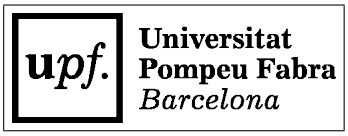
Artificial multicellularity and pattern formation

Salvador Duran Nebreda

TESI DOCTORAL UPF / ANY 2015

DIRECTOR DE LA TESI
Ricard Solé

Departament CEXS



Al meus pares, tant els biològics com els intel·lectuals.

Agraïments

Aquesta tesi mai hauria pogut existir sense la casualitat que em va portar al Complex Systems Lab ara fa ja 5 anys. A les persones implicades (que ja saben qui són), el meu etern agraïment per haver-me ofert aquesta possibilitat. Si no haguéssiu parlat entre vosaltres segurament hagués acabat decapitant ratolins i estudiant una ruta metabòlica durant anys, cosa que ara veig que no és per a mi.

Al mateix temps, em sento profundament agraït a aquelles persones que em van acollir i ensenyar el nou món dels sistemes complexos. Ha estat tot un repte per a mi desenvolupar les eines matemàtiques i conceptuals que sostenen aquesta tesi i en cap cas hauria estat possible sense la vostra guia i suport. En particular, em sento en especial deute amb Carlos Rodríguez-Caso, Javier Macía i Ricard Solé.

Pel Carlos i el Javier: tot i que al final el projecte científic canvies i tres anys d'esforç no es plasmessin en un article, la honestedat intel·lectual que transpireu i la proximitat envers el que seria un subordinat vostre, ha deixat una empremta difícil d'esborrar. En un sistema tan pervers com pot ser a vegades el món acadèmic, vosaltres demostreu dia a dia que poden haver-hi altres valors. Gràcies.

Pel Ricard: gràcies per encoratjar-me des del primer dia a llegir (no sé si estic tan content de que m'encoratgessis a comprar). Les discussions que hem tingut (amb altres membres del CSL també) han estat font d'inspiració i inquietud intel·lectual des del primer dia, el mateix que em vas deixar clar que el més important era fer-se preguntes rellevants. Em sento infinitament agraït per l'oportunitat de desenvolupar les pròpies idees i treballar independentment que he trobat al CSL, ara sé que aquest atribut fa del grup un lloc únic al món.

Pel Sergi: els nostres projectes secrets continuen vigents. Ara més seriosament, gràcies per involucrar-te en la meva formació i proposar tot tipus de reptes estimulants i idees esbojarrades. Algun dia encara ens podrem forrar, sigues pacient.

També és important per a mi fer una menció als altres PhDs, membres del CSL i habitants del despatx que han anant venint i marxant: Max

Carbonell, Eva Garcia-Ramallo, Núria Conde, Amadís Pagès, Ben Shirt-Ediss, Luiño Seoane, Adriano Bonforti, Dani R. Amor, Josep Sardanyés, Raul Montañez, Aina Ollé, Jordi Piñero, José Luis Villanueva i Ana L. *et al*. Les nostres discussions (trivials o científiques) han estat en tot cas un plaer. Seria molt difícil no apreciar l'entorn que s'ha generat al CSL durant aquest doctorat, construit en gran part pel vostre gran humor (potser aquesta darrera part els del despatx del costat no l'hagin apreciat tant com jo). Ara que aquesta etapa s'acaba i els nostres camins potser se separen, sento que abandono un pis compartit amb uns amics estimats.

A la família i als amics, que han hagut d'aguantar 5 anys de tonteries que només em podien interessar a mi, especialment en aquest esprint final a contra rellotge. A aquells que han fet un incessant esforç en intentar comprendre de què anava la meva tesi, quan a vegades ni jo ho sabia. Gràcies pel suport i per oferir-me un petit descans del món científic quan era necessari pel meu bé.

Abstract

Multicellularity, the property of cells to work in a concerted manner to become a new individual, has appeared several times in the history of our biosphere. This collection of events still presents some open questions, specially with regards to the evolution and maintenance of collective structures and new levels of organization and individuality. Taking very different approaches like synthetic biology, artificial life and modeling of artificial evolution, we have tried to gain some insights into particular cases as well as proposing more general scenarios for the evolution of MC and the formation of complex bodily structures. Such endeavor is tied to a very particular step in the evolution of complexity in our biosphere, yet the combined approaches used here might offer some guidance in unraveling the broader picture.

Resum

La multicel·lularitat, o la habilitat de les cèl·lules d'actuar de forma coordinada i formar part d'una entitat major, ha estat inventada diverses vegades al llarg de la història de la nostra biosfera. Aquest repertori d'esdeveniments encara presenta innumerables incògnites, especialment en relació a la construcció d'estructures biològiques i la formació de noves entitats darwinianes. Utilitzant estratègies artificials com la biologia sintètica, la vida artificial o el modelat per ordinador, el treball presentat aquí pretén oferir algunes respostes als temes de la multicel·lularitat i la creació de patrons espaials. Aquestes qüestions estan relacionades amb només un dels grans salts en complexitat que a ha sofert la història de la vida, tanmateix el conjunt d'eines emprades aquí pot servir d'inspiració per a la recerca d'altres transicions en individualitat.

Contents

List of Figures

LIST OF FIGURES

List of Tables

List of abbreviations

UC - Unicellularity
MC - Multicellularity
MTE - Major Transitions in Evolution
GRN - Gene Regulatory Networks
SPS - Stochastic Phenotypic Switching
DPM - Dynamical Patterning Module
ODE - Ordinary Differential Equation
CA - Cellular Automata
RD - Reaction-Diffusion
RAPS - Radially Averaged Power Spectrum
FFT - Fast Fourier Transform

Chapter 1

INTRODUCTION

1.1 Complexity, physics and evolution

When Charles Darwin first proposed a mechanism to explain how might *endless forms most beautiful* come to be, he was also cautious with the explanatory power of evolution. Facing a staunch religious opposition, he was quick to concede that some traits might remain unaccounted for within the canonical selection under variation paradigm. For instance, he though that some human traits like the eye and the mind were too complex and required something else [Darwin, 1859]. Contemporary scholars also voiced their concerns [Hull, 1973], for it was certainly difficult to grasp how such traits might have accrued from the kind of gradual change that Darwin was proposing. To Darwin, in an honest admission of ignorance, the origin of a complex eye was a potential counterexample of the view of evolution grounded on small, continuous changes:

> *To suppose that the eye, with all its inimitable contrivances for adjusting the focus to different distances, for admitting different amounts of light, and for the correction of spherical and chromatic aberration, could have been formed by natural selection, seems, I freely confess, absurd in the highest possible degree. Yet reason tells me, that if numerous gradations from a perfect and complex eye to one very imperfect*

> *and simple, each grade being useful to its possessor, can be shown to exist; if further, the eye does vary ever so slightly, and the variations be inherited, which is certainly the case; and if any variation or modification in the organ be ever useful to an animal under changing conditions of life, then the difficulty of believing that a perfect and complex eye could be formed by natural selection, though insuperable by our imagination, can hardly be considered real.*
>
> *Charles Darwin (1859, pp 186, 187)*

The observed sophistication of the "camera-like" eye appeared indeed too close to a rational design to be compatible with a blind watchmaker. It seemed, as Darwin himself calls it "an absurdity". This example has been repeatedly used to attack the likelihood of evolution in generating complex eyes and many misconceptions have been generated. Not surprisingly, only in recent years it has been possible to show that reasonable estimates of evolutionary rates along with an improved understanding of the phylogenetic relationships between types of eyes among major taxa is fully consistent with a gradual evolution [Dawkins, 1986, Gregory, 2008].

The sequential evolution of the eye as a continuous evolutionary process was described in [Nilsson and Pelger, 1994]. This is a theoretical -not a simulation model that starts from a more or less flat patch of light-sensitive cells and ends in a camera-like eye. The theoretical view is grounded in a sound set of assumptions, and is fully supported by different sources of evidence. Perhaps the more classical -but still powerful- facts support to this picture emerges from the observation that intermediate solutions (all of them) already exist in nature, as illustrated in Figure 1.1. Here several drawings corresponding to different known, current species of animals, from the limp (*Patella*, Figure 1.1B) to the vertebrate eye (cuttlefish, Figure 1.1F) involving increasing, connected levels of complexity.

Here we can add an additional layer of complexity to this problem, which could also be formulated as a potential flaw to the sequential picture. One can argue that we are ignoring the higher-scale phenomena taking place above the gene level. Even if we acknowledge that gene-level

INTRODUCTION

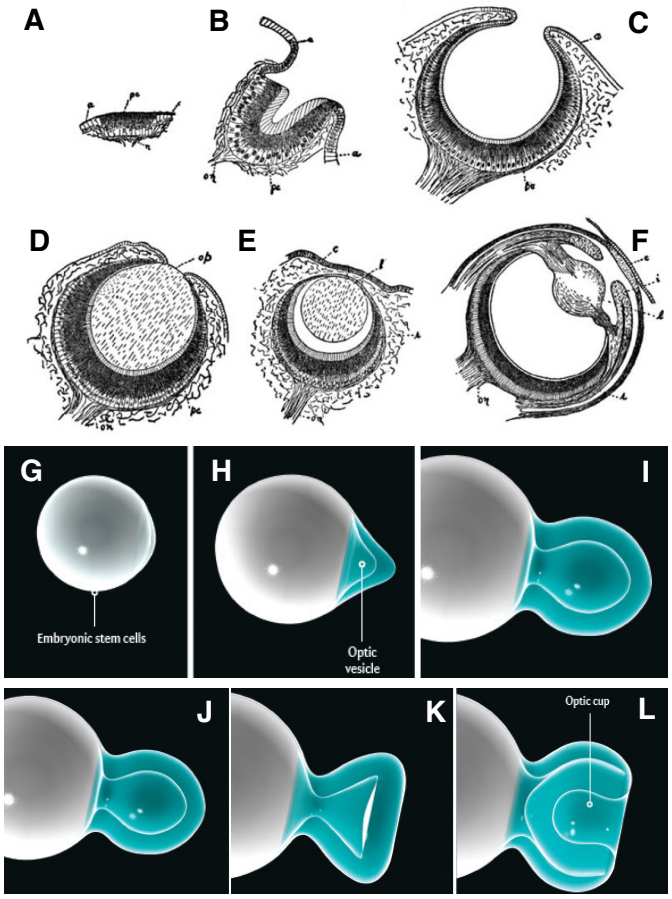


Figure 1.1: The evolution of eyes: from comparative anatomy to stem cells and self-organisation. In (**A - F**) we display several well known examples of the visual organs of different living species (redrawn from Gregory 2008). (**G - L**) Using stem cell-based methods, it has been shown that a whole eye cup gets formed when the right culture conditions are met (redrawn from Nakano *et al.* 2012). It was shown that a whole retinal structure gets self-organised in cell cultures, somewhat recapitulating the morphogenetic process that takes place during embryogenesis. These results indicate that an important part of the construction process is controlled by cell-cell interactions and boundary conditions beyond the purely genetic control metaphor.

changes are taking place over time, can we explain the actual nature of the development of an eye and how such spatial organization process has been shaped through evolution? Here too we could be trapped by an apparent complexity barrier: we need to explain how the eye gets formed as an embodied organ. But here again simple explanations and well-defined experiments help provide an elegant and robust answer. By using embryonic stem cells, it has been shown that they can differentiate under the proper molecular signals into retinal tissue, growing, deforming and shaping in a way that eventually ends in a well-formed optic cup, containing six layers of cells and (after two weeks of development under artificial conditions) resembles like the eye of a developing embryo [Nakano et al., 2012].

Such an achievement as special relevance in our understanding of evolution, and its basis are an essential component of this dissertation. Because cells communicate and interact, changing their local environment and creating feedback loops between gene networks and boundary conditions, self-constructing processes represent a key component of morphogenetic patterns and their evolutionary paths. The fact that such spontaneous processes of pattern formation pervade all similar experiments involving tissue dynamics and organogenesis is a consequence of very universal principles of organization. They provide additional evidence that the eye, as well as other complex biological structures, are largely to be expected.

Darwin's concerns about the evolution of complex organs are far from irrelevant. Natural selection is indeed a major force shaping the evolution of natural populations, and nothing would be really understood in many well-established fields unless under the light of evolution. This is the case, just to give an example, of infectious diseases and the major role played by evolutionary dynamics in the propagation and adaptation of microbes and viruses to their hosts. But it is also known that rapid changes occur through evolution, which appears to be sometimes *punctuated* by episodes of innovation. Partially because of these observations [Eldredge, 1972, Gould and Eldredge, 2000] a new formulation of evolutionary theory, to be named *macroevolution* was proposed. This new field of knowledge was tasked with characterizing common features of

evolution across species, using longer timescales, and remaining largely oblivious to the specifics happening at the molecular level (*microevolution*).

The distinction between micro and macroevolution is nowadays still under discussion. It has been argued that changes in the allele frequency are the ones that ultimately generate, when accumulated over time, the kind of patterns that are observed in larger scales. One could say, for example, that the complex eye is an example of this multiple rounds of microevolution eventually originating an apparently "new" structure. However, many relevant phenomena contradict this assumption [Erwin, 2000, Myers and Saupe, 2013]. The field of evo-devo as well as the realization that species can have a huge ecological -and evolutionary impact due to their role as ecosystem engineers [Jones et al., 1994, Jones et al., 1996] have modified our views of evolutionary patterns and processes.

One of the entrenched concepts since classic times [Lovejoy, 2011] that macroevolution had to deal with was the *scala naturae* (Figure 1.2A), the idea that the natural world could be ordered in a hierarchy of complexity with humans naturally sitting on top of it. Although usually invoked in a beliefs-based setting to justify divine order or the rights of humans over the rest of species, it is undeniable that our modern biosphere presents many more species and elaborated structures than previous epochs, and presumably at the origins of life in our planet as well.

Part of the problems generated by the development of a theory of evolution is clearly associated to multiple scales of complexity. This has led to some authors [Eldredge, 1985] to propose the need for a hierarchical theory of evolution. The definition of *complexity* is itself subject to debate and often presented differently in different case studies. Within complex systems, a generally accepted definition makes use of the presence of emergent phenomena. Specifically, a complex system is characterized by the presence of new, higher-order qualitative features that result form the interactions of smaller-scale entities but that cannot be reduced to the properties of such entities. Ant colonies versus ants or neurons versus brains are two celebrated examples [Miramontes et al., 2001]. In our context, the emergence of multicellular systems and/or pattern formation

processes beyond the cellular level would correspond to this definition.

Lamarck had been one of the first to adventure a scientific hypothesis: the evolutionary process was seemingly a directional ladder, giving rise to increasingly complex species from simpler ones due to the acquisition of traits in his proposed evolutionary theory [Gillispie, 2006]. But within Darwin's framework and its current incarnation, the modern synthesis [Huxley et al., 1942], mutations are considered random, unbiased with respect to complexity.

More recently, several authors have attempted to reconcile these apparently contradictory facts in all encompassing, macroevolutionary theories. The value of such theories resides in not only offering a better understanding of how evolution works but also in providing a satisfactory explanation of why complex creatures are allowed to exist and additionally, if complex life should be expected wherever there is an evolutionary process. In the next sections we will review a non exhaustive list of some of the most relevant.

1.1.1 Gould's statistical approach

In Stephen jay Gould's book *Full House* [Gould, 2011] the famous paleobiologist argued that simple organisms have always dominated the biosphere in terms of cell number and biomass, and that it is only due to a human bias that we have come to regard complexity as a dominant feature of evolving systems. This bias, he exposes, has two origins: limitations in our sampling of species and active anthropomorphic bias. The first one is a consequence of directing the analysis of this trend towards animals and plants, beings that leave traces in the form of fossils that can be easily studied. While most of the earth biosphere, including *eubacteria*, *archea*, viruses and small or soft bodied eukaryotes, go largely unnoticed. The second one, he argues, comes from our own desires to remain on top of the *scala naturae*.

By considering only extreme species (those that fall into the end of the complexity spectrum) instead of average complexity, we make a systematic error in assessing the real complexity of the natural world. A

INTRODUCTION

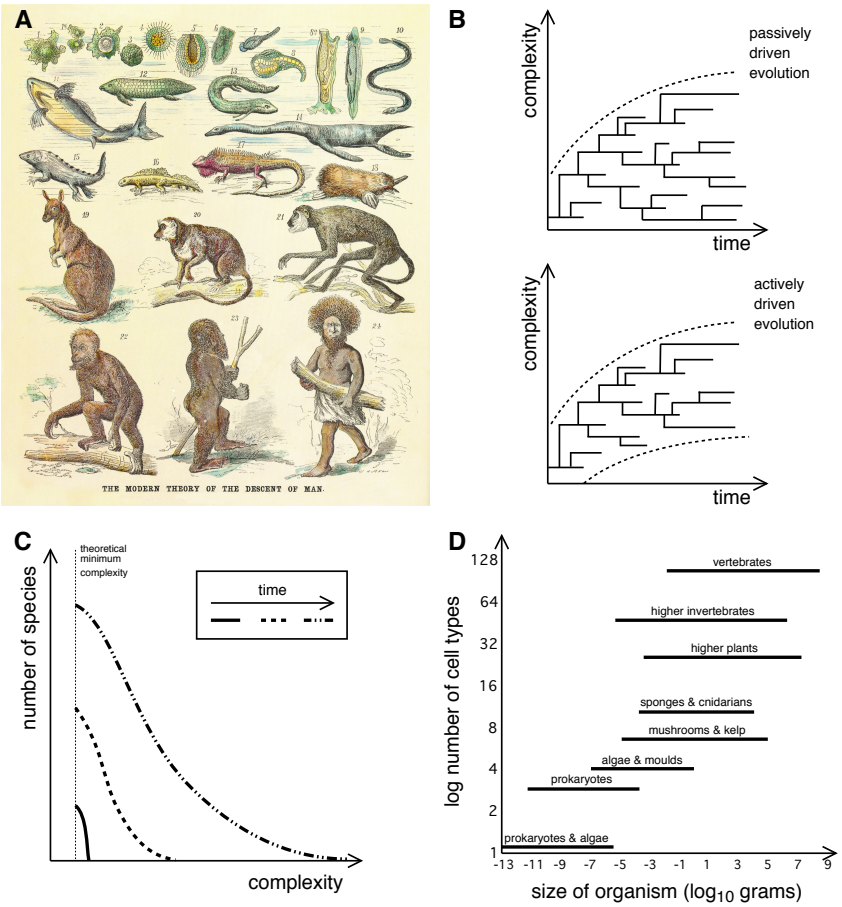


Figure 1.2: Depictions of the evolution of biological complexity. **(A)** Engraving by Haeckel [Haeckel and Lankester, 1883] showing the *scala naturae* or the hierarchical ordering of the natural world from a unicellular amoeba (top) towards man (bottom). **(B)** The role of minimum complexity in identifying if the trend is passively or actively driven according to McShea. In an actively driven scenario both maximum and minimum complexity increase. **(C)** The perspective taken by Gould: as time passes average complexity increases due to the minimum threshold below which cells cannot exist. **(D)** Positive correlation between size and number of different cell types across taxa, Bonner's proposed surrogate measure of complexity (redrawn from bonner 1988).

particularly vivid example of this error is the naming of ages in the history of the biosphere. According to Gould the age of fish or the age of dinosaurs is not an accurate description of the existing biota, but a much more enticing name than a prolonged age of bacteria. However, Gould acknowledges that the average complexity has increased over time, and proposes a mechanism to explain it.

First, he shows that for specific taxa, paleontological analysis reveals that the relation between parent and daughter species' complexity -using shell size as a proxy of morphological complexity- follows a distribution centered at zero. That means, species will generate a smaller sized daughter species and a bigger one with equal probability. This establishes a clear connection between the unbiased mutations at the molecular level and unbiased speciation at the macroevolutionary level.

Then he proposes that a minimum complexity threshold for life exists. A lower bound for structural organization below which cells cannot consistently reproduce and reliably transmit their information to the next generation. Given this constraint, as time progresses new species arising from those close to the boundary are effectively filtered towards higher complexity, and average complexity increases (Figure 1.2C).

According to other authors, Gould's perspective is not devoid of shortcomings [Bonner, 2013]. Specifically, that given the author's attention to the source material to analyze the evolution of complexity, solely using the fossil record of specific taxa is a contradiction. Moreover, some others have exposed that even slight deviations from this zero centered distribution would accumulate over time and generate one of two outcomes: either a world with ever increasing complexity or a world with the most simple organisms, thus suggesting that Gould's view is simplistic.

Finally, I think that ultimately Gould's proposal fails to recognize that there are also physical constrains to the evolution of larger organisms, forming a sort of soft wall in the higher end of species complexity spectrum. Such physical constraints are difficult to observe in the ranges of variation that Gould discusses, but are specially relevant when considering the evolution of new modes of organization [Lavrentovich et al., 2013]. For instance, consider the case of the transition from solitary to gregar-

ious life. Let us and assume that cells feed by transport of high energy chemicals from the environment. For the sake of simplicity, we can consider that cellular volume is closely related to the amount of biomolecules a cell needs to implement function and thus proportional to the energetic cost of living. Following these assumptions, it is clear that the ratio of membrane area to volume of cells is of outmost importance to determine if a cell will be able to grow and divide, for cells will capture nutrient proportionally to their membrane area and will dissipate energy proportional to their volume. Then, if groups of cells develop, the area to volume of the aggregate decreases and so does the average intake per cell supposing a perfect sharing of resources. If these newly formed groups compete with solitary cells for the same resources, they must face a competitive disadvantage, for they surely grow slower and perhaps be outraced by their unicellular siblings. In mathematical terms, this translates to the area of an aggregate scaling with the radius squared while volume scales with an exponent of three. This rapidly becomes unsustainable, new modes of feeding exclusive to clusters must appear, which will enable the existence of bigger aggregates and ensure diminished competition with unicellular relatives.

1.1.2 McShea's zero force evolutionary law

In several articles [McShea, 1996, McShea, 2000, McShea, 2002] and specially in his book *Biology's first law* [McShea and Brandon, 2010], McShea suggests a mechanism for the progressive trend based also on the random nature of the evolutionary process. He proposes a Zero Force Evolutionary Law, akin to laws available in physical systems since the times of Newton. This law states that in the absence of other processes -such as selection- complexity and diversity have a tendency to increase in evolutionary systems.

McShea distinguishes between two definitions of complexity, the colloquial and the "pure". The colloquial meaning of complexity, he argues, is casually related to being structurally intricate or functionally integrated and bears no relation to his law. "Pure" complexity on the other hand, has

a much narrower meaning: the variation and interconnectedness of biological parts. By parts he means any stable biological component: from proteins to genes to organelles and beyond, extending to super-cellular entities. This second definition, which has a clearer meaning, is the variable of interest to the zero force evolutionary law.

His approach resembles that of Gould's in that both require neutrality (or the absence of selective pressures) to attain a net gain in average complexity. Yet differs precisely in what constitutes an increase in complexity, for Gould's proposed metrics are more related to the colloquial meaning (an increase in size doesn't necessarily incur in a diversification of parts). For both, an essentially random drift in terms of complexity defines the basal trend upon which selection acts.

Following this argument, McShea states that the paradigm shift established by punctuated equilibrium [Eldredge, 1972] has to be regarded in the opposite direction. That is, the bursts of diversification are to be expected but the periods of stasis are the ones that do not conform to the null model. Given that the periods of stasis are the norm according to the paleontological records, a serious critique of McShea's proposal could be made in that organisms rarely seem to be free of purifying selection.

A very relevant contribution of the author to the evolution of complexity issue, is the identification of minimum complexity as the key property in characterizing the trend. Specifically, he argues that both in neutral and selection driven frameworks, maximum complexity continuously increases but whether minimum complexity remains stable or not is of critical importance in discerning which one is closer to reality (Figure 1.2B).

1.1.3 Bonner's size dependent niches

Bonner's take on the subject -exposed over an extensive scientific career [Bonner, 1952, Bonner, 1965, Bonner, 1988, Bonner, 1998, Bonner, 2009a, Bonner, 2009b, Bonner, 2013]-, is that an active selective pressure for more complex organisms does not exist and that increased complexity is side product of selection acting in other organism properties. He also argues that instead of the organism size used by Gould, the number of

different cell types is the appropriate parameter to characterize (morphological) complexity.

Regarding size, Bonner analyzes Cope's rule -a law stating that in certain vertebrate lines there is a tendency towards larger size over geological time-, showing evidence of a positive trend in maximum length and weight not only in vertebrates but across taxa. He proposes that natural selection does favor increased size in order to conquer new environments, reduce competition or enforce ecological interactions. Bonner exemplifies this last item by stating that there are ecological roles that can only be performed with particular sizes, such as a predator being necessarily larger than its prey.

Moreover, he exposes that there is a positive -albeit rough- correlation between the size of an organism and the number of different cell types it has (Figure 1.2D). In contrast to the data used by Gould, this trend extends several orders of magnitude and encompasses organisms that we would consider to have *a priori*, radically different complexity (see Figure 1.2D axes and data labels).

According to Bonner, when size increases some cells can become redundant. For instance, they might be no longer exposed to the expected environments and thus might never perform their expected metabolic "duties". A simple example of this proposition can be found in growing avascular masses [Ward and King, 1997, Byrne and Chaplain, 1997], in them cells located the core of a cluster do cease to receive the oxygen and other necessary substances for survival after size reaches a particular threshold. Organisms can seize this opportunity by repurposing such cells with other functions. This introduces new relations between the cells forming the growing aggregate, which is a more important feature than simply scaling up with regards to complexity in Bonner' view. Morphological complexity then, is enabled by increased organismal size.

The possibility of size-dependent fitness has also been explored by other authors. Proposed benefits for organisms with bigger sizes include protection from environmental fluctuations [Gerhart et al., 1997], cooperative metabolism [Dworkin and Bonner, 1972, Koschwanez et al., 2013, Pfeiffer and Bonhoeffer, 2003, Schirrmeister et al., 2013], and evasion of

predator feeding mechanisms [Stanley, 1973, Boraas et al., 1998].

1.2 The major transitions in evolution

Another possible approach to study the evolution of biological complexity is directly analyzing what Gould deemed the high end of the spectrum, that is, to give an account of the evolutionary histories that led to the rise complex organisms. Such undertaking is necessarily more limited in scope, but avoids the burden of having to explain patterns across the whole biosphere. As exemplified by Bonner with his proposed size dependent niches, to assume that the same selective pressures are at work across scales may do a disservice to our capabilities to theorize. This perspective has been taken by several authors, most notably by Stebbin [Stebbins, 1969], Lane [Lane, 2010] and Maynard-Smith and Szathmáry [Maynard-Smith and Szathmary, 1997], who have tried to identify the key innovations in the history of life. By their own nature, these transitions have to be seen as true discontinuities in the otherwise incremental (episodic or not) process of evolution. The deal with truly qualitative novelties, whose nature cannot be reduced to a simple addition of previously defined features. In this context, the study of this innovation events can be considered a theoretical attempt to understand how biological complexity originates.

The lists of events (see below) and the approach taken by different authors differ greatly. Particularly, Lane's list includes events considered "game changers" in evolution, i. e. innovations that forced other biological entities to adapt into complete new landscapes. These seemingly bear no special relation to biological complexity -at least as traditionally understood- but indirectly forced new ways of interaction among organisms and species. Some of his examples are similar to those discussed in the first section, like the evolution of the camera-like eye, which have puzzled countless generations of naturalists and biologists. In contrast, Maynard-Smith and Szathmáry focus in providing insights into the emergence of new levels of organization. They propose the Major Transitions in Evolution (MTEs), which are characterized by severe changes in the information flux between biological entities. Their approach establishes a clearer framework to study the major increases of biological complexity,

and provides guidance to find new MTEs [Szathmáry, 2015].

Event	Description
Origin of Life	The transition from dynamical to information bearing chemistry which could be subject to evolution.
DNA	The invention of a molecule in which information could be reliably stored and retrieved.
Photosynthesis	The innovation in metabolic pathways to collect energy from photons.
Complex cells	The endosymbiotic origins of eukaryotic cells.
Sex	Recombination of genetic information.
Movement	Active motion in predation and evasion.
Sight	The increased awareness of animals by capturing light.
Hot Blood	The invention of homeostasis in body temperature.
Conciousness	Self-awareness and the "private theatre".
Ageing	The evolution of programmed death.

Table 1.1: List of the ten greatest innovations in the history of life according to Lane.

