# From Simple to Complex Multicellularity:

## On the preconditions, design spaces and the evolution of neural agents

Aina Ollé Vila

TESI DOCTORAL UPF / year 2019

THESIS SUPERVISOR
Ricard Solé
Department CEXS

Universitat
Pompeu Fabra
Barcelona

*A la Sílvia, en Jaume, l’Elisabet i en Joan.*

# Agraïments

L'inici del camí que m'ha acabat duent a finalitzar aquesta tesi es remunta a quan era una estudiant de 18 anys, havia escollit Biologia Humana (per atzar, l'altre atractor era – i continua sent –, la física) i vaig tenir la sort de tenir a en Ricard com a professor de matemàtiques. En vaig tenir prou amb la primera classe per saber que allò era el que estava buscant: aprendre que podia estudiar sistemes biològics a través de les matemàtiques va ser un descobriment – ara em costa de creure que encara no n'hagués sentit mai a parlar. Podria estudiar tot el que em fascinava de la biologia amb les eines que més em divertien. Les classes de programació amb en Sergi van ser l'altra revolució. Poder entendre fenòmens biològics a través de la seva simulació va ser el segon descobriment. No havia programat mai i allò era pura diversió i motivació. Aquell va ser l'inici del camí que em va dur a aprendre sobre el camp dels sistemes complexes. Ja no hi va haver volta enrere. Durant el camí que ha seguit des d'aquells dies – que ara es veuen molt llunyans –, se m'ha donat la oportunitat de poder donar les classes de programació durant dos cursos als estudiants de Biologia, on he pogut gaudir d'una de les meves altres passions a part d'aprendre: ensenyar. Gràcies Ricard, Sergi i Javier per aquesta oportunitat impagable.

L'aprenentatge que he adquirit des de l'època en què va començar tot plegat fins ara és difícil de mesurar. Gràcies Ricard per la oportunitat que em vas donar, per tot el que m'has ensenyat, i sobretot per embrancar-nos en totes les preguntes que hem intentat resoldre durant aquests anys, sempre amb la llibertat per aportar idees i propostes. Gràcies també per les oportunitats que m'has donat d'estar al Santa Fe Institute. És difícil de descriure la sensació d'investigar des d'allà. Sens dubte la Summer School i l'estada de recerca amb tu, en Sergi i en Jordi van ser una gran oportunitat per seguir aprenent en el millor entorn possible. Gràcies també per totes les recomanacions d'articles i llibres, una part del teu vast coneixement de la literatura me l'enduc com a regal de tots aquests anys que m'han permès formar-me com a investigadora.

A tots/es els/les companys/es del CSL amb qui he tingut el plaer

de compartir el dia a dia (i bogeries vàries) i, sobretot, d’aprendre de vosaltres. Durant aquests anys es pot dir que he conviscut amb dues generacions-CSL: Max, Salva, Amadís, Luíño, Núria, Sergi, Josep, Raúl, Ben, Dani, Adriano, Carlos, Javier, Eva, Simon, Jordi, Blai, Guim, Víctor, i les recents incorporacions, Arianna, Miguel i Jordi Pla. Al CSL cada dia s’aprenen coses noves, és un caldo de cultiu de nou coneixement. Estar envoltada de gent tant brillant ha estat una oportunitat única durant aquests anys.

Gràcies als meus pares, a l’Elisabet i en Joan. Imprescindibles. A la tieta, en Jordi i la Marina S. Als meus avis i les meves àvies. A la resta de família. A la Cris i la Marina R, la meva família a Barcelona. A la Jès, la Clara, la Flor i l’Elena. A la Paula, la Laura, la Laia, la Txell, la Lídia i la Maria. A la Ilia. Si es pot definir l’amistat, aquesta sou vosaltres. A la Rosa. A la Mireia. A l’Andrea. A en Pep, la Mercè i la Cinta. A en Pau. A en Mike. A en Ra, en Martí, en Suri, en Josep, en Collell, en Genís, en Jeroni i en Pedreta. A tots, gràcies per ser-hi sempre.

# Abstract

The major evolutionary transitions deeply transformed the way information is organized and transmitted, sometimes in the form of a new hierarchical level containing lower, previously individual entities. A specially important case study involves the emergence of multicellularity and in the last decade major advances have been taking place both theoretically and experimentally. Along with standard approaches, new fields and methods have paved the way to a better understanding of the phenomenon and its implications. The rise of organoid technology and the use of *in silico* and experimental models of evolved multicellularity have helped shaping the area beyond natural systems, but little integration of the different approaches has been achieved. Here two instances of such integrative effort are presented in terms of a morphospace analysis of the universe of multicellularity. One based on the developmental-physical-cognitive dimensions and the second grounded in a evodevo approach enriched with the ecological dimension. These spaces provide not only an integrated perspective but also a picture of the limitations imposed by developmental constraints and self-organization, along with a roadmap for synthetic multicellularity. Within the rich repertoire of case studies included in these spaces, two rather unexplored avenues are also presented: (a) the preconditions associated to proto-organism formation and (b) the trade-offs associated to evolved, simple cognitive agents performing information-processing tasks. Both illustrate how theoretical models can provide guidelines to understand the origins of multicellular complexity.

# Resum

Les grans transicions evolutives han transformat profundament la manera com la informació s’organitza i es transmet, a vegades en la forma d’un nou nivell jeràrquic que conté entitats que prèviament es comportaven com unitats individuals. Un cas particular important involucra l’emergència de la multicel·lularitat, sobre la qual en la última dècada s’han fet grans avenços tant des d’una perspectiva teòrica com experimental. Juntament amb aproximacions més estàndard, nous camps i mètodes han aplanat el camí cap a una millor comprensió del fenòmen i les seves implicacions. La posada en escena de la tecnologia dels organoides, juntament amb l’ús de models *in silico* i experimentals d’evolució de la multicel·lularitat han ajudat a ampliar l’àrea d’estudi més enllà dels sistemes naturals. Tot i així, fins avui hi ha hagut pocs avenços en la integració de les diferents aproximacions. En aquesta tesi, dos exemples d’esforç integratiu es presenten en termes d’anàlisi de morfoespais de l’univers de possibilitats que comporta la multicel·lularitat. Un dels morfoespais està basat en dimensions que tenen en compte la biologia del desenvolupament, la física i la cognició, mentre l’altre està basat en una aproximació *evodevo* enriquida amb una dimensió ecològica. Aquests espais no només aporten una perspectiva integrada si no també una avaluació de les limitacions imposades per restriccions lligades al desenvolupament i l’autoorganització, juntament amb un mapa de ruta cap a la multicel·lularitat sintètica. Entre el ric reperotir de casos d’estudi que poden a incloure aquests espais, també es presenten dues rutes bastant poc explorades: (a) les precondicions associades a la formació de *proto-organismes* i (b) els *trade-offs* associats a agents cognitius simples que duen a terme tasques de processament d’informació. Aquestes dues aproximacions il·lustren com els models teòrics poden aportar una sèrie de regles generals per entendre els orígens de la complexitat multicel·lular.

# Contents

# List of Figures

# List of Tables

## List of abbreviations

MET - Major Evolutionary Transitions
MTI - Major (Evolutionary) Transitions in Individuality
DoL - Division of Labour
MLS - Multi Level Selection
MC - Multicellularity

# Chapter 1

# INTRODUCTION

## 1.1 On the origins of biological complexity

> Nothing in biology makes sense
> except in the light of evolution.
>
> Theodosius Dobzhansky

It is generally accepted that an increase in complexity is observed during evolutionary time [Smith and Szathmary, 1995, West et al., 2015] (for a good discussion on the theme, see [McShea, 1991, McShea, 1996]). However, there is no consensus in defining a good measure for biological complexity. Some proposals account for the number of cell types of an organism [Valentine et al., 1994], variety of morphology [Carroll, 2001] or behavior [Szathmáry and Smith, 1995, Bonner, 1988], as well as ecological diversity [Arthur et al., 1993]. More explicit analyses account for the (rough) correlation between the number of cell types and size increase in evolution, linking it to an increase in complexity [Bonner, 1988].

Since complexity typically invoves multiple scales and transitions incorporating noveal features, a relevant question is how the complexity on different scales is related. Here McShea [McShea, 2002] argues that an observed *complexity drain* is at place: higher hierarchies are formed of simpler parts, compared with equivalent 'parts' which live on its own. In a nutshell, achieving higher complexity requires reducing the complexity of the parts. Such situation would be present at the level of cognitive systems, such as ant colonies, where similar trade-offs seem to be at work [Delgado and Solé, 1997].Regarding the information needed to build up complex forms, it is considered that the more complex an organism is, the more information is used to encode it. However, both information and complexity do not seem to have increased gradually during the history of life on earth [Schuster, 1996].

The straightforward question to be answered is the following: why and how has complexity increased? [Szathmáry, 2015]. In particular, an obvious problem is the cost requirements needed to maintain more complex structures. The answer to this question remains elusive, but rapid progress has been taking place in the last decade, particularly in the area of the origins and evolution of multicellularity.

### 1.1.1 Major Evolutionary Transitions (MET)

A particular appealing approach to the problem of complexity increase throughout evolution is the hypothesis that it may have been achieved as a result of a series of *Major Evolutionary Transitions* (METs). This idea was presented in an influential book back in 1995 by John Maynard Smith and Eörs Szathmáry [Smith and Szathmary, 1995], in which they put forward the idea that a series of *major transitions*, namely:

1. Replicating molecules to populations of molecules in compartments
2. Unlinked replicators to chromosomes
3. RNA as gene and enzyme to DNA and protein (genetic code)
4. Prokaryotes to eukaryotes

5. Asexual clones to sexual populations
6. Protists to animals, plants and fungi (transition to multicellularity)
7. Solitary individuals to colonies (non-reproductive castes)
8. Primate societies to human societies (language)

Their thesis was that the increase in complexity in evolution has depended on a small number of major transitions in the way in which genetic information is transmitted between generations. Some of the aforementioned transitions were unique, while others might have occurred several times independently [Smith and Szathmary, 1995]. The division between prokaryotes and eukaryotes is probably the biggest known evolutionary discontinuity [Szathmáry, 2015].

The general accepted definition for the METs is that they are qualitative transitions which involved changes in the way genetic information is stored and transmitted [Smith and Szathmary, 1995, Szathmáry and Smith, 1995]. Interestingly, they identified a series of common features among, *a priori*, very disparate evolutionary shifts:

- Entities that were capable of independent replication before the transition – evolutionar transitions in individuality, see subsection below – can only replicate as parts of a larger unit after it (this is the case for chromosomes, eukaryotes, sexual species, multicellular organisms, social groups) [Smith and Szathmary, 1995]. In order for such a transition to be successful, evolution at the lower level must be somehow constrained by the higher level [Szathmáry, 2015].

- Division of labour (DoL) [Smith and Szathmary, 1995]. The recurrent emergence of the DoL or the combination of functions allows the higher level units to be more efficient under certain conditions, which translates into a fitness advantage. Synergistic fitness interactions are regarded as one of the crucial driving forces behind the major transitions [Szathmáry, 2015].

The origins of deep qualitative changes in complex matter is not new within the context of physics. Profound structural and dynamical changes occur when given external parameters are slowly change- This is the case for example in the transition from water to steam or from a magnetic to a non-magnetic piece of iron as temperature is increased. The relevance of these so called *phase transitions* goes actually far beyond the domain of physics [Solé, 2011]. Importantly, as pointed out in a METs revision made by Szathmáry of the METs 10 years after the publication of the book [Smith and Szathmary, 1995]

> ”it has never been claimed that all transitions would possess all common features or that the possessed features would have uniform weights across all of the transitions”

The main aim of the work was to set an agenda for future research under the determinate idea that insights in any of the transitions could be helpful to understand the others [Szathmáry and Smith, 1995]. Here a specially important feature of phase transition phenomena, namely the presence of universals, could be relevant to substantiate the analogy made here beyond the simple metaphor. In this context, its is well known in the domain of statistical mechanics that very simple rules can fully account for measurable properties of a system undergoing phase transitions. More surprisingly, very different classes of transitions can share exactly the same class of dynamical rules. Such finding strongly reminds us the presence of evolutionary convergence (see Section 1.1.5).

One of the key points in the study of METs is the following: why did natural selection, acting on entities at the lower level, not disrupt integration at the higher level? This is an important question to answer, as there are several examples of conflict arising among the lower level units, such as meiotic drive, the existence of parthenogenesis or the presence of workers laying eggs in an ant colony [Szathmáry and Smith, 1995]. Maynard Smith and Szathmáry proposed that [1.] **contingent irreversibility** (understood as the impossibility of the lower entity levels to reproduce on its own once they have been part of a larger whole for a certain evolutionary time) and [2.] **central control** (suppression of conflict undertaken by the higher unit) could be plausible explanations regarding the maintenance of higher-level entities once they have arisen. Nonetheless, irreversibility is not absolute, then it is not a sufficient explanation for the maintenance of higher level entities [Smith and Szathmary, 1995].

### Forms of heredity

As defined, METs changed the way information was transmitted, in other words, the heredity of this information was changed. A crucial distinction has been made between systems of *limited* heredity, in which only a few distinct states can be transmitted, and *unlimited* heredity, capable of transmitting an indefinitely large number of messages [Smith and Szathmary, 1995]. Evolution transitioned from limited to unlimited heredity in the genetic, epigenetic and linguistic domains [Szathmáry, 2015].

Simple autocatalytic systems, such as the Gánti's *chemoton* [Gánti, 2003, Jablonka and Lamb, 2006, Szathmáry and Smith, 1995] would be an example with limited heredity together with the emergence of a protolanguage in *Homo erectus*. The latter transitioned to unlimited heredity with the emergence of language. On the other hand, the origin of polynucleotide-like molecules, provided too with unlimited heredity.

### The importance of information transmission by non-genetic means

In a follow-up paper of the METs, dedicated to the types of heredity authored by Eva Jablonka and Eörs Szathmáry [Jablonka and Szathmáry, 1995], ideas presented previously by Jablonka [Jablonka, 2002, Jablonka and Lamb, 2005] regarding non-genetic information transmission were put forward linked with the METs proposed by Maynard Smith and Szathmáry [Smith and Szathmary, 1995]. Genetic heredity has itself replaced, and evolved from, other systems.

As pointed out by the authors, evolutionary transitions may have been associated with novel means of information storage and transmission. Catalytic cycles, ribonucleotide replication, translation into proteins, epigenetic inheritance and natural language are all means of information storage and transmission, being all of them critical for METs [Jablonka and Szathmáry, 1995].

In the pre-DNA world, not only transmissible non-DNA information existed but was a prerequisite for the stages that followed later giving rise to a DNA-based genetic system. In the group of transitions giving rise to new levels of organization (evolutionary transitions in individuality), the transmission of non-DNA information is believed to have played a key role, particularly in ensuring the evolutionary stability of the new entity. The iconic example is the role of epigenetics in ensuring the heredity of phenotypes during the differentiation chain: cells with the same DNA content express differentially a set of genes due to the role of epigenetics. Without an efficient transmition of epigenetic information, it would be easy that cells switched to inappropriate states disrupting the structure and function of the whole organism. This is why they argue that the emergence of stable complex MC organisms as well as their development was strongly dependent on epigenetic inheritance.

As Jablonka and Lamb put it [Jablonka and Lamb, 2006],

> We believe that biological information should be seen in terms of the interpretation (or processing) of inputs, rather than as an inherent property of inputs, and is best defined in terms of the receiver system: a source becomes an informational input when an interpreting receiver can react to the form of the source (and variations in this form) in a functional manner. (Jablonka and Lamb, 2006)

Under this more general view, they found straightforward to consider the nervous system as another major transition in evolution, and they give a series of arguments to defend such an idea. Given that METs are defined as changes in the way information is stored, transmitted and processed, then the emergence of the nervous system, a system transmitting a new type of information (neural information), should be considered as one of the most important transitions in evolution. Another argument is that it is precisely the nervous system one of the key distinguishing features of metazoans. This new way of transmitting information has far-reaching evolutionary consequences, as its speed and specificity together with its potential for integration and memory storage were far greater than hormone-based communication systems [Jablonka and Lamb, 2006].

They define the ’neural individual’, which can attain a high level of internal integration and the ability to make rapid adaptive responses, deeply changing the landscape of selection when individuals possess a nervous system. Importantly, the evolution of the nervous system not only emerged as a new way of transmitting information (between cells) but also as a new type of communication between individuals and a new way of transmitting information among subsequent generations. The evolution of these ’neural individuals’ were key for both the transition to social groups and to linguistic communities [Jablonka and Lamb, 2006].

The definition of neural agent raises some questions that are directly connected with the search for the space of the possible discussed in this thesis. In particular, one relevant matter is how cognition itself can be approached and what are the classes of cognitive networks that can be found in biology. Recent studies have suggested that there is a need for a unifying picture of these diverse classes of living (and artificial) systems that depart from standard neural network metaphors [Solé et al., 2019]. Some examples are displayed in Fig. 1.1, which include standard neural-like systems but also microbial assemblies or cell assemblies displaying cohesive spatial dynamics.

In Fig. 1.2, we see the METs enumerated along with extra details regarding their implications for information storage while including the different forms of inheritance systems involved in each of the transitions (genetic, epigenetic and behavioral). The nervous system is also included here as a major transition, while the emergence of sexual individuals has been removed (adapted) due to the general current consensus that it cannot be considered a MET [Szathmáry, 2015]. In this thesis we focus on two specific major evolutionary transitions, the evolutionary transition to multicellularity and the emergence of a nervous system.

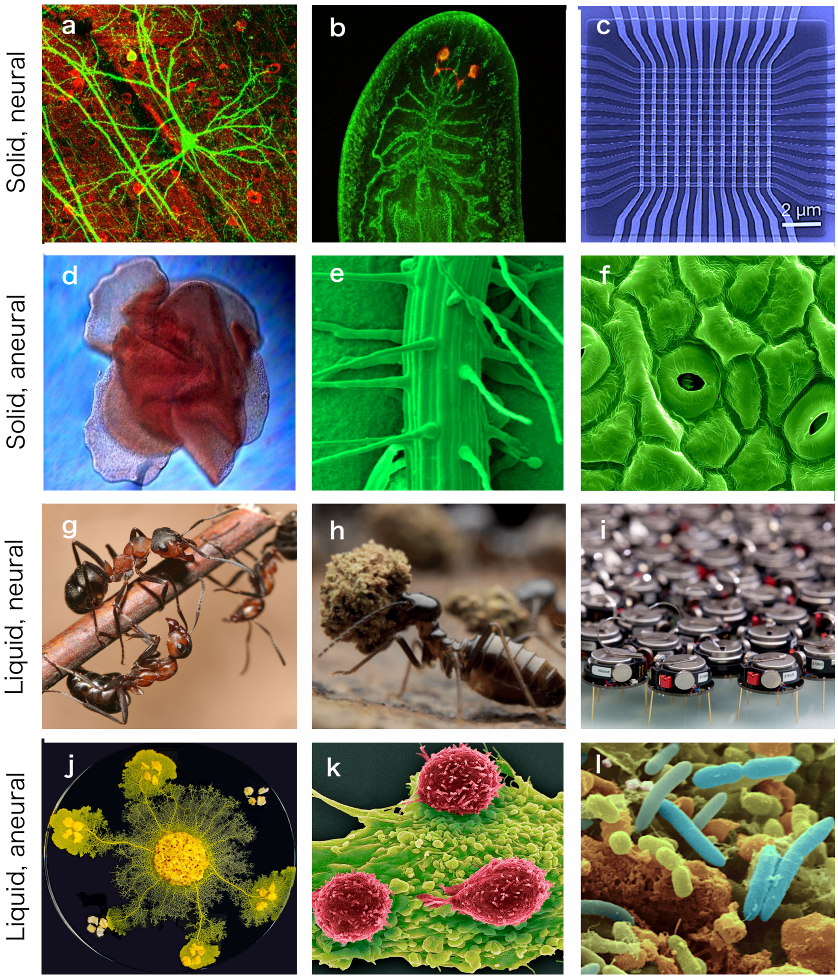
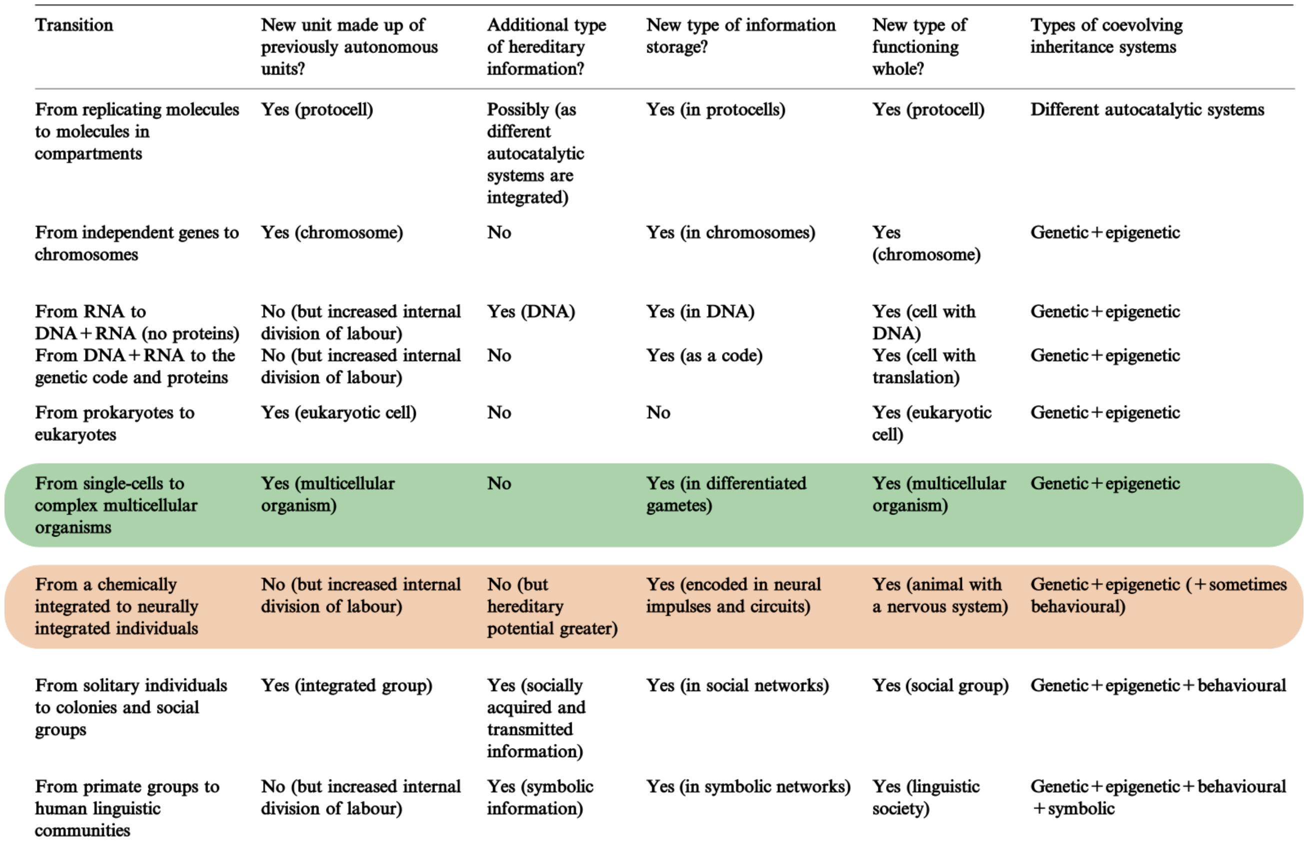


Figure 1.1: What is the space of the possible when dealing with cognitive systems that process information in space and time? Cognition can be found in nature (and its artificial counterparts) in different ways (figure adapted from [Solé et al., 2019]). Along with neural systems (a-c) with neurons located in predictable arrangements that remain stable in time (solid brains). Multiple departures from this can be found, including the lack of true neurons (d-f) to lack of stable cell-cell connections (liquid brains, g-l) that include cellular networks or microbial assemblies.



| Transition | New unit made up of previously autonomous units? | Additional type of hereditary information? | New type of information storage? | New type of functioning whole? | Types of coevolving inheritance systems |
|---|---|---|---|---|---|
| From replicating molecules to molecules in compartments | Yes (protocell) | Possibly (as different autocatalytic systems are integrated) | Yes (in protocells) | Yes (protocell) | Different autocatalytic systems |
| From independent genes to chromosomes | Yes (chromosome) | No | Yes (in chromosomes) | Yes (chromosome) | Genetic + epigenetic |
| From RNA to DNA + RNA (no proteins) | No (but increased internal division of labour) | Yes (DNA) | Yes (in DNA) | Yes (cell with DNA) | Genetic + epigenetic |
| From DNA + RNA to the genetic code and proteins | No (but increased internal division of labour) | No | Yes (as a code) | Yes (cell with translation) | Genetic + epigenetic |
| From prokaryotes to eukaryotes | Yes (eukaryotic cell) | No | No | Yes (eukaryotic cell) | Genetic + epigenetic |
| From single-cells to complex multicellular organisms | Yes (multicellular organism) | No | Yes (in differentiated gametes) | Yes (multicellular organism) | Genetic + epigenetic |
| From a chemically integrated to neurally integrated individuals | No (but increased internal division of labour) | No (but hereditary potential greater) | Yes (encoded in neural impulses and circuits) | Yes (animal with a nervous system) | Genetic + epigenetic (+ sometimes behavioural) |
| From solitary individuals to colonies and social groups | Yes (integrated group) | Yes (socially acquired and transmitted information) | Yes (in social networks) | Yes (social group) | Genetic + epigenetic + behavioural |
| From primate groups to human linguistic communities | No (but increased internal division of labour) | Yes (symbolic information) | Yes (in symbolic networks) | Yes (linguistic society) | Genetic + epigenetic + behavioural + symbolic |

Figure 1.2: The Major Evolutionary Transitions and its implications regarding information (adapted from [Jablonka and Lamb, 2006]). The transitions studied in this thesis: the origins of multicellularity and the nervous system are highlighted in green and orange, respectively. The transition to sexual populations is not shown, in accordance with the revision of the METs done by [Szathmáry, 2015].

### Egalitarian and fraternal transitions

In a very good review of the book from Maynard Smith and Szathmáry [Smith and Szathmary, 1995], David C. Queller [Queller, 1997] pointed out that:

> Insect sociality is just the visible tip of an iceberg of cooperation. To see the bulk of it, we must search below the surface of the organism, and The Major Transitions spends most of its time submerged in this realm. (Cooperators Since Life Began, 1997. David C. Queller)

It was in this same review [Queller, 1997] where David C. Queller put forward the distinction between Egalitarian and Fraternal *Alliances*. In the former, all participants retain reproductive rights [Queller, 2000], as *unlike* units come together, complementing their functions in a higher unit [Szathmáry, 2015]. In the latter, the alliance relies on kinship between its members [Queller, 2000]. In other words, *like* units join or remain joined [Szathmáry, 2015]. The origins of complex MC in general serve as an example for the fraternal transitions, while the origins of the eukaryotec cell for the egalitarian. In Table 1.1, we can check the main features of these two alliances as defined by [Queller, 1997], including also more examples.

### Major Evolutionary Transitions... in Individuality

Leo W. Buss, back in 1987 [Buss, 1987], already presented the ideas that the hierarchies in life history of evolution were to be key for the theory of evolution itself:

> Self-replicating molecules created self-replicating complexes, such complexes created (or became incorporated into) cells, cells obtained organelles, and cellular complexes gave rise to MC individuals. At each transition – at each stage in the history of life in which a new self-replicating unit arose– the rules regarding the operation of natural selection changed utterly. (Leo W. Buss, 1987)

| | Egalitarian | Fraternal |
|---|---|---|
| examples | different molecules in compartments; genes in chromosomes; nucleus and organelles in cells; individuals in sexual unions | same molecules in compartments; same organelles in cells; cells in individuals; individuals in colonies |
| units | unlike, non-fungible | like, fungible |
| reproductive DoL | no | yes |
| control of conflicts | fairness in reproduction; mutual dependence | kinship |
| initial advantage | DoL; combination of functions | economics of scale; later DoL |
| means of increase in complexity | symbiosis | epigenesis |
| greatest hurdle | control of conflicts | initial advantage |

Table 1.1: Two kinds of alliances in major transitions, according to Queller [Queller, 1997].

These ideas for sure fostered the ones on METs by John Maynard Smith and Szathmáry. Buss [Buss, 1987] emphasized the importance of the evolution of individuality, as at each stage in this hierarchical evolution, the units of selection were not only changing but also were affected at both the level of the individual units and the level of the new-formed unit.

The transitions mentioned by Buss [Buss, 1987] included in the proposed METs are also known as Major (Evolutionary) Transitions in Individuality (MTI) [West et al., 2015] or, alternatively, Evolutionary Transitions in Individuality (ETI) [Black et al., 2019], due to the different nature of these transitions with respect to other METs, such as the evolution of the genetic code, sex and language [West et al., 2015]. Therefore, MTI are the ones which, with no exception, consist of lower level entities gathering together and forming a higher level entity with reproductive capacity as a whole, and they can be considered to jointly establish a more concrete research field inside METs [West et al., 2015]

It is in this realm where it is worth it to present the concept of Multi Level Selection (MLS). Okasha [Okasha, 2005] made clear the point that transitions in individuality were deeply linked to the shift from MLS1 to MLS2, that is, from selection acting on lower level units to doing so in higher level ones. For a deeper discussion on the topic (which can lead to confusion), we refer the reader to [Damuth and Heisler, 1988, Okasha, 2001, Okasha, 2005]

In an interesting colloquium paper, Stuart A. West et al. [West et al., 2015], propose a series of questions by which the MTI could be further understood if successfully answered. As they put it,

i. What conditions favor the formation of cooperative groups?

ii. What conditions maintain cooperation during group transformation?

iii. What conditions favor division of labor?

iv. What conditions favor communication that coordinates cooperation at the group level?

v. What conditions lead to negligible conflict within groups?

vi. What conditions favor mutual dependence?

More generally, what they do is to break down MTI into two steps: the formation of a cooperative group (question *i*) and the transformation of that group into an integrated entity (questions *ii-vi*), arguing that all the elements included in the latter group of questions are essential for MTI. However, as they state, major theoretical issues remain unanswered while empirical work has only begun to tackle the issues surrounding the major questions [West et al., 2015]. Interestingly they also provide an extra agenda of research by posing two additional questions:

vii. How are new conflicts of interest suppressed in groups that have already made a major transition?

viii. What conditions favor the breakdown of major transitions?

In the end, what the major transitions approach provides is a unified conceptual framework facilitating comparison across crucial moments in the history of life, suggesting that a similar problem arises at each transition: How are the potentially selfish interests of individuals overcome to form mutually dependent cooperative groups? The potential identification of a few key factors shared among the transitions leading to subsequent increases in biological complexity would potentially unify and simplify our understanding of the evolution of life on earth [West et al., 2015].

### The importance of *preconditions* for METs

Importantly, the concept of *preconditions* was first put forward. If evolution is to engage in a major transitions such as the ones described, a number of (pre)conditions must be fulfilled in order for this transition to be possible. It is the uncovering and identification of such conditions which provides useful insights into the nature of the major transitions. Some examples are [1.] motile cells not being able to divide and mitosing cells not to move are considered preconditions for the division of labor in the *Volvocales*, [2.] if two or more cooperating individuals can achieve something that a similar number of isolated individuals cannot, the preconditions for cooperation exist, as non-additive, or synergistic, fitness interactions are needed. Nevertheless, the dangers of intragenomic conflict remain, [3]. morphological and physiological adaptations of sexual protists could have been preadaptations for simple forms of multicellularity, as alternative phenotypes, specific cell adhesion, cell-to-cell signalling and cell-division arrest play a crucial role in both, [4.] the multiple origins of replication in eukaryotes can be understood as a precondition for the increase in DNA content, necessary for the increase in complexity in multicellular organisms, [5.] meiotic sex can be considered a preadaptation for the subsequent evolutionary radiation of the eukaryotes.

Importantly, METs cannot be explained in terms of the ultimate benefits they entailed [Szathmáry and Smith, 1995]. Therefore, all *preconditions* for METs had a selective advantage on its own, and it is the understanding of the origins of these preconditions and preadaptions of critical importance in order to understand the ultimate origin of METs.

