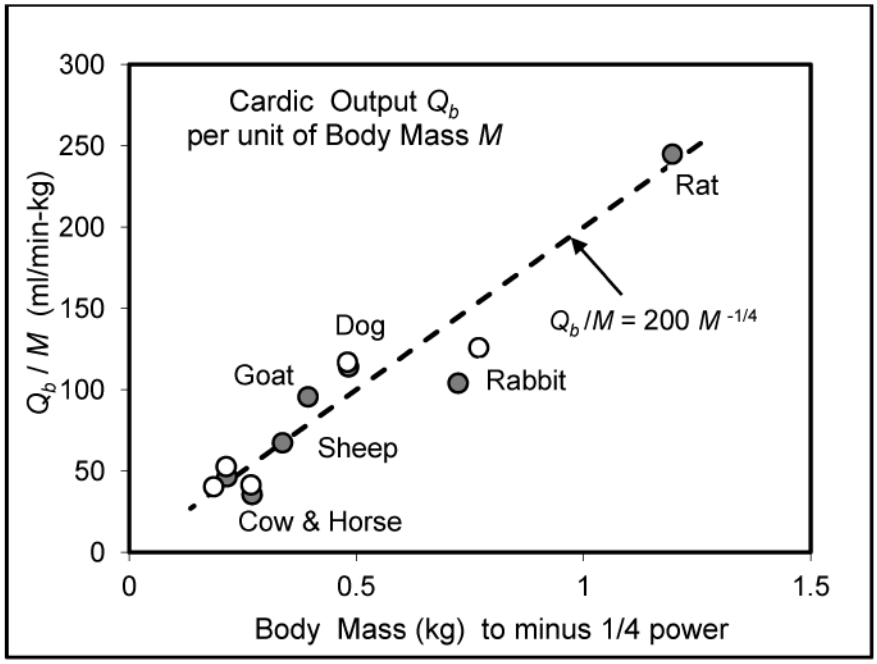


Fig. 4.34 - Relationship between body mass and cardiac output (Dawson 2014).

If we suppose that animals are spherical, and that their surface is proportional to their metabolic rate in order to free heat produced by metabolic activities and maintain a constant body temperature, their ratio of metabolic rate to volume should be 2/3, and not 3/4 (Mitchell 2009): these relationship are to be expected in a four-dimensional space. This strange result suggests that organisms, thanks to "a fractal-like architecture of the hierarchical branching vascular networks that distribute resources within organisms", behave as in a four-dimensional space, increasing their efficiency (West et al. 1999: 1677). Huxley believes that allometric relationships are by-products of underlying mechanisms and have no adaptive significance: the resulting traits are non-adaptive (Huxley 1932:214).

The idea that complex structure can rest on physics laws is behind Cherniak's idea of 'non-genomic nativism' (Cherniak et al. 1999, 2004, Cherniak 2009a, Cherniak 2009b, Cherniak & Rodriguez-Estéban 2015). Applying network optimization theory, these authors claim that the structure of nervous systems is based on the minimization of connection costs among parts, as it appears in the analysis of the macaque visual cortex (Fig. 4.35). This 'pre-formatting' hypothesis postulates "innate abstract internal structure – as opposed to an empty-organism blank-slate account" (2009b: 115): a structure that allows self-organization without contribution from the genome[139].

[139] Note that Cherniak's and D'Arcy-Thompson's critiques have different targets: the first challenges gene-centrism, while the latter challenges adaptationism.

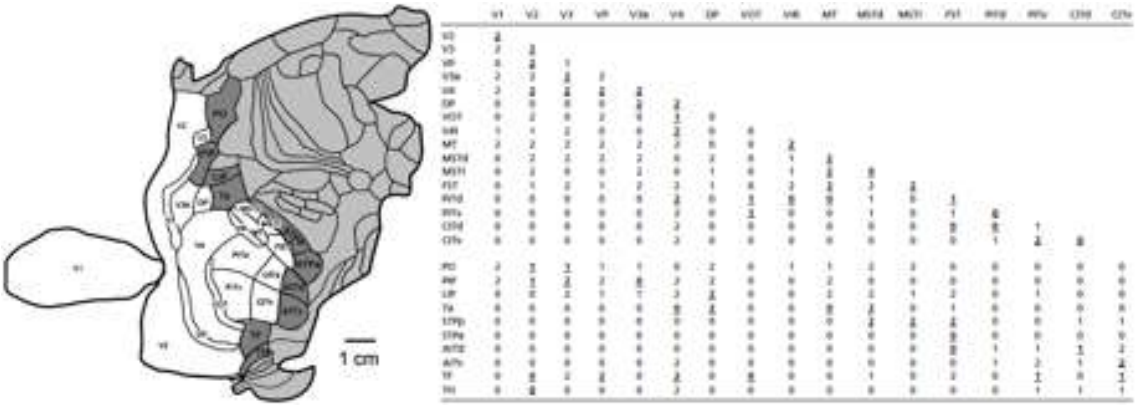


Fig. 4.35 – Analysis of the macaque visual cortex. Figure on the left is a map of the different areas of the cortex. Table on the right details which areas are connected and the correspondent connection cost. Actual layout turns out to be cost-optimal (from Cherniak et al. 2004, fig. 3 and tab. 1)

The main critiques to these ideas, and to structuralism in general, come, as expected, from adaptationist biologists, pointing out, for example, that structuralist explanations usually limit themselves to animals, ignoring the remaining branches of the tree of life: it might be difficult to analyse forms in mushrooms. Ruse (2013) suggests that D'Arcy Thompson is a hidden vitalist because, by ignoring selection, he somehow introduces direction within Nature, via physical laws. One might of course as well blame adaptationism for the same reason, namely the substitution of 'physical laws' with 'selective laws' and of 'Nature' with 'Environment' (capital letters intentional). Adaptationism is not the only area of biology critical of structuralism. Held (2014), an evo-devo biologist, defends that genetic 'tool-kits', and not physical laws, define the developmental routes and constraints that can explain morphology.

4.5.2. Evolution as a thermodynamic process

The order of organic beings could be explained as a thermodynamic consequence of the heat flow from the sun, in the same way as natural structures like snowflakes are. Schrödinger (1944) already pointed out that living systems, by showing high levels of order created from disorder, seem to contradict the second law of thermodynamics. He suggests that the solution to this apparent contradiction to physical laws resides in considering that life exists in far-from-equilibrium systems: such systems take energy from an external environment and use it to reduce their inner entropy, while increasing more than proportionally the external one. The second law is thus preserved.

Prigogine (1984, 1997) studies such systems at far-from-equilibrium conditions, open to material and energy flow, where they can still reach some steady state, although such state

cannot be described, as it happens at equilibrium, simply by the value of its potential[140]. Once the system departs from equilibrium beyond some threshold, it shows a new, and often puzzling, behaviour. A typical example is the transition from laminar to turbulent motion in a fluid[141]. While, at first sight, turbulent motion might appear chaotic at the macroscopic level, it actually is highly organized at the microscopic scale. Instability in a Bénard cell, a horizontal liquid layer between whose surfaces a temperature gradient is applied, is another classical example. Heat is transferred between layers by conduction, till the gradient reaches a threshold: then, convection appears and heat transfer and entropy production increase. Convection consists of the coherent movement of millions of molecules forming hexagonal convection cells of a characteristic size (Fig. 4.36).

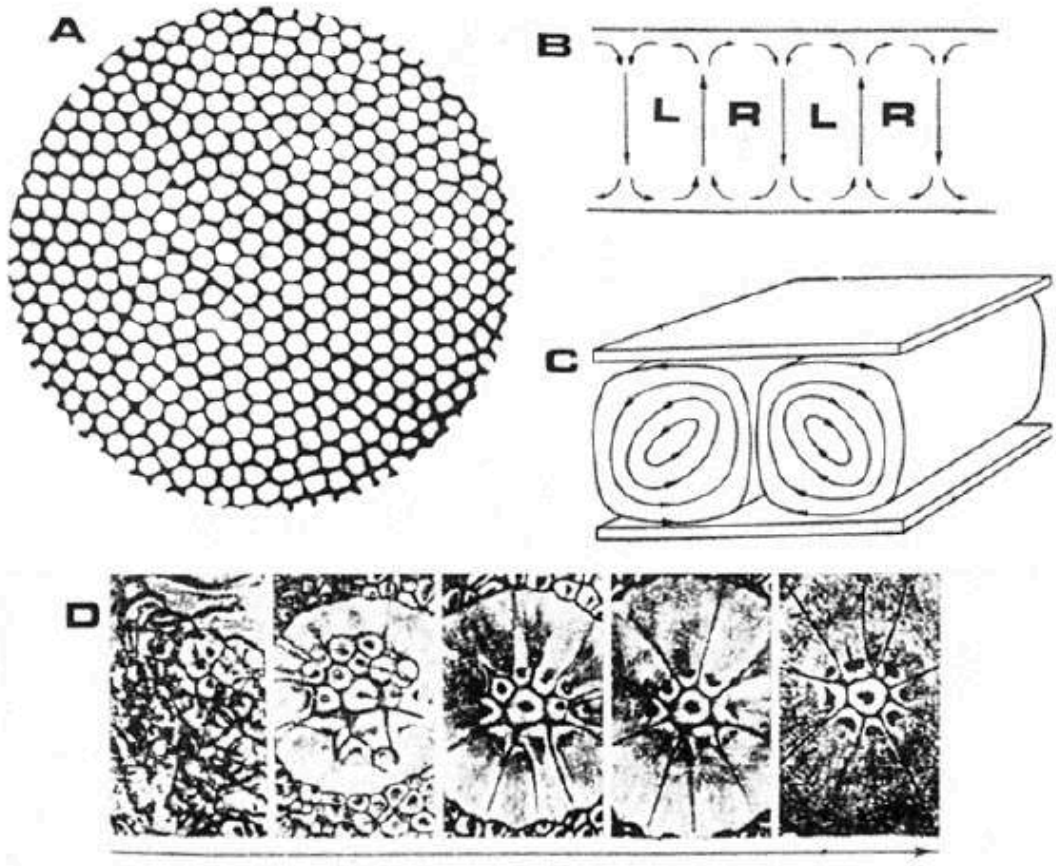


Fig. 4.36 - Bénard cells. A: general view from the top. B, C: schemes of coherent convection streams. D: evolution of Bénard cells patterns under constant heating (from left to right) (Beloussov 2012)

[140] In classical physics, the evolution of some systems can be described by a scalar value that only depends on the initial and final state of the system. For example, the potential energy change of a mass in a gravitational field only depends on its initial and final positions.

[141] A liquid flow is laminar when it is composed by, with no disruption between them; it is turbulent when its pressure and velocity change chaotically. The transition between laminar and turbulent flow usually happens when some threshold velocity is reached.

These dissipative structures consist of a highly improbable (in Boltzmann's terms) supramolecular organizations. Contrary to the classical thermodynamic paradigm, heat transfer (dissipation) is not a source of chaos, but of order (structure): *irreversible processes have a constructive role in nature*. Entropy increase (e.g. heat flow) can lead to order increase: in what is known as 'thermal diffusion', concentration of hydrogen and nitrogen molecules initially homogeneously mixed in a box will show a gradient if one side of the box is heated (Prigognine 1997: 26-27). Likewise, life and organisms are improbable structures at equilibrium conditions, but not at far-from-equilibrium ones, where self-organization can arise: "nonequilibrium systems may evolve spontaneously to a state of *increased complexity*" (op. cit.: 64). As a matter of fact, Prigognine believes that this is the only way complexity can appear and be maintained. An example of such phenomena is proposed by Turing (1952), who suggests a morphogenic mechanism based on the triggering of a disequilibrium in a system of chemical substances initially at equilibrium, by which "the genes of a zygote may determine the anatomical structure of the resulting organism" through known physical laws. Sometimes, Prigognine's view has mystical echoes, as when he claims that "[m]atter at equilibrium is 'blind', but far-from-equilibrium it begins to 'see'" (op. cit.: 67).

Brooks and Wiley (1986: 70-74) believe that thermodynamics and biological evolution are different manifestation of the same principle: entropy production. The second law of thermodynamics is the natural law of history. All biological systems (organisms or species) appear highly ordered in two ways: they are composed of integrated parts and they relate to each other in hierarchical structures. This orderliness is dynamic: reproduction, ontogeny and death, all irreversible processes, change the appearance of organisms. Therefore, irreversible processes are behind the order and hierarchy of nature.

Based on such considerations, the authors propose a theory of evolution applying principles of thermodynamics and information theory in order to make sense of the idea that organisms and species are self-organizing dissipative structures that (i) show *some degree of cohesion*, and (ii) transmit *information* by reproduction in an imperfect way (for a general and short introduction to Brooks and Wiley's theory see Collier 1986). Evolution results from the entropic behaviour of information and cohesion of such structures, constrained by their initial historical and developmental conditions. Natural selection is a boundary condition that eliminates variants unfit for survival, but without any creative power. Organisms and species are identified by their information content, so imperfect reproduction (which entails changes in biological entropy) leads to evolutionary changes in two senses: increase in species organization and decrease in species cohesiveness. These variations affect two kinds of

biological entropy: the entropy of information measures the stability of the organism, its ability to withstand random fluctuations; the entropy of cohesion measures disorder due to imperfect mixing of parts (e.g. imperfect sharing of genetic material among members of a population). Changes in these two kids of biological entropy cause evolutionary events in four basic elements of living systems: metabolism and homeostasis, development, reproduction and speciation. Changes are fuelled by imperfect transmission of four kinds of biological information, defined along two axes:

- stored and potential information. The former defines the phenotype and its functions, and is consistent among all individuals; the latter is not expressed (e.g. recessive alleles) and is different in each individual;
- regulatory (or canalized) and structural (or non-canalized) information. The former leads development and determines traits; the latter determines the basic elements of the organism structure (proteins, enzymes, etc.).

Information transmission can be imperfect and result in:

- potential information becoming stored information;
- new information of any of the four possible kinds due to genetic changes.

Schneider and Kay (1992) take these considerations further by applying to organisms the 'Law of Stable Equilibrium' (Hatsopoulos & Keenan 1965), or 'Unified Principle of Thermodynamics' (Kestin 1966), that defines the direction and end state of any process arising from the elimination of internal constraints in an isolated system. The authors extend the law to non-equilibrium systems, to make it applicable to living beings, and then restate the second law of thermodynamics as follows: a system moved away from equilibrium by an external applied gradient will utilize all avenues available to counter it. According to them, this new formulation avoids recurring to entropy, which is only defined at equilibrium state. They believe that the farther away from equilibrium a system is moved, the more complex the mechanism triggered to resist any further moving. A simple example is the flow of water between two 1,5 litre bottles connected end-to-end, in an hourglass shape. The water gradient between the two bottles is eliminated by water flow in about six minutes; however, if a rotational perturbation is applied, a vortex forms (a highly-organized structure similar to a tornado) and the gradient is dissipated[142] in eleven seconds. Complex systems emerge as

[142] Dissipation of energy means the flow of energy through a system due to an energy gradient. The dissipation might or might not destroy the gradient. If it does, we also have energy degradation: the energy loses the ability to create a gradient.

ways to dissipate gradients through non-equilibrium systems: for this reason, they are called dissipative systems. The authors suggest that life is just a complex way to dissipate the solar energy gradient: living systems are 'sophisticated mini-tornados'. Ecosystems formation and development are examples of extremely complex dissipative structures that cause solar energy gradient to drop in a more efficient way then it would otherwise do.

Vogel (1988) proposes another model based on thermodynamic considerations, which is peculiar because it applies an optimization approach in which entropy production is the maximand, and claims that his model explains, among other phenomena, diversity, extinction, gradualism, increasing biological complexity, neutralism, punctuated equilibrium, r-k selection, and sex: possibly, too long a list.

4.5.3. The laws of complexity

Spontaneous order exists in nature, from the static structure of crystals to the dynamic flows creating tornados[143]. That such spontaneous order should be limited to the non-biological world seems hard to sustain: the cellular membrane is a known example of order rising from physio-chemical properties (Fig. 4.25). D' Arcy Thompson lists many examples of biological forms similar to purely inorganic ones (§4.5.1). As a matter of fact, order is widespread at all levels in the biological world: organisms and their parts are extremely complex but ordered systems. Genes regulating each other's expression, cells and molecules contributing to the immune response, neurons in the brain connecting to each other, insects cooperating through sophisticated kinds of behaviours for the benefit of the colonies are all large networks of elements showing common features: interaction without central control and according to simple rules; complex collective behaviour; sophisticated processing of information coming from external and internal sources; and a starling capacity to adapt through learning and to evolve (Mitchell 2009).

Adaptationism sees selection not only behind adaptedness and variety, but also behind biological complexity (Brown 1991), and considers that complexity cannot arise spontaneously. Others argue that complexity is equally important to explain life's variety: it is thanks to the complex structure of genetic networks (Wagner 2001), for example, that phenotypes otherwise out of reach through mutations become accessible. Complexity constraints, defined as non-random transitions between states of a complex system, may

[143] The definition of 'spontaneous' for such order could of course be challenged by a mechanistic analysis of the phenomena – the same as mutations are not considered to be 'spontaneous' anymore since their complex mechanistic bases were uncovered (thanks to Davide Vecchi for this comment).

have evolutionary implications (Maynard Smith et al. 1985). Rasskin-Gutman and Esteve-Altava (2009) claim that not all hierarchical levels of increasing complexity can be reduced to the sum of the constituent lower level parts. We find breaking points at the level of genome, cell, organism and species: these levels cannot be produced just by assembling lower level parts. Sub-cellular parts (e.g. the membrane) can arise from self-organization of biomolecules, but a cell cannot arise from the assembling of sub-cellular parts: it always arises from another cell, which represents a pre-existing template[144]. In the same way, individuals arise from individuals, not from some kind of self-organization of pre-existing cells. Self-organizing processes explain their generation. An example of solution to adaptive problems that could be linked to complex system laws (together with physical principles) rather than to selection mechanisms is the canids feeding strategy, that evolved from the Eocene exploring different alternatives (from hyper-carnivore to omnivore). It looks like there are three optimal strategies, each with an optimal body mass, that appear several times in canids phylogeny (Erwin 2015).

Kauffman (1991,1993, 2000) goes one step further and makes complexity the main target of selection. Kauffman challenges the hypothesis that natural selection is the only explanation of biological order and identifies the source of novelties in the universal properties and laws governing complex systems: selection just acts on these novelties once they have appeared. Not all complex systems, however, show the same ability to maintain or adapt their structure against changing external conditions. The role of selection, according to Kauffman, is thus not so much to favour the fittest, but to pick up, like Maxwell's demon, from among all potentially self-generated ordered systems, the ones maximizing the properties of survival *and evolvability*. These are the systems poised on the boundary between completely ordered and completely chaotic behaviour.

Complex systems show peculiar chaotic features, namely great sensitivity to initial conditions (the 'butterfly effect'); but also unexpected tendency to order, or anti-chaos. Many of their features (stabilities, discontinuities, non-random probabilities of transitions between states and, in general, their dynamical behaviour) can be predicted without detailed information about the system's elements: a description in terms of the system state variables suffices. Some systems[145], after a walk in the state space, arrive at a point and remain there: such points are called 'attractors' (Kitano 2004, Mitchell 2009, Kauffman 2000). The logical

[144] Template in the sense of set of information used to create a new individual of the same sort.

[145] But not all: the frictionless pendulum is an example of a system not reaching an attractor.

state space is thus scattered with attractors, which constitute a set much smaller than the total logical space, and represent the totality of the possible long-term behaviours of the system. The system, apart from being described by state variables evolving over time, usually has some parameters hold fixed (e.g. temperature in a isothermal system). The attractor is stable if, following a change in some of the parameters of the system, this moves back to it; it is unstable otherwise. It is possible to build, parallel to the state space, a parameter space. Each point of this space identifies a set of attractors in the corresponding state space. By slightly changing one parameter, the set of attractors (also called 'basin of attractors') either remains the same or changes smoothly. Some small parameter changes, however, cause a sudden change in the basin: these changes are called bifurcations, and divide the parameter space into disjoint areas. Most changes in parameters do not lead the system to cross such bifurcation boundaries: the system returns to the same attractor. The systems having this ability are 'structurally stable': their parameter space has bifurcation surfaces much smaller than the volumes they contain. Attractors show a certain degree of intrinsic complexity, and the movements from one to another result in a global 'jump' in complexity, not reducible to the sum of many partial increases.

Attractors were initially studied in 'logistic' models, mathematical formalizations used to analyse non-lineal population growth in the presence of overcrowding. An example of a non-lineal growth phenomenon concerns two groups of twenty rabbits, one located in a single island, and the other distributed between two islands having ten rabbits in each: after one generation, the number of rabbits is not the same in the two cases, even if birth and death rates are (Mitchell 2009).

The general form of the logistic equation is:

$$Xt = R\,(1-Xt)\,Xt$$

where Xt is the value of the variable X at time t and R is a constant. The behaviour of X is generally chaotic, as it heavily depends on the initial conditions X_0. However, it turns out that, regardless of X_0, if $R < 3$, X will converge to a fixed value (*a stable attractor*); if $3 < R < \approx 3.33949$, X 'jumps' between two fixed values (*periodical attractors*, with period = 2); and, by slightly increasing R up to 3.569946, the number of these attractors steadily doubles; for higher R, there are no attractors and the behaviour becomes chaotic (Fig. 4.21) (May 1976).

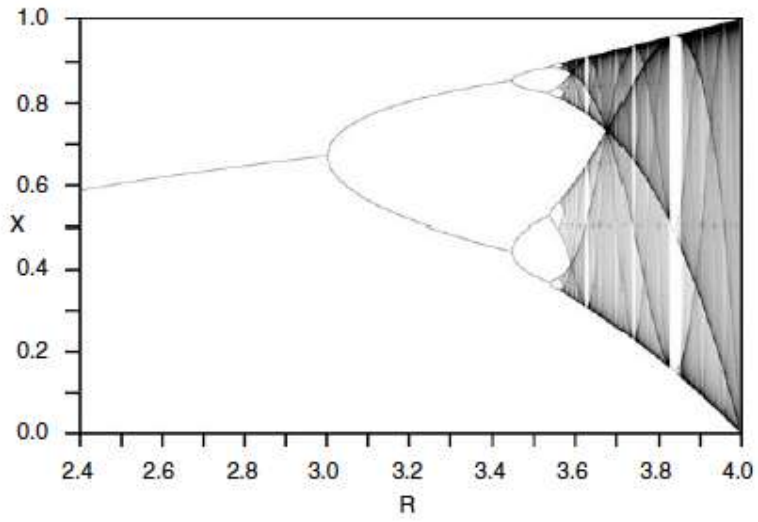


Fig. 4.37 - Value (and number) of attractor(s) for different values of R. Duplication in the number of attractors is clearly visible until the edge of the chaotic behaviour (from Mitchell 2009: 34)

A genome as a set of binary variables (each representing a gene) variously connected to one another through Boolean functions is an example of attractors in biological systems: the value of each gene at time t depends on the value at time (t-1) of all other genes connected to it (Kauffman 1993). The set of values of all genes at a given time t is a *state of the system*. If there are N genes, whose status (active or inactive) depends on K other genes[146], the possible combinations of the inputs of a gene is 2^K. Although it might be a very big figure, it is a finite one: the system must therefore, sooner or later, re-enter an already encountered state. From there, given that the Boolean functions are deterministic, it will proceed through the same cycle of states. Such cycles are called *dynamical attractors* of the network: once the system enters one, it remains there. Attractors have thus strong biological implications, since they represent most of what the system does: the states not included in the cycle of the attractor never happen. Each system has at least one attractor, but they usually have several, and perturbations can move the system from one to another. Some systems recover from any perturbation, others are very sensible and move to a new cycle with any single one. As a general rule, it turns out that the bigger the size of a cycle, the more stable is the attractor.
There are two kinds of perturbations that, affecting a system, might change its trajectory:

- Minimal perturbations: the *value* of some elements is switched (e.g. from 0 to 1).
- Structural perturbation: some *connections* among elements, or some Boolean *functions* describing some connections, are modified (e.g. an 'AND' function is changed to an 'OR' function).

[146] K can be thought of as a measure of the epistatic interaction among genes. For K=0, there are no such interactions and each gene has a pure additive effect on fitness.

Kauffman studies the case of K=N[147]: these systems show maximal disorder and great sensitivity to initial conditions, and the average length of their cycles is more or less the square root of all their possible states (so, it is $2^{K/2}$). With N=200, a complete cycle would take an incredible long time to be completed even if each transition lasted microseconds. However, the number of attractors is just N/e[148], that is, about 74; and few of these include two thirds of all states, being therefore very stable. If K=1, the system becomes modular, the simple sum of isolated subsystems that cannot influence each other.

This chaotic behaviour suddenly changes to an ordered one if K=2: a phase transition occurs. The number and the length of cycles drop to $N^{1/2}$. A system with 100.000 elements and two inputs per element has just about 370 stable cycles, each completed in 370 units of time. The system is thus confined in a very small sub-set of all its possible states. Sensitivity to initial conditions drops dramatically, and each attractor is stable to most perturbations: only in about 10/15% of cases the perturbation moves the system to a different attractor.

Kauffman applies his model to the human genome with startling success. He supposes that a genome contains 100.000 genes and calculates that the number of corresponding attractors is about 370. If each attractor is a cell type, this figure is compatible with the 254 human cell types (the difference might be due to the lack of function of some DNA sequences, or to the postulation of too many genes). The same calculations for other organisms give similarly good predictions (Fig. 4.38).

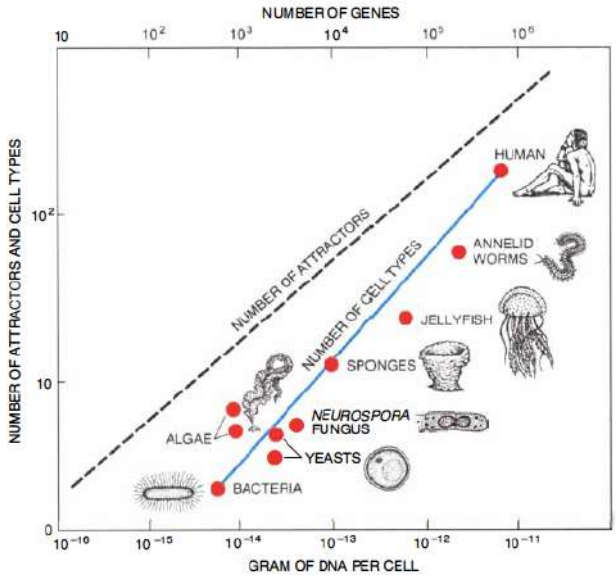


Fig. 4.38 – Relationship between calculated number of attractors and actual number of cell types vs. number of genes for different organisms (Kauffman 1991).

[147] This is the case in which there are N genes and the status of any of them depends on N other genes.

[148] e = 2.718… is the base of natural logarithm

The homeostatic stability of each cell type represents the stability of an attractor. The number of cell types of an organism represents the attractors accessible by changing the state of genes (active/inactive). The interpretation of cell types as attractors also explains why most potentially deleterious mutations end up having limited effects, and the ability of mutations to create new cell types. The hypothesis of K=2 is based on the fact that only a few different cell types (typically one or two) can develop from the same ancestral one (Kauffman 1983).

Some other examples of biological systems and their attractors are (Kauffman 1993: 191):

- immune networks: different immune states
- neural networks: alternative memories by which the network 'knows' its world
- cardiac systems: normal and abnormal rhythms
- ecosystems: alternative stable patterns of species abundances

According to Kauffman, the order of these networks is spontaneous, and "lie[s] at hand for further employment by natural selection" (op. cit.: 192). But, why does this behaviour arise? The system develops a 'frozen core' of elements whose state is fixed (0 or 1) and that creates kinds of 'percolating' walls. These walls separate the system into 'functionally isolated islands', areas that cannot communicate among themselves, as each is unaffected by perturbations coming from any other. The walls appear, for example, thanks to a peculiar feature of some Boolean functions ('canalizing Boolean functions'): once one of their inputs is defined, they are insensitive to the value of the second input. Almost all regulated genes follow such functions.

Kauffman compares an ordered system, possessing a big frozen core, to the solid state, and a chaotic one, made up of many unfrozen islands, to the gaseous state: he claims that living system are placed in between these two extreme states. If the system is too ordered (e.g. K=1), the frozen core is big and the small unfrozen areas are functionally isolated. No complex flow of information among areas can occur. On the other hand, if the system is chaotic (e.g. K>2), the frozen core is very small and ineffective, and any change in an area propagates to most of the other areas. In the boundary between order (K=1) and chaos (K>2), most perturbations (mutations) have small consequences, while few ones cause big avalanches of change: this allows organisms to adapt slowly or quickly to a changing environment. Kauffman's hypothesis is that the most complex, integrated and evolvable behaviours occur in such boundary region. To show this, he studies the structure and

characteristics of fitness landscapes in the extreme cases of K=0 (no epistatic interactions) and K=N-1 (each gene influences and is influenced by the behaviour of all other ones). In the case of no epistatic interactions (K=0), the landscape has a smooth surface, with a single peak accessible form any initial genotype through a path of single-mutants steps, and whose length[149] increases linearly with N. Fitness changes between close points of the landscape are minimal, and the probability of an improvement in fitness by mutation is almost null ('error catastrophe').

On the opposite case of maximum epistatic interactions (K=N-1), the landscape becomes rugged and shows a great number of peaks. Genotypes can access only a small fraction of these peaks, through very short adaptive walks, whose length increases as the logarithmic function of N. The height of peaks shrinks towards the average fitness of the set of genotypes: populations get trapped within a peak and can only escape by long jumps, whose effectiveness shrinks by one half after each successful jump ('complexity catastrophe'). Given that selection, no matter how strong, can only push fitness until the height of the highest peak, many conflicting constraints limit the power of adaptive selection and make its action finally void of any evolutionary meaning.

Systems at the edge between order and chaos are able to escape both catastrophes: by creating a functional modularity thanks to their frozen core (and not an architectural modularity[150] like systems with K<1), they can perform complex tasks, avoid propagation of errors (e.g. lethal mutations) and jump to new attractors when required by changing environmental conditions. If it could be proved that

> "[…] genomic systems of plants to animals, separated for the past 600 million years, are all poised near the edge of chaos, then we would virtually have to conclude that selection has achieved such a poised state. If true, this finding would provide striking evidence that parallel systems with nearly melted frozen components possess the construction requirements which permit complex systems to adapt. Hence such features might be quasi-universals in complex adaptive systems." (Kauffman 1993: 227).

---

[149] The length of a path is the number of single mutations composing it.

[150] Architectural modularity happens when genes are largely independent from one another (k<1): the network is composed by structurally independent islands, the architectural modules. What happens in one module (e.g. a mutation) does not affect other modules.
Functional modularity happens when genes do influence one another, but their Boolean functions are such that functional isolated islands arise, the functional modules. Genes still influence other genes within the module, but not genes of other modules. (Kauffman 1993: 226)

The intermediate case of K=2 and two alleles is an especially interesting example of edge system. The highest peaks in the landscape are close to each another: the space between two local optima is therefore a good place to look for a higher peak. This makes genetic recombination a good search strategy in such landscapes (Kauffman 1993: 60-63).

Kauffman shows how the most efficient adaptive research strategy depends on the structure of the landscape. If the complexity of the landscape is low (N and K low, and K<<N), the best strategy consists of some individuals making long jumps[151] across the space and, in case a better peak is found, the rest of the population following. As complexity increases, however, this strategy turns out to be deleterious, due to the consequences of the 'complexity catastrophe'. The result is that evolution, according to this model, follows three different steps:

- In the early stage, long jumps are more successful than short jumps. If the initial entity is poorly adapted, and its surroundings are just slightly fitter, the probability of finding substantially better fitness with a long jump is high. However, the waiting time to find a fitter place by long jumps doubles each time.
- After a few long jumps, it is therefore more probable that movements will be local and improvements will be slight, until a local optimum is reached.
- To refuel the evolutionary process, a new successful long jump is needed. It turns out that "the waiting time doubles after each fitter variant is found" (Kauffman 19913: 70), so the new successful long jump will occur in very long timescales.

Kauffman applies these results to explain the evolutionary phenomena of radiation and ultimate stasis, the stability that dominates the fossil record (Gould and Eldredge 1993). At early stages, new phyla appear by long jumps, while in later stages evolution occurs within phyla through small jumps that slightly increase fitness, until species freeze in stasis once a local fitness peak is reached. This mechanism would also explain the peculiar differences between the Cambrian explosion[152], when species appeared following a top-down scheme (first phyla, then classes, orders, etc.), and the rapid increase in species following the Permian

[151] A long jump is, for example, a recombination of two genes giving a new protein: e.g. the left half of the gene coding for protein A 1 is recombined with the right half of the gene coding for protein A2, forming a recombined gene (Kauffman 1993: 113). A short jump is, for example, a single nucleotide change in a single gene.

[152] The 'Cambrian explosion' refers to the emergence, around 550 million years ago, of the major animal body forms and phyla.

extinction[153], when the pattern of appearance was bottom-up (with no novelty at class and phyla level). While during the Cambrian explosion new phyla were easily created by long jumps in a still empty landscape, the surviving species after the Permian extinction faced a frozen early ontogeny, and evolved by small jumps in local areas of the landscape (Kauffman 1993: 76-78). Applying the same interpretation, Kauffman explains Von Baer's laws that early embryos of various lineages are more similar than late embryos: early development is caused by distant jumps, whose frequency and probability decreases quickly, while mutants affecting late development accumulate easier.

To summarise, Kauffman's hypothesis about the relationship between self-organizing complex systems and adaptation is two-folded:

1. Much of the *biological order is due to self-organization, not to natural selection*: natural selection just privileges some of the ordered systems already present;
2. Through filtering, natural selection favours those ordered systems more capable of evolving. These are the complex systems placed near the boundary between order and chaos.

This is an interesting view, and one that goes beyond the pure 'negative view' of natural selection: it shares with the negative view the idea that natural selection does not 'create' order, just acts on it; but it also recognizes to natural selection a positive role: the ability of driving evolvability, that is, of favouring these systems that show more potentiality to further evolve. Kauffman believes, however, that the more complex the system, the less able selection becomes to alter its features. "Much of the order in organisms may be spontaneous" (Kauffman 1993: 30): additional order does not appear thanks to, but *despite* selection.

4.5.4. Zero Force Evolutionary Law

McShea and Brandon (2010) postulate that a tendency towards increasing complexity, that they name Zero Force Evolutionary Law (ZFEL), underlies all phenomena of reproduction with heritable variation. The unexpected fact in the history of life is therefore stasis, and not evolutionary change: two evolving entities, when no constraints act, tend to differentiate and diverge from each other. The ZFEL describes "what happens when nothing happens" (McShea et al. 2019: 1). If they are populations, the divergence result in diversity increase

[153] About 200 million years ago, about 96 percent of all species became extinct.

(e.g. new lineages or taxa). For example, allopatric speciation is spontaneous and does not require selection, because "geographically isolated populations will tend to diverge from one another without any input from selection" (McShea & Brandon 2010: 113). If the entities are structures of an organism, the divergence result in complexity increase (McShea et al. 2019).

Complexity is here to be understood as number of different part types at a certain level of analysis of an organism (also called 'horizontal' complexity): a human with 'wisdom teeth' is more complex than another lacking them. This concept of complexity focuses on the origin and number of variations ('how many') and not on their (eventual) function ('what for'): functionless parts also count as complexity. Heritable variations accumulate accidents, and accidents increase the variation, and so on (if there is no inheritance mechanism, ZFEL does not apply): but whether or not some of the new parts will perform a function is a different issue, later decided by selection. For examples, earlobes, appendices, webbing between fingers are products of ZFEL with no current use, but that could become (ex)aptations in the future.

Complexity is 'level-relative' in that different parts, to be considered as forming a complex whole, must be viewed as belonging to a higher structure; at the lower level, they just represent 'diversity', or amount of differentiation among individuals (and not parts within an individual). Considering diversity and complexity as the same phenomenon at different levels helps clarifying why complexity is orthogonal to function: diversity is not expected to have any *a priori* function. Mixing complexity and any form of functionality (including fitness), according to the authors, makes it impossible to understand the relationship between the two.

The opposite of complexity is redundancy, or similarity among parts[154]: the ZFEL tends to destroy redundancy, as it happened with the lobster's two claws, one of which became hypertrophic because of the lack of constraints (e.g. selection for similar claws sizes). The same mechanism is responsible for pseudo-genes, duplicated genes that, initially identical, differentiate later through the accumulation of mutations because they are not under direct selection.

A population will always show increasing diversity and complexity[155]; if not, it is because of some kind of constraint[156]: evolutionary forces (e.g. selection against diversity),

[154] Note that the authors define complexity with reference to parts, not to functions. Redundancy is synonymous of duplicated parts, with no reference to their functions. Two different genes performing the same function are not redundant under this view.

[155] For example, because of phenotypic mutations (mutations in the process of building the phenotype, e.g. in the translation and transcription of DNA to RNA and protein). Although selection

limitations of the inheritance system (e.g. number of genes), genetic mechanisms (linkage, inbreeding, asexual reproduction) hinder diversity generation. This is why the authors define the law as 'zero force'. Their view is a kind of 'Copernican twist': stasis is the result of conflicting forces (the tendency to increasing complexity contrasted by e.g. selection or developmental constraints), and not, as in the classical adaptationist account, the natural outcome of the absence of evolutionary forces. Species stable during long periods of time are expected to be under strong stabilising selection or strong constraints. As in the Punctuated Equilibrium account, what needs explaining are the long periods of stasis, not the quick evolutionary changes.

Several phenomena are listed as putative evidence of ZFEL:

- Nucleotide positions not under selection (as a consequence of the intrinsic unfaithfulness of the DNA replication process[157]) diversify spontaneously (as, e.g., also underlined in Kimura's theory, see §4.2.4);
- Tissues and organs not under selection are more variable among individuals;
- Homologous characters can be maintained by stabilizing selection in multiple lineages and yet the developmental and genetic mechanisms underlying them might diverge;
- Laboratory populations spontaneously drift.

The ZFEL is fundamentally different from drift: while the latter affects the *means* of traits' values, in the same way as selection does, but without any relationship with the trait's fitness contribution, the mechanisms behind the ZFEL only affect their *variance*. It is therefore not directional in the mean, but it is *directional in the variance* (Fig. 4.39). Diversity (as variance of distribution of a trait) increases for any character[158], be the distribution continuous (measurable dimension) or discrete (discrete trait); and at any scale: the organism, the population, the species, the clade,…; and for any property: morphological, physiological, behavioural,… Populations always show drifting means and diffusing variance, and this will result in increased horizontal complexity. Later, selection will grant functions to some of the new parts and eliminate others: simple organisms persist due to strong selection, not to its absence.

---

against genetic mutation is high, selection against phenotypic mutations is low (a puzzling phenomenon) (Bürger et al. 2006).

[156] For an analysis of constraints see the correspondent paragraph in §4.4.1

[157] Thanks to Davide Vecchi for this comment.

[158] Of course, an occasional decrease might sometimes occur by chance.

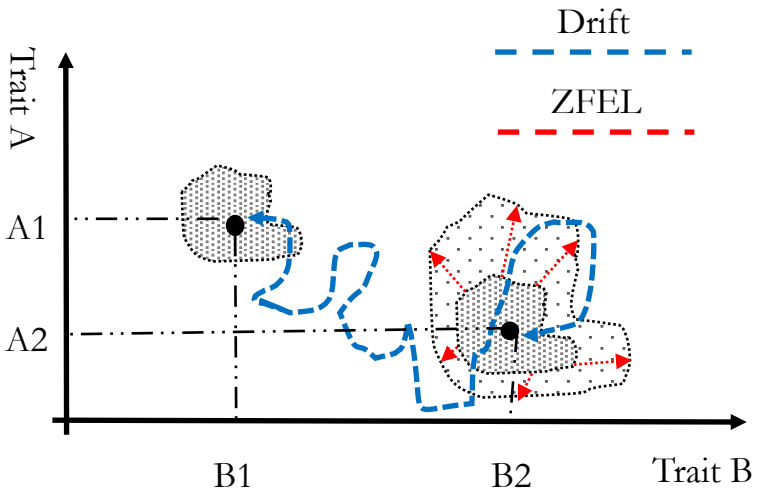


Fig. 4.39 - Effects of drift and ZFEL in a population. Shaded areas represent the individuals of the population. Drift changes the traits' average value in the population (increasing B and decreasing A). ZFEL always increases the traits' variance in the population (expanding the area representing the population's individuals).