The MTEs provided by Marynard-Smith and Szathmáry can be further separated into two lists. The first, including the evolution of DNA, sexual reproduction and language, has to do with how information gets reliably transmitted between biological entities. The second list, harboring the rest of transitions, contains a hierarchy of components coming together to form new biological entities.

This second class of transitions presents the emergence of collective fitness from previously independent replicators. This is the case of chromosomes, eukaryotes, multicellular beings and societies. On the brink of these MTEs, isolated entities are the bearers fitness in their own right.

Event	Description
Origin of Life	From coupled catalytic sets to the first compartment.
Chromosomes	The Transition from isolated information molecules to assemblies of genes.
DNA	The division of labor between the genome and the proteome.
Complex cells	The endosymbiotic origins of eukaryotic cells.
Sex	Recombination of genetic information.
Multicellularity	transition from solitary to gregarious life.
Societies	From solitary individuals to colonies.
Language	From primate to human societies.

Table 1.2: The major transitions in evolution according to Marynard-Smith and Szathmáry.

However, after the transition, the probability of survival and reproduction are heavily reliant on the properties of the collective. This arises unsolved questions regarding how is this fitness transferred between the two entities, whether it is smooth or sudden, and what are the challenges that collectives face when building a new subject of fitness.

When taken together, all instances of MTE seem to profoundly modify the basic elements of evolutionary theory, namely: how information is stored and transmitted to the next generation (inheritance), how modifications are introduced to the genetic material (variation) and who is the subject of fitness (selection). Accordingly, these MTEs define singularities in the evolution of life when the very rules that conform the evolutionary process were changing. It is also worth mentioning that several MTEs (if not all) can be seen as instances of what we know as "phase transitions" in physics [Solé, 2012]. These type of collective phenomena, where new global phenomena emerge as a consequence of interactions

between locally related objects -atoms, molecules, ants or stock brokers- are related to a new form of information propagation beyond the level of the basic units that results into large-scale structures and processes [Miramontes et al., 2001, Ball, 2011]. Perhaps not surprisingly, Szathmáry himself, while reviewing his early work on MTEs concluded that

> *Theories do not have to be predictive but still can have considerable explanatory power. After all, the predictive aspect of evolutionary biology as such is limited as well; and this limitation especially applies to the quantitative aspects. There are two questions that one can raise: (i) Is it possible to tell whether a lineage or a small set of lineages have transited to 20% or 90%? I think this question can be answered in the future if one can show that the evolutionary dynamics of transitions has something in common with phase transitions in physics.*
>
> *E. Szathmáry (2015)*

The search for universals in physics has been a success story: key fundamental regularities have been shown to pervade the complexity of many disparate systems. Within complexity theory such perspective has been also followed by several authors, who had suggested that at least some key features of biological complexity might also display universal properties independent -at least to some extent- of the specific nature of their underlying components [Kauffman, 1993, Miramontes et al., 2001, Bak, 2013].

1.2.1 Fraternal and egalitarian transitions

As argued by Queller [Queller, 2000], the route to these new bearers of fitness has two main paths: the egalitarian and the fraternal. Such distinction has to do with the identity of the lower level particles beginning to form a collective. In the first case, the units that come together are not

alike. This enables greater potential to conflict but also the possibility of increased combination of functions. Examples of egalitarian MTEs are the emergence of the first cell, chromosomes and eukaryotes. In each of them, different individuals that had the capacity to reproduce on their own melded into a new assemblies with their combined capacities.

In the case of fraternal transitions -MC and societies-, conflicts are diminished by the similarity of the constituent particle [Michod, 2000, Bonner, 1998, Grosberg and Strathmann, 2007], but unless division of labor is established no gain in diversification of function is introduced. However, we observe that both societies and multicellular creatures can make use of division of labor [Buss, 2014, Gibbs and Martin, 1962], which introduces the issue of creating and coordinating differential activities. In MC, this entails increased costs for the individual, as well as innovation in the areas of communication and the genetic architecture able to sustain discrete cellular states.

Even after accounting for the reduction in conflict just discussed, fraternal transitions are not entirely devoid of strife, specially when considering the role of cheaters in a forming group. When collective activities require the participation of different elements to achieve a common result and this participation has a cost, individuals can sometimes benefit from letting others do the collective duties. The cells that defect in such cooperative efforts are called cheaters, and represent a major hindrance to the evolution of cooperative behavior [de Vargas Roditi et al., 2013, Okasha, 2005].

1.3 The emergence of multicellularity

> "Now this is the law of the jungle,
> as old and as true as the sky,
> And the wolf that shall keep it may prosper,
> but the wolf that shall break it must die.
> As the creeper that girdles the tree trunk,
> the law runneth forward and back;
> For the strength of the pack is the wolf,
> and the strength of the wolf is the pack."
>
> *The Second Jungle Book* - Rudyard Kipling

The transition towards MC entailed new relations among previously independent units [Buss, 2014, Maynard-Smith and Szathmary, 1997]. As has already been mentioned, the process by which new levels of organization are formed pose serious questions to our understanding of evolution as a purely selfish process. Specially in MC, since it intrinsically implies the formation of new cooperative bonds between entities that can survive and reproduce on their own.

The case of MC is also a remarkable example of convergent evolution [Morris, 2003] in the establishment of cooperative behavior because it is thought to have evolved independently in at least 25 different lineages [Grosberg and Strathmann, 2007, Niklas, 2013]. This collection of events resulted in the myriad of organisms with exquisitely elaborated structures that we observe today (Figure 1.3). However, the origins of multicellular behavior are thought to be much more humble, likely consisting of simple clusters with weak interactions between cells [Bonner, 1998, Boraas et al., 1998, Kirk, 2005b, Pfeiffer and Bonhoeffer, 2003]. After

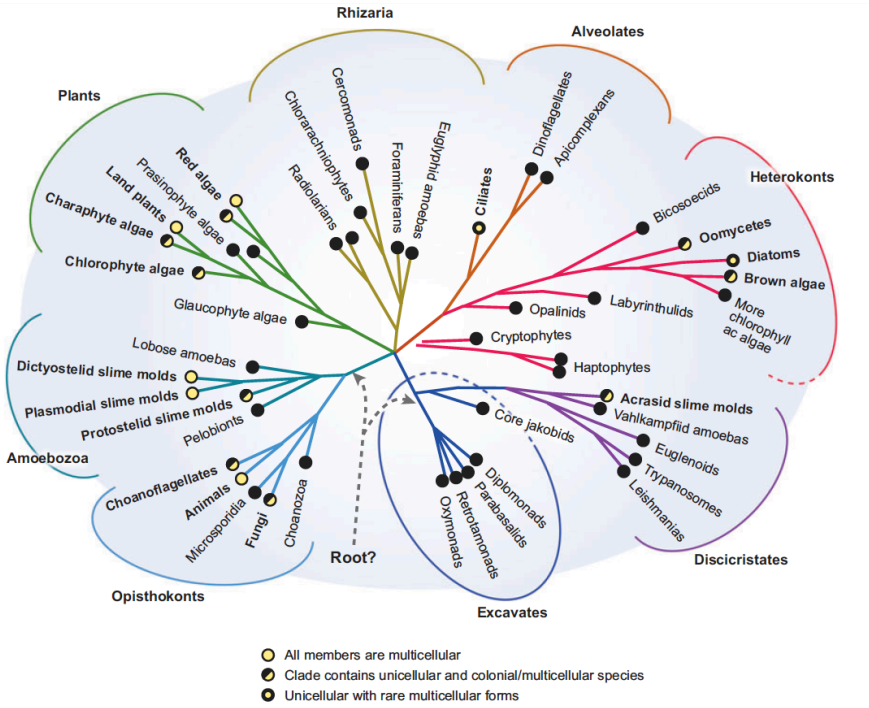


Figure 1.3: Convergent evolution of multicellularity in 25 different lineages. Multicellularity appears to have raised only once in Metazoans, but multiple independent examples -and losses- exist for plants, fungi, and the Eubacteria (from Grosberg and Strathmann, 2007).

cells were able to coexist in an aggregate form, a shift in the level of selection took place [Damuth and Heisler, 1988, Michod, 2005], due to the emergence of group-level traits relevant to the survival and reproduction of both the cellular and collective levels. Nonetheless, this process remains poorly understood, as the last natural transition towards MC happened more than 200 million years ago [Grosberg and Strathmann, 2007, Herron et al., 2009].

To unravel the mysteries of early forms of multicellular organization, some clues have been sought in the study of the fossil record. In particular

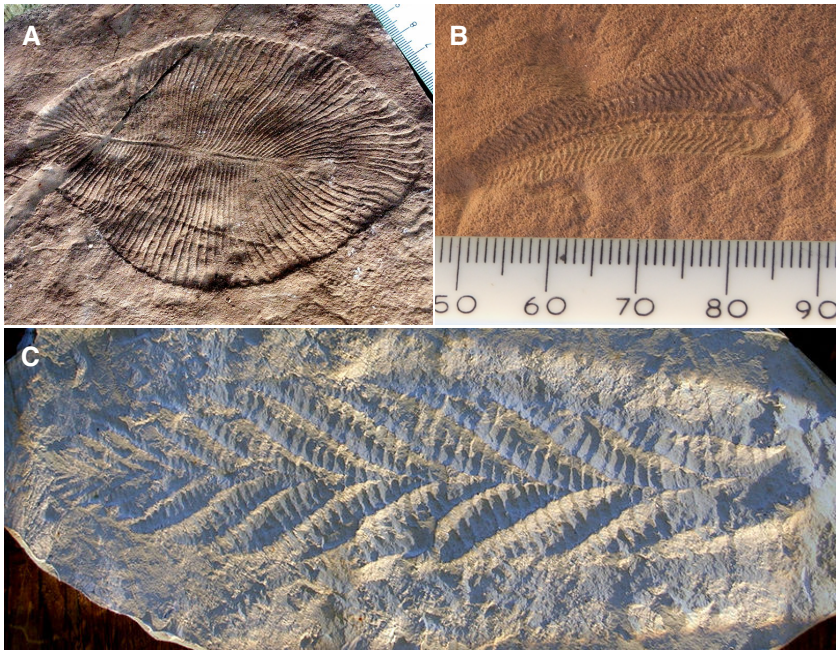


Figure 1.4: Extinct examples of "simple" multicellularity from the fossil record of the Ediacaran period. *Dickinsionia costata* **(A)**, *Spriggina floundensi* **(B)** and *Charnia masoni* **(D)**. (Images courtesy of wikimedia commons).

the scientific community has been deeply fascinated with the Cambrian period, which spans from roughly 540 to 450 million years before present. During this period, a remarkable explosion in morphological diversity is observed, with all currently existing animal body plans originating in a brief interval of approximately 20 million years [Rokas et al., 2005, Morris, 2006]. This has been related by some authors to the dramatic increase in atmospheric concentration of oxygen that took place during the Proterozoic eon [Payne et al., 2009, Bonner, 2013], which supposedly enabled the more dynamic lifestyles of burrowers and predators that appeared during the Cambrian (see also section 1.4.3 for other causes attributed to the Cambrian explosion of diversity).

The mechanisms that facilitated such surge in complexity are certainly enticing, however, the organisms that existed during the Cambrian period were fairly complex to begin with and may not convey a faithful picture of the origins of animal multicellularity. An earlier period, the Ediacaran (635 to 540 millions of years before present), shows a different scenario with much more simple modes of organization. The reconstructed ediacaran biota (Figure 1.4) contains organisms with much more modest features, nonetheless being the first known examples of radial and bilateral symmetry in animal body plans. Going back a little further, the beginning of Ediacaran period was the time when the last common ancestor of all multicellular animals is thought to have existed [Medina et al., 2003, King et al., 2008, Newman and Bhat, 2008]. According to phylogenetic analysis, the first example of animal MC likely possessed non-robust developmental processes, forming sheet-like structures [King et al., 2008].

In this quest for the origins of MC, the study of extant multicellular organisms has also provided relevant insights, especially through phylogenetic analysis. Some current MC organisms are also characterized by simple modes of organization and offer a powerful window into the establishment and implementation of basic collective traits. A remarkable example can be found in cyanobacteria (Figure 1.5), extensively used as models organisms of simple MC and pattern formation. This includes the paradigmatic genus *Volvox*, which builds nested spherical structures, containing the next generation within the current one (Figure 1.5D). In this manner, the progeny is protected from the external conditions until it has matured. Then the "old body" bursts, freeing several propagules at the same time. *Volvox* is also specially interesting for it provides, with the evolutionary history of the closely related species that constitute the genus, a close reconstruction of the processes that led to the establishment of this peculiar kind of reproduction. With the help of phylogenetic reconstruction, a historical gradient of morphological complexity -colony size, structure and cellular specialization- has been reported [Kirk, 2005a, Solari et al., 2006, Herron and Michod, 2008, Coleman, 2012], which can shed some light into the sequential acquisition of traits that encapsulate the transition to complex MC.

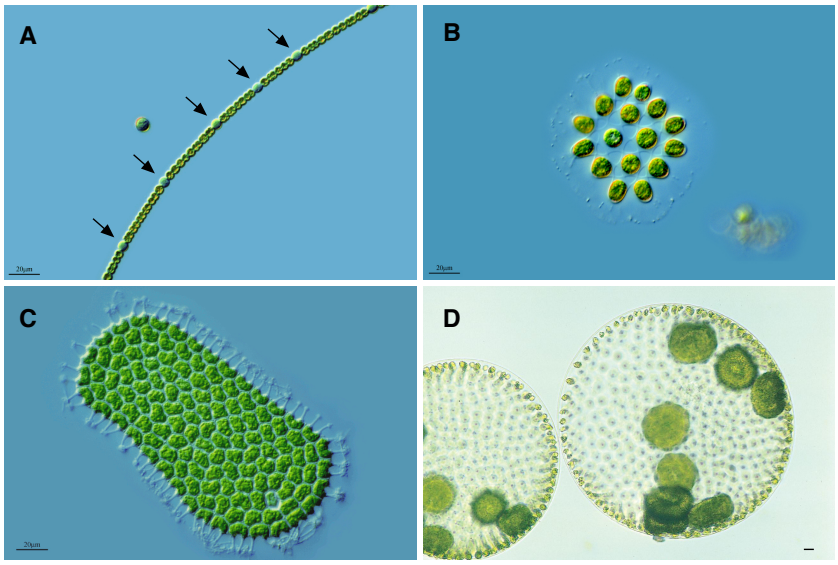


Figure 1.5: Extant examples of "simple" multicellularity in cyanobacteria. *Anabaena circinalis* **(A)**, *Gonium formosum* **(B)**, *Pediastrum ellipticum* **(C)** and *Volvox globator* **(D)**. Notice how the differentiated heterocysts in **(A)**, marked with arrows, are regularly spaced across the filament. (Images courtesy of Antonio Guillén of the Water Project).

Despite the invaluable information fossils and extant MC organisms provide, the picture will not be completed until a case of *de novo* evolution of MC is thoroughly examined. However, the chances of first handedly observing such process are truly slim. This necessity might be partially overcome by the use of artificial methods, both *in vivo* and *in silico*, which are discussed at length in the following sections of this thesis. Despite their intrinsic limitations, they provide the tools to systematically explore the feasibility of the transition towards MC. By examining the conditions in which MC arises artificially, a rationale for the natural cases can be proposed, hopefully identifying along the way common features pervading the collection of events that constitute the emergence of MC.

1.3.1 What is multicellularity?

In order to proceed the discourse of this section, it might be of interest to pin down what MC actually means. This issue has been approached by different authors under many considerations and is not as straightforward as it seems. Even if multicellular organisms are, by definition, entities composed of more than one cell, not all cell clusters can be considered multicellular organisms. This stems from the fact that aggregates might not be always integrated into a coherent organism with coordinated responses, or might be so but only temporally before splitting up into separate unicellular beings. The most commonly alluded property, thought to underpin all multicellular beings, is that the multiple cells must display a group-level fitness, i. e. collective properties must be the main drivers of the chances of survival and reproduction of each individual cell. Then, the establishment of cluster-level adaptation can afterwards result in the evolution of increased complexity [Willensdorfer, 2009, Bonner, 1998]

Michod has explored the possibility of fitness coupling in several works, concluding that most likely the first multicellular constructs experienced this phenomenon by separating the reproductive and somatic functions [Michod, 2000, Michod and Roze, 2001, Michod, 2007]. He argues that, in particular, the abilities to duplicate and be motile are constrained by the existence of a single microtubule organizing center in eukaryotes, and cannot be simultaneously carried in a single cell. Michod proposes that this issue might be overcome in two ways: either cells alternate between these two necessary functions or division of labor takes place. This means that in colonial organisms some specialized cells take responsibility of keeping the cluster in the needed location -which is specially relevant in photosynthetic species- while others are in charge of creating new cells. A cell that only performs one of these functions would not be able to create descendants, and accordingly the fitness of a given cell in a colony depends mostly on the appropriate task allocation among the group.

Shapiro, on the other hand, has championed the idea that bacteria can also be considered multicellular [Shapiro, 1998]. Not only particular groups that are commonly accepted into the MC club like *Anabaena*,

Myxococcus or *Dictyostelium*, Shapiro suggests that the highly regulated collective activities that lead to the controlled growth of colonies can also be considered multicellular. These intricate patterns created by swarming and coordinated growth [Ben-Jacob et al., 1994, Woodward et al., 1995, Xavier, 2011] have puzzled biologists for decades, unable to produce an evolutionary rationale until relatively recent times [Boyle et al., 2013]. The group-level fitness in this case, does not come from a direct separation of reproductive and somatic functions, since all cells in these colonies reproduce, but from the fact that building a collective structure might be beneficial at both the individual and the group level. In particular, it has been shown that swarming strains grow faster and occupy more surface [Boyle et al., 2013] than mutated strain that have lost the necessary cell-cell interactions.

The philosophical perspective taken by Okasha "dismisses" the details of implementation and goes right to the source of fitness in evolving aggregates [Okasha, 2005, Okasha, 2006]. Okasha distinguishes between two types of MC, one -MLS1- characterized by the collective fitness being an average of particle fitness and the second -MLS2- displaying emergent properties and showing a collective fitness unrelated to individual fitness. A common classification related to Okasha's two types of multicellularity is the distinction between simple aggregates and "true multicellular" organisms. The first category typically groups organisms conformed by multiple cells that build aggregates in a reliable manner, but otherwise do not display a clear division of labor. An example being a biofilm of a bacterial species made of equal cells. True multicellularity on the other hand, is used for "higher" organisms with cell differentiation. In them, after countless generations of fine tuning and adaptation to the internal environment individual cell fitness is close to zero.

Queller and Strassman [Queller and Strassmann, 2009] provide another perspective into putative classes of MC. They propose that multicellular organisms can be singled out by increased cooperation among cells and the implementation of conflict mediation mechanisms that mitigate selection within an organism. This classification is not discrete, but organisms might be placed in a continuous cooperation and conflict me-

INTRODUCTION

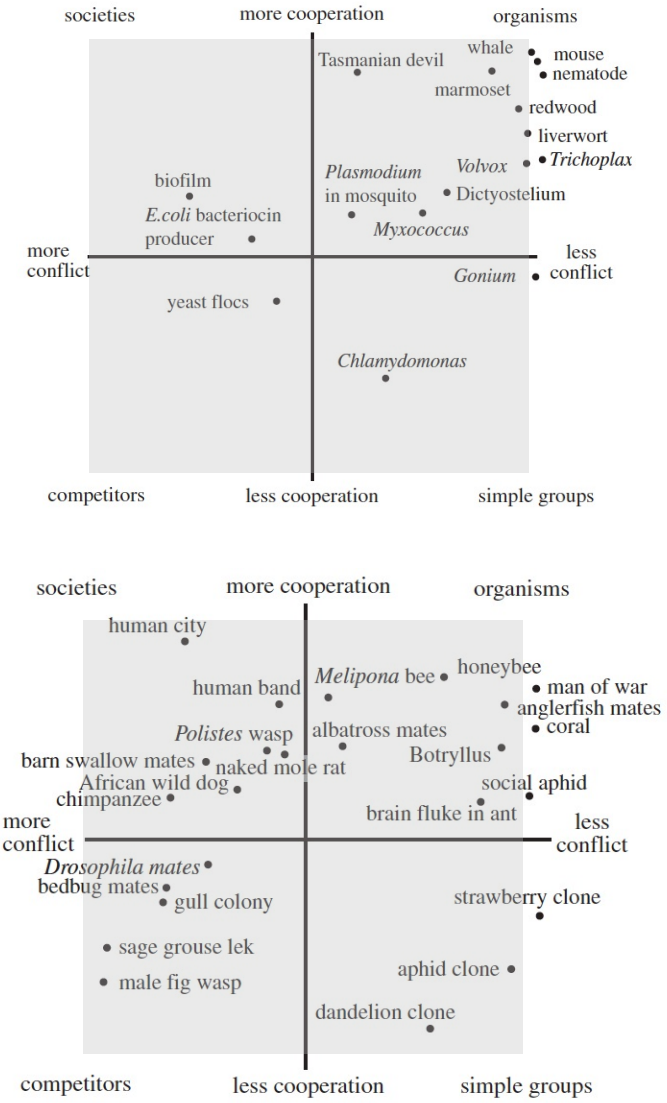


Figure 1.6: Cooperation, conflict and organismality. Top: The space of cooperation and conflict in groups of cells, at the bottom we show the equivalent diagram for multicellular groups (adapted from Queller and Strassman 2009).

diation plot (Figure 1.6). Four domains can be qualitatively distinguished in this plot: competitors, simple groups, societies and organisms. Given that some measures can be proposed for each axis, it would be interesting to monitor the evolution of MC in terms of paths taken in this plot, thus discover if there are preferred directions in which nascent groups are more commonly found to attain organismality.

1.3.2 Functional characterization of multicellularity

A complementary perspective can be taken by analyzing the functional characteristics of MC. Having defined in the previous section the characteristics of a multicellular organism at the theoretical or fitness level, we can now try to make an abstraction and reduce MC to a set of cellular mechanisms necessary to implement it. This approach can be found in the identification of multicellular *toolkits* [King et al., 2008, Sebé-Pedrós et al., 2011, Tweedt and Erwin, 2015]. That is, families of genes related to the implementation of the core mechanisms associated to multicellular traits, like master regulators of development (see section 1.4.2), adhesion molecules and proteins implicated in cell signaling. Moreover, this has enabled researchers to trace and pin down the origin of particular lineages [King et al., 2008, Sebé-Pedrós et al., 2010, Grosberg and Strathmann, 2007]. The interest of such work lies in yielding universal features and perhaps the identification of organisms that can make the transition towards MC.

In particular, it is thought that true MC beings display three key properties stemming from the fitness coupling of division of labor: multistability, communication and mechanisms for spatial organization (Figure 1.7). These are clearly not exclusive to the multicellular domain -as will be further discussed-, as they are also present in plenty of unicellular beings.

Multistability

First and foremost comes multistability, for division of labor is implemented through different cellular states, which are built from distinctive

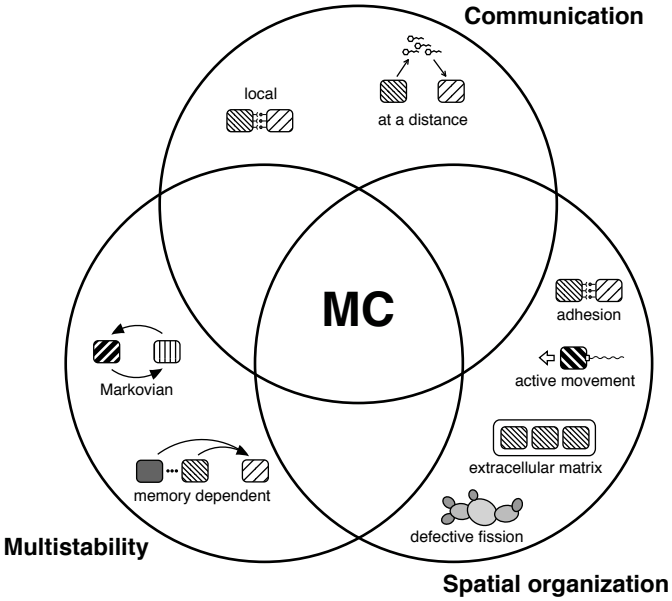


Figure 1.7: Functional characterization of multicellularity. All multicellular organisms display integration though signaling, mechanisms to regulate their spatial distributions and multistable phenotypes in the form of cell types, a particularly relevant feature for division of labor.

patterns of gene expression [de Mendoza et al., 2013]. These have to be enabled -and enforced- through gene regulatory networks (GRNs), whose topology establishes the possibility of mutually exclusive phenotypes. At the same time, GRNs enable the integration of information and computation, using the cellular internal state and the external environment.

Unicellular organisms also show distinctive patterns of gene expression, and accordingly different cell types -from resistance forms like spores [Eichenberger et al., 2004] to movement [Meeks and Elhai, 2002] and reproduction forms [Herskowitz and Oshima, 1981]-. But these are commonly not stable, that is, they fluidly change adapting to the environment. Contrarily, cells that differentiate in current multicellular organisms do not usually revert this transition. This is particularly relevant in animals

through the formation of cellular linages, in which cells are thought to further specialize and rarely regress to an earlier state. When they do so, it is either through a tightly regulated process to complete their life cycle -in the formation of gametes for instance- or because of a complete deregulation of the natural processes of differentiation -in cancer [Mayer et al., 1993]-.

An example of simple differentiation in unicellular prokaryotes is given by the so called phase variation. This phenomena is characterized by the ability of cells to transit between different phenotypes in a way that can be history independent -no memory- nor needs any external chemical cues. The stochastic phenotypic switching (SPS) implies that a population will always be composed of multiple phenotypes, which might be of interest to the population in case there are rapidly changing environments and adaptation is not possible. The rationale commonly given is that this mechanism increases the robustness of the whole: if some phenotypes are selected against and disappear, the original population can be regenerated by the surviving fraction, in a strategy that has been qualified as bet-hedging [Veening et al., 2008]. Phase variation is usually displayed in important features for the survival, from aggregation to antigenicity and pathogenicity [Henderson et al., 1999, Darmon and Leach, 2014].

Cell Signaling

Communication is also commonly regarded as an universal feature, since cell differentiation is normally coordinated for maximum collective benefit. By communication it is usually understood all known methods of information transfer, from mechanical contact and cytoskeleton alteration to long range signaling by diffusible molecules. Some of the hormones and signaling pathways that today are part of convoluted developmental processes are thought to be truly ancient. For instance, paradigmatic signaling pathways like Notch/Delta, JAK/STAT, TGF-β are thought to be present in the last common ancestor of all metazoans, since homologous sequences of can be found in sponges, placozoans and bilaterians [Tweedt and Erwin, 2015].

This is obviously not limited to animals, non-metazoan multicellular organisms also display mechanisms of information transfer. For instance, in the simple *A. circinalis* (Figure 1.5A), heterocysts are formed when there are low levels of nitrogen [Black et al., 1993]. The lack of such chemical is interpreted by cells as an indicator of the surrounding cells' state, and that specialization into a nitrogen fixing cell type is necessary. This creates the simple yet efficient regular distribution of heterocysts across the cyanobacteria filament, based on a lack of specific molecules instead of the presence of protein coded ones.

Unicellular bacteria also organize collective traits with the help of molecular cues. These serve the purpose of not only informing the surrounding population of a particular cell's state, but also as a way to determine the amount of individuals that might be present in the vicinity. The last case is perfectly exemplified by the so called *Quorum sensing* systems: bacterial signaling pathways regulating the expression and sensing of chemicals that diffuse outside the cell [Miller and Bassler, 2001]. In order to operate correctly, cells need to interpret the same molecules that are synthesizing, typically a family of chemicals with similar structure like *lactones* [Fuqua et al., 1994]. If the population density reaches a particular threshold some genes are upregulated, and a coordinated action like pathogenic activity or bioluminiscence begins. Moreover, this communication has also been found cross-species, with such mechanisms mediating complex ecological interactions [Waters and Bassler, 2005].

Spatial organization

Finally, all multicellular creatures show mechanisms of manipulating their spatial distributions to form aggregates in at least in a phase of their life cycle. This property is also thought to precede the origins of multicellular life, adhesion molecules being important for unicellular organisms in evading predators, acquisition of nutrients and spatial location in a particular niche [Newman and Bhat, 2008].