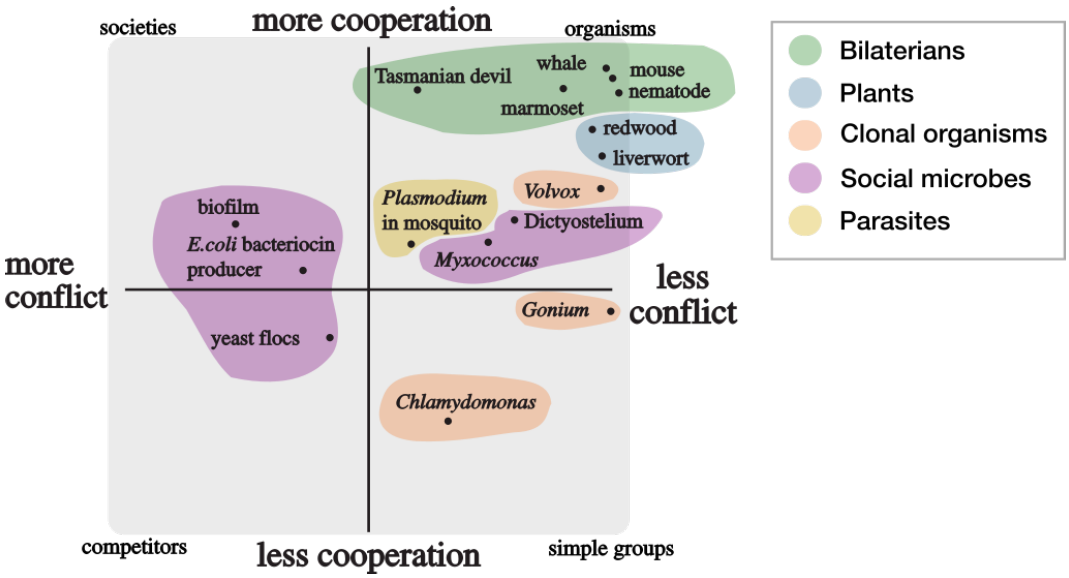


Figure 1.3: To provide a broad view on the concept of *organismality*, coined by Queller and Strassman, they devised a cooperation vs conflict space, in which biological entities of different nature where included. Here we show an adapted version from [Queller and Strassmann, 2009], in which we group biological entities into different categories.

### 1.1.2 What is an organism?

David C. Queller was the one who proposed that the METs might provide another way to think about the concept of organisms [Queller, 1997]. Can each level on the hierarchy of major transitions be considered an organism? What can actually be considered an organism? In a work of the same author together with Joan E. Strassman [Queller and Strassmann, 2009], they coin the term *organismality* to refer to how *organism-like* a biological system can be. The main argument on their definition of organism is that *organismality* is a social process, in the sense that all organims (according to their definition) originated from groups of simpler units that show high cooperation and are nearly free of conflict among them. This feature is what they call ’near-unanimous’ cooperation, and they defend it to be the defining trait of organisms. The qualifying ’near’ is required because some conflicts, like meiotic drive, probably remain in all organisms. Importantly, their definition embraces the concept of adaptation in its core. They suggest that the essence of *organismality* lies in the shared purpose entailed by adaptation: the parts work together for the integrated whole [Queller and Strassmann, 2009].

Notice that their simple definition does not include some features that were to be considered defining for organisms, such as indivisibility, physical-contiguity, high-relatedness, or single-cell origin. And it is precisely the simplicity of their definition what allows for a generality that was lacking until then. In a later review of what can be considered an organism [West and Kiers, 2009], they analyze the work of Queller and Strassman considering that

> The elegance of Queller and Strasmann’s classification is that it generalises from basic evolutionary principles to all levels of biological diversity. (...) The concept of the organism is fundamental to the study of adaptation. By putting adaptation back into the concept of the organism, a firm evolutionary footing is provided to the organism concept, and a slew of interesting questions are raised. (Stuart A. West, 2009)

The paper of Queller and Strassman [Queller and Strassmann, 2009] goes beyond providing a definition of organism: they present a survey of the landscape of actual conflict and cooperation in biological systems. Importantly, as they emphasize, they try to consider only actual conflict rather than potential conflict in pursuing their analyses. This survey is translated in a two dimensional space of cooperation vs. conflict. As they argue, although intuitively high cooperation could be linked to low conflict, the first condition does not imply the other in many examples, as it is shown in their two-dimensional cooperation-conflict spaces [Queller and Strassmann, 2009]. In Fig. 1.3, we show one of their spaces, the one considering groups of cells (the other two gather examples of groups of multicellular individuals and two-species groups. We refer the reader, if interested, to the original source).

As it can be observed, the upper-right quadrant gathers the biological systems that can be considered organisms (high cooperation-low conflict). The opposite one includes what can be considered competitors, with low cooperation and high conflict. The two left quadrants are particularly interesting: the upper-left one indicates an area with high cooperation despite high conflict: the human societies are the canonical examples, together with some mutualist groupings. Finally, the lower-right quadrant contains what can be consider just *simple groups*: they show low conflict but not sufficiently high cooperation to be considered an organism.

In Fig. 1.3 we have adapted the original figure so that related examples are grouped under common categories, easily checking that social microbes, parasite-host relationships or clonal organisms can be considered organisms under the proposed definition. We can observe how all the bilaterians used as examples fall into the organism category, but interestingly they show a high degree of variation when considering conflict: this is due to the conflict originated by mutations. Organisms containing more cells would, a priori, display more conflict (although the Peto's paradox should be taken into account here [Caulin and Maley, 2011]). The Tasmanian devil is placed near the limiting axis because it suffers from a facial cancer that can be transmitted to its progeny, thus entailing certainly higher levels of conflict.

From their work it can be concluded that contiguity is neither necessary nor sufficient for *organismality* (such as in eusocial insects, challenging at the same time the concept of *superorganism* [Hölldobler and Wilson, 2009] commonly used to refer to them), as well as it is not clonality nor even that the entities must be of the same species (being the eukaryote cell the canonical example, but also some mutualisms).

Finally, they rise the point that if evolution seldom crafts new alliances entailing major transitions, it probably crafts new alliances that despite not being truly transformational, they can still be alliances which can be considered to be organismal more often, being the latter ones of critical importance in the understanding of the evolution of *organismality* [Queller and Strassmann, 2009].

### Bacteria as (multicellular) organisms

The first to point out that bacteria could behave as a multicellular organism was James A. Shapiro, back in 1988 [Shapiro, 1988], when amazing multicellular-like behavior performed by species such as *Mycococcus* had already been observed. These ideas did not particularly convince most microbiologists [Shapiro, 1998], but the additional discoveries made in subsequent years make of this idea a nearly undisputed one, mainly due to the identification of several 'quorum sensing' molecules used throughout the eubacterial kingdom to regulate the expression of a wide variety of phenotypes [Shapiro, 1998]. In a review of the same author [Shapiro, 1998] 10 years after the presentation of the first ideas, a set of core concepts of bacterial multicellularity were summarized as:

1. Bacterial cells have communication and decision-making capabilities that enable them to coordinate growth, movement, and biochemical activities.

2. Examples of communication and coordinated behaviors are widespread (possibly ubiquitous) among bacterial taxa and are not limited to a few groups with a specialized multicellular vocation.

3. Bacterial populations derive adaptive benefits from multicellular cooperation and their ability to integrate the diverse activities of different cells. These benefits include (but are not limited to):

(a) More efficient proliferation resulting from a cellular division of labor

(b) Access to resources and niches that cannot be utilized by isolated cells

(c) Collective defense against antagonists that eliminate isolated cells

(d) Optimization of population survival by differentiation into distinct cell types

Examples of multicellularity among the bacteria abound [Shapiro, 1988], such as the spectacular morphologically complex forms of Mycobacteria fruiting bodies (see Fig. 1.4)or the swarm behaviour of *Proteus* colonies. We refer the reader to [Dvorkin and Shapiro, 1997] and [Shapiro, 1998] for a systematic description of examples of the above defined core feature of bacterial behavior.

Bonner himself has suggested that Myxobacteria are a specially interesting candidate to search for the origins of multicellular behaviour. This fascinating group of organisms are well known to display alternative forms of organisation and a developmental process that strongly resembles a MC life cycle. As shown in Fig. 1.4, starvation can shift the dynamics of these colonies (here for *Myxococcus xanthus*) towards a fully organized state where a phenotypically diverse population of bacteria emerges. Resource scarcity triggers a cascade of signals that can include the formation of spiral waves driven by signalling molecules. These waves help organising protrusions (the so called fruiting bodies) that generate spores thus closing the life cycle.

All these features strongly remind us of what to expect from an organism. Are they organisms? All in all, the defining characteristics of bacteria provided by Shapiro do not enter into conflict with the definition provided by Queller and Strassman so that bacteria, acting in groups, can be considered organisms. A relevant question remains to be answered: what is the minimal set of requirements that allow us to talk about organismality? What makes a complex system displaying MC-like properties worth of being labelled as multicellular? And is there anything else that

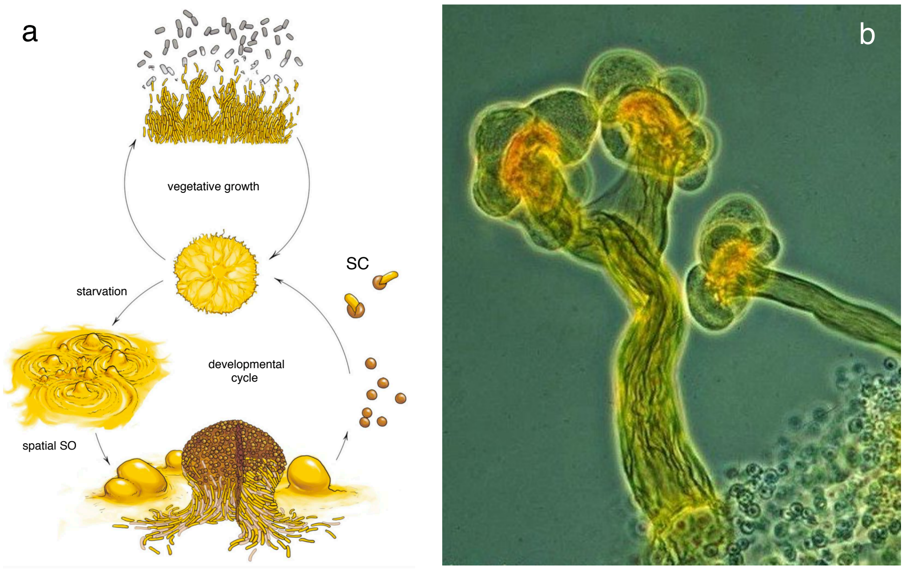


Figure 1.4: What are the simplest living groups representing the most primitive form of multicellular behavior? Myxobacteria are a great candidate and have received considerable attention from multiple perspectives. Here a basic description of the life cycle of *Myxococcus xanthus* (a) is depicted. Among other phenotypic traits, complex structures (b) are generated as part of their life cycle. Although not considered true multicellular organisms, bacteria offer a window to the basic nature of multicellularity. Figures adapted from [Muñoz-Dorado et al., 2016].

pervades the emergence of such minimal conditions? These questions are particularly relevant for the work discussed in this Thesis.

### 1.1.3 The *Synthetic* approach

Eörs Szathmáry and John Maynard Smith discussed in [Szathmáry and Smith, 1995] that a *constructive evolution* approach would be needed so that the understanding of why and how METs occurred was ultimately possible. They understood *constructive evolution* as the experimental recreation of key intermediate stages of evolution which are no longer present in the living biosphere so that a proper comprehension of evolutionary transitions could be attainable.

The term *constructive approach* can be found in a remarkable paper by Phillip W. Anderson [Anderson, 1972], in which he put forward the idea that a reductionist approach did not by any means imply a *constructionist* one. In his words: ”the ability to reduce everything to simple fundamental laws does not imply the ability to start from those laws and reconstruct the universe”. A complete *constructive* approach should take into account the following:

> *The behavior of large and complex aggregates of elementary particles, it turns out, is not to be understood in terms of a simple extrapolation of the properties of a few particles. Instead, at each level of complexity entirely new properties appear, and the understanding of the new behaviors requires research which I think is as fundamental in its nature as any other.* (Anderson. More is different, 1972)

The complexity science has its roots in this idea: the concepts of emergence and self-organization stem from this observation [Kauffman, 1993, Gell-Mann, 1995]. In other words, emergent properties relate to higher-order phenomena that result from the interactions among components belonging to the low-level scale that cannot be reduced to the lower-scale elements’ properties. The METs are a clear example in which these ideas may readily apply. The means to understand the METs, given they occurred a long time ago, resides in experimental and theoretical approaches with the aim to recapitulate the set of (pre)conditions that could drive a system to a major transition. The term *synthetic* approach [Solé, 2016b, Solé, 2016c] is used to describe the exploration of this parallel class of evolutionary transitions (major synthetic transitions, MST) using artificial evolution experiments where alternative paths to innovation can be studied. These artificial means involve synthetic biology approaches, the evolution of robotic agents or neural networks, artificial life systems as well as artificial evolution of natural systems.

In Fig. 1.5, we can see some examples of natural and *synthetic* counterparts regarding different evolutionary transitions. The *synthetic* approaches incorporate artificial means to generate major innovations crossing the boundaries between complexity levels without the need to rely on natural paths [Solé, 2016b].

### Natural vs. *Synthetic* transitions

This approach poses an interesting question: are hard- to-obtain synthetic transitions connected to hard-to-evolve biological novelties? Transitions including multicellularity, symbiosis or different forms of cognitive complexity (which have occurred repeatedly in evolution) have also been achieved in the artificial context, suggesting an hypothetical ease for such transitions to occur. It has been observed that, generally, the basic logic is shared by the natural system and its artificial (or *synthetic*) counterpart. This would suggest the presence of universal rules of organization [Solé, 2016c]. As it was first proposed by Szathmáry [Szathmáry, 2015], the study of METs through the tools given by the physics of phase transitions might be a useful framework to shed light into this hypothesis.

The differences between the natural systems and the synthetic counterpart mainly lies in the fact that the latter are devoid of development - and its strong constraints- (except in the case of natural systems modified using synthetic biology tools). Moreover, the synthetic means of study can also incorporate novel forms of communication or hybrid systems, which implies they can incorporate qualitative features not present in biology. However, these differences can also shed light into the problem of innovation in evolution. The possible differences might also suggest that evolutionary dynamics, in which cost constraints and competition play a critical role, should indispensably be incorporated when using synthetic approaches in order to shed light on the origins of innovation in natural systems [Solé, 2016c].

Among the diverse repertoire of MC systems that are relevant for our approach to this transition, a whole class of synthetic MC structures is provided by the rising area of organoids [Lancaster and Knoblich, 2014].

Briefly, an organoid can be defined as a cohesive group of cell (typically involving different cell types) that can develop from stem cells or organ progenitors through a process of self-organisation and developing into stable structures similar to the ones observed in the *in vivo* development.

#### Preconditions in the *synthetic* domain

As we already pointed out in Section 1.1.1, preconditions are critically relevant because they imply the existence of a landscape of possibilities pervading the emergence of a major qualitative change [Solé, 2016c], in this case, a major evolutionary transition. The usage of *synthetic* approaches can help to shed light into on issue. For instance, computational models can systematically explore the landscape of preconditions that might have predated the chemical space favourable to life, while the usage of robots capable of mirror recognition can be a means to assess if self-recognition can be a precondition for the emergence of consciousness(considered a major transition by some authors) [Solé, 2016c]. On the other hand, cells or, in particular, compartments, might have been a precondition for the emergence of complex life, while the precondition for the origin of eusocial insects might be a ’get together’ rule that should operate after individuals are born [Solé, 2016c]. The so-called *synthetic* approaches can be extremely useful at assessing the likelihood that these hypothesized preconditions might have been the ones leading to the observed major shifts in complexity throughout evolutionary time.

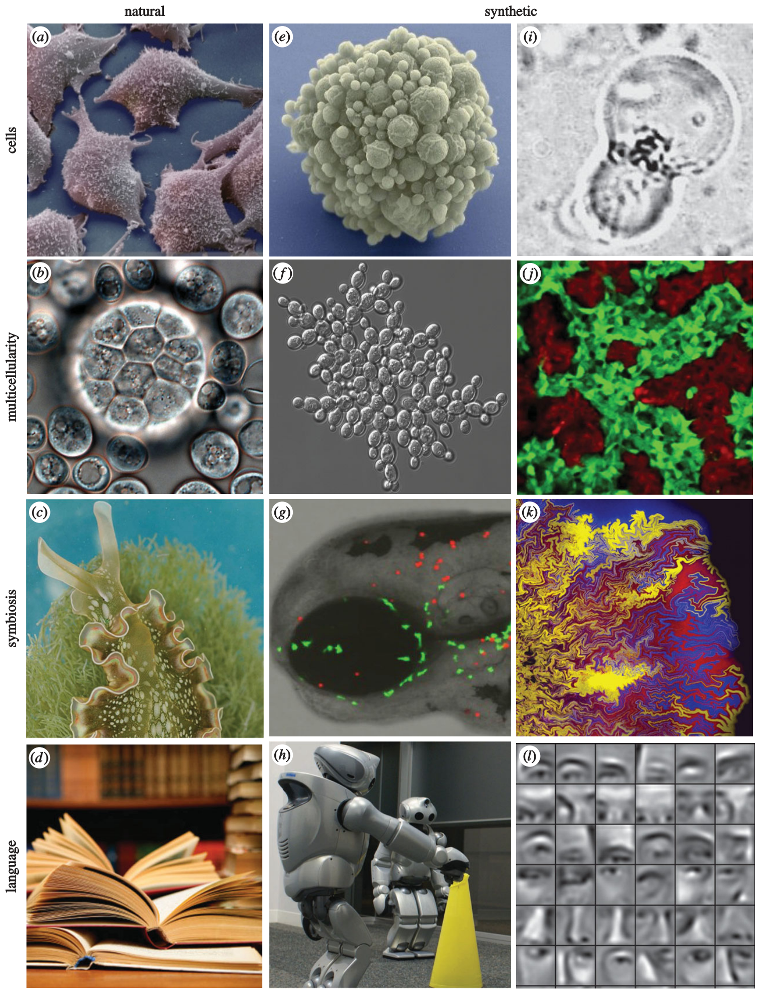


Figure 1.5: METs can be studied through *synthetic* approaches. Natural versus synthetic counterparts are shown. Examples of the transition to cells, multicellularity, symbiosis and language are shown (from top to bottom). The synthetic counterparts to the natural examples shown are: e) synthetic cells using a genome reduction strategy, i) bottom-up protocell approach, f)artificially evolved multicellular system, j) designed multicellular system using synthetic biology, g) and k) engineered cooperation, h) evolved communicating robots and l) artificial pattern recognition. Adapted with permission from reference [Solé, 2016b].

### 1.1.4 Morphospaces

#### Historical acknowledgment

The morphospace concept has its roots in the works of David M. Raup regarding the geometry of coiling in gastropods back in the 60s [Raup, 1961, Raup, 1962, Raup and Michelson, 1965, Raup, 1966]. He was not the first getting his attention attracted by the geometrical form of coiled invertebrate shells, which had already caught the interest of zoologists and mathematicians beforehand. D'Arcy Thompson, in his seminal book 'On Growth and Form' back in 1952 [Thompson, 1952], already dedicated his attention to the problem thoroughly revising the insofar available literature and proposing the idea that the growth of coiled forms could be following a set of (rigid) mathematical laws.

Raup [Raup, 1961] devised a set of measures that he argued were improving the ones proposed by Thompson [Thompson, 1952], defining the basic coiling form of the gastropods through only four parameters. The main goal of his study was to devise a description system general enough to be readily applied to a wide range of gastropod types. In a follow-up work [Raup, 1962] he showed how with the aid of a computer he could obtain graphical reconstructions of the shell forms determined by his mathematical description. This opened a quite important door in the study of biological morphology: the representation of both existent and non-existent biological forms. Raup and Michelson coined the term 'Theoretical morphology' [Raup and Michelson, 1965] to define the study of forms through a conceptual or mathematical model established for some aspect of morphology. This allows to study the set of forms represented by actual species as well as the ones that do not occur in nature, thus having potential importance for evolutionary studies. As Raup and Michelson [Raup and Michelson, 1965] pointed out:

> When the geometries of naturally occurring species are plotted in this space, it becomes evident that it is not evenly filled. Evolution has favored some regions while leaving others essentially empty. In the empty regions we are presumably dealing with forms which are geometrically possible but

> biologically impossible or functionally inefficient. The correct explanation of such empty regions may provide keys to the ultimate interpretation of the morphology of actually occurring shell forms. It is often easier to explain the absence of forms than their presence. (David M. Raup and Arnold Michelson, 1965)

In an actual three-dimensional space (fixing one of the four initially proposed parameters) Raup showed how functional and evolutionary groups were confined to discrete regions of the whole possible spectrum of forms [Raup, 1966]. Indeed, what was clear was that the distribution of actual species was not random at all. This observation leads to quite ambitious questions, as Raup noted:

> Do the relatively unused regions represent physiologically impossible shell forms or has the evolution of these taxa simply not had sufficient time in which to populate the entire block? Are some regions of the block suited only to one set of shell functions? Are swimming forms, for example, necessarily limited to certain discrete regions? (David M. Raup, 1966)

Voids in theoretical morphology spaces pose a challenge to researchers aiming to understand the principles of evolution and its outcomes, observable as extant or extinct (if present in the fossil record) species. A sample theoretical morphospace is shown in Fig. 1.6, in which foraminiferal shells are studied [Tyszka, 2006].

### A deeper conceptualization

George McGhee, a PhD student of David M. Raup, made a valuable synthesis effort for the field [McGhee, 1999]. As he puts it,

> Theoretical morphology is concerned with the simulation of the principal aspects of form with a minimum number of geometric parameters, or with the simulation of the morphogenetic process itself that produced the form under study, and

> its not concerned with the production of a precise mathematical characterization or picture of any given existent form. In fact, the creation and examination of nonexistent form is often of more interest in theoretical morphologic analyses than the examination of existent form.

In his book, he widely reviewed important works using this concept (being some remarkable examples: [McGhee, 1980, McKinney and Raup, 1982, Niklas and Kerchner, 1984, Savazzi, 1987, Ackerly, 1992, Niklas, 1997, McCartney and Loper, 1989, Ellers, 1993, Swan and Kershaw, 1994]), while doing a valuable conceptualization work. He makes clear the distinction between theoretical and functional morphology (the former does not imply the latter, which, in a nutshell, tries to unveil the adaptive significance of the distribution of organic forms in the morphospace), between theoretical and empirical morphospaces (the latter are not aimed to capture nonexistent forms), and between adaptive landscapes [Wright, 1932] and theoretical morphospaces.

The last distinction deserves further attention. Sewall Wright, back in 1932, presented the concept of fitness landscapes in his seminal work *The roles of mutation, inbreeding, crossbreeding, and selection in evolution* [Wright, 1932], in which he envisaged evolution in an $(n-1)$-dimensional space built with all the possible genetic variants organisms could attain, being the $n$th dimension the fitness of every genetic variant combination. Therefore, organisms would be *moving* in this space directed towards the fitness peaks by the processes of mutation and selection. A theoretical morphospace is by no means a fitness landscape, but, if the fitness-related information is available, it can be turned into one, despite maintaining its form-related dimensions instead of the genetic ones proposed by Wright. Of particular interest are the works pursued by Karl J. Niklas in the study of plant morphology, elegantly connecting the theoretical morphology field and the adaptive landscape concept [Niklas, 1994, Niklas, 1999, Niklas, 2004].

McGhee [McGhee, 1999] interstingly highlights the concept introduced by Hickmann [Hickman, 1993], who presented the *design space* concept, which would subsume theoretical morphospaces. The design space concept simply adds the possibility that the space dimensions do not need to be related to morphology, but can refer to ecological, behavioral or physiological properties of the system under study. An interesting work making use of the concept presents a theoretical *ecospace* with the aim to quantify the ecological diversity of Paleozoic and modern marine biotas [Novack-Gottshall, 2007], while another one aiming to study the skeleton space [Thomas and Reif, 1993] combines both morphological and non-morphological dimensions.

McGhee [McGhee, 1999] pointed out the usefulness of adding a time dimension to the theoretical morphology problem: theoretical morphospaces could be of great value in the study of both extinct and extant species, as Stephen Jay Gould was calling for in 1991 [Gould, 1991]. It was also McGhee in a second book [McGhee, 2006], who further formalized the field in distinguishing particular constraints that could be the roots to understand the distribution of forms in theoretical morphospaces, being of particular interest in the attempt to disentangle the reasons for the presence of empty regions in them. He distinguishes between extrinsic constraints –imposed by the laws of physics and geometry– and intrinsic ones –imposed by the biology of the organism–. The former group includes both geometric (possible vs. impossible forms) and functional (non-functional forms prevent the organism from survival) constraints. In the latter, both phylogenetic and developmental constraints are considered. Developmental constraints are considered to play a critical role in evolution, and this theme will be further developed in the next section. However, these extrinsic constraints are not only limited to the ones McGhee proposes. As Gunther J. Eble [Eble, 1999] notes:

> To the extent that morphological evolution expresses the interplay of underlying intrinsic (developmental) and extrinsic (ecological) factors, it is of interest to evaluate the relative importance of such factors in explaining patterns of morphospace occupation. (Gunther J. Eble, 1999)

The aim of McGhee in distinguishing such constraints was to make evident that morphospaces can actually be of great importance as a way to uncover such constraints. As he puts it,

> The thesis of this chapter, and indeed of the entire book, is that the power of the techniques of theoretical morphospace analysis remains largely unrecognized by the biological community. Most evolutionary biologists still think of the adaptive landscape and the epigenetic landscape as heuristic concepts, interesting ways of thinking about life but unusable in the actual analysis of life. The challenge for present and future theoretical morphologists is to dispel this misconception, and to demonstrate to the wider biological community that adaptive landscapes and epigenetic landscapes can be used as actual analytical tools through the creation of theoretical morphospaces. (McGhee, 2006)

Importantly, the dimensions chosen when defining a morphospace should be ideally as orthogonal as possible, in order to avoid interdependencies.

### Ways to explore a morphospace

The initial scope of the theoretical morphospaces under the theoretical morphology research field has been expanded in many ways. Particularly, the objects of study under the idea of a theoretical morphospace can be widely varied, from networks [Avena-Koenigsberger et al., 2015] to possible computations [Solé and Macia, 2013]. Networks can represent a wide variety of systems, from brain connectivity patterns [Avena-Koenigsberger et al., 2014] to language networks [Seoane and Solé, 2018]. The so-called network morphospace field [Avena-Koenigsberger et al., 2015] has been particularly prolific [Solé and Valverde, 2004, Vértes et al., 2012, Corominas-Murtra et al., 2013, Goñi et al., 2013, Avena-Koenigsberger et al., 2014, Esteve-Altava and Rasskin-Gutman, 2014, Seoane and Solé, 2018]. As Avena-Koenigsberger et al. [Avena-Koenigsberger et al., 2015] point out, the common architectural features that have been shown to be shared by many real-world network systems call for compelling explanations. Which is the origin of this phenomena? Do such network attributes reflect common design principles or constraints imposed by selection? The morphospace concept is a particularly adequate approach in the aim to tackle this problem, as it allows to place such networks in a common space which can reveal their relations, helping to unveil which are the evolutionary drivers behind them. In the case of network, either structural traits or, network growth-related parameters can be the dimensions of the network morphospace. The differential filling of the network morphospace can also reveal which are the constraints behind the networked systems.

Particularly relevant is the fact that most of the aforementioned works have used optimization approaches in order to find relevant answers to the questions posed by their particular problems. To be more concrete, optimization can be of great value to explore the possibility that the positioning of empirical networks in particular regions of the morphospace is stemming from the presence of optimization processes [Solé and Valverde, 2004]. Or, on the other hand, to discard this possiblity [Corominas-Murtra et al., 2013] (see Fig. 1.8b). Multi-objective optimization (Pareto-optimality) approaches have proved to be critical for the understanding of the role of conflicting objectives in this realm [Goñi et al., 2013, Avena-Koenigsberger et al., 2014]. This approach also allows to obtain more information regarding the accessibility of certain solutions, as well as allowing the study of trajectories in the morphospace leading to optimality. Finally and importanly, the usage of optimization techniques can be used to test hypothesis about mechanisms underlying network evolution, such as selection pressures or functional and structural constraints [Avena-Koenigsberger et al., 2015].

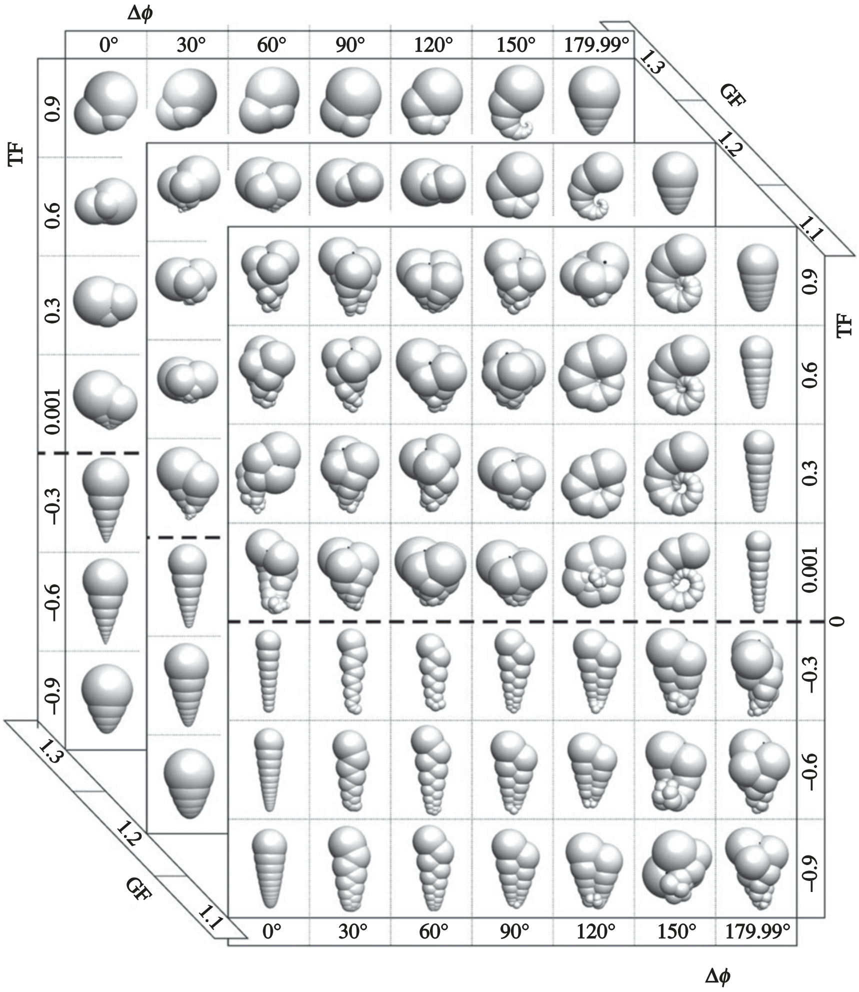


Figure 1.6: Theoretical morphospaces allow organising observed and possible morphologies in a metric space. Here, we show a three-dimensional theoretical foraminiferal morphospace. The potential repertoire of Foraminifera shells is generated by a three-parameter model of form, whose parameters are: Dw, deviation angle; translation factor, TF; growth factor, GF. Adapted with permission from reference [Avena-Koenigsberger et al., 2015].

### Pareto optimality and the trait space, a different approach

A remarkable work in the field of theoretical morphology is the one from Shoval et al. back in 2012 [Shoval et al., 2012], where they elegantly link the trait space (morphospace) with pareto optimality. In a nutshell, pareto optimality is the so-called multiobjective optimization approach in which more than one target is aimed to be optimized at the same time, without neither weighting nor collapsing the distinct targets into a single objective. When the targets get into conflict (the optimization of one occurs in detriment of the second), a set of pareto-optimal solutions is obtained.

The insights provided by Shoval et al. [Shoval et al., 2012] are based under these assumptions: (i) The traits under study underlie a functional trade-off, (ii) the performance at each task decays with distance in trait space from the best phenotype at that task (archetype) and (iii) performances have a single global maximum (just one archetype exists for each task). If this is the case, they find a quite intriguing and compelling result: the best-fitness solutions (termed the Pareto front) lie on simple low-dimensional shapes in trait space: line segments, triangles and other polygons. The vertices of these shapes are always occupied by the archetypes. Under this approach they are able to explain the filling of a various set of morphospaces, from Darwin's ground finches (using beak-related traits) to the leaf-cutter ants (using poison-sac length and head width as traits) and a species of bats (using wing-related traits and body mass) [Shoval et al., 2012]. With these results, they also aim to answer the long-standing observation that most of the morphospace is empty [Raup, 1966, McGhee, 1999, McGhee, 2006]. In a follow-up work, where they relaxed the aforementioned assumptions [Sheftel et al., 2013], they find out quite similar results.