Like drift, the ZFEL depends on patterns of diffusion, and not on how it is produced. It describes a phenomenon, it does not explain it. Increasing diversity and then complexity can spring from several mechanisms. It can be a consequence of drift, in the absence of constraints and of selection[159]. It can result from point mutations, insertions, deletions and other genetic mechanisms, even if they are not random *per se*: it is enough that they are random with respect to each other. Non-random phenomena can also push the ZFEL: selection can multiply the number of parts when acting on different parts in different ways, as in the case of the lobster's claws, or in that of "selection on a bipedal primate's hip to improve walking ability and simultaneous selection on the shoulder to improve throwing ability", that "will tend to make hip and shoulder even more different from each other" (McShea & Hordijk 2013). Selection can also cause populations to undergo ZFEL if they evolve randomly with respect to each other.

The position of the authors with respect to selection is twofold. On the one hand, they seem to adopt an extreme adaptationist position in several senses. They put selection at the core of their argument by stating in the preface that "there is always a tendency for diversity and complexity to increase that does not depend on natural selection" (McShea & Brandon 2010: xii). They insist that only selection can create function (e.g. p. 122). They limit to

[159] Drift reduces variance not in itself, but only in the presence of absorbing boundaries (which are constraints), values where the population gets stuck once reached. Additionally, the same allele in population genetics models can have very different subjacent molecular structures. Therefore, the contradiction between the effects of drift in population genetics models (reducing diversity) and in ZFEL (increasing diversity) is only apparent (p. 104-105).

mutation, migration and non-random mating the evolutionary forces beyond selection; however, they quote almost solely the latter. They mention only selection as the force behind differentiation of identical parts that contribute to ZFEL; they especially quote selection as the evolutionary force behind stasis (as lack of ZFEL), justifying the lower-than-expected presence of ZFEL in nature because "selection might work against [horizontal] complexity" (p. 87). On the other hand, they claim that "it is an empirical question how much important natural selection is in evolution" (p. 102), and that differential reproduction is not necessarily linked to selection (p. 97).

The 'collaborative acting' of ZFEL and selection is behind the suggestion of McShea and Hordijk (2013) of an alternative route to adaptive complexity. Instead of considering complex structures as the result of small, incremental improvements on previous simpler ones, they postulate a process of 'evolution by subtraction'[160]. Under this view, complexity is high at the beginning of the evolutionary phenomenon, with many different part types not always functional; later, selection assigns a function to some of the parts in order to increase the organism's fitness, while eliminating others. Complexity decreases, and does not increase, because of selection. This account has the advantage of dissolving the problem of the non-functionality of intermediate forms.

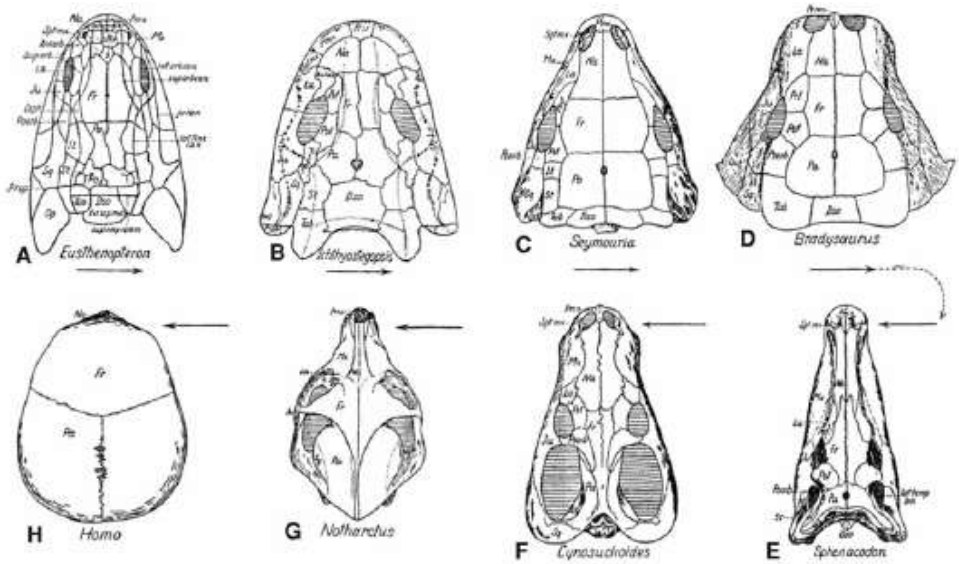


Fig. 10 *Eusthenopteron* is a Devonian lobe-finned fish. *Ichthostegopsis* (now *Ichthyostega*) and *Seymouria* are Permian (labyrinthodont) amphibians, *Bradysaurus*, *Spenacodon*, and *Cynosuchoides* (now *Cynosaurus*) are Permian reptiles, and *Notharctus* is an Eocene primate. From Gregory (1935)

Fig. 4.40 - Decrease of number of skull bones in the transitions from amphibian to reptile to mammals (from McShea & Hordijk 2013)

[160] The idea has points in common with Kimura's view and Stoltzfus' constructive neutral evolution. Selection (and drift) 'subtract' and destroy variation, however generated (Thanks to Davide Vecchi for these comments).

The authors suggest that reduction of skull bones in the transactions from amphibian to reptile to mammals is an example of this mechanism (see Fig. 4.40).

McShea and Anderson (2005) postulate two mechanisms that push organisms towards higher horizontal and vertical complexity[161]. The emergence of higher level entities from lower level ones (e.g. multicellular organisms from protists) might be due to higher efficiencies linked to division of labour among very specialised parts. Differentiation of parts is however limited by constraints (developmental and physiological) and by redundancy requirements to limit risk (some parts are so fundamental that a back-up must always be available). The evolution of higher-level organizations entails:

- on the one hand, the loss of part types in the low level entities (e.g. organelle from protist to multicellular evolutionary transition) for the sake of economy and for the need of higher coordination among previously freer parts (which now need to lose degrees of freedoms, e.g. ability to move independently of other parts). Parts integrated in higher level organisms will have fewer parts and functions than their ´stand-alone' equivalents;
- on the other hand, the lower level parts integrated in the higher level organism differentiate and create intermediate level entities (e.g. organs).

The result is the emergence of new functional higher-level entities with more hierarchical levels, thus showing increased 'vertical' or 'hierarchical' complexity, and with more parts, also structurally and 'horizontally' or 'nonhierarchically' more diverse complexes.

McShea (1991) suggests that complexity increase or stasis can be generated even in the absence of trends (i.e. active processes or 'forces'), if boundaries limiting complexity upward or downward exist, such as minimum complexity of any existing organism, or maximum complexity to avoid chaos. The same idea is behind Gould's 'drunkard's walk' metaphor (Gould 1996: 149-150): a drunkard walking along a wall on his right will show a 'trend to the left' in his (in principle) chaotic wandering, because he cannot go through the wall. In the same way, evolution shows a trend towards increasing complexity because there is a limit to minimal complexity.

[161] See §2.3.2 for definitions of vertical and horizontal complexity.

4.5.5. Final comments on the systemic accounts of variations

The neo-Darwinian view interprets evolution as a process of 'chance and necessity' (Monod 1970): the source of complexity are on the one hand the environment, through the selection of casual mutations (the 'chance') that increase differential reproduction, and on the other hand the genetic system of organisms (the 'necessity'), that irreversibly fixed them within the population. Complexity is thus a result, and not a target, of selection.

The accounts in this section suggest that things can be the other way round: complexity can appear spontaneously and only afterwards be submitted to the judgement of selection. Complex forms can be the result of simple natural laws, like the snowflake. Thermodynamic processes can propel the rise of complex order locally, while increasing entropy at a global level. Systems composed by parts and subparts interacting in multiple and non-linear ways can self-organise according to complexity laws. Individuals in isolated populations free from selective pressure tend to differentiate generation after generation in increasingly more and more complex individuals.

Neo-Darwinism in the empirical adaptationist version can hardly accept any of these proposals, as it does not admit that the phenotype might retain footprints anterior to the selective stage, i.e. complex elements not moulded by selection upon casual inputs.

The last section of this chapter shows how adaptationism usually reacts to non-adaptationist accounts and suggest that this reaction is not based on an inevitable incompatibility.

4.6. Are pluralism and adaptationism compatible?

We have seen how some non-selective forces can be potentially integrated into Neo-Darwinism with ease (as it is the case for genetic accounts §4.2.7), while others would need a re-interpretation to fit within the adaptationit account (as in the case of phenotypic and developmental accounts §4.3.5, §4.4.4), and finally some seem quite incompatible with the classic theory (namely, the systematic accounts §4.5.5)

Regardless of this graded compatibility with adaptationism, mainstream evolutionary biology for the most part denies any importance to non-selective phenomena. Charlesworth et al. (2017) provides a good summary of the usual arguments that advocates of adaptationism brandish to deactivate pluralistic accounts of evolutionary changes:

- They deny the phenomena: "careful genetic studies [..] do not support important roles in adaptation for processes such as directed mutation or the inheritance of acquired characters, and therefore no radical revision of our understanding of the mechanism of adaptive evolution is needed." (Charleswoth et al 2017: 1);
- They admit the phenomena, but claim that they are rare and thus immaterial: "apparently puzzling results in a wide diversity of organisms involve processes that are consistent with Neo-Darwinism" and their effects are "close to neutral, and thus probably not relevant to adaptive evolution." (Íbid., 4);
- They admit the phenomena and their materiality, but claim than it is already included in what the modern synthesis explains: "Finding that epigenetic variation plays a significant role in quantitative trait variability would thus not radically change our understanding of how populations respond to selection." (Íbid., 5);
- They admit the phenomena and their potential materiality, but claim that more data are needed to consider them within the theory: "In contrast with the *rigorous empirical evidence for the role of DNA sequence variants in adaptive evolution* that we have outlined, there is currently *little evidence for effects of epigenetic changes*, although more data are required." (Íbid., 7, italics mine). It is not puzzling that much more data are available for the role of DNA in adaptive evolution, provided that it is the current biological paradigm.

All these answers by empirical adaptationism to pluralistic accounts amount to the denial of the existence of any trait that is not the result of selection. If such a trait appears, empirical adaptationists would claim that selection did not have the time yet to wipe it out or to take advantage of it or to modify it to fit the environment: that is, it would recur to some of its pillars[162] to reject pluralism. On the other hand, non-adaptationist accounts usually question one or more of the adaptationist pillars, as we have shown in §4.1.2. Does this clash make the two approaches incompatible? We believe it does not and, to show that it doesn't, we analyse how, in many cases, the mutual connections among the pillars can help relax and relativize their importance, so that it is possible to put aside some of them without renouncing to adaptationism. We also identify the only fundamental connection between two of the pillars that apparently confront pluralism and adaptationism, and we propose an interpretation to limit its range and significance.

[162] See §3.2

The four pillars of adaptationism are interconnected so that giving up one without affecting the others can be difficult; however, the interpretation of some of the pillars can be loosen thanks to such connections. On the other hand, the link between the two external pillars strengthen both[163]. Let's consider in what sense.

The assumption about ubiquity, for example, can be loosened thanks to natural selection timing: if a trait does not currently show any apparent adaptive value, sooner or later it will. Every trait can eventually become a target of selection, either actual or potential (Mayr 1983). At the level of the whole phenotype, a population can show polymorphism (the contemporary presence of various fit phenotypes) if it is still undergoing the process of selection: however, provided sufficient time has elapsed, any population will get to the fixation of a single, optimal phenotype (Sober 1998a, Grafen 2014).

In the same way, the assumption about the unconstrained randomness of variation is also flexible regarding timing. Adaptationism can admit the existence of constraints on variations in the short-term (e.g. phylogenetic inertia), but these are assumed to be eliminable by mutations and by the cumulated effect of natural selection in the long-term (Parker & Maynard Smith 1990, Ayala 1999, Neander 1995). Dawkins considers pleiotropy unimportant because "there is no reason why selection should not favour modifier genes that detach the two phenotypic effects" (Dawkins 1982: 35). Natural selection will solve any problem regardless of genetic details, provided that the underlying genetic variation is either already present or that genomic change will produce the appropriate genetic variation on which basis selection can then sort and eventually optimize (Mayr 1963). Grafen, an extreme adaptationist, coined the term *phenotypic gambit* (Grafen 2014: 159) for "the research strategy of studying organisms in ignorance of the actual genetic architecture of the trait in question": he claims that our understanding of the design of kidneys, eyes or patterns of camouflage are successful regardless of our ignorance of their underlining genetics. Nesse (1994) provides examples of the same adaptive solution obtained in different organisms through different selective paths. The vertebrate eye has a blind spot due to the nerve fibers that run through the retina, while the cephalopod eye does not, as the optic nerve does not interrupt the view. Natural selection has been able to overcome this constraint by constructing an elaborate work-around of the eyes wobbling back-and-forth to correct for the blind spot. In early vertebrates (sharks, skates, rays) cranial nerves radiate directly to the organs, while in tetrapods they run through a complicated winding path through the

[163] See §3.2 for discussion about pillars of empirical adaptationism and their classification between 'internal' and 'external' ones.

cranium, around structures evolved after the common ancestors with early vertebrate. Natural selection has been able to overcome the constraints linked to the new structures. It is also worth mentioning that adaptationism defines constraints relatively to the optimal phenotype (Amundson 2001): a constraint does not contribute positively to the configuration of a trait but rather prevents the optimal phenotype to appear[164].

In these examples, assumptions about ubiquity and constraints can be loosened recurring to the assumption about timing. However, the assumption about the unconstrained randomness of variation and the belief that order only comes from natural selection are closely linked: it is difficult to reject or loosen one while keeping the other (see also Gould 2002). On the one hand, if variation is not random and/or unconstrained, natural selection could not have an unlimited range of potential variations to choose from and organisms could also be shaped by processes other than natural selection (namely, the ones that limits and constraints variation): therefore, we could detect their action by inspection of the organism alone, without reference to their ecological niche (which is how adaptationism explains traits, Fodor & Piattelli-Palmarini 2010). Given that the variety and complexity of living beings springs "from the war of nature, from famine and death" (Darwin 1859: 490), admitting biased/constrained variation undermines one of the Darwinian core ideas, because some of the variation found in the phenotype could be explained without recurring to selection. On the other hand, if other processes (e.g. developmental laws) shape the organism, their action (and result on the phenotype) need not be, and in general would not be, unconstrained and/or random.

We can therefore conclude that the incompatibility between evolutionary accounts based on non-random and/or constrained variation and adaptationism is not rooted in the whole set of assumptions, but in the external ones[165] and their strong link. However, the issue does not concern the nature of the assumptions themselves, but rather the claim about the scope of their application. Adaptationism focuses on the last stage of the flow leading to variation sorting and fixation of the phenotype (given that it considers natural selection as the only relevant force explaining evolution), but presumes to rule on the whole path by postulating that whatever happens before the selection stage is irrelevant in the long run and cannot leave any enduring markings on the organism: natural selection, provided enough time is granted, wipes all previous markings out and puts its final seal on the phenotype. In a certain

[164] This is a radically different interpretation of constraints than, for example, the one proposed by Evo-Devo, see §4.3.2.

[165] Recall that we identify in the supremacy of selection and the random/unconstrained variations the 'external' pillars of the adaptationist view (see §3.4).

way, adaptationism denies the possibility of any intermediate layer of organization between potential variation (an endless and unconstrained source) and potential evolution (the result of natural selection only). The presence in the phenotype of any footprints due to other causes only means that the adaptationist process has not yet been completed. For example, Brooks and McLennan (1991) consider the existence of both one-horned and two-horned rhinoceros descending from a two-horned ancestor as evidence that the current two-horned version has not yet reached a fitness peak[166]. In an Aristotelian interpretation of adaptationism, we could say that whatever happens before the selection stage is pure matter (e.g. Gould 1976), while the form is provided uniquely by natural selection: no previous form whatsoever can survive the sieve of natural selection.

This extreme view is not at all necessary in order to defend Darwin's theory. Either because of lack of time to respond to change, or because of lack of relevancy of a trait in the 'war for life', or even because of the lack of such war in some occasions of resources abundance, the strength of selection can shrink and, in the same way as the diminishing level of water in a lake makes new details visible, this shrinking can uncover all kinds of pluralistic phenomena that collaborate to the order in life. Pluralism is a polytheistic proposal: instead of changing one god for another (the environment), it suggests to distribute the creative power to a whole pantheon.

4.7. Conclusions

Fig. 4.41, an extension of Fig. 4.4, illustrates the various non-adaptationist views concerning evolution seen in this chapter: developmental processes and the phenotype play an active part in the explanation of evolutionary change, and the phenomena that trigger such changes encompass much more than just mutation, recombination and genetic flows. Natural selection still plays an important role in the configuration of the evolved phenotype; however, other sources of variability accumulate and leave their mark on it. Traits also bear information *about their source* and *not only about the environment*[167]. Adaptive traits coexist with non-adaptive traits: their relative abundance and importance is an empirical matter.

---

[166] The authors claim that the one-horned and two-horned types does not represent two adaptive peaks because they were not achieved independently: the former was derived from the latter, that is present also in extinctic species (Brooks & McLennan 1991: 145-146).

[167] As illustrated in §3.4, empirical, or 'hard', adaptationism claims that the only source of complex traits is selection, through the accumulation of small, random and unconstrained variations. Selection moulds the organism based on the environment, which is thus the final source of any phenotypic

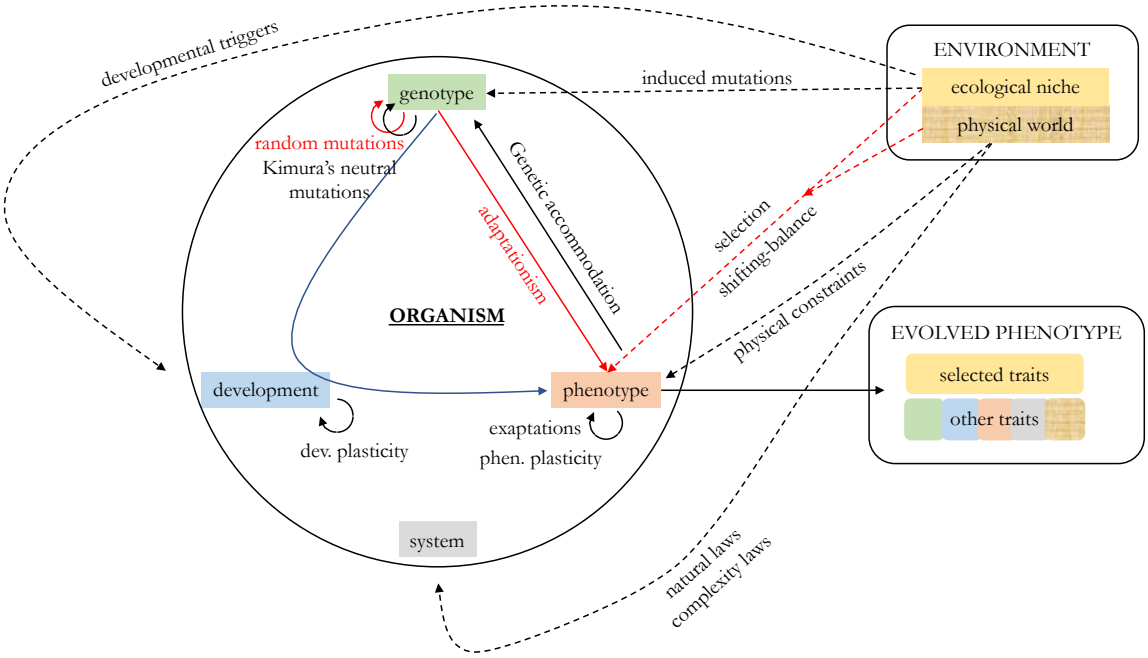


Fig. 4.41 – Pluralistic view of evolutionary changes (compare with Fig. 4.4). Together with adaptationist phenomena (in red), other internal and external processes impact the evolution of the phenotype. The colours in some of the traits of evolved phenotype show how selection cannot always superimpose on every phenotypic effect due to non-adaptationist phenomena .

In the next chapter, we start by showing why adaptationism resists to this change of paradigm, and what tools it brandishes to defend its position, with a special focus on the use of fitness as measure of evolutionary impact: a 'troy' horse that, if accepted, automatically deactivate all non-selective processes of any evolutionary relevance. We then propose to add robustness as a second measure of evolutionary changes to track the impact on the phenotype of non-selective forces, and develop a theoretical approach that unifies the explanations of adaptationism and pluralism in a single model.

trait: any phenotypic detail not mirroring an environmental requirement is either insignificant or provisional. Selection lies not only behind adaptation and variation of life, but also behind its complexity.

Plate 429 - Coquilles Univalves, from Lamarck, J. B, *Histoire naturelle des animaux sans vertèbres, présentant les caractères généraux et particuliers de ces animaux*... London :J.B. Baillière,1835-1845

5. Evolution beyond selection: a logical space for architectural changes

After analysing how adaptationism and other non-adaptationist proposals answer the questions introduced in chapter 2 about the variety and complexity of the biological world, we propose in this last chapter a unifying model to integrate all of them.

The abundance of accounts that aim at explaining how organisms evolve and how organismal complexity increases reflects the wide range of different evolutionary phenomena. We start (§5.1) by showing that, to classify them, fitness alone is an insufficient variable for at least two reasons: it can track, but not explain, the change; and it is by definition a measure of selection, blind to non-selective processes. We then (§5.2) introduce the concept of biological robustness, and suggest how it could offset these two limitations. We finally (§5.3) model evolution through a design space whose dimensions are fitness and robustness, and show how populations, evolutionary forces and evolutionary phenomena can be mapped and classified within it.

5.1. Adaptationism, pluralism and the role of fitness

5.1.1. Evolutionary phenomena and architectural changes

Evolutionary phenomena encompass a great variety of cases, from micro-evolution to speciation, from the configuration of polymorphism in populations to the appearance of new traits and functions. Each phenomenon might be caused by different forces, not always and not only selection[168], and each force might act with a different strength and duration. Let us analyse three cases that are intermediate in this range of possibilities (Fig. 5.1), specifying for each what the change consists of in terms of traits and functions[169], and what the role of natural selection and other forces might have been in shaping it.

[168] We thus advocate, as it should be clear by now, a pluralistic view of evolution.

[169] That is, according to our definition, in terms of architecture (see §2.3.3)

Changes in the wings' colour of *B. betularia* following the increasing pollution due to coal smoke in industrial England is an example of evolution led by natural selection (Fig. 5.1 – i)[170]. Originally, individuals had mainly white wings (*morpha typica*) to camouflage efficiently on the white bark of poplars. Due to the increasing pollution, barks turned dark, and individuals with white wings became easy target for predators, while the mutation correspondent to black wings (*morpha carbonaria*) spread quickly through the population. The trait itself (the wing's colour) and its function (mimicry) did not change. The black version of the trait already existed, either actually, although with a very low frequency; or as a potential recurrent mutation (as it seems to be the case, see Van't Hoff et al. 2016).

Finches' speciation in the Galapagos islands represents a different case (Fig. 5.1 – ii). The opportunity to access new nutritional resources in each island gave raise to new versions of the beak's morphology, allowing to perform the same function of 'feeding' in new ways (e.g. reaching insects in places inaccessible with a short beak).

Finally, new traits and functions, like flight (Fig. 5.1 - iii), or the transformation of an existing trait into a new one often involve the need for a new phenotypic architecture. The turtle carapace, for example, is composed by modified bones such as ribs, vertebrae, and part of the pelvis, and by specialised dermis. The aberrant positions of the ribs compared to all other tetrapods (superficial and dorsal to the limb gridles) result in a novel morphological arrangement that constitutes the unique chelonian Bauplan (Burke 1989).

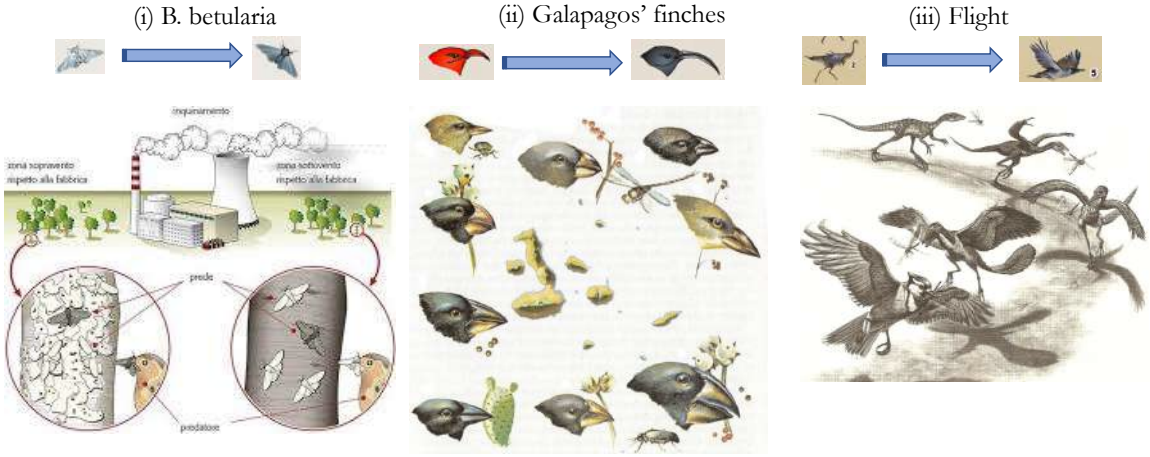


Fig. 5.1 – Varieties of evolutionary phenomena. (i) Changes in *B. betularia* wing's colour. (ii) Speciation of Galapagos' finches. (iii) Appearance of a new function: flight.

As we have seen in chapter 3, adaptationism, both in its narrative accounts and in its formalizations, claims that selection is the only necessary and sufficient force behind all these

[170] See §5.3.2 for a more detailed account.

kinds of evolutionary phenomena. To support its view, adaptationism reduces any phenotypic change to a fitness change. Without taking into account architectural considerations, it borrows from population genetics the modelization of complex phenotypes as correlations among traits and represents them only through scalar fitness (Wagner 2011: 4, see also §3.3 and §3.4). An improvement in phenotypic design is measured as increase in fitness, but fitness (at least as biologists usually apply it in formal models) summarizes the success of a design by measuring its results in competition with others[171], without investigating the reasons behind such success. Under this approach, the organism itself and its architecture remain a black-box (Hall 2003b), whose internal mechanisms transforming alleles' frequencies into average fitness are unknown and, at least in the adaptationist formal models, of little interest (Fig. 5.2).

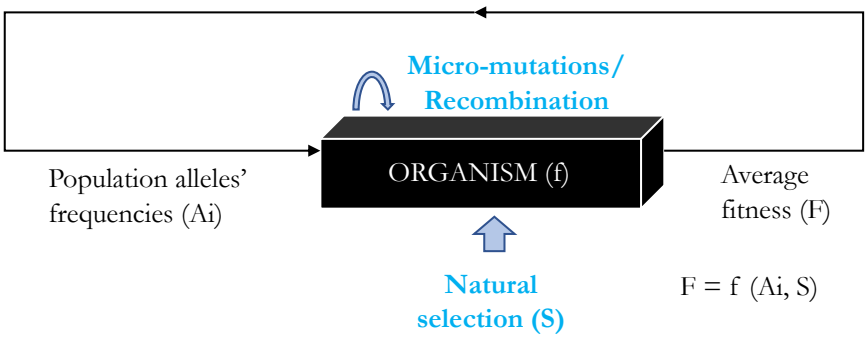


Fig. 5.2 – Formal adaptationism's view of organisms: a black box whose internal architecture (as we define it in §2.4.2) is irrelevant to its evolution and is summarised by the function f (considered as fixed[172]) that transforms alleles' frequencies Ai into average fitness F, based on selection's strength S. Novelties only come through micro-mutations that can introduce new alleles, or recombination, and are immediately submitted to the judgement of selection once appeared.

Such a vision is adequate in textbook cases like the *B. betularia* wings' colour changes, when the evolutionary phenomenon consists just in the modification of the relative frequencies of existing traits' versions (or the appearance of slightly different versions thanks to micro-mutations), led by natural selection: the black wing version becomes more frequent because it grants a higher fitness. The approach, though, appears unfit to capture the complexity of evolutionary phenomena involving new traits: in the finches' speciation example, although it is undeniable that natural selection played an important role in the fixation of the beaks' variants, the transformation of beaks seems to involve more complicated (and non-selective) mechanisms than just the selective shuffle of variant's

[171] In §5.1.2 a second interpretation of fitness is introduced that tries, we believed unsuccessfully, to overcome these limitations: dispositional or propensity fitness.

[172] Note that adaptationism cannot admit any mutations able to substantially change the phenotype. All changes are swift and gradual: a new architecture cannot appear just due to one mutation.

frequencies (e.g. expression patterns of different growth factors in craniofacial development, Abzhanov et al. 2004). This is even more so in the case of novel functions like flight, that consists of cumulative changes in architecture through several intermediate steps[173]: accounts of such phenomena based uniquely on mutations, fitness changes and natural selection are far from being universally accepted. Other processes might play a role in creating such novelties, even though natural selection might determine, once appeared, their fixation in the population[174]. To understand the role of these other processes, it is necessary to 'open' the black-box and to analyse what internal mechanisms might determine, constrain and modify its architecture (Fig. 5.3).

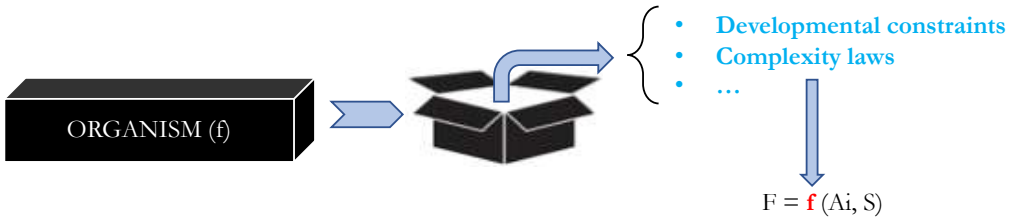


Fig. 5.3 – To understand how non-selective forces impact evolution, it is necessary to 'open the black-box' and understand how and why, under their pressure, the architecture of the organism (summarized by the function f) changes.

To detect and measure the relative effects of all forces at play, we believe that a classification is needed to distinguish the effects of selection upon the population average fitness from the effects of other evolutionary processes that *do not impact primarily and directly on fitness*. In the next section, we develop these ideas: we discuss the importance of fitness for adaptationism, how the concept changed since Darwin's original exposition, and how it must be deactivated in order to integrate non-selective forces into our picture.

5.1.2. Stepping beyond the imperialism of fitness

From Darwin to Fisher

The list of areas analysed by the EES (§4.1.3) should be integrated with another fundamental difference between the MS and the non-adaptationist accounts: the emphasis given to changes in genotype's frequencies and their fitness. As long as these are considered as the only proxy for evolution, any process not directly affecting the genetic pool of a population is dismissed as evolutionary *explanans*. This is not, of course, what Darwin originally claimed,

[173] Additionally, these steps are not the same for the three convergent evolutionary paths that led to pterosaurian (flying dinosaurs), avian (birds), and chiropteran (flying mammals, or bats) flight.
[174] Neutral model also challenge the role of selection in fixing the novelties.

and not just because he knew nothing about genetics. In the *Origin*, he talked about 'individuals' and about 'fit' between organisms and environment:

> The offspring, also, will thus have a better chance of surviving, for, of the many *individuals* of any species which are periodically born, but a small number can survive. I have called this principle, by which each slight variation, if useful, is preserved, by the term Natural Selection, in order to mark its relation to man's power of selection. But the expression often used by Mr. Herbert Spencer of the Survival of the *fittest* is more accurate, and is sometimes equally convenient. (Darwin, *Origin* 1872: 77, italics mine)

Darwin uses the term 'fitness' only twice in the *Origin*, in both cases as a synonymous for 'fit to some purpose', or adaptedness, and not as a scalar variable measuring the individual's success. In the first passage, fitness is a synonymous for 'hability':

> From the highly developed structure of the shoveller's beak we may proceed (as I have learned from information and specimens sent to me by Mr. Salvin), without any great break, as far as *fitness* for sifting is concerned […] (1872: 248, italics mine)

In the second, fitness is applied to biological mechanisms, organs or functions (as we can translate 'contrivances') as a synonymous for perfection:

> Nor ought we to marvel if all the contrivances in nature be not, as far as we can judge, absolutely perfect; as in the case even of the human eye; or if some of them be abhorrent to our ideas of *fitness.* (1872: 538, italics mine)

It was Fisher, one of the MS's fathers, who, sixty years later, while claiming that he was formalizing Darwin's theory, subtly changed the focus. Instead of the individual, he talks about genotypes:

> […] *genotypes* are probably unequally fitted […] to their task of survival and reproduction. (Fisher, The Genetical Theory of Natural Selection 1930: 9-10, italics mine)

And, instead of the property of an individual to be fit for an environment, he coins a new concept, borrowing the idea from Malthusian population dynamics (Malthus 2008): the

fitness m of a gene, a scalar value that summarises all that is relevant with regard to an organism and its evolutionary past and future history[175]:

> Since m measures *fitness* to survive […] (ibid: 33, italics mine)

From Fisher to the Modern Synthesis

Although simplifications are often needed to understand systems way too complex to be completely formalised, it is dangerous to forget that the complexity is still there and to substitute the actual world with the successful model. As Wagner puts it:

> […] whenever a theory is successful, it is also easy to forget its limitations, and this is exactly what happened in the heyday of the modern synthesis, when the grandeur of life's evolution became redefined and demoted to a 'change in allele frequency within a gene pool'. (Wagner 2015: 20-21)

Fisher's fitness re-interpretation has proved formidable to understand *some* of the evolutionary ways in which selection operates, especially in population genetics; however, this success has led to an abuse of the concept of fitness, to the point that any evolutionary phenomenon invisible to fitness changes is considered irrelevant by the MS. The imperialism of (this interpretation of) fitness rests on the transformation of the following simplifying assumptions into supposed empirical outcomes or theoretical pillars of the original Darwinian idea:

- The only relevant evolutionary process is natural selection (this is the element of the synthesis taken from Darwin);
- The only relevant sources of variation are recombination of the existing genetic pool and mutations (this is the element of the synthesis coming from Mendel);
- The only relevant subject of evolution is the genotype (this is an original idea of the synthesis).

[175] Fisher's and Malthus' definitions of fitness are subtly different: Malthusian fitness m refers to the exponential growth rate. With a population of size N(t) at time t, that implies $dN(t)/dt = m\,N(t)$.. Wrightian fitness w is the average number of offspring, and is defined by $N(t + 1) = wN(t)$. For Malthusian fitness, the solution is $N(t) = \exp(mt)N(0)$, and for Wrightian fitness, it is $N(t) = w^t N(0)$. Time is naturally continuous if we use Malthusian fitness, while it is discrete for Wrightian fitness. Both models, however, lead to exponential growth in this most basic form. The relation between the two ways to define fitness is given by $m = \ln(w)$ (see Wu 2013: 1).

From the Modern Synthesis to Adaptationism

Adaptationism takes the dictatorship of fitness one step further. We have seen in chapter 3 that one of the pillars of adaptationism is the idea that only natural selection shapes the phenotype, starting from an endless stock of random/unconstrained variations (micro-mutations). The MS focuses likewise on natural selection, although through its actions on the genotype and without claims about the phenotype. Through this subtle change of focus, adaptationism seizes the concept of fitness handled by population genetics and applies it to its own formalizations in order to support the claim about selection and phenotype. The jump from genotypic to phenotypic fitness is behind what Grafen calls the '*phenotypic gambit*' (Grafen 2014): the idea that, if a (phenotypic) trait is useful to increase (phenotypic) fitness, the trait will sooner or later appear, regardless of its genetic encoding (and it could even appear with different genetic encoding in different species). Once fitness is 'phenotypized', adaptationism can create its own formalizations by identifying the organism with its fitness and by maximizing the latter: Grafen (2007: 1248) claims that "[a]daptation is design, and maximizing fitness is what organisms are designed for", and Waddington (1954) considers fitness as effective design for reproductive survival. Any trait that does not contribute to fitness[176] is anecdotic and will eventually be wiped away by selection: "Traits should be selected for if they increase fitness" (Grafen 2007). The claim about fitness maximization is the second point where adaptationism distances itself from population genetics, given that the latter does not defend such idea.

We can now summarize what we consider to be the two steps through which adaptationism seizes and transforms the concept of fitness introduced by the MS:

- The genotype is largely irrelevant: function (the selective value of a trait) precedes form (the trait and how it is encoded and developed). The only significant evolutionary variable is phenotypic fitness.
- Natural selection maximises phenotypic fitness by selecting traits. Fitness is a comprehensive proxy for the phenotype, which is therefore completely 'functionalised': a function of one variable (fitness).

Fig. 5.4 summarises the conceptual changes that Darwin's theory experienced from *The Origin* until the emergence of adaptationist formal models.

[176] Traits that have no selective value, or even slightly negative value, can of course survive if they are linked in some way to selected traits. An adaptationist would claim that sooner or later a mechanism will appear that separates the useful and useless traits and cause the latter to disappear.

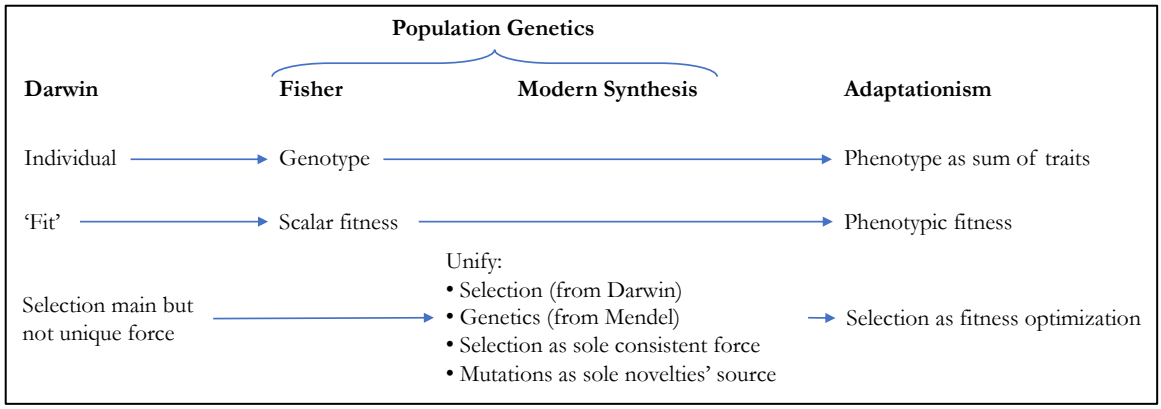


Fig. 5.4 – Subtle and often unrecognized changes from Darwin's original concepts: Fisher talks about genotypes instead of individual, and of scalar fitness instead of 'fit' between individual and environment; the modern synthesis ignores Darwin's flexibility about other possible sources of evolution (see §3.3) and limits to mutations the only source of novelties; population genetics includes all these ideas; adaptationism focuses on the phenotype as the centre of evolution (phenotypic gambit), reduces it to the set of its traits, switches from genotypic to phenotypic fitness and interprets selection as a process of fitness optimization[177].