From the phylogenomic perspective it has been established that integrins, cadherins, fibrillar collagen and laminin likely constituted part of

the adhesion toolkit present in the last common ancestor of metazoans [Tweedt and Erwin, 2015]. These proteins are implicated in various paths to creating cell clusters, from cell-cell adhesions to the secretion of an extracellular matrix that envelopes and keeps together cells. Animal MC has been commonly related to cell-cell adhesion proteins, however, multiple simple cases of filamentous MC in cyanobacteria are actually held together by a tube-like matrix [Helm and Potts, 2012]. This is the case of the aforementioned *A. circinalis*, which creates long strings of cells only held together by a sugar structure. When the cilinder breaks, the filament splits, highlighting the important relation between adhesion processes and group-level division at the simple MC stage.

Another important mechanism to create bodily structures can be found in the aggregative development of slime moulds. These are free roaming unicellular amoeba that, when threatened by a lack of nutrients, collapse into spectacular multicellular structures with different cell types [Bonner, 2009b], all with the purpose of creating spores that might survive the adverse conditions. In order to do so, individual cells secrete a pulsing gradient of cAMP, which serves as chemoattractant to form the nascent multicellular structures. Chemotaxis might not be typically considered in the same category as adhesion molecules, but in the case of facultative aggregative multicellular organisms is a key mechanism in coming together.

Finally, a third route to be considered is simply the failure of separation between two daughter cells. This is not to be confused with formation of syncitiums -as occurs in other slime moulds like the genus *Physarum* [Salles-Passador et al., 1992]-, for syncitiums are constructed by just one membrane for a collection *nuclei*, i.e. they are not multicellular by definition. Contrarily, two cells created with a defective fission conserve their own membranes, but might get stuck with one another by the cell wall or membrane proteins.

1.3.3 Individuality and the levels of selection

A recurrent theme in MC is the concept of *individuality* [Okasha, 2006,

Michod, 2007]. This term refers to the property of collectives of discrete particles or *lower level* entities to act in a coordinated fashion, effectively creating a higher level entity upon which selection can act. This concept is closely related to the *organismality* defined by Queller and Strassman [Queller and Strassmann, 2009], which is tied to the transition from UC to MC.

As has been thoroughly stated by many other authors [Okasha, 2005, de Vargas Roditi et al., 2013], the emergence of this higher level entity is filled with challenges and pitfalls to the group cohesion. When performing a collective action and distributing tasks -for instance with multiple cell types- selection continues to apply to the low level particles besides the group level. The former has been commonly referred to as cell selection, and is the process behind both beneficial [Sprent et al., 1988] and detrimental [Cairns, 2006] developmental processes in animals. When selection is able to act at different levels or classes of entities, conflicts may appear, characterized by the contradictory interests of the lower and the higher levels of organization [Wilson, 1997, Michod and Roze, 2001].

A clear example can be found in the division of reproductive an somatic functions. In most MC organisms only a small fraction of cells are able to transmit their genetic information to the next generation, the germ-line. The rest of the body is made by the soma, lineages of cells that might divide in order to replenish or grow structures but whose information will never be part of the next generation. Not all MC organisms display this division of labor in terms of reproduction versus maintenance, so it is safe to assume that this property evolved over time [Buss, 2014]. The emergence of the germ-line soma divide must have entailed an important conflict, since forgoing the capacity to reproduce is a remarkable individual sacrifice for the good of the collective.

A particularly lucid account of this kind of interaction was articulated by Rudyard Kipling in the quote given at the beginning of this section. In the two last lines of his poem *the law of the jungle*, Kipling declares: *"for the strength of the pack is the wolf, and the strength of the wolf is the pack"*. In the context of multilevel selection this could mean that the success of the lower level entities is tied to the (hunting) strategies of the

whole, which are an emergent property of this new level of organization. At the same time the constituent particles have to submit to the collective, its "strength" depending on the coordinated activities of all individuals, which a rogue element can destroy by following it own interests.

To fully understand the process of hierarchical construction of new beings from an artificial perspective, a framework that allows for both levels to respond to selective pressures is necessary. Such approach must consider and offer opportunities to both levels to optimize fitness-relevant strategies. Failing to do so creates a rigged outcome were only properties of the whole or the parts are relevant to evolution.

1.3.4 Artificial frameworks to study the evolution of multicellularity

Artificial Evolution

Some clues to the origins of MC can be found -at least at the functional level- in living unicellular systems, such as bacteria or yeast (Figure 1.8). Many unicellular species display collective traits associated to the presence of signals, providing examples of coherent population responses [Shapiro, 1998, Ben-Jacob et al., 1994, Bonner, 1998]. Usually as a consequence of stressful conditions to the individual [Ben-Jacob et al., 2000, Marin, 1976], bacterial aggregates can display some degree of spatial organization and specialization. Given that these species already have some sort of integration among individual cells, can some externally set selective pressures force the emergence of collective fitness? And if possible, how do these laboratory conditions relate to the environments experimented by the cells that made the transition towards MC millions of years ago?

These are the kind of questions that can be approached from the field of artificial evolution, which seeks to characterize changes at the functional and genetic levels of organisms in controlled environments that impose unnatural selection pressures. In essence, only an initial population,

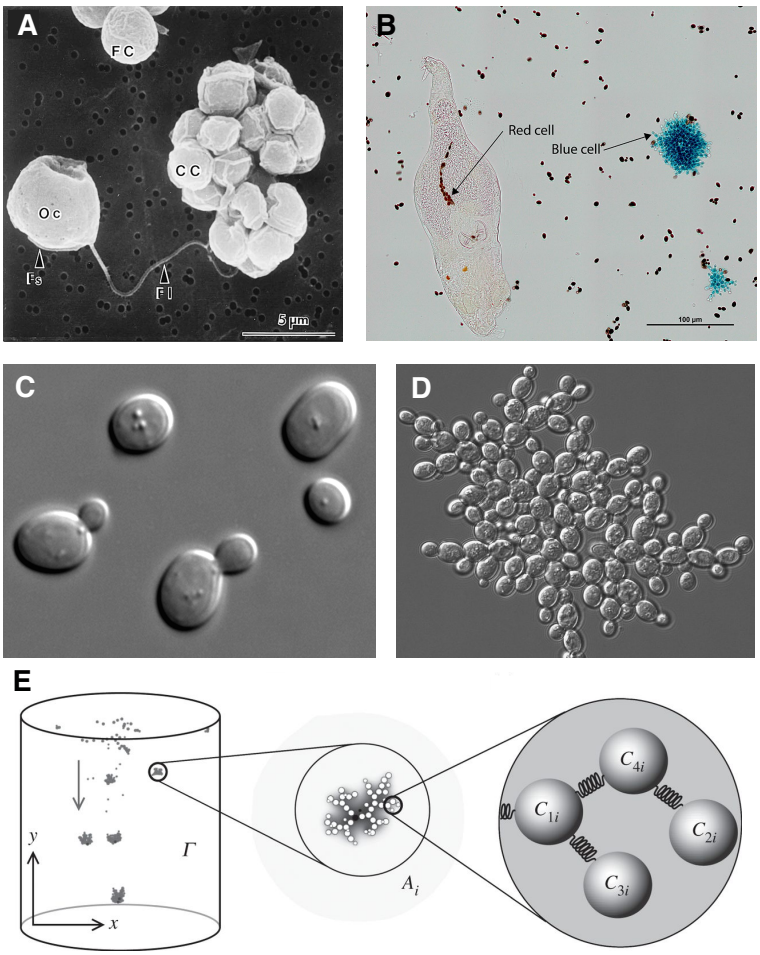


Figure 1.8: Experimental approaches to artificial evolution of multicellular systems provide evidence for a high potential of rapid development of these traits. In (**A - B**) artificial settings with enhanced ecological pressures -like predation- have been used to select for bigger aggregates (from Boraas *et al.* 1998 and Pentz *et al.*) In (**C - D**) we the show single, and evolved, multicellular snowflake structures in yeast derived from experiments that favoured the formation of aggregates by nature of enforcing a sedimentation process (from Ratcliff *et al.* 2012). These experiments provide an elegant framework to formulate and test simple, physically embodied models (**E**).

a selective environment and a method to quantify the changes or characterize new strains are needed. However, the key aspect -and most difficult part- has proven to be the appropriate choice of the selective environment, only limited by the researchers' ingenuity: from past earth conditions to different growth mediums.

Another important point in the design of artificial evolution experiments is the choice of an appropriate temporal scale. The duration can wildly variate from a few iterations of growth [Ratcliff et al., 2012] to thousands of generations [Lenski and Travisano, 1994] and, more importantly sets a limit in the amount of variation that the evolutionary process can generate. However, if the new environment imposes a rather strong selective pressure, strains can be rapidly substituted [Ratcliff et al., 2012, Boraas et al., 1998]. After each iteration of selection, the evolutionary change is usually quantified at the genetic level by means of whole genome resequencing [Elena and Lenski, 2003], or alternatively, a functional analysis if the new strains display clearly distinguishable phenotypes.

The use of successive rounds of mutation and selection to optimize biological processes like metabolic pathways or growth rates is not new. This has usually been achieved in the context of biotechnology in order to explore in parallel multiple variants, which might improve the economy of a biotechnological enterprise. In these attempts at artificial evolution, selection was applied by the researcher, who biochemically evaluated the fitness of the organisms as the increase in biosynthesis of a molecule of interest or in growth rate. In order to speed up the process, chemical agents that induce random mutations have been commonly used.

More recently, speeding evolution by directing mutations to specific regions of DNA has been enabled by different technologies. One of such is the MAGE process developed by Church [Wang et al., 2009]. MAGE stands for multiplex automated genome engineering, and enables precisely that: the systematic modification of different genomic elements in populations of bacteria. These modifications can be directed in location but random in terms of content [Wang and Church, 2011]. This speeds up directed evolution when there is a clear picture of the key genes related to the desired phenotype, but there is no knowledge *a priori* of which

mutations might be beneficial.

In the case of the transition from UC to MC two main approaches have been successfully used to attain aggregates that are fitness coupled: the use of predators and the use of gravity. The former is a very simple setup containing a unicellular prey population and a predator, which can be unicellular or multicellular. The gist of the experiment lies in the incapacity of the predator to consume bigger prey (Figure 1.8A,B). Accordingly, the prey population may display an evolution towards increased size -using aggregation or defective fission- due to the selective advantage of not being efficiently predated, even if clusters of cells grow slower.

Some examples of such strategy to the artificial evolution of MC can be found in the literature [Boraas et al., 1998, Pentz et al., 2015]. In the case of Boraas *et al.* the authors explored the issues an algae face by being cultured with its predator. In essence, what they found is that bigger clusters of the prey began to appear and settled as the dominant variant in this toy ecology. This is specially interesting for it links ecological relations to the evolution of complex traits, similar to the theories of Bonner presented at the beginning of this thesis. If such relation was found to naturally exist, ecological (interactions between species) and organismal complexity (interactions between cells in an organism) might form a closed loop, each one potentially driving the development of the other.

The gravity approach is based in the increased sedimentation rates of aggregates versus their unicellular counterparts. By manually removing the cells that are located at the top of an agitated culture flask, a physical constrain is imposed and a selection for bigger structures ensues. This might seem fairly artificial, in the sense that such selective pressures do not seem to correlate to natural scenarios. However, the presence of nutrient gradients in the sea [Smith et al., 1990, Hayward, 1987] might introduce a similar position-dependent fitness, encouraging cells to develop strategies to remain in particular locations of the water column.

This strategy was put forward in a recent set of experiments involving yeast [Ratcliff et al., 2012] (Figure 1.8C,D), where the authors reported the evolution of aggregative behavior in *Sacharomyces cerevisiae*. Yeast is a specially interesting candidate to explore the potential first steps of the

evolution of multicellularity due to the fact that it already presents some of the discussed toolkit genes and its biology and multicellular states are well enough known so that new emergent phenomena can be easily separated from them.

The authors implemented a selection pressure that led to growing aggregates by subculturing yeast populations, seeding each iteration with the cells sedimented at the bottom of the previous generation growth flask. Remarkably, after just 60 selection events the so-called snowflake phenotype appeared consistently in all replicate experiments. These are roughly spherical clusters of cells formed not by aggregation but by defective separation of cells after division.

The fitness coupling in this case came from the existence of group-level reproduction, in the sense that yeast aggregates would break up in a controlled fashion, creating two daughter clusters with different sizes. This asymmetrical division was reportedly caused by regulated apoptosis in cells located at the core. This was deemed advantageous, for smaller aggregates grow faster thanks to increased diffusion of nutrients to the core and decreased competition for space [Libby et al., 2014]. If correct, this would also entail an interesting example of multi-level selection conflict, for cells that undergo apoptosis necessarily give up their individual fitness for the good of the collective.

In silico models

Another route to explore the possible origins of MC is given by the construction of virtual worlds (Figure 1.9). Similar in perspective to artificial evolution but executed *in silico*, computer simulations enable the reconstruction of potential scenarios and potentially to watch the evolution of complexity unfold. The main interest of building *in silico* models in relation to MC is to provide an initial exploration of minimal sets of premises -both in the properties of cells and environments- that might lead to the establishment of MC behavior.

To be able to observe a transition in individuality, an appropriate framework that permits the emergence of multicellular traits is necessary. Cell-

cell interactions like adhesion or communication provide the basis for such coherent dynamics, and have been typically introduced in models of the evolution of MC. When sufficient potential interactions are put in place, cells have been shown to become metabolically coupled [Kaneko and Yomo, 1999, Pfeiffer and Bonhoeffer, 2003], coordinated in the building of structure [Hogeweg, 2000a] or even display differentiation and split tasks between them (Manuscript 2). In these cases what is sought is to observe an increase cell-cell interactions leading to a fitness integration, that is, cells must fare better as a part of a collective that on their own.

Two main approaches have been commonly used to study the feasibility of becoming multicellular: *phase exploration* and *genetic algorithms*. The former does not deal explicitly with the transitional forms, but exposes conditions in which multicellular organisms with fixed genotypes or traits fare better than their unicellular counterparts. This typically involves defining regions of the parameter space in which a fixed genotype -UC or MC- wins over the other kinds. By contrast genetic algorithms are computer programs that explicitly deals with darwinian dynamics [Forrest, 1993, Mitchell, 1998, Floreano and Mattiussi, 2008]. In them entities can live, mutate and get selected for following some externally set criteria. The specifics behind different implementations can change, yet all include in some way or another the basic aspects of evolutionary theory: inheritance, variation and selection.

A sort of a null model for MC following the first approach is given by Kaneko *et al.* [Kaneko and Yomo, 1999]. In this work the authors explore the possibility of becoming multicellular as a consequence of mere chemical coupling between cells. In particular, they show that cells modeled after chemical reactors that exchange molecules can display collective dynamics, differentiation and pattern formation and reconstruction [Furusawa and Kaneko, 1998, Furusawa and Kaneko, 2003]. These are properties typically attributed to complex regulatory mechanisms, yet the only necessary conditions to obtain these phenomena were shown to be constraints in the connectivity of the chemical networks and the existence of population beyond a given size.

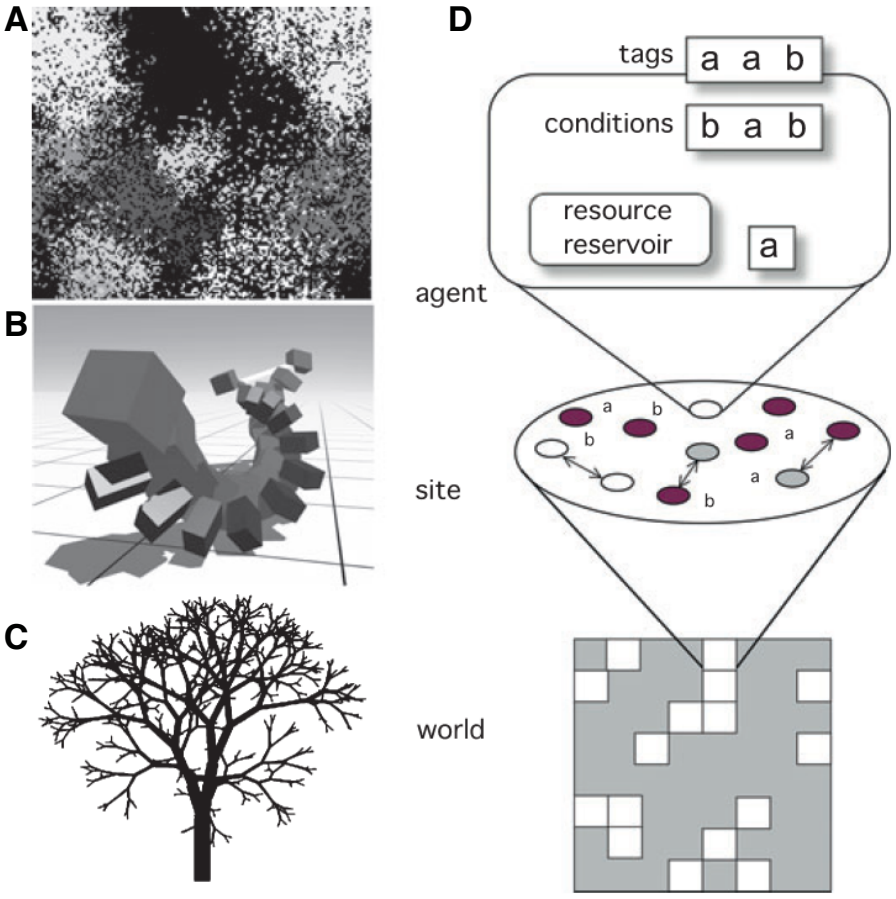


Figure 1.9: In silico models of evolutionary dynamics beyond the gene-based level involve several key ingredients, such as space (**A**) in Avida simulations, (**B**) a physical description of mobile parts used to evolve 3D organisms in a given physical environment under given selection pressures (image by Zach Winkler). Within the context of plant development, structural principles of branching rules (**C**) along with physical constraints associated with efficiency in gathering light or having mechanical resistance evolve to their natural counterparts. Communities of genetic algorithms (**D**) allow to consider the presence of hierarchies.

Remarkably, Kaneko's work shows that perfectly valid unicellular genotypes, meaning coupled sets of chemical reactions able to reproduce on their own, could display collective traits if surrounded by other equal cells. This involved several stages of increasingly complex behavior: first synchronization of oscillatory dynamics, then out-of-phase stabilization and finally the establishment of new states, reinforced through mutual molecular exchanges. This process of "isologous diversification", as the authors dubbed it, contains a powerful lesson: even without invoking fitness unicellular individuals can contain the necessary elements to build complex features.

The first approach was also used in the work by Pfeiffer and Bonhoffer [Pfeiffer and Bonhoeffer, 2003], where the authors explored the possibility of undifferentiated cooperative ensembles of cells forming in a scenario involving different metabolic efficiencies and aggregative behavior. The premises assumed by Pfeiffer and Bonhoffer are quite straightforward, they incorporated cells as regular objects living in a bidimensional world where they could move around and capture nurturing substances in order to reproduce. Cellular movement was constrained by the adhesive properties of cells, in a way that aggregative strains became immobile and created cell mats while unicellular strains moved in a random-walk fashion. At the same time, cells displayed two types of metabolic activity: fermentation and respiration. The main difference between the two was that fermenting cells were considered to elaborate less ATP molecules per unit of nutrient -i.e. were less efficient than respirators-, but were able to grow more rapidly [Pfeiffer et al., 2001]. Within this setting, aggregative respirators were shown to dominate when competing with other strains -free living respirators, aggregative fermentors and free living fermentors- under some parametric conditions. Given that fermenting cells grow faster at the expense of wasting nurturing substances, respiration was considered a cooperative feature and fermentors a sort of cheater that does not invest in the common good of being metabolically efficient. The authors related the success of aggregative cooperators to the natural defense against invasion that characterizes the growth of solid structures, connecting mechanisms to protect cooperative behavior and MC.

A final example of *in silico* analysis of multicellular behavior using the genetic algorithm approach is given by Hogeweg [Hogeweg, 2000a, Hogeweg, 2000b]. In this work a potential link between morphogenesis and the selection for increased cell type diversity was analyzed. In particular, Hogeweg constructed a cellular automaton modeling the development of metazoan-like embryo. The cells conforming the embryo were described as discrete objects of varying size [Graner and Glazier, 1992] with a boolean gene network [Kauffman, 1969, Kauffman, 1993] directing cellular attributes: adhesion, division, death and migration. Interestingly, the author reported the emergence of familiar developmental mechanisms -tissue engulfing, budding, intercalation of layers, programmed cell death and re-differentiation- when selecting for organisms displaying greater numbers of cell types, distinguished by their internal genetic state.

The number of cell types has been suggested as an empirical measure of complexity, and is known to increase through metazoan evolution [Carroll, 2001, Carroll, 2005, Valentine et al., 1994]. Increases in cell type number provide the potential for further evolution of anatomical and functional complexity, essentially enabling division of labor and the formation of specialized tissues. Hogeweg's work takes an opposing perspective to the one proposed by Bonner (section 1.1.3), in which morphological diversity is not selected for but is a consequence of increased organism size. Nonetheless, the remarkable connection reported by Hogeweg, that some mechanisms of development might be favored since they create greater numbers of cell states, is worth considering.

1.4 Pattern formation

"The Scientist must set in order. Science is built up with facts, as a house is with stones. But a collection of facts is no more a science than a heap of stones is a house."

Science and Hypothesis - Henri Poincaré

"Things should be made as simple as possible, but not any simpler."

Albert Einstein (*possibly apocryphal*)

Living multicellular beings are spatially extended, fitness coupled collections of cells. From the thermodynamical point of view they are open systems far away from equilibrium [Nicolis et al., 1977]. This enables them to present highly ordered structures, intimately related to the functions they perform. Thus, a key process in ensuring survivability and reproduction of an organism is the maintenance of spatial distributions of cells and cellular states. This is particularly important when considering the inescapable division of labor present in what we consider true MC.

Accordingly, the study of mechanisms related to the creation and maintenance of order in bodily structures has been of remarkable interest to comprehend MC. In particular, the developmental processes that generate the body plans of modern animals have been thoroughly scrutinized, creating a body of knowledge that extends from the molecular signals that trigger the first stages of patterning to mathematical models able to predict

the outcome of mutations.

Historically, this interest goes back to the classic times, when the first models of morphogenesis were proposed. For instance in the Platonian world view, where the universe was assumed *static* and no evolution of form was possible, it was considered that a preformed, scaled down version of an animal was present in the embryo [Brown and Jones, 1974, Deutsch and Dormann, 2002]. No real development of body structures existed, the organism simply grew until it attained its adult size.

Aristotle, by contrast, held a *dynamical* perspective on development [Deutsch and Dormann, 2002], and considered that there were mechanisms to create structure from an originally uniform mass of cells. However, perhaps due to the impossibility of scientifically assessing the real nature of the developmental mechanisms, he thought that the rules of such processes came only from within the embryo.

Centuries after, Albertus Magnus challenged this idea by developing a theory of plant morphogenesis based in celestial body dynamics [Magnus, 1867]. Despite being certainly a religious explanation, it introduced the idea of a coupling between morphological processes and the physical environment. In particular, he held that the uniform solar irradiation affected the growth of branches, inducing a semi-spherical shape in trees. He also proposed a theory that related the shape of the leaves to differences in "plant fluid" densities. Even if in Magnus' view biological form was the design of a deity, his theories are perhaps the first attempt to link the physical external world to morphogenesis.

Other scholars had also been fascinated by the remarkable regularities in the structure of plants and leaves, like the constant angle at which new leaves grow with respect to one another or the disposition of seeds in a sunflower. Such phenomena hinted at the possibility of mathematically treating the processes of morphogenesis and discovering the subjacent laws that underpinned the construction of form. This view was taken by Hofmeister in his theory of phyllotaxis [Snow and Snow, 1947, Jean, 2009]. He proposed that leaves repel one another and thus tend to appear at the least crowded part of the meristem. Following this proposed mechanism, a fixed angle between leaves must appear, the golden angle.

In essence, his view proposed morphogenesis as an optimization problem, where a minimization of energy -here due repulsion- resulted in the creation of form. This process also has the benefit of avoiding overlapping between leaves, which increases the amount of sunlight that can be captured by them.

Another account of the creation of form by means of simple forces acting on the developing organisms was given by Thompson. In his book *On growth and form* [Thompson et al., 1942], he claimed that simple spatial transformations caused by mechanical processes were the main drivers of morphogenesis. Thompson argued that many creatures share a common basic design and that variations among them could be explained by differential growth rates in the different tissues of the organism. His view might be considered simplistic today, however, it has been demonstrated that many common shapes in different taxa, like the spirals in sea shells or in the horns of goats are indeed caused by this kind of phenomena [Ball, 2011].

A pivotal point in the theory of pattern formation was the introduction of self-organization in biology [Kauffman, 1993, Lehn, 2002]. Self-organized systems are those containing very basic interacting entities, whose simple exchanges cause spatial or temporal dynamics at the collective level. This was a paradigm shift from the previous conceptions linking complex structures in development to complex cellular dynamics. This viewpoint of pattern formation as a natural result of simple exchanges between agents will be discussed in depth in the next sections. More precisely, their role in our current conceptions regarding order and hierarchy in biological pattern formation.

1.4.1 Information transfer in pattern formation

The mechanisms of development create ordered structures, in such a way that more information is needed to describe the organism after the patterning than before. Starting from a small spheroid that is the embryo, successive stages of growth and breakages of symmetry take place and a much more ordered structure is born. Today, it is widely

accepted that such information is stored in the genomes of organisms, even if genome size does not correlate with morphological complexity [Cavalier-Smith, 1978]. Another interesting factor to consider is the actual information exchanges that happen inside the developing embryo. In *Shaping life: genes, embryos and evolution* [Maynard-Smith, 1998] Maynard-Smith proposed that there are two possible origins for the information: it can be inherited from another ordered structure or emerge from within the system.

The first is what the author described as the *rubber stamp* model of pattern formation. In it, an ordered system -such as the previous generation of the organism- produces an outward flux of information towards an unordered system. This procedure is characterized by a *hierarchical transfer of information*, some cells export information while some others are recipients of it.

Contrarily, in the second model there is no hierarchy, all elements of the system communicate in a *decentralized* manner, and patterns emerge from the interplay of dynamical laws of the constituent "particles". A simple physical analogy proposed by Maynard-Smith is the pattern created by a pebble thrown into a body of water: the system is activated by the energy given by the stone, but nowhere in the stone is the recipe for the particular set of ripples and splashes that the system develops into. That configuration is a consequence of the rules by which the water molecules interact, and if different rules were in place other structures might be formed.

Today it is widely accepted that most naturally occurring processes contain their share of both classes [Wolpert et al., 2015, Gilbert, 2000], and that the examples given by Maynard-Smith are useful archetypes. Nonetheless, some iconic developmental theories and mechanisms can be successfully mapped into the former categories. The most straight forward cases are the *french flag model* of positional information developed by Wolpert and the *community effect* by Gurdon and Carnac.

Wolpert championed the model of positional information based around the concept of gradients of signaling molecules [Wolpert, 1969]. Gradients are non-uniform spatial distributions of chemicals, created when there is a source and a sink of signal or, alternatively, when a signal is cre-

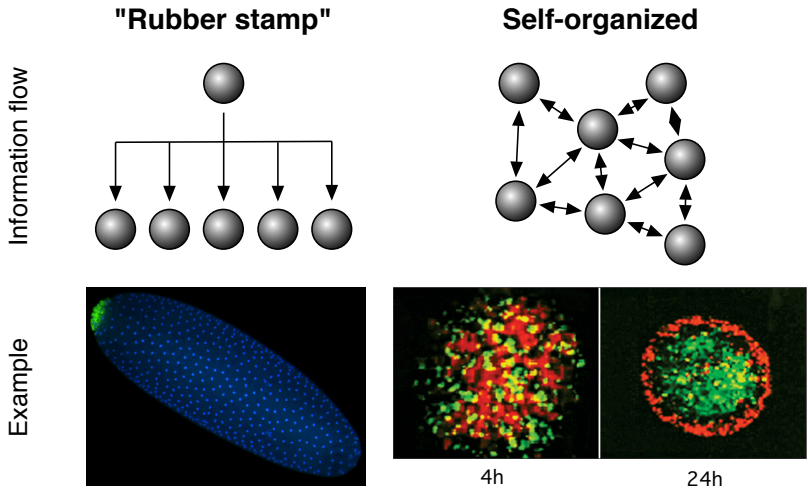


Figure 1.10: The two archetypal modes of pattern formation according to Maynard-Smith. Schematic depiction of information transfer (top), and an example of each process (bottom). On the lower left, differences in gene expression in a *D. melanogaster* embryo following the establishment of a maternal gradient (image provided by Zeiss microscopy. On the lower right side, a time sequence of the sorting process of two preexisting cell types, each labeled with different fluorescent proteins and expressing different levels of Cadherin (from Forgacs *et al.* 2005).

ated in a source and degraded spontaneously. Cells are known to be able to interact with the gradients in a concentration dependent manner and thus extract the local value of signal. According to Wolpert, the amount of different signals coming from different gradients informs each cell of its position in the organism and determines its fate, furthering the developmental process. Following this description, positional information is similar to a "rubber-stamp" process: the gradient provides all necessary information and cells do not dynamically "talk-back" to the sources or sinks.