As a result of these ideas and after checking their robustness to more varied situations [Sheftel et al., 2013], they published a quite relevant paper that opens a new door in the understanding of morphospaces and the aim to unveil the evolutionary drivers behind them [Hart et al., 2015]. Concretely, Hart et al. [Hart et al., 2015] developed a software that is capable of inferring biological tasks from high-dimensional biological data.

These high-dimensional data is described as polytopes, allowing to disentangle the tasks performed by the underlying phenotypes through the exploration of the features which are maximally enriched closest to the vertices (or archetypes). They prove the value of their method through relevant inferences using gene expression spaces, through which they are able to infer different biological functions being optimized.

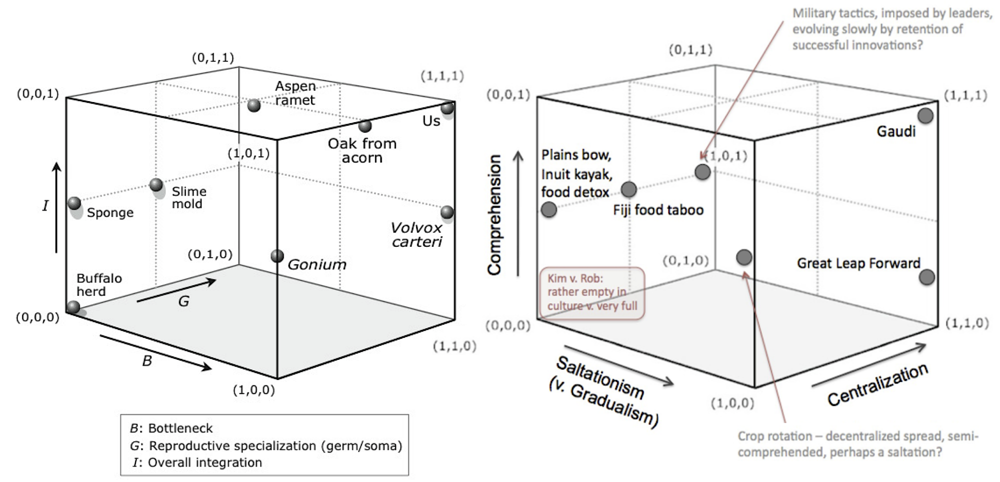


Figure 1.7: Qualitative morphospaces for presence or absence of evolutionary-relevant features, as proposed by Godfrey-Smith [Godfrey-Smith, 2009]. These include a space for reproduction-related dimensions (left) which organise in the highly diverse set of examples from herds (0,0,0) to human societies (1,1,1). Similarly, a qualitative space can be described that allows locating cultural transitions (left).

### Qualitative morphospaces

In section 1.1.2, we have already observed what can be considered a qualitative morphospace: the space of cooperation-competition which located different systems to evaluate its *organismality* [Queller and Strassmann, 2009], as we can observe in Fig. 1.3. Despite the axis are not quantified, the relative position between the systems included in that space and the regions where they fall are already very informative. When the scope of the study is a conceptual one, qualitative morphospaces can be very useful. Godfrey Smith's work on Darwinian dynamics [Godfrey-Smith, 2009] gives a powerful illustration (see Fig. 1.7). Despite we use the morphospace word to define these works, they should fall into the more general category of design space, pointed out by Hickman [Hickman, 1993].

Another example is the landscape of computations [Solé and Macia, 2013], in which a three-dimensional space includes spatial segretation, agent diversity and the degree of parallelism (see Fig. 1.8a). Qualitative and conceputal morphospaces like this and the previous one allow for the inclusion of systems involving several scales of complexity. Solé [Sole, 2017] also used a qualitative morphospace to present a possible cognitive space which included the possible interactions between a human and a robot, in the conceptualisation of a *humanbot* (see Fig. 1.8d). The dimensions considered of relevance in this work included both artificial and human cognitive complexity and the emotional engagement between the two. A work stemming from the latter proposed a morphospace of consciousness[Arsiwalla et al., 2017], including as dimensions autonomous, computational and social complexity (see Fig. 1.8c).

This approach has been also followed in the field of information technology as discussed in [Valverde, 2016] where the author reviewed the major transitions in information technology, revealing the the potential of cultural evolution when understanding the emergence of these transitions. He presented an information technology evolution space including the dimensions of hardware (performance), software (openness) and the rate of diffusion (popularity) in which interestingly the different major transitions are located at different thresholds of these axes, separating phases or groups of distinct designs (see Fig. 1.8e). This would be an example mixing both technological and social dimensions.

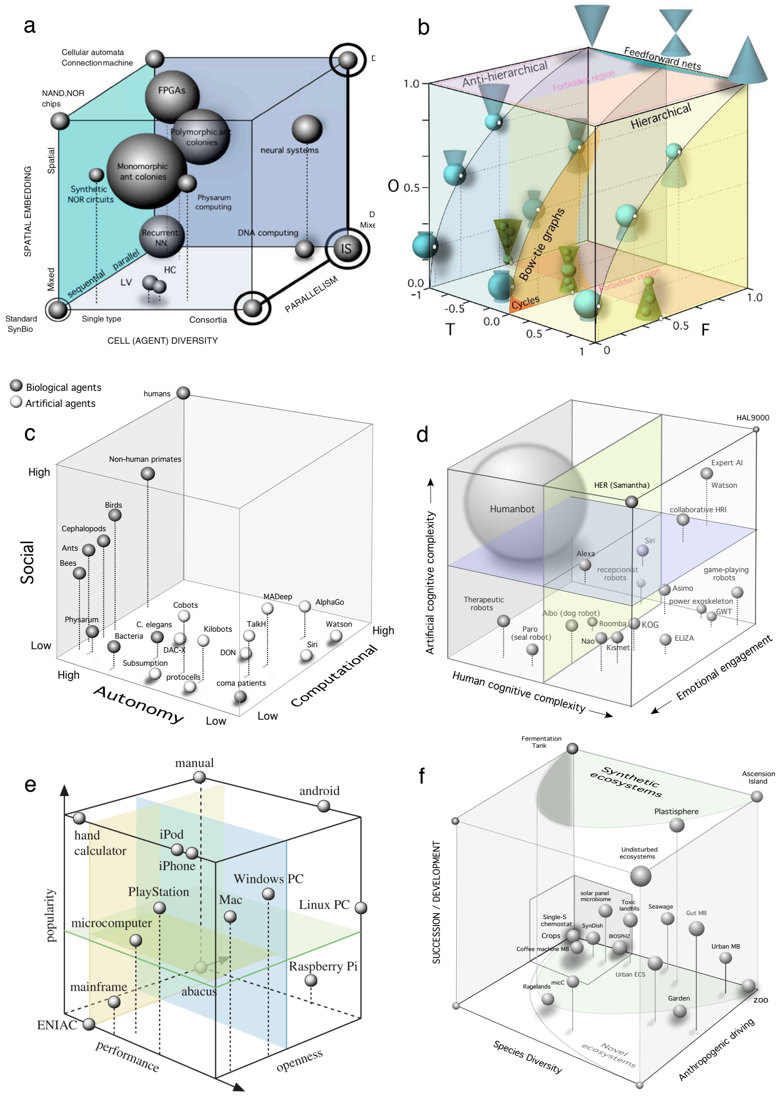


Figure 1.8: Qualitative morphospaces. One way of achieve a global, interactive picture of the space of the possible associated to a class of complex systems is to place them in relative locations in a 3D cube. Here we display examples of these spaces as introduced in (a) computational biosystems, (b) hierarchical networks, (c) cognitive systems, (d) robot-human interactions, (e) technological innovations or (f) synthetic ecosystems.

By using (and expanding) the concept of qualitative morphospace our goal in this dissertation is to move beyond limit cases and trace a consistent and rich space of possible multicellular objects. Along with the definition of such a space and the relative location of known natural, synthetic or *in silico* cases, a properly defined morphospace can help to connect very diverse systems (thus beyond the domain of development) and find the boundaries of the unknown: those empty spaces within the qualitative space.

As discussed in this section, unoccupied domains immediately indicate that either impossible paths exist connecting the observed with the absent or that basic design principles cannot be discovered by evolution. However, that is not necessarily true for engineered systems, which could effectively cope with those limitations. Identifying the empty domains can actually help suggest new conceptual frameworks.

### 1.1.5 Universals: the possible and the actual

> Two possibilities exist:
> either we are alone in the Universe
> or we are not.
> Both are equally terrifying.
>
> Arthur C. Clarke

The hyperdimensionality of the phenotype space still puzzles everyone who takes a few minutes to think about it. Even taking a few number of genes, the number of different proteins they can encode grows to astronomic numbers rapidly [Smith, 1970, Arnold, 2011, Smith and Morowitz, 1982]. How is it possible that evolution has worked out the way to give rise to such an amazing biodiversity of species? We have so far introduced the concepts of major evolutionary transitions and the importance of preconditions, questioned the definition of organism which clearly shows that the limits when talking about biological systems are often more blurred than we expected, explained some of the tools we have to interrogate nature and introduced the interesting concept of the morphospace, which deals with the possible and the actual in nature. The latter concept is the focus of this section.

The a priori possible is astronomic. However, as it has been noted, from the start, not everything is possible. At all. If we keep in mind the powerful analogy proposed by Sewall Wright regarding fitness landscapes [Wright, 1932], we must acknowledge that not all the *walks* [Kauffman and Levin, 1987] on that landscape are possible to undertake. Some empirical examples prove this concept, when studying the possible paths leading from a wild type species to a mutant one [Weinreich et al., 2006, Cervera et al., 2016], showing also the strong effects of epistasis [Kauffman and Weinberger, 1989]. Some of the mutations can lead to the production of non-functional proteins, or even lethal forms of it, while others produce a huge decrease in fitness (aka reproductive success). This already poses a huge constrain on the *walkable* genotype space and the possible. Nevertheless, the numbers are still astronomic.

Developmental processes still puzzle scientists for their *perfect* unfolding in space and time. Morphogenesis is not a linear process, but rather one involving a several number of feedbacks. As morphogenetic processes unfolds, the effects of gene expression on cell properties influence tissue geometry, which feeds back to gene expression patterns [Alberch, 1989], preventing a linear genotype-phenotype mapping. Developmental constraints are defined as a bias in the expression of phenotypic variation due to properties of the developmental process [Alberch, 1982, Oster and Alberch, 1982, Maynard Smith et al., 1985], and they have a strong influence on the space of the possible.

François Jacob, in his seminal paper on 'Evolution and Tinkering' back in 1977 [Jacob, 1977], noted that evolution can only work as a tinkerer,

> Clearly an understanding of the simple is necessary to understand the more complex, but whether it is sufficient is questionable. (...) Can one explain why, among all the possible interactions at one level, only certain are actually observed at the more complex one? (...) Natural selection has no analogy with any aspect of human behavior. However, if one wanted to play with a comparison, one would have to say that natural selection does not work as an engineer works. It works like a tinkerer– a tinkerer who does not know exactly what he is going to produce but uses whatever he finds around him whether it be pieces of string, fragments of wood, or old cardboards; in short it works like a tinkerer who uses everything at his disposal to produce some kind of workable object. (...) Evolution does not produce novelties from scratch. It works on what already exists, either transforming a system to give it new functions or combining several systems to produce a more elaborate one. (François Jacob, 1977)

This again poses another layer of constraints on what is possible during the course of evolution. However, the historical contingencies do play a role. As the same Jacob defended [Jacob, 1977]:

> the interplay of local opportunities -physical, ecological, and constitutional -produces a net historical opportunity which in turn determines how genetic opportunities will be exploited. It is this net historical opportunity that mainly controls the direction and pace of adaptive evolution. This is why the probability is practically zero that living systems, which might well exist elsewhere in the cosmos, would have evolved into something looking like human beings. Even if life in outer space uses the same material as on the earth, even if the environment is not too different from ours, even if the nature of life and of its chemistry strongly limits the way to fulfill certain functions, the sequence of historical opportunities there could not be the same as here. A different play had to be per-

> formed by different actors. Despite science fiction, Martians cannot look like us. And we might as well have looked like one of those 16th-century monsters.

Stephen Jay Gould, a recognized paleontologist who together with Niles Eldredge conceptualized the process of punctuated equilibrium [Gould and Eldredge, 1977], proposed his famous philosophical thought experiment about replaying the tape of life [Gould, 1990]:

> I call this experiment ’replaying life’s tape’. You press the rewind button and, making sure you thoroughly erase everything that actually happened, go back to any time and place in the past–say, to the seas of the Burgess Shale. Then let the tape run again and see if the repetition looks at all like the original. If each replay strongly resembles life’s actual pathway, then we must conclude that what really happened pretty much had to occur. But suppose that the experimental versions all yield sensible results strikingly different from the actual history of life? What could we then say about the predictability of self-conscious intelligence? or of mammals? or of vertebrates? or of life on land? or simply of multicellular persistence for 600 million difficult years?

It is essential to talk about the Burgess Shale and the Cambrian period at this point. The Burgess Shale is the fossil-bearing deposit found in 1909 in British Columbia, well-known because of the highly accurate preservation of soft-bodies animals, one of the richest ever found. Importantly, it contained several body plans not found in actual species. It took several years to understand what was observed, as initially Charles Doolittle Walcott tried to fit all the fossils found to modern groups [Gould, 1990]. It was several years latter, 1971, that Harry Whittington published a reexamination of the fossils, in which it was clear that the diversity found was larger than the initially appreciated, also showing that some of the species found could not be fitted into known groups. His work and that of his students, Simon Conway Morris and Derek Briggs, was crucial

Figure 1.9: Contingency versus convergence and structuralism. Three major contributors to these ideas are displayed here. From left to right: Stephen Jay Gould, Conway Morris and Stuart Kauffman. The first decidedly defended the idea that historical events play a major role in shaping evolution, while the second provided solid evidence for the existence of a dominance of convergent dynamics that would create similar structures under completely independent conditions. The latter proposed the idea of ”order for free” i. e. that many fundamental traits found in biological systems are the result of universal laws beyond the specific features of biological entities.

to understand the importance of that fossil-bearing deposit. The Cambrian period is acknowledged as the one in which a massive radiation of species occurred [Morris, 1989], giving rise to a wide range of body plans and diverse species.

These finding drove Gould to put into question if life followed a cone of increasing diversity, in which further gradual innovation appeared, and propose an inverted cone as the correct view of evolution of forms. The Cambrian explosion would have been the more diverse period in the history of evolution, followed by the extinction of most of the body plans with subsequent evolution of the remaining ones [Gould, 1990]. Giving the difficulty to assess this claim, Gould himself called for an analysis that would unveil if the disparity (a term coined by Gould to define the diversity of body plans) in the Cambrian was effectively higher than the currently observed one. His proposal was that this should be done using the theoretical morphology approach [Gould, 1991]. All in all, Gould defended the huge effect of historical contingencies in the evolution of species. In his retape experiment, he proposes that if done to the Cambrian period, the evolutionary outcome would have been completely different [Gould, 1990].

However, Simon Conway Morris himself, one of paleontologists specialists on the Burgess Shale fauna, puts into question the arguments of Gould. In his book ’The crucible of creation: the Burgess Shale and the rise of animals’ [Morris, 1999], he develops a series of counter-arguments to Gould’s claims. First of all, he acknowledges that during the years in-between the publication of their respective books, much more has been learned. Evidence at the time, he argues, do not support Gould’s metaphor of an ’inverted cone of life’, rather the contrary (see Fig. 1.10 for a schematic view on these ideas). The challenge proposed by Gould in comparing the disparity of modern phyla to the one in the Cambrian seems not to support Gould’s argument [Morris, 1999]. However, the sampling problem is evident from the start, so such a statement is difficult to assess given the inaccessibility to the whole variety of forms of the Cambrian [Morris, 1999]. Another argument Morris uses is that species supposedly isolated on account of their apparently weird anatomy are now beginning to be placed in a sensible framework of evolution [Morris, 1999].

Regarding the contingency argument of Gould, Morris argues that the evidence of convergent evolution might be leading to the fact that evolution is much more constrained than we would expect. According to Morris’ reasoning:

> But at the heart of Wonderful life are Gould’s deliberations on the roles of contingencies in evolution. Rather than denying their operation –and that would be futile– it is more important to decide whether a myriad of possible evolutionary pathways, all dogged by the twists and turns of historical circumstances, will end up with wildly different alternative worlds. In fact the constraints we see on evolution suggest

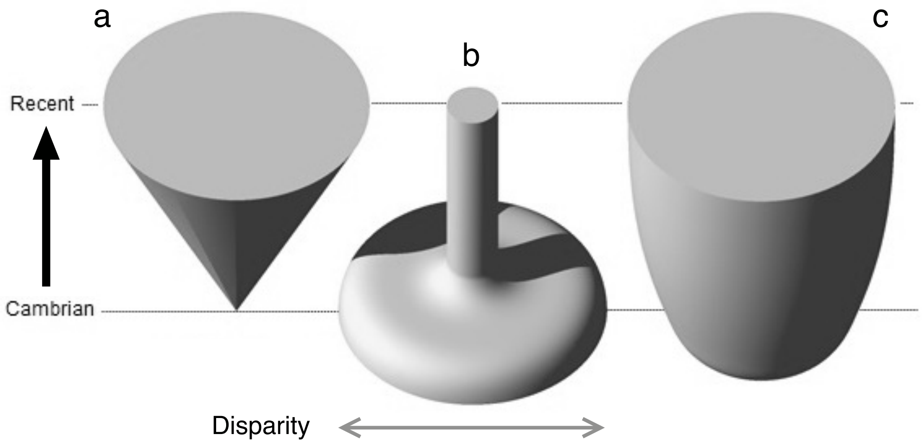


Figure 1.10: Various interpretation of the history of life include: (a) The traditional view whereby disparity steadily increases through geological time. (b) The view presented by S.J. Gould, whereby maximum disparity is in the Cambrian. (c) The view that disparity increased very rapidly in the Cambrian and thereafter stayed much the same or gradually increased, albeit at varying rates. Figure adapted from [Oyston et al., 2015].

> that underlying the apparent riot of forms there is an interesting predictability. This suggests that the role of contingency in individual history has little bearing on the likelihood of the emergence of a particular biological property. (...) *Gould's contingency argument* is based on a basic confusion concerning the destiny of a given lineage, be it of a human family or a phylum, versus the likelihood that a particular biological property or feature will sooner or later manifest itself as part of the evolutionary process. The point is that while the former, say the evolution of the whales, is from the perspective of the Cambrian explosion no more likely than hundreds of other end points, the evolution of some sort of fast, ocean-going animal that sieves sea water for food is probably very

> likely and perhaps almost inevitable. Although there may be a billion potential pathways for evolution to follow from the Cambrian explosion, in fact the real range of possibilities and hence the expected end results appear to be much more restricted. If this is a correct diagnosis, then evolution cannot be regarded as a series of untrammeled and unlimited experiments. On the contrary, I believe it is necessary to argue that within certain limits the outcome of evolutionary processes might be rather predictable. (...) Nevertheless, the point I wish to stress is that again and again we have evidence of biological form stumbling on the same solution to a problem. Consider animals that swim in water. It turns out that there are only a few fundamental methods of propulsion. It hardly matters if we choose to illustrate the method of swimming by reference to water beetles, pelagic snails, squid, fish, newts, ichthyosaurs, snakes, lizards, turtles, dugongs, or whales; we shall find that the style in which the given animal moves through the water will fall into one of only a few basic categories.

### Convergent evolution

Convergent evolution deserves special attention, given it is critical for the predictability arguments of Conway Morris. This same author widely acknowledged the problem of convergent evolution in his book Life’s solution [Morris, 2003], giving a myriad of examples, among the well-known ones and other less recognized. Another worth book on the subject is the one from George McGhee, ’Convergent evolution: limited forms most beautiful’ [McGhee, 2011], paraphrasing Darwin’s quote *’from so simple a beginning endless forms most beautiful... have been, and are being, evolved’* from his fundamental book ’On the origins of species by means of natural selection’ [Darwin, 1859].

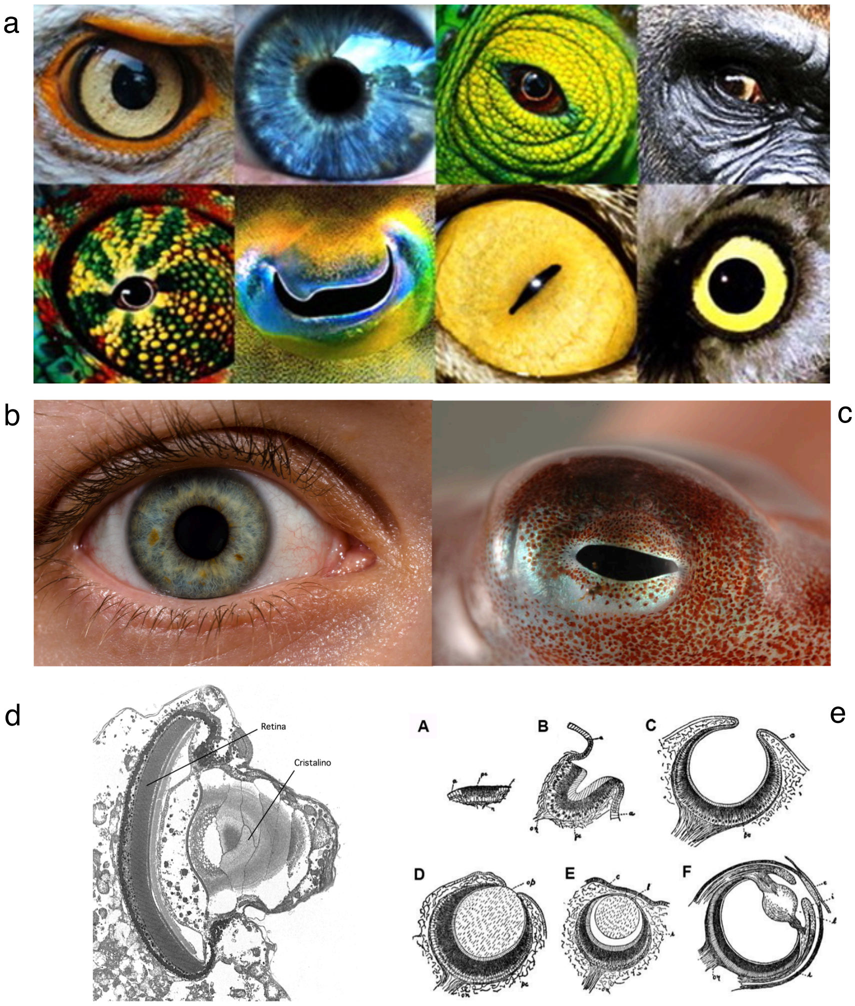


Figure 1.11: Convergent evolution of the camera eye, which display exactly the same design organisation in all the known life forms that use them. The design is externally variable (a), but essentially identical between humans and octopuses (b-c), except for a suboptimal placement of the retinal neurons in the first. The complex eye has been found even in single-celled organisms (d) and the sequence of basic steps for a gradual evolution have been identified (e) (adapted from [Solé, 2016a]). between humans

Convergent evolution is the term used to define the phenomena of evolution of similar traits (in distantly related species) which where not inherited from a common ancestor. These are called homoplastic traits, and there are three ways they may arise in evolution: convergence, parallelism and reversion (see [McGhee, 2011] for a detailed explanation). A crucial requirement for the correct identification of convergent evolution is that the underlying phylogenetic classification is correct and reliable. On the contrary, a trait inherited from a common ancestor could be misidentified as a convergent one.

The observation is that evolution has produced the same form –or a very similar one– over and over again in many independent species lineages, repeatedly, on timescales of hundreds of millions of years. Examples of convergent evolution exist from organic molecules to entire ecosystem of species, so it seems to be a rather universal principle [McGhee, 2011]. Convergent evolution is highly linked to evolutionary constraints (functional, developmental), which cause a reduction in the number of available evolutionary paths.

Examples of convergent traits abound. Considering locomotion, due to its functional constraints, there are a series of interesting examples, such as the high similarity between marine mammals and some extinct Mesozoic marine reptiles as well as the present day fast-swimming fishes. Streamlined, fusiform morphologiesl seem then to be convergent traits between these species. Regarding the ability to fly, one of the most well-known examples is the evolution of flying in land animals. No less than three separate groups of tetrapods have modified their forelimbs into wings (pterosaurs, living birds and bats). On the other hand, insect wings developed from the modification of gill branches originally present in the larval stages [McGhee, 2011].

The most recognized example is the convergent evolution (or parallel, as it has been argued recently) of the camera eye, which has been independently evolved in deuterostome chordates, protostome molluscs and arthropods, and nonbilaterian cnidarians [McGhee, 2011], adopting slightly different designs. A more dissimilar design, the compound eyes, has independently evolved eight different times within the three major types of protostomes, but not outside this clade. Actually, the evolution of eyes provides several illustrations of the importance of the concepts described so far. One one hand, compound eyes display a considerable wide variation in their external traits (see Fig. 1.11a) suggesting path dependent traits while sharing a remarkable set of common design principles shared by all known case studies.

The role played by accident is present in a subtle but revealing way. The human (Fig. 1.11b) and the octopus (Fig. 1.11b) eyes are essentially identical, except for an important design failure affecting the first: the light entering the eye crosses first a layer of neurons before it reaches the photoreceptors, to send back the signals and creating a well known anatomical anomaly: the blind spot. The Octopus eye instead shares the same optical design but the layers are properly placed. The universal design is present in jellyfish and even in unicellular organisms (Fig. 1.11d) where different groups of endogenous bacteria create an easily identifiable structure. Finally, compared anatomy and molecular phylogeny reveal the basic sequence of increasing complexity followed by complex eyes towards optimal function (Fig. 1.11e).

Given the evidence of convergent evolution, McGhee argues [McGhee, 2011],

> (...) the number of ecological roles or niches available for Earth organisms is demonstrably limited, in that species from many different phylogenetic lineages have been constrained in their evolution to filling the same ecological niche, even if that ecological role is extremely specialized. Ecological niche convergence is the rule, rather than the exception, in evolution. (George McGhee, 2011)

Conway Morris [Morris, 2003], meanwhile, extensively develops the argument of predictability at several levels in evolution in his book 'Life's solution: Inevitable Humans in a Lonely Universe', accounting for a broad range of examples, such as studies from the assessed optimality of the genetic code [Freeland and Hurst, 1998], as well as several examples of convergent features: from the different sensory systems to eusociality. An specially interesting example reviewed in the book is the ’experiment’ occurring in the island of New Zealand, a landmass which the mammals have failed to colonize. Remarkably, a bird showing a varied set of mammalian-typical traits have evolved: the kiwi. They are flightless, have fur-like feathers, are nocturnal (like many equivalent-sized mammals) and strongly dependent on the sense of smell (among other similar traits), being another compelling example on convergent evolution. The kiwis are called the ’honorary mammals’ [Morris, 2003].

### Final remarks

All in all, two primary positions are confronted: contingency vs. inevitability. The *actual* could be pointing to the second one due to the paramount examples of convergent evolution, but life history constraints do play a role. The debate between the role played by evolutionary, path-dependent processes versus the constraints associated to convergent evolution can accommodate to some extent within a broader picture emerging from the early work by Stuart Kauffman who has been a strong advocate for the concept of ”order for free” [Kauffman, 1993] .In Kauffman’s picture, there are very strong limitations to the dynamical processes emerging from networks of connected elements (such as gene regulatory networks). Some of them are deeply related to connectivity: lose or high levels of interactions produce frozen or chaotic behavior, while values close to a critical point separating these two phases provides a source of reliable behavior. A byproduct of criticality is the existence of generic properties that would pervade the observed properties of living matter.

Is it possible to answer Gould’s question about the tape of evolution? Evolution is a single run experiment and there seems to be no way to repeat it, so the only way we have to do an experiment such as the one proposed by Gould is to use simple organisms. These organisms can be grown in determinate environments exerting selection pressures, and the final outcome can be of significant value. The asexuality of bacteria is perfect for these experiments, as all the replicas can be performed with the same cloned genotype. Then, an assessment of the evolutionary trajectories is of critical interest. However, experiments in different contexts have been performed, with different conclusions. Some obtain remarkable repeatibility of the experiments, with the populations achieving the same final fitness [Travisano et al., 1995], suggesting the presence of convergent evolution, while others show high influence of chance and history [Lenski and Travisano, 1994].These experiments are interesting and not conclusive yet. However, any conclusion from this simple, unicellular systems needs to be taken with a grain of salt. It was the rise of multicellularity that predates the Cambrian Explosion what clearly makes this event an experiment in morphological, functional and ecological complexity. Is there any other alternative?

Interestingly, a computer experiment aimed at explore the origins of ecological diversity provided an unexpectedly affirmative answer to the last question. Tom Ray, a field ecologist, created a model based on a set of self-replicating programs competing for the RAM memory of the computer and having the potential to made mistakes while replicating [Ray, 1991]. With these artificial setup, programs started to evolve and several major transitions took place. Early selection events led to new generations dominated by shorter programs, which replicated faster than the original ones. The next evolutionary innovation was the emergence of parasites: computer codes unable to replicate themsleves but able to exploit the information carried by other programs. Hyper-parasites followed and afterwards, sex was invented: programs became capable of exchanging pieces of code to escape from the recognition by parasites. Eventually, several small programs that were not much fit as replicators were able to replicate faster than others by cooperating [Ray, 1991, Ray, 1998]. This kind of in silico evolution has been successfully used and expanded in the last decades [Adami, 1998, Ofria and Wilke, 2004, Solé and Valverde, 2013b].

The take home message of Ray's results is that *qualitative* universals might be inevitable as a generic, universal outcome of a large class of evolutionary processes. It is remarkable to see that such a simple comouter model lacking any kind of embodied realism is capable to deliver the sequence of evolutionary events that we can obviously map into observable life forms. It is likely that the introduction of a more detailed description of digital genomes or virtual agents creates the conditions for a more variable, perhaps path-dependent process. But nevertheless Ray's work suggests that a whole class of essential traits of real biology will inevitably emerge. If true that would mean that Gould's conjecture is, at the level of the logic of life, essentially wrong. And also supports the view that qualitative features of living complexity might be more relevant that some measurable ones. This is an important point in relation to the problem of defining qualitative morphospaces, which will be central to our approach in this dissertation.

In sections 1.1.1 and 1.1.3 we presented the Major Evolutionary Transitions, which are acknowledged to share important properties. Their study through the presented *synthetic* approach can lead to important contributions regarding the contingency or inevitability of these transitions, while assessing how likely they are to occur [Solé, 2016b]. Looking for universals, in trying to understand the relation between the possible and the actual is a rather appealing intellectual problem that for sure will bring novel ideas at the stage in the years to come.

### 1.1.6 Robustness in biological systems

Robustness is defined as a property that allows a system to maintain its functions despite external and internal perturbations, and it is argued to be a fundamental feature of evolvable complex systems [Kitano, 2004, Kitano, 2007, Whitacre, 2012, Stelling et al., 2004]. Evidence shows that robustness is a property widely observed across many species and across different levels of complexity, from gene transcription to the level of systemic homeostasis [Kitano, 2004]. In development, robustness is described as 'canalization' [Waddington, 1942] toward a specific outcome from uncertain starting conditions [Stelling et al., 2004]. Back in the 50s, John Von Neumann published his seminal work on 'Probabilistic logics and the synthesis of reliable organisms from unreliable components' [Von Neumann, 1956], precisely inspired by the paramount evidence of the robustness of biological systems, while engineering circuits faced a remarkable problem of failure when their components stopped working. Other works inspired on Von Neumann's one were published in the following years [Moore and Shannon, 1956, Winograd and Cowan, 1963].

Importantly, as Von Neumann already noticed, robustness is a systems-level phenomena that cannot be understood by the properties of the system individual components [Kitano, 2004]. As it is schematically shown in Fig. 1.12, it can work in either two ways: whether a system returns to its original state (attractor) after a perturbation, or it changes its state (another attractor) which maintains the system's functionality [Kitano, 2004]. As Kitano points out [Kitano, 2004], the tradeoffs that presents robustness with other properties, such as fragility, might be key to understand it. The microbiome behavior is a good example, as it is characterized for its ability to transition between alternative states, a property which has been recognized as fundamental for its robustness and fragility [Solé, 2016c].