Problems with adaptationist fitness

These implicit changes in the content of the concept of fitness are not harmless and entail at least two serious consequences: the transformation of Darwinism in a truism, and the rejection of pluralism. Let's consider the first point.

For the MS, and population genetics, fitness is a *predictive* concept: it is a tool to forecast what the population's gene pool will look like in the next generation, given the gene frequencies distribution in the current one. It does neither promote any claim about optimization of traits, nor any narrative about the past history of the trait. If a gene guarantees higher fitness, it will spread across the population, as long as the genotypic constraints allow: why the gene has higher fitness is not a question targeted by population genetics in this sense. It is a 'forward-looking' concept: it just measures the success of a gene. Adaptationist fitness is a totally different matter: it is an *explicative*, 'backward-looking' concept. It aims to justify why a trait exists: it exists because it granted higher fitness. Ecological opportunity is the driving force for any novel traits (Erwin 2015). Fitness becomes therefore at the same time the measure *and* the reason of evolutionary success: that is, cause and effect of differential reproduction (Rosenberg & Williams 1986). Some authors consider that to interpret fitness as *explanans* of evolutionary traits makes it impossible to define independently the concepts of fitness, survival and reproduction, henceforth

[177] This is of course our own interpretation of a limited set of complex historical facts covering a century and a half. It does not pretend to be exhaustive, but only to underline what we consider to be some of the main conceptual changes from Darwin's original ideas to current biology.

transforming Darwinian theory in a truism[178]: if the fittest survives [effect] and the fittest is the one who survives [cause], the theory just affirms that the one who survives, survives (Mills & Beatty 1979). Kauffman (1993: 17) claims that all theories have some circular relationships among their fundamental concepts, and Darwinism is no exception. For example, in Newton's second law of dynamics, while acceleration is independently defined as changing velocity, force and mass are defined one in terms of the other. Many other authors, though, tried to escape the truism by looking for a principle outside of the theory that can act as cause of traits (as it exists in artificial selection: who will reproduce the most is decided by the breeder based on external considerations). Sober (2000) proposes an analysis of fitness that distinguishes two steps in the life-cycle of organisms: survival till reproductive age and reproduction. This approach could in principle help overcome the 'truism' problem, if we suppose that surviving implies traits whose causes are not linked to the actual fact of reproducing. Sober divides fitness between the ability to survive in an environment (or ability to pass from zygote to adult, that he calls 'viability'), and the ability to reproduce (or ability to pass from fertile adult to zygote, that he calls 'reproductive fitness'). Unfortunately, its distinction, although clarifying, retains the limitations of fitness *tout-court*, as both concepts measure evolutionary dynamics *ex-post*. Using a metaphor, it is like saying that either the best designed car is the one that won the race (viability), or that it was the one with the most copied technology (reproductive fitness): while probably true, these sentences cannot help explaining why the car won, or why the technology is so widely copied. Mills and Beatty (1979) suggest considering cause-fitness[179] as a reproductive propensity to have an expected number of descendants. Such propensity is defined by analysing relevant traits based on considerations external to the theory of natural selection (e.g. engineering analysis of traits). Millstein (2006) proposes to change cause-fitness with 'causally relevant physical differences'[180], where this relevance is to be defined likewise in terms of other types of considerations concerning the physical basis of adaptedness. This dependence of the definition of fitness on external theories has been used to claim that fitness is a 'primitive

[178] It might as well be claimed that it is the double-sided interpretation of the concept that creates the truism. If we accept fitness as a cause, but discharge it as effect, the truism disappears: but so does the theory of selection that claims the survival of the fitness.

[179] In order to avoid confusion, we will talk about cause-fitness when referring to the adaptationist concept and about effect-fitness when referring to the population genetic concept. This terminology is however not used in the literature.

[180] The extent of this definition should anyhow be carefully defined: the physical location of the organism, for example, cannot be considered 'causally relevant', or otherwise accidental death due, e.g., to lighting would compute as element of fitness and not of drift.

theoretical term', undefinable within the theory of natural selection (Rosenberg 1982, Rosenberg & Williams 1986). Although these proposals are *ex-ante*, they still focus on reproductive success, thus on the 'selection step' of evolution.

The second effect of the adaptationist reinterpretation of fitness is the putative demonstration that pluralism is explanatory void. Pluralism is the idea that "the basic patterns and regularities of the evolutionary pathways of life" are caused, together with natural selection, by "a set of additional laws, as well as a large role for history's unpredictable contingencies" (Gould 1997). Fitness, in its original formulation, is a measure of the effect of natural selection on a population. As such, other processes and laws are not expected to produce any change in effect-fitness, because they are not a function of the fitness properties of individuals (Lynch 2007): the MS simply ignores them in its formalizations. Successful novelties, however, "may arise without accompanying diversification and there may be lags between the origin of a novelty and its successful exploitation as an innovation" (Erwin 2015: 937). This means that changes in the architecture may appear without any immediate fitness effect (or 'diversification' of the population). Adaptationism, on the other hand, brandishes the lack of effect on fitness as a proof that a non-selective process is evolutionary insignificant. The reasoning is invalid because it is based on the polysemy of the term 'fitness'. Adaptationism reasoning goes as follows: traits are explained by their contribution to fitness, so any evolutionary process should impact fitness or else be irrelevant in shaping the phenotype; but fitness is by definition a measure for selection, so no other process other than natural selection is responsible for traits. The invalidity of the reasoning is easily seen if the polysemy is removed: traits are explained by (cause-)fitness, so, even if it is true that only selection impacts (effect-)fitness, this does not exclude that other processes can shape traits by increasing (cause-)fitness.

Accepting fitness as a unique and sufficient proxy for phenotype is like letting the Troy horse enter within the walls of Pluralism. Non-selective processes cannot be analysed with selective criteria. The following quote is quite significant of the fitness-centred view pervading current evolutionary biology, as it excludes that any other concept could be relevant and it expands the ability of fitness to account not only for reproduction, but also for survival: "Fitness is *the* central concept in evolutionary theory. It measures a phenotype's ability to *survive and reproduce*" (Wu et al. 2013, italics mine).

If we limit ourselves to measuring evolutionary changes in terms of fitness, (1) we will never be able to understand why they cause an improvement/decrease of fitness and (2) we

will never detect the changes that are fitness-neutral and might have, eventually, fitness-effects. Effect-fitness should be integrated with some other measure of the impact of non-selective evolutionary phenomena. Cause-fitness should be dismissed and substituted by another measure of the potential success of a trait. In the next section, we propose to integrate the unidimensional measure of organismic design as reproductive fitness with a second variable able to fulfil both requirements: reflecting *ex-ante* the capacity of the organism to face and survive new environmental challenges, and being independent from the organism subsequent reproductive success. This dimension is robustness.

5.2. Robustness as second dimension of organic design

5.2.1. Evolution as a two steps process

The need for an additional measure of organismic architecture is evident when one considers the flaws of the standard adaptationist view, which explains novelties as the final output of adaptive radiation (Schluter 2000). Erwin (2015) claims that radiations, in the first place, consist of the adaptive exploitation of existing variation, but do not themselves generate novelties; moreover, there are usually macroevolutionary lags between the appearance of a novelty and its exploitation as an innovation (e.g. grass, that spread millions of years after its appearance, Stromberg 2005), or its disappearance (e.g. the 'weird wonders' of the Cambrian Burgess Shale fauna[181], Fig. 5.5). Novelties may arise and persist for long times without accompanying diversification, i.e. without spreading and giving rise to new species (Erwin 2015). The constant rate of the appearance of phenotypic novelties and the variable rate of their wide adoption and spreading also point to a separation of the two processes (Jablonski & Bottjer 1990).

[181] The Burgess Shale is a deposit of fossils in British Columbia (Canada), dated 508 million years old (middle Cambrian). Fossils show an uncommon level of preservation of soft parts' imprints, what has allowed reconstructing many organisms with great detail.

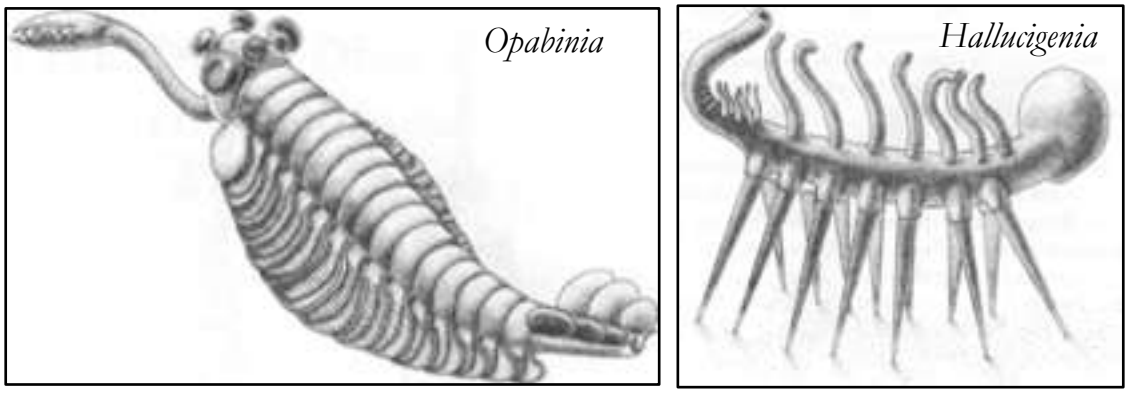


Fig. 5.5 - The great majority of reconstructed animals from the Cambrian Burgess Shale fauna do not correspond to modern groups and show strange, otherwise unknown, anatomical features: e.g. *Opabinia*, with five eyes and a snout like a vacuum cleaner hose (picture from Gould 1989: 126); *Hallucigenia*, that walked on bilaterally symmetrical spines (picture from Gould 1989: 154).

These considerations suggest the importance of clearly differentiating two steps in all evolutionary phenomena: the appearance of architectural change (as defined in §2.3.3), and the (eventual) submission of the change to the judgment of selection[182]. To understand where change springs from, and why it triumphs, it is crucial to consider the step before selection takes place: "[b]y divorcing novelty [*the appearance of a new trait or function*] from innovation [*its spread and fixation*], we can explore the mechanistic basis of each process" (Erwin 2015: 936, italics mine). Fitness tells us a story about triumph, but the seeds of this triumph are hidden among the alternative organisms that did not succeed: and this is a story of robustness, not of fitness. The relationship between fitness and robustness can be also illustrated with a metaphor from the car industry. There are thousands of different models of cars. The variety of cars springs from many sources: new technologies applied to automobile industry; fashion; innovative ideas of solitary inventors; chance… Nevertheless, all cars share some characteristics defined by the physics of the medium they are built from and by the physiology of the humans they have to carry. Some models are very rare, others are common; some are produced for short periods, the production of others continues for decades with only slight changes. We can explain this situation by assuming that only the best models are kept in production -i.e. survive-. How is, however, a car better than another? It might be faster, or consume less fuel per kilometre, or make less noise, or be more reliable, or all of these things together. Our choice of the better car is not automatically based on any of these traits: it depends on our needs and preferences. If we want to take part to a race, we

[182] The idea of two different processes linked to the evolution of organisms dates back at least to Lamarck, whose theory involves an 'adaptive' force (use and disuse and inheritance of acquired traits), responsible for the variety of species and genera; and an inner tendency of organism towards more complex structures ('orthogenesis'), responsible for new phyla (Gould 2001).

might prefer a faster car, while considerations about noise and fuel consumptions might be secondary. If we need a car to go to the office every day, we might look for low fuel consumption and high reliability. The issue of which car we select is thus only partially related to the issue of what characteristics each car has and how these came to be.

We think that the same criteria can be applied to the realm of biology: we should consider "the processes generating phenotypic novelty separately from the ecological and evolutionary processes that regulate their success" (Erwin 2015: 931). Environmental, ecological and evolutionary factors decide which traits of an organism are relevant for survival and grant a higher fitness to the individual possessing them. The set of such traits, however, is defined by laws, mechanisms and processes that act logically prior to the selection stage. The organism shaped by these laws, mechanisms and processes has the ability to self-maintain within a more or less wide range of external variables: which sub-set of these abilities is currently relevant is then decided by environmental selection[183]. The potential for survival granted by these abilities with respect to other possible environments (and thus to new survival challenges) is likewise important, although indifferent for current selection, or even deleterious, if these unused abilities absorb resources (e.g. energy for their maintenance) or interfere with other currently relevant ones (e.g. the male fiddler crab has two asymmetrical claws, one of them much bigger than the other; the big claw can potentially increase its ability in sexual competition, but greatly reduces its ability to escape predator, Jordão & Oliveira 2001). The set of the currently used abilities measures adaptedness, or how fit the organism is for the environment[184]. *The set of the unused ones measures the organism's robustness*. Fig. 5.6 (i) illustrates the logical space of the current environment and the position of two organisms with respect to it: one with unexploited abilities, and one without them.

[183] The idea that the environment can be described by sets of abilities sufficient to survive in it has an adaptationist ground: adaptationism reduces the organism to the set of functions it performs and the environment to the optimization program maximizing these functions.

[184] Note that fitness is here considered, as in Darwin, as a measure of the 'fit' between the organism and the environment, and not, as in Fisher, as the scalar figure measuring the number of offspring.

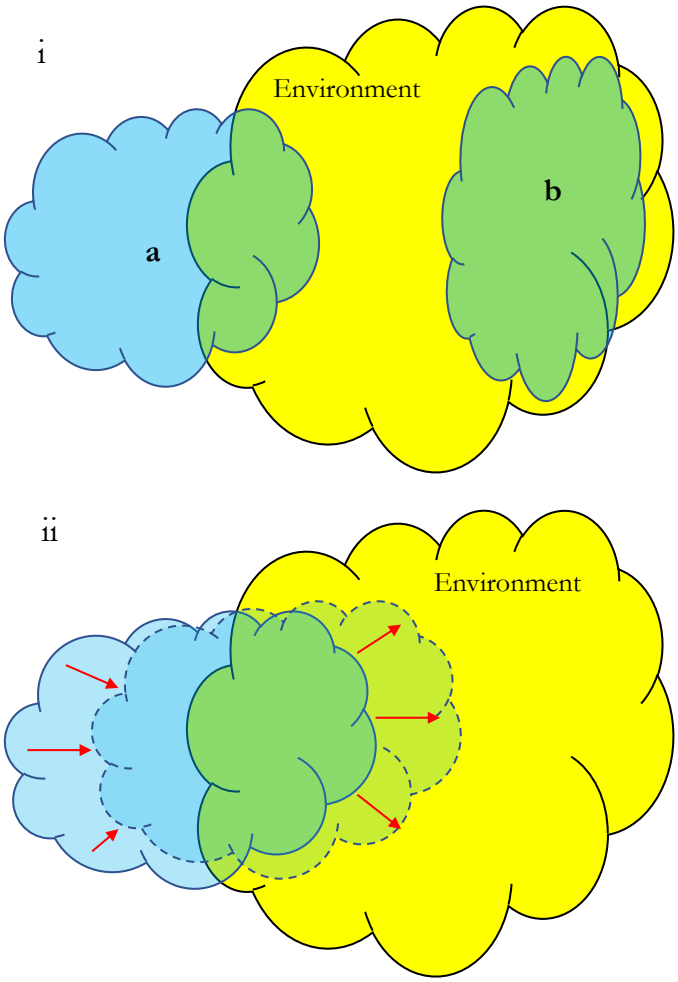


Fig. 5.6 – Logical space of abilities. (i) In yellow, the set of all abilities sufficient to survive in the current environment. Organism (b) possesses a sub-set of these abilities -that measures its fitness (in green)- and no currently unused ones. Organism (a) possesses a smaller subset of current useful abilities (in green) but also some unused ones (in blue). Organism (a) shows therefore a higher robustness than organism b, although organism (b) has a higher current fitness. So, R(a) > R (b) and F(b) > F(a), where R stay for robustness and F for fitness. (ii) In stable conditions and strong stabilizing selection, unused abilities are a burden and the average phenotype tends to lose them and to acquire immediately useful ones, thus reducing its robustness to increase its fitness. This is the 'empirical adaptationism' view, which claims that stabilizing selection is always strong on all traits.

In a stable and unchanging world with strong stabilizing selection on current useful traits, excess robustness in this sense (i.e. potentially useful but unused abilities) is superfluous or even deleterious for the current performance of an organism (Fig. 5.6 - ii). Unused abilities tend to disappear for several reasons (Lahti et al. 2009), among which:

- Their cost, either in terms of energy or in terms of negative correlation with other traits' fitness: they are selected against. Duplicating a (part of a) system is expansive in terms of resources, although it optimizes robustness. Mammals are highly robust but at a high cost; bacteria are simple systems with no particularly complex

robustness mechanisms, but they are able to reproduce massively thanks to their low resource needs (Kitano 2007)[185].

- Recurrent mutations, together with drift, that are not opposed by stabilizing selection (they are not selected for). Mutation are usually deleterious, and tend to cause loss of functions; drift can likewise eliminate traits that contribute to adaptedness. If selective forces do not eliminate the degraded phenotypes, these could expand and get fixed in the population. The cave form of the teleost fish *Astyanax mexicanus* is an example of such degeneration (Fig. 5.4).

Cave form Surface form

Fig. 5.7 – The cave form of the *Astynax mexicanus* (on the left) has lost eyes and pigmentation with respect to the surface form (on the right). Its robustness is thus decreased and it would be worse off in a surface environment (pictures from Lahti et al. 2009).

The actual fate of unused abilities, though, depends on several factors (Fig. 5.8).

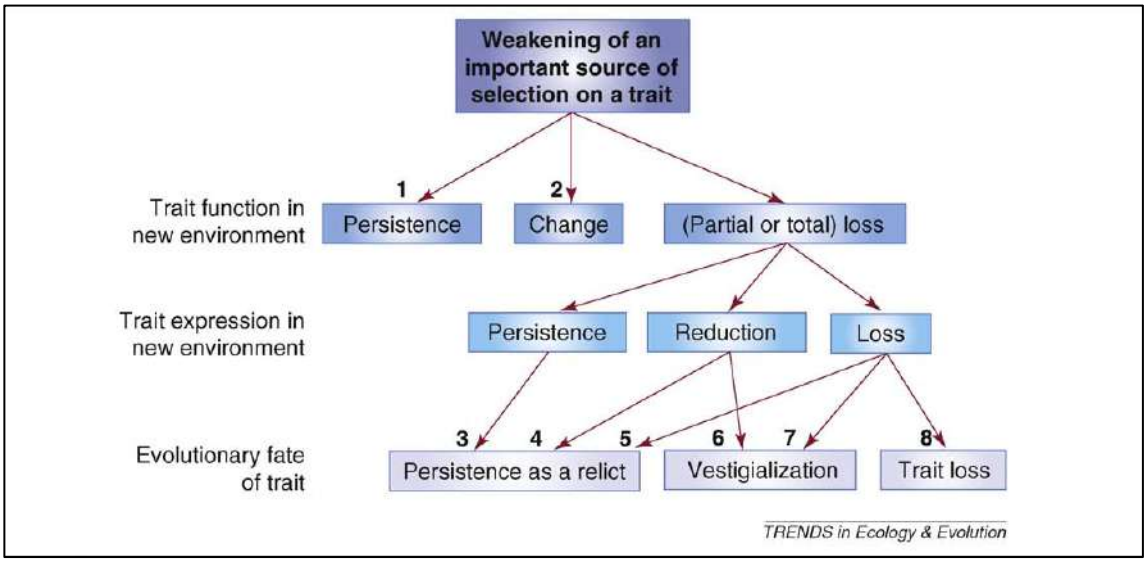


Fig. 5.8 – Possible fate of unused abilities (from Lathi et al. 2009). See text for explications.

[185] Kitano refers here to the robustness of bacteria as individuals. The bacterial high reproductive and mutation rates result, on the other hand, in high lineage robustness, which allow bacteria to survive in a wide range of environmental conditions (thanks to Davide Vecchi for this comment).

In absence of selective pressure, a trait can experience different fates: persistence, modification and partial or total loss. Note that, in a typical adaptationist view, a trait is implicitly considered uniquely as output of a selective pressure, and therefore its evolutionary fate is triggered by the relaxation of such pressure (the logical tree starts with the weakening of a source of selection). The possibility that a trait might appear for some reason other than selection and that it is afterwards judged by the existence or not of a selective pressure is not considered.

Stability is however a theoretical issue: the biological world is far from being stable at any level of analysis or at any time span. As a consequence, robustness is ubiquitous in the realm of life (Truchetet & Pradeu 2018), and "one of the fundamental characteristics of biological systems" (Kitano 2007: 1): mechanisms behind robustness give a fundamental contribution to viability and functionality of biological organizations (Bich 2018). This richness of organismal potential is reflected in their 'overdesign', protecting organisms from all kinds of instabilities (Fig. 5.9).

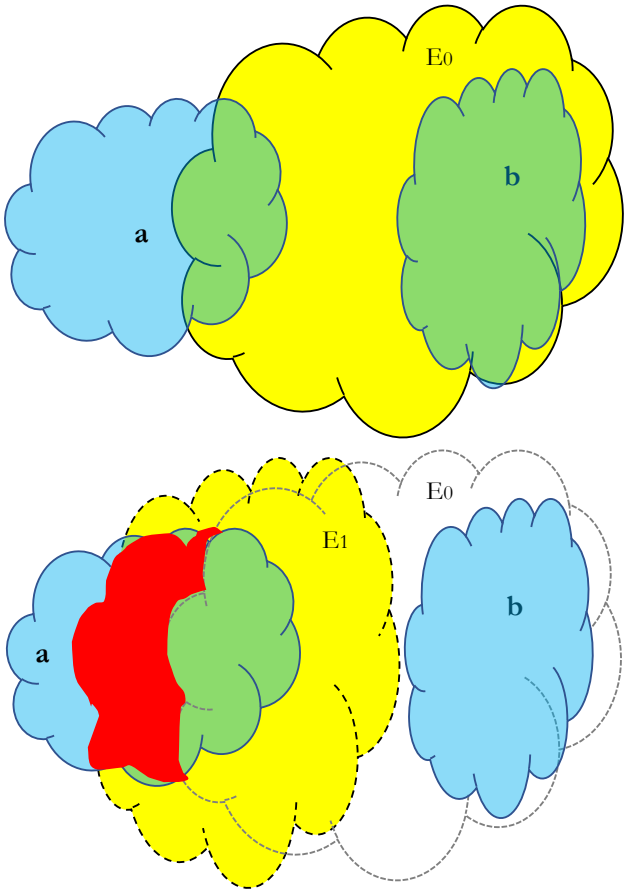
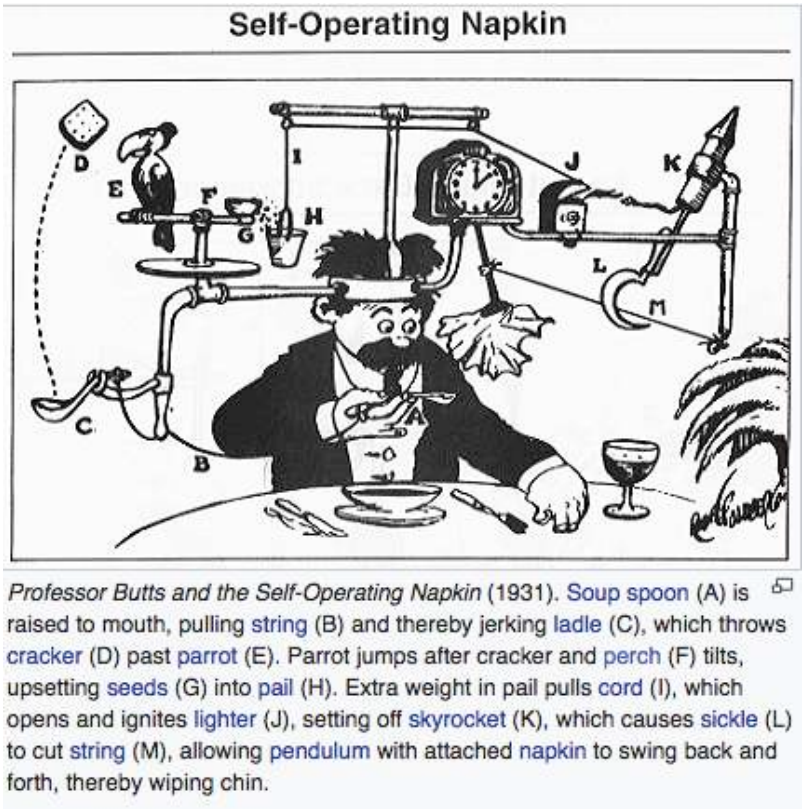


Fig. 5.9 – An environmental change (from E0 to E1) radically modifies the required abilities: organism (b), very adapted to E0 but not to E1, disappears; while organism (a), although not so adapted to E0, survives the change thanks to its higher robustness, that is, to the previously unused, and now useful, abilities (area shaded in red): its 'over-design' saves it.

Contrary to human-designed artefacts, based on the principle of economy of resources, the 'baroque extravagance' of biological design reminds us of a Rube Goldberg cartoon (Fig. 5.10) "with device mounted upon device to accomplish the simplest of tasks" (Krakauer & Pltokin 2005: 98).



Professor Butts and the Self-Operating Napkin (1931). Soup spoon (A) is raised to mouth, pulling string (B) and thereby jerking ladle (C), which throws cracker (D) past parrot (E). Parrot jumps after cracker and perch (F) tilts, upsetting seeds (G) into pail (H). Extra weight in pail pulls cord (I), which opens and ignites lighter (J), setting off skyrocket (K), which causes sickle (L) to cut string (M), allowing pendulum with attached napkin to swing back and forth, thereby wiping chin.

Fig. 5.10 - Rube Goldberg description of a self-operating napkin machine: an example of 'overdesign' (from Wolfe 2000).

Note that robustness is not the same as complexity: it can increase while complexity decreases, as in the case of the column of aquatic mammals, simpler than the column of their terrestrial ancestors (McShea 1991). Robustness refers to the architecture of the organism: a lesser complex organism can be more robust than another one, as 'what is not there, cannot break'.

But, what is robustness, how can we formally define and measure it? The next section provides some answers to these questions.

5.2.2. Definition of robustness

We have informally identified biological robustness with the set of available, but currently unused abilities, that an organism possesses. It turns out to be difficult to give a more formal and universally accepted definition. Waddington's concept of canalization (1957) can be

interpreted as a first approximation to the idea of robustness. Waddington noted that organisms suffering unusual environmental shocks during development show higher phenotypic variation, and that these variations, if the shock is maintained throughout several generations, tend to get fixed and to appear in following generations even if the shock is finally eliminated. He claims that this phenomenon is linked to the emergence of mechanisms safeguarding the developing organism against repeated turbulences, mechanisms that insure that the same phenotype appears notwithstanding their association with different genotypes[186]. Moczek (2008) likewise resorts to a similar idea, even though in order to explain the appearance of novelties: the canalization of developmental processes, i.e., the ability to maintain the same phenotype against changes in the genetic and environmental input, allows the accumulation of cryptic genetic variation (i.e. phenotypically unexpressed) which, over some threshold, leads to new phenotypes.

[186] See §4.3.2 for a detailed account of this phenomenon.

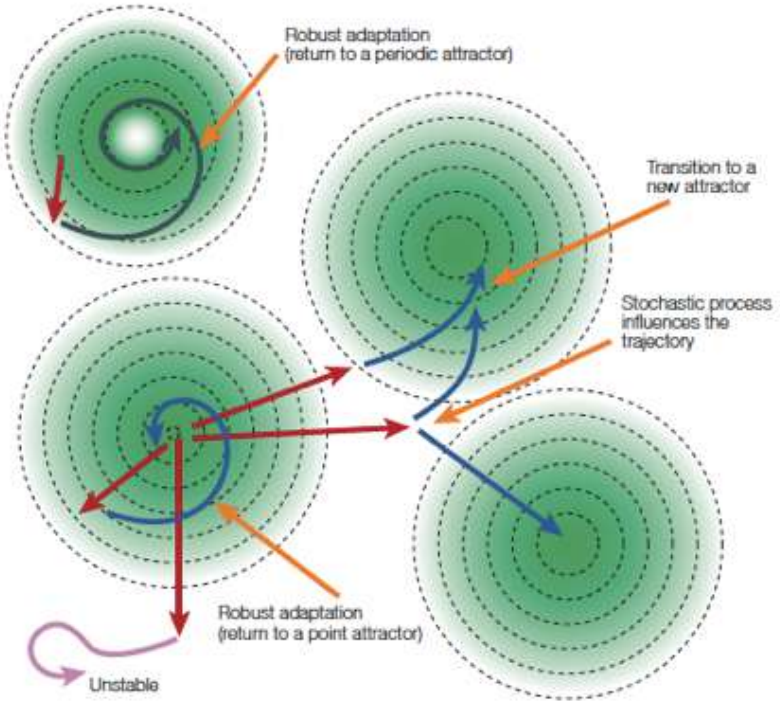


Fig. 5.11 – Kitano's attractors in the state space (figure and text from Kitano 2004). Robust reactions of the system: to stay or to change. The state of a system can be shown as a point in the state space. In this case, the state space is simplified into two dimensions. Perturbations forcefully move the point representing the system's state. The state of the system might return to its original attractor by adapting to perturbations, often using a negative feedback loop. Bacterial chemotaxis is an example. There are basins of attractions in the state space with which the state of the system moves back to that attractor. If the boundary is exceeded, the system might move into an unstable region or move to other attractors. Positive feedback can either move the system's state away from the current attractor, or push the system towards a new state. The cell cycle involves a combination of positive and negative feedback that facilitate transition between two attractors (G1 and S/G2/M) creating a bistable system. Often, stochastic processes affect transition between attractors, as seen in λ-phage fate decision, but maintenance of a new state has to be robust against minor perturbations.

Kitano (2004) defines robustness as "a property that allows a system to maintain its functions despite external and internal perturbations" (2004: 826). A robust system does not remain unchanged, but modify its structure or way of operation in order to guarantee the continuity of its performance: pushed by perturbations, it moves back to the initial equilibrium state (that Kitano calls 'attractor'[187]) or to a new one, instead of dissolving (Fig. 5.11). Truchetet and Pradeu (2018: 45) follow Kitano when they define robustness as "the maintenance of specific functionalities of a given system against internal and external perturbations". Stelling et al. (2004) give a similar definition but focused on 'performance' (or the actual of more or less efficient use of a function) instead than on functionalities themselves.

[187] See also §4.5.3

Pigliucci (2008) shifts the focus from current to new features, and sees robustness as a fundamental piece of 'evolvability' in the strict sense, or the propensity to develop new traits and functions, that also includes characteristics such as modularity and the genotype-phenotype mappings.

Wagner (2015, 2011) in some way unifies Pigliucci's and Kitano's views: he considers robustness at the same time as the persistence of the biosystem[188] in the face of change[189], and as the ability to develop new features.

It is worth underlining the difference between robustness and the apparently similar concepts of homeostasis and stability. *Homeostasis*, a key concept in biology, is about *maintaining a system's variable* (internal energy, e.g. body temperature; or internal medium, e.g. concentration of glucose in the blood) within a range of values, while *robustness* is about *maintaining a system's function* (e.g. bacterial chemiotaxis) against perturbations (Truchetet & Pradeu 2018): robustness is reached also through, but not only through, changes in the system's variables. Homeostasis can therefore be a mechanism of robustness, although robustness includes a wider set of mechanisms. *Stability* is the actual maintenance of the current functions of a dynamical system against small internal and external changes (Jen 2005), and can be insured by homeostasis (Reimann 1996, Duncan et al. 2018, Nijhout & Reed 2014) or other mechanisms (Ponizovskly 2013), depending on the kind and degree of the perturbation. Robustness can also insure stability, but stability can be deleterious for robustness, especially when the latter is level-dependent: as systems are usually composed by many inter-acting subsystems, stability of a part (that is, of a lower level) can be detrimental to the robustness of whole system, as in tumour tissues (Truchetet and Pradeu 2018). Some species gain robustness by increasing instability in a part of its system: for example, the HIV-1 virus is robust against different therapies because of its high mutation rate (Kitano 2007). This trade-off between robustness and stability is particularly evident in systems with high robustness against certain pre-defined kinds of perturbations, that might involve high fragility (that is, low stability) against other, unexpected ones: patterns of forest buffering areas and tree planting are optimal against some kinds of fire, but fragile against others (Carlson & Doyle 1999, 2002, Kammash 2016, Kitano 2007). Modern airplanes are robust against component failures and perturbations, but unstable against total power failure: the

[188] Meaning the trans-temporal identity of that particular self-maintenance biological system.

[189] In this sense, human language is robust, while computer language is not: loss or corruption of information during the transmission of the former does not usually invalidate the message content, while can have disastrous consequences in the latter.

'Wright Flyer' was a non-robust system free from such problem, as it did not have any electric circuit (Kitano 2007).

Robustness and stability must be defined with respect to a *system*, a *function* of the system, and a *perturbation* affecting that function (Table 5.1). Depending on the characteristics of this triad system/function/perturbation, stability can be interpreted as an instance of robustness (Kitano 2007). We talk of robustness instead of stability (Fig. 5.12) when (Jen 2005):

- The perturbation entails *changes in the architecture* of the system, or a radical new environment;
- The feature (e.g. a phenotypic trait) is difficult to quantify;
- The system shows a multi-hierarchical organization of modules. The level of analysis must be specified, given that a system can be robust at one level and not robust at any lower or upper one. In such systems, robustness can spring from the interrelations between architecture and dynamics (e.g. molecular/cellular buffering mechanisms decoupling mutations and phenotype).

Table 5.1

System	Context	Perturbation	Robust property
Protein folding assisted by chaperones	Cytosol	Hydration shell, protein interactions, temperature	Conformation dynamics
Circadian clock	*Drosophila* (Gonze et al., 2002)	Molecular noise	Cycle period
Cell cycle	Budding yeast (Li et al., 2004)	Protein concentrations	Protein concentration pattern
Signal transduction subnetwork	Bacterial chemotaxis	Biochemical parameters *in silico* (Barkai and Leibler, 1997), *in vivo* (Alon et al., 1999)	Tumbling frequency
Metabolic subnetwork	*Escherichia coli* simulation (436 metabolites and 736 reactions; Edwards and Palsson, 2000)	48 central reaction pathways (loss-of-function mutations in enzyme-coding genes)	Metabolic flux ratios that are optimal for growth
Gene regulatory network	Cell nucleus	Signaling, oxidative stress, chromatin remodeling	Gene expression pattern
Multi-cellular development	*Drosophila* (von Dassow et al., 2000; Eldar et al., 2002)	Kinetic parameters	Cell fate patterning
	Arabidopsis (Espinosa-Soto et al., 2004)	Kinetic parameters	
	Caenorhabditis elegans (Félix and Wagner, 2008)	Molecular noise, environmental variation, and loss-of-function gene mutations	
Cell	*Escherichia coli* (Isalan et al., 2008)	Modified regulatory regions in genes	Cell survival
Animal	Tardigrade	Temperature, pressure, hydration	Animal survival
Deme/species	*Arabidopsis* (Fu et al., 2009)	500,000 single nucleotide polymorphisms	Transcript, protein, and metabolite abundance

Robustness is always referred to a system, to a perturbation affecting the system and to the function of the system maintained following the perturbation (here named 'robust property'). The table (from Whitacre 2012) lists several examples of robustness for different biological systems at different hierarchical levels.

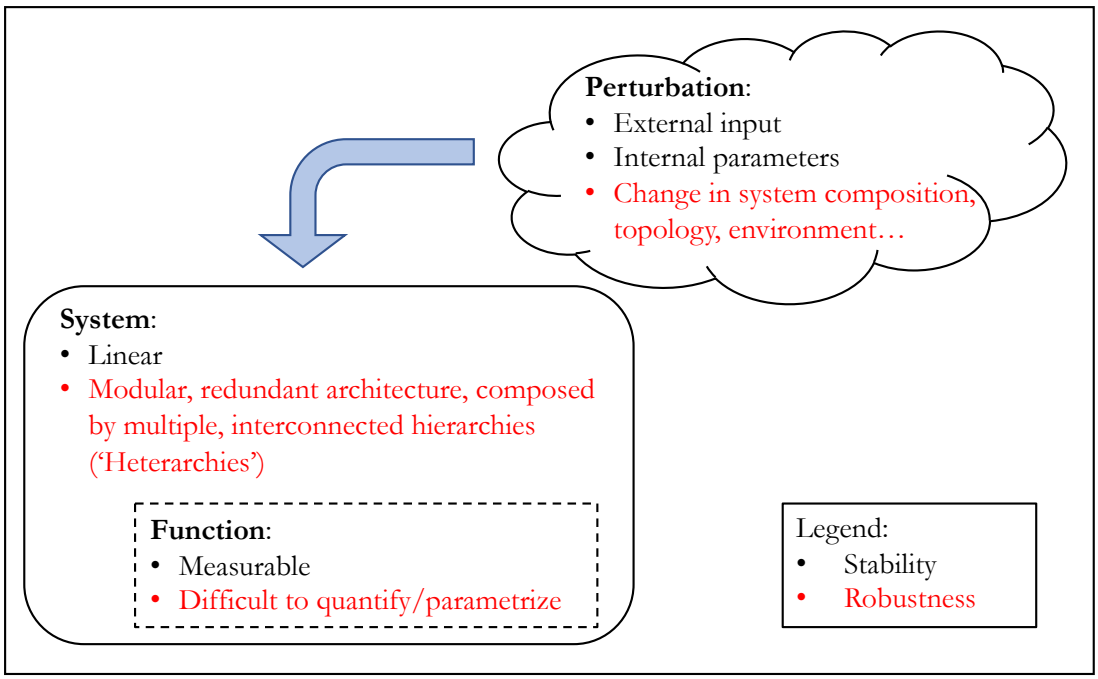


Fig. 5.12 – The characteristics of the elements of the triad system/function/perturbation indicate if the system shows stability (in black) or robustness (in red).

We end this section with two final caveats. The first is to specify the limit of the evolutionary phenomena we consider. The lifecycle of an organism includes various phases, and robustness can appear in each of them: it has to deal with internal and external environmental changes; it has to compete with other individuals of the same or different species for resources; it has to reproduce. We limit our considerations to the relationship between robustness and external environmental changes (including the opening of new niches). The second is to underline that robustness can also decrease following an evolutionary phenomenon, as it will be clear in §5.3.2.

5.2.3. The two aspects of robustness

Kitano's definition of robustness entails two slightly different concepts: a robust system can either return to a same attractor or move to a new one. Robustness under the first alternative is close to the concept of stability, and we will refer to this kind as *stability robustness* or *Rs*. The second concept of robustness entails a change in the state or even in the architecture of the system, and we will refer to this kind as *architectural robustness* or *Ra*. If we apply the concept of robustness to a population, stability robustness is somehow linked to fitness, particularly to Sober's survival fitness: the higher the population average survival fitness, the more stable the population is against environmental changes. Truchetet and Pradeu (2018) use a similar distinction in their study about robustness of the immune system: they talk

about functional and structural robustness. Structural robustness (what we call stability robustness, Rs) entails the restoration of the initial phenotypic structure (e.g. tissue structure after an injury), while functional robustness (what we call architectural robustness, Ra) is reached by restoring the function with a new structure (e.g. a scarf).

The difference between stability, stability robustness and architectural robustness can be grasped by looking at a simple graph mapping changes of some of the system's state variables through time (Fig. 5.13). Note that the three possible fates of the system undergoing the perturbation reflect the ones in Kitano's state space with regards to attractors (Fig. 5.11): the perturbed system either returns to the initial attractor, moves to a new one, or dissolves. The equilibrium ranges of the state variable can be thus interpreted as attractors in Kitano's sense (Kitano 2004).