By contrast, the community effect proposed by Carnac and Gurdon describes a pattern formation process similar to the self-organized class. In this model, cells secrete and incorporate a collection of chemicals, which serve to locally inform cells of the neighborhood state. Cells might react to these signals by changing their state or by changing the signals they produce. In the end, a specific arrangement of cells with specific expression patterns emerges from the interplay of the rules of synthesis and interpretation of the chemicals [Carnac and Gurdon, 1997].

1.4.2 Evolutionary developmental biology

The evolutionary perspective of pattern formation is given by a recently formed field in biology, the so called *Evo-Devo*. This discipline, born with the advent of molecular techniques to probe genetic information and developmental processes, is interested in characterizing the evolution of genotypes and how do they translate into the building of a phenotype. In particular, identifying universal properties in developmental processes and how did those evolve across multicellular organisms.

According to proponents of Evo-Devo, the connection between development and evolutionary processes, frequently overlooked in the mainstream theories, is of utmost importance in determining the possible outcomes of evolution. Development -and by extension pattern formation- can be considered an active source of selection, concluding then that fitness not only comes from the externally imposed environment but also has an internal component, emerging from the constraints in what can be

attained during development.

This realization materialized with the discovery of homeotic (or *Hox*) genes and their central role in the formation of anatomical structures in *D. melanogaster* [Lawrence et al., 1992]. It was observed that some congenital conditions were related to alterations in these *Hox* genes. One of the most famous mutations is Antennapedia, which causes legs grow from the head of a fly instead of the expected antennae [Scott et al., 1983]. It was found that *Hox* genes are master transcriptional regulators of development, switched on in particular segments of the organism. Following this pattern of activation, a cascade of other transcription factors is recruited [Krumlauf, 1994], which serves to further attribute identity to the segments that conform the organism.

Interestingly, other species that do not have segments (including *S. cerevisiae*, *D. rerio* and other model organisms) were shown to contain homologs of these genes in their genomes [Akam, 1989, Krumlauf, 1994]. In these organisms, Homeotic genes are also master regulators of developmental processes, yet with different activation patterns. Remarkably, it was found that *Hox* genes follow a common genetic architecture, with a genomic ordering that follows the activation pattern in the longitudinal axis of the animal. This exemplifies the kind of common features that are thought to underpin the evolution of form in MC creatures and are of interest to Evo-Devo.

Moreover, it brings to the general attention the modular properties of developmental processes in metazoans. The existence of these master regulators enables few mutations to create whole new beings using the sequential recruitment of functional genetic modules. If the genetic instructions to build legs in *D. melanogaster* required contextual information present in the segment where they normally appear, they would not be able to grow such appendages in the place of antennae, and thus would not be self-contained instructions.

1.4.3 Dynamical patterning modules

Other authors have proposed that beyond the common features at the

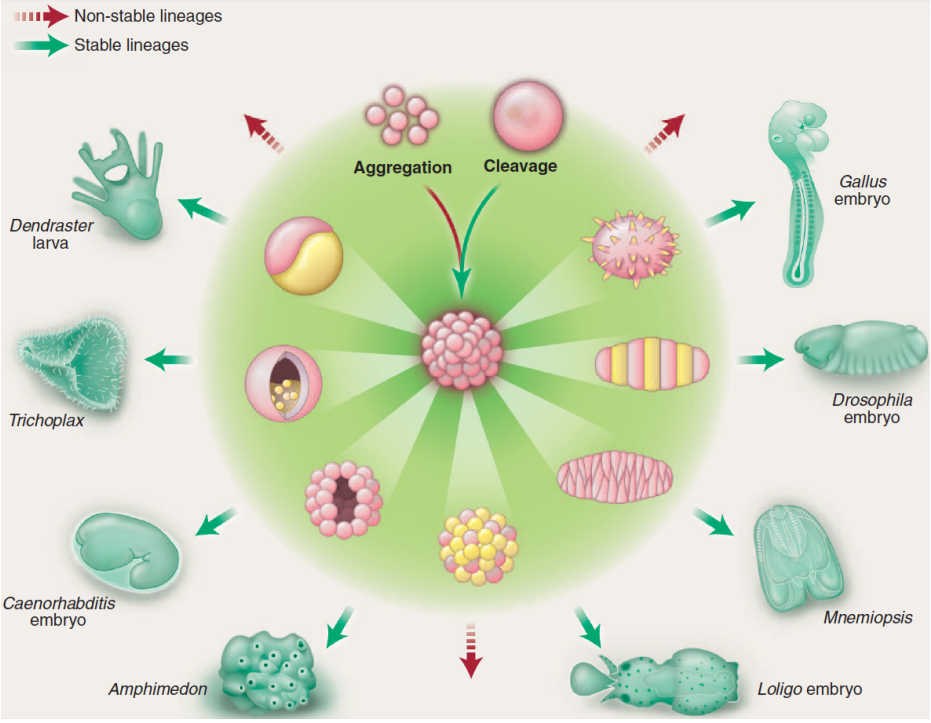


Figure 1.11: A core set of physico-genetic modules underlies the morphological evolution of animals. Multicellular entities (center) were formed by the aggregation of unicellular organisms (red curved arrow) or the cleavage of enlarged cells (green curved arrow). The green inner circle shows morphological motifs generated by some of the key DPMs: physical forces and effects relevant to the multicellular scale, mobilized by certain ancient single-cell gene products and pathways. Emergent motifs include appendages, segments, elongated bodies and primordia, coexisting alternative cell types, interior cavities, dispersed cells, and multiple layers. Genetically uniform clusters produced stable lineages (straight green arrows), whereas chimeric clusters did not (broken red arrows). Contemporary organisms containing some or all of these motifs are shown in the outer circle (From Newman 2012).

genetic level, stemming perhaps from convergent evolution [Morris, 2003] or the constraints of tinkering [Jacob, 1977], universal properties of development can be found in the physical embedding in which it takes place. Namely, that due to the physico-chemical nature of the substrate in which patterns are created, some mechanisms come more readily in the development of multicellular creatures, and specifically metazoans.

This is the "physicalist" perspective of pattern formation taken by Newman and Bhat [Newman and Bhat, 2008], who hold that at the mesoscopic scales of embryos and early multicellular animals, small aggregates -perhaps in the order of a hundred cells- can be understood as viscoelastic and excitable matter [Beysens et al., 2000, Newman et al., 2006]. In this kind of material, morphogenetic processes can be driven by simple, generic mechanisms operating at the cellular level, the so called dynamical patterning modules (DPMs). These include, among others, cell-cell adhesion, phase separation of differentially adhesive populations, oscillations in cellular state, lateral inhibition and structural anisotropy.

For instance, Newman and Bhat show that the processes of adhesion can readily create embryo-like aggregates, with layers organized in a concentric manner. The lateral inhibition mechanisms also affect cell type distributions, but generates regularly spaced islands of a particular type instead. By contrast, anisotropy is shown to be more closely related to the problem of the form of the aggregate, for changes in shape at the cellular level can force the collective to grow in particular directions or bend and buckle in a repeatable way. As their name indicates, the DPMs are thought to be modular, i.e. independent from each other. Accordingly more than one might be applied at the same time, giving rise to more elaborated patterns.

The authors propose that such processes are typically mobilized by the pressence of a small number of molecules, ussually proteins. For instance, cellular adhesion and cell sorting are orchestrated by transmembrane proteins like cadherins or lectins [Takeichi, 1991], oscillations can be attained with particular GRN topologies [Elowitz and Leibler, 2000, Danino et al., 2010], lateral inhibition can be incorporated by the action of signalling molecules like *Notch/Delta* and cellular anisotropy is ob-

tained by the action of *Wnt* [Merks et al., 2006].

In all these cases, the origin of such molecules can be traced to the UC ancestry of metazoans [Tweedt and Erwin, 2015]. However, it seems certain that the functions they performed in single-celled organisms must have been different from the ones that perform today. That is, they were necessarily co-opted into new roles with the change of spatial scale that entailed the evolution of multicellularity, specially at the size of a small colony.

According to Newman and Bhat, the preexistence of these *toolkit* molecules is what enabled the burst of diversification at the cambrian period, when all current animal body plans first appeared. Mainly because no protein innovation was really required, which would have caused a more gradual evolution, and only regulatory sequence mutations were necessary. After this explosion in terms of morphological complexity - which barely lasts 20 million years- no new body plans have appeared [Passamaneck and Halanych, 2004], which the authors conflate to the exhaustion of new dynamical patterning modules to explore.

In current MC organisms, the DPMs have been thoroughly validated as sources of structure [Foty et al., 1996]. This tests come from the manipulation and characterization of animal cells, showing that at the scale suggested by Newman and Bhat such physical mechanisms are responsible for the kind of patterns expected from them. Moreover, since some of these mechanisms seem to strongly constrain the repertoire of potential structures that can be generated, they also offer a powerful framework to understand the origins of convergent designs [Alberch, 1980]. However, the connection to the origins of MC and the central role in causing the cambrian explosion that the authors suggest are still open questions.

1.4.4 Mathematical models of morphogenesis

Another possible line of enquiry into pattern formation besides genetic analysis and experimental characterization resides in the usage of mathematical and computer models to simulate putative mechanisms of morphogensis. This approach requires the formalization in mathemati-

cal terms of the problem being analyzed, which involves proposing dependencies among the variables of the system and relating them to the progression of time. The modeled system can then be "solved", in the sense of generating the time evolution of the variables given some initial conditions.

Consequently, mathematical models provide a unique setting to test concepts and make useful predictions, which can latter be validated in an experimental setup. This cycle might be closed by returning the results of new experiments to the simulation approach, in the form of modified initial conditions, more refined parameter values or new relations among the variables.

Furthermore, given the consistent increase in computer's speed, simulations also provide the modeler with the ability to explore ample parametric values. This is important for understanding the conditions in which the patterning processes might fail, as well as characterizing the whole set of possible configurations that a system can achieve, also known as the morphospace. In turn, this is useful in comprehending the evolvability of a mechanism, in the sense of the capacity to easily switch from one output configuration to another given slight parameter variations.

Several frameworks have been successfully used to model morphogenetic processes, most notably ordinary differential equations (ODEs) and cellular automata (CA). Here, the general considerations regarding these frameworks will be discussed before deepening into the specific aspects that lead to the manuscripts included in this thesis.

Ordinary differential equations

The ODE approach entails the coupling of variables to their derivative in time. For instance, given a system with two variables:

$$\frac{dx}{dt} = f(x, y)$$

$$\frac{dy}{dt} = g(x, y),$$

where f and g are functions that depend on the values of both x and y, which might interact linear or non-linearly. Then, by inputing some initial values for x and y, one can in principle analytically evaluate the system at any given time in the future by integration. However, non-linear systems are notoriously difficult to treat analytically, and numerical methods are commonly used instead.

In the context of morphogensis, the variables x and y usually refer to some biological products, like proteins or chemicals. Insights to define the interactions between x and y can come from the laws of mass action acting at the chemical level. The variables are then considered to decide cellular states by connecting the chemical level and genetic regulation.

Moreover, when working with interacting cells inside an aggregate, the description of the system might have to be extended to a series of coupled compartments, with movement of biological products across them. This can be introduced in different ways, depending on the kind of movement assumed in the system. The most frequent and simple one is passive diffusion. At the lowest description level free small particles in a fluid are known to behave as random-walkers, in what is commonly referred to as brownian motion. This assumes isotropy in the movement of particles, which can shift direction in a stochastic fashion. If enough density of biomolecules is present, a useful simplification can be made by not tracking each individual element but only applying the collective fluxes. When all this considerations are in place, a coupled set of ODEs now stands as:

$$\frac{dx}{dt} = D_x \bigtriangledown^2 x + f(x, y)$$
$$\frac{dy}{dt} = D_y \bigtriangledown^2 y + g(x, y),$$

where D_n is the diffusion coefficient of a given molecular species n and $\bigtriangledown^2$ is the Laplace operator evaluated over the spatial dimensions of the system. For a system with two spatial dimensions (i and j), the Laplace operator is:

$$\bigtriangledown^2 = \frac{\partial^2}{\partial i^2} + \frac{\partial^2}{\partial j^2}$$

Another possibility is directed movement. This usually incurs at the cellular level though the usage of chemotaxis, where cells might be attracted or repelled by a of non-uniform distribution of biomolecules. A general expression for this kind of displacement can also be used, which is similar to the previously shown yet it links the flow of cells to other variables with a new term. For instance, lets consider that in the previous set of coupled ODEs, variable x represents the concentration of a population of cells that are able to respond to the gradient of y:

$$\frac{dx}{dt} = D_x \bigtriangledown^2 - \bigtriangledown C_x(y)x(\bigtriangledown y) + f(x,y)$$

$$\frac{dy}{dt} = D_y \bigtriangledown^2 y + g(x,y),$$

where $C_x(y)$ stands for the chemotactic coefficient, which is usually a function of the signal concentration, and with its sign determines whether cells will be attracted or repelled by y. Equations of this type were first introduced in the context of bacterial movement by Keller and Segel [Keller and Segel, 1970], and over the years have been extensively modified and analyzed [Grindrod et al., 1989, Oster and Murray, 1989]. Such family of equations are known to be able to reproduce collective growth in bacteria [Ben-Jacob et al., 1994, Woodward et al., 1995] and can be further expanded to include several chemotactic processes and more complicated interaction maps.

Cellular automata

CA were developed by von Neuman and Ulam in the 40s as a useful tool to analyze spatially extended systems with discrete particles and simple rules of interaction between them. They were initially modeled after physical processes, such as the growth of crystals and the motion of liquids [Deutsch and Dormann, 2002], were the interacting entities were atoms or inorganic molecules. However, soon researchers began to realize its potential as a framework for the study of biological processes, like Wiener's work on excitable media [Wiener and Rosenblueth, 1946] such

as cardiac tissue and Conway's game of life [Conway, 1970]. This required a change in scale -the interacting particles became cells- and what the rules would stand for. In particular CA have been widely successful in incorporating cell-cell, cell-medium and cell-medium-cell interactions.

Formally, CA assume a collection of sites, also known as *cells*, in a discretized spatial domain. This spatial domain might be regularly partitioned like the square lattice of a chess board or have an irregular placement of sites. The definition of the lattice is important in providing the neighborhood, that is, how many other sites can be interacted with from a given position.

Each cell is considered to have a finite amount of states, commonly just active and inactive. Then, a set of rules is defined for all sites, which establish when the cellular state should change in accordance to the neighbor's states. A discretized time is considered to advance when rules are enacted, this includes different modes *synchronous* -applying rules in all positions of the lattice at once- and *asynchronous* -choosing a random subset of positions and evaluating the rules only in them-.

In the next sections two examples of models of pattern formation relevant to the presented manuscripts are discussed in depth: Turing patterns and cell sorting. Following the classification proposed by Maynard-Smith, these clearly sit in the self-organized domain, for these processes entail bidirectional communication among the cells.

Turing patterns

When Turing approached the problem of morphogenesis in 1953 it had long been known that animal cells could communicate with each other in order to create patterns by secreting diffusible molecules. In particular, Spemann had established 30 years prior the key role of organizers in the induction of structure [Spemann and Mangold, 1924]. These are specific regions of embryonic tissue that decide the fate of the neighboring cellular population, even when transplanted to other parts of the embryo. This was thought to be caused by the action of specific diffusible signals invading the surrounding tissue from the organizers. Such secreted

molecules could be used as a patterning signal if they induced changes in cellular state in a concentration or time dependent manner. Their molecular identity was far from known, but whether proteins or simple organic compounds these signals came to be called morphogens.

Even if some regions were know to induce local changes in cellular state, this did not solve the problem of how did the initial asymmetrical roles -organizer versus non-organizer tissue- arose in the first place. Turing's contribution to the field of developmental biology then, was the suggestion of a revolutionary mechanism by which non-uniform distributions of morphogens might be created and maintained in a system with initial compositional symmetry.

Turing's proposal included two morphogens that interacted in a non-linear way and diffused through the spatial domain. This constitutes the requirements for the so called reaction-diffusion systems (RD), which may be described using the ODE formulation given before. Prior to this seminal work, diffusion was considered to be a purely homogenizing process, which necessarily would drive the system into a final uniform state. However, Turing showed that under certain parametric conditions RD systems could display a diffusion-driven instability, that is the existence of stable states in the local system that become unstable in the presence of diffusion.

Under such conditions, diffusion enables the amplification and the spatial propagation of infinitesimal variations in morphogen concentration, which are assumed as the initial condition in the Turing mechanism. As time progresses distinct stable regions develop, some with high concentrations of both morphogens and some were only one of them is present. These domains also display higher order properties, for they create periodic regular patterns: spots, stripes and mazes (Figure 1.12). These changes in morphogen concentration are thought to induce different biological structures, like pigmentation distributions in animals [Murray, 1988, Kondo and Asai, 1995, Kondo and Miura, 2010], follicles [Sick et al., 2006], teeth structures [Jernvall, 2002], establishment of palate ridges [Economou et al., 2012] and the creation of limb skeletal primordia [Hentschel et al., 2004, Sheth et al., 2012, Raspopovic et al., 2014]. Also

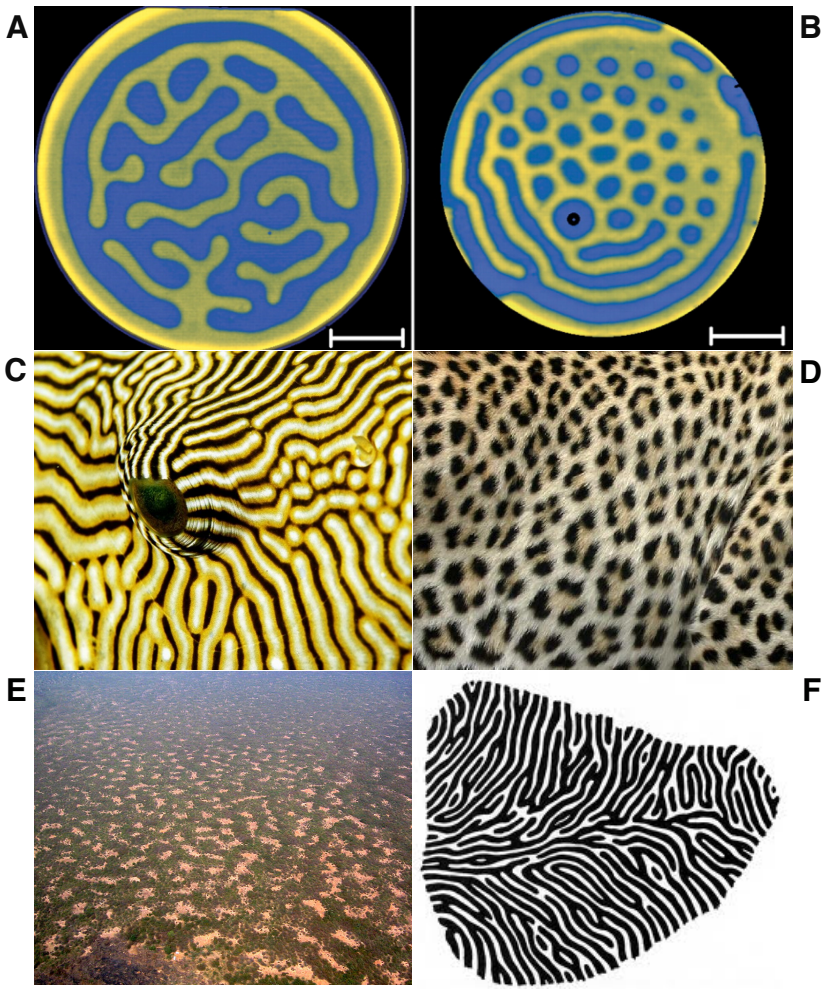


Figure 1.12: Examples of patterns thought to be created by a Turing-type mechanism. Engineered chemical systems displaying regularly spaced structures from [Horváth et al., 2009] **(A, B)**. Patterns surrounding the eye of a giant puffer fish **(C)** with particular interest to Kondo and coworkers. The iconic spots of the leopard **(D)**. Patches of tiger bush in Niger **(E)**. Ocular dominance in the primary visual cortex **(F)**. Unless stated otherwise, all images are provided by Wikimedia commons foundation.

chemical [Ouyang and Swinney, 1991, Kapral and Showalter, 2012], and ecological settings [Lejeune and Tlidi, 1999, HilleRisLambers et al., 2001, Alonso et al., 2002, Thar and Kühl, 2005, Baurmann et al., 2007] contain plausible cases of Turing mechanisms, where the morphogens might be entities of radically different nature.

Periodic structures generated in these systems are commonly characterized with Fourier Transforms, more precisely with its numerical counterpart Fast Fourier Transforms (FFT). Using this method, spatial distributions of a signal (whether continuous or discrete) are transformed and represented in the frequency domain. If the signal intensity has any regularly spaced correlation, the Fourier Transform will capture it by displaying a dominant spatial frequency. Turing patterns in particular have a specific signature in the Radially Averaged Power Spectrum (RAPS): a decaying correlation with distance overlapped with a single peak located at the dominant wavelength of the pattern.

Despite the wide success in modeling biological phenomena, some criticisms have been raised against Turing's initial approach. In particular, the fact that the interaction terms in Turing's RD bear little resemblance to the ones expected in biochemical systems. These regularly comprise exponential decays or dissipation, regulated synthesis and degradation and multimerization in the form of hill functions [Diambra et al., 2014], among others. Another precondition for Turing patterns that has been subject of concern is the required difference in diffusion rates between the morphogens, typically in several orders of magnitude, and how could it be achieved and evolved in a biological setting.

Later, Gierer and Meinhardt revisited the subject of the Turing mechanism and sorted out this issue by fleshing out the role of the two morphogens. They proposed that one of them could be interpreted as an *activator* with autocatalytic activity while the other should be regarded as an *inhibitor* [Gierer and Meinhardt, 1972, Meinhardt and Gierer, 1980, Meinhardt and Gierer, 2000]. The former would increase its own synthesis and that of the inhibitor, while the latter would somehow interfere with the activator. Given the more rapid diffusion of the inhibitor, the authors proposed that Turing's proposal was fact is a particular case of Local Ac-

tivation Lateral Inhibition (LALI) mechanisms. A common feature of LALI is the "Mexican hat" potential, regardless of the specific subjacent mechanisms activation applies only locally while at some characteristic distance of an activated region a predominantly negative interaction is found.

Many other authors have since proposed their equations for Turing patterns in the context of multicellular organisms, yet identification of the actual morphogens and the interactions between them has only been a recent possibility [Economou et al., 2012, Raspopovic et al., 2014]. Turing's grand contribution to the field of pattern formation resides in inspiring countless scientist to seriously consider the role of modeling and self-organization in developmental biology. Despite this, it is worth noting that the real Turing patterns might be implemented by a myriad of mechanisms beyond RD systems, as will be seen in the manuscripts and the discussion.

Steinberg's differential adhesion hypotheses

Another kind of problem related to the establishment of particular configurations of cell types arises when considering that cell states do not change as time progresses, but instead that cells move around inside a tissue to achieve a target spatial distribution. The mechanisms of motion can be multiple: from directed migration -as in the case of development neural networks in the brain [Lois and Alvarez-Buylla, 1994]- to differential adhesion -briefly introduced previously when discussing the DPMs-.

The first case necessarily introduces a signal gradient of some kind that cells follow up to their destination. This gradient can be a diffusible molecule -very much like the morphogens of last section- or a physical non-homogeneous property of the tissue [Lois and Alvarez-Buylla, 1994]. However, in differential adhesion [Steinberg, 1970, Steinberg, 1975] motion appears due to a process of free energy minimization (Figure 1.13). This stems from the fact that cell-cell interactions have different associated strengths, mediated by the attachment of adhesion proteins located at the membrane, like cadherins and lectins [Mayer et al., 1993]. These can

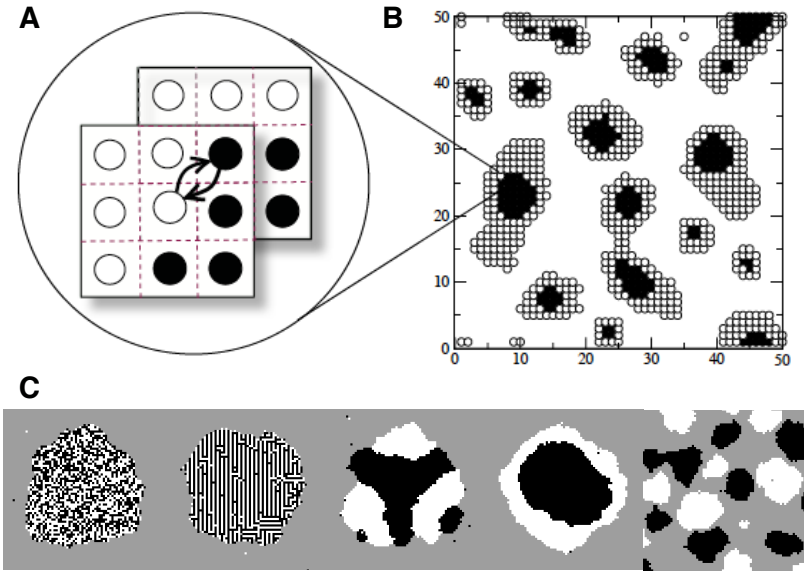


Figure 1.13: The differential adhesion hypothesis defines a energy minimization process **(A)** in which individual cell dynamics is driven by the local preferences. This generates regular and repeatable at the global scale **(B)**. Depending on the relations between the values of the adhesion matrix different kinds of structures can be obtained **(C)**.

preferentially bind to equal molecules and form homodimers or to other adhesion molecules creating heterodimers.

If these interactions are weak and thus reversible, cells can move around inside a tissue or an aggregate in an almost random fashion, similarly to brownian motion. This makes tissues behave like liquid matter [Steinberg and Poole, 1982] and exhibit properties equivalent to viscosity and surface tension [Beysens et al., 2000]. Then, differential adhesion would be the equivalent of phase separation in a purely physical setting. This is the process by which a mixture of liquids with different polarity becomes segregated in distinct phases with sharp boundaries. According to the differential adhesion hypotheses, things are no different in a cellular environment, certain cellular arrangements -those that put together cells that do preferentially attach between them- incur in less free energy and are stabilized.

In order to model this process the cellular automaton approach can be used. In it, cells are considered discrete entities, simplified as solid objects with equal size. Cells exist in an equally discretized spatial domain, including the possibility of particular sites being empty -consisting of culture medium or extracellular matrix-. Each position (i, j) in the spatial domain (Ω), whether empty or occupied by a cell, will be characterized by a state $(S_{(i,j)})$: 0 if empty and $\{1, ..., n\}$ if occupied, where n is the finite number of cell types assumed in the model.

Then, the interaction energy between a site with state a and one with state b might be expressed as $J_{(a,b)}$, and matrix consisting of all possible pairwise interactions between types can be constructed:

$$\mathcal{J} = \begin{pmatrix} J_{(0,0)} & J_{(0,1)} & \cdots & J_{(0,n)} \\ J_{(1,0)} & J_{(1,1)} & \cdots & J_{(1,n)} \\ \vdots & \vdots & \ddots & \vdots \\ J_{(n,0)} & J_{(n,1)} & \cdots & J_{(n,n)} \end{pmatrix}$$

If interactions can only happen among nearest neighbors, the energy $(\mathcal{H})$ associated to a given position (i, j) in the lattice can be defined as:

$$\mathcal{H}_{ij} = \sum_{(k,l)\in\Gamma_{(i,j)}} \mathcal{J}_{(S_{(k,l)},S_{(i,j)})}$$

where $\Gamma_{(i,j)}$, in a square lattice, is the set defined by the eight nearest neighbors of a given cell in position (i, j) (Moore's neighborhood), each of them occupying a position (k, l), and having a defined state $S_{(k,l)}$.