Several mechanisms have been proposed in providing systems with robustness: system control , alternative (or fail-safe) mechanisms, modularity and decoupling [Kitano, 2004]. System control is mainly achieved through negative and positive feedbacks [Cinquin and Demongeot, 2002, Ferrell Jr, 2002, Angeli et al., 2004]. Regarding fail-safe mechanisms, they are the ones which, in case of failure of a given mechanism, rescue the functionality of the system. This can happen through redundant mechanisms (with the same structure as the failed one) or through degenerate mechanisms (structurally different). Degeneracy term is considered to be a ubiquitous biological property, and it encompasses examples such as the DNA code, gene duplication, cells in a given tissue, pathways in organismal development, immune responses or neural networks connectivity, among many others [Edelman and Gally, 2001].

Considering modularity, it is an efficient mechanism to contain perturbations locally, preventing their spread to the whole system. For instance, in a study of regulatory networks it was shown that links between highly connected proteins were systematically suppressed, while those between a highly connected and low-connected pairs of proteins were favored. This effect is clearly in the lines of decreasing the cross talk be-

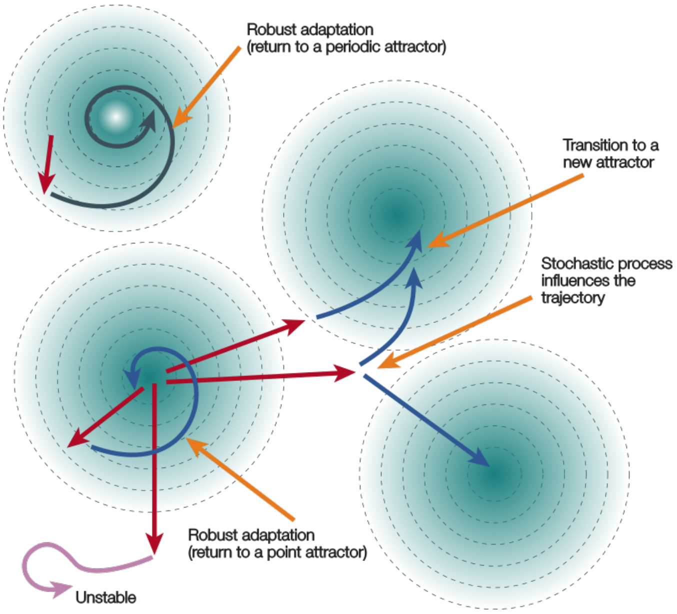


Figure 1.12: The concept of robustness depicted. State space simplified in two dimensions, in which a point represents the state of the system. Perturbations displace the state of the system, which might return to its original attractor or move to an alternative one. The perturbation can also lead to an unstable state. Adapted from [Kitano, 2004].

tween different functional modules so that deleterious perturbations get localized, increasing overall network robustness [Maslov and Sneppen, 2002]. Modularity is widespread in biological systems, so there exist major efforts in trying to understand its origins.

An hypothesis is that modularity might be the result of direct selective pressures. However, some works would refute the latter hypothesis, as it has been shown that modularity emerges as a by-product of the intrinsic dynamics of network growth by duplication and diversification [Solé and Valverde, 2007] or as a consequence for direct selection to reduce connectivity costs in the evolution of networks [Clune et al., 2013]. Remarkably, the breakdown of modularity also occurs, conveying a loss in systems robustness that can lead to ill-functioned systems, i.e. causing disease outbreak, thus a proper understanding is needed. Sergi Valverde has provided a worth effort in this direction [Valverde, 2017]. Decoupling, by definition, is a mechanism which isolates low-level variations from high-level functionalities. The most recognized example is the decoupling between genotype and phenotype with the mechanisms to cope with mutations [Kitano, 2004]. As it is pointed out by Kitano [Kitano, 2004], all mechanisms of robustness are ultimately linked to system control.

Hiraoki Kitano calls for the need of a mathematical theory of robustness, providing a valuable first step in his paper 'Towards a theory of biological robustness' [Kitano, 2007], in which he poses the importance of identifying the system, the function, and the particular perturbation to get valuable insights.

## 1.2 Multicellularity: origins and evolution

One of the METs, as stated in Section 1.1.1, is the transition to multicellularity (MC), in which we particularly focus in this thesis. Concretely, the transition from single cells with autonomous reproduction to an ensemble of cells in which the reproduction now occurs at the group level. The transition to MC is under the category of transitions in individuality. Those, as stated above, which consist of lower level entities gathering together and forming a higher level entity with reproductive capacity as a whole. In this case, the lower level entities are the individual cells, while the higher level entity is the ensemble of cells giving rise to a MC organism. The difficulties in studying this transition abound, but several approaches are starting to shed light into the question of how this transition occurred, given the *a priori* unlikelihood due to the loss of reproduction capacity of some of the cells (in most instances of MC) once the higher level entity is formed. One of the most important issues is to unveil which were the potential preconditions that, once at place, would readily enable the transition to MC to occur.

### 1.2.1 Origins

The rise of multicellularity has been reported to occur at least 25 times independently during the course of evolution when defined simply as cellular aggregation [Bonner, 2000, Grosberg and Strathmann, 2007, Niklas and Newman, 2013]. These several transitions occurred not only as separate events but at different times [Bonner, 2000] and in the three domains of life: Eukaryotes, Archaea and Bacteria. See Fig. 1.13 to check some of the independent MC origins in a simplified tree of life. The evidence, coming from fossils of prokaryotic filamentous and mat-forming Cyanobacteria-like organisms [Grosberg and Strathmann, 2007], situates the transition to MC $\sim$ 3.5 billion years ago [Grosberg and Strathmann, 2007, Bonner, 2000], while MC in eukaryotes dates back 1 billion years ago [Grosberg and Strathmann, 2007]. MC is likely to have originated only once in Metazoa [King, 2004], but multiple times in plants, fungi, and Eubacteria (with secondary losses) [Grosberg and Strathmann, 2007].

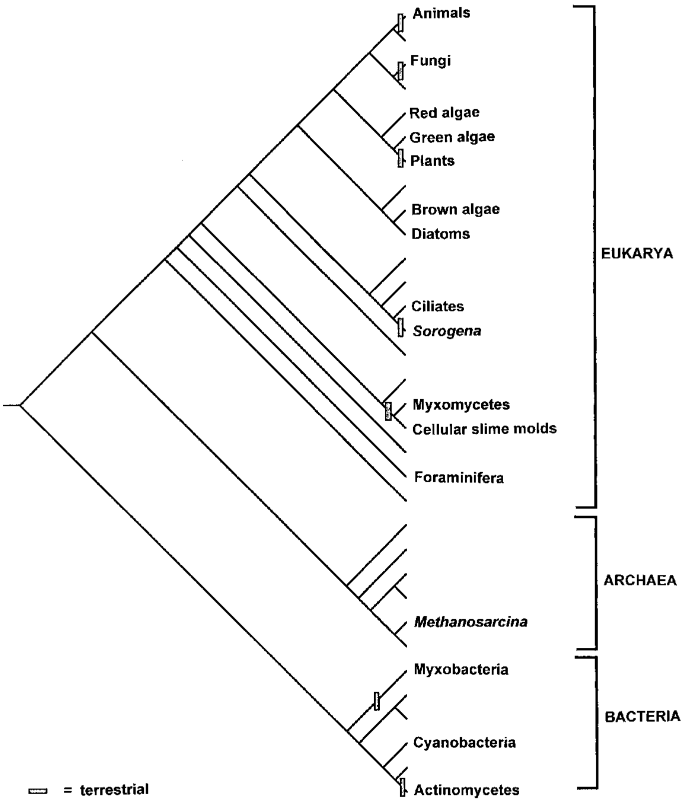


Figure 1.13: Multiple origins of multicellularity. Adapted from [Bonner, 2000]. A molecular phylogeny of the major groups of organisms, showing that multicellularity evolved independently a number of times.

John Tyler Bonner, a renown evolutionary biologist, points out the following:

> *'Indeed, there is nothing to rule out the possibility that at this very moment multicellularity is in the process of being invented by some single-cell form somewhere on our earth'.*
> [Bonner, 2000]

### A bottom-up approach to understand the origins of MC

In Sections 1.1.1 and 1.1.3 we have already developed the importance of a constructive approach with special focus on the properties emerging from the interactions of individual entities when trying to understand the formation of new hierarchical levels in the course of evolution. In the study of the transition to MC, this approach is critically important, as pointed out by Jordi Van Gestel and Corina Tarnita [Van Gestel and Tarnita, 2017], who highlight the importance of a bottom-up approach for the study of the transitions in individuality (MTI), which they coin Hierarchical Evolutionary Transitions (HET), with special focus on the life cycle (see Fig. 1.14), instead of the more frequently used top-down approach. The key importance of the life cycle in the evolution of individuality was already put forward by Leo W. Buss [Buss, 1987], while John Tyler Bonner [Bonner, 1965, Bonner, 1974] considered the life cycle to be the unit of selection in biology.

Van Gestel and Tarnita argue that the most generally undertaken top-down approach, starting with paradigmatic biological units and identifying their properties –i.e. cooperation, reduced conflict, distinct cell types, etc.– and then try to answer how groups might have evolved these properties, albeit remarkable contributions have been made in this way, it is not the most appropriate path to take to get profound insights on the origins of MC, in the lines of what has already been pointed out for the study of METs and MSTs. As they put it:

> This type of top-down approach to the study of HET runs into two critical problems. First, by focusing on properties of groups that qualify as paradigmatic examples of biological units, studies largely ignore the ancestor, including its internal organization and properties, the ecological context, and the mechanisms that gave rise to the primitive instantiations of those groups. As a consequence, it often remains unclear how the organization of the group –including the properties of interest– originated from that of the ancestor, making it impossible to fully unravel the evolutionary trajectory from

> the solitary ancestor to a new unit of biological organization. Which organizing principles and properties (e.g., differentiation, conflict suppression, metabolic specialization, cooperation) evolved de novo and which appeared as by-products due to strong interdependencies? What was the order in which organizing principles evolved? How did the organization at one point in time constrain or potentiate the evolution of new organizing principles? What is the relative importance of various factors (e.g., ecological context, conflict avoidance, development/physiology/ life history traits) for the evolution of new organizing principles? What types of organizing complexity can emerge from different ancestral properties and evolutionary trajectories?
>
> Second, in addition to ignoring the ancestral properties, by fixating on certain properties common to the known paradigmatic examples of HET, the top-down approach fails to explore the full potential of evolutionary trajectories and transitions, not only the paradigmatic but also the peripheral, and not only the actual (i.e., realized) but also the possible. This likely paints an incomplete picture of HET and precludes a valuable comparison across potential evolutionary transitions: only by comparing their full spectrum can we determine the causal factors that explain why certain trajectories did result in new units of biological organization and others did not. (J. Van Gestel and C. Tarnita, 2017)

In the light of these observations, they propose six questions to be answered in order to properly understand the HET:

i. When/how does a group originate that has the potential to undergo a HET?

ii. What emergent properties do these groups have?

iii. How does selection act on these properties?

iv. How does this affect the ancestral developmental program(s) and change group properties?

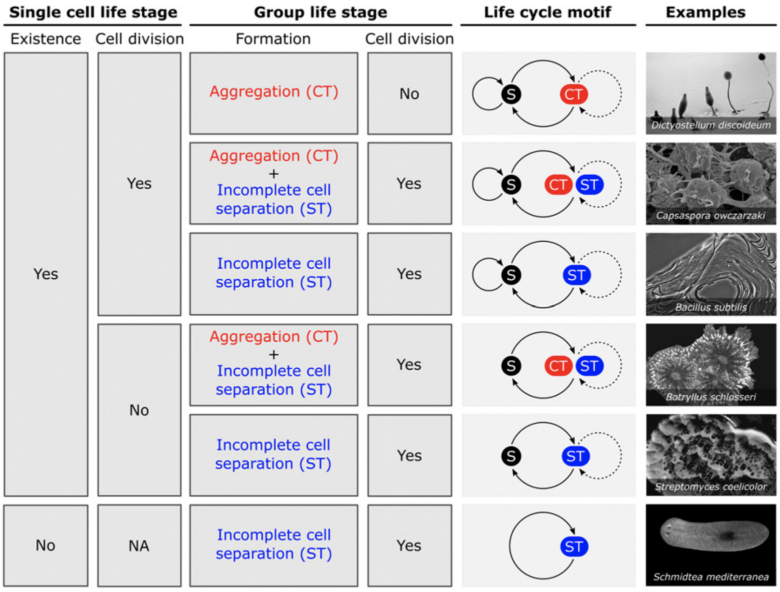
v. When/how does this lead to novel organizing/developmental principles within the new unit?

vi. What kinds of organizing complexity can evolve?

The problem of the origins of MC will be tackled through this bottom-up approach, either in the guiding thread of the explanations as well as in the interpretations.

### ***Routes*** **to Multicellularity**

Roughly, two ways to achieve MC have been described, either through (i) Clonal Development/MC or through (ii) Aggregative Development/MC [Bonner, 2000, Grosberg and Strathmann, 2007, Brunet and King, 2017, Sebé-Pedrós et al., 2017]. Clonal MC refers to a single cell origin, in which the cells resulting from successive divisions of the founder cells fail to separate, forming a cluster of genetically identical cells [Sebé-Pedrós et al., 2017]. All aquatic organisms developing MC structures used this clonal route to achieve it [Bonner, 2000] as well as some of the terrestrial, particularly the ones with complex body plans. On the other hand, aggregative MC refers to the attachment of initially independent cells, which can be genetically distinct, to form a MC entity [Sebé-Pedrós et al., 2017]. Aggregative MC evolved at least seven times in eukaryotes as well as in some bacterial lineages. In most cases of aggregative MC, cells respond to adverse conditions, such as nutrient depletion or hypoxia, by migrating toward each other and aggregating into a resistant mass of propagules (spores or cysts), or (if a stalk is present) a fruiting body. Aggregative MC is observed in terrestrial microorganisms (check Fig. 1.13) such as ciliates, myxomycetes, myxobacteria or cellular slime molds [Bonner, 2000]. In aggregative development, only subsets of cells present in the MC stage of the life cycle are dispersed for reproduction [Olson, 2013].



| Single cell life stage | | Group life stage | | Life cycle motif | Examples |
|---|---|---|---|---|---|
| Existence | Cell division | Formation | Cell division | | |
| Yes | Yes | Aggregation (CT) | No | S CT | *Dictyostelium discoideum* |
| | | Aggregation (CT)<br>+<br>Incomplete cell separation (ST) | Yes | S CT ST | *Capsaspora owczarzaki* |
| | | Incomplete cell separation (ST) | Yes | S ST | *Bacillus subtilis* |
| | No | Aggregation (CT)<br>+<br>Incomplete cell separation (ST) | Yes | S CT ST | *Botryllus schlosseri* |
| | | Incomplete cell separation (ST) | Yes | S ST | *Streptomyces coelicolor* |
| No | NA | Incomplete cell separation (ST) | Yes | ST | *Schmidtea mediterranea* |

Figure 1.14: Potential multicellular life cycles that could emerge upon the formation of the first MC groups. Categorization is based on (i) existence of single cell stage (S), (ii) mechanism of group formation (CT/ST), and (iii) life stage where cell division occurs. Adapted from [Van Gestel and Tarnita, 2017]. ST means ’stay together’, an analogy to refer to clonal development, while CT means ’come together’, used to refer to aggregative development.

These two alternative routes to MC have several implications, as we will see in the following sections. Both clonal and aggregative MC have scattered distributions in the eukaryotic phylogenetic tree (see Fig. 1.15), suggesting that they both evolved several times independently.

## Core concepts

The top-down approach mentioned by Van Gestel and Tarnita [Van Gestel and Tarnita, 2017], despite its drawbacks, is still useful to explain the core concepts related to MC organisms. For instance, in Metazoa, it is clear that the individual cells display cooperative traits (see Table 1.2) that improve the fitness of the whole organism, which is to be reproduced (giving rising to another alike organism) using only a particular cell lineage specialized in reproduction. Therefore, the remaining cell types have endured the so-called reproductive altruism (see Table 1.2) (they forgo reproductive capacity in the benefit of the higher entity they are part of). Moreover, there are many well-known mechanisms trying to avoid selfish behavior among these individual cells, as for example cell cycle control to avoid any dysfunction. However, some cells escape control mechanisms, leading to, i.e. cancer. Cancer cells are considered cheaters (see Table 1.2), as they have uncontrolled reproduction and can lead to the disruption of the whole organism as a fatal end point.

This example is aimed to have a clear idea on the cooperation and conflict concepts, which are the main important ones but we have listed in Table 1.2 other concepts that will be useful to properly understand the following sections. These concepts are rooted in the field of social biology, which is aimed at understanding the different relationships arising between biological entities. From an evolutionary perspective, social behaviours are defined as the ones having fitness consequences for both the individual that performs a particular behaviour and another individual [West et al., 2007], and it has applications at other levels, such as human societies.

Returning to the origins of MC, it is also clear that Metazoa have already evolved much more features since early MC organisms arose. The first MC aggregates could either be undifferentiated or have distinct phenotypes. In any case, cooperation between the units was likely to occur, and the threat of cheaters was already at play. William D. Hamilton [Hamilton, 1963, Hamilton, 1964], a pioneer in the field, hypothesized that genetic relatedness can be a key driver to maintain cooperation. Under this scenario, clonal MC would have higher possibilities to be maintained through kin selection compared to aggregative MC, more prone to conflict due to the (generally) lower genetic relatedness among its components. However, it has been argued that kin selection due to relatedness is not enough to maintain cooperation and that it is likely that early conflict-mediation mechanisms had to emerge to stabilize the groups. Mathematical models have been developed to account for this policy mechanisms [Frank, 1995].

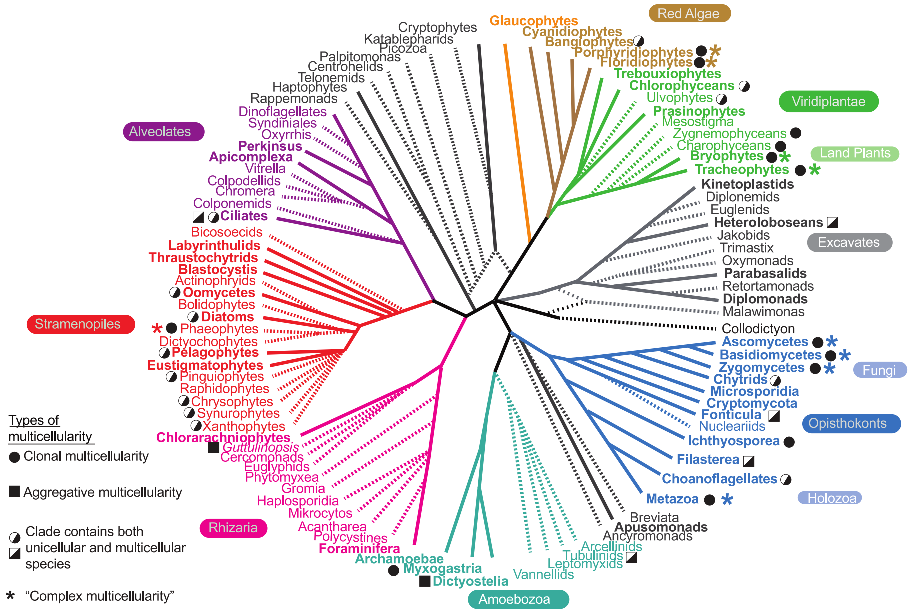


Figure 1.15: Eukaryotic phylogenetic distribution of clonal and aggregative multicellularity. Complex multicellular forms are also shown. Adapted from [Brunet and King, 2017].

However, cooperation (and probably control of conflict) is a necessary condition for a transition in individuality, but not sufficient: group reproduction must also evolve and, hence, a life cycle with at least a stage involving a MC form. This life cycle can take many different forms, as suggested by [Van Gestel and Tarnita, 2017] (see Fig. 1.14). In the case that not all the individuals in the group take part in reproductive tasks, it is the most clear example that a transition in individuality has occurred. Cooperation has benefits at the higher entity level, reporting an indirect fitness benefit to individual cells through the group reproduction. However, this is an intricate matter in which is easy to commit inconsistencies in describing it, as the germ-soma distinction is again one of the features that the top-down approach normally uses and does not represent all the possible routes to MC.

The Multi Level Selection theory tries to put a common framework for the study of transitions in individuality [Okasha, 2005], which are discussed to be divided into two levels, MLS1 and MLS2 [Okasha, 2005]. The former considers early stages in the formation of groups, describing conditions for the origin and maintenance of cooperation. The fitness of these groups is the average fitness of the component cells. MLS2 applies to groups that already exist as Darwinian individuals, that is, they are already the unit of selection and a reproductive strategy for the whole group has emerged. In this case, group fitness is defined as the number of collective offspring each group produces. It is likely that this fitness is already decoupled from the individual fitness of cells comprising the group, hence the transition in individuality has been completed. As discussed in [Libby and Rainey, 2013], this theory fails to explain how the transition from MLS1 to MLS2 occurs. A conceptual framework has been proposed, trying to operate free of assumptions such the ones the MLS theory impose [Libby and Rainey, 2013].

Richard Michod has been particularly involved in trying to understand the transitions in individuality [Hanschen et al., 2015] , developing a framework which encompasses various issues surrounding the problem of the origins of MC. The basic steps in a MTI would be: (i) formation of groups, (ii) increase of cooperation within groups, (iii) cheating and conflict, (iv) conflict mediation leading to enhanced cooperation, (v) division of labor in the basic components of fitness leading to (vi) fitness decoupling and individuality of the group. As he argues, altruism and other forms of cooperation lead to the transfer of fitness from the lower level (the costs of altruism) to the group level (the benefits of altruism). This so-called ’transfer of fitness’ does not need to be a conserved quantity, and the difficulties in defining the connections between individual and group fitness are evident [Hanschen et al., 2015] .

| Concept | Definition |
|---|---|
| Actor | Focal individual who performs a behaviour. |
| Recipient | An individual who is affected by the behaviour of the focal individual. |
| Cooperation | A behaviour which provides a benefit to another individual (recipient). Cooperation can be mutually beneficial or altruistic, depending on the effect on the actor (see below). |
| Direct fitness | The component of fitness gained through the impact of an individuals behaviour on the production of offspring. |
| Indirect fitness | The component of fitness gained from aiding the reproduction of related individuals. |
| Mutual benefit | A behaviour which is beneficial to both the actor and the recipient. |
| Altruism | A behaviour which is costly to the actor and beneficial to the recipient. |
| Relatedness | A measure of genetic similarity. |
| Kin selection | Process by which traits are favoured because of their beneficial effects on the fitness of relatives. |
| Mutualism | Cooperation between species. |
| Cheaters | Individuals who do not cooperate (or cooperate less than their fair share), but are potentially able to gain the benefit of others cooperating. Cheaters endure ’exploitation’ of a given trait or resource provided by other individuals which are cooperating. |
| Public goods | A resource that is costly to produce, and provides a benefit to all the individuals in the local group or population. |
| Tragedy of the commons | A situation when individuals would do better to cooperate, but cooperation is unstable because each individual gains by selfishly pursuing their own short-term interests. |

Table 1.2: Social behaviour concepts’ definitions [West et al., 2006, West et al., 2007].

### Advantages of Multicellularity?

Having explained the core concepts to understand the implications of studying the origins of MC, the key question is precisely this one: which are the advantages entailed by MC?

*Size increase-related advantages*

John Tyler Bonner has widely defended the view that size increase is a clear advantage of MC [Bonner, 1988, Bonner, 2000, Bonner, 2006]. An increase in size in the aquatic medium can imply the possibility of remaining fixed at one spot instead of being swept away through the currents. This could be advantageous in some environmental conditions, and it is the possible origins of colonial stalked ciliates and diatoms [Bonner, 2000]. An increase in size also gives an advantage in motility through the water for flagellated cells, as well as a way to avoid predation [Bonner, 2000].

The predation hypothesis deserves further attention, as it has been experimentally tested with positive results [Boraas et al., 1998] (explained in more details in Section 1.2.3). As Steven M. Stanley put forward back in 1973 [Stanley, 1973], the emergence of phagotrophic organisms, with a possible increase in size by preys to diminish its mortality due to predation, could have been the trigger for an arms race in size in prey-predator species, giving rise to a great species diversification, accounting for the fauna observed in the Ediacaran and Cambrian periods [Stanley, 1973]. In these lines, Bonner argues that there is always an open niche at the top of the size spectrum, which can report a selective advantage as it is the one realm that is ever available to escape competition. He argues that it is equally possible to evolve smaller as well as larger depending on the ecological conditions and the niches available, and evidence from phylogenomic studies supports this idea [Bonner, 2000].

Other advantages to size increase have been suggested, such as the expansion of feeding opportunities, generating an internal environment protected by an external layer of cells or allowing novel metabolic opportunities, among others [Grosberg and Strathmann, 2007]. In Section

1.2.3 we provide a compelling example of the evolutionary drivers behind a transition to undifferentiated MC [Pfeiffer and Bonhoeffer, 2003].

*Division of Labour (DoL)*

It is well-known the capacity of unicellular organisms to produce diverse phenotypes in response to determinate environmental conditions. However, MC organisms can simultaneously perform distinct tasks if they are shared among its individual cells. This could have been a clear advantage for MC, accounting for several examples. The best-known is the germ-soma division, with cells accounting for reproduction-altruism (which is also one of the most difficult changes to explain). The Bacteria domain has several examples, such as the metabolic cooperation occurring in *Cyanobacteria*. Some key metabolic processes cannot occur simultaneously within a cell, for instance photosynthesis and nitrogen fixation. A solution is to do photosynthesis during the day and nitrogen fixation at night, but MC did appear in these bacteria in the form of heterocyst, with a division of labour among the cells performing the two processes simultaneously. On the other hand motility-mitosis trade-offs might have been another evolutionary driver for the emergence of MC with DoL [Grosberg and Strathmann, 2007].

### The threat of cheaters

We have already discussed that the emergence of conflict among cooperating cells is a likely outcome, in the form of cheater cells exploiting i.e. common resources of the group. The most commonly suggested mechanisms that evolved in order to deal with the presence of cheaters are various policing mechanism to control directly the emergence of cheaters, the ’unicell bottleneck’ strategy, which accounts for clonal reproduction, ensuring the minimal genetic conflict in subsequent generations, germ-line early sequestration (to avoid either mutation in the germ line and at the same time segregating it from other cells to avoid competition) and programmed cell death, selectively performed by cells in order to not disrupt the MC organism [Grosberg and Strathmann, 2007].

An interesting hypothesis regarding the cheaters will be further developed in Section 1.2.3, which consists on considering cheaters as a primitive germ line which, if appropriately switching to the MC cooperating phenotype, could have been a very primitive form giving rise to a life cycle [Rainey, 2007, Rainey and Kerr, 2010, Hammerschmidt et al., 2014].

Finally, it is important to mention parasites, which are conjectured to be one of those inevitable outcomes of evolved systems [Solé, 2016b]. The presence of parasites can destabilize cooperative systems, but it has been suggested that compartments can strongly constrain their impact [Solé, 2016c].

### 1.2.2 Landscapes and preconditions

The core concepts regarding the origins of MC have been presented. However, a lot of issues remain to be tackled. Which is the landscape of possibilities regarding the alternative evolutionary paths to MC? Do we have evidences of such transitions in nature? There exist model organisms? Which were the set of conditions predating the origins of MC? How can we access them? Here we will try to give a broad view on what is known and the ways/tools we have to answer these kind of questions. It will be continuously seen that the environment had a key role in shaping the preconditions to MC.

#### Insights from the study of UC relatives and model organisms

Deciphering the evolutionary transitions to multicellularity is not an easy task, requiring a robust deep phylogeny of eukaryotes to shed light into the relationships amongst multicellular groups and determine their closest unicellular relatives [Sharpe et al., 2015]. The lack of model organisms close to the rise to MC difficults the study of this transition, but there exist some exceptions. The best known model organism is Volvox (see Fig. 1.16), from the family of Volvocales algae. The *Volvocales* range from their single-cell *Chlamydomonas* –the ancestral type– to the large *Volvox* made up of many thousands of cells, while there also exist a series

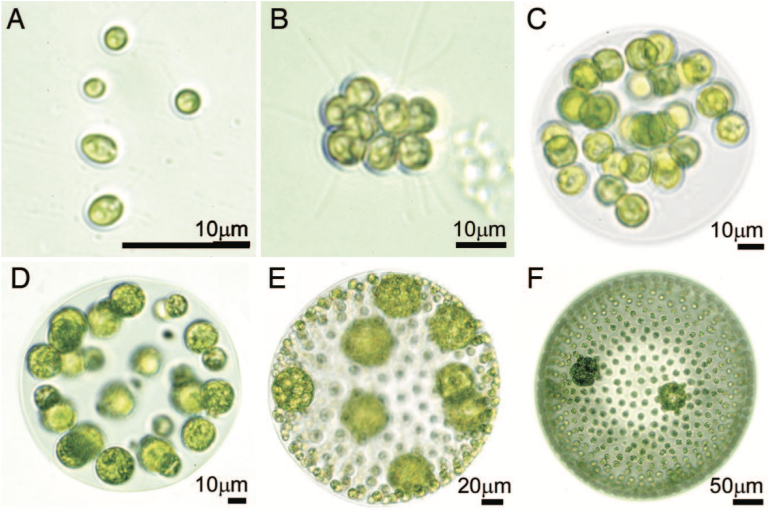


Figure 1.16: Examples of Volvocine species varying in cell number, colony volume, degree of specialization, and proportion of somatic cells. (A) *C. reinhardtii*, a unicell. (B) *Gonium pectorale*, a flat or curved sheet of 8-32 undifferentiated cells. (C) *Eudorina elegans*, a spherical colony of 16-64 undifferentiated cells. (D) *Pleodorina californica*, a spherical colony with 30-50 percent of somatic cells. (E) *Volvox carteri*. (F) *Volvox aureus*. Where two cell types are present (D-F), the smaller cells are somatic cells and the larger cells are reproductive cells. Adapted from [Michod, 2007]

of genera with intermediate-size colonies. In Volvox, large size is accompanied by a division of labor: most of the cells remain vegetative and concern themselves solely with photosynthesis and locomotion, while a few cells are either asexual or sexual reproductive cells and are able to start the next generation. In the smaller species, all the cells manage both functions [Bonner, 2000].

Therefore, *Volvox* would be a perfect example of simple division of labour, with the germ-soma differentiation. Somatic cells improve the fitness of the MC organism while having endured reproduction altruism. The *Volvox carteri* genome differs only in minor ways from that of its close unicellular relative *Chlamydomonas reinhardtii* [Knoll, 2011], and some studies have even proposed what processes had to be added to the ancestral developmental repertoire of *Chlamydomonas* in order to evolve a developmental repertoire like that of modern *Volvox carteri* [Kirk, 2005, Sachs, 2008]. Approaches like this can be important contributions in unveiling the preconditions necessary, in this case, for the division of labour emergence.

However, most of the efforts have concentrated on the study of the origins of Metazoa (animals). Phylogenomic studies have considerably changed the understanding of the tree of life (see Fig. 1.15), confirming that the Metazoa are sister to the choanoflagellates, but also two additional independent lineages, the filastereans and the ichthyosporeans, have been confirmed to be closely related to Metazoa. These three unicellular lineages (choanoflagellates, filastereans and ichthyosporeans) form a clade with Metazoa, called Holozoa [Sebé-Pedrós et al., 2017].

The approach followed is to determine the genes and genetic pathways that are shared between animals and their relatives, to then infer which genes and genetic pathways were present in the ancestor [Sebé-Pedrós et al., 2017]. It has been found that despite the gene innovation at the onset of Metazoa, the unicellular ancestor of animals already had a rich repertoire of genes that are required for cell adhesion, cell signalling and transcriptional regulation in modern animals [King, 2004, Richter and King, 2013, Sebé-Pedrós et al., 2017] .

This fact would support the hypothesis that this ancestral genes repertoire was subsequently repurposed for multicellularity through co-option. Together with the evolution of novel animal genes and a substantial expansion and diversification of some ancestral gene families, they configured the gene toolkit for animal MC [Sebé-Pedrós et al., 2017].