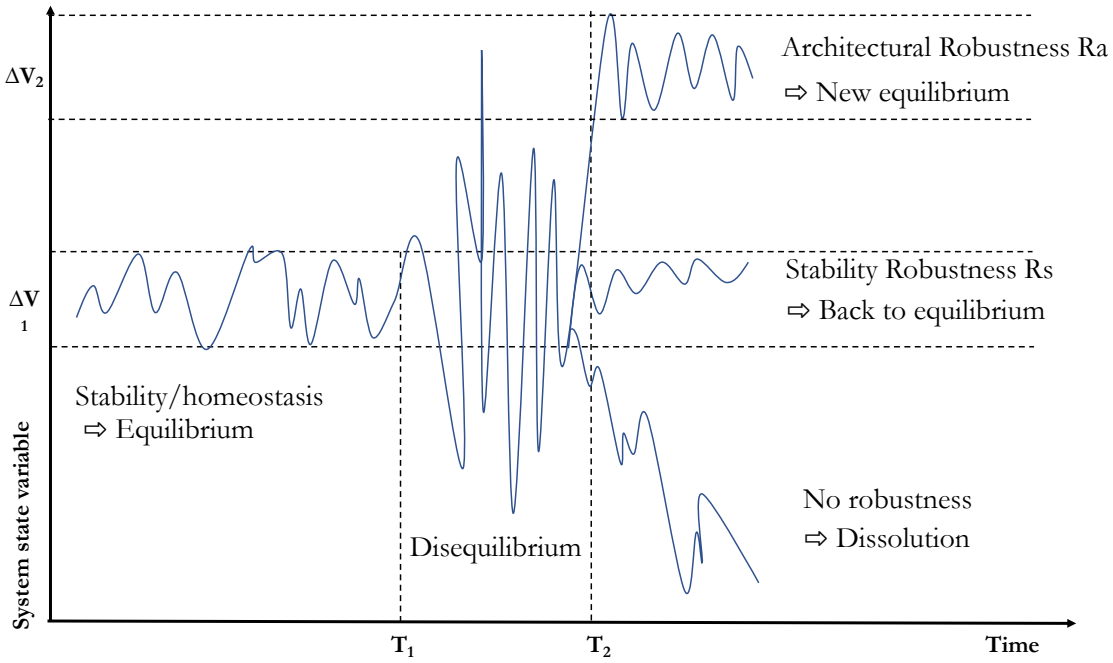


Fig. 5.13 – Graphical interpretation of the concepts of stability, stability robustness and architectural robustness. Before T_1, the system is stable and homeostatic: it responds to small internal and external perturbations by oscillating within a limited range of values $\triangle V_1$, acceptable for the system maintenance. Between T_1 and T_2, the system suffers unusual perturbations that cause a severe displacement from its equilibrium and magnify the range of its oscillations. As a consequence, after T_2, the system can either dissolve, if not robust enough; or return to the state it had before T_1, showing stability robustness; or switch to a different range of stable values $\triangle V_2$ and stay there, thanks to its architectural robustness. The Y axis describes the system through any of its state variables, and the ranges $\triangle V_1$ and $\triangle V_2$ can be interpreted as attractors in Kitano's sense.

The duality of stability robustness and architectural robustness reflects two different and apparently contradictory aspects of robustness: on the one hand, the ability to resist change (Kitano 2004, Carlson & Doyle 2002); on the other hand, the disposition to develop new traits – thus to generate change (Moczek 2008, Wagner 2015, Masel & Trotter 2010). This

duality is captured by Wagner's definition (Ch. 8 in Wagner 2011) describing genetic robustness at the same time as:

(a) the ability to resist changes in the current environment. In this sense, robustness keeps organisms alive. This is what we call system robustness, or robustness *sensu* Kitano;

(b) the disposition to develop new traits, functions and architectures to adapt to new environments (e.g. a new niche). In this sense, robustness fosters changes in the genotype and in the phenotype allowing to explore new solutions to the problem of 'making a living', without losing the current ones. This is what we call architectural robustness, or robustness *sensu* Pigliucci.

Krakauer and Plotkin (2004) underline the same duality when they talk about mutation and robustness against mutation as opposite forces leading at the same time to novel phenotypic variance (required by selection) and to loss of adaptation of preadapted variants. They underline that redundancy (one possible component of robustness) is implicit in the steepness of the fitness landscape in the neighbourhood of its fitness peak. Again, robustness against mutations[190] (which is a kind of stability robustness Rs) reduces selective variance, and thus architectural robustness Ra (Fig. 5.14).

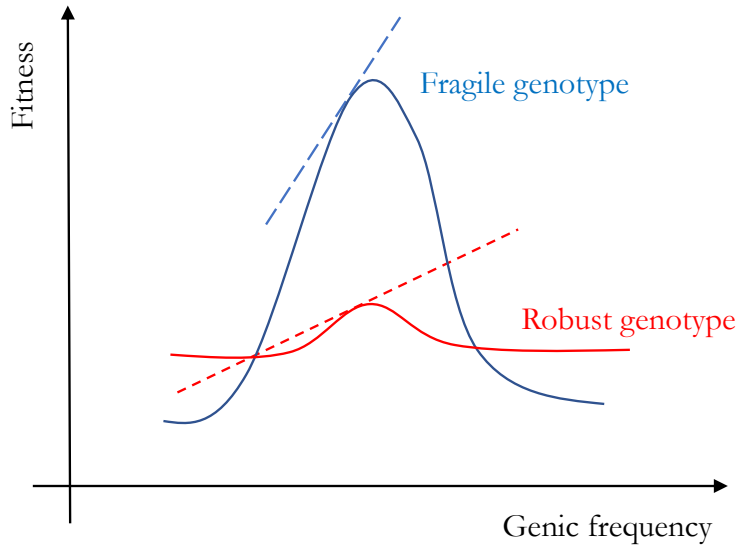


Fig. 5.14 – A way to interpret stability robustness in fitness landscapes: the peak of the blue landscape corresponds to a fragile phenotype because fitness dramatically decreases for small perturbations that moves genic frequency away from the optimum; the peak of the red landscape is robust for the opposite reason. On the other hand, the blue phenotype shows a much higher variance against genic frequencies than the red one, and it might therefore show higher architectural robustness.

[190] For instance, presence of mechanisms to resist somatic mutational changes.

At first sight, thus, increase in robustness and increase in variation are opposite phenomena: the lower the robustness of a genotype, the higher the phenotypic effects of any mutation, thus the phenotypic variability within the population. Mutations, however, are usually deleterious: a little robust genotype allows exploring just a few alternatives, and exclude the possibility of accumulating mutations until a new viable genotype appears[191]. Conversely, a mutation in a robust genotype leads to a (functionally) equivalent genotype in the same genetic network: the potentially explorable genetic space is much wider and, although limiting phenotypic variation in the short run, it could allow new variants in the long run. In other words, a robust genotype reduces the immediately accessible phenotypes, but allows the accumulation of cryptic variations that, in the end, could give rise to radical phenotypic novelties (like Goldschmidt's 'hopeful monsters', Masel & Siegal 2009). The clash between robustness and mutation rate results in a trade-off between reducing and increasing accessible phenotypes, an issue decided case by case (Ch. 18 in Wagner 2005, Ch. 8 in Wagner 2011; Lenski et al. 2006).) Population genetics models describe how mutational robustness and adaptation are linked depending on population size, mutation rate and fitness landscape structure: in populations with low mutation rates, selection favours high fitness peaks with steep cliffs, because mutations pushing offspring off the cliff are rare; in population with high mutation rates, low peaks are favoured for the opposite reason (Draghi et al. 2010, Lenski et al. 2006). Masel and Trotter (2010) focus on the recombination rate as the fundamental driver deciding which of the two robustness tendencies wins.

5.2.4. Mechanisms behind robustness

Biological robustness (either stability and architectural robustness) results from employment of different architectural characteristics, among which:

- Mechanisms of system control, such as positive and negative feed-back (Kitano 2004, Kauffman 1969, Khammash 2016). A feed-back is a loop mechanism by means of which a system responds to perturbations either by magnifying (positive feed-back) or contrasting (negative feed-back) them. In a biological system, feed-back involves some signal that accelerates or inhibits a process. For example, in bacterial chemiotaxis, the movement of bacteria is driven by some chemical stimulus, regulating organismal behaviour either towards a source of energy (e.g. glucose), or

[191] See §4.2.5 for the concepts of genetic networks and equivalent genotypes.

away from poison (e.g. phenol). Negative feed-back mechanisms can also enhance the robustness of a system element, as in self-regulating gene expression circuits.

- Redundancy of some system elements (Kitano 2004, Whitacre 2012). When the same element is present more than once, and each copy can perform the same function (e.g. duplicated organs), the system is protected against the failure of the element, and its robustness is increased. Redundancy has been identified as a fundamental element of robustness in the immune system (Truchetet & Pradeu 2018).
- Diversity of system functionalities, or 'functional redundancy' (Kitano 2004, Whitacre 2012). A lost function (e.g. because of the failure of a system element) can be recovered by performing it in an alternative way (e.g. alternative metabolic routes, such as glycolysis and oxidative phosphorylation to produce ATP). The majority of biochemical networks are remarkably stable against changes in external conditions and internal reaction parameters (Daniels et al. 2008): their robustness might be linked to functional redundancy.
- Modularity (Kitano 2004, Bich 2018, Pigliucci 2008a). Wagner (2015) underlines many common points between nature and technology: among them, the existence of a few standards (e.g. the universal DNA encoding) that allows combinations and recycling of existing modules. Innovation, in life as well as in technology, is basically combinatorial. The majority of biological systems consist of morphological (e.g. cells), functional, developmental and temporal modules, where perturbations and damage of a module are limited locally, and do not spread to other areas (Wagner et al. 2005, Callebaut & Rasskin-Gutman 2005). When the damage involves too many disconnected modules, however, the system might collapse (such as what occurred to the Titanic with its watertight modules).
- Decoupling among system levels (Kitano 2004, Rutherford 2000, Masel and Siegal 2009). By isolating low-level variations in the system elements from high level functionalities, these can be protected against many perturbations affecting the elements, as in, for instance, in genetic and regulatory buffering. The Hsp90 protein, for example, constitutes a buffering that insures that environmental shocks and new mutations do not affect development (Jablonka and Lamb 2005). Similar processes protect the individual and the population against random mutations and environmental insults (e.g. changes of external temperature during transcription linked to competing activation/repression factors), and maintain genetic diversity.
- 'Cooperativity between parts': "[f]or example, at the biochemical level, one step

molecular processes will always suffer a high error rate, but when multiple simultaneous or sequential events are required, it is highly unlikely that they all contain errors" (Masel & Siegal 2009: 397).

- Switch-like responses or 'transistor effect'[192], instead of graded ones. An external perturbation does not cause any change until it reaches a threshold, thus protecting the organism in most cases (e.g. the response to several signal of the Saccharomyces cerevisiae mitogen-activated protein kinase (MAPK) cascade, Masel & Siegel 2009).

Robustness is especially relevant during developmental processes, which are often regulated by peculiar mechanisms insuring the reduction of phenotypic variance against perturbations (Krakauer and Plotkin 2004):

- Canalization, or adaptive suppression of variation during development: it describes the ability of the developmental system to suppress phenotypic variability occurring as a consequence of mutations or environmental insults.
- Neutrality, or selective equivalence of phenotypes given selective neutrality of alleles. It does not mean that different alleles produce the same phenotype, but only that they do not produce differences with any selective significance.
- Redundancy, or the overlap of gene function: it captures the dependency of some trait or function of the phenotype from more than one gene, so that the removal of one of them has no selective consequence for the individual.

All these architectural characteristics can be identified, and measured, before the selection stage: for instance, the robustness of different architectures can be estimated and, on this basis, reasonable comparative judgement can be made[193], something that the propensity interpretation of fitness, for example, cannot do[194].

[192] A transistor is an electronic element whose output voltage is null until a threshold input voltage is reached.

[193] The issue of how to practically measure robustness is of course not trivial, and it remains outside of the scope of our research.

[194] Propensities (applied to fitness or to probabilities in general) present several problems related to their real cognitive content, e.g. calling a property a 'propensity' does not explain what the property is (Hitchcock 2002); propensities are as explanatory as Molière's 'dormitive virtue' (Sober 2000), and propensity based accounts are metaphysical rather than scientific (Gillies 2000).

5.2.5. Types of robustness

There are two crucial questions concerning robustness, linked to the mentioned triad system/function/perturbation: *what* is robust, and *against what* is it robust? Because "a system is not robust in general; rather, it is robust *vis a vis* a certain kind of perturbations that can occur at a given level (or at a limited number of levels)" (Truchetet & Pradeu 2018: 45). Different answers to these questions produce different possible dual classifications of robustness.

Individual and systemic robustness

Robustness can be referred, the same as fitness[195], to a single trait of an organism (be it genetic or phenotypic), to a collection of interacting traits, to the whole organism or even to supra-organismal biological system such as, for instance, a whole ecosystem[196] (Lenski et al. 2006) (Table 5.2).

Table 5.2

Robustness of	Fitness equivalence	Example		Measure
		Phenotype	**Genotype**	
a trait performing some function	ecological (or trait)	single metabolic enzyme robust against changes of temperature	proofreading and repairing systems	range of external variable in which the trait's function is maintained
collection of interacting traits/whole organism	reproductive (or individual)	alternative metabolic routes	genomic redundancy	probability of survival of the system to an external perturbation

Trait and system robustness.

The two kinds reflect Wagner's distinction (2005, Ch. 15) between redundancy (e.g. duplicated elements) and distributed robustness, or alternative ways to perform the same function (e.g. when a metabolic reaction is lost, another takes its place): the first is typical of trait robustness, while the second is a character of a whole system. Wagner gives distributed robustness a much greater importance than redundancy. The distinction between trait and

[195] See §3.6.2 for the concepts of trait fitness and individual fitness.

[196] "An ecosystem might be robust if it resists the extinction of some keystone species or, if extinction does occur, because surviving species can compensate over physiological, demographic, or evolutionary time scales." (Lenski et al. 2006: 2190). Such robustness at the ecosystem level could be based, for instance, on species redundancy.

system robustness also mirrors the two complementary engineering approaches for improving robustness: the one that focuses on individual functions and performances, and the system-oriented one that focuses on collective properties (Bich 2018).

As a property of a collection of interacting traits, robustness measures the probability that the system survives external perturbations. For example, metabolism, a collection of biochemical reactions facilitated by enzymes, is robust with respect to the failure of some of its reactions or lack of some metabolite thanks to the existence of alternative metabolic pathways (van Dongen et al 2011). A genotype can be robust to mutations thanks to redundant genes.

As a property of an individual trait, robustness measures its ability to continue performing its functions against a range of values of an external variable. For example, a gene is robust against duplication errors thanks to the existence of proofreading mechanisms improving fidelity by several orders of magnitude[197] (Bębenek & Ziuzia-Graczyk 2018), or an enzyme is robust if it maintains its functions against a range of environmental temperature changes. Robustness of individual elements is widespread at all organizational levels, including "protein folding, gene expression, metabolic flux, physiological homeostasis, development, organism survival, species persistence, and ecological resilience" (Whitacre 2012: 1). Examples of trait's robustness increase (taken from Wagner 2015) are found in the artic code *Boregadus saida*, a species living 100 meters beneath the surface at 84º north latitude: a modification of its antifreeze proteins allowed a new version to survive 100 miles further north; and in the *Anser indicus* (bar-headed goose), that is able to fly over the Himalaya when migrating, at altitudes higher than 10,000 meter, thanks to a special haemoglobin molecule.

Genetic and phenotypic robustness

Another alternative classification (Table 5.3) distinguishes between robustness of the genotype and or the phenotype[198] (Daniels 2008), a distinction that, again, mirrors the analogous definitions of fitness (as seen in §3.7).

[197] The level of fidelity of course depends on which kind of proofreading process one takes into account, with DNA replication's proofreading being more faithful than transcriptional proofreading which in turn is more faithful than translational proofreading etc.

[198] Classification which also reflects the double interpretation of fitness as genetic and phenotypic.

Table 5.3

	Focus	Measure	Interpretation
Phentypic Robustness	maintain function against internal and external perturbations	average effect of perturbation	ability to resist change
Genotypic Robustness	maintain function against genetic perturbations	probability of neutral mutations fraction of neighbour genotypes in the genotype network that are phenotypically neutral	disposition to develp new traits

Phenotypic and genotypic robustness.

A similar, but subtler, distinction is suggested by Jen (2005), who talks about mutational robustness, or the ability of a system to withstand perturbations in structure without changes in function, and about phenotypic robustness, or the ability of a system to perform different functions when needed without permanent changes in structure (a concept similar to phenotypic plasticity). While echoing the functionalist-structuralist clash, these definitions can be applied to our architectural matrix[199] to describe horizontal or vertical evolutionary phenomena (Fig. 5.15).

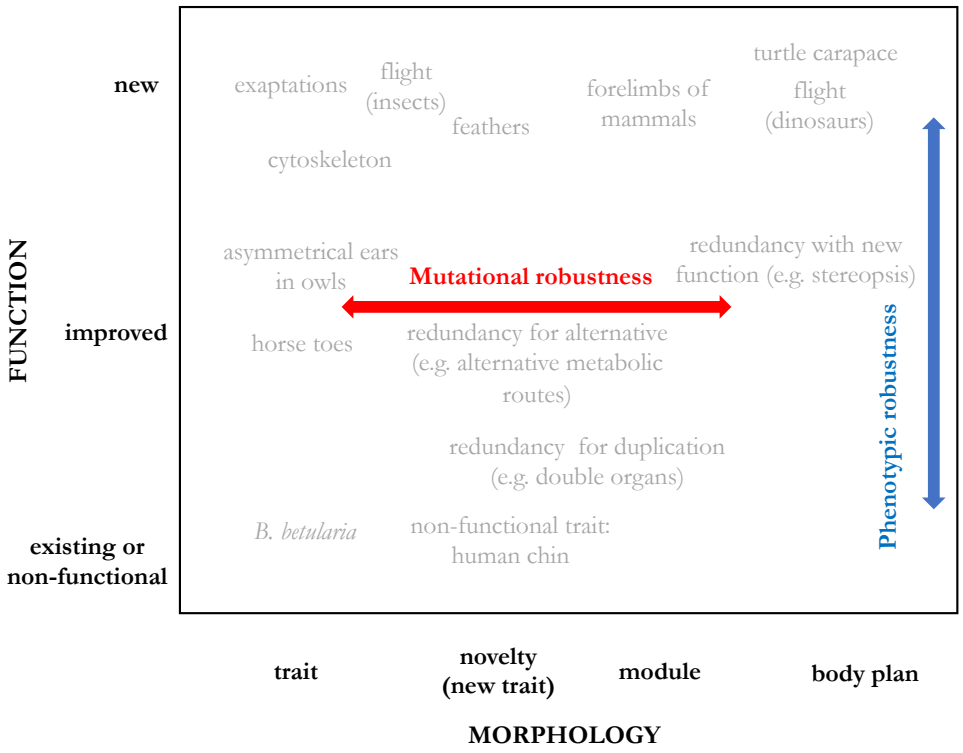


Fig. 5.15 – Robustness and architecture. Mutational (or genetic) robustness is the ability to maintain the same function against changes in structure. Phenotypic robustness is the ability to perform different functions with the same structure (Jen 2005).

[199] See §2.4.2 and §2.4.3

Hermisson & Wagner (2005) define phenotypic robustness as the sensitivity of the phenotype to changes in the genotype (due to mutations and re-combinations), in the external environment (temperature and other macroscopic variables) and in the internal development system (developmental noise, such as stochastic gene expression and other sources of cellular noise[200]). Each aspect is independent from the other, thus we could define three different kinds of phenotypic robustness. Moreover, a reference point is needed against which robustness can be measured. Several measures have been proposed. Kitano (2007) suggests measuring it in terms of the average effect of the perturbation on the population individuals. Being a measure of the phenotype's ability to resist change, it matches the first side of Wagner's dual characterization of robustness.

Genetic robustness can be defined as a state of reduced variability of the genotype's expression pattern with respect to mutations and re-combinations (Masel & Trotter 2010). In this sense, it is similar to the concept of 'genetic canalization'[201], and it is part of the mechanism allowing the accumulation of cryptic genetic changes and exploration of neutral network (Hermisson and Wagner 2005), matching the second side of Wagner's characterization of robustness: the disposition to develop new traits and functions. It can be measured in terms of the fraction of neighbour genotypes in the genotype network that are phenotypically neutral (Wagner 2005), or probability of neutral mutations (Draghi et al. 2010), or through the function linking the percentage of offspring retaining some trait to the global mutation rate: the smaller the slope of the function, the more robust the genotype (Voigt et al. 2005) (Fig. 5.16).

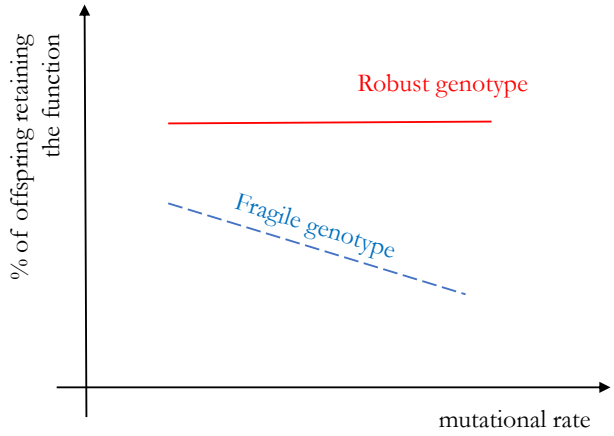


Fig. 5.16 – Genetic robustness as sensitivity of function to mutation.

[200] The importance of noise in development is testified, e.g., by the differences in fingerprints between genetically identical twins. The first cloned cat displayed very different coat patterns and personality than the original (Shin et al. 2002).

[201] See §4.3.2 for a definition and examples of genetic canalization.

Based on the combination of these two dyadic classifications, we can identify several kinds of robustness (Table 5.4).

Table 5.4

	individual	systemic
genetic	duplicated gene	alternative metabolic routes
phenotypic	duplicated organ	phenotypic plasticity

Examples of robustness.

One last remark: the organism, regardless of whether we focus on it as a whole or as a collection of individual traits, on its genotype or on its phenotype, represents an integrated, complex system composed of interacting parts, and all these partial views are thus linked among them. Traits' robustness is obviously a part of system robustness, just as the ecological fitness of traits is related to reproductive fitness . Phenotypic and genetic robustness are likewise linked. One could then wonder whether these distinctions have any cognitive content at all. The answer lies in the importance of the structure of the genotype-phenotype map, of the developmental system, and of the organismic architecture. Traits can interfere with each other, or result in emergent properties not deducible from their characteristics: once again, the whole organism is more than just the sum of its parts.

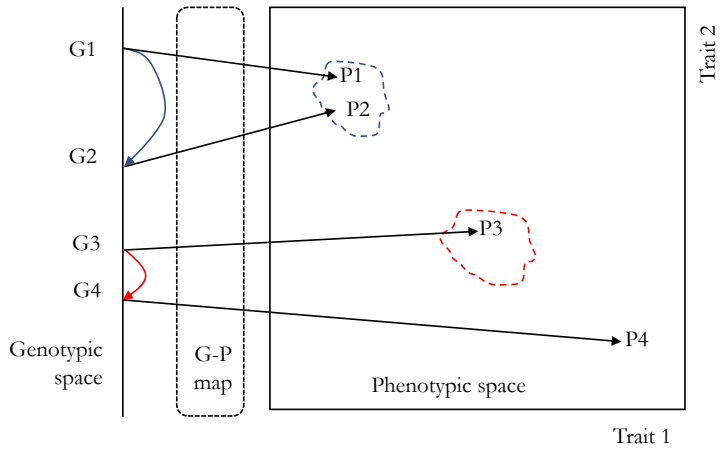


Fig. 5.17 – Robustness and the genotype-phenotype map. A mono-dimensional genotypic space expresses a bi-dimensional phenotypic space. The great change from G1 to G2 due, e.g., to mutation, causes only a small displacement from phenotype P1 to P2: both can perform the same function (all phenotypes included within the blue dotted line can). The genotype is robust. The opposite happens for the change from G3 to G4: the genotype is fragile and phenotypic function from P3 to P4 is lost (as P4 lies outside of the red dotted line) (Figure partially based on Taylor 2018).

Phenotypic robustness, for example, can be more or less linked to genetic robustness (which, by definition, is part of it) depending on these structures: a robust genotype can result in fragile phenotype and the other way round (Fig. 5.17).

5.3. Design Space

After showing the limitations of fitness as a sole measure of architectural changes in some evolutionary phenomena (§5.1), and after introducing robustness as a second dimension of biological design (§5.2), we present in this section a model that applies these two variables to map populations' evolution. It is a generic model in the sense suggested by Parker and Maynard Smith (1990): contrary to specific models, that are designed to be applied to a particular species, and include parameters easily measured, generic ones have a heuristic function and give qualitative insights on a problem.

We firstly present the elements of the model (§5.3.1), and then describe how the evolutionary forces listed in chapters 3 and 4 impact on these elements, alone or in combination (§5.3.2). After showing in what sense the design space can be interpreted as an extension of Fisher's fundamental theorem (§5.3.3) and of fitness landscapes (§5.3.4), we apply the developed concepts to provide a classification of evolutionary phenomena (§5.3.5).

5.3.1. Dimensions and elements

The main elements of the proposed bi-dimensional design space are sketched in fig. 5.18**.**

The X-axis represents fitness F, while robustness is mapped on the Y-axis: each point of the area thus represents the fitness and robustness of an individual, or of a trait of the individual, at a given time[202]. A population is represented by a cloud of points (Provine 2001)

In this design space, we can write:

$$\text{Design} = f\,[\text{fitness, robustness}] \qquad (1)$$

[202] We talk here of individual fitness and robustness as opposed to average values. 'Individual' has thus here a different semantic content than in 'individual fitness' as opposed to 'trait fitness': we can have individual fitness of an individual organism and trait fitness of an individual. The same is valid for robustness.

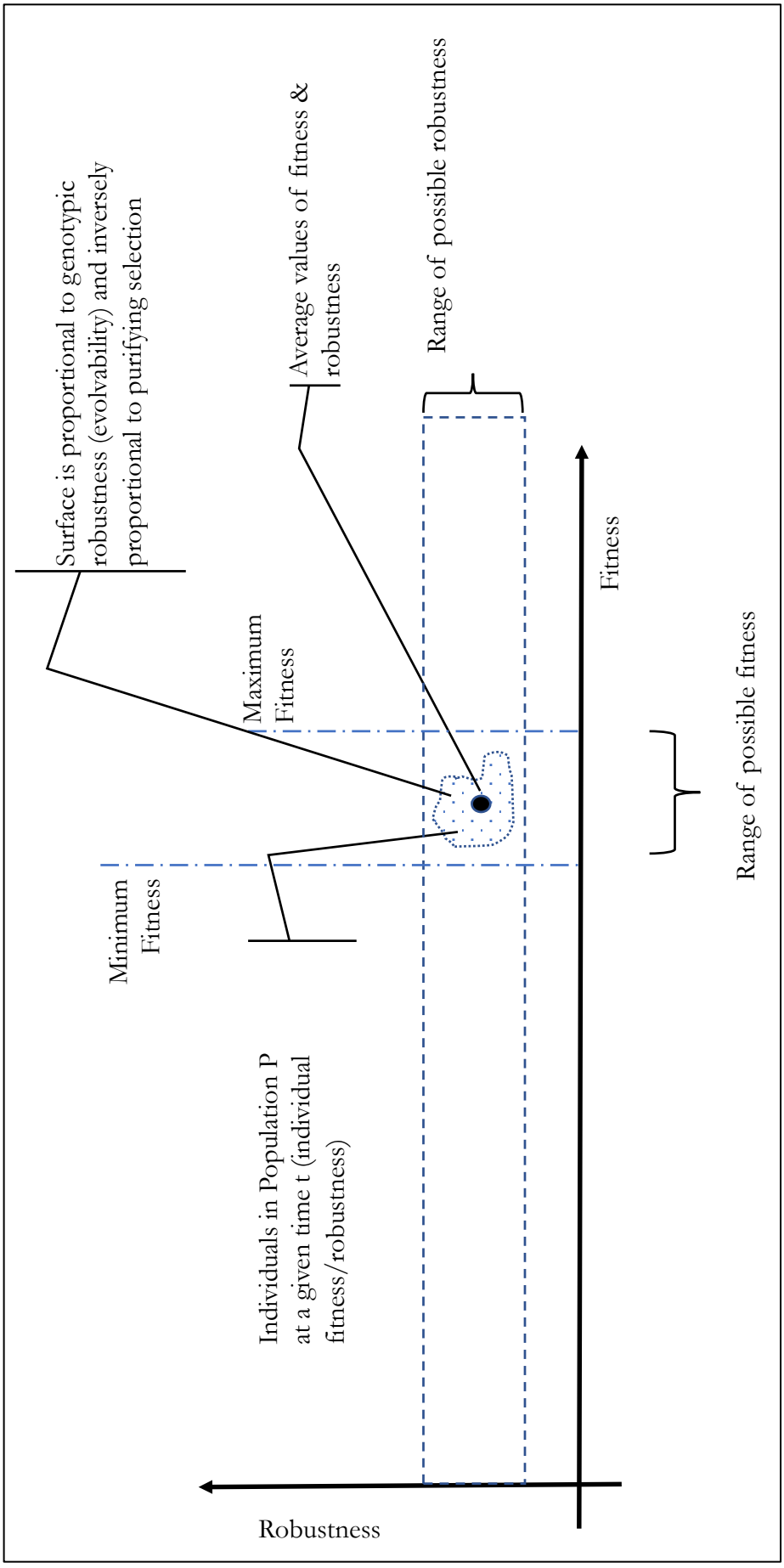


Fig. 5.18 – The design space (see text for detailed explanation of elements).

Fitness depends on the current traits of the organism and their mix, while robustness on the organism architecture (as we have defined it in §2.4.2):

$$\text{Design} = f\ [\ \text{fitness (current traits)}, \text{robustness (architecture)}\ ] \qquad (2)$$

In general, individuals belonging to the same population[203] have similar values of fitness and robustness, and aggregate in limited geographical areas. These areas representing populations at a given time show some peculiar characteristics:

- An average fitness and an average robustness. Their positions depend on how they are defined. Averages are a good tool to track changes in a population and to estimate how these changes are statistically linked to evolutionary forces. The tendency to see populations as homogeneous, or the 'tyranny of the golden mean' (Bennet 1987), reveals however to be a poor way for understanding the reasons behind these links, that only the analysis of the individual organism can provide. The importance of average values in the design space is therefore mainly to provide a translation of the results of population genetics models and of optimization programs to the design space.
- A range of fitness. Maximum theoretical fitness of any individual is determined by genetic, historical and developmental constraints. Note that these are different from the constraints to fitness maximization found in population genetics, which involve averages. For example, in a diploid population with homozygote advantage, the population will never be composed uniquely by homozygotes, and theoretical average fitness will be lower than homozygote fitness (which is the maximum theoretically possible). Homozygote fitness, however, is limited by individual constraints of a different nature.

  Minimal fitness is not set by constraints but by the selective process; individuals with any value of fitness below the maximum can appear, e.g. thanks to mutations, but we can assume that they are rare and scattered on the left of the main population area. The higher the strength of selection, the smaller the range of admitted fitness values. The lower point in the range of admitted values of fitness mirrors the strength of purifying selection: the lower the selection strength, the lower the admitted fitness values ('non-survival of the not-enough fit' instead than 'survival of the fittest').

[203] We mean by "population" a group of interbreeding individual organisms of the same species sharing the same geographical territory at the same time.

- Upper and lower limit of robustness are linked to the trade-off between the advantages of a higher robustness and its costs (Wagner 2015, Lenski et al. 2006). Mammals and bacteria, for example, have opposite strategies: the former dedicate many resources to maintain high robustness during development and daily life, the latter adopt simpler robustness mechanisms demanding little resources, but can reproduce much quicker (Kitano 2004). As we have already discussed (§5.2.1), an individual with too low a robustness does not survive, while one with too high a robustness spends energy in traits useless in that particular environment: costly traits are lost quickly (Lahti et al. 2009). Proofreading is good, but the excess cost of making it perfect to avoid any mutation is not worth the cost (Lewontin 1987). Both too high and too low robustness tend to be eliminated by selection.
- It is important to bear in mind that the maximum and minimum values of fitness and robustness are theoretical: the actual region occupied by the population might be smaller, i.e. there are not necessarily individuals showing these values.
- The actual region somehow depends on the genetic robustness of the population: the highest it is, the farther the population can spread on the Y-axis around the average robustness. This also represents in some sense the 'evolvability' of the population (in the sense of Pigliucci 2008a), or the population potential to evolve towards new phenotypes. The area also depends on purifying selection. A population subjected to weak purifying selection would allow for a wider range around average fitness on the X-axis. "[…] populations with a low intensity of purifying selection […] possess the highest evolutionary potential" (Koonin 2012: 283): which translates, in our model, in a wider area, as it will become clear in the next section.
- The theoretical region is somehow linked to the concept of an attractor in the sense of complex system theories (Kaufmann 1991, 1993, 2000) or in Kitano's sense (Kitano 2004, 2007): individual organisms within this area are stable, thus viable. The idea of stable states as pillars for bridges allowing movement in a space of possible but unstable states is also implicit in Simon's description of evolutionary complex hierarchical systems (Simon 1962). Kitano's attractors also echoes the idea of 'possible morphologies' proposed by the evo-devo perspective: "[…] the structure of a developmental process may be such as to give rise to a series of *bounded discrete morphologies without intermediates*" (Maynard Smith et al. 1985: 266, italics mine). Only regions of the theoretical space of possible morphologies that have a modular design are filled by actual organisms (Rasskin-Gutman 2005). All such morphologies can be

considered as belonging to the same attractor. In the design space, an attractor is therefore an area within which a population can exist at equilibrium, performing a set of functions, or that can explore when pushed by evolutionary forces, provided that these do not cause any change in functions (recall that Kitano defines functional stability as a property of an attractor). It is a neutral space that includes a collection of equivalent solutions to the same biological problem (Wagner 2005, Ch. 13). If phenotypic function changes (or a new function appears), the population moves away from the original attractor to a new one. In §5.3.4, we develop the idea of attractors in the design space and relate it to fitness landscape, while in §5.3.5 we relate different kinds of evolutionary phenomena with movements within and between attractors.

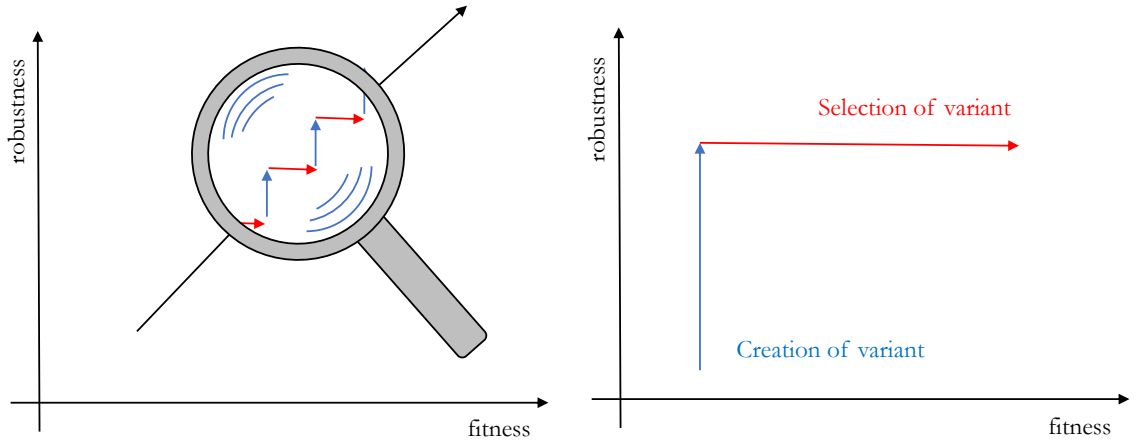
5.3.2. Dynamics of evolutionary movements in design space

The area representing a population will usually change generation after generation: its surface, borders and shape, as well as the position of its average fitness and robustness, will move around the logical design space. In this section, we analyse different types of movements and what they represent in evolutionary terms. We will also analyse how selection and other non-selective forces cause these movements, given that each of them has a peculiar impact on the average and the variance of one or of both dimensions: reproductive fitness increases if reproductive success augments; phenotypic robustness increases if any modification in the architecture of the phenotype improves survival.

Our model is based on two postulates: that changes in fitness and change in robustness are independent, and that the first are mainly due to selection while the later do non-selective processes. Let us illustrate each postulate and their rationale.

In the first place, we postulate that changes in fitness and changes in robustness are quite independent, because we interpret evolution by natural selection as a two-steps process: creation of new variations followed by selection of new variations A new trait or a new function can leave the organism better or worse off with regards to its ability to survive, that is, with regards to its robustness: this has nothing to do with the environment. Fitness only comes into play at the (logically) later stage of selection, when the usefulness of the trait is judged against the current or a new environment. Evo-Devo, for example, aims at explaining phenotypic changes due to alterations in developmental processes regardless of whether they are adaptive or not (Müller 2007). Independence is usually neglected in simulation studies,

that at best assume some kind of equivalence between fitness and phenotypic stability (Whitacre 2012). The assumption of independence is neutral with regards to the positive/negative view debate[204]. Neader (1995), for example, arguing in favour of the former view, thinks that the cumulative process of creation/selection is itself a fundamental piece of the generation of novelties, but does not deny that the two steps exists: novelties are built by a series of vertical and horizontal movements (increase in robustness and in fitness), although each very small. Wagner (2015: 176) defends a similar position (in the sense of distinguishing the two steps of creation/selection, not in defending the positive view) when he claims that genotype networks allow the exploration of new phenotypes, not their conservation: it is selection that preserves good variants (Fig. 5.19).



(i) Positive view (ii) Negative view

Fig. 5.19 – The difference between the positive and negative view of selection appears to be just a 'scale issue' in the design space. Both approaches consider that selection is a sorting process from an existing set of variants, whose origin is prior to selection. The difference is that the positive view (fig. i) claims that these variants differs very little from each other and from the previous generations ones, and are due mainly to random mutations and recombination. Evolution advances as a swift line with no perceptible steps. The negative view (fig. ii) defends that variants also appear thanks to mechanisms other than simple random mutations, and that can differ substantially from each other and from the previous generation ones. Some of these mechanisms are described in Ch. 4.

Secondly, and coherently with our discussion about the nature of fitness as a measure of selection[205], we also postulate that consistent and directional[206] changes in fitness are mainly

[204] Recall that the positive view claims that selection has creative power, while the negative view denies this (see §3.3).

[205] See §5.1.2

[206] By consistent change of one of the variables by a force, we mean a change that always happens whenever the force is present; by directional, a change that always affects the variable in the same way (increasing or decreasing it).

linked to selection, while changes in robustness are mainly due to non-selective processes. Again, this postulate is based on our interpretation of fitness and robustness, and on the view of evolution as a two steps process. What impacts the robustness of an organism is a change in its architecture, regardless of whether the change is in the sense of increasing or decreasing it. Selection has no role in the architectural change itself, and fitness, which is a measure of selection, not of creation, of variants, is unaffected by processes affecting robustness. For example, a new trait (e.g. a duplicated organ) might appear as a consequence of cryptic genetic changes, thus affecting phenotypic robustness; but whether the trait will get fixed or not will be decided by natural selection (Gregory 2008). Moreover, the same change might get fixed in one environment and might be eliminated in a different one. Movements along the horizontal axis towards higher fitness are linked to an *efficiency* increase[207] (Fig. 5.20): the organism fulfils the same functions in the same way, with the same traits or modified traits, and the individuals performing it in the most efficient way are selected. The case of the *B. betularia* is paradigmatic in this sense: the change in the wings' colour improve the efficiency of the mimicry function. Black wings allow the same function (mimicry) than white wings, but in a much more efficient way. Movements along the vertical axis towards higher robustness are linked to an *efficacy* increase: the organism realizes the same function in new ways, or new functions thanks to new traits' versions, new traits, or new uses of existing traits (as in exaptations). Darwin's finches are an example of the development of new trait's versions (different beaks in different islands) that allow the individual to realise the same function (feeding) in new ways (accessing previously inaccessible foods). The new beak form is more efficacious in the new environment.