The algorithm of the CA proceeds as follows, given an initial configuration -typically a random distribution of cells and states- a position in the grid is randomly chosen. Then a random position from the eight nearest neighbors is picked. The energy function just presented here is evaluated for the current configuration of the system and for a proposed new configuration in which the states of the two chosen sites have been swapped. The change is applied if the new configuration has less energy, if not the system remains as it is. The cycle is started over with a new random position and a random neighbor until a number of iterations have taken place.

With these simple rules, patterns similar to the ones coming from experimental settings are observed. That is, cells will form regular structures commonly forming layers with the cells with higher self-cohesion sitting at the core of the aggregate (Figure 1.10 bottom right). Beyond its existence in natural cell types [Foty and Steinberg, 2005], the differential adhesion hypothesis has been tested in engineered populations of cells consisting of two types, each expressing a different adhesion protein [Forgacs and Newman, 2005]. It has also been found that differential adhesion can be achieved using a single cadherin, as long as there are significant differences in the level of expression in the two or more involved cell types [Forgacs and Newman, 2005].

This gives aggregates, and supposedly early multicellular animals, a powerful tool to develop bodily structures. Even in the face of an uncertain environment, where mechanical effects can split the aggregate or random cells are destroyed, the pattern can rebuild itself. Despite this, the ratio between cell types might shift, for cell types cannot change and therefore cannot adapt to external manipulations. This stands as a basic difference between this model of patterning and the Turing mechanism

just discussed, which can dynamically assign cellular states.

1.4.5 Synthetic pattern formation

In the last few years, the possibility of engineering biological systems in order to investigate patterning mechanisms has gained some traction [Vinson, 2011, Miyazawa, 2012, Kong et al., 2014, Fu and Huang, 2014]. Biological matter began to be regarded as a possible target of engineering after techniques for extensive genetic manipulation and a vast body of knowledge at the molecular were available [Benner and Sismour, 2005]. This is the basis of synthetic biology, a new field in biology that intends to splice basic biologic components to create new functions or reproduce existing ones in unrelated systems [Endy, 2005]. The connection to engineering disciplines comes from the systematic way of characterizing and building biological devices [Heinemann and Panke, 2006], in an approach that has been casually referred as "lego-like" [Endy, 2005].

The premise is quite simple, to functionally characterize simple biological components and modularly build more complicated functions from them, frequently putting together proteins and DNA sequences that do not naturally coexist. A key aspect then for the successful engineering of biological matter is orthogonality. This term refers to the capacity of biological components to operate without interfering with each others function, i.e. to act as isolated modules. Orthogonality also implies that the efforts in characterization of a particular element are simplified, once characterized on its own the same behavior should be expected regardless of the biological context.

An important source of inspiration to synthetic biology has come from electric engineering. In that field, electronic components with well known functionality are put together in order to manipulate electric currents, with expected and predictable results. However, the medium or context in which biological components operate is quite different from a circuit board [Regot et al., 2011]. The "dense soup" that is the interior of a cell does not resemble the neat organization of a motherboard, which imposes serious obstacles to the scalability of biological designs through the noto-

rious wiring problem.

This has not deterred synthetic biologists from constructing fairly complicated devices, although currently there is a cap to design complexity, set at 6 codifying genes [Purnick and Weiss, 2009]. This might be due to lack orthogonality between the used pieces or interference with the host natural processes, like in the so called metabolic load [Glick, 1995].

In the field of pattern formation, several milestones in constructing and testing simple patterning modules have been achieved. This is the case of Basu *et al.* [Basu et al., 2005] who elegantly built a bacterial example of the *french flag* model [Wolpert et al., 2015]. This required putting together three transcription factors in an incoherent feed-forward loop [Mangan and Alon, 2003], one of the simplest motifs to create a band pass filter. It also required the external input of a gradient of signal, in this case given by the extensively used and characterized quorum sensing system of *Vibrio fisherii* [Fuqua et al., 1994].

Another interesting example of a hierarchical transfer of information mechanism implemented in the context of synthetic biology can be found in Levskaya *et al.* [Levskaya et al., 2005]. In this work, the authors proposed and implemented a clever solution to edge detection, a relevant problem in computer vision [Perona and Malik, 1990] which also might have important implications for development. Here, the target behavior is to drive the expression of some biological product in edges of a given shape flashed into a 2D culture as light. This means that crucial information for the formation of the final pattern comes straight from the environment, but cells do not just store the information of the flash, they have to make some distributed computation and display the transformed shape.

The mechanism was implemented with two modules, one for light sensing and another for proximity sensing. The first was introduced with an engineered sensor that enabled cells to distinguish between light and dark regions. Without light, the engineered bacteria were designed to produce a diffusible chemical from the quorum sensing genes of *Vibrio fisherii* that can reach the illuminated sections of the petri dish. The logic then was to express some biological marker wherever both light and the diffusible signals were present with an AND gate.

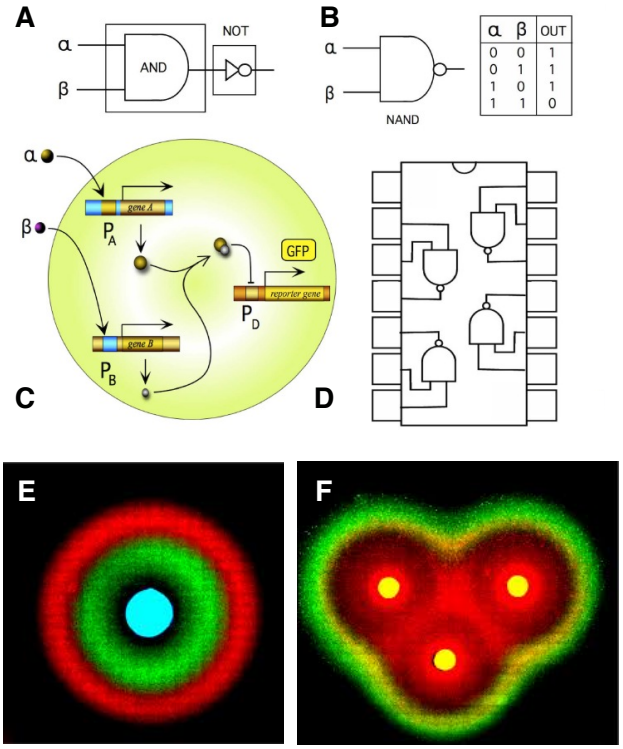


Figure 1.14: Synthetic biology provides an embodied design system of living cells capable of performing predefined functions. An example is provided by the implementation of a given logic element, such as the NAND gate, indicated in (**A,B**) both as a logic unit and as a truth table. In (**C**) we show an example of a potential implementation based on an engineered interaction between input signals, genes and a reporter (GFP). Ideally, a combination of these gates, as it is done in electronic chips (**D**) could allow designing arbitrary living computational devices. (**E**) Pattern formation by Basu *et al.* 2005 Here cells compute a band-detector of a gradient with synthetic circuit and express different fluorescent proteins following a simple french flag model: absence of signal, green and red fluorescence proteins. (*F*) If more "organizing centers" are present, more complicated structures can be attained.

Beside these examples of "rubber-stamp" pattern formation, there are also a few cases of self-organization in the synthetic biology literature. A very interesting one can be found in Liu *et al.* [Liu et al., 2011], were the authors modified the natural processes of bacterial movement in order to achieve periodic ring formation. Although considered unicellular systems, bacteria have long been known to generate intricate structures of cell density [Woodward et al., 1995, Ben-Jacob et al., 2000, Xavier, 2011]. These have been known to depend on several physical parameters, like agar stiffness or concentration of nutrients and biological ones, like bacterial or chemoattractant concentration.

This phenomenon is what inspired the authors to manipulate the gene *CheZ*, linking its expression to bacterial density. The connection was again given by a diffusible signal constitutively synthesized by cells from the quorum sensing system. If enough signal was present, bacterial movement was decreased by overexpressing CheZ. When a population was grown from a small aliquot at the center of a petri dish, the engineered cells created periodic stripes of high density, in stark contrast to the uniform distribution that the natural population generates when grown in the same conditions.

Chapter 2

OBJECTIVES

Given the discussed properties of evolutionary transitions in general and the transition to multicellularity in particular, in this PhD thesis we have aimed to shed some light into the organizational principles of multicellular organisms. In particular we have directed our attention towards:

- The development of new models of spatial patterning in cellular collectives.
- The creation of clear frameworks to understand the emergence of MC arising from darwinian entities.
- To provide general scenarios in which differentiated multicellularity with division of labor should be expected.

Several manuscripts are included in the results of this thesis in the following chapter, one of which ("Emergence of multicellularity in a model of cell growth, death and aggregation under size-dependent selection") has been published in the Journal of the Royal Society Interface. The rest have been submitted for publication at the time of writing this PhD thesis.

Chapter 3

RESULTS

Duran-Nebreda S, Solé R. Emergence of multicellularity in a model of cell growth, death and aggregation under size-dependent selection. J R Soc Interface. 2015 Jan 6;12(102):20140982. DOI: 10.1098/rsif.2014.0982

Duran-Nebreda S, Bonforti A, Montañez R, Valverde S, Solé R. Emergence of proto-organisms from bistable stochastic differentiation and adhesion. J R Soc Interface. 2016 Apr;13(117). pii: 20160108. DOI: 10.1098/rsif.2016.0108

Turing patterns in an engineered cell population model with chemotaxis

Salva Duran-Nebreda[†,‡] and Ricard V. Solé[*,†,‡,¶]

†*ICREA-Complex Systems Lab, Universitat Pompeu Fabra, 08003 Barcelona*

‡*Institute of Evolutionary Biology, UPF-CSIC, 08003 Barcelona*

¶*Santa Fe Institute, 1399 Hyde Park Road, 87501 Santa Fe, New Mexico, USA*

E-mail: ricard.sole@upf.edu

Abstract

A major force shaping form and patterns in biology is based in the presence of amplification mechanisms able to generate ordered, large-scale spatial structures out of local interactions and random initial conditions. Turing patterns are one of the best known candidates for such ordering dynamics, and their existence has been proved in both chemical and physical systems. Their relevance in biology, although strongly supported by indirect evidence, is still under discussion. Extensive modeling approaches have stemmed from Turing's pioneering ideas, but further confirmation from experimental biology is required. An alternative possibility is to engineer cells so that self-organized patterns emerge from local communication. Here we propose a potential synthetic design based on the interaction between population density and a diffusing signal, including also directed motion in the form of chemotaxis. The feasibility of engineering such a system and its implications for developmental biology are also assessed.

Introduction

Living systems are spatially-extended, out of equilibrium systems with highly organized structures. In order to achieve such complex organization and specially during development, molecules must get distributed in ordered ways within cells and tissues. At the organ level, spatially differentiated groups involving different cell types need to exchange signals capable of modifying the molecular responses of other distant cells, ensuring proper coherent responses to the environment.[1] This requires information storage and processing, connecting different scales and thus needs the involvement of long-range dynamical processes.[2,3] A specially relevant problem in synthetic biology is the understanding of how spatially organized patterns emerge from local interactions among neighboring cells.[4]

This problem was approached by the great British mathematician Alan Turing in his classical 1952 work "On the chemical basis of morphogenesis".[5,6] In that paper, Turing presented a groundbreaking approach to the problem by using a model involving two interacting molecules (the *morphogens*) which could also diffuse over space. Under an appropriate conditions,[7] these *reaction-diffusion systems* can be shown to display spatial instabilities: the combination between local amplification due to reactions together with the propagation of such amplifications due to diffusion -which plays here the role of a transmission system- is able to display periodic arrangements of morphogens concentration.

Reaction-diffusion (RD) systems are usually described in terms of a system of coupled partial differential equations. In its simplest form, two morphogens M_1 and M_2 would interact through a nonlinear model:

$$\frac{\partial M_1}{\partial t} = f_1(M_1, M_2) + D_1 \nabla^2 M_1$$
$$\frac{\partial M_2}{\partial t} = f_2(M_1, M_2) + D_2 \nabla^2 M_2$$

where the functions $f_n(M_1, M_2), n = 1, 2$ define the specific form of the molecular interactions taking place between M_1 and M_2 and that can involve different kinds of chemical

reactions: synthesis, degradation, catalysis and multimerization among other possibilities.[8] Turing implicitly considered morphogens as Brownian particles, which allows the use of a diffusion term $D_k\nabla^2 M_n$ where D_n is the standard diffusion coefficient, weighting how fast the particles move on average. One successful way of finding Turing-generating instabilities is considering that one of the morphogens acts as an activator (hereafter A) and the other as an inhibitor (I).[9–11] Specifically, the most common interaction map includes autocatalytic activity and cross activation for the activator, and it is often assumed that the I inhibits A by either repressing its synthesis or accelerating its degradation (Figure 1A). Moreover, a widely accepted condition for instability is provided by the assumption that $D_I \gg D_A$. In other words, that the inhibitor diffuses much faster than the activator.

The RD approach has been widely successful in reproducing the patterning in different scales and systems: from skin pigmentation in animals[12,13] to organism distribution in ecological scales.[14–17] However, the identification of the molecular mechanisms underpinning some of these processes have only been recently identified,[18,19] and are still under discussion.

Another perspective can be taken by asking if one or both of the morphogens can be discrete, active agents such as cells.[20–22] While not included in Turing's approach, this possibility appears as an interesting line of inquiry to consider, specially under the light of recent experimental results.[23,24] In them, it was reported that pigment cell rearrangement during development is required for proper pattern formation in zebrafish. This raises important questions regarding the role of cell movement in naturally occurring Turing patterns, whether its just a transmission system as mentioned before or can have a more relevant activity,[25,26] but it also opens up new opportunities in the synthetic pattern formation domain.

Here we will explore the possibility of creating periodic structures with finite wavelength arising from a new interaction map (figure 1B), containing also two entities: a population of cells capable of chemotactic movement (x) and a freely diffusing signaling molecule and chemoattractant (Q). Essentially, we will consider that cells constitutively synthesize Q and are drawn to it following the steepest gradient, while at the same time, the chemoattractant

drives an apoptotic response in x and diffuses as a brownian particle.

Given the properties assumed for each morphogen, we think that a suitable pair of candidates would be fibroblasts and TNF-α (Figure 1C). It has been established that fibroblasts are capable of chemotactic movement towards several key hormones[27,28] including our proposed signaling molecule. Additionally, these eukaryotic cells move in a way that closely matches our model, specially when compared to the swim and tumble movement of model prokaryotic organisms.[29,30] It has also been described that fibroblasts do naturally synthesize and respond to several hormones including this tumor necrosis factor.[31–33] Accordingly, since all the relevant proteins for synthesis and transport of this hormone are in place, a simple integrative or episomal vector harboring a synthetic gene that constitutively expresses TNF-α would introduce the desired behavior. On the other hand, TNF-α activates antagonistic signaling pathways that can either increase survival or force apoptotic responses depending on the genetic background of cells and other signaling processes.[34–36] Fortunately, extensive efforts have been put on identifying the key effectors of both pathways and blocking the proliferative effect of TNF-α with a simple knock out is possible.

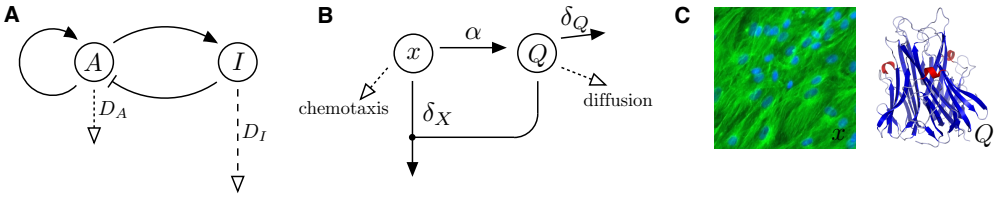


Figure 1: The basic logic of Turing's diffusion-driven instabilities. The two main players are E. coli cells and lactone **(A)** whose spatial abundances are expected to organize due to Turing instabilities. In **(B)** we display the schematic representation of interactions taking place in the classical setting, A and I standing for activator and inhibitor respectively. In contrast, **(C)** shows a schematic depiction of the mechanism used in this work, x and Q being the cells and the chemoattractant which take the roles of activator and the inhibitor respectively.

The model

In order to test our hypothesis regarding the possibility of building periodic arrangements of cells using a chemotactic process, we built an agent based, embodied computational *Netlogo* model incorporating the two entities previously described: cells (x) and the chemoattractant (Q). The first one is introduced as a collection of point objects, placed in a bidimensional domain Ω, each one characterized by its position and velocity. The second one is modeled as a regular square lattice with periodic boundary conditions.

Following the interactions displayed in Figure 1c, we assume that cells are introduced at a fixed rate -or alternatively could be considered to differentiate from a subjacent stem cell population- with essentially randomized positions. At the same time, cells are killed -or detach from the culture surface- if the signaling molecule Q surpasses a given threshold (Θ). Moreover, cells react to the local values of signal concentration -i.e. they have access to the concentration of Q in their eight nearest neighbors- and move uphill following the chemical gradient.

More precisely, our cells decide their velocity by integrating the gradient in the local neighborhood, accounting also for other cells that might be in their vicinity (in order to avoid overlaps). The first part is computed by increasing the chemotactic pull in a particular direction following a Michaelis-Menten like formulation, namely:

$$\vec{v}_{i,j} = \sum_{(i',j')\in\Gamma_{i,j}} \vec{u}_{(ij\to i'j')} \frac{\beta(Q_{i',j'} - Q_{i,j})}{(k_Q + (Q_{i',j'} - Q_{i,j}))},$$

where $\vec{v}_{i,j}$ is the velocity a cell would experience if it stood in the position (i,j), $\Gamma_{i,j}$ is a standard Moore's neighborhood of size one centered at (i,j) and $\vec{u}_{(ij\to i'j')}$ is the unitary vector from (i,j) to each of the neighbors (i',j'). The second part, is incorporated by checking if there is any other cell at a radius of 1.5 cell widths, then a repulsive component to the velocity is applied to both cells which keeps them separated.

On the other hand, we consider that Q is locally synthesized at a constant rate by cells,

decays exponentially and diffuses into a Moore's neighborhood of size one, computed using the standard discretization:

$$D_Q \bigtriangledown^2 Q_{i,j} = D_Q \sum_{(i',j') \in \Gamma_{i,j}} [Q_{i',j'} - Q_{i,j}],$$

Taking all these processes into account, a simple ODE formulation can be obtained for the finite element (i, j) in Ω, which reads as:

$$\frac{dx_{i,j}}{dt} = B(x, Q) + \rho - \delta_x x_{i,j} \Phi(\Theta, Q)$$
$$\frac{dQ_{i,j}}{dt} = D_Q \bigtriangledown^2 Q_{i,j} + \alpha x - \delta_Q Q_{i,j},$$

where $B(x, Q)$ is the cellular flux associated to the chemotactic and repulsive processes, which depend on the fields of cells and signaling molecule as described before, ρ is the constant input of cells, $\Phi(\Theta, Q)$ is a step function that equals 1 if $Q \geq \Theta$ and implements the signal mediated death, α is the parameter regulating cell-dependent synthesis of Q and $\delta_Q Q_{i,j}$ is a standard exponential decay for the chemotactic signal.

Results and discussion

Phasespace of pattern formation

First we wanted to assess the existence of domains in parameter space that yielded distinguishable phases in terms of the formed patterns. More precisely, under what conditions periodic structures where formed in the distributions of cells and morphogens. This periodicity has been commonly analyzed by the existence of a peak in the frequency domain of the Fourier Transform of spatial data, corresponding to dominant wavelength of the pattern. In order to examine this possibility we run the model with varying parameter sets and found that D, ρ and β (Diffusion coefficient, cell input and chemotaxis coefficient respectively)

where the main drivers behind the establishment of periodic structures.

Figure 2 shows the boundary between a non pattern forming region in white -where cells are scattered in a random fashion across the lattice as well as the chemotactic signal- and a periodic structure forming region in grey -where a peak in frequency domain exists-. This suggests that a trade-off between cell input and chemotaxis strength decides the fate of the system: if too many cells are introduced and the Q mediated death cannot remove them fast enough or the chemotactic pull does not pull them together fast enough, the pattern dissolves into a random distribution.

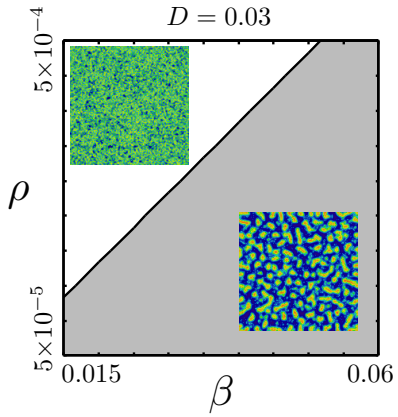


Figure 2: Phase space analysis of the system. For a constant $D = 0.03$, $\beta = 0.3$, $\Theta = 1.1$, $\alpha = 0.075$, $\delta_Q = 0.05$ and $\delta_x = 0.05$ the parameters ρ and β were linearly explored (21×21 independent parameter sets, 10 replicates each), finding a clear boundary between the periodic and non-periodic pattern forming regions. These were discriminated by the existence of a peak in the frequency domain after the treatment of the spatial distribution of the chemotactic signal Q with a FFT algorithm (Matlab 2013b, data not shown). As insets we show some examples of patterns formed in each side of the boundary, representing in color scale the distributions of Q.

Wavelength determination

Next we wanted to assess the capacity of the system to display the typical structures that a normal reaction-diffusion system can construct, namely spots, stripes and labyrinths.[37–39]

These are higher order properties of the distributions of morphogens, i.e. how the activated regions spatially connect and extend. Figure 3A shows how parameter variation can force the creation of the aforementioned types of structures in the distribution of cells. In particular, for low chemotactic pull and high input of cells, increasingly longer structures are formed, eventually reaching an almost completely interconnected phase. On the other hand, when cells are added more less frequently to the simulation space and move slower, spots tend to be created, arranging themselves in quasi-hexagonal lattice (Figure 3B). Moreover, the rules upon which this model is based readily enable intercalation and fusion phenomena described both in other models and natural systems (arrows in Figure 3B). These refer to the apparition of new structures between distant activated regions and the collapse of two activated regions into a single spot when they are too close.

An important issue regarding the pattern specification is the dominant wavelength of the spatial distribution. This typically depends on the various parameters directing the non-linear processes of interaction among morphogens and the ratio between diffusivities. For our proposed model involving chemotactic movement, we found that three parameters are mainly responsible for defining the dominant wavelength, the diffusion coefficient (D), the chemotactic pull parameter (β) and the cellular input (ρ). Figure 3C shows for a fixed value of cellular input how β and D affect the dominant wavelength (λ).

Impact of saturation in pattern formation

One interesting property observed in natural systems is the formation of not only periodic structures but linear arrangements of high morphogen concentration, a well documented outcome of fish pigmentation patterning across different species.[38–41] This effect can be obtained in canonical reaction-diffusion systems with the introduction of particular boundary conditions, which create a pre-pattern and force the organization of the morphogen structures in accordance with that pre-existing information. However, another possibility arises in our model when considering that chemotactic movement is necessarily mediated by signaling

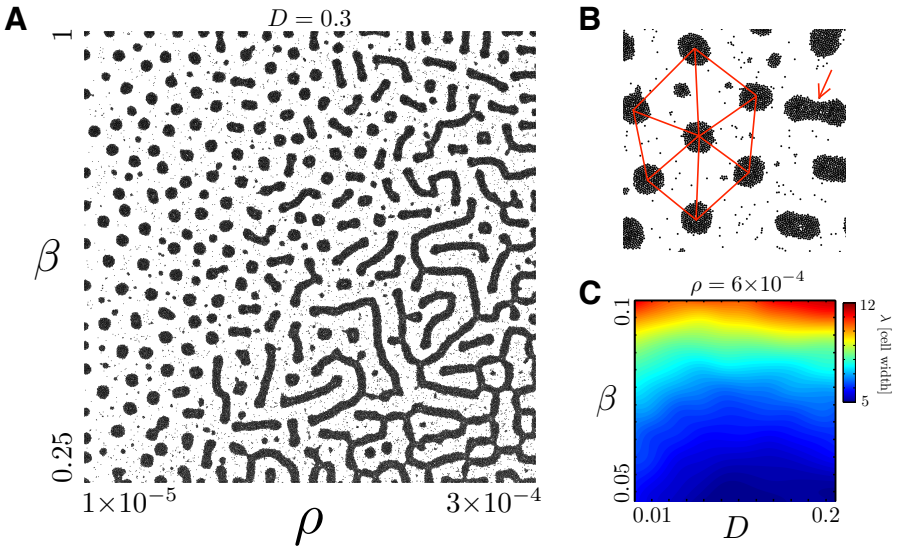


Figure 3: Phases characterized by spots, stripes and mazes in relation to the parameters used. In (**A**) we display the cellular distribution after 5×10^4 algorithm iterations under varying β and ρ values but otherwise the same parameter set as before. The whole simulation was carried out in a single non-periodic lattice with locally assigned parameters, varying linearly in each axis. (**B**) is a zoomed region of the previous lattice, showcasing intercalation, fusion and hexagonal lattice distribution of spots, as well as the different scales -from cell to spot- involved in this model. (**C**) impact of β and D on the dominant wavelength in cell widths (bicubic interpolation of a lattice consisting of 21×21 independent simulations 10 replicates each).

pathways and that receptors can become saturated.

In particular, we tested the impact of such effect by modifying the semi saturation constant (k_Q) of the chemotactic pull function (Figure 4A). At low values of k_Q (i.e. $\Delta Q >> k_Q$) cells lose the ability to discriminate small differences in chemotactic signal and cease to move uphill the chemical gradient. What this remarkably entails is the creation of more regular structures, with stripes extending in a straight manner and creating domains of parallel lines (Figure 4B).

This higher order property of the pattern can be also characterized by changes in the frequency domain, namely now the two dimensional power spectrum will not be radially

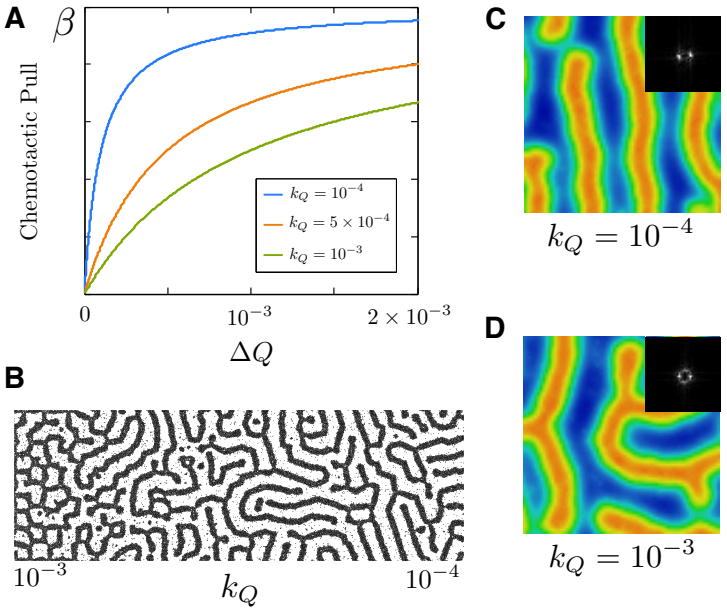


Figure 4: Pattern "straightness" dependence on the saturation parameter k_Q of the chemotactic process. (**A**) shows the chemotactic pull as a function of the chemoattractant difference between lattice sites ΔQ for different values of k_Q, which saturates at value β following our Michaelis-Menten like formulation previously introduced. In (**B**) we display the effect of varying k_Q linearly on cell distributions for an otherwise constant parameter set ($D = 0.15$, $\rho = 6 \times 10^{-4}$, $\beta = 0.3$), when the signal gradient saturates more readily the striped domains are more likely to extend without twisting. (**C-D**) shows the chemotactic signal distribution for two values of k_Q (10^{-4} top, 10^{-3} bottom) and their respective two dimensional Fourier Transforms.

homogenous (Figure 4C). This reflects that the regularly spaced correlation in morphogen concentration is only found in a particular direction, in stark contrast to the previously found radial symmetry when k_Q has a larger value (Figure 4D).