However, the gene content alone is only an additional layer of knowledge but not sufficient to shed light into the cell biology, life cycle and

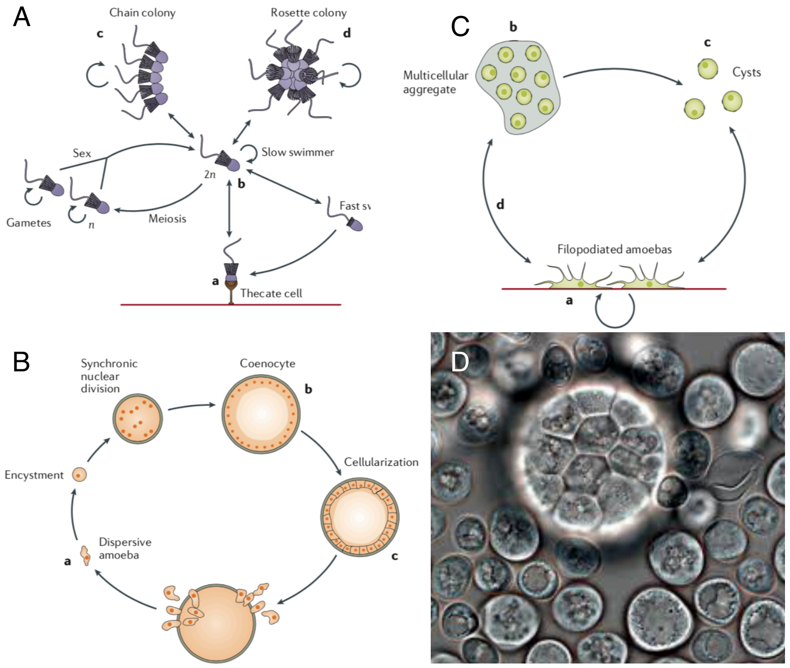


Figure 1.17: The life cycle of UC relatives of Metazoa: (A) the choanoflagellate *Salpingoeca rosetta*, the ichthyosporean *Creolimax fragrantissima* and the (C) filasterean *Capsaspora. owczarzaki*. An image of Capsaspora is displayed in (D). Adapted from [Sebé-Pedrós et al., 2017].

regulation capabilities of the unicellular ancestor. This requires analysis of the biology of the extant unicellular relatives of animals, with several interesting behaviors observed. A choanoflagelate species, *S. rosetta*, forms colonies by clonal cell division, and its formation is triggered by the presence of its bacterial prey, suggesting a deep evolutionary link between the origins of early MC and prey capture. Furthermore, it is observed that cadherins do not have a cell-cell adhesion role in these colonies, but putatively developed an active role in prey capture [King, 2004, Sebé-Pedrós et al., 2017], supporting the hypothesis of co-option of cell adhesion genes in Metazoa.

On the other hand, the study of *C. owczarzaki*, a filasterean, unveils its aggregative MC, being the only example of this route to MC observed in Holozoa. This has led to new hypothesis regarding the origins of Metazoa, suggesting that clonal development arising from an initial aggregative MC cannot be ruled out in explaining the origins of MC in Metazoa [Olson, 2013]. The study of *C. owczarzaki* life cycle showed temporally regulated cell differentiation, triggered by starvation conditions. Altogether, these findings suggest that the unicellular ancestor of Metazoa was a bacterivor displaying sexual reproduction and multiple temporally differentiated cell types. The transitions between cell types were likely to be triggered by environmental conditions such as lack of nutrients and prey presence [Sebé-Pedrós et al., 2017]. These findings suggest that the first Metazoa evolved from unicellular ancestors with an already complex life cycle displaying distinct phenotypes through a transition from temporally to spatiotemporally regulated cell differentiation [Sebé-Pedrós et al., 2017].

This hypothesis is confronted with a possible division of labour emerging after MC aggregates were at place. Thibaut Brunet and Nicole King name these two hypothesis as temporal-to-spatial transition (TST) and division of labour (DoL) [Brunet and King, 2017]. Therefore, TST hypothesis proposes that cell differentiation predated MC in the light of the observation of temporally-regulated cell types in the unicellular relatives, while the DoL hypothesis suggests that cell differentiation evolved after MC by differential loss of function from multifunctional ancestral cell types. As they discuss, the two hypothesis are complementary rather than mutually exclusive [Brunet and King, 2017], and probably both phenomena have occurred in the multiple independent origins of MC.

As a final remark, recent analysis of sponge and unicellular holozoan cell transcriptomes, development and behaviour provides support to the hypothesis that the ancestral metazoan cell type had the capacity to exist in and transition between multiple cell states in a manner similar to modern transdifferentiating and stem cells [Sogabe et al., 2019].

### Common features among independent MC origins?

It is important to mention the efforts focused on unveiling commonalities between the origins of MC in different clades in eukaryotes. The studies undertaken in this regard conclude that the extent to which multicellularity is achieved using the same toolkits and modules (and thus the extent to which multicellularity is homologous among different organisms) differs among clades and even among some closely related lineages [Niklas and Newman, 2013].

However, different molecular systems have evolved to serve similar functions in different multicellular lineages [Sebé-Pedrós et al., 2017], being a case of molecular convergence. A typical example is the one of cell adhesion molecules [Abedin and King, 2010], which is mediated by different molecules in different multicellular taxa: in plants is largely mediated by pectins and hemicelluloses (’glues’), in fungi it involves extracellular glycoproteins while in animals it is mediated by cadherins and integrins (transmembrane proteins), and in brown algae it involves a series of polymers [Sebé-Pedrós et al., 2017]. What can explain the existence of widely different modes of cell adhesion in each of the different MC lineages? [Abedin and King, 2010].

What is observed is that each of the MC lineages is more similar in terms of their TFs or their signalling repertoires to their unicellular relatives than to other MC groups. Therefore, despite some common trends, there is no universal gene toolkit for multicellularity [Sebé-Pedrós et al., 2017], but the interest relies on the fact that despite differences in the identity of molecules, they end up developing very similar functions.

### Oxygen as a precondition for MC?

Several examples of environmental-driving forces for the origins of MC have already been mentioned: predation and starvation are the most clear examples. However, if we go 2.7 billion years ago, we find one of the most important revolutions on Earth: the rise of oxygen in the Archean through water-splitting oxygenic photosynthesis. The rise of oxygen in the atmosphere led to the ozone layer formation, protecting the DNA-bearing organisms from the UV radiation damage. However, reactive oxygen was all around, and other types of defences had to be developed to cope with it. This led to a stratisfied biosphere, with obligate anaerobes living at the anoxic waters of the deep oceans. However, oxygen meant energy, as the product of oxygenic photosynthesis was both oxygen and carbon. Therefore, aerobic respiration could become a dominant way of breaking down organic matter into energy in surface environments. This is proposed as a clear environmental-driving force for the origins of larger and more complex cells which ultimately would have given rise to MC organisms [Lenton et al., 2011]. The rise in oxygen levels after the colonization of land and after the great glaciations is though to have facilitated the Edicaran metazoan evolution and the subsequent radiation during the Cambrian period [Johnston et al., 2012]. Besides, predation and the arms race stemming from it together with ecological factors were for sure catalyst for the diversification of life forms [Johnston et al., 2012, Lenton et al., 2011].

### What can be learned from bacteria?

As already pointed out in Section 1.1.1, bacteria display a wide range of MC behavior, most of them triggered by certain environmental conditions, such as stressful conditions, giving rise to cooperative behaviors. However, another interesting behaviour observed in bacteria is the stochastic phenotypic switching. This is characterized by the change to distinct stable cell types without the need of a particular trigger, the change occurs randomly. This has been suggested to be the result of an exposition to highly fluctuating environments as a bet-hedging strategy. Bet-hedging, by definition, is a risk spreading strategy to diversify phenotypes with the aim to increase fitness in temporally variable conditions. To stochastically transition among multiple phenotypes can have fitness advantages in these conditions, as it is ensured that some cells are always prepared for an unforeseen environmental fluctuation [Acar et al., 2008, Veening et al., 2008].

### 1.2.3 The *synthetic* transition

In the previous section, a thorough revision of systems that can give insights in the transition to MC and, more specifically, to the preconditions of the transition, has been made. We observe how the influence of environmental-driving forces must have been a key factor. A powerful way to provide more insights into the open questions regarding this MTI is to use the *synthetic* approach presented in Section 1.1.3. Some of the most relevant examples providing significant insights in the origins of MC are presented below. From experimentally evolved life cycles to cooperative behavior emerging in *in silico* models, it can be envisaged that, once the appropriate conditions are at place, the transition to MC might have not been a difficult one. One of the aims of the study of METs through a *synthetic* approach is to explore all the possible routes to MC, which can give insights on how restricted (or not) might have been this transition.

#### Experimentally evolved MC

*Bacteria as the perfect testbed*

Bacteria are prokaryotes easy to manipulate in the laboratory with rapid reproduction and also highly controllable at the genetic level due to its clonal development. Therefore, they have been widely used to test several hypothesis regarding the origins of MC. Is the rise of MC an easily accessible transition?

Paul B. Rainey and Michael Travisano developed a set of experiments using *Pseudomonas fluorescens* [Rainey and Travisano, 1998] that served as a testbed for multiple hypothesis in the subsequent years. In this first work [Rainey and Travisano, 1998], they tested the if adaptive radiation could be a response to an heterogeneous environment. Adaptive radiation is defined as a rapid diversification of organisms due to the availability of new resources or niches. In their experiments, they observed a rapid diversification of the bacteria when exposed to a spatially structured environment (heterogeneous conditions), providing with ecological opportunity. When ecological opportunity was restricted (homogeneous conditions), such diversification did not occur or even reverted. Therefore, in a simple experiment they provided a clear example that adaptive radiation can be triggered (in a few number of generations) by the exposition to heterogeneous conditions (ecological opportunity).

In a follow-up experiment, Paul B. Rainey and Katrina Rainey used the same experimental setup to shed light into more details regarding a particularly frequent phenotype that was observed when exposing *Pseudomonas fluorescens* in a spatially heterogeneous environment, the so-called wrinkly spreader (WS). The WS phenotype colonized the liquid-air interface by the over-production of an adhesive polymer (see Fig. 1.18a). This behavior reported an advantage to cells as they had a better accessibility to both nutrients and oxygen, but the production of the polymer was costly. It is a clear example of cooperative behaviour giving rise to an undifferentiated MC aggregate (WS phenotype), for which the selective advantage is under debate (further contributions will be presented in the *In silico* approaches). In the study it was shown that this cooperative behaviour aligned the interests of individuals with those of the group. However, the group was susceptible to cheaters: reverting the mutation and not producing the adhesive polymer permitted cheater cells to exploit the advantages of remaining at the liquid-air interface. The reproductive advantage of defectors eventually caused the disruption of the MC aggregate. Remarkably, after each WS mat collapse, the WS phenotype emerged afresh maintained by negative ferquency-dependent selection [Rainey and Rainey, 2003]. Therefore, it was put forward a compelling example of a robust formation of undifferentiated aggregates, even if susceptible to defectors.

It was precisely the outcome of the mentioned experiments that led Paul B. Rainey to formulate an attractive hypothesis for the origins of a primitive life cycle [Rainey, 2007, Rainey and Kerr, 2010]. As we have already pointed out, cooperation is not a sufficient condition for a complete route to MC. A key issue is how variation in lower-level individuals generates a corporate (collective) entity with Darwinian characteristics. Of central importance to this process is the evolution of a means of collective reproduction, however, the evolution of a means of collective reproduction is not a trivial issue, requiring careful consideration of mechanistic details [Rainey and Kerr, 2010]. Control of conflict is another hallmark for the rise of MC, however, he challenges this view and proposed that cheaters in a cooperating cell ensemble can precisely be the founders of this same ensemble. In other words, conflicts between selective forces might fuel, rather than hinder, the transition to higher levels of complexity [Rainey, 2007]. The group, if unable to replicate itself, is analogous to the soma. The cheater, however, can be considered to be loosely analogous to a primordial germ line. The life cycle would be closed if the cheater switched back to the cooperating phenotype able to produce the group. It is not an implausible scenario given the evidence shown in the previous experiments [Rainey and Travisano, 1998, Rainey and Rainey, 2003].

He achieved a compelling proof-of-concept on this idea together with his colleagues in a set of experiments presented in 2014 [Hammerschmidt et al., 2014]. Using again *Pseudomonas fluorescens*, they observed that collectives reproduced via life cycles that either embraced, or purged, cheating types. When embraced, the life cycle alternated between phenotypic states through what could be considered a selected developmental switch. They show how the fitness of the collective, when using this route, became decoupled from the fitness of constituent cells. Therefore, they obtained in the laboratory a testable example of a transition in individuality through a non-widely recognized route: a life cycle alternating between MC and unicellular states, represented by cooperators and defectors, respectively.

The latter contribution to the field using *Pseudomonas fluorescens* was the experimental evolution of bet-hedging. Facing *Pseudomonas fluorescens* to a set of fluctuating environments, they obtained bacteria adopting a rapid phenotypic switching mechanism that they found to be linked to a particular mutation (interestingly, the evolution of bet-hedging was contingent upon earlier mutations that affected the particular mutation enabling the phenotypic switch). Therefore, these experiments elegantly capture the evolution of a bet-hedging mechanism in the face of fluctuating environments, suggesting the likelihood of being one of the earliest evolutionary solutions to life in this kind of environments.

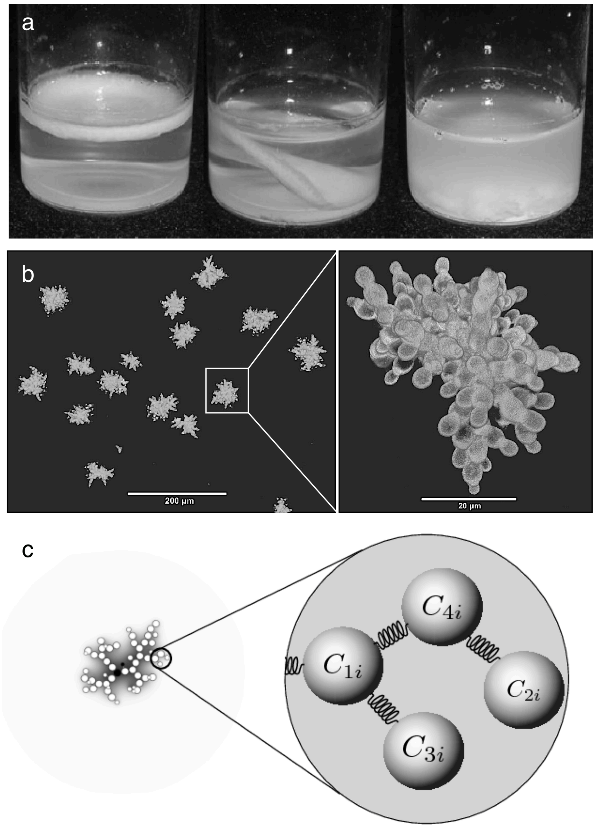


Figure 1.18: Examples of experimentally evolved multicellularity. a) Cooperative strategy developed by the WS phenotype *Pseudomonas fluorescens*. Adapted from [Rainey and Kerr, 2010, Hammerschmidt et al., 2014]. b) Evolution of MC using gravity as a selective pressure, the ’snowflake’ phenotype arises. Adapted from [Ratcliff et al., 2012] c) Simple models allow to synthetically evolve *in silico* these structures by using embodied descriptions of cells and their interactions [Duran-Nebreda and Solé, 2015].

All in all, these experiments show how the diversification into distinct phenotypes is possible when the appropriate environmental conditions - and selective pressures- are at place, hence confirming the ’easy’ access to one of the proposed preconditions for MC: cell differentiation in a given species. The life cycle example shows also that this precondition can lead to a primitive life cycle.

*Experimentally evolved life cycles in yeasts and algae*

Other examples of experimentally-evolved life cycles had already been achieved before the one just described, but the organisms used were not bacteria but yeast and a species of unicellular algae. William Ratcliff, Michael Travisano and colleagues [Ratcliff et al., 2012] developed a simple but elegant experiment using *Saccharomyces cerevisiae* in which a MC life cycle evolved. They imposed a particular selective pressure: gravity (an experimentally tractable method to select for larger size). They selected for the most rapidly settling cell clusters. In a few generations, they obtained MC aggregates (which settled faster) that also evolved a primitive life cycle through selective apoptosis of certain cluster cells: division of labour through reproductive altruism had evolved, as some cells committed programmed cell death in the benefit of the MC aggregate, freeing MC propagules that would increase again in size. Remarkably, the clusters reached a characteristic size, having developed determinate growth, and 'juvenile' and 'adult' clusters could be distinguished. The initial formation of clusters due to gravity were uniclonal, diminishing genetic conflict.

Again Ratcliff, Travisano and colleagues [Ratcliff et al., 2013] experimentally evolved a life cycle in a unicellular algae, *Chlamydomonas reinhardtii*. The selection pressure was the same as in the previous experiment, gravity, but in this case the life cycle evolved was of a different nature, as it reproduced via motile unicellular propagules. The relevance of the work relies upon the fact that this single cell propagules are found to be adaptive even in absence of intercellular conflict, maximizing cluster-level fecundity (which would not support the hypohtesis of a single-cell bottleneck in order to limit among-cell conflict).Therefore, this would be a compelling example of a protolife cycle in which the unicellular bottleneck can arise rapidly via co-option of the ancestral unicellular form [Ratcliff et al., 2013].

These examples [Ratcliff et al., 2012, Ratcliff et al., 2013] are remarkable on their own because using a simple selective pressure they achieved two hallmark features of MC: division of labour and group reproduction, suggesting that given the appropriate ecological conditions, the rise of MC need not to be a particularly difficult transition to achieve.

*Origins of MC due to predatory pressure: a compelling example*

One of the hypothesized advantages for MC has been mentioned to be predation avoidance. A powerful experiment proving the validity of this hypothesis is the one provided by Martin E. Boraas and colleagues [Boraas et al., 1998], in which a green unicellular algae, *Chlorella vulgaris*, a well-studied eukaryote which retained its unicellular form in cultures for thousands of generations, developed MC growth when a predator (a phagotrophic flagellated protist, *Ochromonas vallescia*) was inoculated in the culture. With less than 100 generations, the MC form of *Chlorella vulgaris* became dominant in the culture. Remarkably, the MC aggregates reached a characteristic size of eight cells: virtually immune to predation due to its size but small enough so that the nutrient in the medium was available for all the cells in the aggregate. These MC colonies were indeed self-replicating [Boraas et al., 1998]. The hypothesis is that this kind of adaptation could trigger an arms race in size in prey-predator species, giving rise to a great species diversification [Stanley, 1973].

### ***In silico* MC**

*A possible origin for undifferentiated MC*

When following a bottom-up approach for the origins of MC one must analyse the most primitive conditions in which MC might have been advantageous. The emergence of initially undifferentiated MC aggregates is a clear possibility for the origins of MC, but the advantages in the lack of division of labour can be harder to find. Thomas Pfeiffer and Sebastian Bonhoeffer provided a compelling example in which undifferentiated MC can provide an advantage [Pfeiffer et al., 2001, Pfeiffer and Bonhoeffer, 2002, Pfeiffer and Bonhoeffer, 2003]. They analysed if the existent tradeoff between yield and rate of ATP production might have any evolutionary consequences. Species displaying high yield of ATP (the production of ATP with respect to the resource quantity consumed is high) and low rate of production (the production of ATP is very efficient in terms of resource consumption, but low in time) are the ones using in a more efficient manner the resources available. However, an opposite strategy would be to perform with low yield of ATP (very few ATP units / unit of resource) but at a high rate (ATP production is fast in time). This second strategy ensures a faster ATP production but the rapid depletion of resources is the principal drawback.

In their model, they found that species displaying the first strategy clustered, forming undifferentiated aggregates, in order to cooperate in the use of external energy resources. The first strategy will be disadvantageous if surrounded by individuals displaying the second, but surrounding with similar individuals ensures a safer strategy in resource consumption and ATP production. They also showed how further advantages of the clustering strategy could arise from the exchange of resources between the cells of the cluster.

*Different life cycles emerging in an artificial life model*

M. Staps, J. van Gestel and C. E. Tarnita [Staps et al., 2019] developed the first theoretical approach in which, starting from the unicellular ancestor, the multicellular life cycles and life histories that can originate were systematically explored. In their work, they evolved a regulated phenotypic switching in individual cells when exposed to two distinct environments. The particular phenotype was advantageous regarding reproduction in each of the separate environments. Then, they assumed that an adhesion gene was co-opted with the activation of one of the genes that gave rise to a specific phenotype. Furthermore, they introduce a mechanism of defense of MC, as the mortality of cells due to predation was reduced if they reached a critical size. Under this framework, and permitting the evolution of a simple neural network regulating the cell phenotypes, they kept changing the input environment with different values of cell stickiness and the critical size needed to avoid predation. Up to six different life cycles emerged, some of them having counterparts in natural systems. Some where expected under the conditions tested in the model while three of them were quite the contrary, showing the strength of the approach in sampling varied modes of life under environmental pressures once cell adhesion was a possibility [Staps et al., 2019].

*The niche concept and the importance of embodiment*

The niche concept can have different interpretations [Pocheville, 2015], but here we understand a niche as an environment with particular conditions that might be advantageous (or not) for certain species survival. From the aforementioned examples, it is already clear that the exploitation of new niche opportunities can drive evolutionary transitions and particularly the rise of MC, the matter at hand. However, a commonly forgotten ingredient for the transition to MC are the physical properties of the given system, as well as the role of embodiment, which was already pointed out to be a key ingredient in the DPMs framework [Newman et al., 2003] and references therein. The role of physical embodiment cannot be ignored, as boundary conditions can play a significant role, as it was demonstrated in an artificial life setup where the physical context, ecological factors and cell-cell adhesion were considered with the aim to give insights on the preconditions needed for MC to emerge [Solé and Valverde, 2013a]. This computational model of evolving virtual organisms showed the importance of the physical embodiment and ecological factors in the evolution of individual, free-moving cells. In this computational model of evolving virtual organisms, free-moving cells evolved to exploit a given ecological niche through an evolutionary transition in their adhesion parameters which ultimately induced a change in the nutrient flow that triggered the emergence of a new population of detritivors, with individuals now separated in two well-defined compartments. This is a compelling example of ecosystem engineering, not previously found in an artificial life setup [Solé and Valverde, 2013a], which clearly demonstrates the importance of the physical embodiment of cells. Moreover, it is also an example were multiple scales become tangled, which is critical when trying to uncover universal patterns in macroevolutonary dynamics

[Solé and Valverde, 2013b].

As it has been clearly exemplified with this artificial life example, organisms influence their environment, and this can ultimately modify the selective regime of the builder (a phenomena called ’niche construction’) while also can influence the ecological success of other species (a phenomena called ’ecosystem engineering’, as the one described in [Solé and Valverde, 2013a]) [Erwin, 2008]. It is suggested that the influence of these processes has been increasing over time, with positive feedback through environmentally mediated selection being a critical factors for the enhanced biodiversity observed through the Phanerozoic eon [Erwin, 2008].

*Emergence of proto-organisms in a resource-waste context*

One of the mechanisms susceptible to predate simple MC forms is stochastic phenotypic switching, which as discussed above emerges as an adaptive strategy to fluctuating environments. As it has been discussed, cell adhesion is another key ingredient for the origins of MC, as well as the influence of the environment triggering certain responses. These ingredients were studied in a spatial model in which cells exposed to diffusing resources and waste (entailing cell death) were able to switch back and forth between two given phenotypes; one of them able to detoxify the environment as a tradeoff with reproductive success [Duran-Nebreda et al., 2016]. The results of the evolution of parameters such as cell adhesion strengths, switching probabilities or the ability to detoxify waste led to the formation of *proto-organisms* incorporating both species while displaying nested substructures and the creation of an internal environment with lower levels of waste.

It is shown that waste degradation is in fact a public good (check Table 1.2) in the system, prone to be exploited by cheaters – individuals not investing reproductive costs to this cooperative trait – (which do occur). It is found that the evolution of cell adhesion properties ensures a minimal impact of cheaters on the fitness of cooperators by shaping the local genetic relatedness, hence a conflict-mediation strategy evolved in this system.

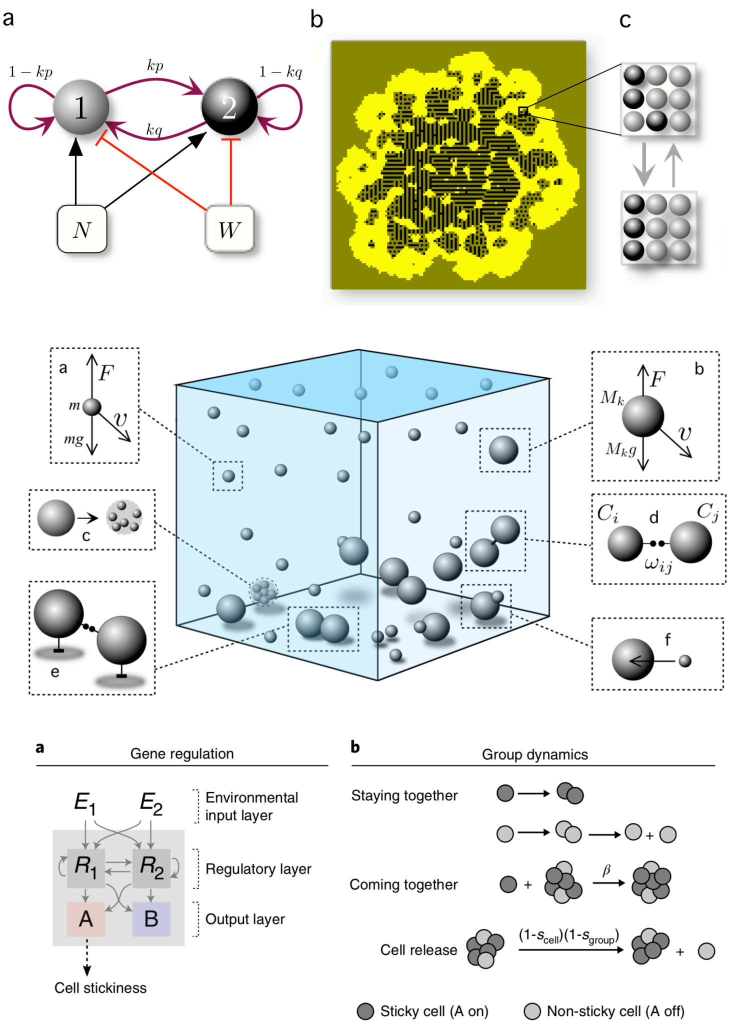


Figure 1.19: Examples of *insilico* studies of multicellularity involving embodied cells. (a) evolved *proto-organisms* on a 2D spatial lattice with metabolic trade-offs, stochastic phenotypic switching and cell-cell differential adhesion [Duran-Nebreda et al., 2016]. (b) the Chimera model, which implements a 3D simulation approach with physical forces, adhesion, metabolic trade-offs and a localized source of diverse resources [Solé and Valverde, 2013a]. In (c) the basic rules from [Staps et al., 2019] for evolved life cycles are displayed.

**Engineered MC**

Synthetic biology offers a unique opportunity for testing theories concerning the origin of MC [Solé, 2016c] as well as gaining insights into particular problems that poses its origins. A particularly interesting study is one in which the cooperation of two strains is engineered enabling the study of their dynamics in space, as well as evaluating the influence of a third parasitic species [Amor et al., 2017], concluding that different contexts allow mutualism, competition or parasitism to succeed or even transition from one to the other in a spatially extended context, emphasizing again that environmental conditions are a key driver for the emergence of social interactions between previously unrelated entities.

The engineering of a prey-predator system has also been achieved [Balagaddé et al., 2008] and even the design of synthetic ecosystems based on airborne inter- and intrakingdom communication [Weber et al., 2007], which allow to mimic fundamental coexistence patterns in nature, including symbiosis, parasitism, or predator-prey interactions.

### 1.2.4 Complex multicellularity I

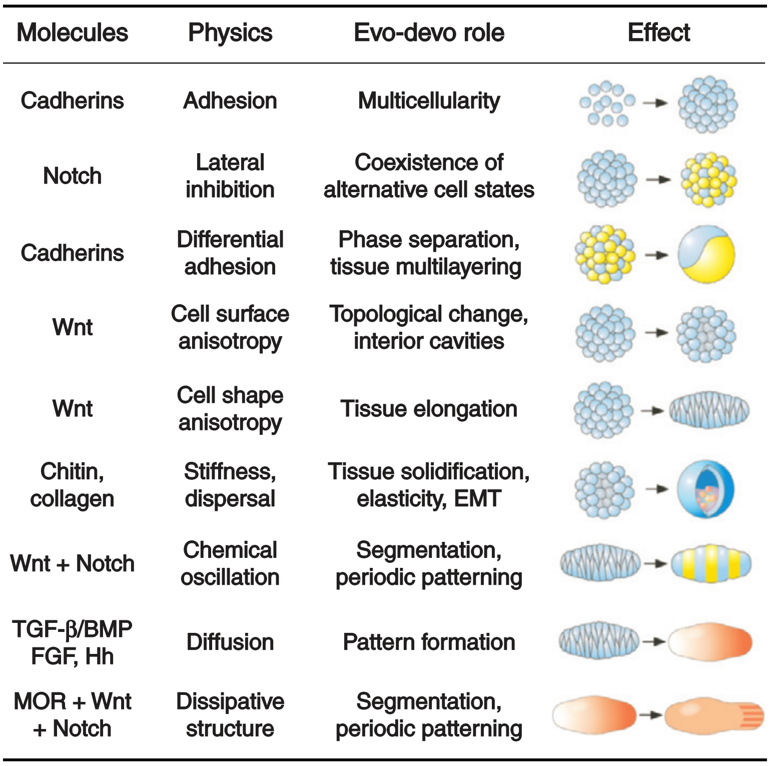
As already mentioned in Section 1.2.1 , when defined simply as cellular aggregation, a conservative estimate is that MC evolved over 25 times [Bonner, 2000, Niklas and Newman, 2013, Grosberg and Strathmann, 2007]. However, complex MC organisms are only found in eukaryotic lineages, including Metazoa, embryophytic land plants, some red and brown algae and in two groups of fungi [Knoll, 2011] (see Fig. 1.15). What do we understand by complex MC? As it has been introduced at the beginning of Section 1.1, the definition of biological complexity remains elusive albeit significant contributions for a measure have been made. However, the distinction between ’simple’ and ’complex’ MC’ is generally accepted [Knoll, 2011, Brunet and King, 2017, Niklas and Newman, 2013]. Simple MC includes filaments, balls or sheets of cells; and while differentiation of somatic and reproductive cells is common, complex patterns of differentiation is not. Adhesive molecules are also observed, but communication between cells and the transfer of resources among cells is commonly limited, seemingly a feature reserved for complex forms of MC [Knoll, 2011].

One of the key features of complex MC is that it displays a three-dimensional organization with just some cells in direct contact with the environment. This entails an important problem regarding the availability of oxygen, nutrients and signaling molecules, that occurs through diffusion in simple MC forms. The success in the circumvention of the diffusion problem is considered a physiological key to the evolution of complex MC [Knoll, 2011]. All in all, complex MC organisms are the ones fulfilling a more stringent definition, requiring sustained cell-to-cell interconnection and communication [Niklas and Newman, 2013].

With these conditions fulfilled, new functional capabilities are available, contributing to the diversity and ecological prominence seen today in plants and animals: bilaterians comprise at least 99% of all animal species, whereas vascular plants make up at least 90% of all species in the streptophyte/ embryophyte clade [Knoll, 2011]. This variability is also tight to the diversity of body plans and patterning capabilities of complex MC organisms, as intercellular signalling underpins MC development [Knoll, 2011]. Therefore, when dealing with complex MC the question is no more how a transition in individuality occurred, but how a finely-regulated developmental program came about instead, giving rise to determinate forms with stable distinct phenotypes. Here, the role of non-genetic information transmission explained in Section 1.1.1 has a key role through the epigenetic mechanisms of cellular memory [Sebé-Pedrós et al., 2017].

### From 'simple' to 'complex' MC

In Section 1.2.2 it was pointed out the evidence found regarding multiple key genes for MC to already exist in UC ancestors and also the presence of cell differentiation triggered by environmental conditions. As mentioned, the main question is how temporally-regulated differentiation came to be regulated in time and space and, ultimately, regulated by a gene regulatory network rather than by environmental triggers (in other words, the emergence of a developmental program).



| Molecules | Physics | Evo-devo role | Effect |
|---|---|---|---|
| Cadherins | Adhesion | Multicellularity | |
| Notch | Lateral inhibition | Coexistence of alternative cell states | |
| Cadherins | Differential adhesion | Phase separation, tissue multilayering | |
| Wnt | Cell surface anisotropy | Topological change, interior cavities | |
| Wnt | Cell shape anisotropy | Tissue elongation | |
| Chitin, collagen | Stiffness, dispersal | Tissue solidification, elasticity, EMT | |
| Wnt + Notch | Chemical oscillation | Segmentation, periodic patterning | |
| TGF-β/BMP FGF, Hh | Diffusion | Pattern formation | |
| MOR + Wnt + Notch | Dissipative structure | Segmentation, periodic patterning | |

Figure 1.20: The Dynamical Patterning Modules (DPMs) are displayed. The combined effect of certain molecules and physical rooted phenomena can entail certain evodevo roles, which ultimately cause effects related to biological forms emergence. Adapted from [Newman and Bhat, 2008].