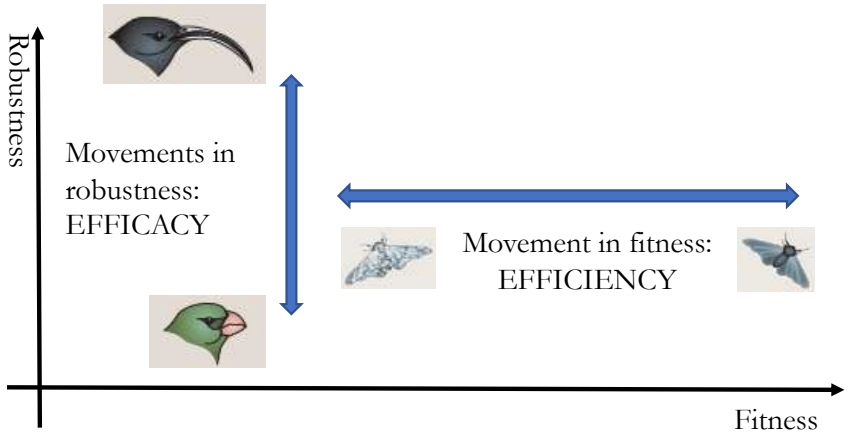


Fig. 5.20 – Efficacy and efficiency increase in the design space.

[207] We define *efficacy* the ability to solve a problem, and *efficiency* the ability to do so getting the best result using the least possible resources.

Natural selection will decide whether the new functions bring any competitive advantage or not. We can borrow Dennet's idea about the 'shift and drag' of organisms caused by selection (Dennet 1995): selection does 'drag' them towards higher fitness (increasing efficiency), but the 'shifting' (increasing efficacy) is mainly due to other evolutionary forces and is reflected in robustness increase.

The time scales across the two axis are also very different. Movement along the fitness axis represents 'short-term evolution', i.e., the output of the recombination of the existing genetic pool, of mutations, migrations, etc. Movement on the robustness axis represents 'long-term' evolution. Short-term evolution takes over after long-term evolution has produced a new evolutionary context (Eshel & Feldman 1984).

Adaptationism could of course deny the validity of both postulates. Regarding the independence of robustness and fitness, adaptationism would claim that the former is irrelevant as long as it is fitness that decides what traits pass to the following generation. This way of considering fitness has been already criticised in the first section of this chapter (§5.1), as it focuses on the results of the 'selective game', without inquiring on the reasons behind them: that is, the reasons behind the successful design. Regarding the second postulate, it could be accused of being trivial: if robustness is a property of organisms and changes in robustness are, at least initially, fitness-neutral (e.g. developmental changes), then selection cannot be involved in their production. Its putative triviality, however, only appears once we admit robustness as relevant variable, an idea that the adaptationist would hardly buy.

In the rest of this section, we shall analyse how selection and non-selective forces impact movements along each axis, and how they affect the average and the variance of each variable (Table 5.5). We also suggest a way to interpret, through an example, how multiple forces can combine to cause complex evolutionary phenomena.

Table 5.5

Force	fitness		robustness		Directionality: always increasing average F/R?	Continuous vs discrete increase of average F/R?
	average	variance	average	variance		
selection	↑	↓	0	0	yes, F	continuous
drift	↑↓	0	↑↓	0	no	continuous
ZFEL	0	↑	0	↑	no	N/A (1)
ex-ante exaptations	N/A	N/A	↑	N/A	yes, R (2)	discrete
complex systems laws	N/A	N/A	↑	↑	yes, R	discrete

(1) averages do not increase
(2) robustness increases, or it would not be an exaptation

Impact on the design space of different evolutionary forces. For each evolutionary force, impact on the values of average and variance of fitness and robustness is shown: upward arrow indicates that the value increases, downward arrow indicates that it decreases, and double arrow indicates that the value can either increase or decrease; a '0' indicates that the value does not change; N/A indicates that changes in the value are possible but casual, and not linked primary to the force. Directionality refers to the direction of the change in averages: some forces always cause some of the averages to increase, while others can either increase or decrease them. The last column shows whether the changes are continuous or happen as 'jumps'. See text for detailed explanation of the table for each force.

Natural selection

We suppose that consistent[208] movements towards higher average fitness are basically due to selection. According to Fisher's fundamental theorem[209] (Fisher 1930, Price 1972), when no other forces act and in a constant environment[210], the average fitness of a given population never decreases. Grafen's model[211] suggests that fitness tends to increase until the maximum allowed by genetic, historical, and developmental constraints. At the same time, fitness variance decreases: this is immediately clear if we consider that the number of individuals with lower fitness decreases and the ones with fitness close to the maximum increases.

Regarding robustness, we suppose that, under the action of natural selection alone, it remains constant. We follow here Kitano (2004), who suggests that robustness is a conserved quantity whenever the organization, regardless of the changes it undergoes, maintains its

[208] Fitness changes due to other phenomena (e.g. drift) might of course increase the population average fitness: but these increases are casual, and thus non-consistent.

[209] See §2.2.3

[210] If these hypotheses are not fulfilled, average fitness might of course also decrease; for example because of changes in the environment, as in the example of the change of *B. betularia*'s wings' colour, discussed hereafter, or in case of frequency-dependent selection, when fitness of a trait depends on how frequent is the trait in the population (see §2.2.3).

[211] See §3.6

functions (a situation typical of 'short-term evolution'). When robustness remains constant, architecture does not change either, and evolution consists of the optimization of the mix of existing traits. This is the case when evolution is fuelled by selection alone. Equation (2) reduces to:

$$\text{Design} = f\,[\,\text{fitness (current traits)}\,] \quad (3)$$

that is, the basic equation of all adaptationist models: fitness can in these cases satisfactorily capture the essence of evolutionary phenomena (compare with Fig. 5.2). This is the reason why robustness can be ignored in traditional models of population genetics and in optimization programs (e.g. Grafen's project) (Eshel Feldman 2001, Wilkins & Godfrey Smith 2009), where design is reduced to a simple optimal mix of existing traits produced by selection through the optimization of their combined fitness. Some typical phenomena characterised by increasing fitness and constant robustness are the spread of a new, beneficial trait (however first appeared), the fixation in a population of the best mix of existing traits, or the optimization of an existing quantitative trait, that Erwin (2015) calls 'adaptive refinement', e.g. the dimension of the rhino's horn (Fig. 5.21). The rhino's horn fulfils several functions, among which intimidation, sexual attractiveness, defence, and can be used as a tool for digging to search for eatable roots and water, for testing thickness of mud before entering a mud hole, for steering the young ones (Owen-Smith 1984). All these functions can be performed with horns of different lengths, so, once the horn has appeared, its morphological changes are linked to improvement of efficiency at constant robustness: adaptationist models could estimate, if information about the link between fitness and morphology is complete, the horn's optimal dimensions[212].

[212] If a new function could be assigned to the horn, robustness would of course increase following the exaptation; but the eventual subsequent improvement in fitness could not be forecasted a priori by the adaptationist model. New historical information should be gathered a posteriori to update the fitness function, thus changing the original architectural 'black box' with a new one, and a new narrative should be created to justify the appearance of the new function as triggered and driven by selection.

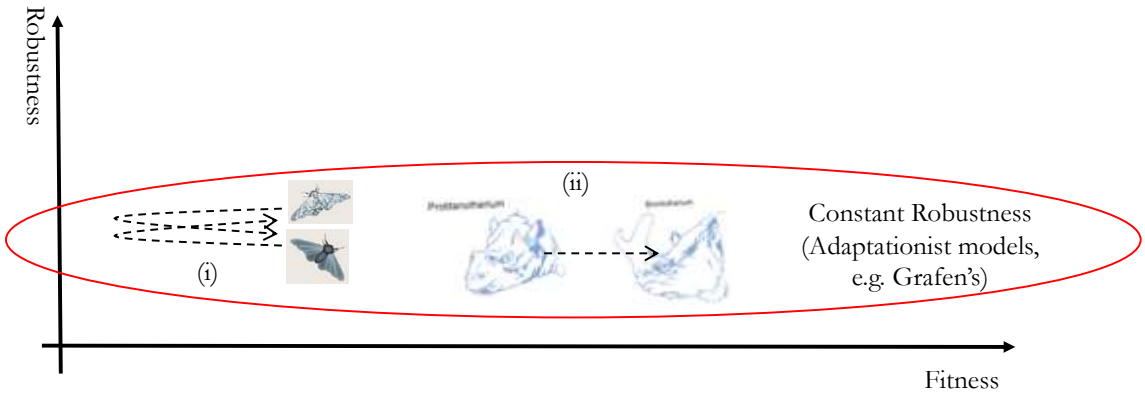


Fig. 5.21 – Mapping in the design space of evolutionary phenomena with constant robustness. (i) Answer to environmental challenge: the *B. betularia* wings' colour switch. (ii) Quantitative trait's optimization: the Rhino horn.

We shall now analyse in some detail the case of the change in the wing's colour of *Biston betularia*. The *B. betularia*, or peppered moth, represents one of the most studied cases of rapid evolutionary change in a changing environment (Cook & Saccheri 2013, Van't Hoff et al. 2016). The typical phenotype in Britain before industrialization had white wings with a sprinkling of black marks, while the melanic form (*carbonaria*), with almost uniform black wings, and the intermediate version (*insularia*) were residual. By the end of the nineteenth century, the original type had almost disappeared in some industrialised areas in favour of the melanic form, to reappear in recent time following de-industrialization and environmental policies[213], while the melanic form has always remained marginal outside of industrial regions. Among the many reasons proposed to explain this evolutionary pattern, selective predation is considered the main factor: darkening of poplars' bark due to pollution makes the white form more visible to predators, while it allows the melanic form enhanced camouflaging. This text-book case is a good example of an evolutionary change that can be described and forecasted through population genetics equations and fitness alone. The trait under selection and its versions (white and melanic) already existed at the beginning of the evolutionary phenomenon, and were not modified by it. The phenotypic architecture (form and functions) did not undergo any change: robustness was thus maintained. Each morphotype showed a fitness that determined its relative frequency. The evolutionary phenomenon was triggered by an external factor: an environmental change in some regions of Britain (i.e., the darkening of poplars' bark due to pollution caused by industrialization). All hypotheses behind population genetics, adaptationism and optimization programs are

[213] Heterozygote advantage limits the power of selection to eliminate one of the morphs, avoiding that either reaches 100% of the population.

met: externalism (the environment determines the fitness of the phenotype); natural selection as the unique process behind the fixation of the change; fitness alone as a full proxy of the change (thanks to the maintenance of architecture); evolution as change of genetic frequencies in a population.

Let us analyse these changes in the design space. The external change caused the fitness of the white type to decrease, and the fitness of the melanic type to increase. Consequently, in a few generations, the relative frequencies of the two types changed as well. In theory, abstracting from details and additional elements affecting the process, we can identify three steps in this phenomenon (Fig. 5.22 shows the steps in the proposed design space; numbering of the following list corresponds to numbering in the picture):

i. Initial equilibrium: the white type (*typica*) has a frequency F and the melanic type (*carbonaria*) a much lower frequency f, maintained by mutation rates, migration and the like. Relative frequencies are determined by relative fitness. Average fitness of the whole population is Fi.

ii. Temporary disequilibrium. The environmental change due to pollution causes the darkening of poplars' bark, triggers the radical decrease of the white type fitness and the concomitant increase of the melanic type. For the argument's sake, let us suppose there is a fitness swap. Consequently, white type and melanic frequencies switch: the first changes from F to f, and the second from f to F. The initial consequence of the new situation consists in a drastic reduction of the total number of individuals of the population: the more abundant white type is easier to spot for birds; therefore, *typica* individuals are eliminated in greater quantity. At the same time, the melanic type survives better, but it takes time before spreading. Because of the quick shrinking of the volume of white type, average fitness of the whole population decreases to Fii < Fi.

iii. Final equilibrium. Under the simplifying assumptions of specular fitness of the two types, after a few generations the population will go back to the initial size, split into two groups of white and melanic form. The relative abundance of the types, however, will be switched. Average fitness will go back to its initial value: Fiii = Fi.

All movements are limited to the horizontal (fitness) axis: functions, architecture and robustness do not change. The cloud representing the population shrinks in the second step, reflecting the reduction in population due to sudden lower fitness of white type individuals. The last step is identical to the first one, but with switched colours. This identity shows that the whole process is reversible: a new environmental change entailing less dark barks (e.g.

environmental policies, or deindustrialization) has pushed the population back to the initial situation (at least as long as mutation, migration and other mechanisms avoid the fixation of one of the types and the definitive disappearance of the other). We will talk more about the reversibility and irreversibility of evolutionary phenomena in §5.3.3.

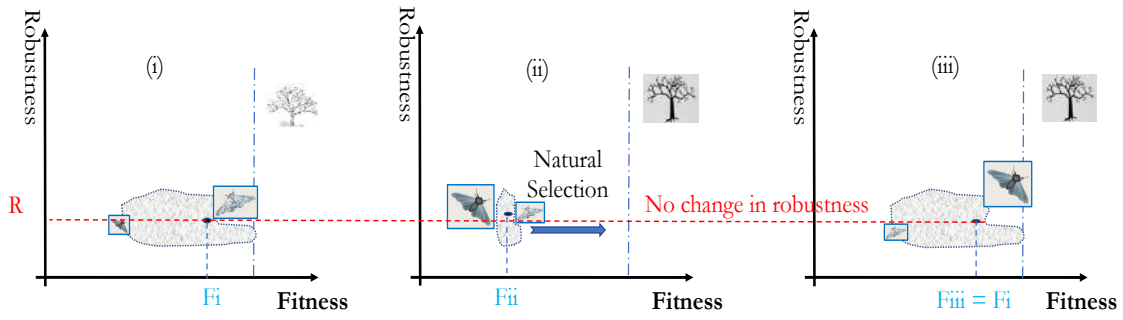


Fig. 5.22 – Adaptation of *B. betularia* wings' colour to change in the environment. Areas of *B. betularia*'s pictures are proportional to the percentage of the phenotype in population. Black points represent averages. Grey area represents the population (the smaller the area, the less numerous the population).

We can generalize these results by saying that selection, when acting alone in a constant environment, tends to increase average fitness and to reduce fitness variance, while it does not constantly impact on robustness (Fig. 5.23). It is thus a directional and continuous force with regards to fitness.

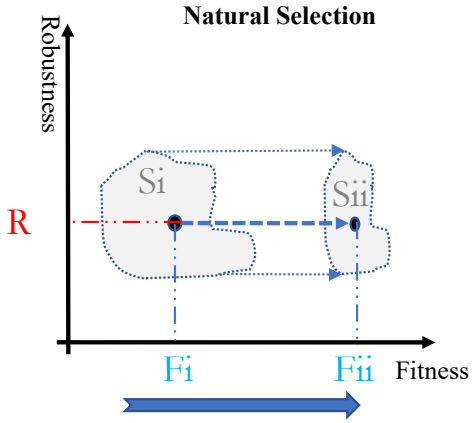


Fig. 5.23 - Natural selection pushes the population toward higher average fitness (Fii > Fi) and reduces fitness variance (final area Sii < initial area Si), while having limited impact on robustness (ΔR=0).

Drift

Drift can defined as sorting without differential birth and death: while selection picks up the individuals most able to survive and reproduce in a given environment thanks to the traits and functions they possess, drift picks up individual casually, so that the resulting subset can

be better or worse off than the original one, or remain unaffected if the sorting eliminates or fixes selectively indifferent traits. Consider the case of the land snail *Cepaea nemoralis*. The species is polymorphic: both within and among population, the individuals' shells differ in colour (pink, yellow, and brown) and presence, type and number of bands (from none to five) (Fig. 5.24).

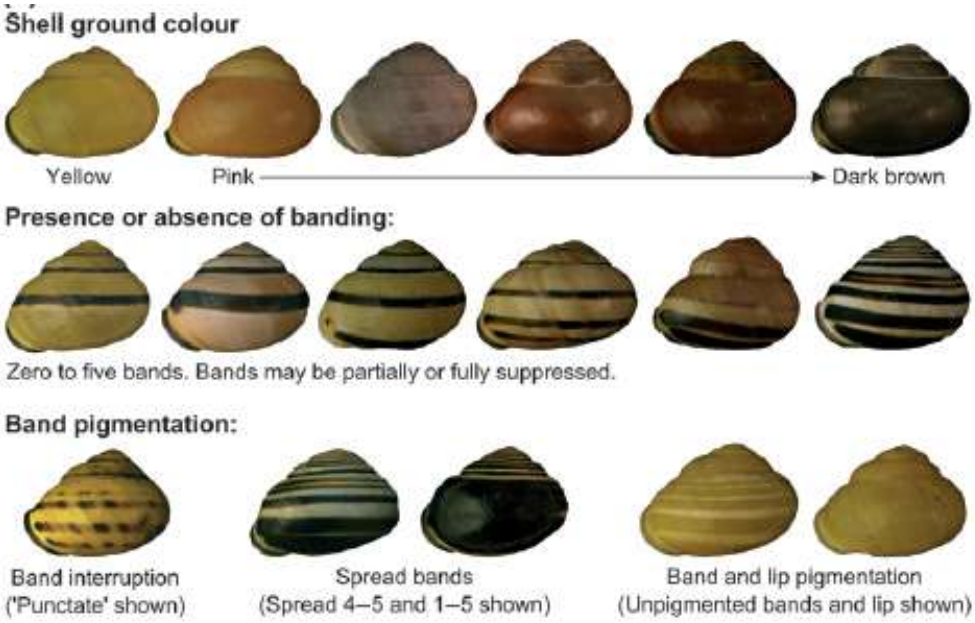


Fig. 5.24 – Polymorphism in *C. nemoralis* (from Richards et al. 2013).

While "the proportions of different varieties in each colony are strongly affected by selective visual predation [*that is, by natural selection*]" (Cain & Sheppard 1954: 89, italics mine), establishments of new colonies are "of considerable importance because of the chance variations in the composition [*i.e. shell types*] of the first colonizers [*that is, because of drift*]" (Lamotte 1959: 80, italics mine). We have on the one hand the creation of variety and its fixation in each colony due to drift (for example, through the founder effect[214]), and on the other hand its subsequent reduction due to selection. Drift is a random process and can thus change the population average of both fitness[215] and robustness, although, being a non-directional force, it is impossible to forecast if they will increase or decrease (Fig. 5.25). On

[214] The 'founder effect' is the loss of genetic variation due to the establishment of a new colony by a reduced group of individual from a bigger population: the consequent sampling of alleles is not linked to their fitness.

[215] The fact that drift and other forces can reduce average fitness does not contradict Fisher's theorem, given that this applies when selection is the only force in place (Price 1973).

the other hand, drift does not act upon variances[216] (Brandon 2006), thus leaving the population area unaffected (Fig. 5.25). Drift is thus a non-directional force, and a continuous one, as any increase/decrease of averages is possible.

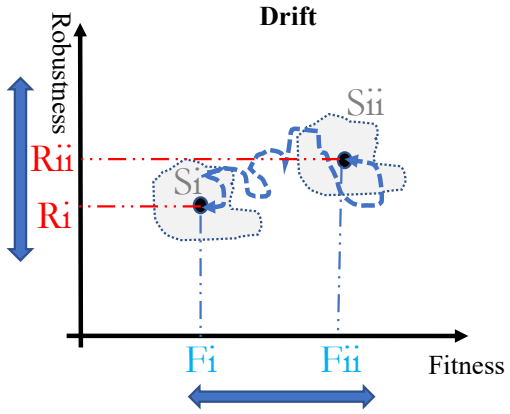


Fig. 5.25 - Drift causes random changes in the averages of fitness and robustness ($\Delta F \gtrless 0$, $\Delta R \gtrless 0$), leaving variances unchanged ($\Delta S = 0$). (Note that the evolutionary phenomenon can lead from i to ii or the other way round; also, due to drift, one of the two averages might increase and the other might decrease at the same time.)

The Zero Force Evolutionary Law (ZFEL)

The ZFEL[217], like drift, is non-directional: it entails a variety of phenomena whose effects increase the variance of both dimensions of the design space, but leave averages unaffected (McShea & Brandon 2010). Consider the asymmetry between claws in the male phenotype of fiddler crabs (genus *Uca*). The large claw appears to be the result of sexual selection. The two claws "can be said to evolve randomly with respect to each other, and their differentiation is therefore properly an instance of ZFEL" (Íbid., 72). The fact that, in most but not all species, half of males have a major claw on the left and half on the right, while, in *U. vocans vomeris*, less than two percent are left-clawed (Backwell et al. 2007) also suggests a role of ZFEL in initiating the size differentiation: indeed, the phenomenon cannot be due to drift as it happens in all species and population but at different degrees in each; and it cannot be due to selection alone as selection clearly acts separately on each claw once the claws are already differentiated. We can imagine that the initial asymmetry due to ZFEL appeared and augmented at different degrees in different individuals of the population, making these individuals better or worse off at surviving, thus impacting on their robustness, and at reproducing, this impacting on their fitness. These impacts were nevertheless casual, and, as

[216] Note that we are talking here about fitness and robustness variances, and not about the population genetic variance, that is usually reduced following drift, because of the casual loss and fixation of alleles.

[217] See §4.5.4

a whole, the population's averages did not change. What was not casual was the subsequent spread and fixation of some of the variants thanks to selection. For example, individuals with a bigger left claw could have a lower fitness than individual with a bigger right claw (because better in sexual fights, Backwell et al. 2007), and both types could have a lower robustness because less able to escape from predators (Jordão & Oliveira 2001). Individual with both claws of big size could have the highest fitness and the lowest robustness for the same reasons (Fig. 5.26).

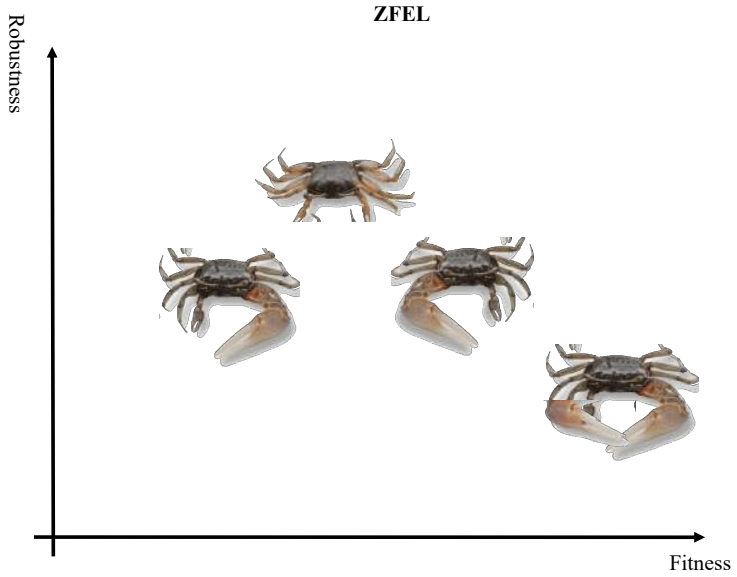


Fig. 5.26 – Possible scattering due to ZFEL of male individuals of a population of fiddle crabs away from the initial phenotype with two symmetrical claws.

The result of differentiation due to ZFEL is thus a progressive increase of the population area (Fig. 5.27).

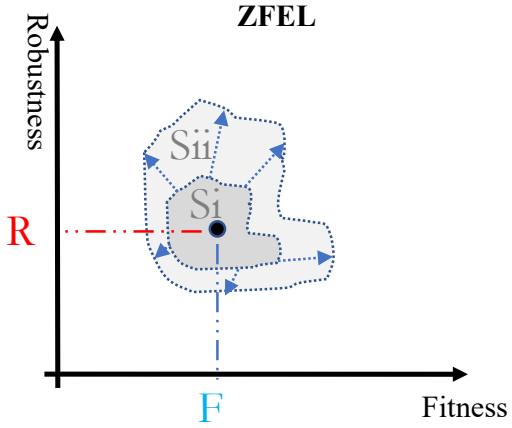


Fig. 5.27 - The 'Zero Force Evolutionary Law' leaves averages unchanged ($\Delta F = 0$, $\Delta R = 0$), while increasing fitness and robustness variance (Sii > S1, so $\Delta S > 0$).

Exaptations

Exaptations are traits that, however appeared (as previous adaptations, or as output of any of the non-selective forces and processes presented in Ch. 4), increase the fit between the organism and the environment, either because they help coping with a change in the environment, or because they permit some new function allowing the exploitation of a new niche in the current environment[218].

Before showing the effects of an exaptation on the design space, we must distinguish between two different aspects of the concept: the concept of the exaptation *before* it contributes to increase the fitness of the organism (we will refer to this concept as 'ex-ante exaptation') and the concept *after* that happens (ex-post exaptation)[219]. If we focus on the ex-post concept, we will fall once again in the pitfall of fitness: we will analyse the organism too late in the flow 'from-variation-to-selection', and we will lose sight of the potential ex-ante exaptations that have not yet become useful. Ex-post exaptations have an impact similar to classical adaptations, and their mapping in design space is the same as the one from selection.

Let us consider the case of calcium storage, exapted as bones. The ability of calcium to supply a rigid structure to the organism represents the ex-ante exaptation; the fact that this ability turned out to be useful in a given environment is the ex-post exaptation. The ex-post exaptation can be measured through fitness, but the ex-ante should be measured for its contribution to robustness. The potentiality of an ex-ante exaptation might or might not become an effectively useful ex-post exaptation.

We can therefore conclude that ex-ante exaptations increase robustness, but do not impact fitness. Note that the effect of exaptation on robustness cannot be negative, or they would not be exaptations: in this sense, they are directional on robustness, the same as selection is directional on fitness. If all individuals possess the exapted trait, robustness variance remains in principle unaffected, as the new environment affects all individuals equally. If not, it will increase (Fig. 5.28).

---

[218] See §4.3.1

[219] None of these names is used in the literature and we introduce them here just to simplify the discussion and avoid confusions.

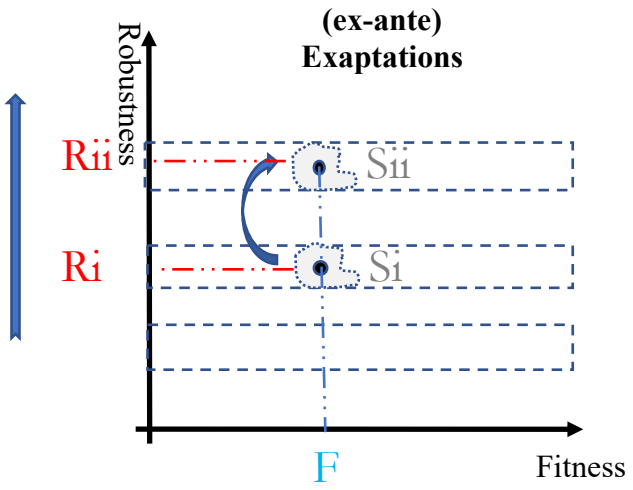


Fig. 5.28 – 'Ex-ante' exaptations have no immediate effect on fitness (ΔF), while increasing average robustness (Rii > Ri). If all individuals possess the exaptations, robustness variance is unaffected (ΔS =0); if not, it increases (ΔS >0).

Complex systems self-organizing laws

The laws underlining self-organizing systems[220] affect both robustness average and variance, given that they tend to increase the complexity of the organism in the sense of exhibiting new parts, new connections among parts and new functions. Although the average fitness change is null in principle, some individuals might experience an increase in fitness due to such changes. The population fitness average, however, can consistently change only under selection[221] (Fig. 5.29).

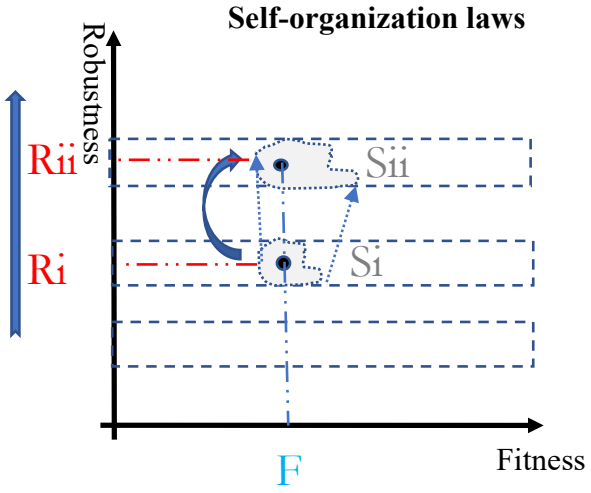


Fig. 5.29 - Complex systems' laws affect robustness average and variance, without consistently affecting fitness.

[220] See §4.5.3

[221] See note 206 for a definition of 'consistent'.

The mechanisms considered in the previous sections usually act incrementally. The complexity laws, evo-devo laws and the like, on the contrary, could cause discreet changes in the value of robustness. Evo-Devo research, for example claims that "[…] the structure of a developmental process may be such as to give rise to a series of *bounded discrete morphologies without intermediates*" (Maynard Smith et al. 1985: 266, italics mine), each with a radically different robustness.

Mix of evolutionary forces

Several evolutionary forces might act at the same time, although typically with different strengths and operating at different time scales.

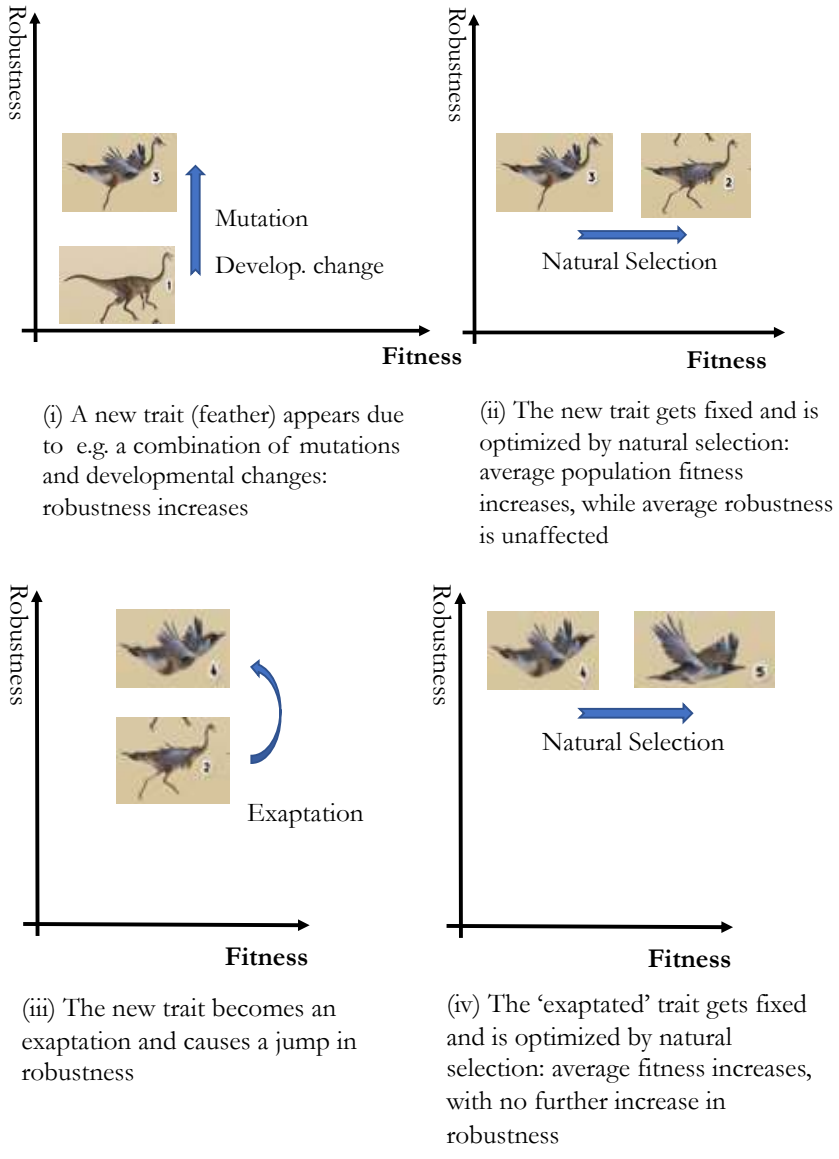


Fig. 5.30 – Evolution of flight as sum of several evolutionary phenomena triggered by different evolutionary forces, each affecting robustness and fitness in a different way (see text for detailed account of each step).

Evolutionary phenomena can be factorised in movements along both axis, linked to the independent impact of each intervening force. As an example, we propose a possible explanation of the appearance of feathers and of the function of flight (Fig. 5.30).

Prum (1999) proposes a functionally neutral model of the origin and evolutionary diversification of bird feathers. This author identifies several stages in the evolution of novelties: the initial stages of feather evolution could have fulfilled functions in communication, defence, thermal insulation, or water repellence, while the aerodynamic function could have appeared only in the last stages. These different stages reflect different movements in the design space:

i. The first step (Fig. 5.30.i) consists in the appearance of the new trait: feathers are "a complex evolutionary novelty characterized by structural diversity and hierarchical development" (Prum 1999: 291). Although their evolutionary origin remains controversial and poorly understood (Brush 1996), Prum (1999) suggests that they originated by a combination of mutations and developmental changes.

   The new trait 'feathers' has a potential for improving robustness in several ways (Aparicio et al. 2003, Biewener & Patek 2003, Dial 2003): by acting as thermal insulation and heat shielding (allowing the organism to survive to wider ranges of external temperatures), for display (increasing sexual attractiveness), as water repellent, and as stabilizers during running (allowing the individual to bend forward without losing equilibrium).

ii. The actual impact of the new trait on fitness is decided by selection. If it is immaterial, the trait will disappear or remain marginal (as in the 'weird wonders' of Fig. 5.5). If it is positive, the trait will spread and get fixed among the individuals of the population, and its configuration will be optimized, in terms, for example, of shape and quantity of the feathers: the fossil record of the transition from theropod dinosaurs to early birds shows a stepwise trend of miniaturization (Erwin 2015). It is a 'primary adaptation', as it develops an answer to the original environmental challenge. Different aspects of robustness might be relevant in different environments, leading to the evolution of different forms and functions: for example, sexual attractiveness might be more important than flight, in absence of predators (Fig. 5.31).

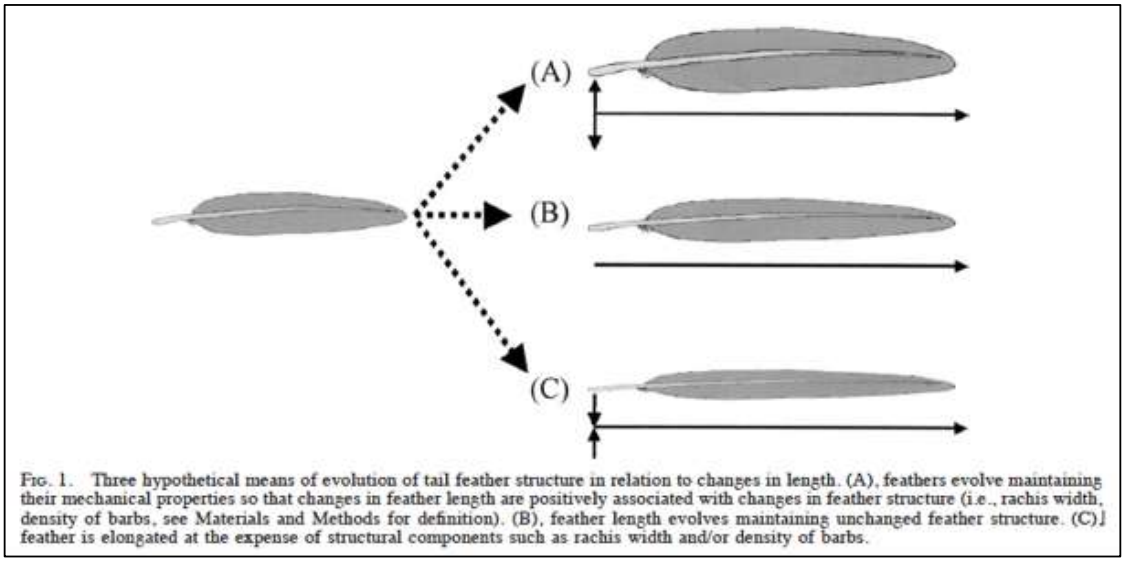


FIG. 1. Three hypothetical means of evolution of tail feather structure in relation to changes in length. (A), feathers evolve maintaining their mechanical properties so that changes in feather length are positively associated with changes in feather structure (i.e., rachis width, density of barbs, see Materials and Methods for definition). (B), feather length evolves maintaining unchanged feather structure. (C), feather is elongated at the expense of structural components such as rachis width and/or density of barbs.

Fig. 5.31 – Alternative evolutionary changes in feather structure, according to different selectively-relevant uses (from Aparicio et al. 2003).

iii. At some point, the new trait, which has been optimized by selection for any of the functions listed in (ii), becomes an exaptation: it allows a primitive and rudimentary flight. This new function increases the organism's abilities, and it thus increases its robustness[222]. Its relevancy for fitness, nevertheless, can only be decided by selection. If flight does not give any competitive advantage in a particular environment, selection could eliminate it (because, for example, it uselessly consumes energy resources), as it was the case with ostriches, emus, kiwis and other flightless birds (Baker et al. 2014). In theory, there might exist species with feathers belonging to lineages that were never able to fly and had remained at stage (ii) of our account. The mainstream adaptationist literature, however, seems to ignore this possibility and contends that feathers must have led to flight, and that flightless birds must be the result of a loss of function. An exception is the analysis of the evolution of the function of insects' wings. Short wings have no aerodynamic effects, but they have significant thermo-regulating capacity, while long wings show the opposite behaviour. Once wings appeared, thus, they potentially improved the organism's robustness in several ways. Recall that we defined robustness as the set of current abilities of the organism, regardless of whether they are actually used or not. Wings can improve the organism abilities in several ways (e.g. acting as thermoregulating mechanism, allowing mating-dance in sexual selection, etc.) thus improving its robustness. Only

[222] See §5.2.1

one of these ways was subsequently picked up and improved by selection in each environment (Kingsolver and Koel 1985).

iv. If flight does grant some advantage, selection tends again to its optimization, for example, in terms of shape or number of feathers, or dimensions of wings: it's a 'secondary adaptation'. Additionally, the organism undergoes other secondary adaptations that modify order, arrangement, and shape of limbs and muscles to enhance flight (Gould & Vrba 1982, Dial 2003).

This example can be analysed also through the conceptual framework suggested by Erwin (2015), who identifies four aspects in the appearance of novelties and innovations. The first step ('potentiation') is linked to genetic mechanisms allowing changes in genotype (mutations, re-combinations, etc.) and their accumulations. The second step ('generation of novel phenotype') involves new traits, and it can be mapped in our design space as a big upward jump in robustness. The third step ('adaptive refinement') entails changes due to small mutations/re-combinations and selection, and it is mapped as a series of small movements either upward or rightwards[224]. The final step ('exploitation') transforms the novelty into an innovation (using Erwin's vocabulary) thanks to environmental, ecological and evolutionary conditions that greatly increase the fitness of the organism possessing the given trait. This is mapped as a big rightward movement in design space in line with what predicted by population genetics equations.