Acknowledgement

We thank the members of the Complex Systems Lab for useful discussions. This work was supported by MINECO BES-2010-038940 fellowship (SDN), by the Fundación Botín through

its global Universities program and the Santa Fe Institute.

References

(1) Goodwin, B. C. *Comptes Rendus de l'Académie des Sciences - Series III - Sciences de la Vie* **2000**, *323*, 15–21.

(2) Kondo, S.; Miura, T. *Science* **2010**, *329*, 1616–20.

(3) Green, J. B. a.; Sharpe, J. *Development* **2015**, *142*, 1203–1211.

(4) Basu, S.; Gerchman, Y.; Collins, C. H.; Arnold, F. H.; Weiss, R. *Nature* **2005**, *434*, 1130–4.

(5) Turing, A. *Philosophical Transactions of the Royal Society* **1952**, *237*.

(6) Ball, P. *Philosophical Transactions of the Royal Society of London B: Biological Sciences* **2015**, *370*, 20140218.

(7) Murray, J. D. *Journal of theoretical biology* **1982**, *98*, 143–163.

(8) Diambra, L.; Senthivel, V. R.; Barcena Menendez, D.; Isalan, M. *ACS synthetic biology* **2014**,

(9) Koch, A.; Meinhardt, H. *Reviews of modern physics* **1994**, *66*, 1481.

(10) Gierer, A.; Meinhardt, H. *Kybernetik* **1972**, *12*, 30–39.

(11) Meinhardt, H. *Interface Focus* **2012**, rsfs20110097.

(12) Kondo, S.; Asai, R. *Nature* **1995**, *376*, 765–768.

(13) Nakamasu, A. *Proceedings of the National Academy of Sciences* **2009**, *106*, 8429–8434.

(14) Kawasaki, K.; Mochizuki, a.; Matsushita, M.; Umeda, T.; Shigesada, N. *Journal of theoretical biology* **1997**, *188*, 177–185.

(15) Lejeune, O.; Tlidi, M. *Journal of Vegetation science* **1999**, *10*, 201–208.

(16) Thar, R.; Kühl, M. *FEMS Microbiology Letters* **2005**, *246*, 75–79.

(17) Deng, P.; de Vargas Roditi, L.; van Ditmarsch, D.; Xavier, J. B. *New journal of physics* **2014**, *16*, 015006–15006.

(18) Sheth, R.; Marcon, L.; Bastida, M. F.; Junco, M.; Quintana, L.; Dahn, R.; Kmita, M.; Sharpe, J.; Ros, M. a. *Science (New York, N.Y.)* **2012**, *338*, 1476–80.

(19) Raspopovic, J.; Marcon, L.; Russo, L.; Sharpe, J. *Science* **2014**, *345*, 566–570.

(20) Cocho, G.; Pérez-Pascual, R.; Rius, J. *Journal of theoretical biology* **1987**, 419–435.

(21) Bonabeau, E. *Artificial life* **1997**, *3*, 191–211.

(22) Adamatzky, A.; Martínez, G. J.; Mora, J. C. S. T. *International Journal of Bifurcation and Chaos* **2006**, *16*, 2985–3005.

(23) Frohnhöfer, H. G.; Krauss, J.; Maischein, H.-m.; Nüsslein-volhard, C. *Development* **2013**, *3007*, 2997–3007.

(24) Mahalwar, P.; Walderich, B.; Singh, A. P. *Science* **2014**, *345*, 1362–1364.

(25) Edelstein-Keshet, L.; Ermentrout, G. *Journal of mathematical biology* **1990**, 33–58.

(26) Chen, T.-H.; Guo, C.; Zhao, X.; Yao, Y.; Boström, K. I.; Wong, M. N.; Tintut, Y.; Demer, L. L.; Ho, C.-M.; Garfinkel, A. *Interface focus* **2012**, *2*, 457–64.

(27) Seppa, H.; Grotendorst, G.; Seppa, S.; Schiffmann, E.; Martin, G. R. *J Cell Biol.* **1982**, *92*, 584–588.

(28) Postlethwaite, a. E.; Seyer, J. M. *The Journal of experimental medicine* **1990**, *172*, 1749–1756.

(29) Kuo, S. C.; Koshland, D. *Journal of bacteriology* **1987**, *169*, 1307–1314.

(30) Alon, U.; Surette, M. G.; Barkai, N.; Leibler, S. *Nature* **1999**, *397*, 168–171.

(31) Vilcek, J.; Palombella, V. J.; Henriksen-DeStefano, D.; Swenson, C.; Feinman, R.; Hirai, M.; Tsujimoto, M. *The Journal of experimental medicine* **1986**, *163*, 632–643.

(32) Battegay, E. J.; Raines, E. W.; Colbert, T.; Ross, R. *The Journal of Immunology* **1995**, *154*, 6040–6047.

(33) Mansbridge, J. N.; Liu, K.; Pinney, R. E.; Patch, R.; Ratcliffe, A.; Naughton, G. K. *Diabetes, Obesity and Metabolism* **1999**, *1*, 265–279.

(34) Miura, M.; Friedlander, R. M.; Yuan, J. *Proceedings of the National Academy of Sciences* **1995**, *92*, 8318–8322.

(35) Mohan, R. R.; Mohan, R. R.; Kim, W.-J.; Wilson, S. E. *Invest Ophthalmol Vis Sci* **2000**, *41*, 1327–1335.

(36) Wajant, H.; Pfizenmaier, K.; Scheurich, P. *Cell Death & Differentiation* **2003**, *10*, 45–65.

(37) Nakamasu, A.; Takahashi, G.; Kanbe, A.; Kondo, S. *Proceedings of the National Academy of Sciences* **2009**, *106*, 8429–8434.

(38) Shoji, H.; Mochizuki, A.; Iwasa, Y.; Hirata, M.; Watanabe, T.; Hioki, S.; Kondo, S. *Developmental Dynamics* **2003**, *226*, 627–633.

(39) Shoji, H.; Iwasa, Y. *Journal of Theoretical Biology* **2005**, *237*, 104–116.

(40) Painter, K.; Maini, P.; Othmer, H. *Proceedings of the National Academy of Sciences* **1999**, *96*, 5549–5554.

(41) Ball, P.; Borley, N. R. *The self-made tapestry: pattern formation in nature*; Oxford University Press Oxford, 1999; Vol. 198.

Bonforti A, Duran-Nebreda S, Montañez R, Solé R. Spatial self-organization in hybrid models of multicellular adhesion. Chaos. 2016 Oct;26(10):103113. DOI: 10163/1.4965992.

Self-organized pattern formation in a synthetic microbial population

Salva Duran-Nebreda[†,‡] and Ricard V. Solé[*,†,‡,¶]

†ICREA-Complex Systems Lab, Universitat Pompeu Fabra, 08003 Barcelona

‡Institute of Evolutionary Biology, UPF-CSIC, 08003 Barcelona

¶Santa Fe Institute, 1399 Hyde Park Road, 87501 Santa Fe, New Mexico, USA

E-mail: ricard.sole@upf.edu

Abstract

Modern multicellular entities are characterized by exquisite structures, intimately related to the functions they perform. The controlled construction and maintenance of form is paramount to ensure survival in these beings and have been subject of great interest to comprehend the constraints and potentials underlying multicellular evolution. An example mechanism of pattern formation for the breakage of symmetry present in many biological systems can be found in Turing's model, involving decentralized communication through morphogens. Here we report a synthetic self-organized mechanism stemming from the embodiment of bacterial cells, leading to the establishment of periodic branching structures from initial compositional symmetry. The artificial construction of form approach taken here can help validate developmental theories and identify universal properties underpinning biological pattern formation.

Introduction

An important issue in the creation of form in multicellular metazoans is the breakage of symmetry that characterizes the early stages of embryo development.[1–3] Starting from a spheroid composed of equal cells, the pattern formation processes quickly establish non-homogenous distributions of cellular states -i.e. distinct profiles of gene expression-, that transform into the major structures of the developing animal.[4,5]

An example mechanism that provides the basis for the creation of non-uniform distributions of signaling molecules -or morphogens- from essentially deterministic rules with random initial fluctuations was formulated by Turing.[6,7] In his seminal paper *On the chemical basis of morphogenesis*, Turing suggested that a system composed of two diffusing and interacting molecules could explain present innate *diffusion-driven instabilities* under some parametric conditions. Later expanded by others,[8–11] Turing patterns have come to be regarded as a paradigmatic example of simple systems yielding complex features, and are thought to underpin structures across scales and domains, including skin pigmentation in animals,[12–14] primordia of skeletal elements,[15–18] palatal ridges,[19] teeth formation,[20] establishment of hair follicles[21] and chemical,[22–24] and ecological[25–29] structures among others.

Nonetheless, experimental evidence for the existence of systems following Turing's proposal has been scarce, with most advances based on the modeling approach to pattern formation. Only recently, some experiments have begun to unravel the picture of Turing-type mechanisms in the formation of skeletal limb primordia,[17,18] suggesting that more morphogens and more complicated interaction maps might me implicated than originally thought.

Another possibility for inquiry into pattern formation is given by synthetic biology, which tries to artificially construct -or reconstruct- complex functions and features by splicing together coding and regulatory sequences that do not naturally coexist.[30–32] Such approach has enabled researchers to mimic gradient-based positional information developmental systems,[33,34] however self-organized systems -characterized by simple interacting units and hor-

izontal exchanges of information like in Turing's mechanism- have proven to be more elusive to engineer.[35]

Here we report a novel way of synthetically constructing a self-organized process in bacteria that involves communication, filamentous growth, adhesion and growth inhibition. Our engineered cells are able to create branching structures characterized by higher order regularities in cellular densities, similar to those created by Turing-type mechanisms. This research provides an important milestone in the establishment of mutual feedbacks between experimental embryology, modeling of pattern formation and synthetic strategies to reconstruct putative mechanisms and interactions.

Materials and Methods

DNA constructs and plasmids

Final genetic constructs used in this work were generated using the standard biobrick cloning techniques and enzymes:EcoRI/XbaI/SpeI/PstI restrictases and T4 DNA ligase (New England Biolabs, USA). Some DNA sequences were provided by the iGEM 2010 spring collection, including *LuxR* (C0062), *LuxI* (C0161), *pLux* (R0062), *MinC* (K299806), *GFP* (E0040), constitutive promoter (J23100), bidirectional terminator (B0014) and RBSs (B0033, B0034). *JunA* was formatted to biobrick standard 10 from a coding sequence kindly provided by L.A. Fernández. The inefficient Lux promoter *pL40* was created *de novo* by primer hybridization (Sigma Aldrich, USA). See supplementary materials for sequences of all used DNA pieces.

Genetic devices were split between two plasmids pSB1AC3 and pSB3K5, also obtained from the iGEM 2010 distribution, with high and high-intermediate copy numbers respectively. The two plasmids harbor different origins of replication, and can coexist inside a single cell. All final constructs were sequenced by the PRBB core facilities.

Bacterial strains and growth conditions

Cloning procedures were carried out in *E. coli* Top10 strain (Invitrogen, USA). Final essays were performed in *E. coli* UT5600 kindly provided by L.A. Fernández.

Colony essays were performed as follows: UT5600 cells harboring each device were fresh plated overnight from a glycerinate stored at -80C, a single colony was then grown in Lisogenic Broth (Sigma Aldrich, USA) supplemented with Chloramphenicol and Kanamycin (Sigma Aldrich, USA) for 5 hours and diluted to $Abs_{660} = 0.2$. A small volume (2 μL) of the density adjusted cultures was dropped in the center of 5.5 cm petri dishes, filled with 5.5 mL of LB Eiken agar (Eiken Chemical, Japan) at 0.4% w/v again supplemented with Chloramphenicol and Kanamycin and, when necessary, 10^{-8} M N-[β-ketocaproyl]-L-homoserine lactone (Cayman Chemical Company, USA). Inoculated plates were dried for 5 minutes and grown 1 4h at 37C, then stored at 22C for 7 days, were data capture took place.

Data capture and processing

Assessment of lactone concentration impact on strain growth was carried out in Synergy MX microplate reader (BioTek Instruments, USA), similarly to our previously described protocol.[36] Photographs of colony pattern were taken daily with a Canon EOS with diffuse illumination. Initial and final state of the pattern formation process were captured by bright field and fluorescence microscopy with a Leica DMI6000B (Leica Mycrosystems, USA). Regularities in colony boundaries were characterized with Matlab 2013b polar transformation and FFT algorithms (MathWorks, USA). All images were processed with a background subtraction and brightness adjustment.

Results

Non-homogeneous spatial distributions of cells arise at the intersection between adhesion, cell signaling and filamentous growth

In this study we have assessed the impact of four genes (Figure 1A) in the colony growth of *E. coli* UT5600 strain, which typically develops into uniform circular colonies as time progresses. The genes synthetically introduced in the model organism comprise different effects in cell morphology, growth rates cell-cell adhesion and cell-cell signaling through quorum sensing. The first two items are given by *MinC*, a protein involved in *E. coli* segmentation.[37,38] In particular, exogenous expression of this protein has been linked to the elongation of cells with the formation of filaments that can attain two orders of magnitude the typical cell length and entails a diminished biomass growth.[39] Cell-cell adhesion is introduced by a chimeric protein composed of the animal *JunA* coupled with an autotranslocator domain.[40] This chimeric sequence is able to target the membrane and homodimerize with equal proteins expressed by other cells, increasing the sedimentation rates in liquid cultures of bacteria. Finally, communication was incorporated by the expression of two components from quorum sensing system of *V. fisherii*, widely used by the synthetic biology community.[33–35,41,42] This is typically composed of a receptor protein (*LuxR*), able to enhance expression in specific promoter sequences in the presence of the ligand homoserin lactone (hereafter HSL), and the *LuxI* gene, able to synthesize the cognate molecule from preexisting substrates. This family of ligand molecules can passively diffuse across the cellular membranes, reaching high concentrations naturally when cellular density surpasses a threshold.

In order to explore the landscape of pattern formation capabilities of these genes implicated in cell morphology, adhesion and communication, all possible combinations of them were constructed (Figure 1A) and their impact in colony growth was tested (Figure 1B). Those conditions lacking the ability to synthesize the signaling molecule were externally supplemented with HSL in the petri dish to ensure similar levels of *MinC* and *JunA* expres-

sion, but otherwise lack the spatial information given by the quorum sensing mechanism. Only when all three capabilities were included in the synthetic cells spatial non-homogenous distributions of cell densities were created, in stark contrast to the other conditions were circular growth took place.

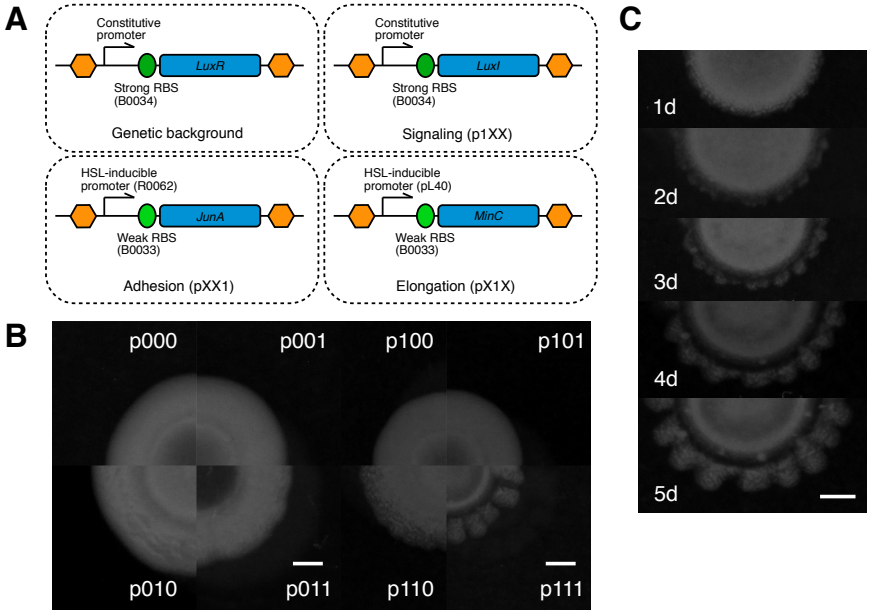


Figure 1: (**A**) Basic scheme of the constructs used in this work. Receiver gene (*LuxR*) was present in all tested strains while the other three were introduced in all possible combinations. Notation goes as follows, each bit indicates the presence or absence of a particular feature, from left to right: signaling, elongation and adhesion. (**B**) The eight strains evaluated in this work after 5 days of culture. The group on the right contains all combinations capable of cell-cell communication, while those on the left were exogenously supplemented with the signaling molecule. (**C**) Time series of p111 colony growth, daily progession. Scale bar in all pictures is 2 mm.

Branch distributions are characterized by a dominant wavelength

As it can be seen in Figure 1B, when all four genes are present the symmetry of the colony is broken in a regular fashion. In order to characterize the higher order properties displayed by the growing structure, we analyzed the distribution of bacterial concentrations in a cir-

cumference centered at the origin of the colony (Figure 2A,B). After transforming the data from cartesian to polar coordinates and applying a FFT algorithm, bacterial density was found to display a fixed wavelength correlation at approximately 0.2 cm (Figure 2C).

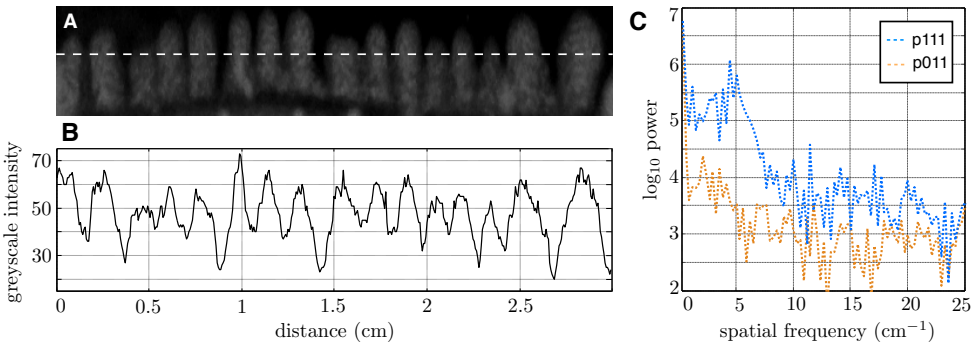


Figure 2: Characterization of the regular structures created by bacterial density. (**A**) 5 day colony picture transformed into polar coordinates. (**B**) Greyscale intensity (a.u.) from the previous data is used as a surrogate of bacterial density close to the edge of the colony. (**C**) Fourier Transform analysis of the radial profile of the data, showing the existence of a peak in the power spectrum located at approximately 0.2 cm for the p111 strain. Single colony assessment, for comparison we show the same analysis for a unstructured p011 colony.

Branch primordia are formed by the extension of cohesive bundles of filaments

In order to assess the micro-scale properties of the pattern formation process we captured the first stages of colony growth with bright field and fluorescence microscopy (Figure 3A). In particular, we observed that the starting symmetry imposed by the circular droplet of cells is broken as soon as 24 hours when the whole set of genes is present. However, these primordia of the branching pattern do not have the regularity displayed by the colony at later stages, and are characterized by the formation of cohesive bundles of cells with the same orientation. Given that *E. coli* segmentation occurs perpendicular to the longest axis of the cell, the establishment of a collective orientation forcefully imposes growth in a preferred direction, which cells maintain in the following days.

Moreover, after the fifth day of culture some high-density regions split into smaller

outward-growing branches (Figure 3B). These are also characterized by the formation of coherent bundles of cells at the edge of a branch, which otherwise form swirls of filaments without a particular direction. The formation of periodic structures that are able to self-regulate into domains of a particular size suggests that a lateral inhibition or Turing-type mechanism might underlie the patterns reported here.

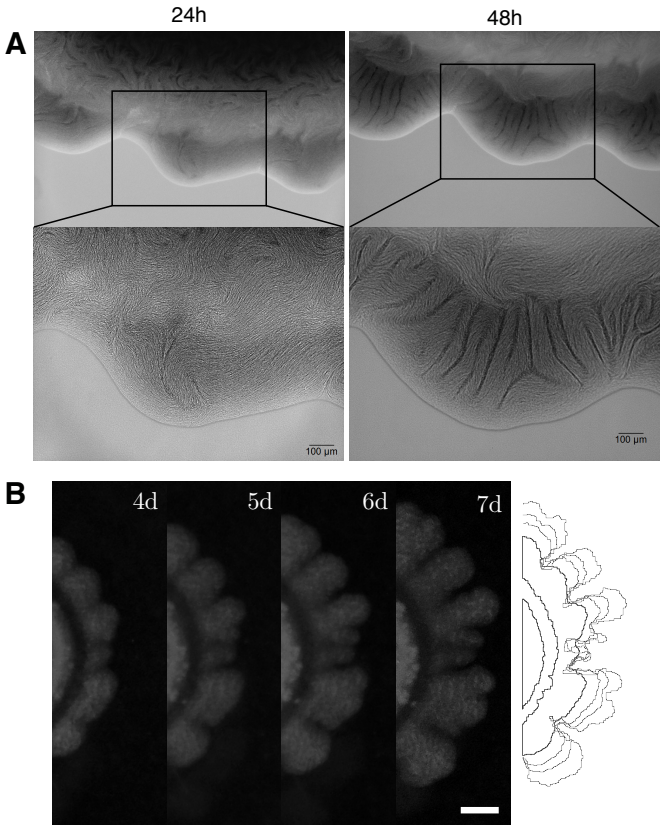


Figure 3: Panel exploring the temporal dynamics of the pattern formation process. (**A**) displays the same region of a p111 strain colony after 24 and 48 hours of growth in bright field microscopy. On the top, a wider region containing several branches, below the detail of cell bundles forming at the edge of the growing colony. In (**B**) we show the daily development after the fourth day of a p111 colony and the aggregate colony outline for the whole set, scale bar here is 2 mm.

Lactone-induced expression of *MinC* and *JunA* inhibits growth of synthetic cells

Following the previous examination we characterized the impact of the secreted quorum sensing molecule in the growth rate of cells. This was tested by monitoring bacterial density for different constructs at Abs_{660} in a liquid culture, with increasing amounts of externally introduced lactone. Figure 4 shows how the p011 strain (able to express *MinC* and *JunA* in the presence of inducer but unable to constitutively synthesize the signaling molecule) displays varying rates of growth and carrying capacities dependent on the concentration of the signal. This necessarily implies that both a decrease in growth speed and a cell death response are mediated by *MinC* and *JunA*, suggesting that this mechanism might be supported by a lateral inhibition between branches and local activation due to cellular division.

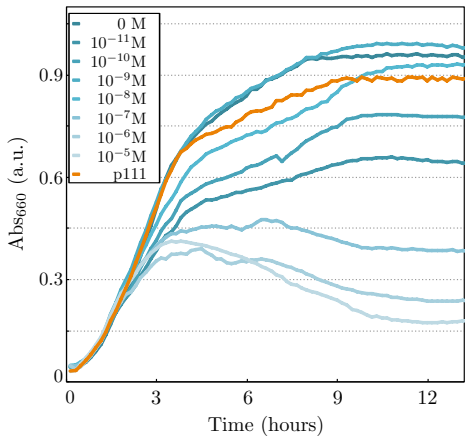


Figure 4: Characterization of the effect of *MinC* and *JunA* expression on the growth of synthetic cells. For different concentrations of the quorum sensing molecule, the construct p011, harboring constituive *LuxR* and conditional expresion of *MinC* and *JunA*, shows different growth rates and carrying capacity (shades of blue). As a control p111, capable of endogenously synthesizing lactone, displays a growth rate similar to 10^{-9}M of the previous case (orange).

Discussion

Models of pattern formation in biological systems have provided a solid theoretical background upon which research in the developmental processes can base further inquiry. Given the existing wealth of knowledge provided by the theoretical and characterization enterprises, synthetic biology can be used to consolidate our understanding of such complex processes by reconstructing the molecular relations thought to underlie the mechanisms of pattern formation with unrelated pieces and functions.

The case of Turing-type mechanisms is specially enticing since it was formulated before evidence to support it had appeared, which spurred great controversy regarding their relevance in biological settings. Our work enlarges this picture by providing patterns similar to those created by reaction-diffusion systems in the synthetic domain, suggesting that more complicated interactions including the specifics of cellular embodiment may give rise to the family of periodic structures that are now considered to be caused by Turing mechanisms.

Acknowledgement

We thank the members of the Complex Systems Lab for useful discussions. We would like to thank LA Fernandez and the Registry of Biological parts (MIT) for the source material provided. This work was supported by MINECO BES-2010 fellowship (SDN), by the Fundación Botín through its global Universities program and the Santa Fe Institute.

References

(1) Leptin, M. *The EMBO journal* **1999**, *18*, 3187–3192.

(2) Forgacs, G.; Newman, S. A. *Biological physics of the developing embryo*; Cambridge University Press, 2005.

(3) Blum, M.; Schweickert, A.; Vick, P.; Wright, C. V.; Danilchik, M. V. *Developmental biology* **2014**, *393*, 109–123.

(4) Lemaire, L.; Kessel, M. *Mechanisms of development* **1997**, *67*, 3–16.

(5) Durston, A.; Jansen, H.; Wacker, S. *Genomics* **2010**, *95*, 250–255.

(6) Turing, A. *Philosophical Transactions of the Royal Society* **1952**, *237*.

(7) Ball, P. *Philosophical Transactions of the Royal Society of London B: Biological Sciences* **2015**, *370*, 20140218.

(8) Gierer, A.; Meinhardt, H. *Kybernetik* **1972**, *12*, 30–39.

(9) Murray, J. D. *Journal of theoretical biology* **1982**, *98*, 143–163.

(10) Koch, A.; Meinhardt, H. *Reviews of modern physics* **1994**, *66*, 1481.

(11) Meinhardt, H. *Interface Focus* **2012**, rsfs20110097.

(12) Kondo, S.; Asai, R. *Nature* **1995**, *376*, 765–768.

(13) *Science* **2010**, *329*, 1616–20.

(14) Murray, J. D. *Scientific American* **1988**, *258*, 80–87.

(15) Hentschel, H.; Glimm, T.; Glazier, J. A.; Newman, S. A. *Proceedings of the Royal Society of London B* **2004**, *271*, 1713–1722.

(16) Newman, S. A.; Bhat, R. *Birth Defects Research Part C: Embryo Today: Reviews* **2007**, *81*, 305–319.

(17) Sheth, R.; Marcon, L.; Bastida, M. F.; Junco, M.; Quintana, L.; Dahn, R.; Kmita, M.; Sharpe, J.; Ros, M. a. *Science* **2012**, *338*, 1476–80.

(18) Raspopovic, J.; Marcon, L.; Russo, L.; Sharpe, J. *Science* **2014**, *345*, 566–570.

(19) Economou, A. D.; Ohazama, A.; Porntaveetus, T.; Sharpe, P. T.; Kondo, S.; Basson, M. A.; Gritli-Linde, A.; Cobourne, M. T.; Green, J. B. *Nature genetics* **2012**, *44*, 348–351.

(20) Jernvall, J. *Proceedings of the National Academy of Sciences* **2002**, *99*, 8116–8120.

(21) Sick, S.; Reinker, S.; Timmer, J.; Schlake, T. *Science* **2006**, *314*, 1447–1450.

(22) Castets, V.; Dulos, E.; Boissonade, J.; De Kepper, P. *Physical Review Letters* **1990**, *64*, 2953.

(23) Ouyang, Q.; Swinney, H. L. *Nature* **1991**, *352*, 610–612.

(24) Kapral, R.; Showalter, K. *Chemical waves and patterns*; Springer Science & Business Media, 2012; Vol. 10.

(25) Lejeune, O.; Tlidi, M. *Journal of Vegetation science* **1999**, *10*, 201–208.

(26) HilleRisLambers, R.; Rietkerk, M.; van den Bosch, F.; Prins, H. H.; de Kroon, H. *Ecology* **2001**, *82*, 50–61.

(27) Alonso, D.; Bartumeus, F.; Catalan, J. *Ecology* **2002**, *83*, 28–34.