An hypothesis proposed by Stuart A. Newman and colleagues is rooted in the so-called dynamical patterning modules (DPMs), a term coined by the authors [Newman et al., 2003, Newman and Bhat, 2008, Newman et al., 2009, Newman and Bhat, 2009, Newman, 2012]. They propose that sets of molecules already present in the UC ancestor mobilized a series of physical effects so as to generate an aspect or alteration in the cluster's form or pattern. ’Form’ relates to shape, size and topology, while ’pattern’ refers to particular cell type arrangements. The importance of their hypothesis is that MC form and patterns did not need to rely on complex regulatory networks in the early stages of MC but on physical and non-programmed determinants instead.

In other words, they suggest that the ’DPM-molecules’ evolved in single-celled organisms prior to the evolution of MC organisms and only took on their DPM-associated roles with the change of spatial scale that was a consequence of MC. Later, this initially physical-induced changes would have been finely regulated through the interplay of regulatory networks and the spatial scale, but they were not the responsible of their origins [Newman et al., 2003, Newman and Bhat, 2008, Newman et al., 2009, Newman and Bhat, 2009, Newman, 2012]. Under this framework, the mobilization of newly relevant physical effects and self-organizing dynamics are thought to have provided the early basis for spatiotemporal regulation in multicellular organisms [Niklas and Newman, 2013]. If true, this hypothesis might explain the origins of complex MC forms.

On the other hand, the environmental context of animal diversification deserves further attention as a possible trigger for complex forms. As explained in Section 1.2.2, the rise of oxygen is hypothesized to be a precondition for MC, not only of simple MC forms but specially for complex ones, with larger size and higher energetic requirements. Oxygen itself would not have caused animals to evolve. Rather, it would have removed an environmental barrier to the evolution of large, metabolically active animals, creating ecological opportunities that were accompanied by the evolution of morphological innovations. This hypothesis is aimed to explain the fossil record of the Proterozoic Eon, in which large animals were first found [Knoll and Carroll, 1999, Canfield et al., 2007].

Finally, returning to the perspective given by the transitions in individuality, it is worth to mention Leo W. Buss ideas regarding ontogeny. He argued that the explicitly hierarchical perspective on evolution predicts that the myriad complexities of ontogeny, cell biology, and molecular genetics are ultimately penetratable in the context of an interplay of synergisms and conflicts between different units of selection (that is, both at the level of individual units and the level of the new-formed unit) [Buss, 1987].

### How do we study complex MC?

The complexity of the developmental process giving rise to complex MC structures is still an unresolved puzzle, despite the paramount advances of the scientific community. Developmental biology started as a descriptive science but progressively became more experimental. Karl Ernst von Baer, in the 1800s, was the first descriptive embryologists to see and describe the mammalian egg as well as the fundamental germ layers of embryos. Later in the century, the experimental approach took hold, using amphibian embryos but also vertebrates as the chick and many invertebrates, with a special focus on the sea urchin [Bonner, 2000].

Remarkably, there is one aspect of experimental developmental biology that has an ancient history: the study of regeneration. In the 18th century, Abraham Trembley, in a wonderful treatise (1744), showed that the freshwater polyp Hydra had quite extraordinary powers of regeneration [Bonner, 2000]: when a polyp is cut into small pieces, each piece except for tentacles and basal disk can regenerate into a whole hydra within 2 to 3 days, while also reaggregates from dissociated single cells can also regenerate [Fujisawa, 2003]. Hydra, and other hydroids (phylum Cnidaria), have made important contributions to the study of development since then [Bonner, 2000].

Since Trembley's time, the experimental study of regeneration has continued to be a field that paralleled embryology. One aspect that received special attention over the years has been limb regeneration –of arthropods, newts, salamanders, and other vertebrates. The term developmental biology gradually came to supplant the more restricted embryology term with the advent of the experiments, focused on understanding the sequential steps of development and its causal links [Bonner, 2000].

The advent of molecular biology supposed a change of paradigm: a myriad of morphogens started to be identified together with its signaling role in the sequential steps of development [Bonner, 2000]. However, Bonner argues that this reductionist approach endangers a proper under-

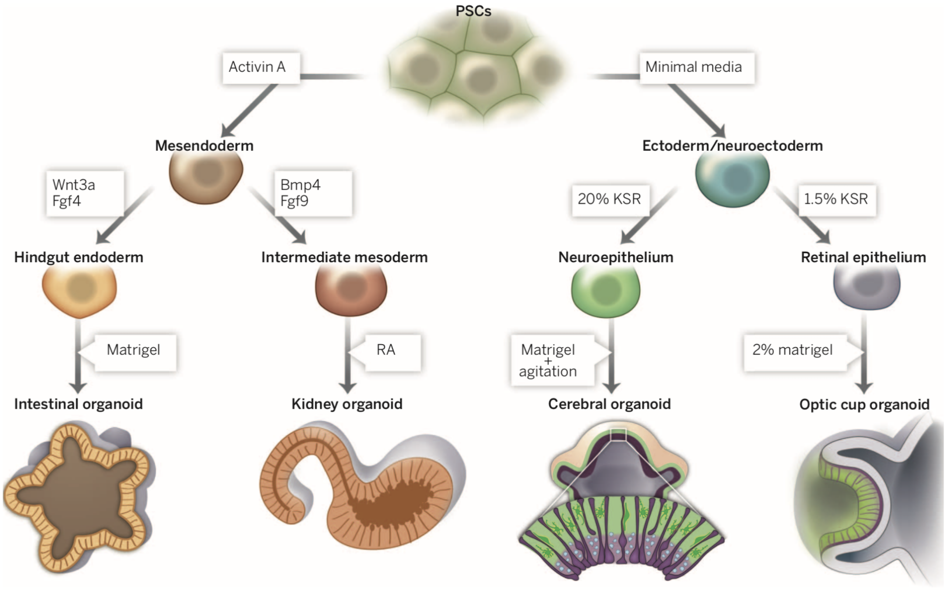


Figure 1.21: A major avenue to explore complex multicellularity is provided by organoid technology, particularly well developed within the context of human iPSCs. Under diverse sets of conditions and growth factors, different classes of organoids can be obtained. Sometimes, physical constraints (such as boundary shapes of the presence of flows) are needed for a reliable development. Adapted from [Lancaster and Knoblich, 2014].

standing of development [Bonner, 2000]. As it was pointed out in Section 1.1.5, it is not only the gene regulatory networks which are responsible of morphogensis, but rather the interplay between them and the spatial scale of the emerging structure [Alberch, 1989]. The self-organization properties in morphogenesis are widely recognized [Sasai, 2013, Bozorgmehr, 2014], and it is of paramount importance to understand them to gain proper insights in developmental biology.

The application of mathematics to the problem of development is a success story in the chase for general laws driving development. Alan

Turing was the pioneer of this field, showing that developing pattern in organisms could be described by reaction-diffusion equations [Turing, 1990]. Since then, several theoretical efforts have been made in the understanding of many developmental mechanisms, with special focus on pattern formation [Wolpert, 1969, Wolpert, 1971, Goodwin and Trainor, 1985, Cummings, 1990, Koch and Meinhardt, 1994, Cummings, 2003], the implications of differential cell adhesion [Foty and Steinberg, 2005, Cummings, 1996] –which has a critical role on tissue architecture and morphogenesis [Gumbiner, 1996] –, as well as the interplay between morphogens and geometry [Cummings, 2001]. The study of homeostasis (the capacity to maintain a given system state in the front of perturbations –related to the concept of robustness presented in Section 1.1.6 –) has also caught attention, related to the regenerative capabilities of tissues [Basanta et al., 2008, Gerlee et al., 2011], as well as the evolution of morphogenesis mechanisms [Hogeweg, 2000a, Hogeweg, 2000b].

As Bonner acknowledges, mathematical models tell us that simple explanations are possible and therefore serve as a guide, a beacon for our search. As he puts it,

> *'the models are a perfect example of how an overarching explanation can gather together and organize all the pieces of the micro–explanations'* [Bonner, 2000].

Currently, the study of development in general and that of particular organs is being boosted through the study of organoids (see Fig. 1.21), which are obtained through a mix of progenitor, ES or iPS cell cultures in particular 3D contexts and biochemical conditions. 3D structures resembling tissues and real organs are obtained, and some even display some of the characteristic functions of their natural counterparts [Willyard, 2015]. Importantly, bioengineering strategies can be used to steer the cell composition and the 3D organization within stem-cell based organoids [Yin et al., 2016].

Furthermore, synthetic biology has proved to open a new door in the study of development building up systems that display morphogenetic behaviours [Basu et al., 2005, Liu et al., 2011, Miller et al., 2012, Chau et al., 2012, Park et al., 2014, Cachat et al., 2014, Cachat et al., 2016], allowing to get insights in the underlying possible paths to generate it. Synthetic biology can be used to interrogate the design principles of systems of different nature [Bashor et al., 2010], so the exploration of alternative synthetic paths and novel functionality might be readily available.

Finally, it is important to mention the success story of the study of development together with evolution. The emerging field in the 1980s is the so-called evolutionary developmental biology (evodevo) [Müller, 2007, Carroll, 2008].

### 1.2.5 Complex multicellularity II

The building of the two morphospaces presented in this thesis required including two main unusual dimensions: ecology and cognition. The later is a specially ignored one in most modelling approaches to the evolution of biological complexity. Modelling developmental paths and including embodiment in an ecological context reveals how spatially explicit rules led to unexpected outcomes. But what about cognition as part of the evolution of multicellular systems? What kind of minimal model would be relevant and what basic components to be incorporated?

In Section 1.1.1, we presented the idea of considering the emergence of the nervous systems as a MET [Jablonka and Lamb, 2006]. The nervous system is a key differential feature in Metazoa, showing a wide variety of organizational forms: from diffuse nerve nets with a radial organization found in Cnidaria, Cthenophores and Echinoderms to central collections of neuron cell bodies linked to the periphery by nerves [Moffet, 2012, Moroz, 2009].

Nervous systems are found in all Metazoa except Porifera and Placozoa [Bucher and Anderson, 2015, Moroz, 2009]. Remarkable facts are that some echinoderms have its nervous systems origins either in the ectoderm, mesoderm or endoderm layers [Moroz, 2009], while the cephalopods (invertebrates from the phylum Mollusca) possess the most complex nervous system of invertebrates, with features which are rather unusual or even unique among invertebrates strongly resembling equiva-

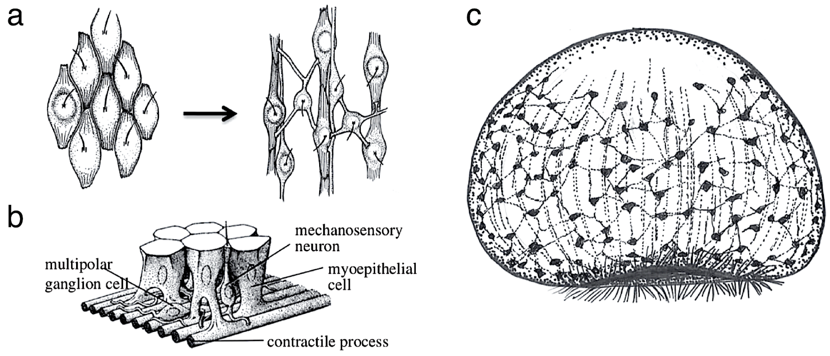


Figure 1.22: Evolution of the nerve net. (a) Evolution of a sensory-contractile network of neurons and muscle cells by division of labour. Here individual sensory-motor neurons innervate muscle cells. (b) Cell types of the cnidarian (primitive) nerve net. (c) The neuromuscular orthogon, a very simple neural network with true sensory, inter- and motor neurons. These minimal neural agents provide a baseline for modelling the emergence of biological computation in multicellular systems. Adapted from [Arendt et al., 2015].

lent parts of the vertebrates nervous system [Budelmann, 1995] (an even with brain size relative to body size exceeding that of many fish and reptiles [Morris, 2003]) and displaying highly complex behavior [Budelmann, 1995, Godfrey-Smith, 2016]. Al these are minimal multicellular systems that provide inspiration for modelling evolution of cognition. Here in particular we will keep the basic feed-forward structure that was likely to be in place when the first nervous systems.

### Origins of the nervous system

The evolutionary origins of the nervous system and the selective pressures behind it are still object of debate [Bucher and Anderson, 2015, Moroz, 2009, Budd, 2015, Wray, 2015, Liebeskind et al., 2016]. Most of the debate is focused on whether it originated in a single ancestral lineage or if it arised multiple times in evolution [Tanaka and Ferretti, 2009, Moroz, 2009], although recent works support the latter option [Moroz et al., 2014].

It is hypothesized that the origins of the nervous system –or more specialized ones– must lay in the era between late Ediacaran and the Cambric period [Budd, 2015, Wray, 2015]. We have already pointed out in previous sections that these were periods of ecological change (accompanied by ecological opportunities), which possibly drove the so-called Cambrian explosion and the appearance of macropredators [Wray, 2015]. Nervous systems could have been the result of the evolutionary pressures that predator-prey relationships posed [Budd, 2015, Wray, 2015].

In Section 1.1.6 we have presented the robustness concept. The nervous system entails complex functionality and must deal with perturbations (e.g. as a result of aging or direct damage)and still maintain its function, being a clear case where robustness mechanisms must be playing a critical role. A possible strategy to maintain the nervous system robustness is through its regenerative capacities, which present a distribution that lacks any straightforward explanation. Replacement of all or part of the nervous system has been documented in a few invertebrate phyla, including coelenterates, flatworms, annelids, gastropods and tunicates [Moffet, 2012]. The regenerating capacities of Hydra, already mentioned in Section 1.2.4, also entail the complete regeneration of its nervous system. Regarding vertebrate species, all of them can produce new neurons postnatally in discrete regions of their nervous system, but only some lower vertebrates (fish and amphibians) can significantly repair several neural structures. Some regenerative ability, however, is found also in reptiles and birds, and even in mammals[Ferretti, 2011]. Probably, ecological factors might have been the drivers of the evolution of regeneration capabilities in nervous systems [Bely and Nyberg, 2010, Kaslin et al., 2007]. Other relevant strategies to provide robustness to the nervous system are probably in the lines of the ones outlined in Section 1.1.6, e.g. redundancy, degeneracy or modularity as a means to maintain fault-tolerant systems.

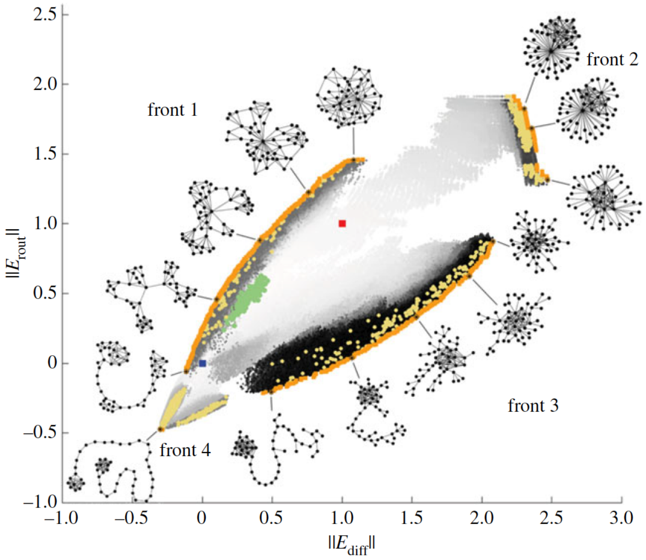


Figure 1.23: Communication-efficiency morphospace. Every point represents a network generated by the optimization algorithm (from Avena-Koenigsberger et al 2014).

### Evolving reliable computation in neural circuits

Which are the evolutionary drivers shaping neural circuitry? Which role does robustness have? Theoretical efforts have been directed to this direction, with special focus on optimization cues. The object of study which attracts more attention is the human brain. Graph theory is vastly used to analyze the properties of brain neural networks [Bullmore and Sporns, 2009], while efforts in deciphering the driving forces behind neural circuitry topology suggests wiring optimization to be a clear candidate [Chklovskii et al., 2002, Chen et al., 2006]. Pareto-optimization has also been used, as explained in Section 1.1.4, with special focus on the trade-off between communication efficiency and the underlying costs [Avena-Koenigsberger et al., 2014, Goñi et al., 2013]. Approaches like these allow to get deeper insights on network topologies emerging under determinate selection pressures, which are putative to have shaped the network archictecture of human brains. Importantly, some of the geometric, biophysical, and energy constraints that have governed the evolution of cortical networks are also being envisaged [Laughlin and Sejnowski, 2003]. In Fig. 1.23 we show an example of a morphospace of communication efficiency, an approach used to unravel evolutionary pressure behind network topologies.

What are the main drivers shaping the complexity of the first cognitive networks? As complex living forms emerged around the time of the Cambrian explosion, cognition and the sensor-actuator networks. Although behavioral patterns were already present before the advent of neurons [Jékely, 2010], the appearance and rapid expansion of neurons and their interaction webs created the basis for true behavior. Information became an essential part of how complex organisms adapted to new conditions within their life spans, allowing much faster and more flexible responses. At some point, the ”first” brains were formed [Pagán, 2014] and became the first prediction machines [Llinás, 2001]. In this context, an important trigger of the development of neural prediction networks was based on the need of exploring a spatial world. The *moving hypothesis* proposed that exploration of an organism’s spatial environment was crucial in the evolutionary trajectory that produced brains [Llinás, 1987]. Under this viewpoint, prediction based on neural networks is central to the understanding of early evolution of MC cognition [Miguel-Tomé, 2018].

Evolving neural networks has been an important part of the studies of artificial systems. A natural approach to the problems raised above would be to explore a simple description of a neural network capable of (a) performing predefined computations i. e. ”predicting” a given set of outputs associated to a given set of inputs, (b) do the computation in a reliable way, i. e. being robust against noise and (c) incorporating mechanisms of repair, in particular the regeneration of lost connections among cells. Because regeneration is a crucial component of developmental robustness, including it in the evolution of neural complexity is a natural step. More importantly, it also allows us to consider another largely ignored element of MC modelling: the existence of a well-defined lifespan of organisms.

In this dissertation, a minimal model of computation in a feed-forward neural network experiencing damage and capable of evolving regeneration is presented. This model provides a baseline for exploring the relative role played by the interplay between cost, connectivity and reliability.

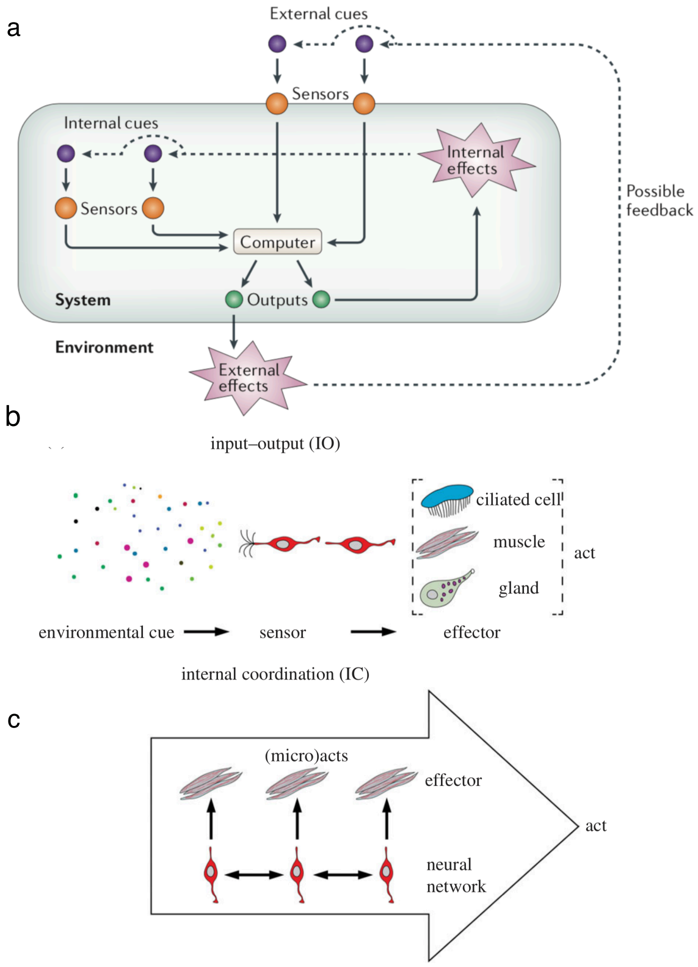


Figure 1.24: Cognitive substrates for evolution. In order to evolve cognitive traits the standard approach is using a simple input-output scheme as those based on the early cybernetics picture, where sensors, comparators and actuators are considered. This translates into a simplified scheme that incorporates sensing units, intermediate processing units and response signals. In (a) this scheme is shown for single cells (adapted from [Benenson, 2012]). In (b) a simplified diagram of a feed-forward cellular network (equipped with early cell types capable of sensing and propagating signals) is shown (adapted from [Jékely et al., 2015]).

# Chapter 2

# OBJECTIVES

- Define a design space aimed at guiding the study of the origins of simple (and complex) multicellularity through a synthetic approach
- Define a design space aimed at guiding the study of complex forms of multicellularity: synthetic organs and organoids
- Provide new insights on the possible and the actual forms of multicellularity through the defined design spaces
- Provide new insights on the preconditions of simple forms of multicellularity through a particular case study
- Study the potential evolutionary tradeoffs and selective pressures regarding the robustness of information processing in neural-based agents

# Chapter 3

# RESULTS

## 3.1 A design space for synthetic multicellularity

This work is aimed to thoroughly review what we define *synthetic* multicellularity. That is, any means to achieve multicellularity through a *synthetic* approach, including artificial life models (*in silico*), artificial evolution experiments of natural organisms or synthetically modified organisms. Taking the state-of-the-art on the topic, we propose a morphospace to study *synthetic* multicellularity defined through three (ideally orthogonal) dimensions: [1.] Evolution, [2.] Development and [3.] Ecology. In a nutshell, we propose an *evo-devo-eco* design space which allows us to disentangle the possible and the actual in terms of the study of the origins of MC using a synthetic approach, as well as the possible existent constraints. The proposed design space is aimed to provide an integrated view to guide future research on the field. *In silico* approaches prove to be the most effective in exploring the landscape of possibilities, while experimentally evolved MC would be in a second position in this regard. Meanwhile, the engineered systems occupy a narrow position in the morphospace, only exploring the ecological axis. Interestingly, none of the entities in the design space occupies the region where the three dimensions were to be maximized, pointing towards constraints (still) not by-passed by current *synthetic* approaches. In the light of the myriad of open questions regarding the origins of simple MC (and its transition to more complex forms), we could speculate that an eventual reach of such unexplored morphospace area would take us closer to answering them. We suggest that the synthetic path to MC is the fourth wave of research required to fully understand the origins of MC, along with the study of conserved toolkits, the definition of modules responsible for morphological diversity and the experimental testing of predictions drawn from the two previous waves.

Solé, R., Ollé-Vila, A., Vidiella, B., Duran-Nebreda, S. and Conde-Pueyo, N. (2018). The road to synthetic multicellularity. Current Opinion in Systems Biology, 7, 60-67. DOI: 10.1016/J.COISB.2017.11.007

## 3.2 A design space for synthetic organs and organoids

Organs and organoids are well known scenarios to explore the structural and evolutionary constraints associated to a specific level of organisation below the organismal level. This work attempts to fulfill two goals. First, to systematically review and unify the fields of synthetic biology and tissue engineering, which are to be the key tools to explore and understand the design principles of organs and organoids Second, tit seeks to provide a theoretical framework through the basis of a morphospace in which existent organs and organoids, but also totally different complex biological systems (such ant colonies or microbiomes) can be qualitatively located on it (in relative terms). It is suggested that a good approximation for a morphospace of organs and organoids should contain three basic properties: [1.] Developmental complexity, [2.] Physical state and [3.] Cognitive complexity. While the first axis is a well known dimension connected to developmental programs, the other two are novel ingredients in the analysis of multicellular complexity. They provide a physics-level approach to the ”state of matter” associated to these systems (with a potentially quantifiable axis from liquid to solid) and a third dimension that incorporates the associated information processing levels. When positioning both real and synthetic systems on the morphospace, the observation is that a large volume located between the liquid-solid limits is apparently empty. This conveys interesting questions: what is the reason to find such empty space? It is due to fundamental physical reasons? Or, alternatively, evolution has not been able to access this part of a morphospace? In this work, we provide some insights regarding the possible directions to be undertaken so that this morphospace could be further explored through merging synthetic biology and tissue engineering.

Ollé-Vila, A., Duran-Nebreda, S., Conde-Pueyo, N., Montá ñez, R.,and Solé, R. (2016). A morphospace for synthetic organs and organoids: the possible and the actual. Integrative Biology, 8(4), 485-503. DOI: 10.1039/C5IB00324E

## 3.3 On the preconditions for multicellularity: cooperation and metabolic trade-offs

This work explores a set of preconditions that plausibly predated the origins of very simple forms of multicellularity. A previous spatial model including both resources and a toxic waste causing direct mortality to cells showed that *proto-organisms* displaying non-trivial internal structure were formed. Stochastic phenotypic switching and differential adhesion were two key ingredients of this model. Because of the complex and predictable organisation in nontrivial structural patterns, these systems were dubbed ”protoorganisms” and a space of possible resource-toxic levels was found to allow for protoorganisms to emerge. A key question beyond the structural complexity emerging in space, was how a population of cells reacting in different ways to resources and waste can actually persist in time. This work aims to explore the most fundamental features of the proto-organismal model through a deterministic mathematical approach together with a discrete stochastic implementation. We found and characterized mathematically the conditions needed for the stability of the heterogeneous system, namely: [1.] asymmetry in toxic sensitivities of the species and [2.] a nonlinear positive correlation between the parameters regulating the efficiency of an existing cooperative link. Our system is characterized in terms of the social interaction between the species, finding phases of both altruism and mutual benefit regarding the present cooperative link. This allows us to speculate on the putative evolutionary stability of such interaction, which would lead to the formation of simple forms of MC.

Ollé-Vila A and Solé R. (2019). Cellular heterogeneity results from indirect effects under metabolic trade-offs. Royal Society open science, 6(17), 190281. DOI: 10.1098/rsos.190281

## 3.4 On the evolution of neural agents

The exploration of the cognitive axis described in the previous papers (and the introduction) can be explored in several ways, in particular in connection with developmental and regeneration processes. Despite the previous work on evolved neural networks, no previous study considered the impact of regeneration processes and it connection with aging. This work aims to shed light into the possible reasons behind the broad levels of regeneration capacity observed in nature, while providing a very general framework capable of exploring the effects of the life span length as well as the aging process endured in any living system (inflicted through direct damage on the network connections). Digital organisms are represented by feed-forward networks evolved to reliably perform given computations under persistent damage. They were shown to attain network robustness through determinate connectivity patterns entailing higher density of connections or, alternatively, through higher levels of regeneration. Using a multi-objective optimization approach (MOO) we are able to explore the tradeoffs intrinsic of such an evolutionary framework. The effects of the length of the life span as well as the damage inflicted to networks throughout an aging process is quite interesting: both features exert a clear evolutionary pressure on the networks, finding a more diverse repertoire of strategies. Importantly, the severity of the aging process widely affects the region of the morphospace that can be explored due to the strong imposed constraints. The nature of MOO approaches allow to retrieve a solution linking the diversity of optimal strategies in a mathematical object, the so-called Pareto optimal front, whose geometry has been linked to phase transitions and criticality. We show how transitioning between damage regimes (severity of the aging process) cause an overall shape change in the geometry of the Pareto-optimal front, which gives us some insights about how accessible the range of optimal solutions are. The implications of such observations are discussed.

Ollé-Vila A, Seoane, LF, and Solé R. (2019). Aging, computation and the evolution of neural regeneration processes. bioRxiv, 780163. DOI: 10.1101/780163

# Chapter 4

# GLOBAL DISCUSSION

Biological complexity has experienced profound changes throughout the course of evolution. A way to study these innovations in a unified fashion is provided by the so called Major Evolutionary Transitions (METs). These qualitative shifts in living organisation involved new forms of reproduction associated to novel forms of complexity. These transitions are associated to changes in the way information is stored, processed and transmitted, acquiring higher complexity at each transition, while some of them also convey developmental changes. The nature of these transitions has been often approached with a reductionist perspective, where a gene-central picture has been taken.

This thesis focuses in the emergence and evolution, but also the engineering, of multicellular complexity under a complex systems approach. It is thus connected with one of the most important evolutionary events of evolution, namely the rise of multicellular life forms. Multicellularity (MC) effectively incorporates a new level of organisation where single celled entities get together to form more costly, but better adapted assemblies where division of labor helped to build increasingly higher levels of complexity. A focus on the preconditions for the emergence of MC structures is proved to be key in unraveling the very first steps in this transition in individuality, with ecological factors playing a decisive role. However, it is less appreciated in the existing literature the fact that embodiment has been an essential part in creating complexity. On the other hand, and also considered here as part of our integrative view, MC is a multi-scale feature of biology. Tissues and organs define sub-scales of internal complexity within organisms. What are the intrinsic limits and design principles? In this context, the rise of organoid technology has been informative in creating a novel opportunity of interrogating biology at a higher scale beyond cells. Similarly, synthetic biology has provided new opportunities to test the design limits of biology.

A global picture has been developed that is grounded in an expansion of the concept of qualitative morphospace. Qualitative morphospaces are a unique way of building a global picture of the possible and the actual by properly locating very diverse systems in a spatially constrained volume characterised by meaningful axes that are aimed at capturing the relevant dimensions of complexity. Here, the construction and analysis of these spaces has been done with a very ambitious perspective that goes beyond the standard developmental or cellular view tied to the organismal level. Instead, it has moved towards a whole variety of systems that, for different reasons, exhibit dynamical states and structural principles that connect them with true multicellular entities.

Two MC morphospaces have been proposed and a dedicated effort has been made towards locating a very diverse array of known case studies to fill them. The first morphospace (see **Result 3.1**) was proposed to define the boundaries and constraints associated to synthetic models of MC organisation. These *synthetic* examples include those designs obtained from genetic engineering, artificial experimental evolution as well as models of MC based on artificial life approximations (*in silico*). Along the consideration of *synthetic* approaches, the morphospace has been constructed by including one very important novelty (as compared with the state of the art of developmental biology and *evodevo* studies). We include ecology as the third axis along with two axes that capture the relative role played by development (to what extent a developmental program is at work) and evolution (to what extent the final design results from an evolutionary process).

Our suggestion is that the following variables: 'Ecology', 'Evolution' and 'Development', taken together, have been critical to define the origins of both simple and complex MC forms. Therefore, a design space built upon these three variables is presented, which helps at providing insights on what are the relative contributions of each variable in the study of the origins of MC. Beyond the *synthetic* approaches included in the morphospace, we also include some natural examples in the ecodevo wall – as we only consider 'nonzero' evolution for those systems which have undergone evolutionary changes in the experimental timescale.

The space of the possible is filled with varying contributions of the different systems included, with *in silico* approaches being the most effective in the exploration of this space: from the unfolding of determinate developmental programs through a morphogenetic engineering approach (but lacking ecology while displaying low degree of evolution), to an opposite extreme, occupied by the Chimera model, which can be considered to display a high degree of both evolution and ecology, with a more reduced degree of Development. This is due to the clear innovation of the entities in the model that through ecosystem engineering give rise to a new niche, which is ultimately filled with a new evolved species. This example shows the power of embodiment as a prerequisite for innovation. The Chimera model is close to other remarkable examples, the well-known AVIDA [Ray, 1991] and ECHO [Holland, 1994] models, which allow to explore high degrees of the ecological and evolution dimension, while lacking any means to increase in developmental complexity.

Interestingly, experimentally evolved MC using bacteria or simple eukaryotes as a test bed are also successful in the exploration of the morphospace, with the experiments evolving i) bet-hedging [Beaumont et al., 2009] and ii) life cycles (through the embracement of cheaters) [Hammerschmidt et al., 2014] as regulated strategies being the most successful in exploring a considerable degree of the three axis. Finally, synthetic biology-related examples occupy a narrow space, mainly exploring the ecological axis, but here a sampling issue is at place. Other examples not included in the morphospace also falling in this discipline, regarding synthetic pattern formation [Cachat et al., 2016, Basu et al., 2005, Schaerli et al., 2014], would be considered to have some degree of development.