5.3.3. The design space and Fisher's fundamental theorem: the thermodynamic metaphor extended

The design space can be interpreted as an extension of Fisher's fundamental theorem and of the idea of fitness landscape; or, alternatively, these can be seen as mono-dimensional sections of the design space, under the tacit hypothesis of constant robustness. In this section, we develop the first of these extensions, while in the following section we show the relationship between design space and fitness landscapes.

Fisher compares his famous fundamental theorem of selection[225] to the second law of thermodynamic[226], given that

[224] This is how evolution according to the positive view appears in the design space, see Fig. 5.19.
[225] See §2.2.2

> [B]oth are properties of populations, or aggregates, true irrespective of the nature of the units which compose them; both are statistical laws; each requires the *constant increase of a measurable quantity*, in the one case the entropy of a physical system and in the other the fitness […] of a biological population. (Fisher 1930: 36; italics mine)

In Fisher's model, a population average fitness in a constant environment never decreases, provided that selection is the only acting force (Price 1973): it always either increases or, if the population genetic variance is zero, remains constant. In the same way, entropy of an isolated thermodynamic system (a system that does not exchange matter nor energy with the outside) never decreases. Fisher limits his analysis to the variables fitness/entropy: having introduced a second dimension into the evolutionary picture, we propose to extend it to architectural robustness/energy (we will clarify below why architectural robustness).

In an isolated evolutionary system (IES), selection is the only acting force and the environment is constant: its thermodynamic equivalent is an isolated system (ITS). The biological variables fitness and robustness of a IES are the equivalent of the state variables of an ITS (e.g. any two out of the triad pressure, volume and temperature). Once the main points of the analogy are so determined, we can say that fitness never decreases in an IES as entropy never decreases in an ITS (this is Fisher's original analogy). Moreover, architectural robustness is conserved in an IES[227], the same as energy is conserved in an ITS. This is straightforwardly deduced from our hypothesis that selection only acts on fitness, but not on robustness: and in an IES, by definition, only selection acts.

The analogy can be further extended: evolutionary processes, the same as thermodynamic ones, can be reversible or irreversible (Fig. 5.29).

---

[226] Wright (1949) also compares average fitness to the potential function of a conservative field, whose gradient determines the system changes, and that is maximized at equilibrium. We do not include this analysis here because Wright later discharged the idea (Edwards 1994).

[227] Stability robustness might increase, if the function is maintained through new traits; the increase is not bit enough, though, to cause the population to jump to a new attractor (see §5.3.5 for a detailed discussion on the evolutionary differences between stability and architectural robustness changes).

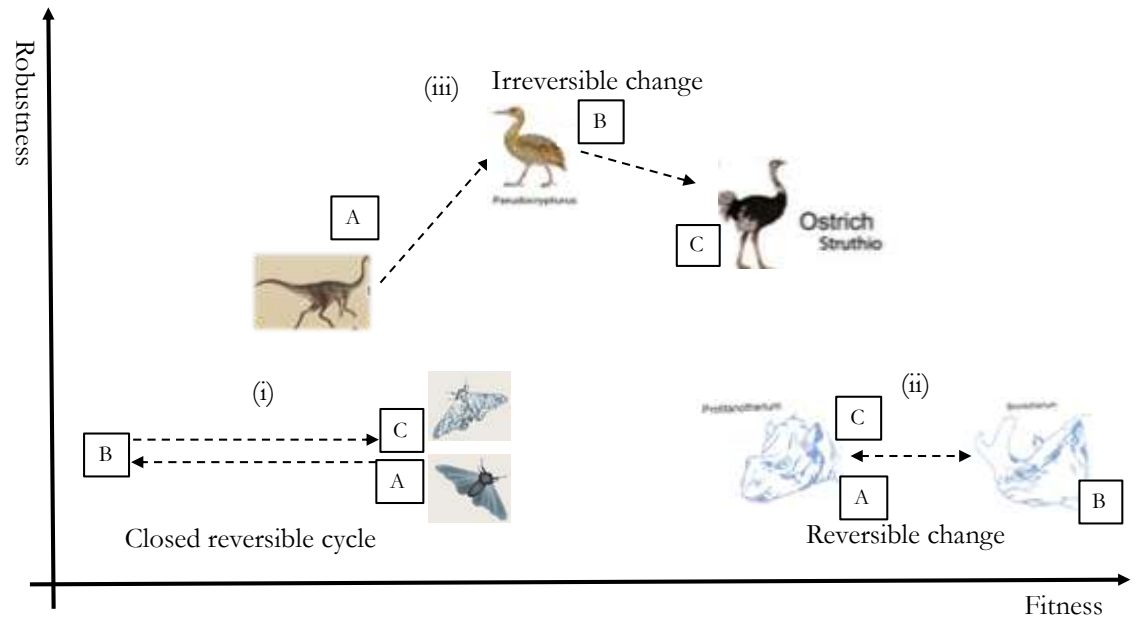


Fig. 5.32 – Reversible and irreversible evolutionary phenomena (see text for detailed explanation). (i) Closed reversible cycle: the change from white to black wings and back of the *B. betularia* is on-going. (ii) Reversible change: although not happened, the rhino's horn could undergo reverse evolution and disappear. (iii) Irreversible change: even in the case of radical environmental changes making flight a fundamental function, it is very unlikely that ostrich could recuperate it. In (i) and (ii), it is impossible to deduce which phenotypic version appeared first without further information; in (iii), the phenomenon itself also suggests the direction of the 'arrow of time'.

Reversibility of a thermodynamic cycle (where entropy change is null) mirrors closed evolutionary changes where average *fitness is the same at the beginning and at the end* of the change. A simple modification of genetic frequencies in a population due to a change in the environment is reversible if the environmental change is reversed. In these phenomena, traits and functions, and therefore architectural robustness, are maintained, the same as energy in thermodynamic reversible cycles. In the mentioned case of the *B. betularia* wings' colour (Fig. 5.22), the poplars' barks switched back from black to white following environmental policies, and the pre-industrial population's phenotypic distribution was restored. But the *B. betularia* is not the only known case of reversible evolution. Populations of the black-bellied seed cracker (*Pyrenestes ostrinus*), a finch found throughout sub-Saharan Africa, exhibit a bimodal distribution in bill size that is unrelated to sex, age class, or geographic origin (Smith 1987); rather, the polymorphism is related to the efficiency that the morphs exhibit while feeding on the seeds of two sedge species found throughout the bird's range. The seeds of the two sedge species are identical in size but differ in their hardness. Bill morphology is heritable and controlled by a single locus (Smith 1993). The large-billed finch morph can crack and consume seeds of the hard-seeded species (*Scleria verrucosa*) more efficiently than the small-billed form. The small-billed morph (*S. goossensii*), in turn, is more efficient at cracking and

consuming seeds of the soft-seeded species. During periods of high food abundance, following the rainy season, the diets of both morphs overlap considerably, with the large-billed morph consuming soft as well as hard seeds. As seed abundance declines, the diet of the large-billed morph becomes more restricted; it specializes on harder seeds. The large-billed form can consume the softer seeds, but is presumably excluded by intraspecific competition when resources are scarce (Smith 1990). During wet years, the proportion of soft seeds increases, at which point smaller-billed individuals experience enhanced survivorship and the mean bill size declines (from Halama & Reznick 2011). The reversibility of the process is even clearer in this case, as it regularly occurs after changes in the dryness of the environment. The evolution of beak dimensions of Galápagos finches is yet another example of a reversible cycle: between 1976 and 1987, opposite extreme climate favoured large bills and small beaks alternatively (Hendry & Kinnison 1999). This kind of reversibility matches the third criteria for reversibility stated by Hollinger and Zenzen (1982): a process is reversible if the process and its reverse both naturally occur. In these cases, fitness changes cannot provide evidence concerning an 'arrow of time' of evolution. Note that reversible evolutionary phenomena are not necessarily actually closed cycles: it is enough that they can be closed. Rhino's horn evolution (Fig. 5.21 - ii) is not a cycle, but it is reversible because it could become a cycle in the future: the horn could go back to its original shape and also be entirely lost. Fitness changes are not significant clues of an 'arrow of time' because, without information about the history of the species, we could not tell which version of the horn appeared first.

In *irreversible* evolutionary changes, architectural robustness is not maintained and, if selection acts, fitness increases irreversibly at the end of the change. Flightless birds have little probability to regain the lost function, given the complex genetic and morphological changes they have undergone (described in Baker 2014). Their robustness is substantially different than in their ancestors because the phenotypic modifications are not just a new mix of existing traits, nor optimizations of quantitative traits. Brooks and Wiley (1986: 4) classify under this heading the evolution of biological diversity and claim that it requires a macroscopic cause, thus excluding that selection alone can explain it. What is known as 'Dollo's law of irreversibility', or the idea that "an organism cannot return, even partially, to a previous state already realized in its ancestral series" (Dollo 1893: 165), has often been interpreted as an absolute impossibility, but there are opinions (Dawkins 1996, Gould 1970) and examples (Pagel 2004) that suggests a statistical interpretation. Reverse evolution is simply very improbable whenever substantial genetic and developmental modifications are

involved. It is thus mainly applicable to complex traits, especially when they involve new body plans (Gould 2007): it is hard to reverse a 'major transition' (Maynard Smith & Szathmáry 1999: 25).

The main concepts of our proposal are summarised in the following table.

Table 5.6

	Termodynamics	Evolution
State description	state variables	fitness, robustness
1st law	conservation of energy	conservation of architectural robustness
2nd law	entropy never decreases	fitness never decreases
reversible cycle	entropy and energy changes are null	fitness and architectural robustness changes are null
irreversible change	entropy change is positive energy change is not null	fitness change is positive architectural robustness change is not null

Analogy between thermodynamics and evolutionary concepts and phenomena in isolated systems (see text for details).

5.3.4. The design space and fitness landscapes: architectures and attractors.

Fitness landscapes[228] can be seen as mono-dimensional versions of the corresponding design spaces for the obvious reason that both graphical representations share fitness as one of their variables. We suggest, however, a more subtle link, showing that both representations include attractors in Kitano's sense: in the design space, attractors correspond to areas that can theoretically be occupied by populations; in fitness landscapes, attractors are placed around fitness peaks. In this section, we resort to this interpretation of fitness landscapes as unidimensional section of design spaces to show that they provide insufficient information for the complete analysis of evolutionary phenomena involving architectural changes.

Let us consider the simplest possible landscape, where the relationship between genetic frequencies and the individual fitness is summarized in population genetics as a function:

[228] Remember the difference between fitness landscape (when individual values are considered) and adaptive landscape (when averages are considered): see §4.2.2

$$\text{Fitness} = f\ (\text{genetic frequencies}) \qquad (4)$$

The reasons behind this relationship are not explicit: we do not know, for example, why a frequency has a better fitness than another. As we have discussed in §5.1.2, the organism, summarised by the function f, is considered as a black box, and we have no information about its inner architecture (Fig. 5.2). The only relevant feature of each type of organism is the value of its fitness.

We can make an analogy with the graph showing the relationship between the number of revolutions per minute (rpm) and the corresponding torque (the rotational equivalent of a linear force) of a car engine. The graph gives us relevant information about the performance of an engine, but no tip about how this performance is reached. Trying to understand why an engine performs better at a given value of rpm, or why it is better than another, just by analysing these graphs, would be a hopeless task (Fig. 5.33).

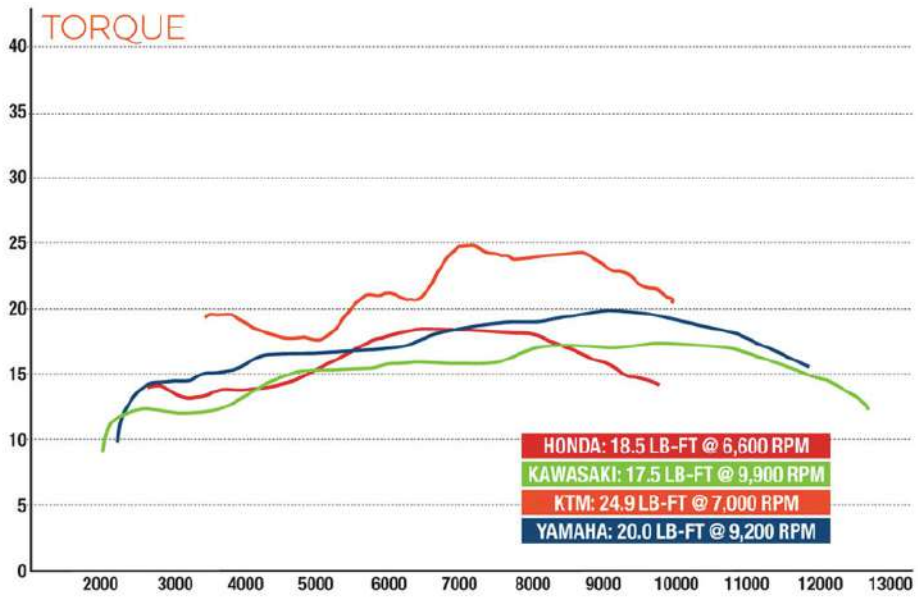


Fig. 5.33 – Relationship between rpm and torque for different car engines. Yamaha (blue curve) performs better than Kawasaki (green curve) at almost any rpm, but by analysing the graph itself we get no information about why it is so.

Consider now a fitness landscape with two peaks, and two populations placed in each of them (Fig. 5.34).

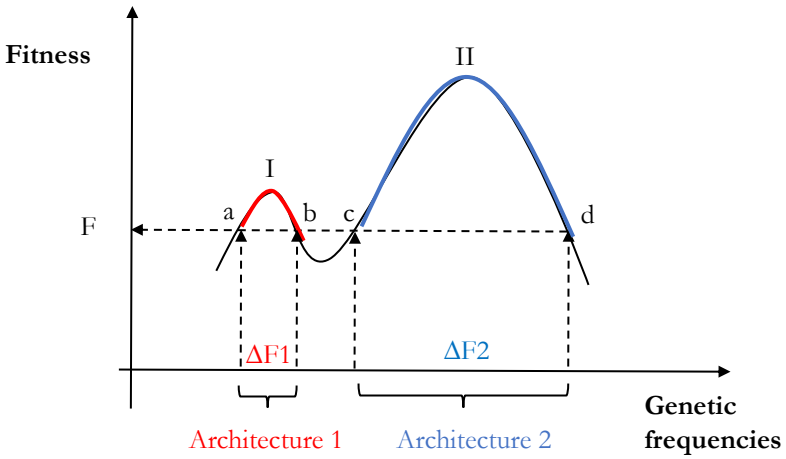


Fig. 5.34 – Architectures behind fitness landscape. F is the lower value of fitness of two different populations, highlighted in red and blue, spread around peaks I and II. The individuals at points a, b, c, d have all the same fitness F, but possibly different subjacent architectures: something that the fitness landscape alone cannot reveal.

Each peak represents an equilibrium state: in an ideal situation, all individuals of each population would place themselves at the peaks. Due to all kinds of constraints, however, not all individuals can possess the maximum theoretical fitness: they are rather spread around the peaks, producing a range of fitness values with F as the lower limit. Not just the peak, but the whole curves a-b and c-d thus represent the possible equilibrium states of the populations. From an adaptationist perspective, points a, b, c, d are identical in that their genetic constitution result in the same fitness. From an evolutionary standpoint, on the other hand, there is more than that: c and d have a potentially much higher fitness, although constraints do not let them reach it.

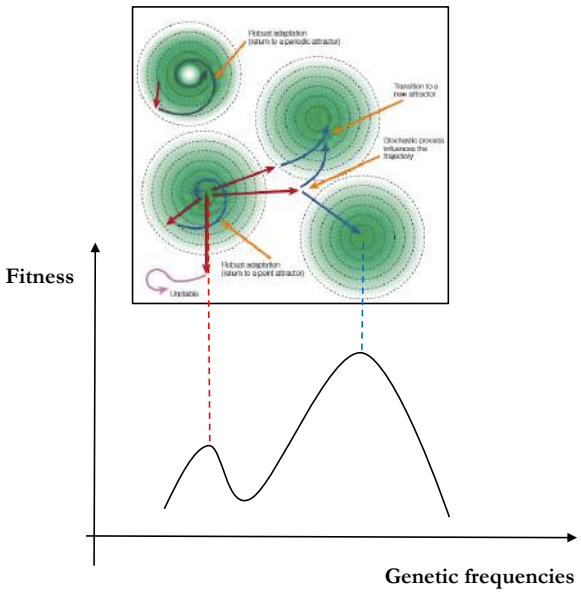


Fig. 5.35 – Relationship between fitness landscape and attractors (drawing of attractors' space from Kitano 2004): the space of attractors is similar to a fitness landscape seen from above.

We can interpret each peak and the surrounded areas (a-b and c-d) as attractors in the sense proposed by Kitano (2004). Kitano's illustration of the space of attractors is similar to a fitness landscape seen from above, where each concentric circle represents an 'iso-fitness' curve around a common peak (Fig. 5.35).

The state of the system in Kitano's space corresponds to the genetics' frequencies in the landscape. Average fitness and a peak can split up, for example, because of an environmental change modifying the landscape, or because of drift (as in Wright's shifting balance hypothesis). After the split, the population tends to return to the initial peak or to move towards a new one. In the same way, according to Kitano, a system moved away from an attractor tends to move back to it or to find a new one. Attractors can be thus interpreted biologically as complex organic systems with a given level of fitness: as such, Kitano's representation has little to add to fitness landscapes. We propose to consider attractors as possessing a level of robustness too. As entities defined by both fitness and robustness, they cannot be fully mapped into fitness landscapes, but can be represented by areas in design space, areas with a peculiar characteristic: when a population enters such an area, it tends to remain in it; if displaced away by some accident, it tends to move back to it, or to move to another attractor's area.

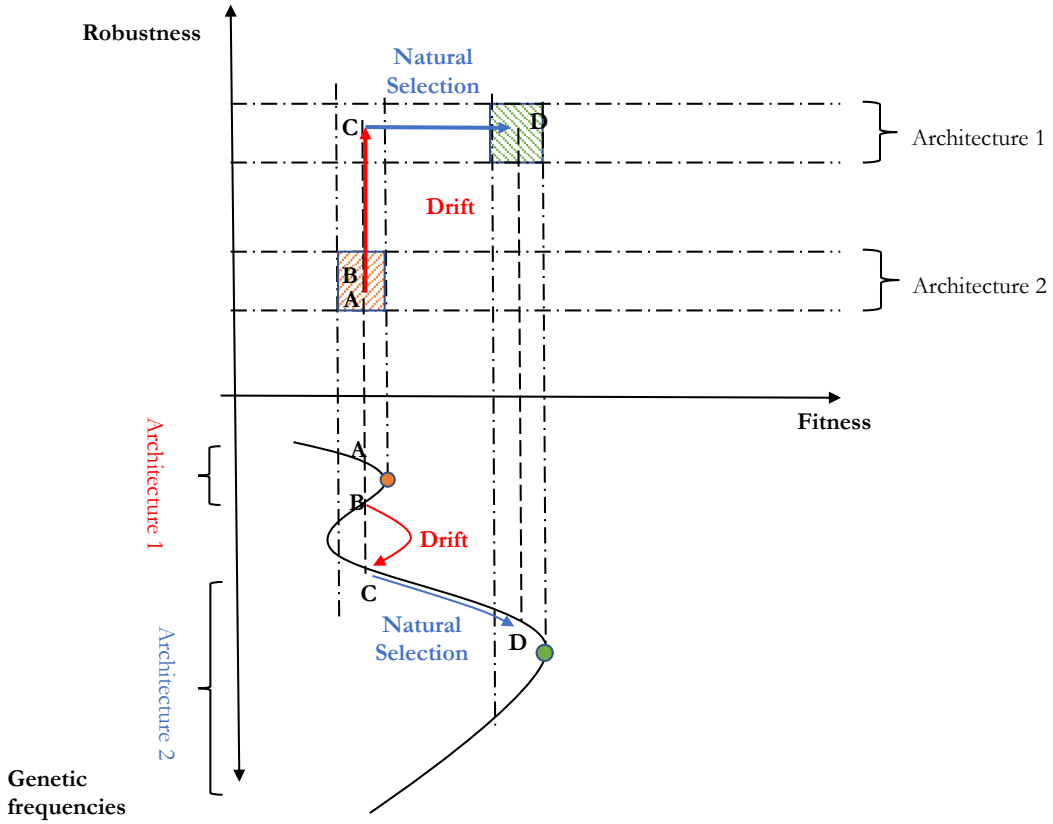


Fig. 5.36 – Expansion of fitness landscape (lower graph) into design space (upper graph). Robustness movements at constant fitness are invisible in the fitness landscape.

We can expand a fitness landscape into a design space to illustrate these concepts. Fig. 5.36 shows a fitness landscape and a design space sharing the same fitness axis.

The fitness landscape and its elements (lower part of the figure) are the same as in Fig. 5.34 (the same figure rotated 90 degrees clockwise): the curve A-B identifies a population at equilibrium with an underlying architecture 1. Following drift, a sub-set of the population shifts from B to C: moved away from the attractor, it arrives to the area of influence of the new attractor 2, where the sub-set finally moves, pushed by selection, until positioning in D.

We can follow the same movements in the design space (upper part of the figure). The shaded red area corresponds to attractor 1, included between the maximum and minimum fitnesses of the population as mapped in the fitness landscape and between the maximum and minimum robustness of this population, the latter *information not explicit in the fitness landscape* representation. Drift moves the population to a higher level of robustness, *leaving fitness unaffected*: point C has the same fitness as points A and B, but this fitness is based on a different architecture possessing higher robustness. Recall (from §4.2.2) Wright's belief that drift, by eliminating and sorting alleles in a way that selection would not allow, opens the way to new recombinations that were impossible, and that the corresponding phenotypes, although worse off at the beginning, are potentially better off in the long run and can reach higher fitness after a new selective process. This 'potentiality' lies in the new architecture: an information that is absent in the fitness landscape, where we only know that a higher pick is available, but *we do not know why it* is higher. This higher robustness of the new architecture is selectively relevant, and thus the population is moved to higher fitness until arriving at point D, in the area of attractor 2.

5.3.5. A classification of evolutionary phenomena

In the previous sections, we have proposed a way to characterise biological design and to understand how evolutionary forces, alone or in combination, might cause it, generation after generation, to move around design space. We have also suggested how the proposed design space can be interpreted as an extension of Fisher's fundamental theorem and of fitness landscapes, and why such an interpretation is informationally relevant. Thanks to all this theoretical background, we can now give an answer to the question concerning the epistemological and ontological need - underlined at the beginning of this chapter – to provide a classification of the great variety of evolutionary phenomena (Fig. 5.1), with the ultimate aim to understand the actual causal role selection and non-selective forces play in

each of them. In this section, we will suggest how to build such a classification, borrowing some concepts about types of novelties form discussions on evolvability (Pigliucci 2008a) and open-endedness[229] (Banzhaf et al. 2016, Taylor 2018), and applying them to attractors in the design space. We believe these discussions are useful to illustrate our model because they recognize, explicitly or implicitly, that evolutionary phenomena cover a wide range of cases and they propose some categories to classify them.

Evolvability is a 'hot' term in current philosophy of biology, but it turns out to be a polysemous concept with no unique definition. Pigliucci (2008a) lists the many senses in which the concept is used in the literature and classifies them under three supra-categories (Fig. 5.37): heritability, evolvability in the strictest sense, innovation. Heritability measures the standard genetic variance of a population and is thus considered as a proxy of the ability of the population to respond to selection. The concept of evolvability moves the focus from the actual variance of the population's gene pool to the propensity of population traits to vary through mutation and recombination, a propensity that could or not actually be realized, and that depends, among other factors, on the genotype-phenotype map and developmental details. Finally, innovation is the potential to evolve radically new structures and architectures.

Suggested term	Scale	Description	Effects	Example
Heritability (*sensu* Houle)	Within populations	Standing pool of genetic variation and covariation	Determines the response to natural selection within populations	
Evolvability (*sensu* Wagner & Altenberg)	Within species	Includes variability (*sensu* Wagner & Altenberg), depends on genetic architecture and developmental constraints	Affects long-term adaptation, channels evolution along non-random trajectories, allows mid-term exploration of phenotypic space	
Innovation (*sensu* Maynard-Smith & Szathmary)	Within clades	As for within species, but includes the capacity to overcome standing genetic and developmental constraints, opening new areas of phenotypic space for further evolution	Generates major phenotypic (morphological, behavioural or physiological) breakthroughs (novelties)	

Fig. 5.37 – Possible senses of the concept 'evolvability' (from Pigliucci 2008)

Pigliucci's classification of the different semantic contents of the term 'evolvability' can be used as a taxonomy for evolutionary phenomena, if we classify them based on the kind of change they entail. A more formalised way to draw such a taxonomy is found in

[229] "Loosely defined, an open-ended evolutionary system is one that is capable of producing a continual stream of novel organisms rather than settling on some quasi-stable state beyond which nothing fundamentally new occurs." (Taylor et al. 2016: 409)

discussions about open-endedness (Banzhaf et al. 2016, Taylor 2018), or the ability of a system (be it an individual, a population, an ecosystem or a generic set of interlinked entities) to continuously produce novelties. Such discussions have identified three levels of change producing novelties[230]: variation, innovation, and emergency. The three levels mirrors Pigliucci's classification. At the level of *variation*, the change just affects the value of a variable of the system, or the combination of variables, without modifying the set of entities of the system nor its structure: e.g. a switch of a gene to a new allele, or a change in the value of a phenotypic trait. Variation is exploratory novelty: the same function is performed with the same traits, but slightly modified. Variations correspond to Pigliucci's heritability level: once the optimum is reached, the system needs some kind of 'shock' (a change in the adaptive landscape, in the genetic space or in the genotype-phenotype map) to produce a novel phenotype. At the level of *innovation*, new entities are introduced in the system, but without changing the system's general structure: e.g. a gene is duplicated, a new chemical species is synthetized in metabolism, or a new species migrates into an ecosystem. Innovation is expansive novelty: the same function is performed in a new way by a new trait. Innovation corresponds to Pigliucci's evolvability level. At the level of *emergency*, the change affects the system structure: e.g. a new mechanism for variation like sexual recombination, or a new level of organization like an organ, or a new locomotion mean like flight appear. Emergency is transformational novelty: a new function appears, supported by new traits. Emergency corresponds to Pigliucci's innovation level[231] (Table 5.7).

[230] We maintain here the original article's terminology, although bearing in mind that the literature is full of different uses of the term innovation, novelty, etc. (see §2.4.2 for a brief presentation and discussion of some of them; see also Brigandt & Love 2012).

[231] Unfortunately, terminology in this area is not only polysemic, but often contradictory among authors: innovation for Pigliucci and for Banzhaf or Taylor means very different things.

Table 5.7

		Function	
		same	**new**
Trait	**new**	a) evolvability b) innovation c) expansive novelty	a) innovation b) emergency c) transformational novelty
	same	a) heritability b) variation c) exploratory novelty	(exaptation)

Changes in traits and functions for different kinds of evolutionary phenomena: terminology from (a) Pigliucci 2008a, (b) Banzhaf et al. 2016, (c) Taylor 2018.

We can now map on the design space these types of evolutionary changes in terms of fitness, robustness, and attractors. At the level of *heritability*, or *variation/exploratory novelty*, the set of traits does not change before and after the evolutionary phenomenon: what changes is their optimal values and combination. Functions do not change, and robustness is therefore maintained: after the shock triggering the change, the population remains within the limit of the original attractor. Fitness might remain constant (as in the case of *B. betularia*) or increase thanks to selection. Given that traits are fixed and do not change before and after the evolutionary phenomenon, the target of optimization is individual fitness[232]. Changes at this level are usually reversible and selection is the main force behind them. Optimization models (e.g. Grafen's project) apply to this level thanks to robustness' invariance.

At the level of *evolvability*, or *innovation/expansive novelty*, new traits' versions or new traits can appear, increasing average fitness. For example, the form of the beak of finches depends on the expression of two genes: although it is not a qualitative new feature, but a quantitative modification of an existing one, it does entail a new way of performing the same function 'feeding': the beak's novel version of *G. magnirostris* can crush the hardest seeds, something that the previous beak's versions could not do (Wagner 2011). These examples show that selection can lead to speciation without necessarily any radical change in the architecture of the organism. As it is a 'blind' force with regards to robustness, adaptation can be equally successful in a new environment either through morphological specialization [*where robustness remains constant*] or through improved biomechanical efficiency [*where robustness increases*] (Gould 1977, comments in italics on robustness are mine). The possibility of such different

[232] Recall our discussion about trait and individual fitness and how they can be inter-linked (§3.7).

evolutionary answers to the same environmental challenge shows that there must be other processes in place apart from selection. Moreover, subsequent selective opportunities could define some of these answers as a new attractor, thus increasing the size of organisms and fine-tuning the correspondent new traits through fitness optimization. What drives the change at the base of the evolutionary phenomenon is not fitness optimization by selection of the mix of traits, but other non-selective forces that impact on stability robustness. These changes are usually irreversible and are led by non-selective mechanisms. The population remains within the same attractor, because functions do not change: most problems that living beings have solved have many alternative solutions (Wagner 2005: 5). The difference with the previous case is that robustness increases, but not enough as to cause the entrance in a new attractor, as it happens with the next and last case.

Finally, at the level of *innovation*, or *emergency/transformational novelty*, new traits can result in new functions: some of the many equivalent solutions to a biological problem can harbour the seeds of innovation to solve other problems (Wagner 2005). New functions and new architectures arise that can include new characteristics like higher modularization or new restrictions (in the sense of Mossio, Bich and Moreno 2013): when architectural robustness increases, the organism moves to a new attractor in an irreversible transformation.

Table 5.8 summarises the main concepts of the classification of evolutionary phenomena. Fig. 5.38 shows how the three kinds of phenomena appear in design space. Finally, Fig. 5.39 suggests where each kind of phenomenon is most likely to be found in the architectural matrix of form and function proposed in §2.4.2.

Table 5.8

type of change			architecture		Example (trait/fuction)	design space			
Pigliucci 2008	Bazhof et al. 2016	Taylor 2018	morphology	function		focus of change	robustness space	mechanims generating novelties	transformation
heritability	**variation**	**explorative novelty**	same traits, new versions of same traits (standing genetic variation)	same through same, or modified, traits	B. betularia: wings/mimecry	fitness	same attractor	optimization of mix of existing traits by selection	usually reversible
evolvability	**innovation**	**expansional novelty**	new traits	same with new traits	Darwin finches: beak/feeding	stability robustness	same attractor	non-selective mechanisms: - phenotypic - developmental - systemic	unsually irreversible
innovation	**emergency**	**transformational novelty**	new traits	new	Turtle: carapace/defence	architectural robustness	new atractor		irreversible

Different kinds of evolutionary change (see text for detailed explanation).

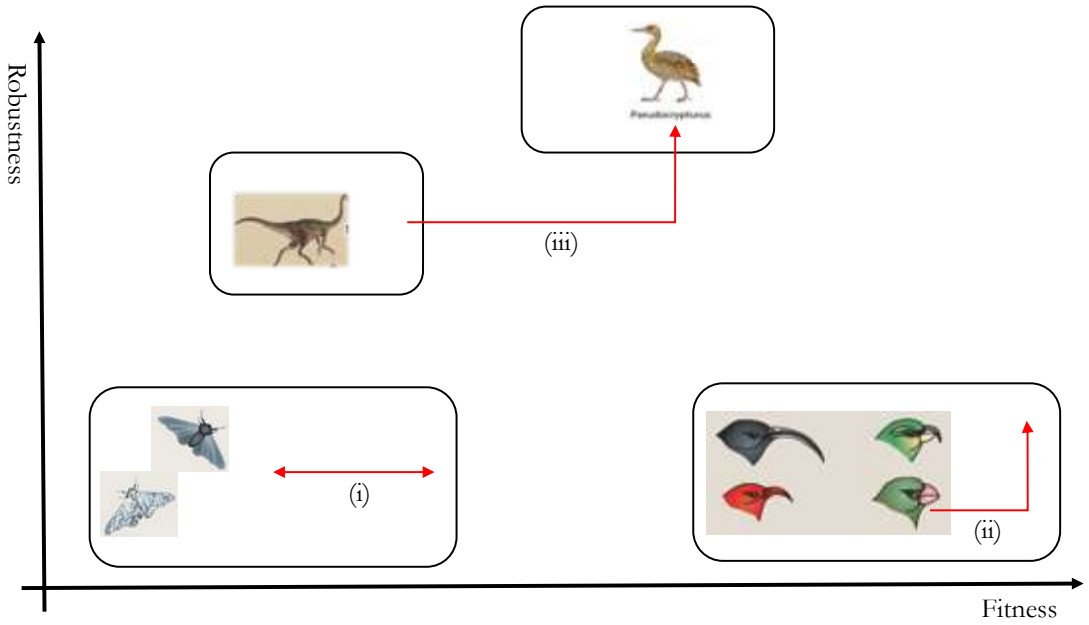


Fig. 5.38 – Mapping in design space of three different kinds of evolutionary phenomena. (i) Exploratory novelty: robustness is constant before and after the change. (ii) Expansive novelty: stability robustness increases after the change, but the population remains in the same attractor, as the phenotype does not gain new functions. (iii) Transformational novelty: architectural robustness increases and the population moves to a new attractor, corresponding to the new phenotypic function.

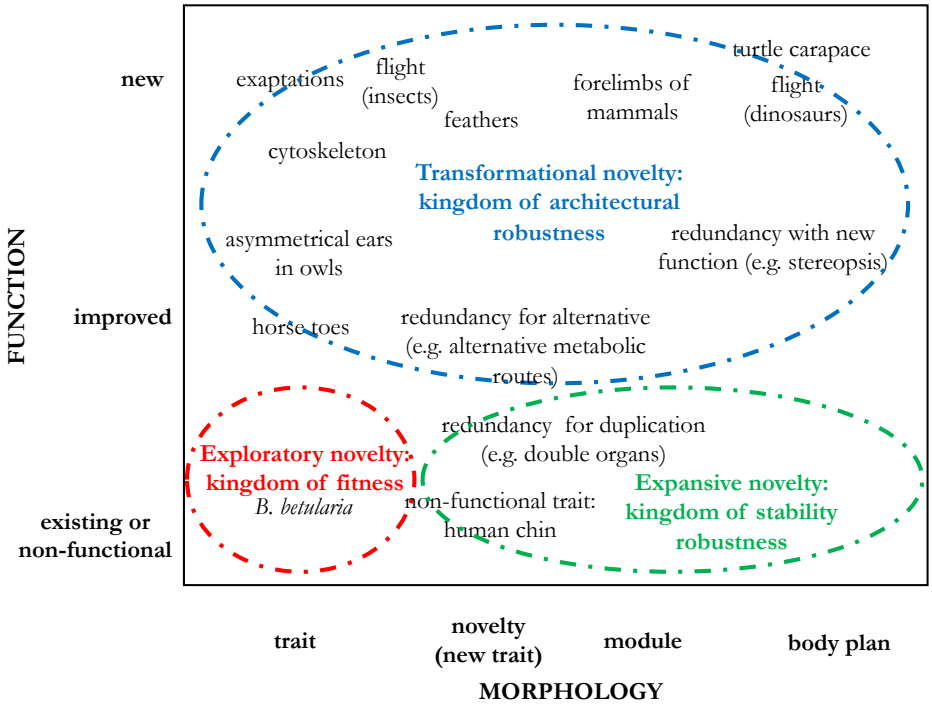


Fig. 5.39 – Mapping of kinds of evolutionary phenomena in the architectural matrix.

5.4. Conclusions

In this chapter, we have systematised much of the material introduced in the previous ones. We have shown how fitness, the main variable of population genetics and adaptationist models, cannot fully describe nor explain evolutionary phenomena involving substantial changes in organismal architecture due to non-selective forces. This limitation is twofold: on the one hand, because fitness is by definition a measure of selection, and blind to non-selective processes; on the other hand, because, even with respect to selection, fitness can predict future changes but cannot explain past ones. We have then proposed a conceptual framework that, including robustness as a second dimension of biological design, can overcome the limitations of fitness. Finally, we have introduced a logical model -the design space- that potentially unifies all the evolutionary accounts discussed in the previous chapters, by showing how each of them affects fitness and robustness, and how any evolutionary phenomenon can be interpreted as the sum of many such contributions. Additionally, this logical model has been shown to be coherent with classical contributions such as Fisher's fundamental theorem and fitness landscapes, that turn out to be particular cases of the design space. Finally, we have proposed a classification of evolutionary phenomena based on what the main target of change is (fitness, stability robustness or architectural robustness), we have integrated the classification with ideas about attractors, and we have shown how these different kinds of phenomena can be mapped into design space and into our architectural matrix.

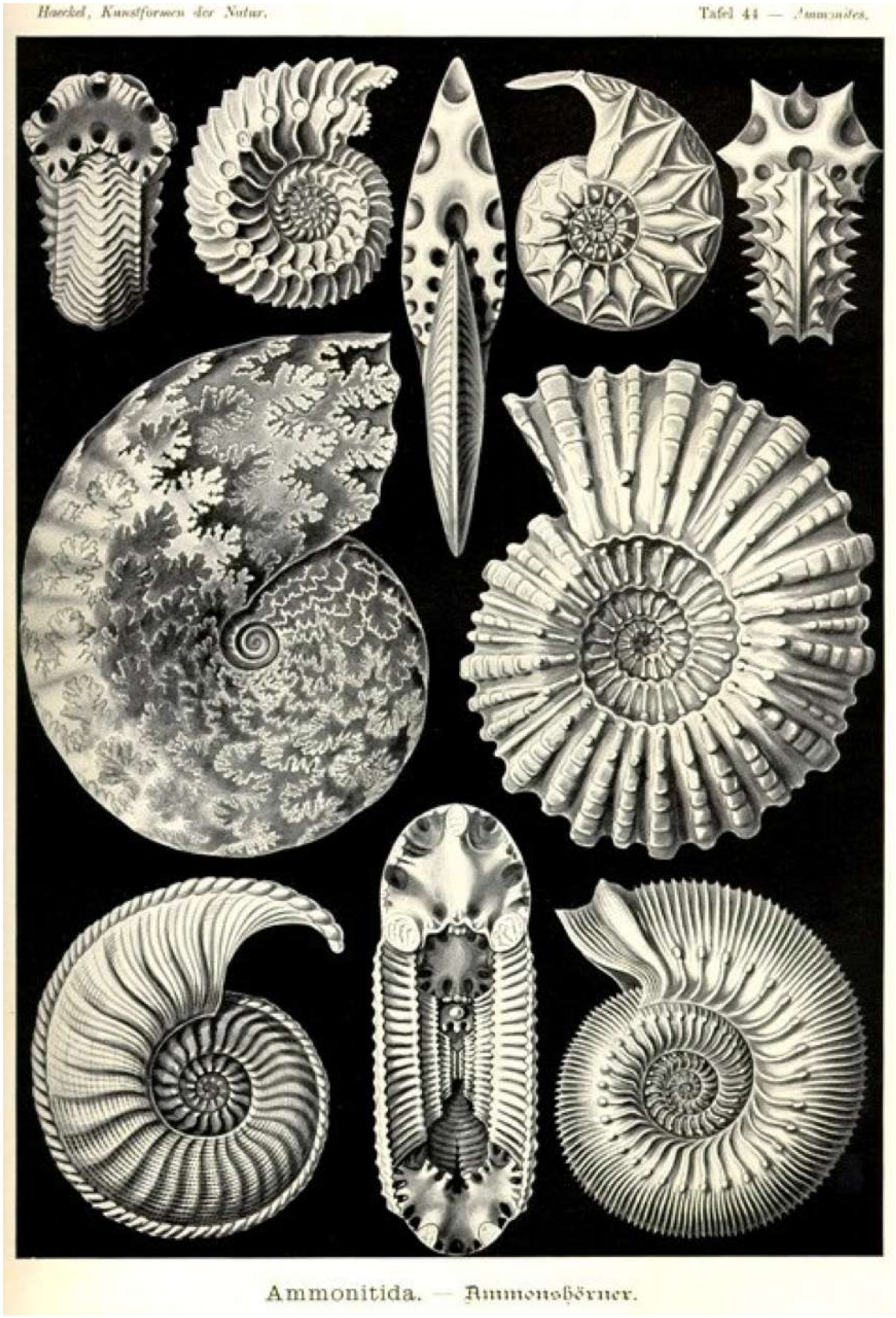


Plate 41 – Ammonitida, from Haeckel, E., *Kunstformen der Natur*. Leipzig und Wien: Verlag des Bibliographischen Instituts,1899-1904.