(28) Thar, R.; Kühl, M. *FEMS Microbiology Letters* **2005**, *246*, 75–79.

(29) Baurmann, M.; Gross, T.; Feudel, U. *Journal of Theoretical Biology* **2007**, *245*, 220–229.

(30) Andrianantoandro, E.; Basu, S.; Karig, D. K.; Weiss, R. *Molecular systems biology* **2006**, *2*.

(31) Purnick, P. E.; Weiss, R. *Nature reviews Molecular cell biology* **2009**, *10*, 410–422.

(32) Khalil, A. S.; Collins, J. J. *Nature Reviews Genetics* **2010**, *11*, 367–379.

(33) Basu, S.; Gerchman, Y.; Collins, C. H.; Arnold, F. H.; Weiss, R. *Nature* **2005**, *434*, 1130–4.

(34) Levskaya, A.; Chevalier, A. A.; Tabor, J. J.; Simpson, Z. B.; Lavery, L. A.; Levy, M.; Davidson, E. A.; Scouras, A.; Ellington, A. D.; Marcotte, E. M. *Nature* **2005**, *438*, 441–442.

(35) Liu, C.; Fu, X.; Liu, L.; Ren, X.; Chau, C. K.; Li, S.; Xiang, L.; Zeng, H.; Chen, G.; Tang, L.-H. *Science* **2011**, *334*, 238–241.

(36) Carbonell-Ballestero, M.; Duran-Nebreda, S.; Montañez, R.; Solé, R.; Macía, J.; Rodríguez-Caso, C. *Nucleic acids research* **2014**, *42*, 14060–14069.

(37) Erickson, H. P. *Trends in cell biology* **1997**, *7*, 362–367.

(38) Hu, Z.; Lutkenhaus, J. *Journal of bacteriology* **2000**, *182*, 3965–3971.

(39) Bhomkar, P.; Materi, W.; Wishart, D. *PloS one* **2011**, *6*, e27559.

(40) Veiga, E.; de Lorenzo, V.; Fernández, L. A. *Journal of bacteriology* **2003**, *185*, 5585–5590.

(41) Fuqua, W. C.; Winans, S. C.; Greenberg, E. P. *Journal of bacteriology* **1994**, *176*, 269.

(42) Tamsir, A.; Tabor, J. J.; Voigt, C. A. *Nature* **2011**, *469*, 212–215.

Chapter 4

DISCUSSION

The work presented here has a tried to shed some light into the origins of multicellularity. As discussed in previous chapters, this collection of events in the evolution of our biosphere still presents some open questions, specially with regards to the evolution and maintenance of collective structures and new levels of organization and individuality. Taking very different approaches like synthetic biology, artificial life and modeling of artificial evolution, we have tried to gain some insights into particular cases as well as proposing more general scenarios for the evolution of MC and the formation of complex bodily structures. Such endeavor is tied to a very particular step in the evolution of complexity in our biosphere, yet the combined approaches used here might offer some guidance in unraveling the broader picture.

In the first part of this thesis, we have modeled two different systems containing an evolutionary transition to MC. One of them, the microcosm created by Ratcliff and colleagues [Ratcliff et al., 2012], is a recent example of artificial evolution of MC in the substrate of a facultative, but mostly UC organism. Since this work was published, other promising efforts have been directed towards only UC species, with equally remarkable results [Ratcliff et al., 2013]. In all these instances of emergence of MC, externally imposed constraints are the main drivers of the process. Specifically, the researchers make use of differential sedimentation rates

to select for larger objects.

This approach might come as rather trivial -selecting for bigger entities is bound to yield either bigger cells or the creation of aggregates-, yet the real value of these works lies in providing a *tempo* for the evolution of simple MC. In remarkably few rounds of selection -far from the geological times that are commonly assumed when thinking about the evolution of complex traits-, a repeatable and stable multicellular phenotype was reached. The particular mutations leading to this morphological changes are yet to be understood, however some clues might be found in the mechanisms used by cells. Namely defective fission, which as introduced in our functional description of MC is an effective method of manipulating spatial structuring.

Another feature that the authors discussed in their article is the evolution of programmed cell death. It was shown that apoptotic cells were present in their evolved yeast strain and were non-randomly distributed within the cluster. This was deemed advantageous, for it might enhance collective fitness through the facilitation of group-level division and faster growth [Libby et al., 2014]. If true, it would imply a clear division of labor: some cells forgo their individual fitness and in doing so promote survival and reproduction at the group level.

However, in our modeling efforts of this microcosm, we have found that including accepted metabolic features of yeast generates a similar patterning of localized mortality. Ratcliff and coworkers had previously argued that the observed non-linear trend between size and the prevalence of apoptosis ruled out direct causality between these two variables, and that an active signaling process had to be behind it. Yet our model showed that the null hypothesis -that no signaling process is directly responsible for this trend-, still stands after incorporating an extended description of yeast biology. This does not diminish the relevance of this experimental setting, quite the contrary, it brings to the general attention that features normally considered detrimental might operate in a beneficial manner under the right circumstances.

The second manuscript deals with a different scenario than the experimental set up elaborated by Ratcliff. Incorporating the possibility of

multistability, regulation of spatial arrangement and a fitness-relevant division of labor we showed that an appropriately challenging environment can drive the evolution of differentiated MC. Our proposed setup includes the existence of nourishing as well as toxic substances in the environment. This was designed to be as general as possible while still resembling current [Pattus et al., 1990, Ben-Jacob et al., 2000, Hibbing et al., 2010] and past [Schirrmeister et al., 2013] situations that UC organisms might have faced.

In particular, we included multistability in a rather simple form, using a stochastic transition between cell types. These memory-less dynamics are similar to the mechanisms of phase variation in bacterial species [Henderson et al., 1999]. In these unicellular organisms the stochastic transition between cell types imposes phenotypic plasticity at the population level, creating cells with discrete variations in key properties such as antibiotic resistance or the antigenic display [Lewis, 2007]. The discreteness does not need to be enforced by multistable GRNs, unregulated genome rearrangements -inversions in particular- are known to switch off some coding sequences while activating others, providing the basis of SPS.

This mode of phenotypic plasticity has been linked to bet hedging strategies that enable survival in the face of a changing hostile environment [Veening et al., 2008]. However, as with the previously discussed DPMs, this underlying mechanisms can be co-opted for pattern formation when coupled with an energy minimization process.

When taken together, these two projects deal with the two possible origins of MC: aggregative and colonial. As discussed in section 1.2.1, the origin of the lower level entities determines the reach of potential conflicts arising in a group. Even if the constituent particles are from the same species -as in fraternal MTEs-, more disparity translates into increased competition. Then, aggregative modes of MC are bound to display greater conflicts than clonal ones, for in clonal development collectives begin as genetically uniform and cheater strategies must appear from mutations, while aggregates can be non-homogeneous from the start.

However, as shown in the second manuscript, cells in an aggregate

can make use of the tools provided to them to reduce potential conflict. This is readily seen in the second example of the genetic algorithm, where a particular strain evolves the capacity to degrade waste. This population -which coexists with other species that can behave as cheaters-, also displays higher self-cohesion than the other strains, which causes them to be spatially segregated. That is, in our model these cells manipulate their local relatedness and evade cheaters with the mechanism of cell sorting. This ensures that the investment in the common good that is the degradation of waste has a bigger impact on cells close in genetic distances.

This introduces an interesting twist to the theme of relatedness and the evolution of cooperative behavior started by Hamilton [West et al., 2002], similar to the paradigm shift that supposed the introduction of explicit space in meanfield ecological models [Bascompte and Solé, 1996]. In particular, the fact that cells can evolve strategies to correlate euclidean with genetic distances enables the propagation of mutually beneficial interactions among the population. Obviously, this surely must also enable cheaters to evolve the inverse strategy -being surrounded by cooperators- in order to maximize their fitness. This was not observed in our simulations yet we confer that a race in adhesion properties is what should be expected in the long run, providing an interesting testbed for the so called red queen dynamics [Dieckmann et al., 1995].

In the second part of this thesis, pattern formation, we have provided two new minimal sets of rules capable of generating spatial regularities. This method of building structure, commonly referred as Turing patterns, is widely regarded as very relevant in the development of unrelated systems across scales [Maini, 1997] (Figure 1.12).

Particularly, we have demonstrated that a coupling of chemotactic movement and programmed cell death reliably generates the kind of structures naturally observed in mammalian and fish skin patterns. This endeavor has also led us to identify the relevant physico-chemical properties of the system (diffusion, speed of movement and signal saturation) with regards to pattern specification. That is, whether spots, stripes or mazes are created, but also the higher order properties of these patterns -whether they are more rounded or straight-.

DISCUSSION

Additionally, we have shown that SPS opposes cell sorting in the minimization of free energy of our model lattice system. At the interplay between these two mechanisms periodic structures were created, remarkably without the inclusion of the standard morphogens or diffusing molecules. Moreover, our proposed system with basic rules linked the capacity of cells to arrange themselves in particular structures to increased chances of survival at the individual level, linking collective properties to individual fitness.

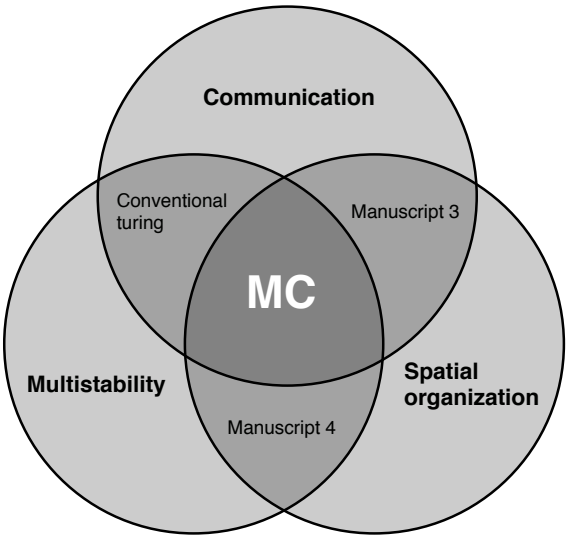


Figure 4.1: Location of our proposed models of periodic pattern formation in the functional characterization of multicellularity diagram.

These two projects have been inspired by the recent discovery that cell movement might be of some relevance in the creation of periodic structures in zebrafish [Frohnhöfer et al., 2013, Mahalwar et al., 2014]. At the same time, some authors have suggested that the approach taken by Turing -to consider that both morphogens *diffuse*- might be inappropriate when one of such morphogens is a cell-sized entity. This issue is avoided in most formulations of Turing mechanisms by considering that both the activator and the inhibitor are sub-cellular entities in great numbers. Then,

whether small chemicals or proteins, their motion can be aptly described by Fick's law. However, as we and others have shown [Babloyantz, 1977, Bullara and Decker, 2015], this does not mean that periodic structures cannot be created using cell-sized morphogens.

We have also explicitly dealt with the fact that it is cells, with discrete states, that create the pattern with their pigments. In reaction-diffusion models, the connection to cellular types is either lost or given by a simple threshold function, applied to one or the two morphogens in order to determine the particular state of a cell. In real systems, the cells that conform the pattern are either pigmented or not -melanophores and iridiophores in zebrafish for instance-, which typically yields sharp boundaries between domains. If naturally occurring Turing systems make use of discrete cell types, we can conclude that they must lie at the intersection between multistability and communication in our previously shown functional characterization of MC (Figure 4.1). In this figure we also provide the locations of other works presented here, and observe that remarkably, the potential for generating periodic structures can be found at each of the pairwise intersections of these three properties.

Taking advantage of embodiment

A particularly relevant theme discussed throughout this text, is the impact of cellular embodiment in implementing pattern formation and its connection to the emergence of MC. By cellular embodiment it is usually meant the purely physical description of cells, an aspect that is sometimes overlooked, most notably after extreme simplifications that collude cells to point objects or static compartments. This simplifications are sometimes necessary and make sense, yet as shown by Newman and Bhat in the DPMs, the physico-chemical embedding in which the processes of MC take place are of utmost relevance in determining possible outcomes and modes of patterning. Even under very simplistic assumptions, complex structures are often obtained as soon as physical interactions are put in place [Niklas, 1994, Eggenberger, 1997, Kaandorp et al., 2008, Solé et al., 2013]

DISCUSSION

Our contribution to this subject can be found in different sections of this thesis. In the modeling of pattern formation processes and the emergence of MC we have shown how important it is to use the appropriate level of description and how the purely physical structures of cells can introduce inevitable constraints. These constraints can sometimes be "boons" at other levels of description, in the sense that they can enable new ways of patterning or induce higher order properties in the system. This is the case of the introduction of chemotactic movement and signal saturation in the third manuscript, the inclusion of an extended spatial description of yeast cells and their metabolic processes in the first manuscript of the same experimental setting- and the link between differential adhesion and the evolution of cooperation in the second manuscript.

In the case of our synthetic biology approach to periodic patterns, we have shown how relevant is the fact that cells become more elongated as the signal increases. Our bacteria end up being flexible strings with strong binding forces between them, which is far from the kind setting a typical RD system can deal with. In particular, as mentioned in the manuscript, this enables the coherent growth of bundles of cells, which form the primordia of the branches. Without accounting for this physically process, for instance by using spherical cells, it would be difficult to faithfully reproduce the mechanism that creates this branching structures.

All these aspects come to reinforce the important role that physical processes -specially gravity, adhesion and diffusion- could have had in the evolution of early MC or simple pattern formation [Alberch, 1980, Newman and Bhat, 2008, Bonner, 2013, Solé and Valverde, 2013], even asserting dominance over complex regulatory processes -such as those found in current developmental plans-. At the same time, this sets up a path to research the origins of MC: by including richer interactions among cells and by providing them the tools to recruit physical processes, increasingly complex forms should be expected.

Although physics and embodiment are usually discussed in the context -or at the level of- organisms or tissues, there is another level of embodiment that requires attention: the external world, whose fluctuations and properties influence the repertoire of available adaptations. Thus,

other factors playing a role in the early stages of multicellularity, including the ecological context and the physics of the environment should also be taken into account.

Artificial approaches to evolutionary change should be a natural component of the exploration of macroevolutionary patterns and the tempo and mode of the major transitions. Despite their limitations, they offer what no other approach can: an opportunity to recreate the past and how complexity developed over time.

Opportunities in the synthetic domain

Another avenue to address the dawn of multicellular systems is provided by synthetic biology [Davies, 2008, Chuang, 2012, Maharbiz, 2012]. By engineering unicellular systems it is possible to build novel forms of cell-cell interaction, whether physical or communication mediated. This will enable scientists to create, or perhaps re-create, forms of multicellular assemblies, able to perform novel functions and even complex computations [Regot et al., 2011].

This will require greater efforts in characterizing *effectors* instead of transcription factors [Davies, 2008]. With a quick glance at the repositories of genetic parts it can be observed that the mostly used and characterized parts in synthetic biology are genetic regulators. However, the path suggested here to attain simple MC deals with interactions between cells, not the relaying of information or the ability to perform complex computations. This is the approach we have tried in the fifth manuscript: to use very simple modes of communication and entrust the patterning process to physical phenomena.

Given the potential of genetic engineering techniques to alter the logic of cell-cell exchanges, synthetic biology offers a unique opportunity to explore the landscape of transitions from UC to MC forms of organization. Even more after considering the potential synergies between artificial evolution and synthetic biology. Today it is possible to provide simple unicellular organisms with unexpressed, unregulated coding sequences as genomic background. These could mobilize simple physical processes

like adhesion or cell polarity, as pointed out by Newman and Bhat.

The key experiment then would be to force the evolution of MC by exposing cells with challenging environments and watch how the rearrangement of genes drives new modes of pattern formation and manipulates individual and collective fitness. Such experiment might enable us to reconstruct a scenario similar to the cambrian explosion, for cells can be introduced the necessary pre-adaptations to achieve MC, and a wealth forms and species might be rapidly selected for. This constitute a relevant first hand account of the evolution of complexity, even if in an artificial setting.

Chapter 5

CONCLUSIONS

Regarding our first objective of developing of new models of spatial patterning in cellular collectives, we have found that:

- A system composed of the energy minimization process of differential adhesion driven outside equilibrium configurations by stochastic phenotypic switching generates spatial regular structures in periodic arrangement.
- The wavelength of these patterns depends on the model parameters, specially the scaling factor between the two mechanisms.
- In a model system with a population of cells following a chemical gradient that they themselves synthesize, periodic structures can be found when the chemical signal induces cell death.
- Higher order properties of such system, like pattern straightness, can be incorporated by the action of saturation in the chemotactic process.
- The interplay between adhesion processes, cell morphology and signalling leading to inhibited growth yields also regular distributions of cell density in an engineered bacterial colony.

Regarding the second objective of creating of clear frameworks to understand the emergence of MC arising from darwinian entities:

- We have created different simulations platform incorporating relevant physicochemical elements and evolution, linking both internal and exogenous selective pressures to the evolution of multicellularity.

Regarding our final interest of providing general scenarios in which differentiated multicellularity with division of labor should be expected, we have found that:

- The presence of toxic products, both of endogenous and exogenous origin, can force the emergence of cellular collectives with differentiation from darwinian populations.
- In the case of the exogenous waste, we have found that this multicellular strain is linked by group-dependent alterations to individual fitness.
- In the case of the exogenous waste scenario, cells manipulate their spatial relatedness with increased strain cohesion.

Chapter 6

ANNEX

Carbonell-Ballestero M, Duran-Nebreda S, Montañez R, Solé R, Macía J, Rodríguez-Caso C. A bottom-up characterization of transfer functions for synthetic biology designs: lessons from enzymology. Nucleic Acids Res. 2014 Dec 16;42(22):14060-9. DOI: 10.1093/nar/gku964

Solé R, Duran-Nebreda S, Insilico transitions to multicellularity. Dins de: Ruiz-Trillo AM, Nedelcu AM (eds.) Evolutionary transitions to multicelular life. (Advances in marine genomics ; 2) 2015 Mar 28 : 245-266. DOI: 10.1007/978-94-017-9642-2_13

Bibliography

[Akam, 1989] Akam, M. (1989). Hox and hom: homologous gene clusters in insects and vertebrates. *Cell*, 57(3):347–349.

[Alberch, 1980] Alberch, P. (1980). Ontogenesis and morphological diversification. *American Zoologist*, 20(4):653–667.

[Alonso et al., 2002] Alonso, D., Bartumeus, F., and Catalan, J. (2002). Mutual interference between predators can give rise to turing spatial patterns. *Ecology*, 83(1):28–34.

[Babloyantz, 1977] Babloyantz, A. (1977). Self-organization phenomena resulting from cell-cell contact. *Journal of theoretical biology*, 68(4):551–561.

[Bak, 2013] Bak, P. (2013). *How nature works: the science of self-organized criticality*. Springer Science & Business Media.

[Ball, 2011] Ball, P. (2011). *Shapes: nature's patterns: a tapestry in three parts*, volume 1. Oxford University Press.

[Bascompte and Solé, 1996] Bascompte, J. and Solé, R. V. (1996). Habitat fragmentation and extinction thresholds in spatially explicit models. *Journal of Animal Ecology*, pages 465–473.

[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):1130–1134.

[Baurmann et al., 2007] Baurmann, M., Gross, T., and Feudel, U. (2007). Instabilities in spatially extended predator–prey systems: spatio-temporal patterns in the neighborhood of turing-hopf bifurcations. *Journal of Theoretical Biology*, 245(2):220–229.

[Ben-Jacob et al., 2000] Ben-Jacob, E., Cohen, I., Golding, I., Gutnick, D. L., Tcherpakov, M., Helbing, D., and Ron, I. G. (2000). Bacterial cooperative organization under antibiotic stress. *Physica A: Statistical Mechanics and its Applications*, 282(1):247–282.

BIBLIOGRAPHY

[Ben-Jacob et al., 1994] Ben-Jacob, E., Schochet, O., Tenenbaum, A., Cohen, I., Czirok, A., and Vicsek, T. (1994). Generic modelling of cooperative growth patterns in bacterial colonies. *Nature*, 368(6466):46–49.

[Benner and Sismour, 2005] Benner, S. A. and Sismour, A. M. (2005). Synthetic biology. *Nature Reviews Genetics*, 6(7):533–543.

[Beysens et al., 2000] Beysens, D., Forgacs, G., and Glazier, J. (2000). Cell sorting is analogous to phase ordering in fluids. *Proceedings of the National Academy of Sciences*, 97(17):9467–9471.

[Black et al., 1993] Black, T. A., Cai, Y., and Wolk, C. P. (1993). Spatial expression and autoregulation of hetr, a gene involved in the control of heterocyst development in anabaena. *Molecular microbiology*, 9(1):77–84.

[Bonner, 1952] Bonner, J. T. (1952). *Morphogenesis: an essay on development*. Princeton University Press Princeton NJ.

[Bonner, 1965] Bonner, J. T. (1965). Size and cycle: an essay on the structure of biology. *American Scientist*, pages 488–494.

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

[Bonner, 1998] Bonner, J. T. (1998). The origins of multicellularity. *Integrative Biology Issues News and Reviews*, 1(1):27–36.

[Bonner, 2009a] Bonner, J. T. (2009a). *First signals: the evolution of multicellular development*. Princeton University Press.

[Bonner, 2009b] Bonner, J. T. (2009b). The social amoebae. *Princeton UniversityPress, Princeton*.

[Bonner, 2013] Bonner, J. T. (2013). *Randomness in evolution*. 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.

[Boyle et al., 2013] Boyle, K. E., Heilmann, S., van Ditmarsch, D., and Xavier, J. B. (2013). Exploiting social evolution in biofilms. *Current opinion in microbiology*, 16(2):207–212.

[Brown and Jones, 1974] Brown, J. E. and Jones, K. (1974). Basic concepts of differentiation and growth of cells. In *Differentiation and growth of cells in vertebrate tissues*, pages 1–51. Springer.

BIBLIOGRAPHY

[Bullara and Decker, 2015] Bullara, D. and Decker, Y. D. (2015). Pigment cell movement is not required for generation of Turing patterns in zebrafish skin. *Nature Communications*, 6(May):1–7.

[Buss, 2014] Buss, L. W. (2014). *The evolution of individuality*. Princeton University Press.

[Byrne and Chaplain, 1997] Byrne, H. and Chaplain, M. A. (1997). Free boundary value problems associated with the growth and development of multicellular spheroids. *European Journal of Applied Mathematics*, 8(06):639–658.

[Cairns, 2006] Cairns, J. (2006). Mutation selection and the natural history of cancer. *Science of Aging Knowledge Environment*, 2006(10):cp1.

[Carnac and Gurdon, 1997] Carnac, G. and Gurdon, J. B. (1997). The community effect in xenopus myogenesis is promoted by dorsalizing factors. *International Journal of Developmental Biology*, 41:521–524.

[Carroll, 2001] Carroll, S. B. (2001). Chance and necessity: the evolution of morphological complexity and diversity. *Nature*, 409(6823):1102–1109.

[Carroll, 2005] Carroll, S. B. (2005). Evolution at two levels: on genes and form. *PLoS biology*, 3(7):1159.

[Cavalier-Smith, 1978] Cavalier-Smith, T. (1978). Nuclear volume control by nucleoskeletal dna, selection for cell volume and cell growth rate, and the solution of the dna c-value paradox. *Journal of Cell Science*, 34(1):247–278.

[Chuang, 2012] Chuang, J. S. (2012). Engineering multicellular traits in synthetic microbial populations. *Current opinion in chemical biology*, 16(3):370–378.

[Coleman, 2012] Coleman, A. (2012). A compartive analysis of the volvocaceae (chlorophyta) 1. *Journal of phycology*, 48(3):491–513.

[Conway, 1970] Conway, J. (1970). The game of life. *Scientific American*, 223(4):4.

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

[Danino et al., 2010] Danino, T., Mondragón-Palomino, O., Tsimring, L., and Hasty, J. (2010). A synchronized quorum of genetic clocks. *Nature*, 463(7279):326–330.

[Darmon and Leach, 2014] Darmon, E. and Leach, D. R. (2014). Bacterial genome instability. *Microbiology and Molecular Biology Reviews*, 78(1):1–39.

[Darwin, 1859] Darwin, C. (1859). On the origins of species by means of natural selection. *London: Murray*, page 247.

BIBLIOGRAPHY

[Davies, 2008] Davies, J. A. (2008). Synthetic morphology: prospects for engineered, self-constructing anatomies. *Journal of anatomy*, 212(6):707–719.

[Dawkins, 1986] Dawkins, R. (1986). *The blind watchmaker: Why the evidence of evolution reveals a universe without design*. WW Norton & Company.

[de Mendoza et al., 2013] de Mendoza, A., Sebé-Pedrós, A., Šestak, M. S., Matejcic, M., Torruella, G., Domazet-Loso, T., and Ruiz-Trillo, I. n. (2013). Transcription factor evolution in eukaryotes and the assembly of the regulatory toolkit in multicellular lineages. *Proceedings of the National Academy of Sciences of the United States of America*, 110(50):E4858–66.

[de Vargas Roditi et al., 2013] de Vargas Roditi, L., Boyle, K. E., and Xavier, J. B. (2013). Multilevel selection analysis of a microbial social trait. *Molecular systems biology*, 9(1):684.

[Deutsch and Dormann, 2002] Deutsch, A. and Dormann, S. (2002). Cellular automaton modeling of biological pattern formation: Characterization, applications, and analysis.

[Diambra et al., 2014] Diambra, L., Senthivel, V. R., Menendez, D. B., and Isalan, M. (2014). Cooperativity to increase turing pattern space for synthetic biology. *ACS synthetic biology*, 4(2):177–186.

[Dieckmann et al., 1995] Dieckmann, U., Marrow, P., and Law, R. (1995). Evolutionary cycling in predator-prey interactions: population dynamics and the red queen. *Journal of Theoretical Biology*, 176(1):91–102.

[Dworkin and Bonner, 1972] Dworkin, M. and Bonner, J. (1972). The myxobacteria: new directions in studies of procaryotic development. *CRC critical reviews in microbiology*, 1(4):435–452.

[Economou et al., 2012] Economou, A. D., Ohazama, A., Porntaveetus, T., Sharpe, P. T., Kondo, S., Basson, M. A., Gritli-Linde, A., Cobourne, M. T., and Green, J. B. (2012). Periodic stripe formation by a turing mechanism operating at growth zones in the mammalian palate. *Nature genetics*, 44(3):348–351.

[Eggenberger, 1997] Eggenberger, P. (1997). Evolving morphologies of simulated 3d organisms based on differential gene expression. In *Proceedings of the fourth european conference on Artificial Life*, pages 205–213.

[Eichenberger et al., 2004] Eichenberger, P., Fujita, M., Jensen, S. T., Conlon, E. M., Rudner, D. Z., Wang, S. T., Ferguson, C., Haga, K., Sato, T., Liu, J. S., et al. (2004). The program of gene transcription for a single differentiating cell type during sporulation in bacillus subtilis. *PLoS Biology*.

BIBLIOGRAPHY

[Eldredge, 1972] Eldredge, N. (1972). Punctuated equilibria: an alternative to phyletic gradualism. *Models in Paleobiology*, pages 82–115.

[Eldredge, 1985] Eldredge, N. (1985). *Unfinished synthesis: biological hierarchies and modern evolutionary thought*. Oxford University Press New York.

[Elena and Lenski, 2003] Elena, S. F. and Lenski, R. E. (2003). Evolution experiments with microorganisms: the dynamics and genetic bases of adaptation. *Nature Reviews Genetics*, 4(6):457–469.

[Elowitz and Leibler, 2000] Elowitz, M. B. and Leibler, S. (2000). A synthetic oscillatory network of transcriptional regulators. *Nature*, 403(6767):335–338.

[Endy, 2005] Endy, D. (2005). Foundations for engineering biology. *Nature*, 438(7067):449–453.

[Erwin, 2000] Erwin, D. H. (2000). Macroevolution is more than repeated rounds of microevolution. *Evolution & Development*, 2(2):78–84.

[Floreano and Mattiussi, 2008] Floreano, D. and Mattiussi, C. (2008). *Bio-inspired artificial intelligence: theories, methods, and technologies*. MIT press.

[Forgacs and Newman, 2005] Forgacs, G. and Newman, S. A. (2005). *Biological physics of the developing embryo*. Cambridge University Press.