Despite major successes have been achieved by the genetic approach based on the definition of early toolkits and phylogenomic analyses (which also help in unveiling the hypothetical preconditions for MC) and an increase understanding of some developmental processes has been obtained by means of the analysis of genetic modules, we propose that the *synthetic* path to MC will be essential to the future developments of MC research. A whole picture requiring embodiment necessarily implies using synthetic biology, experimental evolution or *in silico* surrogates. Besides, the synergy among the approaches embedded in the morphospace is likely to boost the advances in the field. Remarkably, the morphospace obtained in our study, in spite of its qualitative nature, suggests that the most relevant understanding of the evolutionary dynamics of MC will be achieved in the currently empty corner of this space of the possible. In the paper we discuss a possible roadmap to reach that empty space. Therefore, in this contribution we provide an integrative picture which is aimed at providing interesting avenues on the research agenda of the next years.

The second study on MC morphospaces was instead focused in the structural organisation of multicellular complexity along three axis that incorporate both physical features and functional traits associated to cognition (see **Result 3.2**). The latter is a very important axis when dealing with both the origins of neural agents as well as with the sub-organismal structures that also perform decisions and adaptive control and that are easily identified as modular structures. The other two axes include again the role played by developmental processes (degree of unfolding of a given developmental program)for a physics picture of MC (as defined by the growing field of active matter) but seldom considered relevant within developmental biology: the actual, physical state of these entities. That means to explicitly consider the ordering of different candidate systems within a liquid-solid axis.

The paper makes a systematic effort of including several components required to characterise, create or model multicellular systems, with particular attention to tissue and organ engineering as well as to the minimal models of cell sorting and differentiation described by discrete automata. Organs and organoids appear organised along one face of the morphospace characterised by a ”solid” phase, meaning that a spatial coherence is at work, with cells and cell layers remaining organised in a stable fashion. Since known organs perform different kinds of functionalities by means of gathering information (about the organism state) and responding to it, they occupy diverse locations along the cognitive axis and at the top on the developmental range (all have been created by means of whole developmental processes). The brain would be the limit case with large cognitive/computational complexity. Organoids would occupy lower levels within the developmental dimension and (so far) displaying little or moderate computational potential.

Once we move towards the ”liquid” face of the morphospace many interesting observation can be made. Some limit cases are, here too, easy to identify. The immune system is also the result of a developmental hierarchy of events and exhibits a high computational complexity. It would thus be located at the opposite side of the brain vertex. This particular finding is a good reminder that cognitive phenomena can take place within very different physical phases. Red blood cells instead would be a good example of a ”liquid” cellular population with very simplified organisation and rather limited in terms of computational decisions. Between these two faces, many interesting things can happen, but it turns that (as far as we know) little is there that truly belongs to the organ/organoid context. Intermediate forms of tissue/ organ complexity are not easily found, and as a consequence a large volume of space appears essentially empty. Is it really?

Looking elsewhere in natural systems, we can provide an interesting insight to answer this question. Systems like ant colonies, which (as discussed in the paper) experience a developmental process including growth and differentiation (both in terms of nest complexity and task/caste allocation) appear to be in a middle domain: nests are ”solid” while the colony itself is very fluid, with individuals moving around while they perform their functions and maintain the nest in a dynamical way. The microbiome, on the other hand, can also be seen as an intermediate state since miroorganisms do not remain in stable positions indefinitely and are dependent on (and can help create) a spatial context. Beyond these examples the void observed that dominates the center of the morphospace asks for a theoretical understanding. In the paper, several interpretations and implications are discussed. A very important consequence of this work and the implications of the MC void was the development within the Complex Systems Lab of a new research area labelled "Liquid Brains" which triggered the organisation of a workshop at the Santa Fe Institute (New Mexico) on that new topic and recently led to a theme Issue in Philosophical Transactions of the Royal Society B: *'Liquid brains, solid brains: How distributed cognitive architectures process information'*, compiled and edited by Ricard Solé, Melanie Moses and Stephanie Forrest. Therefore, our work has already led to new research avenues, which suggest that there might be novel, and functionally very interesting instances of engineered designs that could help fill the empty region observed in the proposed morphospace.

These two morphospaces rise many important questions regarding the multiverse of potential forms of organisation of MC systems and their synthetic counterparts, while providing clear examples that the study of biological complexity can be achieved by means of design spaces of this nature. In this thesis, beyond the integrative approaches discussed above, two specific, but we believe that relevant problems, have been fully analysed, providing insights both on the preconditions of simple forms of MC as well as on the evolution of simple neural agents.

Our approach on the study of the preconditions of simple MC forms (see **Result 3.3**) is a model of heterogeneous cell populations living in an environment characterized by the presence of a common resource and a toxic field that damages their potential for reproduction through direct cell mortality. The reason for choosing this problem and exploring in detail its consequences is the previous work developed by [Duran-Nebreda et al., 2016] that developed the concept of *proto-organism* as a precursor structural organisation that could predate the emergence of more complex forms of MC. It was shown that a spatially constrained population of cells (in an *in silico* setting) displaying two phenotypes (linked through a phenotypic switch; one able to process the toxic waste in the media) and a metabolic trade-off (reduced reproduction capacity proportional to the waste degradation investment – displayed by one of the phenotypes) could result into very well organized structures with internal structures and protective layers enclosing them. The surprising complexity of these *proto-organisms* was a rather unexpected result, and a whole explanation for the origins of their persistence was lacking.

This problem has been addressed in this dissertation by means of a further simplification of the original model. The goal was not so much explaining the embodied structures emerging from the original discrete model, but finding the fundamental basis of how heterogeneity could be maintained in the first place. To this goal a minimal model was developed, containing the most fundamental features of the *protoorganisms* one, while not connecting our cell populations by any developmental means, providing a more general scenario.

Related studies had already shown that growth inhibitors can trigger cell heterogeneity in the context of antibiotic resistance [Lenski and Hattingh, 1986, Hsu and Waltman, 1992, Hsu and Waltman, 2004]. However, the nature of the *protoorganisms* (PRO) model was different from those previously studied scenarios, so taking its most fundamental features set up a previously unstudied scenario. In this work we unravel the hypothetical preconditions allowing for coexistence of the characteristic cell populations in the PRO model, namely: an asymmetry in toxic-sensitivity and a positive nonlinear correlation between the efficiency of the parameters regulating a cooperative trait. These two conditions have been characterised mathematically.

The cooperative trait is displayed by the species enduring a metabolic tradeoff between reproduction and toxic degradation (the *actor* species [West et al., 2006, West et al., 2007]), while the second species reaps benefit from this cooperative behavior (being the *recipient* [West et al., 2006, West et al., 2007]). To our knowledge, the nonlinear correlation between the parameters regulating the cooperative trait was not previously described in related models – where a similar cooperative trait was present in toxic-resistant species –, while we further show that the sensitivity to toxic of the *actor* – the toxic-degrading species– influences this domain.

A discrete stochastic model allows to confirm the basic results form the mean field as noise is introduced. Moreover, we find that species coexistence is present under local dispersal irrespective of the asymmetry in toxic-sensitivity (in determinate regions of the parameter space), being an example where the spatial dimension opens new niches for heterogeneous populations [Bascompte and Solé, 1998].

Interestingly, to assess the evolutionary stability of the presented interaction, we characterized the domains in which the cooperative trait is mutually beneficial or altruistic [West et al., 2006, West et al., 2007], showing that the model accounts for both possibilities. Therefore, we have characterized a particular case study for the origins of simple MC forms with potential to be evolutionary stable due to the mutually beneficial domain. One possible route from this primordial scenario would be the one studied in the PRO model, where the physical properties of cell-cell adhesion and a simple developmental mechanism – through a phenotypic switch – were already at play. Interestingly, in that model it is shown how the cooperative trait displayed by the toxic-degrading phenotype was in the altruistic regime. In spite of this, coexistence with the other phenotype was possible through the emergence of a conflict-mediation mechanism (green beard effect), in which toxic-degrading cells preferentially attach alike phenotypes, diminishing the fitness reap by the cheaters. Therefore, despite the a priori unlikelihood of evolutionary stability of the altruism domain, the PRO work shows how evolvable cell adhesion and phenotypic switching properties might be the sufficient conditions to stabilize such interaction. This particular route involves cell colonies mainly emerging from single cells – clonal development –, which differentiate into the two phenotypes through the phenotypic switch mechanism.

However, our model can be extended to more general cases where the phenotypic switch is absent. In this case, the route to MC would be through aggregative development. Albeit this particular origin of MC seems to be reserved to simpler MC forms (such as Bacteria, Archaea or slime molds), the recent observations that *C. owczarzaki*, a close relative of Metazoa, displays this kind of behavior [Sebé-Pedrós et al., 2017] opens the door to new hypothesis for the initial steps predating the origins of eventually more complex forms (like Metazoa). Concretely, it has been hypothesized that clonal development arising from an initially aggregative MC cannot be ruled out in explaining the origins of MC in Metazoa [Olson, 2013].

The detailed characterization of scenarios predating the origins of MC is quite interesting, as it suggests potential pathways for designing the postulated model using synthetic biology. Our results actually give a solid guideline for engineering two cell populations mapping the requirements analysed here, that could be extended to microbial populations living in bioreactors and evolving spatial organisation. Therefore, this research contributes to the synthetic MC field devised in our first contribution.

Finally, the fourth research contribution (see **Result 3.4**) to be discussed here involves the problem of how to explore the cognitive dimension previously discussed within the context of MC spaces. This cognitive dimension becomes also important in the context of the origins of the nervous system, which is also considered to be a major evolutionary transition (MET) due to the change it entailed in the way information is stored, transmitted and processed. Concretely, the 'neural individual' can attain a high level of internal integration and the ability to make rapid adaptive responses, deeply changing the landscape of selection when individuals possess a nervous system. This would have had far-reaching evolutionary consequences due to its speed and specificity together with its potential for integration and memory storage [Jablonka and Lamb, 2006].

A full model including an embodied organism, perhaps also incorporating its *in silico* development, is a major challenge that would require sophisticated computational techniques. But there are relevant questions concerning simple neural networks performing decisions that can be addressed within our context. The one studied here deals with a choice for simplicity that could allow taking into account both minimal cognitive traits and the incorporation of aging as well as regeneration as a mechanism to endure it.

A crucial part of complex multicellular systems is their potential for dealing with external stresses and damage that jeopardize a proper perception of the world and a correct response. Our choice provides a satisfactory approach to these requirements: we consider the evolution *in silico* of simple feed-forward neural networks that computes simple Boolean functions. Since biological computation, as discussed at the introduction, is necessarily plagued with perturbations of diverse nature, the model takes into account the possibility of damage and repair that takes place along the digital life of each agent. If a reliably computation, connectivity and regeneration are costly, the first question to be considered here is what kind of evolved networks are expected to be found. Are there multiple solutions compatible with a complex landscape or are there instead global optima?

The possibility of having agents of different ’ages’ came immediately given the potential for simulating a shorter or a longer lifetime. In this way, the model approach also considers a very important element of life: the age of each organism and the extent of the aging process endured by it. Despite the limits of our model, the study reveals that the evolved networks display a broad range of solutions with a marked trade off: the neural networks can either display high regeneration and exhibit a minimal number of components (close in some cases with early studied models of evolved networks [Miller et al., 1989, Angeline et al., 1994, Yao, 1999, Floreano et al., 2008, Stanley and Miikkulainen, 2002, Stanley et al., 2009]) or instead rely on densely connected webs where alternative mechanisms involving determinate connectivity patterns, e.g redundancy, allows to minimise the impact of repeated damage.

The study offered important insights regarding the evolutionary dynamics exhibited by organisms of different life spans. Longer-lived agents led to a better exploration of the optimal space of solutions (as defined by Pareto optimality) suggesting an interesting consequence of extended life spans. Increasing damage regimes show even stronger evolutionary pressures. Concretely, besides also forcing the networks to choose among Pareto-optimal designs, it severely constraints the space of the possible when considering reliable computing networks, whereas it induces more rugged shapes in the Pareto-optimal tradeoff.

The ruggedness of optimal tradeoffs can tell us something about how accessible our space of optimal solutions is. Cavities and singular points in the Pareto-optimal front have been linked to phase transitions [Seoane and Solé, 2013, Seoane and Solé, 2015a, Seoane and Solé, 2015b, Seoane and Solé, 2016, Seoane and Solé, 2018] and critical phenomena [Seoane and Solé, 2015b, Seoane and Solé, 2018]. The phase transitions observed in our evolutionary setup for increasing damage rates suggests that this evolutionary pressure might cause a discrete phenotypic space, being less accessible as well as prone to suffer drastic changes when it varies. Interestingly, the relation between organismal lifespan and that of its components is also hypothesized to be acting as an evolutionary pressure in our system, pointing towards an interesting research avenue if we wonder at what level (organismal versus component part) can a Darwinian process store the information gathered as evolution proceeds, which would have implications on the problem of Multi Level Selection [Okasha, 2005]. Albeit the studied networks do not have natural counterparts, this study could also be considered an example of a network morphospace [Avena-Koenigsberger et al., 2015] which is explored through optimization techniques.

Several potential extensions emerge from this study and relevant implications emerge when connecting computations and aging. Future work should extend this model to more complex functional tasks and embodied implementations. Once again, the full integration of development, evolution and cognition (in an ecological context) remains a future task that can be inspired by these toy models.

In summary, this thesis involves several complementary approximations to general and specific aspects of multicellular complexity grounded in an integrative view of multicellularity far beyond the more standard approximations. The more general picture emerging from the MC morphospaces suggests a roadmap for future research while opening relevant questions concerning the limits of evolution and engineering associated to empty volumes. The two case studies presented are two examples of the many open avenues still open for exploration, as well as being two means of the study of biological complexity focusing on two of the major evolutionary transitions (METs), namely the rise of MC organisms and the nervous system, which is a distinguishing feature of complex MC organisms (Metazoa). Beyond the answers obtained from them, many other questions and exciting alternative paths emerge from these contributions.

# Chapter 5

# CONCLUSIONS

*Objective 1: Define a design space aimed at guiding the study of the origins of simple (and complex) multicellularity through a synthetic approach*

- The development of a three-dimensional design space has been done, defining Evolution, Development and Ecology as its dimensions.
- This supposes a further exploration of the common acknowledged *evodevo* field with the addition of the ecological dimension. In the light of the evidence regarding the importance of this factor in triggering transitions in individuality, and specially the origins of multicellularity, this can provide a more global view on what has been explored.
- The location of experiments of different nature through *synthetic* approaches provides an integrative view on the field not previously presented.

*Objective 2: Define a design space aimed at guiding the study of complex forms of multicellularity: synthetic organs and organoids*

- The development of a three-dimensional design space has been done, defining Developmental Complexity, Physical State and Cognitive Complexity as its dimensions.

- The addition of cognitive complexity as a dimension can provide deep insights on the complexity of organismal structure performing decisions and adaptive control, while a physical state allows to include systems of very varied nature in the morphospace.

*Objective 3: Provide new insights on the possible and the actual forms of multicellularity through the defined design spaces*

- Through the accomplishment of *Objective 1*, the observation of a largely unexplored area of the morphospace is observed, namely the one where the three dimensions are maximized. The origin of this void is tight to the current constraints even imposed in the *synthetic* domain.

- We propose that the reach of this area might unravel a more profound understanding of the evolutionary dynamics of multicellularity, going beyond simple MC forms. The means to achieve it are outlined in the presented paper.

- Through the accomplishment of *Objective 2*, we uncover the presence of a large void in the defined morphospace. We propose an avenue combining synthetic biology and tissue engineering to explore such a void.

*Objective 4: Provide new insights on the preconditions of simple forms of multicellularity through a particular case study*

- We define and characterize two sufficient conditions for the maintenance of cell heterogeneity in a resource-waste environment, where both cell populations suffer from direct cell mortality due to toxic waste. Namely, an asymmetry in toxic-sensitivity and a positive nonlinear correlation between the efficiency of the parameters regulating a cooperative trait.
- A discrete stochastic model allows to confirm the basic results form the mean field. Moreover, we find that species coexistence is present under local dispersal irrespective of the asymmetry in toxic-sensitivity (in particular regions of the parameter space), being an example where the spatial dimension opens new niches for heterogeneous populations.
- The characterization of parameter space regions where the cooperative trait is mutually beneficial or altruistic allows to speculate on the evolutionary stability of the interaction.
- The characterization of such two-species interaction might be of applicability in the synthetic biology domain aiming to study scenarios predating the origins of MC.

*Objective 5: Study the potential evolutionary tradeoffs and selective pressures regarding the robustness of information processing in neural-based agents*

- We show how a very minimal model including neural agents (in the form of feed-forward neural networks) under the appropriate evolutionary setup allows to retrieve meaningful conclusions regarding hypothetical evolutionary pressures.
- Both the life span and the extend of damage received by the neural agents –during an aging process– have been shown to be strong evolutionary pressures shaping the landscape of possible network designs

- A tradeoff between connectivity and regeneration capabilities is observed as a means to achieve reliable information processing (through the computation of simple Boolean functions)
- First order phase transitions are observed in the Pareto-optimal front of the evolutionary setup for increasing damage rates, suggesting that this evolutionary pressure might cause a discrete phenotypic space, being less accessible as well as prone to suffer drastic changes when this evolutionary pressure varies.
- The relation between organismal lifespan and that of its components is also hypothesized to be acting as an evolutionary pressure in our system.
- This study is an example of a network morphospace which is explored through optimization techniques. Albeit networks have not natural counterparts, it still serves as a powerful approach to elucidate possible evolutionary pressure shaping the landscape of possible strategies to achieve reliable computation.

# Bibliography

[Abedin and King, 2010] Abedin, M. and King, N. (2010). Diverse evolutionary paths to cell adhesion. *Trends in cell biology*, 20(12):734–742.

[Acar et al., 2008] Acar, M., Mettetal, J. T., and Van Oudenaarden, A. (2008). Stochastic switching as a survival strategy in fluctuating environments. *Nature genetics*, 40(4):471.

[Ackerly, 1992] Ackerly, S. C. (1992). The structure of ontogenetic variation in the shell of pecten. *Palaeontology*, 34:847–867.

[Adami, 1998] Adami, C. (1998). *Introduction to artificial life*. Springer Science & Business Media.

[Alberch, 1982] Alberch, P. (1982). Developmental constraints in evolutionary processes. In *Evolution and development*, pages 313–332. Springer.

[Alberch, 1989] Alberch, P. (1989). The logic of monsters: evidence for internal constraint in development and evolution. *Geobios*, 22:21–57.

[Amor et al., 2017] Amor, D. R., Montañez, R., Duran-Nebreda, S., and Solé, R. (2017). Spatial dynamics of synthetic microbial mutualists and their parasites. *PLoS computational biology*, 13(8):e1005689.

[Anderson, 1972] Anderson, P. W. (1972). More is different. *Science*, 177(4047):393–396.

[Angeli et al., 2004] Angeli, D., Ferrell, J. E., and Sontag, E. D. (2004). Detection of multistability, bifurcations, and hysteresis in a large class of biological positive-feedback systems. *Proceedings of the National Academy of Sciences*, 101(7):1822–1827.

[Angeline et al., 1994] Angeline, P. J., Saunders, G. M., and Pollack, J. B. (1994). An evolutionary algorithm that constructs recurrent neural networks. *IEEE transactions on Neural Networks*, 5(1):54–65.

[Arendt et al., 2015] Arendt, D., Benito-Gutierrez, E., Brunet, T., and Marlow, H. (2015). Gastric pouches and the mucociliary sole: setting the stage for nervous system evolution. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 370(1684):20150286.

[Arnold, 2011] Arnold, F. H. (2011). The library of maynard-smith: my search for meaning in the protein universe. *Microbe*, 6(7):316.

[Arsiwalla et al., 2017] Arsiwalla, X. D., Sole, R., Moulin-Frier, C., Herreros, I., Sanchez-Fibla, M., and Verschure, P. (2017). The morphospace of consciousness. *arXiv preprint arXiv:1705.11190*.

[Arthur et al., 1993] Arthur, W. B. et al. (1993). On the evolution of complexity. Technical report.

[Avena-Koenigsberger et al., 2014] Avena-Koenigsberger, A., Goñi, J., Betzel, R. F., van den Heuvel, M. P., Griffa, A., Hagmann, P., Thiran, J.-P., and Sporns, O. (2014). Using pareto optimality to explore the topology and dynamics of the human connectome. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 369(1653):20130530.

[Avena-Koenigsberger et al., 2015] Avena-Koenigsberger, A., Goñi, J., Solé, R., and Sporns, O. (2015). Network morphospace. *Journal of the Royal Society Interface*, 12(103):20140881.

[Balagaddé et al., 2008] Balagaddé, F. K., Song, H., Ozaki, J., Collins, C. H., Barnet, M., Arnold, F. H., Quake, S. R., and You, L. (2008). A synthetic escherichia coli predator–prey ecosystem. *Molecular systems biology*, 4(1).

[Basanta et al., 2008] Basanta, D., Miodownik, M., and Baum, B. (2008). The evolution of robust development and homeostasis in artificial organisms. *PLoS computational biology*, 4(3):e1000030.

[Bascompte and Solé, 1998] Bascompte, J. and Solé, R. V. (1998). *Modeling spatiotemporal dynamics in ecology*. Springer.

[Bashor et al., 2010] Bashor, C. J., Horwitz, A. A., Peisajovich, S. G., and Lim, W. A. (2010). Rewiring cells: synthetic biology as a tool to interrogate the organizational principles of living systems. *Annual review of biophysics*, 39:515–537.

[Basu et al., 2005] Basu, S., Gerchman, Y., Collins, C. H., Arnold, F. H., and Weiss, R. (2005). A synthetic multicellular system for programmed pattern formation. *Nature*, 434(7037):1130.

[Beaumont et al., 2009] Beaumont, H. J., Gallie, J., Kost, C., Ferguson, G. C., and Rainey, P. B. (2009). Experimental evolution of bet hedging. *Nature*, 462(7269):90.

[Bely and Nyberg, 2010] Bely, A. E. and Nyberg, K. G. (2010). Evolution of animal regeneration: re-emergence of a field. *Trends in ecology & evolution*, 25(3):161–170.

[Benenson, 2012] Benenson, Y. (2012). Biomolecular computing systems: principles, progress and potential. *Nature Reviews Genetics*, 13(7):455.

[Black et al., 2019] Black, A. J., Bourrat, P., and Rainey, P. B. (2019). Ecological scaffolding and the evolution of individuality: the transition from cells to multicellular life. *BioRxiv*, page 656660.

[Bonner, 1965] Bonner, J. T. (1965). *Size and cycle: an essay on the structure of biology*.

[Bonner, 1974] Bonner, J. T. (1974). *On development: the biology of form*, volume 2. Harvard University Press.

[Bonner, 1988] Bonner, J. T. (1988). *The evolution of complexity by means of natural selection*. Princeton University Press.

[Bonner, 2000] Bonner, J. T. (2000). *First signals: the evolution of multicellular development*. Princeton University Press.

[Bonner, 2006] Bonner, J. T. (2006). *Why size matters: from bacteria to blue whales*. Princeton University Press.

[Boraas et al., 1998] Boraas, M. E., Seale, D. B., and Boxhorn, J. E. (1998). Phagotrophy by a flagellate selects for colonial prey: a possible origin of multicellularity. *Evolutionary Ecology*, 12(2):153–164.

[Bozorgmehr, 2014] Bozorgmehr, J. E. H. (2014). The role of self-organization in developmental evolution. *Theory in Biosciences*, 133(3-4):145–163.

[Brunet and King, 2017] Brunet, T. and King, N. (2017). The origin of animal multicellularity and cell differentiation. *Developmental cell*, 43(2):124–140.

[Bucher and Anderson, 2015] Bucher, D. and Anderson, P. A. (2015). Evolution of the first nervous systems–what can we surmise?

[Budd, 2015] Budd, G. E. (2015). Early animal evolution and the origins of nervous systems. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 370(1684):20150037.

[Budelmann, 1995] Budelmann, B. (1995). The cephalopod nervous system: what evolution has made of the molluscan design. In *The nervous systems of invertebrates: An evolutionary and comparative approach*, pages 115–138. Springer.

[Bullmore and Sporns, 2009] Bullmore, E. and Sporns, O. (2009). Complex brain networks: graph theoretical analysis of structural and functional systems. *Nature reviews neuroscience*, 10(3):186.

[Buss, 1987] Buss, L. W. (1987). *The evolution of individuality*. Princeton University Press.

[Cachat et al., 2014] Cachat, E., Liu, W., Hohenstein, P., and Davies, J. A. (2014). A library of mammalian effector modules for synthetic morphology. *Journal of biological engineering*, 8(1):26.

[Cachat et al., 2016] Cachat, E., Liu, W., Martin, K. C., Yuan, X., Yin, H., Hohenstein, P., and Davies, J. A. (2016). 2-and 3-dimensional synthetic large-scale de novo patterning by mammalian cells through phase separation. *Scientific reports*, 6:20664.

[Canfield et al., 2007] Canfield, D. E., Poulton, S. W., and Narbonne, G. M. (2007). Late-neoproterozoic deep-ocean oxygenation and the rise of animal life. *Science*, 315(5808):92–95.

[Carroll, 2001] Carroll, S. B. (2001). Chance and necessity: the evolution of morphological complexity and diversity. *Nature*, 409(6823):1102.

[Carroll, 2008] Carroll, S. B. (2008). Evo-devo and an expanding evolutionary synthesis: a genetic theory of morphological evolution. *Cell*, 134(1):25–36.

[Caulin and Maley, 2011] Caulin, A. F. and Maley, C. C. (2011). Peto's paradox: evolution's prescription for cancer prevention. *Trends in ecology & evolution*, 26(4):175–182.

[Cervera et al., 2016] Cervera, H., Lalić, J., and Elena, S. F. (2016). Effect of host species on topography of the fitness landscape for a plant rna virus. *Journal of virology*, 90(22):10160–10169.

[Chau et al., 2012] Chau, A. H., Walter, J. M., Gerardin, J., Tang, C., and Lim, W. A. (2012). Designing synthetic regulatory networks capable of self-organizing cell polarization. *Cell*, 151(2):320–332.

[Chen et al., 2006] Chen, B. L., Hall, D. H., and Chklovskii, D. B. (2006). Wiring optimization can relate neuronal structure and function. *Proceedings of the National Academy of Sciences*, 103(12):4723–4728.

[Chklovskii et al., 2002] Chklovskii, D. B., Schikorski, T., and Stevens, C. F. (2002). Wiring optimization in cortical circuits. *Neuron*, 34(3):341–347.

[Cinquin and Demongeot, 2002] Cinquin, O. and Demongeot, J. (2002). Roles of positive and negative feedback in biological systems. *Comptes rendus biologies*, 325(11):1085–1095.

[Clune et al., 2013] Clune, J., Mouret, J.-B., and Lipson, H. (2013). The evolutionary origins of modularity. *Proceedings of the Royal Society b: Biological sciences*, 280(1755):20122863.

[Corominas-Murtra et al., 2013] Corominas-Murtra, B., Goñi, J., Solé, R. V., and Rodríguez-Caso, C. (2013). On the origins of hierarchy in complex networks. *Proceedings of the National Academy of Sciences*, 110(33):13316–13321.

[Cummings, 1990] Cummings, F. (1990). A model of morphogenetic pattern formation. *Journal of theoretical biology*, 144(4):547–566.

[Cummings, 1996] Cummings, F. W. (1996). A model of growth and form based on adhesion molecules. *Journal of theoretical biology*, 178(3):229–238.

[Cummings, 2001] Cummings, F. W. (2001). The interaction of surface geometry with morphogens. *Journal of theoretical biology*, 212(3):303–313.

[Cummings, 2003] Cummings, F. W. (2003). On the origin of pattern and form in early metazoans. *International Journal of Developmental Biology*, 50(2-3):193–208.

[Damuth and Heisler, 1988] Damuth, J. and Heisler, I. L. (1988). Alternative formulations of multilevel selection. *Biology and Philosophy*, 3(4):407–430.

[Darwin, 1859] Darwin, C. (1859). *On the origin of species, 1859.*

[Delgado and Solé, 1997] Delgado, J. and Solé, R. V. (1997). Collective-induced computation. *Physical Review E*, 55(3):2338.

[Duran-Nebreda et al., 2016] Duran-Nebreda, S., Bonforti, A., Montanez, R., Valverde, S., and Solé, R. (2016). Emergence of proto-organisms from bistable stochastic differentiation and adhesion. *Journal of The Royal Society Interface*, 13(117):20160108.

[Duran-Nebreda and Solé, 2015] Duran-Nebreda, S. and Solé, R. (2015). Emergence of multicellularity in a model of cell growth, death and aggregation under size-dependent selection. *Journal of the Royal Society Interface*, 12(102):20140982.

[Dvorkin and Shapiro, 1997] Dvorkin, M. and Shapiro, J. A. (1997). *Bacteria as Multicellular Organisms*. Oxford University Press.

[Eble, 1999] Eble, G. J. (1999). Developmental and non-developmental morphospaces in evolutionary biology. *Morphospace concepts and applications*.

[Edelman and Gally, 2001] Edelman, G. M. and Gally, J. A. (2001). Degeneracy and complexity in biological systems. *Proceedings of the National Academy of Sciences*, 98(24):13763–13768.

[Ellers, 1993] Ellers, O. (1993). A mechanical model of growth in regular sea urchins: predictions of shape and a developmental morphospace. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, 254(1340):123–129.

[Erwin, 2008] Erwin, D. H. (2008). Macroevolution of ecosystem engineering, niche construction and diversity. *Trends in ecology & evolution*, 23(6):304–310.

[Esteve-Altava and Rasskin-Gutman, 2014] Esteve-Altava, B. and Rasskin-Gutman, D. (2014). Theoretical morphology of tetrapod skull networks. *Comptes Rendus Palevol*, 13(1):41–50.

[Ferrell Jr, 2002] Ferrell Jr, J. E. (2002). Self-perpetuating states in signal transduction: positive feedback, double-negative feedback and bistability. *Current opinion in cell biology*, 14(2):140–148.

[Ferretti, 2011] Ferretti, P. (2011). Is there a relationship between adult neurogenesis and neuron generation following injury across evolution? *European Journal of Neuroscience*, 34(6):951–962.

[Floreano et al., 2008] Floreano, D., Dürr, P., and Mattiussi, C. (2008). Neuroevolution: from architectures to learning. *Evolutionary Intelligence*, 1(1):47–62.

[Foty and Steinberg, 2005] Foty, R. A. and Steinberg, M. S. (2005). The differential adhesion hypothesis: a direct evaluation. *Developmental biology*, 278(1):255–263.

[Frank, 1995] Frank, S. A. (1995). Mutual policing and repression of competition in the evolution of cooperative groups. *Nature*, 377(6549):520.

[Freeland and Hurst, 1998] Freeland, S. J. and Hurst, L. D. (1998). The genetic code is one in a million. *Journal of molecular evolution*, 47(3):238–248.

[Fujisawa, 2003] Fujisawa, T. (2003). Hydra regeneration and epitheliopeptides. *Developmental dynamics*, 226(2):182–189.

[Gánti, 2003] Gánti, T. (2003). *The principles of life*. Oxford University Press.

[Gell-Mann, 1995] Gell-Mann, M. (1995). *The Quark and the Jaguar: Adventures in the Simple and the Complex*. Macmillan.

[Gerlee et al., 2011] Gerlee, P., Basanta, D., and Anderson, A. R. (2011). Evolving homeostatic tissue using genetic algorithms. *Progress in biophysics and molecular biology*, 106(2):414–425.

[Goñi et al., 2013] Goñi, J., Avena-Koenigsberger, A., de Mendizabal, N. V., van den Heuvel, M. P., Betzel, R. F., and Sporns, O. (2013). Exploring the morphospace of communication efficiency in complex networks. *PLoS One*, 8(3):e58070.

[Godfrey-Smith, 2009] Godfrey-Smith, P. (2009). *Darwinian populations and natural selection*. Oxford University Press.

[Godfrey-Smith, 2016] Godfrey-Smith, P. (2016). *Other minds: The octopus and the evolution of intelligent life*. William Collins London.

[Goodwin and Trainor, 1985] Goodwin, B. C. and Trainor, L. (1985). Tip and whorl morphogenesis in acetabularia by calcium-regulated strain fields. *Journal of theoretical biology*, 117(1):79–106.