6. Final thoughts

I conclude this work with some reminders concerning the main issues that I have treated and some final considerations around the problems that I have addressed. I started in chapter 0 by noticing that, of the three explananda of evolutionary biology -variety, complexity and adaptedness of organisms- selection partially explains the first (drift is also a candidate) and is the only known explanation of the last, while it struggles to explain the second. This weakness dates back to Darwin's times, who himself acknowledged that the laws of variation were "various, unknown, or but dimly understood" (1872: 26). He was also convinced that selection of numerous, successive and slight modifications had to account for the appearance of any complex organ, or his whole theory would collapse (1872: 204). Although this claim underlines the role of selection in explaining adaptation without circumscribing to selection the source of variations (1859: 5-6), Darwin's interpretation was vindicated by the confluence of the theory of natural selection and of genetics into the modern synthesis. The modern synthesis crystallised, on one hand, the view of the absolute, and solitary, centrality of selection in evolution:

> [S]election is a creative force that initiates and governs change, providing shape and direction, while variation is merely the source of fuel or raw materials -never a source of initiative, creativity, discontinuity, or direction, as it is in all non-Darwinian theories. (Stoltzfus 2017: 2)

On the other hand, it declared settled the issue of the source of variation, limited to genetic mutations and re-combinations:

> The basic processes of evolution are five: (1) mutation and (2) genetic re-combination which are the sources of variability, but do not provide direction. […] (Stebbins 1966: 12)

But: is it so? The modern synthesis transformed evolutionary biology into a formal science, able to explain and forecast the fate of populations through falsifiable hypotheses and repeatable experiments, such as Lenski's famous *E. coli* long-term evolution investigation (Lenski & Janick 2003). While showing the power of the mechanisms of mutation plus selection to increase adaptedness and define variability, such research seem to

have little to say about complexity: and indeed, they rarely commit themselves with accounts regarding phenotypic design. It is adaptationism in its empirical version, as I have shown in chapter 3, that talks about the latter and declares it a result of adaptedness: behind complexity, there lies selection

only this, and nothing more.[233]

Following this approach, complexity becomes (in Aristotelian terminology) a potency hidden in the environment, and selection the *Deus-ex-Machina* able to actualise it. This view is clear in many adaptationist metaphors. For instance, Dawkins (2006, chapter 3) imagines a monkey in front of a typing machine as an example of a mechanism generating random information (strings of letters). An incredible long amount of time would be required for the monkey to write the sentence "Methinks it is like a weasel"[234] just by typing random strings of 28 letters. The required time, however, dramatically decreases if, each time the monkey types a correct letter in the correct place (for example, an 'e' in the second place), this letter is retained in the following tries. Dawkins got the correct sentence in just 43 'rounds' by programming his PC with this monkey-type mechanism. Indeed, chance generation of variations and selection of correct variations does work great if the target *is already 'written' somewhere*! But: is biological complexity 'written' in the environment, and in the environment only? In other words, is biological complexity just a by-product of adaptedness, and variation a by-product of mutations? And what about the complex structures (such as snow crystals), connected through complex relationships (such as the evaporation, condensation, sublimation in the water cycle) shaping complex systems (such as the global climate on earth) that appear everywhere in the natural world? In chapter 4, I have gone through a great deal of potential sources of variations, from chemical-physical processes, to genetic and phenotypic mechanisms, to complex systems laws. Some of them might seem a bit too imaginative, but it is difficult to discharge all of them as either real but immaterial, or as mere fantasies. These phenomena are determined by physical, chemical, thermodynamic, and complex systems' laws. Can selection really overcome all of them? And above all: why should it? Why not taking advantage of elements that come ready-made, without going through long series of trial-and-error?

Chapter 5 has been devoted to answer these questions. The issue of the relative evolutionary importance of selection and non-selective forces is, no doubt, empirical; but all

---

[233] E. A. Poe, *The Raven*

[234] Shakespeare, Hamlet, Act 3, Scene 2

empirical researches rest on a theory (Popper 2000) and entail some metaphysics (Roger 2005): as a consequence, they might become 'blind-by-design' to phenomena that lay outside of their theoretical frameworks. Fitness is one of the main theoretical concepts embedded within evolutionary biology: while this prominence is justified in studying the sources and effects of selection, it limits beforehand studies around potential sources of variations, that become invisible because they have, by definition, no fitness-effect.

Apart from this blindness, the use of fitness for tracking both survival and reproduction, for both forecasting the future of traits and explaining their present existence, and as a measure of both appearance and selection of variations also leads to a second problem: it transforms the theory of selection into a truism, condensed in the motto "The most adapted is so because it is the one who survives, and the one who survives does so because it is the most adapted".

The roles of fitness as *prima-donn*a in evolutionary biology is, nevertheless, hard to question, as shown by the strong opposition to an extended evolutionary synthesis. One reason is that scalar fitness is a concept easy to measure and define, and therefore allows to easily model populations and their evolution. Another reason, perhaps more relevant, is that fitness is so intrinsically bound to selection that any proposal aimed to limit its relevance sounds like an attack to the explicative power of the latter, something that, to some ears, might suspiciously sound like the endorsement of anti-Darwinian theories, or even creationism.

To escape the truism, and in order to correctly measure the putative role of non-selective forces to the shaping of the phenotype, I suggested that a re-thinking of the theoretical framework of evolutionary research is needed, one that transfers to robustness some of the explanatory roles currently assigned to fitness. The use of two variables allows to expand the classical unidimensional vision of evolutionary phenomena into a bidimensional space, to formalize the two-steps view of evolution as variation-creation plus variation-selection, and to take into account otherwise invisible processes that act logically (and/or temporally) before the selection stage. One immediate result of such a bidimensional account of evolution is the unfolding of interesting connections and unexpected echoes between concepts, e.g. between Wright's adaptive landscape and Kitano's attractors' space. Another interesting result is the possibility to map populations and their evolution on a bidimensional space, and to show how selective and non-selective forces impact on evolutionary phenomena. I have shown this impact for some of the forces

introduced in the previous chapters, and I have suggested how they might have acted in combination, to give rise to a complex new function like flight.

Explaining biological complexity and phenotypic architectures requires knowledge from many fields of research, from molecular to developmental biology, to palaeontology, to name a few: I am well aware that such an ambitious task is titanic and that I have just scratched the surface of the issue. My contribution is inevitably incomplete in scope and limited in perspective, my hypothesis questionable and my conclusions disputable. The choice of robustness as a second dimension of evolutionary change, for example, can be challenged, but I do believe that a second dimension is needed: which one, and how to practically measure it, are tasks that I have not addressed; nor have I addressed the empirical issue of how to measure the actual impact of non-selective forces on robustness. Even with all these limitations, I hope that this work has awaken in the reader the idea that a re-thinking of the theoretical framework of evolutionary research is needed; and I would deem myself satisfied if some of my suggestions contribute in some not completely irrelevant way to the attainment of this goal.

The beginning of the Gutenberg Bible: Volume 1, Old Testament, Epistle of St. Jerome. 1454/55

7. References

Abrams, M. (2012). Measured, modelled, and causal conceptions of fitness. *Frontiers in Genetics*, 3, 196. doi:10.3389/fgene.2012.00196

Abrams, P. (2001). Adaptationism, Optimality Models, and Tests of Adaptive Scenarios. in Orzack, S. H., Sober, E. (eds.). *Adaptationism and Optimality*. Cambridge Studies in Philosophy and Biology. Cambridge: Cambridge University Press:273-302

Abzhanov A., Protas M., Grant B.R., Grant P.R,., Tabin C.J. (2004). Bmp4 and morphological variation of beaks in Darwin's finches. *Science*, 305(5689):1462-5.

Adami, C., Ofria, C., Collier, T. C. (2000). Evolution of biological complexity. *PNAS*, 97(9): 4463-4468

Alberch, P. (1982). Developmental constraints in evolutionary processes. In T. Bonner (ed.), *Evolution and Development*, 313-332, Springer-Verlag.

———, (1991). From genes to phenotypes: dynamical systems and evolvability. *Genetica* 84, 5-11

Altenberg (2005). Modularity in Evolution: Some Low-Level Questions, in Callebaut, W., Rasskin-Gutman, eds. (2005). *Modularity. Understanding the Development and Evolution of Natural Complex Systems*. The MIT Press: 99-128.

Amundson, R. (1994). Two concepts of constraint: adaptationism and the challenge from developmental biology. *Phylosophy of Science*, 61:556-578

———, (2001). Adaptation and Development, in Orzack, S. H., Sober, E. (eds.). *Adaptationism and Optimality*. Cambridge Studies in Philosophy and Biology. Cambridge: Cambridge University Press:303-334

Ancel, L. W., Fontana W. (2000). Plasticity, evolvability, and modularity in RNA. *J Exp Zool.*, 288(3):242-283.

Aparicio, J. M, Bonal, R., Cordero P. J. (2003). Evolution of the structure of tail feathers: implications for the theory of sexual selection. *Evolution*, 57(2):397-405

Appel, T. (1987). *The Cuvier-Geoffroy Debate: French Biology in the Decades before Darwin*. Oxford University Press. ISBN 0-19-504138-0.

Asma, S. T. (1996). *Following form and function: A philosophical archaeology of life science*. Northwestern University Press. ISBN 9780810113978.

Ayala, F. J. (2007). Darwin's greatest discovery: Design without designer. *Proceedings of the National Academy of Sciences*, 104 (1):8567-8573; DOI: 10.1073/pnas.0701072104

Backwell, P. R., Matsumasa, M., Double, M., Roberts, A., Murai, M., Keogh, J. S., Jennions, M. D. (2007). What are the consequences of being left-clawed in a predominantly right-clawed fiddler crab?. *Proceedings. Biological sciences*, 274(1626), 2723-2729. doi:10.1098/rspb.2007.0666

Badyaev, A. V. (2009). Evolutionary significance of phenotypic accommodation in novel environments: an empirical test of the Baldwin effect. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences*, 364(1520), 1125-41.

Baker, A. J., Haddrath, O., McPherson, J. D., Cloutier, A. (2014). Genomic Support for a Moa–Tinamou Clade and Adaptive Morphological Convergence in Flightless Ratites, *Molecular Biology and Evolution*, (31)7: 1686-1696 DOI: 10.1093/molbev/msu153

Baldwin, J. M. (1896). A new factor in evolution. *The American Naturalist*, 30 (354):441-451

Banzhaf, W., Baumgaertner, B., Beslon, G., Doursat, R., Foster, J., McMullin, A., White, L. (2016). Defining and simulating open-ended novelty: Requirements, guidelines, and challenges. *Theory in Biosciences,* 135(3):131-161.

Barve, A., Wagner, A. (2013). A latent capacity for evolutionary innovation through exaptation in metabolic systems*, Nature* 500:203–206

Bębenek, A., Ziuzia-Graczyk, I. (2018). Fidelity of DNA replication-a matter of proofreading. *Current genetics*, 64(5):985-996.

Beloussov, L. (2012). Self-Organization, Symmetry and Morphomechanics in Development of Organisms, in *Embryology - Updates and Highlights on Classic Topics*. OpenAccess, DOI10.13140/2.1.1908.3204

Bennett, A. F. (1987). Interindividual variability: An underutilized resource, in Feder, M. E., Bennett, A. F., Burggren, W. W., Huey, R. B. *New directions in physiological ecology*: Cambridge University Press:147–166.

Bethel, T. (1976). Darwin's Mistake. *Harper's Magazine*, February issue: 70-5.

Bich, L. (2018). Robustness and autonomy in biological systems: how regulatory mechanisms enable functional integration, complexity and minimal cognition through the action of second-order control constraints. In M. Bertolaso, S. Caianiello & E. Serrelli (Eds.), *Biological Robustness. Emerging Perspectives from within the Life Sciences* :123-147. New York: Springer. Doi: 10.1007/978-3-030-01198-7_6.

Bidder, G. P. (1923). *The relation of the form of a sponge to its currents*. Quarterly Journal of Microscopical Science 67: 293–325.

———, (1937). The perfection of sponges. *Proceedings of the Linnean Society* 149: 119–146.

Biewener, A. A., Patek, S. N. (2003). Animal locomotion. Oxford University Press.

Birch, J. (2012), The negative view of natural selection, *Studies in History and Philosophy of Biological and Biomedical Sciences* 43:569–573

———, (2014). Has Grafen formalized Darwin?, *Biol Philos*, 29: 175-180

———, (2018). Fitness Maximization. In Richard Joyce (ed.), *The Routledge Handbook of Evolution and Philosophy*. London: Routledge: 49-63

Blair D. F., Dutcher S. K. (1992). Flagella in prokaryotes and lower eukaryotes. *Curr Opin Genet Dev.,* 2(5):756-67.

Bonner, J. T. (1988). *The Evolution of Complexity by Means of Natural Selection*. Princeton University Press.

———, (1998), The origins of multicellularity. *Integr. Biol.*, 1: 27-36.

Boon, E. (2017). In what sense is Lateral Gene Transfer random, and why does it matter?, abstract of the conference *On the nature of variation: random, biased and directional*, 3 & 4 October 2017, University of Lisbon

Bourke, A. F. G. (2014), "The gene's-eye view, major transitions and the formal Darwinism project", *Biol Philos*, 29: 241-248

Bowler, P. J. (1992). Darwinism and Victorian Values: Threat or Opportunity? *Proceedings of the British Academy*, 78:129-147

Brandon R. N. (1990). *Adaptation and environment*. Princeton University Press

———, (2005), Evolutionary Modules: Conceptual Analyses and Empirical Hypotheses, in Callebaut, W., Rasskin-Gutman, eds. (2005). *Modularity. Understanding the Development and Evolution of Natural Complex Systems*. The MIT Press:51-60

———, (2006). The principle of drift: Biology's first law. *Journal of Philosophy*, 103(7), 319-335. doi:10.5840/jphil2006103723

Brigandt, I., Love, A. C. (2012). Conceptualizing evolutionary novelty: moving beyond definitional debates. *J. Exp. Zool. (Mol. Dev. Evol.)* 318B:417-427.

Brooker, R. J. (1999). *Genetics. Analysis and Principles*. Benjamin Cummings, Addison Wesley Longman

Brooks, D. R., Wiley, E. O. (1986). *Evolution as entropy*. The University of Chicago Press.

Brooks, D. R., McLennan, D. A. (1991). *Phylogeny, ecology, and behavior*. The University of Chicago Press.

Brown, J. S. (2001). Fit of Form and Function, Diversity of Life, and Procession of Life as an Evolutionary Game, in Orzack, S. H., Sober, E. (eds.). *Adaptationism and Optimality*. Cambridge Studies in Philosophy and Biology. Cambridge: Cambridge University Press:114-160

Brush, A. H. (1996). On the origin of feathers. *J Evol Biol* 9:131-142

Buchanan, A. V., Sholtis, S., Richtsmeier, J., & Weiss, K. M. (2009). What are genes "for" or where are traits "from"? What is the question?. *BioEssays : news and reviews in molecular, cellular and developmental biology*, 31(2):198-208.

Bürger R. (2011). Some Mathematical Models in Evolutionary Genetics. In: Chalub F., Rodrigues J. (eds) *The Mathematics of Darwin's Legacy. Mathematics and Biosciences in Interaction*. Springer, Basel:67-89

Bürger, R., Willensdorfer, M., Nowak, M. A. (2006). Why are phenotypic mutation rates much higher than genotypic mutation rates?. *Genetics*, 172(1):197-206.

Burke, A. C. (1989). Development of the Turtle Carapace: Implications for the Evolution of a Novel Bauplan. *Journal of Morphology*, 199:363-378

Cain, A. J. (1989). The perfection of animals. *Biological Journal of Linnean Society*, 36:3-29

Cain, A. J., Sheppard, P. M. (1954). Natural selection in *Cepae*. *Genetics*. 39:89-116

Cairns J., Foster P.L. (1991). Adaptive reversion of a frameshift mutation in Escherichia coli. *Genetics* 128:695–701.

Calcott, B. (2008). Assessing the fitness landscape revolution. *Biol Philos*, 23:639–657

———, (2014). Engineering and evolvability. *Biol Philos*, 29(3):293-313. doi:10.1007/s10539-014-9425-3

Callebaut, W., Rasskin-Gutman, eds. (2005). *Modularity. Understanding the Development and Evolution of Natural Complex Systems*. The MIT Press

Canfield, J. V. (1990). The concept of function in biology. Philosophical Topics, 18(2):29-53

Carlson, J. M., Doyle, J. (1999). Highly Optimized Tolerance: A Mechanism for Power Laws in Designed Systems. *Phys. Rev.* E 60:1412

———, (2002). Complexity and Robustness. *PNSA*, vol. 99, 2538-2545

Chapuis, A., Droz, E. (1949). *Les automates. Figures artificielles d'hommes et d'animaux. Histoire et technique*. Neuchâtel: Editions du Griffon.

Charlat, S. et al. (2007). Extraordinary Flux in Sex Ratio, Science, 317(5835):214

Charlesworth, B., Lande, R. (1982). A neo-darwinian commentary on macroevolution. *Evolution,* 36: 474-498

Charlesworth, D., Barton N. H., Charlesworth B. (2017). The sources of adaptive variation. *Proc. R. Soc.* B284: 20162864. http://dx.doi.org/10.1098/rspb.2016.2864

Charnov, E. L. (1976). Optimal Foraging, the Marginal Value Theorem, *Theoretical Population Biology*, Vol. 9, No. 2, April 1976

Cherniak, C. (2009a). Minimal rationality and optimal brain wiring. In Glymour, C; Wei, W; Westerstahl, D. (eds.) *Logic, Methodology and Philosophy of Science: Proceedings of 13th International Congress.* College Publications. pp. 443–454

Cherniak, C., (2009b). Brain wiring optimization and non-genomic nativism. In M. Piattelli-Palmarini, M., Uriagereka, J., Salaburu, P. (eds), *Of Minds and Language: a Dialogue with Noam Chomsky in the Basque Country*, Oxford, Oxford University Press, 2009:108-19

Cherniak, C.; Changizi, M; Kang, Du Won (1999). Large-scale optimization of neuron arbors. *Physical Review E.* 59: 6001–6009

Cherniak, C.; Mokhtarzada, Z; Rodriguez-Esteban, R; Changizi, B (2004). Global optimization of cerebral cortex layout. *Proc. Natl. Acad. Sci. U.S.A.* 101: 1081–1086

Cherniak, C., Rodriguez-Esteban, R. (2013). Body maps on the human genome. *Mol. Cytogenet.* 6 (1): 61

Chouard, T. (2010). Revenge of the hopeful monster. *Nature*, 463(18): 864-867

Christensen, W. D., Bickhard, M. H. (2002). The Process Dynamics of Normative Function, *The Monist*, 85(1):3-28.

Clausen, J., D. D. Keck, and W. W. Hiesey (1958). Environmental Responses of Races of Achillea. *Carnegie Inst. Wash. Pub.*, 581:1-129.

Colleen F., (1997). Did Lungs and the Intracardiac Shunt Evolve to Oxygenate the Heart in Vertebrates?. *Paleobiology. Paleontological Society.* 23 (3): 358–372

Colless, D. H. (1985). On "character" and related terms, *Systematic Biology*, 34(2):229–233

Collier, J. (1986). Entropy in Evolution, *Biol. Philos.* 1: 5-24

Cook L. M., Saccheri I. J. (2013). The peppered moth and industrial melanism: evolution of a natural selection case study, *Heredity*, 110, 207–212

Cope, E. D. (1887). The Origin of the Fittest. D. Appleton & Co.

Craig, L.R. (2010). The So-Called Extended Synthesis and Population Genetics. *Biol Theory*, 5(2):117-123. DOI: 10.1162/BIOT_a_00035

Crampton H. E. (1916). Studies on the variation, distribution and evolution of the genus Partula. The species inhabiting Tahiti. *Carnegie Institution of Washington*, 228: 1-311.

———, (1932). Studies on the variation, distribution and evolution of the genus Partula. The species inhabiting Moorea. *Carnegie Institution of Washington*, 410: 1-335.

Craver, C. F. (2001). Role Functions, Mechanisms, and Hierarchy, *Philosophy of Science,* 68:53-74

Crick, F. H. C. (1958). On Protein Synthesis. In Sanders, F. K. (ed.). *Symposia of the Society for Experimental Biology, Number XII: The Biological Replication of Macromolecules*. Cambridge University Press:138–163.

Crispo, E. (2007). The Baldwin effect and genetic assimilation: revisiting two mechanisms of evolutionary change mediated by phenotypic plasticity. *Evolution* 61-11:2469-2479

Daniels, B. C., Chen, Y. J., Sethna, J. P., Gutenkunst, R. N., Myers, C. R. (2008). Sloppiness, robustness and evolvability in systems biology, *Current Opinion in Biotecnology*, 19: 389-395

Darwin, C. R. (1859). *On the Origin of Species*. John Murray, London.

———, (1872, 6th Edition). *On the Origin of Species*. . John Murray, London.

———, (1883). Essay on Instinct. In Romanes, G. J. *Mental evolution in animals. With a posthumous essay on instinct by Charles Darwin*. London: Kegan Paul Trench & Co.:355-384

Davis, G. K., Patel, N. H. (1999). The origin and evolution of segmentation. *Trends Cell Biol.*, 9(12):M68-72

Davies J. A. (2016) Machines for living in: connections and contrasts between designed architecture and the development of living forms. *Architectural Research Quarterly* 20(1):45-50

Davies, J. A. (2017, March 17). Biological Architecture. Retrieved 2019, April 16, from www.ed.ac.uk/biomedical-sciences/postgraduate-studying/biomedical-sciences/structure/themes/biological-architecture

Davies, N. B., Krebs, J. R., and West, S. A. (2012). *An Introduction to Behavioural Ecology*. Wiley-Blackwell, Hoboken, NJ, 4th edition.

Dawkins, R. (1982). *The extended phenotype*. Oxford University Press

———, (1986). *Blind Watchmaker*, W.W. Norton and Company, New York.

———, (2006). *The Selfish Gene*. Oxford University Press

———, (2009). *The Greatest Show on Earth*. Free Press

Dawson, T. H. (2014). Allometric Relations and Scaling Laws for the Cardiovascular System of Mammals. *Systems*, 2:168-185; doi:10.3390/systems2020168

Den Boer, P. J. (1999). Natural Selection or the non-survival of the non-fit. *Acta Biotheoretica* 47:83-97

Dennet, D. C. (1995). *Darwin's Dangerous Idea*, Penguin Books

Denton, M. (1985). *Evolution: A Theory in Crisis*. Burnett Books, London.

Depew, D. J. (2003), Baldwin Boosters, Baldwin Skeptics. In: Weber, B. H., Depew, D. J., *Evolution and learning: The Baldwin effect reconsidered*. MIT Press

Dial, K. P. (2003). Wing-assisted incline running and the evolution of flight. *Science*, 299:402-404. doi: 10.1126/science.1078237.

Dietrich, M. R. (2003). Richard Goldschmidt: hopeful monsters and other 'heresies'. *Nature Reviews, Genetics*, 4:68-74

Dobzhansky, Th. (1958). *Genetics and the Origin of Species*. Third Edition. Columbia University Press, New York.

-————, (1968). On Some Fundamental Concepts of Darwinian Biology. In Dobzhansky, Th.; Hecht, M. K.; Steere, W. C. *Evolutionary Biology*, 2. Appleton-Century-Crofts.

Dollo, L. (1893). Les lois de l'évolution. *Bull. Soc. Belge Geol. Pal. Hydr.* VII:164-166

Dover, G. (2002). Molecular Drive. *Trends in Genetics* 18(11): 587-589

Draghi, J. A., Parsons, T. L., Wagner, G. P., Plotkin, J. P. (2010). Mutational robustness can facilitate adaptation, *Nature*, 463: 353-355

Duncan W., Best J., Golubitsky M., Nijhout H. F., Reed (2018). Homeostasis despite instability. *Math Biosci.* 300:130-137. doi: 10.1016/j.mbs.2018.03.025

Dupont C., Armant D. R., Brenner C. A. (2009). Epigenetics: definition, mechanisms and clinical perspective. *Seminars in Reproductive Medicine*. 27 (5): 351–357

Dupré, J. (2003). *Darwin's Legacy: What Evolution Means Today*. Oxford University Press.

Dwyer, P. (1984). Functionalism and Structuralism: Two Programs for Evolutionary Biologists. *The American Naturalist*, 124(5):745-750

Edwards, A. W. F., (1994). The fundamental theorem of natural selection. *Biological Reviews*, 69:443-474

————, (2007). Maximisation principles in evolutionary biology. In Matthen, M., Stephens, C., (eds.), *Handbook of the Philosophy of Science: Philosophy of Biology*. North-Holland, Amsterdam: 335-347.

Eldredge, N., Gould, S. J. (1972), Punctuated equilibria: an alternative to phyletic gradualism, in *Models in paleobiology*, Shopf, TJM Freeman, Cooper / Co.

———, (1977). Evolutionary models and biostratigraphic strategies. In Kauffman, E. G. & Hazel, J. E. (eds.), Concepts and Methods of Biostratigrpahy. Stroudsburg Pa: Dowden, Hutchinson & Ross: 25-40

Endler, J. (1986). *Natural Selection in the Wild*. Princeton: Princeton University Press.

Erwin, D. H. (2015). Novelties and Innovations in the History of Life. *Current Biology* 25, R930–R940

Erwin, D. H., Anstey, R. L. (1995). *Speciation in the fossil record*. In *New approaches to speciation in the fossil record*. Edited by Douglas H. Erwin and Robert L. Anstey (eds.):11-38. New York: Columbia Univ. Press.

Eshel, I., Feldman, M. W. (1984). Initial rate of increase of new mutants and some continuity properties of ESS in two-locus systems. *American Naturalist* 124: 631–640

———, (2001). Optimality and evolutionary stability under short-term and long-term selection. In: Orzack, S. H. and Sober, E. (eds), *Adaptationism and Optimality,* Cambridge University Press:161-190.

Ewens, W. J. (2004). *Mathematical population genetics, I: theoretical introduction*. Springer, New York.

———, (2014), Grafen, the Price equations, fitness maximization, optimization and the fundamental theorem of natural selection, *Biol Philos*, 29: 197-205

Fisher, R. A. (1930). The genetical theory of natural selection. Oxford Clarendon Press.

———, (1941). Average excess and average effects of a gene substitution, *Annals of Eugenics*, 11:53-63.

Fodor, J., Piattelli-Palmarini, M. (2010). *What Darwin got wrong*. Profile Books

Forber, P. (2005). On the explanatory roles of natural selection. *Biology and Philosophy*, 20:329–342.

Foster, P. L. (2007). Stress-Induced Mutagenesis in Bacteria. *Critical Reviews in Biochemistry and Molecular Biology*, 42(5):373–397

Freeman, A. S., Byers, J. E. (2006). Divergent Induced Responses to an Invasive Predator in Marine Mussel Populations, *Science* 313: 831-833

Futuyma, D. J. (1987). On the role of species in anagenesis. *Am Nat* 130(3):465–473

Gardner, A. (2008), "The Price equation", *Current Biology*, Vol. 18 No 5, 198-202

———, (2009). Adaptation as organism design. *Biology Letters*, 5(6), 861. doi:10.1098/rsbl.2009.0674

———, (2014a), Life, the universe and everything, *Biol Philos*, 29: 207-215

———, (2014b), Adaptation of individuals and groups, in Bouchard, F., Huneman, P. (eds.), *From Groups to Individuals*, The MIT Press: 99-116

Gardner, A., Grafen, A. (2009). Capturing the superorganism: a formal theory of group adaptation. *J Evol Biol* 22:659-671.

Gehring W.J. (1996). The master control gene for morphogenesis and evolution of the eye. *Genes Cells*. (1):11-5.

Gilchrist, G. W., Kingsolver, J. G. (2001). Is Optimality Over the Hill?, in Orzack, S. H., Sober, E. (eds.). *Adaptationism and Optimality*. Cambridge Studies in Philosophy and Biology. Cambridge: Cambridge University Press:45-63

Gillies, D. (2000). Varieties of Propensity, *British Journal for the Philosophy of Science*, 51:807-835.

Gillespie, J. H. (1981). Mutation modification in a random environment. *Evolution*, 35(3):468-476.

Godfrey-Smith, P. (1994). A Modern History Theory of Functions, *Nous*, 28: 344-362.

———, (1999). Adaptationism and the Power of Selection. *Biology and Philosophy* 14:181-194

———, (2001). Three kinds of adaptationism. In: Orzack S. H. and Sober E. (eds), *Adaptationism and Optimality*, Cambridge University Press: 335-357

Goldschmidt, R. (1940). *The Material Basis of Evolution*. Yale University Press

Gould, S. J. (1970). "Dollo on Dollo's law: irreversibility and the status of evolutionary laws". *Journal of the History of Biology*. 3:189-212. doi:10.1007/bf00137351.

———, (1973). Betting on Chance – and No Fair Peeking. In Eight Little Piggies: Reflections in Natural History, Norton:396-408

———, (1976). Darwin's Untimely Burial. Nature, Vol. 85, No. 8: 24-30

———, (1977). Eternal Metaphors of Paelontology, in A. Hallam (ed.), *Patterns of Evolution*, Elsevier: 1-26

———, (1980), The Return of the Hopeful Monsters, from *The Panda's Thumb*, Norton & Co.:186-193

———, (1988). Kropotkin was no crackpot. *Natural History,* 97:12-21.

———, (1989). *Wonderful Life. The burgess shala and nature of history*. W. W. Norton & Co.

———, (1990). Speciation and sorting as the source of evolutionary trends, or 'things are seldom what they seem'. In McNamara, J. C. (ed.) *Evolutionary Trends*. University of Arizona Press:3-27

———, (1993). *Eight little piggies*. W. W. Norton & Company

———, (1996). *Full house: The Spread of excellence from Plato to Darwin*. New York: Harmony Books; 1996.

———, (1997). Evolution: The pleasures of pluralism. *The New York Review of Books* 44 (June 26): 47–52.

———, (2001). *The lying stones of Marrakech : penultimate reflections in natural history*. Vintage:119–121. ISBN 978-0-09-928583-0.

———, (2002). *The Structure of Evolutionary Theory*. The Belknap Press of Harvard University Press

Gould, J. S., Eldredge, N. (1977). Punctuated equilibria: the tempo and mode of evolution reconsidered. *Paleobiology* 3:115-151

———, (1993). Punctuated equilibrium comes of age. *Nature*, 534:102-105

Gould S.J., Lewontin S.J. (1979), The spandrels of San Marco and the Panglossian paradigm: a critique of the adaptationist program, *Proc. R. Soc. Lon.*, B 205, 581-598

Gould, J.S., Vrba, E. S. (1982), Exaptation – A Missing Term in the Science of Form, *Paleobiology*, Vol. 8,No. 1: 4-15

Grafen, A. (1985) Geometric view of relatedness, *Oxford Surveys in Evolutionary Biology*, 2: 28-90

———, (1999) Formal Darwinism, the individual-as-maximizing-agent and bet-edging, *Proc. R. Soc. Lond. B*, 266, 799-803

———, (2000) Development of the Price equation and natural selection under uncertainty *Proc. R. Soc. Lond. B* , 267, 1223-1227

———, (2002) A first formal link between the Price equation and an optimization program, *J. theor. Biol.* 217, 75-91

———, (2003), Fisher the evolutionary biologist, *The Statistician*, 52, Part 3, 319-32

———, (2006a) Optimization of inclusive fitness. *J. theor. Biol.* 238, 541-563

———, (2006b) A theory of Fisher reproductive value. *J. Math. Biol.* 53, 15-60

———, (2007) The Formal Darwinism project: a mid-term report, *J. Evol. Biol.,* 1243-1254

———, (2008) The simplest formal argument for fitness optimization. *Journal of Genetics*, Vol. 87 No. 4

———, (2009) Formalizing Darwinism and inclusive fitness theory, *Phil. Trans. R. Soc. B* 364, 3135-3141

———, (2014a) The formal Darwinism project in outline, *Biol Philos*, 29: 155-174

———, (2014b) The formal Darwinism project in outline: response to commentaries, *Biol Philos*, 29: 281-292

———, (2015) Biological Fitness and the Fundamental Theorem of Natural Selection, *The American Naturalist*, 186(1):1-14

Grant, P.R., Grant, B.R. (1995) Predicting microevolutionary responses to directional selection on heritable variation, *Evolution* 49:241–251

Gregory, T.R. (2008a). Evolutionary trends. *Evo Edu Outreach*, 1:259–273.

-———, (2008b). The Evolution of Complex Organs. *Evo Edu Outreach*, 1: 358-389. DOI 10.1007/s12052-008-0076-1

Grinnell, J. (1904). The origin and distribution of the chest-nut-backed chickadee. *The Auk*, 21 ,364–382.

Grodwohl, (2016). "The Theory was Beautiful Indeed": Rise, Fall and Circulation of Maximizing Methods in Population Genetics (1930–1980). *Journal of the History of Biology*, 50(3):571-608 DOI 10.1007/s10739-016-9449-4

Haig, D. (2014), "Genetic dissent and individual compromise", *Biol Philos*, 29: 233-239

Hairston N.G. Jr, Ellner S.P., Geber M.A., Yoshida T., Fox J.A. (2008). Rapid evolution and the convergence of ecological and evolutionary time. *Ecol Lett.*;8:1114–1127.

Halama, K. J., Reznick, D. N. (2001). Adaptation, Optimality, and the Meaning of Phenotypic Variation in Natural Populations, in Orzack, S. H., Sober, E. (eds.). *Adaptationism and Optimality*. Cambridge Studies in Philosophy and Biology. Cambridge: Cambridge University Press:242-272

Haldane, J. B. (1932). *The Causes of Evolution*. Longmans Green.

———, (1958). *The Unity and Diversity of Life*. (Sardar Vallabhbhai Patel Lectures). Delhi: Ministry of Information and Broadcasting, Government of India.

Hall, B. K. (2003a). Evo-Devo: evolutionary developmental mechanisms. *Int. J. Dev. Biol.* 47: 491-495.

———, (2003b). Unlocking the Black Box between Genotype and Phenotype: Cells and Cell Condensations as Morphogenetic (modular) Units. *Biol. & Philos.*, 18: 219-247.

Hamilton, W. D. (1964). The genetical evolution of social behaviour. *J Theor Biol* 7:1–52

Hamilton, M. B. (2009). *Population Genetics*. John Wileys & Sons.

Harvey, P. H., Pagel, M. D. (1991). *The comparative method in evolutionary biology*. Oxford: Oxford University Press.

Hatsopoulos, G., Keenan, J. (1965). *Principles of General Thermodynamics*. John Wyley

Hendry, A.P. & Kinnison, M.T. (1999). The pace of modern life: measuring rates of contemporary microevolution. *Evolution*, 53, 1637–1653.

Hegel, G. W. F. (1986). *Werke*, 20 vols. Moldenhauer, E. and Michel, K. M. (eds.). Frankfurt am Main: Suhrkamp.

Heimans, J. (1962). Hugo de Vries and the Gene Concept. *The American Naturalist*, 96(887), 93-104. Retrieved from http://www.jstor.org/stable/2458335

Held, L. I. (2014). *How the snake lost its legs: curious tales from the frontier of evo-devo.*: Cambridge University Press.

Hermisson, J. and Wagner, G. P. (2005). Evolution of Phenotypic Robustness, in Jen, E. (ed.). *Robust Design: A Repertoire of Biological, Ecological, and Engineering Case Studies*. Oxford University Press: 47-70

Herrel A, Huyghe K, Vanhooydonck B, et al. (2008). Rapid large-scale evolutionary divergence in morphology and performance associated with exploitation of a different dietary resource. *Proceedings of the National Academy of Sciences of the United States of America*. 105(12):4792-4795. doi:10.1073/pnas.0711998105.

Hines, W. G. S., Bishop, D. T. (1983). Evolutionarily stable strategies in diploid populations with general inheritance patterns. *Journal of Applied Probability* 20: 395–399.

Hinton, G.E., Nowlan, S.J. (1987). How Learning Can Guide Evolution. *Complex Systems*, 1:495-502.

Hitchcock, C. (2002). Probability and Chance. In the *International Encyclopedia of the Social and Behavioral Sciences* (Volume 18), Elsevier:12089-12095.

Hollinger, H. B., Zenzen, M. J. (1982). An interpretation of macroscopic irreversibility within the Newtonian framework. *Philos. Sci.* 49:309-354

Houle, D. (1992). Comparing evolvability and variability of quantitative traits, *Genetics* 130, 195-204

Hull, D. L., Ruse, M. (2007). *The Cambridge Companion to Philosophy of Biology*. Cambridge University Press.

Huneman, P. (2014), "Formal Darwinism as a tool for understanding the status of organism in evolutionary biology", *Biol Philos*, 29: 271-279

Hutchinson, G. E. (1957). Concluding Remarks. *Cold Spring Harbor Symposia on Quantitative Biology*, 22 , 415–427.

Huxley, J. (1932). *Problems of Relative Growth*. MacVeagh

———, (1936). Natural selection and evolutionary progress. *Nature*, 3:571-573

———, (1942). *Evolution, the modern synthesis*. Allen and Unwin.

Jablonka, E., Lamb, M. J. (2005). *Evolution in Four Dimensions. Genetic, Epigenetic, Behavioral, and Symbolic Variation in the History of Life*. The MIT Press.

Jablonka, E., Lamb, M. J., Avital, E. (1998). Lamarckian' mechanisms in darwinian evolution. *Trends in Ecology & Evolution*, Volume 13 , Issue 5: 206-210

Jablonski, D., Bottjer, D.J. (1990). The origin and diversification of major groups: environmental patterns and macroevolutionary lags. In *Major Evolutionary Radiations*, P.D. Taylor, and G.P. Larwood, (eds.), Clarendon Press:17-57.

Jacob, F. (1977). Evolution and tinkering. *Science,* 196(4295):1161-6.

Jacob, F., Monod, J. (1961). Genetic Regulatory Mechanisms in the Synthesis of Proteins. *J. Mol. Biol.* 3:318-356

Janis, C. M. (1982). Evolution of horns in ungulates - ecology and palaeoecology. *Biological Review (Cambridge Philosophical Society),* 57 (2):261-317

Jeffrey, W. R. (2005). Adaptive evolution of eye degeneration in the Mexican blind cavefish. *J. Hered.* 96:185-196

Jen, E. (2005). Stable or Robust? What's the Difference?, in Jen, E. (ed.). *Robust Design: A Repertoire of Biological, Ecological, and Engineering Case Studies*. Oxford University Press: 7-20

Johnson, J. G. (1975) Allopatric Speciation in Fossil Brachiopods. *J. Paleontol,* 49:646–661

———, (1982). Occurrence of phyletic gradualism and punctuated equilibria through geologic time. *J. Paleontol.* 56:1329-1331

Johnson, N. (2010). *Simply Complexity: a Clear Guide To Complexity Theory*. Oneworld Publications

Jordão, J. M., Oliveira, R. F. (2001). Sex differences in predator evasion in the fiddler crab Uca tangeri (decapoda: ocypodidae), *Journal of crustacean biology*, 21(4): 948-953

Kaplan, J. (2008). The end of the adaptive landscape metaphor? *Biol Philos* 23:625–638

Karn, M. N., Penrose, L. s. (1951). Birth weight and gestation time in relation to maternal age, parity and infant survival. *Theor. Pop. Biol.* 6:355-382

Kauffman, S. A. (1969). Metabolic stability and epigenesis in randomly produced genetic nets. *J. Theor. Biol.* 22:437-467

———, (1983). Developmental constraints: Internal factors in evolution. In Goodwin, I., Holder, N., and Wylie, C. C. (eds.), *Development and Evolution*, Cambridge University Press:195-225

———, (1991). Antichaos and Adaptation. *Scientific American*, August:78-84

———, (1993). *The Origins of Order. Self-Organization and Selection in Evolution.*. Oxford University Press.