[Forrest, 1993] Forrest, S. (1993). Genetic algorithms: principles of natural selection applied to computation. *Science*, 261(5123):872–878.

[Foty et al., 1996] Foty, R. A., Pfleger, C. M., Forgacs, G., and Steinberg, M. S. (1996). Surface tensions of embryonic tissues predict their mutual envelopment behavior. *Development*, 122(5):1611–1620.

[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.

[Frohnhöfer et al., 2013] Frohnhöfer, H. G., Krauss, J., Maischein, H.-m., and Nüsslein-volhard, C. (2013). Iridophores and their interactions with other chromatophores are required for stripe formation in zebrafish. *Development*, 3007:2997–3007.

[Fu and Huang, 2014] Fu, X. and Huang, W. (2014). Quantitative measurement and analysis in a synthetic pattern formation multicellular system. In *Engineering and Analyzing Multicellular Systems*, pages 137–149. Springer.

[Fuqua et al., 1994] Fuqua, W. C., Winans, S. C., and Greenberg, E. P. (1994). Quorum sensing in bacteria: the luxr-luxi family of cell density-responsive transcriptional regulators. *Journal of bacteriology*, 176(2):269.

BIBLIOGRAPHY

[Furusawa and Kaneko, 1998] Furusawa, C. and Kaneko, K. (1998). Emergence of multicellular organisms with dynamic differentiation and spatial pattern. *Artificial Life*, 4(1):79–93.

[Furusawa and Kaneko, 2003] Furusawa, C. and Kaneko, K. (2003). Robust development as a consequence of generated positional information. *Journal of theoretical biology*, 224(4):413–435.

[Gerhart et al., 1997] Gerhart, J., Kirschner, M., and Moderbacher, E. S. (1997). *Cells, embryos, and evolution: Toward a cellular and developmental understanding of phenotypic variation and evolutionary adaptability*. Blackwell Science Malden.

[Gibbs and Martin, 1962] Gibbs, J. P. and Martin, W. T. (1962). Urbanization, technology, and the division of labor: International patterns. *American Sociological Review*, pages 667–677.

[Gierer and Meinhardt, 1972] Gierer, A. and Meinhardt, H. (1972). A theory of biological pattern formation. *Kybernetik*, 12(1):30–39.

[Gilbert, 2000] Gilbert, S. F. (2000). *Developmental Biology. Sunderland, MA*. Sinauer Associates, Inc.

[Gillispie, 2006] Gillispie, C. C. (2006). The formation of lamarck's evolutionary theory. *Transactions of the American Philosophical Society*, 96(5):47.

[Glick, 1995] Glick, B. R. (1995). Metabolic load and heterologous gene expression. *Biotechnology advances*, 13(2):247–261.

[Gould, 2011] Gould, S. J. (2011). *Full house*. Harvard University Press.

[Gould and Eldredge, 2000] Gould, S. J. and Eldredge, N. (2000). Punctuated equilibrium comes of age. *Shaking the Tree: Readings from Nature in the History of Life*, 17.

[Graner and Glazier, 1992] Graner, F. and Glazier, J. A. (1992). Simulation of biological cell sorting using a two-dimensional extended potts model. *Physical review letters*, 69(13):2013.

[Gregory, 2008] Gregory, T. R. (2008). The evolution of complex organs. *Evolution: Education and Outreach*, 1(4):358–389.

[Grindrod et al., 1989] Grindrod, P., Murray, J., and Sinha, S. (1989). Steady-state spatial patterns in a cell-chemotaxis model. *Mathematical Medicine and Biology*, 6(2):69–79.

[Grosberg and Strathmann, 2007] Grosberg, R. K. and Strathmann, R. R. (2007). The evolution of multicellularity: a minor major transition? *Annual Review of Ecology, Evolution, and Systematics*, pages 621–654.

BIBLIOGRAPHY

[Haeckel and Lankester, 1883] Haeckel, E. and Lankester, S. E. R. (1883). *The history of creation*. Kegan Paul London.

[Hayward, 1987] Hayward, T. L. (1987). The nutrient distribution and primary production in the central north pacific. *Deep Sea Research Part A. Oceanographic Research Papers*, 34(9):1593–1627.

[Heinemann and Panke, 2006] Heinemann, M. and Panke, S. (2006). Synthetic biology—putting engineering into biology. *Bioinformatics*, 22(22):2790–2799.

[Helm and Potts, 2012] Helm, R. F. and Potts, M. (2012). Extracellular matrix (ecm). In *Ecology of Cyanobacteria II*, pages 461–480. Springer.

[Henderson et al., 1999] Henderson, I. R., Owen, P., and Nataro, J. P. (1999). Molecular switches-the on and off of bacterial phase variation. *Molecular microbiology*, 33(5):919–932.

[Hentschel et al., 2004] Hentschel, H., Glimm, T., Glazier, J. A., and Newman, S. A. (2004). Dynamical mechanisms for skeletal pattern formation in the vertebrate limb. *Proceedings of the Royal Society of London B*, 271:1713–1722.

[Herron et al., 2009] Herron, M. D., Hackett, J. D., Aylward, F. O., and Michod, R. E. (2009). Triassic origin and early radiation of multicellular volvocine algae. *Proceedings of the National Academy of Sciences*, 106(9):3254–3258.

[Herron and Michod, 2008] Herron, M. D. and Michod, R. E. (2008). Evolution of complexity in the volvocine algae: transitions in individuality through darwin's eye. *Evolution*, 62(2):436–451.

[Herskowitz and Oshima, 1981] Herskowitz, I. and Oshima, Y. (1981). Control of cell type in saccharomyces cerevisiae: mating type and mating-type interconversion. *Cold Spring Harbor Monograph Archive*, 11:181–209.

[Hibbing et al., 2010] Hibbing, M. E., Fuqua, C., Parsek, M. R., and Peterson, S. B. (2010). Bacterial competition: surviving and thriving in the microbial jungle. *Nature Reviews Microbiology*, 8(1):15–25.

[HilleRisLambers et al., 2001] HilleRisLambers, R., Rietkerk, M., van den Bosch, F., Prins, H. H., and de Kroon, H. (2001). Vegetation pattern formation in semi-arid grazing systems. *Ecology*, 82(1):50–61.

[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.

BIBLIOGRAPHY

[Horváth et al., 2009] Horváth, J., Szalai, I., and De Kepper, P. (2009). An experimental design method leading to chemical turing patterns. *Science*, 324(5928):772–775.

[Hull, 1973] Hull, D. (1973). Darwin and his critics. the reception of darwin's theory of evolution by the scientific community. *Cambridge: Mass., Harvard University Press xii, 473p.. Geog*, 1:1809–1882.

[Huxley et al., 1942] Huxley, J. et al. (1942). Evolution. the modern synthesis. *Evolution. The Modern Synthesis.*

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

[Jean, 2009] Jean, R. V. (2009). *Phyllotaxis: a systemic study in plant morphogenesis*. Cambridge University Press.

[Jernvall, 2002] Jernvall, J. (2002). A gene network model accounting for development and evolution of mammalian teeth. *Proceedings of the National Academy of Sciences*, 99(12):8116–8120.

[Jones et al., 1994] Jones, C. G., Lawton, J. H., and Shachak, M. (1994). Organisms as ecosystem engineers. *Oikos*, 69:373–386.

[Jones et al., 1996] Jones, C. G., Lawton, J. H., and Shachak, M. (1996). Organisms as ecosystem engineers. In *Ecosystem management*, pages 130–147. Springer.

[Kaandorp et al., 2008] Kaandorp, J. A., Blom, J. G., Verhoef, J., Filatov, M., Postma, M., and Müller, W. E. (2008). Modelling genetic regulation of growth and form in a branching sponge. *Proceedings of the Royal Society of London B: Biological Sciences*, 275(1651):2569–2575.

[Kaneko and Yomo, 1999] Kaneko, K. and Yomo, T. (1999). Isologous diversification for robust development of cell society. *Journal of theoretical biology*, 199(3):243–256.

[Kapral and Showalter, 2012] Kapral, R. and Showalter, K. (2012). *Chemical waves and patterns*, volume 10. Springer Science & Business Media.

[Kauffman, 1969] Kauffman, S. A. (1969). Metabolic stability and epigenesis in randomly constructed genetic nets. *Journal of theoretical biology*, 22(3):437–467.

[Kauffman, 1993] Kauffman, S. A. (1993). *The origins of order: Self organization and selection in evolution*. Oxford university press.

[Keller and Segel, 1970] Keller, E. F. and Segel, L. A. (1970). Initiation of slime mold aggregation viewed as an instability. *Journal of Theoretical Biology*, 26(3):399–415.

BIBLIOGRAPHY

[King et al., 2008] King, N., Westbrook, M. J., Young, S. L., Kuo, A., Abedin, M., Chapman, J., Fairclough, S., Hellsten, U., Isogai, Y., Letunic, I., et al. (2008). The genome of the choanoflagellate monosiga brevicollis and the origin of metazoans. *Nature*, 451(7180):783–788.

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

[Kirk, 2005b] Kirk, D. L. (2005b). *Volvox: a search for the molecular and genetic origins of multicellularity and cellular differentiation*, volume 33. Cambridge University Press.

[Kondo and Asai, 1995] Kondo, S. and Asai, R. (1995). A reaction-diffusion wave on the skin of the marine angelfish pomacanthus. *Nature*, 376(6543):765–768.

[Kondo and Miura, 2010] Kondo, S. and Miura, T. (2010). Reaction-diffusion model as a framework for understanding biological pattern formation. *science*, 329(5999):1616–1620.

[Kong et al., 2014] Kong, W., Celik, V., Liao, C., Hua, Q., and Lu, T. (2014). Programming the group behaviors of bacterial communities with synthetic cellular communication. *Bioresources and Bioprocessing*, 1(1):1–11.

[Koschwanez et al., 2013] Koschwanez, J. H., Foster, K. R., and Murray, A. W. (2013). Improved use of a public good selects for the evolution of undifferentiated multicellularity. *Elife*, 2:e00367.

[Krumlauf, 1994] Krumlauf, R. (1994). Hox genes in vertebrate development. *Cell*, 78(2):191–201.

[Lane, 2010] Lane, N. (2010). *Life ascending: the ten great inventions of evolution*. Profile books.

[Lavrentovich et al., 2013] Lavrentovich, M. O., Koschwanez, J. H., and Nelson, D. R. (2013). Nutrient shielding in clusters of cells. *Physical Review E*, 87(6):062703.

[Lawrence et al., 1992] Lawrence, P. A. et al. (1992). *The making of a fly: the genetics of animal design.* Blackwell Scientific Publications Ltd.

[Lehn, 2002] Lehn, J.-M. (2002). Toward self-organization and complex matter. *Science*, 295(5564):2400–2403.

[Lejeune and Tlidi, 1999] Lejeune, O. and Tlidi, M. (1999). A model for the explanation of vegetation stripes. *Journal of Vegetation science*, 10(2):201–208.

[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.

BIBLIOGRAPHY

[Levskaya et al., 2005] Levskaya, A., Chevalier, A. A., Tabor, J. J., Simpson, Z. B., Lavery, L. A., Levy, M., Davidson, E. A., Scouras, A., Ellington, A. D., Marcotte, E. M., et al. (2005). Synthetic biology: engineering escherichia coli to see light. *Nature*, 438(7067):441–442.

[Lewis, 2007] Lewis, K. (2007). Persister cells, dormancy and infectious disease. *Nature Reviews Microbiology*, 5(1):48–56.

[Libby et al., 2014] Libby, E., Ratcliff, W., Travisano, M., and Kerr, B. (2014). Geometry shapes evolution of early multicellularity. *PLoS Computational Biology*.

[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.

[Lois and Alvarez-Buylla, 1994] Lois, C. and Alvarez-Buylla, A. (1994). Long-distance neuronal migration in the adult mammalian brain. *Science*, 264(5162):1145–1148.

[Lovejoy, 2011] Lovejoy, A. O. (2011). *The great chain of being: A study of the history of an idea*. Transaction Publishers.

[Magnus, 1867] Magnus, A. (1867). De vegetabilibus. *Libri VII.(Ed. by EHF Meyer and C. Jessen). G. Berolini, Reimeri*, 752.

[Mahalwar et al., 2014] Mahalwar, P., Walderich, B., and Singh, A. P. (2014). Local reorganization of xanthophores fine-tunes and colors the striped pattern of zebrafish. *Science*, 345(6202):1362–1364.

[Maharbiz, 2012] Maharbiz, M. M. (2012). Synthetic multicellularity. *Trends in cell biology*, 22(12):617–623.

[Maini, 1997] Maini, P. K. (1997). Bones, feathers, teeth and coat marking: a unified model. *Science progress*, 80(3):217–229.

[Mangan and Alon, 2003] Mangan, S. and Alon, U. (2003). Structure and function of the feed-forward loop network motif. *Proceedings of the National Academy of Sciences*, 100(21):11980–11985.

[Marin, 1976] Marin, F. T. (1976). Regulation of development in dictyostelium discoideum: I. initiation of the growth to development transition by amino acid starvation. *Developmental biology*, 48(1):110–117.

[Mayer et al., 1993] Mayer, B., Johnson, J. P., Leitl, F., Jauch, K. W., Heiss, M. M., Schildberg, F. W., Birchmeier, W., and Funke, I. (1993). E-cadherin expression in primary and metastatic gastric cancer: down-regulation correlates with cellular dedifferentiation and glandular disintegration. *Cancer research*, 53(7):1690–1695.

BIBLIOGRAPHY

[Maynard-Smith, 1998] Maynard-Smith, J. (1998). *Shaping life: genes, embryos and evolution*. Yale University Press.

[Maynard-Smith and Szathmary, 1997] Maynard-Smith, J. and Szathmary, E. (1997). *The major transitions in evolution*. Oxford University Press.

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

[McShea, 2000] McShea, D. W. (2000). Functional complexity in organisms: parts as proxies. *Biology and Philosophy*, 15(5):641–668.

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

[McShea and Brandon, 2010] McShea, D. W. and Brandon, R. N. (2010). *Biology's first law: the tendency for diversity and complexity to increase in evolutionary systems*. University of Chicago Press.

[Medina et al., 2003] Medina, M., Collins, A. G., Taylor, J. W., Valentine, J. W., Lipps, J. H., Amaral-Zettler, L., and Sogin, M. L. (2003). Phylogeny of opisthokonta and the evolution of multicellularity and complexity in fungi and metazoa. *International Journal of Astrobiology*, 2(03):203–211.

[Meeks and Elhai, 2002] Meeks, J. C. and Elhai, J. (2002). Regulation of cellular differentiation in filamentous cyanobacteria in free-living and plant-associated symbiotic growth states. *Microbiology and Molecular Biology Reviews*, 66(1):94–121.

[Meinhardt and Gierer, 1980] Meinhardt, H. and Gierer, A. (1980). Generation and regeneration of sequence of structures during morphogenesis. *Journal of Theoretical Biology*, 85(3):429–450.

[Meinhardt and Gierer, 2000] Meinhardt, H. and Gierer, A. (2000). Pattern formation by local self-activation and lateral inhibition. *Bioessays*, 22(8):753–760.

[Merks et al., 2006] Merks, R. M., Brodsky, S. V., Goligorksy, M. S., Newman, S. A., and Glazier, J. A. (2006). Cell elongation is key to in silico replication of in vitro vasculogenesis and subsequent remodeling. *Developmental biology*, 289(1):44–54.

[Michod, 2000] Michod, R. E. (2000). *Darwinian dynamics: evolutionary transitions in fitness and individuality*. Princeton University Press.

[Michod, 2005] Michod, R. E. (2005). On the transfer of fitness from the cell to the multicellular organism. *Biology and Philosophy*, 20(5):967–987.

[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.

BIBLIOGRAPHY

[Michod and Roze, 2001] Michod, R. E. and Roze, D. (2001). Cooperation and conflict in the evolution of multicellularity. *Heredity*, 86(1):1–7.

[Miller and Bassler, 2001] Miller, M. B. and Bassler, B. L. (2001). Quorum sensing in bacteria. *Annual Reviews in Microbiology*, 55(1):165–199.

[Miramontes et al., 2001] Miramontes, O., Sole, R. V., and Goodwin, B. (2001). Neural networks as sources of chaotic motor activity in ants and how complexity develops at the social scale. *International Journal of Bifurcation and Chaos*, 11(06):1655–1664.

[Mitchell, 1998] Mitchell, M. (1998). *An introduction to genetic algorithms*. MIT press.

[Miyazawa, 2012] Miyazawa, S. (2012). Striping by striding: synthetic approaches toward pattern formation. *Pigment cell & melanoma research*, 25(2):124–125.

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

[Morris, 2006] Morris, S. C. (2006). Darwin's dilemma: the realities of the cambrian 'explosion'. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 361(1470):1069–1083.

[Murray, 1988] Murray, J. D. (1988). How the leopard gets its spots. *Scientific American*, 258(3):80–87.

[Myers and Saupe, 2013] Myers, C. E. and Saupe, E. E. (2013). A macroevolutionary expansion of the modern synthesis and the importance of extrinsic abiotic factors. *Palaeontology*, 56(6):1179–1198.

[Nakano et al., 2012] Nakano, T., Ando, S., Takata, N., Kawada, M., Muguruma, K., Sekiguchi, K., Saito, K., Yonemura, S., Eiraku, M., and Sasai, Y. (2012). Self-formation of optic cups and storable stratified neural retina from human escs. *Cell stem cell*, 10(6):771–785.

[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 et al., 2006] Newman, S. a., Forgacs, G., and Muller, G. B. (2006). Before programs: the physical origination of multicellular forms. *The International journal of developmental biology*, 50(2-3):289–99.

[Nicolis et al., 1977] Nicolis, G., Prigogine, I., et al. (1977). *Self-organization in nonequilibrium systems*, volume 191977. Wiley, New York.

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

BIBLIOGRAPHY

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

[Nilsson and Pelger, 1994] Nilsson, D.-E. and Pelger, S. (1994). A pessimistic estimate of the time required for an eye to evolve. *Proceedings of the Royal Society of London B: Biological Sciences*, 256(1345):53–58.

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

[Okasha, 2006] Okasha, S. (2006). *Evolution and the Levels of Selection*, volume 16. Clarendon Press Oxford.

[Oster and Murray, 1989] Oster, G. F. and Murray, J. D. (1989). Pattern formation models and developmental constraints. *Journal of Experimental Zoology*, 251(2):186–202.

[Ouyang and Swinney, 1991] Ouyang, Q. and Swinney, H. L. (1991). Transition from a uniform state to hexagonal and striped turing patterns. *Nature*, 352(6336):610–612.

[Passamaneck and Halanych, 2004] Passamaneck, Y. J. and Halanych, K. M. (2004). Evidence from hox genes that bryozoans are lophotrochozoans. *Evolution & development*, 6(4):275–281.

[Pattus et al., 1990] Pattus, F., Massotte, D., Wilmsen, H., Lakey, J., Tsernoglou, D., Tucker, A., and Parker, M. (1990). Colicins: prokaryotic killer-pores. *Experientia*, 46(2):180–192.

[Payne et al., 2009] Payne, J. L., Boyer, A. G., Brown, J. H., Finnegan, S., Kowalewski, M., Krause, R. A., Lyons, S. K., McClain, C. R., McShea, D. W., Novack-Gottshall, P. M., et al. (2009). Two-phase increase in the maximum size of life over 3.5 billion years reflects biological innovation and environmental opportunity. *Proceedings of the National Academy of Sciences*, 106(1):24–27.

[Pentz et al., 2015] Pentz, J. T., Limberg, T., Beermann, N., and Ratcliff, W. C. (2015). Predator escape: An ecologically realistic scenario for the evolutionary origins of multicellularity. *Evolution: Education and Outreach*, 8(1):1–8.

[Perona and Malik, 1990] Perona, P. and Malik, J. (1990). Scale-space and edge detection using anisotropic diffusion. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 12(7):629–639.

[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.

BIBLIOGRAPHY

[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.

[Purnick and Weiss, 2009] Purnick, P. E. and Weiss, R. (2009). The second wave of synthetic biology: from modules to systems. *Nature reviews Molecular cell biology*, 10(6):410–422.

[Queller, 2000] Queller, D. C. (2000). Relatedness and the fraternal major transitions. *Philosophical Transactions of the Royal Society 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.

[Raspopovic et al., 2014] Raspopovic, J., Marcon, L., Russo, L., and Sharpe, J. (2014). Digit patterning is controlled by a bmp-sox9-wnt turing network modulated by morphogen gradients. *Science*, 345(6196):566–570.

[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.

[Regot et al., 2011] Regot, S., Macia, J., Conde, N., Furukawa, K., Kjellén, J., Peeters, T., Hohmann, S., de Nadal, E., Posas, F., and Solé, R. (2011). Distributed biological computation with multicellular engineered networks. *Nature*, 469(7329):207–211.

[Rokas et al., 2005] Rokas, A., Krüger, D., and Carroll, S. B. (2005). Animal evolution and the molecular signature of radiations compressed in time. *Science*, 310(5756):1933–1938.

[Salles-Passador et al., 1992] Salles-Passador, I., Moisand, A., and Wright, M. (1992). Cell cycle-dependent assembly and disassembly of cytoplasmic microtubules in the plasmodium of the myxomycete physarum polycephalum. *European journal of cell biology*, 58(1):136–148.

[Schirrmeister et al., 2013] Schirrmeister, B. E., de Vos, J. M., Antonelli, A., and Bagheri, H. C. (2013). Evolution of multicellularity coincided with increased diversification of cyanobacteria and the great oxidation event. *Proceedings of the National Academy of Sciences*, 110(5):1791–1796.

BIBLIOGRAPHY

[Scott et al., 1983] Scott, M. P., Weiner, A. J., Hazelrigg, T. I., Polisky, B. A., Pirrotta, V., Scalenghe, F., and Kaufman, T. C. (1983). The molecular organization of the antennapedia locus of drosophila. *Cell*, 35(3):763–776.

[Sebé-Pedrós et al., 2011] Sebé-Pedrós, A., de Mendoza, A., Lang, B. F., Degnan, B. M., and Ruiz-Trillo, I. (2011). Unexpected repertoire of metazoan transcription factors in the unicellular holozoan capsaspora owczarzaki. *Molecular biology and evolution*, 28(3):1241–1254.

[Sebé-Pedrós et al., 2010] Sebé-Pedrós, A., Roger, A. J., Lang, F. B., King, N., and Ruiz-Trillo, I. (2010). Ancient origin of the integrin-mediated adhesion and signaling machinery. *Proceedings of the National Academy of Sciences*, 107(22):10142–10147.

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

[Sheth et al., 2012] Sheth, R., Marcon, L., Bastida, M. F., Junco, M., Quintana, L., Dahn, R., Kmita, M., Sharpe, J., and Ros, M. a. (2012). Hox genes regulate digit patterning by controlling the wavelength of a Turing-type mechanism. *Science*, 338(6113):1476–80.

[Sick et al., 2006] Sick, S., Reinker, S., Timmer, J., and Schlake, T. (2006). Wnt and dkk determine hair follicle spacing through a reaction-diffusion mechanism. *Science*, 314(5804):1447–1450.

[Smith et al., 1990] Smith, R. E., Harrison, W. G., Harris, L. R., and Herman, A. W. (1990). Vertical fine structure of particulate matter and nutrients in sea ice of the high arctic. *Canadian Journal of Fisheries and Aquatic Sciences*, 47(7):1348–1355.

[Snow and Snow, 1947] Snow, M. and Snow, R. (1947). On the determination of leaves. *New Phytologist*, pages 5–19.

[Solari et al., 2006] Solari, C. A., Ganguly, S., Kessler, J. O., Michod, R. E., and Goldstein, R. E. (2006). Multicellularity and the functional interdependence of motility and molecular transport. *Proceedings of the National Academy of Sciences of the United States of America*, 103(5):1353–1358.

[Solé, 2012] Solé, R. (2012). *Phase Transitions*. Princeton University Press.

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

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

BIBLIOGRAPHY

[Spemann and Mangold, 1924] Spemann, H. and Mangold, H. (1924). Über induktion von embryonalanlagen durch implantation artfremder organisatoren. *Development Genes and Evolution*, 100(3):599–638.

[Sprent et al., 1988] Sprent, J., Lo, D., Gao, E.-K., and Ron, Y. (1988). T cell selection in the thymus. *Immunological reviews*, 101(1):173–190.

[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.

[Stebbins, 1969] Stebbins, G. L. (1969). *The basis of progressive evolution*. UNC Press Books.

[Steinberg and Poole, 1982] Steinberg, M. and Poole, T. (1982). Liquid behavior of embryonic tissues. *Cell behavior*, pages 583–607.

[Steinberg, 1970] Steinberg, M. S. (1970). Does differential adhesion govern self-assembly processes in histogenesis? equilibrium configurations and the emergence of a hierarchy among populations of embryonic cells. *Journal of Experimental Zoology*, 173(4):395–433.

[Steinberg, 1975] Steinberg, M. S. (1975). Adhesion-guided multicellular assembly: a commentary upon the postulates, real and imagined, of the differential adhesion hypothesis, with special attention to computer simulations of cell sorting. *Journal of theoretical biology*, 55(2):431–443.

[Szathmáry, 2015] Szathmáry, E. (2015). Toward major evolutionary transitions theory 2.0. *Proceedings of the National Academy of Sciences*, page 201421398.

[Takeichi, 1991] Takeichi, M. (1991). Cadherin cell adhesion receptors as a morphogenetic regulator. *Science*, 251(5000):1451–1455.

[Thar and Kühl, 2005] Thar, R. and Kühl, M. (2005). Complex pattern formation of marine gradient bacteria explained by a simple computer model. *FEMS Microbiology Letters*, 246:75–79.

[Thompson et al., 1942] Thompson, D. W. et al. (1942). On growth and form. *On growth and form.*

[Tweedt and Erwin, 2015] Tweedt, S. M. and Erwin, D. H. (2015). Origin of metazoan developmental toolkits and their expression in the fossil record. In *Evolutionary Transitions to Multicellular Life*, pages 47–77. Springer.

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

BIBLIOGRAPHY

[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.

[Vinson, 2011] Vinson, V. K. (2011). How escherichia coli got its stripes. *Science Signaling*, 4(195):ec290–ec290.

[Wang and Church, 2011] Wang, H. H. and Church, G. M. (2011). Multiplexed genome engineering and genotyping methods applications for synthetic biology and metabolic engineering. *Methods Enzymol*, 498:409–426.

[Wang et al., 2009] Wang, H. H., Isaacs, F. J., Carr, P. A., Sun, Z. Z., Xu, G., Forest, C. R., and Church, G. M. (2009). Programming cells by multiplex genome engineering and accelerated evolution. *Nature*, 460(7257):894–898.

[Ward and King, 1997] Ward, J. P. and King, J. (1997). Mathematical modelling of avascular-tumour growth. *Mathematical Medicine and Biology*, 14(1):39–69.

[Waters and Bassler, 2005] Waters, C. M. and Bassler, B. L. (2005). Quorum sensing: cell-to-cell communication in bacteria. *Annu. Rev. Cell Dev. Biol.*, 21:319–346.

[West et al., 2002] West, S. A., Pen, I., and Griffin, A. S. (2002). Cooperation and competition between relatives. *Science*, 296(5565):72–75.

[Wiener and Rosenblueth, 1946] Wiener, N. and Rosenblueth, A. (1946). The mathematical formulation of the problem of conduction of impulses in a network of connected excitable elements, specifically in cardiac muscle. *Archivos del instituto de Cardiología de México*, 16(3):205–265.

[Willensdorfer, 2009] Willensdorfer, M. (2009). On the evolution of differentiated multicellularity. *Evolution*, 63(2):306–323.

[Wilson, 1997] Wilson, D. S. (1997). Altruism and organism: Disentangling the themes of multilevel selection theory. *The American Naturalist*, 150(S1):s122–S134.

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

[Wolpert et al., 2015] Wolpert, L., Tickle, C., and Arias, A. M. (2015). *Principles of development*. Oxford university press.

[Woodward et al., 1995] Woodward, D., Tyson, R., Myerscough, M., Murray, J., Budrene, E., and Berg, H. (1995). Spatio-temporal patterns generated by salmonella typhimurium. *Biophysical journal*, 68(5):2181.

[Xavier, 2011] Xavier, J. B. (2011). Social interaction in synthetic and natural microbial communities. *Molecular systems biology*, 7(1):483.