[Gould, 1990] Gould, S. J. (1990). *Wonderful life: the Burgess Shale and the nature of history*. WW Norton & Company.

[Gould, 1991] Gould, S. J. (1991). The disparity of the burgess shale arthropod fauna and the limits of cladistic analysis: why we must strive to quantify morphospace. *Paleobiology*, 17(4):411–423.

[Gould and Eldredge, 1977] Gould, S. J. and Eldredge, N. (1977). Punctuated equilibria: the tempo and mode of evolution reconsidered. *Paleobiology*, 3(2):115–151.

[Grosberg and Strathmann, 2007] Grosberg, R. K. and Strathmann, R. R. (2007). The evolution of multicellularity: a minor major transition? *Annu. Rev. Ecol. Evol. Syst.*, 38:621–654.

[Gumbiner, 1996] Gumbiner, B. M. (1996). Cell adhesion: the molecular basis of tissue architecture and morphogenesis. *Cell*, 84(3):345–357.

[Hamilton, 1963] Hamilton, W. D. (1963). The evolution of altruistic behavior. *The American Naturalist*, 97(896):354–356.

[Hamilton, 1964] Hamilton, W. D. (1964). The genetical evolution of social behaviour. ii. *Journal of theoretical biology*, 7(1):17–52.

[Hammerschmidt et al., 2014] Hammerschmidt, K., Rose, C. J., Kerr, B., and Rainey, P. B. (2014). Life cycles, fitness decoupling and the evolution of multicellularity. *Nature*, 515(7525):75.

[Hanschen et al., 2015] Hanschen, E. R., Shelton, D. E., and Michod, R. E. (2015). Evolutionary transitions in individuality and recent models of multicellularity. In *Evolutionary Transitions to Multicellular Life*, pages 165–188. Springer.

[Hart et al., 2015] Hart, Y., Sheftel, H., Hausser, J., Szekely, P., Ben-Moshe, N. B., Korem, Y., Tendler, A., Mayo, A. E., and Alon, U. (2015). Inferring biological tasks using pareto analysis of high-dimensional data. *Nature methods*, 12(3):233.

[Hickman, 1993] Hickman, C. (1993). Theoretical design space: a new program for the analysis of structural diversity. *Neues Jahrbuch fur Geologie und Palaontologie, Abhandlungen*, 190:169–182.

[Hogeweg, 2000a] Hogeweg, P. (2000a). Evolving mechanisms of morphogenesis: on the interplay between differential adhesion and cell differentiation. *Journal of theoretical biology*, 203(4):317–333.

[Hogeweg, 2000b] Hogeweg, P. (2000b). Shapes in the shadow: evolutionary dynamics of morphogenesis. *Artificial life*, 6(1):85–101.

[Holland, 1994] Holland, J. H. (1994). Echoing emergence: Objectives, rough definitions, and speculations for echo-class models, complexity: Metaphors, models and reality (ga cowan, d. pines, and d. meltzer, eds.). *Santa Fe Institute Studies in the Sciences of Complexity, Proceeding*, 19.

[Hölldobler and Wilson, 2009] Hölldobler, B. and Wilson, E. O. (2009). *The superorganism: the beauty, elegance, and strangeness of insect societies*. WW Norton & Company.

[Hsu and Waltman, 1992] Hsu, S.-B. and Waltman, P. (1992). Analysis of a model of two competitors in a chemostat with an external inhibitor. *SIAM journal on applied mathematics*, 52(2):528–540.

[Hsu and Waltman, 2004] Hsu, S.-B. and Waltman, P. (2004). A survey of mathematical models of competition with an inhibitor. *Mathematical Biosciences*, 187(1):53–91.

[Jablonka, 2002] Jablonka, E. (2002). Information: Its interpretation, its inheritance, and its sharing. *Philosophy of science*, 69(4):578–605.

[Jablonka and Lamb, 2005] Jablonka, E. and Lamb, M. J. (2005). *Evolution in four dimensions: Genetic, epigenetic, behavioral, and symbolic variation in the history of life*. MIT press.

[Jablonka and Lamb, 2006] Jablonka, E. and Lamb, M. J. (2006). The evolution of information in the major transitions. *Journal of Theoretical Biology*, 239(2):236–246.

[Jablonka and Szathmáry, 1995] Jablonka, E. and Szathmáry, E. (1995). The evolution of information storage and heredity. *Trends in ecology & evolution*, 10(5):206–211.

[Jacob, 1977] Jacob, F. (1977). Evolution and tinkering. *Science*, 196(4295):1161–1166.

[Jékely, 2010] Jékely, G. (2010). Origin and early evolution of neural circuits for the control of ciliary locomotion. *Proceedings of the Royal Society B: Biological Sciences*, 278(1707):914–922.

[Jékely et al., 2015] Jékely, G., Keijzer, F., and Godfrey-Smith, P. (2015). An option space for early neural evolution. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 370(1684):20150181.

[Johnston et al., 2012] Johnston, D., Poulton, S., Goldberg, T., Sergeev, V., Podkovyrov, V., Vorob’Eva, N., Bekker, A., and Knoll, A. (2012). Late ediacaran redox stability and metazoan evolution. *Earth and Planetary Science Letters*, 335:25–35.

[Kaslin et al., 2007] Kaslin, J., Ganz, J., and Brand, M. (2007). Proliferation, neurogenesis and regeneration in the non-mammalian vertebrate brain. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 363(1489):101–122.

[Kauffman and Levin, 1987] Kauffman, S. and Levin, S. (1987). Towards a general theory of adaptive walks on rugged landscapes. *Journal of theoretical Biology*, 128(1):11–45.

[Kauffman, 1993] Kauffman, S. A. (1993). *The origins of order: Self-organization and selection in evolution*. OUP USA.

[Kauffman and Weinberger, 1989] Kauffman, S. A. and Weinberger, E. D. (1989). The nk model of rugged fitness landscapes and its application to maturation of the immune response. *Journal of theoretical biology*, 141(2):211–245.

[King, 2004] King, N. (2004). The unicellular ancestry of animal development. *Developmental cell*, 7(3):313–325.

[Kirk, 2005] Kirk, D. L. (2005). A twelve-step program for evolving multicellularity and a division of labor. *BioEssays*, 27(3):299–310.

[Kitano, 2004] Kitano, H. (2004). Biological robustness. *Nature Reviews Genetics*, 5(11):826.

[Kitano, 2007] Kitano, H. (2007). Towards a theory of biological robustness. *Molecular systems biology*, 3(1).

[Knoll, 2011] Knoll, A. H. (2011). The multiple origins of complex multicellularity. *Annual Review of Earth and Planetary Sciences*, 39:217–239.

[Knoll and Carroll, 1999] Knoll, A. H. and Carroll, S. B. (1999). Early animal evolution: emerging views from comparative biology and geology. *science*, 284(5423):2129–2137.

[Koch and Meinhardt, 1994] Koch, A. and Meinhardt, H. (1994). Biological pattern formation: from basic mechanisms to complex structures. *Reviews of modern physics*, 66(4):1481.

[Lancaster and Knoblich, 2014] Lancaster, M. A. and Knoblich, J. A. (2014). Organogenesis in a dish: modeling development and disease using organoid technologies. *Science*, 345(6194):1247125.

[Laughlin and Sejnowski, 2003] Laughlin, S. B. and Sejnowski, T. J. (2003). Communication in neuronal networks. *Science*, 301(5641):1870–1874.

[Lenski and Hattingh, 1986] Lenski, R. E. and Hattingh, S. E. (1986). Coexistence of two competitors on one resource and one inhibitor: a chemostat model based on bacteria and antibiotics. *Journal of Theoretical Biology*, 122(1):83–93.

[Lenski and Travisano, 1994] Lenski, R. E. and Travisano, M. (1994). Dynamics of adaptation and diversification: a 10,000-generation experiment with bacterial populations. *Proceedings of the National Academy of Sciences*, 91(15):6808–6814.

[Lenton et al., 2011] Lenton, T., Watson, A., and Watson, A. J. (2011). *Revolutions that made the Earth*. Oxford University Press.

[Libby and Rainey, 2013] Libby, E. and Rainey, P. B. (2013). A conceptual framework for the evolutionary origins of multicellularity. *Physical biology*, 10(3):035001.

[Liebeskind et al., 2016] Liebeskind, B. J., Hillis, D. M., Zakon, H. H., and Hofmann, H. A. (2016). Complex homology and the evolution of nervous systems. *Trends in ecology & evolution*, 31(2):127–135.

[Liu et al., 2011] Liu, C., Fu, X., Liu, L., Ren, X., Chau, C. K., Li, S., Xiang, L., Zeng, H., Chen, G., Tang, L.-H., et al. (2011). Sequential establishment of stripe patterns in an expanding cell population. *Science*, 334(6053):238–241.

[Llinás, 1987] Llinás, R. R. (1987). *’Mindedness’ as a functional state of the brain.* Oxford, UK: Basil Blackwell.

[Llinás, 2001] Llinás, R. R. (2001). *I of the vortex: From neurons to self*, volume 50. MIT press Cambridge, MA.

[Maslov and Sneppen, 2002] Maslov, S. and Sneppen, K. (2002). Specificity and stability in topology of protein networks. *Science*, 296(5569):910–913.

[Maynard Smith et al., 1985] Maynard Smith, J., Burian, R., Kauffman, S., Alberch, P., Campbell, J., Goodwin, B., Lande, R., Raup, D., and 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.

[McCartney and Loper, 1989] McCartney, K. and Loper, D. E. (1989). Optimized skeletal morphologies of silicoflagellate genera dictyocha and distephanus. *Paleobiology*, 15(3):283–298.

[McGhee, 1980] McGhee, G. R. (1980). Shell form in the biconvex articulate brachiopoda: a geometric analysis. *Paleobiology*, 6(1):57–76.

[McGhee, 1999] McGhee, G. R. (1999). *Theoretical morphology: the concept and its applications*. Columbia University Press.

[McGhee, 2006] McGhee, G. R. (2006). *The geometry of evolution: adaptive landscapes and theoretical morphospaces*. Cambridge University Press.

[McGhee, 2011] McGhee, G. R. (2011). *Convergent evolution: limited forms most beautiful*. MIT Press.

[McKinney and Raup, 1982] McKinney, F. K. and Raup, D. M. (1982). A turn in the right direction: simulation of erect spiral growth in the bryozoans archimedes and bugula. *Paleobiology*, 8(2):101–112.

[McShea, 1991] McShea, D. W. (1991). Complexity and evolution: what everybody knows. *Biology and Philosophy*, 6(3):303–324.

[McShea, 1996] McShea, D. W. (1996). Perspective metazoan complexity and evolution: Is there a trend? *Evolution*, 50(2):477–492.

[McShea, 2002] McShea, D. W. (2002). A complexity drain on cells in the evolution of multicellularity. *Evolution*, 56(3):441–452.

[Michod, 2007] Michod, R. E. (2007). Evolution of individuality during the transition from unicellular to multicellular life. *Proceedings of the National Academy of Sciences*, 104(suppl 1):8613–8618.

[Miguel-Tomé, 2018] Miguel-Tomé, S. (2018). The influence of computational traits on the natural selection of the nervous system. *Natural Computing*, 17(2):403–425.

[Miller et al., 1989] Miller, G. F., Todd, P. M., and Hegde, S. U. (1989). Designing neural networks using genetic algorithms. In *ICGA*, volume 89, pages 379–384.

[Miller et al., 2012] Miller, M., Hafner, M., Sontag, E., Davidsohn, N., Subramanian, S., Purnick, P. E., Lauffenburger, D., and Weiss, R. (2012). Modular design of artificial tissue homeostasis: robust control through synthetic cellular heterogeneity. *PLoS computational biology*, 8(7):e1002579.

[Moffet, 2012] Moffet, S. B. (2012). *Nervous system regeneration in the invertebrates*, volume 34. Springer Science & Business Media.

[Moore and Shannon, 1956] Moore, E. F. and Shannon, C. E. (1956). Reliable circuits using less reliable relays. *Journal of the Franklin Institute*, 262(3):191–208.

[Moroz, 2009] Moroz, L. L. (2009). On the independent origins of complex brains and neurons. *Brain, behavior and evolution*, 74(3):177–190.

[Moroz et al., 2014] Moroz, L. L., Kocot, K. M., Citarella, M. R., Dosung, S., Norekian, T. P., Povolotskaya, I. S., Grigorenko, A. P., Dailey, C., Berezikov, E., Buckley, K. M., et al. (2014). The ctenophore genome and the evolutionary origins of neural systems. *Nature*, 510(7503):109.

[Morris, 1989] Morris, S. C. (1989). Burgess shale faunas and the cambrian explosion. *Science*, 246(4928):339–346.

[Morris, 1999] Morris, S. C. (1999). *The crucible of creation: the Burgess Shale and the rise of animals*. Peterson's.

[Morris, 2003] Morris, S. C. (2003). *Life's solution: inevitable humans in a lonely universe*. Cambridge University Press.

[Müller, 2007] Müller, G. B. (2007). Evo–devo: extending the evolutionary synthesis. *Nature reviews genetics*, 8(12):943.

[Muñoz-Dorado et al., 2016] Muñoz-Dorado, J., Marcos-Torres, F. J., García-Bravo, E., Moraleda-Muñoz, A., and Pérez, J. (2016). Myxobacteria: moving, killing, feeding, and surviving together. *Frontiers in microbiology*, 7:781.

[Newman, 2012] Newman, S. A. (2012). Physico-genetic determinants in the evolution of development. *Science*, 338(6104):217–219.

[Newman and Bhat, 2008] Newman, S. A. and Bhat, R. (2008). Dynamical patterning modules: physico-genetic determinants of morphological development and evolution. *Physical Biology*, 5(1):015008.

[Newman and Bhat, 2009] Newman, S. A. and Bhat, R. (2009). Dynamical patterning modules: a" pattern language" for development and evolution of multicellular form. *International Journal of Developmental Biology*, 53(5-6):693–705.

[Newman et al., 2009] Newman, S. A., Bhat, R., and Mezentseva, N. V. (2009). Cell state switching factors and dynamical patterning modules: complementary mediators of plasticity in development and evolution. *Journal of biosciences*, 34(4):553–572.

[Newman et al., 2003] Newman, S. A., Forgacs, G., and Muller, G. B. (2003). Before programs: the physical origination of multicellular forms. *International Journal of Developmental Biology*, 50(2-3):289–299.

[Niklas, 1994] Niklas, K. J. (1994). Morphological evolution through complex domains of fitness. *Proceedings of the National Academy of Sciences*, 91(15):6772–6779.

[Niklas, 1997] Niklas, K. J. (1997). *The evolutionary biology of plants*. University of Chicago Press.

[Niklas, 1999] Niklas, K. J. (1999). Evolutionary walks through a land plant morphospace. *Journal of Experimental Botany*, 50(330):39–52.

[Niklas, 2004] Niklas, K. J. (2004). Computer models of early land plant evolution. *Annu. Rev. Earth Planet. Sci.*, 32:47–66.

[Niklas and Kerchner, 1984] Niklas, K. J. and Kerchner, V. (1984). Mechanical and photosynthetic constraints on the evolution of plant shape. *Paleobiology*, 10(1):79–101.

[Niklas and Newman, 2013] Niklas, K. J. and Newman, S. A. (2013). The origins of multicellular organisms. *Evolution & development*, 15(1):41–52.

[Novack-Gottshall, 2007] Novack-Gottshall, P. M. (2007). Using a theoretical ecospace to quantify the ecological diversity of paleozoic and modern marine biotas. *Paleobiology*, 33(2):273–294.

[Ofria and Wilke, 2004] Ofria, C. and Wilke, C. O. (2004). Avida: A software platform for research in computational evolutionary biology. *Artificial life*, 10(2):191–229.

[Okasha, 2001] Okasha, S. (2001). Why won't the group selection controversy go away? *The British journal for the philosophy of science*, 52(1):25–50.

[Okasha, 2005] Okasha, S. (2005). Multilevel selection and the major transitions in evolution. *Philosophy of science*, 72(5):1013–1025.

[Olson, 2013] Olson, B. J. (2013). Multicellularity: from brief encounters to lifelong unions. *Elife*, 2:e01893.

[Oster and Alberch, 1982] Oster, G. and Alberch, P. (1982). Evolution and bifurcation of developmental programs. *Evolution*, pages 444–459.

[Oyston et al., 2015] Oyston, J. W., Hughes, M., Gerber, S., and Wills, M. A. (2015). Why should we investigate the morphological disparity of plant clades? *Annals of botany*, 117(5):859–879.

[Pagán, 2014] Pagán, O. R. (2014). *The first brain: the neuroscience of planarians*. Oxford University Press.

[Park et al., 2014] Park, J. S., Rhau, B., Hermann, A., McNally, K. A., Zhou, C., Gong, D., Weiner, O. D., Conklin, B. R., Onuffer, J., and Lim, W. A. (2014). Synthetic control of mammalian-cell motility by engineering chemotaxis to an orthogonal bioinert chemical signal. *Proceedings of the National Academy of Sciences*, 111(16):5896–5901.

[Pfeiffer and Bonhoeffer, 2002] Pfeiffer, T. and Bonhoeffer, S. (2002). Evolutionary consequences of tradeoffs between yield and rate of atp production. *Zeitschrift für Physikalische Chemie*, 216(1):51.

[Pfeiffer and Bonhoeffer, 2003] Pfeiffer, T. and Bonhoeffer, S. (2003). An evolutionary scenario for the transition to undifferentiated multicellularity. *Proceedings of the National Academy of Sciences*, 100(3):1095–1098.

[Pfeiffer et al., 2001] Pfeiffer, T., Schuster, S., and Bonhoeffer, S. (2001). Cooperation and competition in the evolution of atp-producing pathways. *Science*, 292(5516):504–507.

[Pocheville, 2015] Pocheville, A. (2015). The ecological niche: history and recent controversies. In *Handbook of evolutionary thinking in the sciences*, pages 547–586. Springer.

[Queller, 1997] Queller, D. C. (1997). Cooperators since life began.

[Queller, 2000] Queller, D. C. (2000). Relatedness and the fraternal major transitions. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, 355(1403):1647–1655.

[Queller and Strassmann, 2009] Queller, D. C. and Strassmann, J. E. (2009). Beyond society: the evolution of organismality. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 364(1533):3143–3155.

[Rainey, 2007] Rainey, P. B. (2007). Unity from conflict. *Nature*, 446(7136):616.

[Rainey and Kerr, 2010] Rainey, P. B. and Kerr, B. (2010). Cheats as first propagules: a new hypothesis for the evolution of individuality during the transition from single cells to multicellularity. *Bioessays*, 32(10):872–880.

[Rainey and Rainey, 2003] Rainey, P. B. and Rainey, K. (2003). Evolution of cooperation and conflict in experimental bacterial populations. *Nature*, 425(6953):72.

[Rainey and Travisano, 1998] Rainey, P. B. and Travisano, M. (1998). Adaptive radiation in a heterogeneous environment. *Nature*, 394(6688):69.

[Ratcliff et al., 2012] Ratcliff, W. C., Denison, R. F., Borrello, M., and Travisano, M. (2012). Experimental evolution of multicellularity. *Proceedings of the National Academy of Sciences*, 109(5):1595–1600.

[Ratcliff et al., 2013] Ratcliff, W. C., Herron, M. D., Howell, K., Pentz, J. T., Rosenzweig, F., and Travisano, M. (2013). Experimental evolution of an alternating uni-and multicellular life cycle in chlamydomonas reinhardtii. *Nature communications*, 4:2742.

[Raup, 1961] Raup, D. M. (1961). The geometry of coiling in gastropods. *Proceedings of the National Academy of Sciences of the United States of America*, 47(4):602.

[Raup, 1962] Raup, D. M. (1962). Computer as aid in describing form in gastropod shells. *Science*, 138(3537):150–152.

[Raup, 1966] Raup, D. M. (1966). Geometric analysis of shell coiling: general problems. *Journal of Paleontology*, pages 1178–1190.

[Raup and Michelson, 1965] Raup, D. M. and Michelson, A. (1965). Theoretical morphology of the coiled shell. *Science*, 147(3663):1294–1295.

[Ray, 1991] Ray, T. (1991). Documentation for the tierra simulator. *Tierra Simulator*, 4.

[Ray, 1998] Ray, T. S. (1998). Selecting naturally for differentiation: Preliminary evolutionary results. *Complexity*, 3(5):25–33.

[Richter and King, 2013] Richter, D. J. and King, N. (2013). The genomic and cellular foundations of animal origins. *Annual review of genetics*, 47:509–537.

[Sachs, 2008] Sachs, J. L. (2008). Resolving the first steps to multicellularity. *Trends in ecology & evolution*, 23(5):245–248.

[Sasai, 2013] Sasai, Y. (2013). Cytosystems dynamics in self-organization of tissue architecture. *Nature*, 493(7432):318.

[Savazzi, 1987] Savazzi, E. (1987). Geometric and functional constraints on bivalve shell morphology. *Lethaia*, 20(4):293–306.

[Schaerli et al., 2014] Schaerli, Y., Munteanu, A., Gili, M., Cotterell, J., Sharpe, J., and Isalan, M. (2014). A unified design space of synthetic stripe-forming networks. *Nature communications*, 5:4905.

[Schuster, 1996] Schuster, P. (1996). How does complexity arise in evolution? In *Evolution and Progress in Democracies*, pages 147–159. Springer.

[Sebé-Pedrós et al., 2017] Sebé-Pedrós, A., Degnan, B. M., and Ruiz-Trillo, I. (2017). The origin of metazoa: a unicellular perspective. *Nature Reviews Genetics*, 18(8):498.

[Seoane and Solé, 2015a] Seoane, L. F. and Solé, R. (2015a). Phase transitions in pareto optimal complex networks. *Physical Review E*, 92(3):032807.

[Seoane and Solé, 2015b] Seoane, L. F. and Solé, R. (2015b). Systems poised to criticality through pareto selective forces. *arXiv preprint arXiv:1510.08697*.

[Seoane and Solé, 2016] Seoane, L. F. and Solé, R. (2016). Multiobjective optimization and phase transitions. In *Proceedings of ECCS 2014*, pages 259–270. Springer.

[Seoane and Solé, 2018] Seoane, L. F. and Solé, R. (2018). The morphospace of language networks. *Scientific reports*, 8.

[Seoane and Solé, 2013] Seoane, L. F. and Solé, R. V. (2013). A multiobjective optimization approach to statistical mechanics. *arXiv preprint arXiv:1310.6372*.

[Shapiro, 1988] Shapiro, J. A. (1988). Bacteria as multicellular organisms. *Scientific American*, 258(6):82–89.

[Shapiro, 1998] Shapiro, J. A. (1998). Thinking about bacterial populations as multicellular organisms. *Annual review of microbiology*, 52(1):81–104.

[Sharpe et al., 2015] Sharpe, S. C., Eme, L., Brown, M. W., and Roger, A. J. (2015). Timing the origins of multicellular eukaryotes through phylogenomics and relaxed molecular clock analyses. In *Evolutionary transitions to multicellular life*, pages 3–29. Springer.

[Sheftel et al., 2013] Sheftel, H., Shoval, O., Mayo, A., and Alon, U. (2013). The geometry of the p areto front in biological phenotype space. *Ecology and evolution*, 3(6):1471–1483.

[Shoval et al., 2012] Shoval, O., Sheftel, H., Shinar, G., Hart, Y., Ramote, O., Mayo, A., Dekel, E., Kavanagh, K., and Alon, U. (2012). Evolutionary trade-offs, pareto optimality, and the geometry of phenotype space. *Science*, 336(6085):1157–1160.

[Smith, 1970] Smith, J. M. (1970). Natural selection and the concept of a protein space. *Nature*, 225(5232):563.

[Smith and Szathmary, 1995] Smith, J. M. and Szathmary, E. (1995). *The major transitions in evolution*. Oxford University Press.

[Smith and Morowitz, 1982] Smith, T. and Morowitz, H. (1982). Between physics and history. *J. Mol. Evol*, 18:265–282.

[Sogabe et al., 2019] Sogabe, S., Hatleberg, W. L., Kocot, K. M., Say, T. E., Stoupin, D., Roper, K. E., Fernandez-Valverde, S. L., Degnan, S. M., and Degnan, B. M. (2019). Pluripotency and the origin of animal multicellularity. *Nature*, page 1.

[Solé, 2011] Solé, R. (2011). Phase transitions. princeton u. *Press. Princeton*.

[Solé, 2016a] Solé, R. (2016a). *La lógica de los monstruos*. Tusquets.

[Solé, 2016b] Solé, R. (2016b). The major synthetic evolutionary transitions.

[Solé, 2016c] Solé, R. (2016c). Synthetic transitions: towards a new synthesis. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 371(1701):20150438.

[Sole, 2017] Sole, R. (2017). Rise of the humanbot. *arXiv preprint arXiv:1705.05935*.

[Solé et al., 2019] Solé, R., Moses, M., and Forrest, S. (2019). Liquid brains, solid brains.

[Solé and Macia, 2013] Solé, R. V. and Macia, J. (2013). Expanding the landscape of biological computation with synthetic multicellular consortia. *Natural Computing*, 12(4):485–497.

[Solé and Valverde, 2004] Solé, R. V. and Valverde, S. (2004). Information theory of complex networks: on evolution and architectural constraints. In *Complex networks*, pages 189–207. Springer.

[Solé and Valverde, 2007] Solé, R. V. and Valverde, S. (2007). Spontaneous emergence of modularity in cellular networks. *Journal of The Royal Society Interface*, 5(18):129–133.

[Solé and Valverde, 2013a] Solé, R. V. and Valverde, S. (2013a). Before the endless forms: embodied model of transition from single cells to aggregates to ecosystem engineering. *PloS one*, 8(4):e59664.

[Solé and Valverde, 2013b] Solé, R. V. and Valverde, S. (2013b). Macroevolution in silico: scales, constraints and universals. *Palaeontology*, 56(6):1327–1340.

[Stanley et al., 2009] Stanley, K. O., D'Ambrosio, D. B., and Gauci, J. (2009). A hypercube-based encoding for evolving large-scale neural networks. *Artificial life*, 15(2):185–212.

[Stanley and Miikkulainen, 2002] Stanley, K. O. and Miikkulainen, R. (2002). Evolving neural networks through augmenting topologies. *Evolutionary computation*, 10(2):99–127.

[Stanley, 1973] Stanley, S. M. (1973). An ecological theory for the sudden origin of multicellular life in the late precambrian. *Proceedings of the National Academy of Sciences*, 70(5):1486–1489.

[Staps et al., 2019] Staps, M., van Gestel, J., and Tarnita, C. E. (2019). Emergence of diverse life cycles and life histories at the origin of multicellularity. *Nature ecology & evolution*, 3(8):1197–1205.

[Stelling et al., 2004] Stelling, J., Sauer, U., Szallasi, Z., Doyle III, F. J., and Doyle, J. (2004). Robustness of cellular functions. *Cell*, 118(6):675–685.

[Swan and Kershaw, 1994] Swan, A. R. and Kershaw, S. (1994). A computer model for skeletal growth of stromatoporoids. *Palaeontology*, 37(2):409–424.

[Szathmáry, 2015] Szathmáry, E. (2015). Toward major evolutionary transitions theory 2.0. *Proceedings of the National Academy of Sciences*, 112(33):10104–10111.

[Szathmáry and Smith, 1995] Szathmáry, E. and Smith, J. M. (1995). The major evolutionary transitions. *Nature*, 374(6519):227.

[Tanaka and Ferretti, 2009] Tanaka, E. M. and Ferretti, P. (2009). Considering the evolution of regeneration in the central nervous system. *Nature Reviews Neuroscience*, 10(10):713.

[Thomas and Reif, 1993] Thomas, R. D. and Reif, W.-E. (1993). The skeleton space: a finite set of organic designs. *Evolution*, 47(2):341–360.

[Thompson, 1952] Thompson, d. W. (1952). *On growth and form*, volume 1. Cambridge university press.

[Travisano et al., 1995] Travisano, M., Mongold, J. A., Bennett, A. F., and Lenski, R. E. (1995). Experimental tests of the roles of adaptation, chance, and history in evolution. *Science*, 267(5194):87–90.

[Turing, 1990] Turing, A. M. (1990). The chemical basis of morphogenesis. *Bulletin of mathematical biology*, 52(1-2):153–197.

[Tyszka, 2006] Tyszka, J. (2006). Morphospace of foraminiferal shells: results from the moving reference model. *Lethaia*, 39(1):1–12.

[Valentine et al., 1994] Valentine, J. W., Collins, A. G., and Meyer, C. P. (1994). Morphological complexity increase in metazoans. *Paleobiology*, 20(2):131–142.

[Valverde, 2016] Valverde, S. (2016). Major transitions in information technology. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 371(1701):20150450.

[Valverde, 2017] Valverde, S. (2017). Breakdown of modularity in complex networks. *Frontiers in physiology*, 8:497.

[Van Gestel and Tarnita, 2017] Van Gestel, J. and Tarnita, C. E. (2017). On the origin of biological construction, with a focus on multicellularity. *Proceedings of the National Academy of Sciences*, 114(42):11018–11026.

[Veening et al., 2008] Veening, J.-W., Smits, W. K., and Kuipers, O. P. (2008). Bistability, epigenetics, and bet-hedging in bacteria. *Annu. Rev. Microbiol.*, 62:193–210.

[Vértes et al., 2012] Vértes, P. E., Alexander-Bloch, A. F., Gogtay, N., Giedd, J. N., Rapoport, J. L., and Bullmore, E. T. (2012). Simple models of human brain functional networks. *Proceedings of the National Academy of Sciences*, 109(15):5868–5873.

[Von Neumann, 1956] Von Neumann, J. (1956). Probabilistic logics and the synthesis of reliable organisms from unreliable components. *Automata studies*, 34:43–98.

[Waddington, 1942] Waddington, C. H. (1942). Canalization of development and the inheritance of acquired characters. *Nature*, 150(3811):563.

[Weber et al., 2007] Weber, W., Daoud-El Baba, M., and Fussenegger, M. (2007). Synthetic ecosystems based on airborne inter-and intrakingdom communication. *Proceedings of the National Academy of Sciences*, 104(25):10435–10440.

[Weinreich et al., 2006] Weinreich, D. M., Delaney, N. F., DePristo, M. A., and Hartl, D. L. (2006). Darwinian evolution can follow only very few mutational paths to fitter proteins. *science*, 312(5770):111–114.

[West et al., 2015] West, S. A., Fisher, R. M., Gardner, A., and Kiers, E. T. (2015). Major evolutionary transitions in individuality. *Proceedings of the National Academy of Sciences*, 112(33):10112–10119.

[West et al., 2007] West, S. A., Griffin, A. S., and Gardner, A. (2007). Social semantics: altruism, cooperation, mutualism, strong reciprocity and group selection. *Journal of evolutionary biology*, 20(2):415–432.

[West et al., 2006] West, S. A., Griffin, A. S., Gardner, A., and Diggle, S. P. (2006). Social evolution theory for microorganisms. *Nature reviews microbiology*, 4(8):597.

[West and Kiers, 2009] West, S. A. and Kiers, E. T. (2009). Evolution: what is an organism? *Current Biology*, 19(23):R1080–R1082.

[Whitacre, 2012] Whitacre, J. M. (2012). Biological robustness: paradigms, mechanisms, and systems principles. *Frontiers in genetics*, 3:67.

[Willyard, 2015] Willyard, C. (2015). The boom in mini stomachs, brains, breasts, kidneys and more. *Nature News*, 523(7562):520.

[Winograd and Cowan, 1963] Winograd, S. and Cowan, J. D. (1963). *Reliable computation in the presence of noise*. Number 22. Mit Press Cambridge, Mass.

[Wolpert, 1969] Wolpert, L. (1969). Positional information and the spatial pattern of cellular differentiation. *Journal of theoretical biology*, 25(1):1–47.

[Wolpert, 1971] Wolpert, L. (1971). Positional information and pattern formation. In *Current topics in developmental biology*, volume 6, pages 183–224. Elsevier.

[Wray, 2015] Wray, G. A. (2015). Molecular clocks and the early evolution of metazoan nervous systems. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 370(1684):20150046.

[Wright, 1932] Wright, S. (1932). *The roles of mutation, inbreeding, crossbreeding, and selection in evolution*, volume 1. na.

[Yao, 1999] Yao, X. (1999). Evolving artificial neural networks. *Proceedings of the IEEE*, 87(9):1423–1447.

[Yin et al., 2016] Yin, X., Mead, B. E., Safaee, H., Langer, R., Karp, J. M., and Levy, O. (2016). Engineering stem cell organoids. *Cell stem cell*, 18(1):25–38.