———, (2000). *Investigations*. Oxford University Press.

Kelly, J. K. (2011) The breeder's equation. *Nature Education Knowledge* 4(5):5

Kestin, J. (1966). *A Course in Thermodynamics*. Blaisdell

Kettlewell, H. B. D. (1955). Selection experiments on industrial melanism in the Lepidoptera. *Heredity* 10: 287–301.

Khammash, M. (2016). An engineering viewpoint on biological robustness. *BMC Biology*. DOI 10.1186/s12915-016-0241-x

Kimura, M. (1960). Optimum mutation rate and degree of dominance as determined by the principle of minimum genetic load. *Journal of Genetics*, 57:21-34.

———, (1977). Preponderance of synonymous changes as evidence for the neutral theory of molecular evolution. *Nature* 267: 275-276

———, (1983). *The neutral theory of molecular evolution*. Cambridge University Press

King, N. (2004). The unicellular ancestry of animal development. *Dev Cell.* 7(3):313-25.

Kingsolver, J., Koehl, M. (1985), Aerodynamics, Thermoregulation, and the Evolution of Insect Wings – Differential Scaling and Evolutionary Change. *Evolution*, 39:488-504

Kitano, H. (2004). Biological robustness. *Nature Reviews Genetics*, 5(11), 826-837. doi:10.1038/nrg1471

———, (2007). Towards a theory of biological robustness. *Molecular Systems Biology*, 3(1), 137-n/a. doi:10.1038/msb4100179

Kolmogorov, A. N. (1965). Three approaches to the concept of the amount of information. *IEEE Prob. Info. Trans.* 1: 1-7

Konishi, M., Volman, S. F. (1990). Comparative Physiology of Sound Localization in Four Species of Owls, *Brain Behavior and Evolution*, 36:196-215

———, (1994). Adaptations for bi-coordinate sound localization in owls. In: Schildberger, K. And Elsner, N. (Eds.) , Neural Basis of Behavioral Adaptations. *Progress in Zoology*;39:1–11.

Koonin, E. V. (2009). Darwinian Evolution in the light of genomics, *Nucleic Acids Res,* 37 (4): 1011-1034

———, (2012a). Does the Central Dogma Still Stand? *Biology Direct* 7:27

———, (2012b). *The logic of chance*. FT Press Science.

Koonin, E. V., Wolf, Y. I (2009). Is evolution Darwinian or/and Lamarckian? *Biology Direct*, 4:42

Korzybski, A- (1933). *Science and Sanity. An Introduction to Non-Aristotelian Systems and General Semantics*. The International Non-Aristotelian Library Pub. Co.:747–761.

Krakauer, D. C., Plotkin, J. B. (2004). Principles and Parameter of Molecular Robustness, in Jen, E. (ed.). *Robust Design: A Repertoire of Biological, Ecological, and Engineering Case Studies*. Oxford University Press: 7-20

Krebs, J. R., McCleery, R. H. (1984). Optimization in behavioral ecology. In Krebs, J. R., N. B. Davies, N. B. (eds.). *Behavioural ecology: An evolutionary approach*: 91–121. Oxford: Blackwell Scientific Publications.

Krimbas, C. B. (2004). On fitness. *Biol Philos*, 19(2), 185-203. doi:10.1023/B:BIPH.0000024402.80835.a7

Kropotkin, P. (2006). *Mutual Aid. A Factor of Evolution*. Dover Publications, Inc.

Kuhn, T. (1962). *The Structure of Scientific Revolutions*. Chicago: University of Chicago Press

Laland K.N., Uller T., Feldman M.W., Sterelny K., Müller G.B., Moczek A., Jablonka E., Odling-Smee J. (2015). The extended evolutionary synthesis: its structure, assumptions and predictions. *Proc. R. Soc.* B 282: 20151019. http://dx.doi.org/10.1098/rspb.2015.1019

Lamb, B. C. (1975), Cryptic mutations: Their predicted biochemical basis, frequencies and effects on gene conversion, *Molec. gen. Genet.*, 137:305-314

Lamotte, M. (1959). Polymorphism of Natural Populations of *Cepaea Nemoralis*. *Cold Spring Harbor Symposia on Quantitative Biology*, 24:65-84.

Lande, R. (1976). Natural selection and random genetic drift in phenotypic evolution. *Evolution* 30: 314–334.

Lahti, D. C., Johnson, N. A, Ajie, B. C., Otto, S. P., Hendry, A. P, Blumstein, D. T., Coss, R. G., Donohue, K., Foster, S. A. (2009). Relaxed selection in the wild. *Trends in Ecology & Evolution*, 24(9):487-496. DOI:10.1016/j.tree.2009.03.010

Lee D.E., Cavener D.R., Bond M.L. (2018) Seeing spots: quantifying mother-offspring similarity and assessing fitness consequences of coat pattern traits in a wild population of giraffes (Giraffa camelopardalis), *PeerJ*, 6:e5690 https://doi.org/10.7717/peerj.5690

Lees, D. R. (1981). Industrial melanism: Genetic adaptation of animals to air pollution. In *Genetic consequences of man made change*, Bishop, J. A., Cook, L. M. (eds.). London: Academic Press:129–176

Lehmann, L., Rousset F. (2014), Fitness, inclusive fitness, and optimization, *Biol Philos*, 29: 181-195

Leigh, E. G. jr. (2001). Adaptation, Adaptationism, and Optimality, in Orzack, S. H., Sober, E. (eds.). *Adaptationism and Optimality*. Cambridge Studies in Philosophy and Biology. Cambridge: Cambridge University Press:358-387

Lenski, R. E., Barrick, J. E., Ofria, C. (2006). Balancing robustness and evolvability. *PLoS Biology*, 4(12), e428. doi:10.1371/journal.pbio.0040428

Lenski, R. E., Janick, J. (ed.) (2003) Phenotypic and Genomic Evolution during a 20,000-Generation Experiment with the Bacterium Escherichia coli. *Plant Breeding Reviews*. New York: Wiley. 24 (2): 225–65. doi:10.1002/9780470650288.ch8.

Lewens, T. (2007), Adaptation, in Hull, D. L., Ruse, M. (eds.), *The Cambridge Companion to Philosophy of Biology*. Cambridge University Press: 1-21

Lewontin, R. C. (1967). Spoken remark, in Moorhead, P. S., Kaplan, M. (eds.) *Mathematical Challenges to the Neo-Darwinian Interpretation of Evolution*. Wistar Institute Symposium Monograph, 5:79

———, (1968). The concept of evolution. *The International Encyclopedia of the Social Sciences*, 5:202–209.

———, (1974). *The Genetic Basis of Evolutionary Change*. Columbia University Press.

———, (1979), Sociobiology as an adaptationist program. *Syst. Res.*, 24: 5-14.

———, (1983), The Organism as the Subject and Object of Evolution. *Scientia* 77 (18):65

———, (1984). Adaptation. In Sober, E. (ed.), *Conceptual Issues in Evolutionary Biology*, MIT Press:235-251

———, (1987). The shape of optimality. In *The latest on the best*, Dupré, J. (ed.), MIT Press: 151-159

Lewontin, R. C. , Goss P. J. E. (2005). Developmental Canalization, Stochasticity, and Robustness, in Jen, E. (ed.). *Robust Design: A Repertoire of Biological, Ecological, and Engineering Case Studies*. Oxford University Press: 21-46

Lieberman, B. S., Eldredge, N. (2014). What is punctuated equilibrium? What is macroevolution? A Response to Pennell et al. *Trends Ecol. Evo.*, 29:185-186

Linde-Medina, M. (2011). Adaptation or exaptation? The case of the human hand. *J Biosci.*, 36(4):575-85.

Linde-Medina M., Newman, S.A. (2014). Limb, tooth, beak: three modes of development and evolutionary innovation of form. *J Biosci.*, 39(2):211-23.

Löfgren, L. (1977). *Complexity of decription of systems: A foundational study. Int. J. Gen. Sys.* 3: 197-214

Losos, J. B. (2009). *Lizards in Evolutionary Tree: Ecology and Adaptive Radiation of Anolis*, University of California Press

Luque, V. (2016). The Principle of Stasis: Why drift is not a Zero-Cause Law. *Studies in History and Philosophy of Biological and Biomedical Sciences* 57:71-79

Lynch, M. (2007). The frailty of adaptive hypothesis for the origins of organismal complexity. *PNAS*, 104 (1):8597-8604

MacArthur, R. H., Pianka, E. R. (1966). On Optimal Use of a Patchy Environment. *The American Naturalist*, 100,(916):603-609.

Maeda, R. K., Karch, F. (2006). The ABC of the BX-C: the bithorax complex explained. *Development* 133: 1413-1422; doi: 10.1242/dev.02323

Maisnier-Patin S., Roth J,R. (2015). The origin of mutants under selection: how natural selection mimics mutagenesis (adaptive mutation). *Cold Spring Harb. Perspect. Biol.* 7(7), a018176. doi: 10.1101/cshperspect.a018176.

Malthus, T. (2008). *An Essay on the Principle of Population* (Oxford World's Classics). Oxford University Press.

Mark, L. S., & Jun-Yi Leu. (2014). On the nature and evolutionary impact of phenotypic robustness mechanisms. *Annual Review of Ecology, Evolution and Systematics,* 45, 495.

Martin, E., Ruse, M., Holmes, E. (1996). *Dictionary of Biology*. Oxford University Press.

Masel, J., Siegal, M. L. (2009). Robustness: Mechanisms and consequences. *Trends in Genetics*, 25(9), 395-403. doi:10.1016/j.tig.2009.07.005

Masel, J., Trotter, M. V. (2010). Robustness and evolvability. *Trends in Genetics,* 26(9), 406-414. doi:10.1016/j.tig.2010.06.002

Mather, K. (1955). Polymorphism as an outcome of disruptive selection. *Evolution*. 9 (1):51–61. doi:10.2307/2405357

Matthen, M., and Ariew, A. (2002). *Two ways of thinking about fitness and natural selection. Journal of Philosophy*, 99: 55–83.

Maturana, H., & Varela, F. J. (1980). *Autopoiesis and cognition: The realization of the living.* Dordrecht: D. Reidel.

May, R. M. (1976). Simple mathematical models with very complicated dynamics. *Nature*, 261: 459-467, doi:10.1038/261459a0

Maynard Smith, J. (1982). *Evolution and the theory of games*. Cambridge University Press.

———, (1987). Natural selection: when learning guides evolution. *Nature* 329: 761-762.

Maynard Smith, J., Burian, R., Kauffman, S., Alberch, P., Campbell, J., Goodwin, B., Lande, R., Raup, D., Wolpert, L. (1985). Developmental Constraints and Evolution: A Perspective from the Mountain Lake Conference on Development and Evolution. *The Quarterly Review of Biology*, (60) 3: 265-287

Maynard Smith, J., Price, G.R. (1973). The logic of animal conflict. *Nature*. 246 (5427):15–18

Maynard Smith, J., Szathmáry, E. (1999). *The origins of life. From the Birth of Life to the Origins of Language*. Oxford University Press.

Mayr, E. (1942). *Systematics and the origin of species*. Columbia University Press

———, (1960). The emergence of Novelty. In Tax, S. (ed.),*The Evolution of Life*, , Univ. of Chicago Press: 349–380.

———, (1963). *Animal Species and Evolution*. Cambridge, MA: Harvard University Press.

———, (1983). How to carry out the adaptationist program?, *The American Naturalist*, 121: 324–334

McCandlish, D. M., Stoltzfus, A., & Dykhuizen, H. E. D. E. (2014). Modeling Evolution Using the Probability of Fixation: History and Implications. *Quarterly Review of Biology*, 89(3):225–252. https://doi-org.ezproxy.uned.es/10.1086/677571

McClintock, B (1950). The origin and behavior of mutable loci in maize. *Proc Natl Acad Sci USA*, 36 (6):344-355.

McCoy, J. W. (1977). Complexity in organic evolution. *J. Theor. Biol.* 69: 457-458

McFadden, B. J. (2005). Fossil horses – evidence for evolution. *Science*, 307:1728-1730

McHorse B. K., Biewener A. A., Pierce S. E. (2017). Mechanics of evolutionary digit reduction in fossil horses (Equidae). *Proc. R. Soc. B* 284

McKinney, M. L. (1990), Classifying and analyzing evolutionary trends. In McNamara, K. J. (ed.), *Evolutionary Trends*, University of Arizona Press:28-58

McLaughlin, R. N., Malik, H. S. (2017). Genetic conflicts: the usual suspects and beyond. *The Journal of experimental biology*, 220(1):6-17.

McShea, D.W. (1991). Complexity and evolution: What Everybody Knows. *Biol. Philos.* 6: 303-324

———, (1996). Perspective: Metazoan Complexity and Evolution: Is there a Trend?. *Evolution*, Vol. 50, No. 2: 477-492

———, (2001). The hierarchical structure of organisms: a scale and documentation of a trend in the maximum. *Paleobiology*, 27(2): 405-423

———, (2005). The evolution of complexity without natural selection, a possible large-scale trend of the fourth kind. *Paleobiology*, 31(Suppl.):146-156

McShea, D.W., Anderson, C. (2005). The Remodularization of the Organism. In Callebaut, W., Rasskin-Gutman (eds.), *Modularity. Understanding the Development and Evolution of Natural Complex Systems*. The MIT Press:185-205

McShea, D.W., Brandon, R. (2010). *Biology's First Law: the tendency for diversity and complexity to increase in evolutionary systems*. The University of Chicago Press.

McShea, D.W., Hordijk, W. (2013). Complexity by Subtraction. *Evol. Biol.* 40: 504-520

McShea, D.W., Venit, E. P., (2001). What is a Part? In Wagner, G. P. (ed.), *The Character Concept in Evolutionary Biology*. Academic Press:259-283

McShea, D.W., Wang. S. C., Brandon, R. N. (2019). A quantitative formulation of biology's first law. *Evolution, International Journal of Organic Evolution*, 13:1-15.

Mills, S., Beatty, J. (1979), The Propensity Interpretation of Fitness, *Philosophy of Science,* 46: 263-286

Millstein, R. L. (2006). Natural selection as a population-level causal process. *British Journal for the Philosophy of Science*, 57(4):627–653.

Mitchell, M. (2009). Complexity. A Guided Tour. Oxford University Press.

Moczek, A. P. (2008). On the origins of novelty in development and evolution. *BioEssays* 30: 432-477

Monod, J. (1970). *Le Hasard et la Nécessité : Essai sur la philosophie naturelle de la biologie moderne*, Éditions du Seuil.

Morgan, T. H. (1925). *Evolution and Genetics*. Princeton University Press.

———, (1932). *The scientific basis of evolution*. New York: W. W. Norton and Co.

Morris, J. J., Lenski, R. E., Zinser, E. R. (2012). The Black Queen Hypothesis: Evolution of Dependencies through Adaptive Gene Loss. *mBio* 3 (2):e00036-12; DOI: 10.1128/mBio.00036-12

Mossio, M., Bich, L., Moreno, A. (2013). Emergence, closure and inter-level causation in biological systems. *Erkenntnis,* 78, 153-178. doi: 10.1007/s10670-013-9507-7

Mossio, M., Saborido, C., Moreno, A. (2009). An Organizational Account of Biological Functions. *Brit. J. Phil. Sci.* 60: 813-841

Muller, H. J. (1949). The Darwinian and modern conceptions of natural selection. *Proceedings of the American Philosophical Society*, 93:459-470

Müller, G. B. (1990). Developmental mechanisms at the origin of morphological novelty. In *Evolutionary Innovations*, M. H. Nitechi ed., University of Chicago Press, 99-130

———, (2007). Evo–devo: extending the evolutionary synthesis. *Nat Rev Genet*. 8(12):943-9.

———, (2008). Evo-Devo as a discipline. In Minelli A., Fusco G. (eds.), *Evolving Pathways: Key Themes in Evolutionary Developmental Biology*. Cambridge University Press:5-30

Müller, G. B. and Wagner, G. P. (1991). Novelty in evolution: restructuring the concept. *Annu. Rev. Ecol. Syst*. 22, 229-256

Neander, K. (1995). Pruning the Tree of Life. *The British Journal for the Philosophy of Science*, 46(1):59-80

Nesse, R. (1994). *Why We Get Sick*. Random House

Newman, S. A., Forgacs, G., Müller, G.B. (2006). Before programs: the physical origination of multicellular forms. *Int. J. Dev. Biol.* 50: 289-299

Nityananda, V., & Read, J. C. A. (2017). Stereopsis in animals: evolution, function and mechanisms. *The Journal of Experimental Biology*, 220(14):2502–2512. http://doi.org/10.1242/jeb.143883

Nijhout H. F., Reed M. C. (2014). Homeostasis and dynamic stability of the phenotype link robustness and plasticity. *Integr Comp Biol.* 54(2):264-75. doi: 10.1093/icb/icu010.

Ochocinska, D., Taylor, J. R. E. (2003). Bergmann's rule in shrews: geographical variation of body size in Palearctic Sorex species. *Biological Journal of the Linnean Society*, 78(3):365-381, https://doi.org/10.1046/j.1095-8312.2003.00150.x

Okasha, S. (2006). *Evolution and the levels of selection*. Oxford University Press, New York.

———, (2018). *Agents and goals in evolution*. Oxford University Press.

Okasha, S., Paternotte. C. (2014), Adaptation, fitness and the selection-optimality links, *Biol Philos*, 29: 225-232

Orr, H. A. (2005). The genetic theory of adaptation: a brief history. *Nature Reviews Genetics*, 6: 119-127

Orzack, S. H. (2014), A commentary on the 'Formal Darwinism Project': there is no grandeur in this view of life, *Biol Philos*, 29: 259-270

Orzack, S. H., Forber, P. (2010). Adaptationism, in E. N. Zalta (ed.), *Stanford Encyclopedia of Philosophy,* <plato.stanford.edu/archives/fall2010/entries/ adaptationism/>.

Orzack, S. H., Sober, E. (2001a). Introduction, in Orzack, S. H., Sober, E. (eds.). *Adaptationism and Optimality*. Cambridge Studies in Philosophy and Biology. Cambridge: Cambridge University Press:1-23

———, (2001b). Adaptation, Phylogenetic Inertia, and the Method of Controlled Comparisons, in Orzack, S. H., Sober, E. (eds.). *Adaptationism and Optimality*. Cambridge Studies in Philosophy and Biology. Cambridge: Cambridge University Press:45-63

———, (eds.) (2001). *Adaptationism and Optimality*. Cambridge Studies in Philosophy and Biology. Cambridge: Cambridge University Press.

Osborne, M. J., Rubinstein, A. (1994). *A Course in Game Theory*. Cambridge, MA: MIT.

Owen-Smith, Norman (1984). Macdonald, D., ed. *The Encyclopedia of Mammals*. New York: Facts on File:490–495. ISBN 978-0-87196-871-5.

Pagel, M. (2004). Limpets break Dollo's Law. *Trends in Ecology & Evolution*. 19: 278. doi:10.1016/j.tree.2004.03.020.

Paley, W. (2008). *Natural Theology*. Eddy, M. D., Knight, D. (eds.). Oxford University Press.

Parker, G., Maynard Smith, J. (1990). Optimality theory in evolutionary biology, *Nature*, 348: 27-33

Parkhurst, D. F., and O. L. Loucks. (1972). Optimal leaf size in relation to environment. *Journal of Ecology* 60: 505–537.

Payne, Roger S. (1971). Acoustic Location of Prey by Barn Owls (Tyto alba). *J Exp Biol.* 54:535-573

Pence, C., Ramsey, G. (2013). A new Foundation for the Propensity Definition of Fitness. *British Journal for the Philosophy of Science*, 64: 851–881.

Phillips, B. L., G. P. Brown, and R. Shine (2010). Evolutionarily accelerated invasions: the rate of dispersal evolves upwards during the range advance of cane toads. *Journal of Evolutionary Biology,* 23:2595-2601.

Pigliucci M (2007) Do we need an extended evolutionary synthesis? *Evolution*, 61(12): 2743–2749.

———, (2008a). Is evolvability evolvable? *Nature Reviews Genetics*, 9(1), 75-82. doi:10.1038/nrg2278

———, (2008b). Sewell Wright's adaptive landscapes: 1932 vs. 1988, *Biol Philos*, 23: 591-603

———, (2010). Genotype-phenotype mapping and the end of the 'genes as blueprint' metaphor. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences*, 365(1540): 557-66.

Pigliucci, M., Kaplan, J. (2006). Making Sense of Evolution: The Conceptual Foundations of Evolutionary Biology. University Of Chicago Press.

Pigliucci M, Müller GB (2010). Elements of an extended evolutionary synthesis. In: Pigliucci M, Müller GB, (eds.). *Evolution—The Extended Synthesis*:3–17. Cambridge, MA: MIT Press.

Pigliucci, M., Murren, C. J., Schlichting, C. D. (2006). Phenotypic plasticity and evolution by genetic assimilation. *Journal of Experimental Biology*, 209:2362-2367

Plutynski, A. (2008). The rise and fall of the adaptive landscape?. *Biol Philos*, 23:605–623

Pocheville, Arnaud (2015). The Ecological Niche: History and Recent Controversies. In Heams, T.;,Huneman, P., Lecointre, G., et al. (eds.) *Handbook of Evolutionary Thinking in the Sciences*. Dordrecht: Springer:547–586.

Ponizovskiy M.R. (2013). The Central Regulation of all Biophysical and Biochemical Processes as the Mechanism of Maintenance Stability of Internal Energy and Internal Medium both in a Human Organism and in Cells of an Organism. *Mod Chem appl*, 1:1-2, http://dx.doi.org/10.4172/mca.1000e101

Popper, K. (2002). *The Logic of Scientific Discovery*. Routledge

Preer, J. R. (1997). Whatever happened to paramecium genetics?. *Genetics*. 145 (2):217-225

Price, G. R. (1972). Fisher's 'fundamental theorem' made clear. *Annals of Human Genetics*, 36(2), 129.

Price, T. D., Qvarnström, A., Irwin, D. E. (2003). The role of phenotypic plasticity in driving genetic evolution. *Proc. R. Soc. Lond. B* 270:1433-1440

Prigogine, I. (1997). *The End of Certainty. Time, Chaos, and the New Laws of Nature*. The Free Press.

Prigogine, I., Stengers, I. (1984). *Order out of chaos: man´s new dialogue with nature*. Bantam Books.

Prothero, D. R., Shubin, N. (1989). *The Evolution of Perissodactyis* (eds. Prothero, D. R. and Schoch, R. M.), Oxford University Press: 142-175

Provine, W. B. (1986). *Sewall Wright and Evolutionary Biology*. University of Chicago Press

———, (2001). *The Origins of Theoretical Population Genetics*. University of Chicago Press.

Prum, R. O. (1999). Development and evolutionary origin of feathers. *J. Exp. Zool. B Moi. Dev. Evol.* 285: 391-306

Prusiner, S. B. (1998). Prions. *Proceedings of the National Academy of Sciences of the United States of America*, 95(23), 13363–13383.

Pulliam, H. R. (1989). Individual behavior and the procurement of essential resources. In Roughgarden, J., May, R. M., Lev, S. A. (eds.) *Perspectives in ecological theory*:25–38. Princeton, NJ: Princeton University Press.

Queller, D. C. (2017). Fundamental Theorems of Evolution. *The American Naturalist*, 189 (4):345-353

Rasskin-Gutman (2005), Modularity: Jumping Forms within Morphospace, in Callebaut, W., Rasskin-Gutman, eds. (2005). *Modularity. Understanding the Development and Evolution of Natural Complex Systems*. The MIT Press:207-219

Rassking-Gutman, D., Esteve-Altava, B. (2008). The multiple directions of evolutionary change. *BioEssays* 30: 521-525

———, (2009). Modeling Evo-Devo: Broken Hierarchies and Multiple Scales of Organization and Complexity. *AIP Conference Proceedings,* 167:43-56

Razeto-Barry, P., & Frick, R. (2011). Probabilistic causation and the explanatory role of natural selection. *Studies in History and Philosophy of Biological and Biomedical Sciences* 42(3):344-55.

Razeto-Barry, P., Vecchi, D. (2016). Mutational randomness as conditional independence and the experimental vindication of mutational Lamarckism, *Biological Reviews* 92(2):673-683 doi: 10.1111/brv.12249

Reeve, H. K., and Sherman, P. W. (1993). Adaptation and goals of evolutionary research. *Quarterly Review of Biology* 68: 1–32.

Reimann, S. (1996). Homeostasis and stability. Doi: 10.1.1.47.5383

Rice, S. H. (2004) *Evolutionary Theory: Mathematical and Conceptual Foundations*. Sinauer Associates, Sunderland, MA.

Richards, P. M., Liu, M. M., Lowe, N., Davey, J. W., Blaxter, M. L., Davison, A. (2013). RAD-Seq derived markers flank the shell colour and banding loci of the *Cepaea nemoralis* supergene. *Molecular ecology.* 22: 3077-3089.. 10.1111/mec.12262.

Robson, G. C., Richards, O. W. (1936). *The Variations of Animals in Nature*. Longmans Green

Roff, D. (1981). On being the right size. *American Naturalist* 118: 405–422.

Rogers, K. (2005). *On the Metaphysics of Experimental Physics*. Palgrave Macmillan UK. DOI: 10.1057/9780230505100

Rosenberg, A. (1982). On the Propensity Definition of Fitness. *Philosophy of Science*, Vol. 49, No. 2:268-273

Rosenberg, A., Williams, M. (1986). Fitness as Primitive and Propensity. *Philosophy of Science*, Vol. 53, No. 3:412-418

Rouse Ball, W. W. (1960). *A Short Account of the History of Mathematics*. Dover.

Ruse, M. (2013). From Organicism to Mechanism-and Halfway Back?. In Henning, B. G. & Scarfe, A. *Beyond Mechanism: Putting Life Back Into Biology*. Lexington Books: 419ss

Rutherford, S. L. (2000). From genotype to phenotype: buffering mechanisms and the storage of genetic information. *Bioessays*. 22(12):1095-105. DOI: 10.1002/1521-1878(200012)22:12<1095::aid-bies7>3.0.co;2-a

Rutherford, S. L.; Lindquist, S. (1998). Hsp90 as a capacitor for morphological evolution. *Nature*. 396: 336–342

Saborido, C., Moreno, A. J., & Mossio, M. (2010). La dimensión teleológica del concepto de función biológica desde la perspectiva organizacional. *Teorema: Revista Internacional de Filosofía*, 29(3):31-56.

Saunders, P. T., Ho, M. W. (1976). On the Increase of Complexity in Evolution. *Journal of Theoretical Biology* 63:375-384

Santos, M., Szathmáry, E., Fontanari, J. F. (2015). Phenotypic Plasticity, the Baldwin Effect, and the Speeding up of Evolution: the Computational Roots of an Illusion. *Journal of Theoretical Biology* 371:127–136

Sarkar, S. (2005). Maynard Smith, optimization, and evolution. *Bio Philos*, 20:951-966

———, (2008), A note on frequency-dependence and the levels/units of selection, *Biol Philos*, 23: 217-228

———, (2014), Formal Darwinism: Some Questions, *Biol Philos*, 29: 249-257

———, (2015), The Genomic Challenge to Adaptationism, *Brit. J. Phil. Sci. 66: 505-546*

Schlosser, G. (2005), Amphibian Variations: The Role of Modules in Mosaic Evolution, in Callebaut, W., Rasskin-Gutman, eds. (2005). *Modularity. Understanding the Development and Evolution of Natural Complex Systems*. The MIT Press:143-179

Schluter, D. (2000). *The Ecology of Adaptive Radiation*. Oxford: Oxford University Press.

Schoener, T. W. (1971). Theory of feeding strategies. *Annual Review of Ecology and Systematics* 2: 369–404.

Schrödinger, E. (1944). *What is Life? The Physical Aspect of the Living Cell*. Cambridge University Press.

Sheiner, S. M. (1993). Genetics and evolution of phenotypic plasticity. *Annual Reviews in Ecology and Systematics*, 24. 35-68

Sheldon, P. R. (1990). Shaking up evolutionary patterns. *Nature*, 345:772

Shelton, D. E., Michod R.E. (2014), Levels of selection and the formal Darwinism project, *Biol Philos*, 29: 217-224

Shin T., Kraemer D., Pryor J., Liu L., Rugila J., Howe L., Buck S., Murphy K., Lyons L., Westhusin M. (2002). A cat cloned by nuclear transplantation. *Nature*, 415(6874):859.

Shubin N.H., Alberch P. (1986) A Morphogenetic Approach to the Origin and Basic Organization of the Tetrapod Limb. In: Hecht M.K., Wallace B., Prance G.T. (eds) *Evolutionary Biology*. Springer

Simon, H. A. (1962). The architecture of complexity. *Proceedings of the American Philosophical Society*, (106) 6: 467-482

Simpson, G. G. (1944). *Tempo and mode in evolution*. New York: Columbia University Press.

Slatkin, M. (1970). Selection and Polygenic Characters. *Proceedings of the National Academy of Sciences*, 66(1):87-93

Smith, S. A., Shine, R. (1997). Intraspecific Variation in Reproductive Mode within the Scincid Lizard Saiphos equalis, *Australian Journal of Zoology* 45(5):435-445

Smith, P. (2007). *An Introduction to Gödel's Theorems*. Cambridge University Press.

Sober, E. (1984). *The Nature of Selection: Evolutionary Theory in Philosophical Focus*. MIT Press.

———, (1995). Natural selection and distributive explanation: A reply to Neander. *British Journal for the Philosophy of Science*, 46:384–387.

———, (1998a). Six Sayings about Adaptationism. In D. Hull and M. Ruse (eds). *The Philosophy of Biology*. Oxford: Oxford University Press. ISBN 0-19-875213-X.

———, (1998b). Three Differences between Deliberation and Evolution. In: Danielson, P. A. (ed.), *Modelling Rationality, Morality, and Evolution*, Oxford University Press:408-422.

———, (2000). *Philosophy of biology*. New York: Westview Press.

———, (2013). *Trait fitness is not a propensity, but fitness variation is. Studies in History and Philosophy of Biological and Biomedical Sciences,* 44:336–341

Stebbins, G.L. (1966). *Processes of Organic Evolution*. Prentice Hall

———, (1982). *Darwin to DNA, Molecules to Humanity*. W.H. Freeman and Company.

Stegmann, U. E. (2010). What can natural selection explain? *Studies in History and Philosophy of Biological and Biomedical Sciences* 41:61–66

Stephens, D. W., Krebs, J. R. (1986). *Foraging theory*. Princeton, NJ: Princeton University Press.

Sterelny. K. (2008). Macroevolution, minimalism, and the radiation of animals, in Hull, D. and Ruse, M. (eds.), *Cambridge Companion to the Philosophy of Biology*, Cambridge: Cambridge University Press, pp. 182–210.

Stoltzfus, A. (1999). On the Possibility of Constructive Neutral Evolution. *J Mol Evol*, 49:169–181

———, (2006). Mutationism and the dual causation of evolutionary change. *Evol Dev* 8:304–317.

———, (2017). Why we don't want another "Synthesis". *Biology Direct*, 12:23

Stromberg, C.A.E. (2005). Decoupled taxonomic radiation and ecological expansion of open-habitat grasses in the Cenozoic of North America. *Proc. Natl. Acad. Sci. USA* 102:11980-11984.

Sultan, s. E. (2017). Developmental plasticity: re-conceiving the genotype. Interface Focus 7(5). 20170009 DOI: 10.1098/rsfs.2017.0009

Swarthmore. (2010). *Optimal Foraging Theory*. Image retrieved from: http://wikis.swarthmore.edu/mathbio/index.php?title=File:Optimal_Foraging_Theory.jpg&limit=50

Snyman, Jan A. (2005). *Practical Mathematical Optimization. An Introduction to Basic Optimization Theory and Classical and New Gradient-Based Algorithms*. Springer.

Sznajder, B., Sabelis, M. W., Egas, M. (2012). How Adaptive Learning Affects Evolution: Reviewinf Theory on the Baldwin Effect. *Evol Biol*, 39:301-310

Taylor, T. (2018). Evolutionary Innovations and Where to Find Them: Routes to Open-Ended Evolution in Natural and Artificial Systems. Web.

Taylor, T., Bedau, M., Channon, A. (2016). Open-Ended Evolution: Perspectives from the OEE Workshop in York. *Artificial Life,* 22:408-423 doi:10.1162/ARTL_a_00210

Thomson, Ann (ed.) (1996). *La Mettrie: Machine Man and Other Writings*. Cambridge University Press

Thompson, D. (1894). Some Difficulties of Darwinism. *Nature*. 50: 435ff.

———, (1942). *On Growth and Form*, 2nd rev. ed., Cambridge University Press.

Thompson J.N. (1998). Rapid evolution as an ecological process. *Trends Ecol Evol.*;13:329–332.

Tinbergen, N. (1963). *On aims and methods of Ethology*. Zeitschrift für Tierpsychologie. 20(4):410-433. doi:10.1111/j.1439-0310.1963.tb01161.x.

True, J. R., Haag, E. S. (2001). Developmental system drift and flexibility in evolutionary trajectories. *Evolution & Development* 3(2):109-19, DOI10.1046/j.1525-142x.2001.003002109.x

Truchetet, M.-E.., Pradeu, T. (2018). Re-thinking our understanding of immunity: Robustness in the tissue reconstruction system. *Seminars in Immunology*, 36:45-55

Turing, A. M. (1952). The Chemical Basis of Morphogenesis. *Phil. Trans. Roy. Soc. London. Series B,* 237(641):37-72.

Valentine, J. W. (2004). *On the Origin of Phyla*. University of Chicago Press. ISBN 978-0226845487.

van Dongen, J. T., Guptaab, K. J., Ramírez-Aguilara, S. J., Araújob, W. L., Nunes-Nesib, A., Fernieb, A. R. (2011). Regulation of respiration in plants: A role for alternative metabolic pathways. Journal of Plant Physiology (168)12:1434-1443. doi: 10.1016/j.jplph.2010.11.004

Van't Hoff, Campagne P., Rigden D.J., Yung C.J., Lingley J., Quail M.A., Hall N., Darby A.C., Saccheri I.J. (2016). The industrial melanism mutation in British peppered moths is a transposable element. *Nature*, 534:102-105

Varela, F., Maturana, H. (1972). Mechanism and biological explanation. *Philosophy of Science*, 39(3), 378-382. doi:10.1086/288458

Void, C. A. (2005). Directing the Evolvable: Utilizing robustness in Evolution, in Jen, E. (ed.). *Robust Design: A Repertoire of Biological, Ecological, and Engineering Case Studies*. Oxford University Press:105-134

Vogel, J. H. (1988). Evolution as an entropy driven process: An economic model. *System Research and Behavioral Science*, Volume 5, Issue 4: 299–312

Waddington, C. H., (1956). Genetic assimilation of the *Bithorax* phenotype. *Evolution* 10:1-13

———, (1957). *The Strategy of the Genes*. George Allen & Unwin

———, (1967). Discussion. In Moorehead, P. S., Kaplan, M. M. (eds.), *Mathematical Challenges to Neodarwinian Interpretation of Evolution*. Philadelphia Pa: Wistar Institute.

Wagner, A., (2005). *Robustness and Evolvability in Living Systems*, Princeton University Press

———, (2008). Robustness and evolvability: A paradox resolved. *Proceedings of the Royal Society B: Biological Sciences*, 275(1630): 91-100. doi:10.1098/rspb.2007.1137

———, (2011), *The Origins of Evolutionary Innovations*, Oxford University Press

———, (2015). *Arrival of the Fittest*, Oneworld Publications

Wagner, G. P. (1989), Multivariate mutation-selection balance with constrained pleiotropic effects. *Genetics* 122: 223-234

———, (2014), *Homology, Genes, and Evolutionary Innovation*. Princeton University Press

Wagner, G. P., Altenberg, L. (1996). Complex adaptations and the evolution of evolvability, *Evolution* 50, 967-976

Wagner, G. P., Mezey, J., Calabretta, R. (2005), Natural Selection and the Origin of Modules, in Callebaut, W., Rasskin-Gutman, eds. (2005). *Modularity. Understanding the Development and Evolution of Natural Complex Systems*. The MIT Press:33-50

Weiss, K. M., Fullerton S. M. (2000). Phenogenetic Drift and the Evolution of Genotype–Phenotype Relationships. *Theoretical Population Biology*, 57 (3): 187-195

West G. B., Brown J.H., Enquist B. J. (1997). A general model for the origin of allometric scaling laws in biology. *Science*, 276:122-126.

———, (1999). The fourth dimension of life: fractal geometry and allometric scaling of organisms. *Science*, 284:1677-1679.

West-Eberhard, M. J. (1992). Adaptation, Current Usage. In *Keywords In Evolutionary Biology,* Keller, E. and Lloyd, E. A. (eds.), Harvard University Press, Cambridge, Mass.:13-18.

———, (2003). *Developmental Plasticity and Evolution*. Oxford University Press

———, (2005a). Developmental Plasticity and the origin of species differences. *PNAS,* 102:6543-6549

———, (2005b), Phenotypic accommodation: adaptive innovation due to developmental plasticity. J. *Exp. Zool.*, 304B: 610-618. doi:10.1002/jez.b.21071

Whitacre J. M. (2012). Biological robustness: paradigms, mechanisms, and systems principles. *Frontiers in genetics*, 3, 67. doi:10.3389/fgene.2012.00067

Wickstead, B., Gull, K. (2011). The evolution of the cytoskeleton. *The Journal of Cell Biology*, 194 (4)513-525; DOI: 10.1083/jcb.201102065

Wilkins, J. F., Godfrey-Smith, P. (2009). Adaptationism and the adaptive landscape. *Bio Phil*, 24:199.214

Williams, G. C. (1966). *Adaptation and Natural Selection: A Critique of Some Current Evolutionary Thought*. Princeton: Princeton University Press.

Willmore, K.E. (2012). Evo Edu *Outreach,* 5:219-230.

Wolfe, M. F. (2000). *Rube Goldberg: Inventions*. Simon and Shuster.

Wright, B. E. (2000). A biochemical mechanism for nonrandom mutations and evolution. *Journal of Bacteriology* 182:2993–3001.

Wright, B. E., A. Longacre, and J. M. Reimers. (1999). Hypermutation in derepressed operons of Escherichia coli K12. *Proceedings of the National Academy of Sciences of the United States of America* 96: 5089–5094.Wright, S. (1931). Evolution in Mendelian populations. *Genetics* 16:97-159

Wright, S. (1932), The roles of mutation, inbreeding, crossbreeding and selection in evolution. In: *Proceedings of the sixth international congress of genetics*: 356-366

———, (1935). Evolution in populations in approximate equilibrium. *Journal of Genetics* 30: 257–266.

———, (1949). Adaptation and selection. In: Mayr, E.. (ed.), *Genetics, Paleontology, and Evolution*. Princeton University Press: 365-389

———, (1982), The shifting balance theory and macroevolution, *Ann. Rev. Genet.* 16:1-19

Wu, B., Gokhale, C. S., van Veelen, M., Wang, L., & Traulsen, A. (2013). Interpretations arising from Wrightian and Malthusian fitness under strong frequency dependent selection. *Ecology and Evolution*, 3(5):1276–1280.