Eco-evolutionary dynamics of collective behaviour

Silvia de Monte

Eco-evolutionary dynamics of collective behaviour

SILVIA DE MONTE

CNRS, Institut de Biologie de l'ENS, Paris, France
Max Planck Institute for Evolutionary Biology, Plön, Germany

Memoire de soutenance
HABILITATION A DIRIGER DES RECHERCHES

Ecole Normale Supérieure
Université Paris Sciences et Lettres, Paris

Defended on June 29, 2022 in front of a jury composed by :

Pr.	Sara Mitri	Reviewer
Pr.	Allyson Sgro	Reviewer
Pr.	Joshua Weitz	Reviewer
Pr.	Vincent Hakim	President of the jury
Pr.	Massimo Vergassola	Examiner

Contents

To my mitochondrial lineage, and its carriers.

Acknowledgments

Cleary, I waited far too long before I convinced myself of the need to submit this memoir. Therefore, I have hosts of persons to thank: colleagues, friends, and all the others who stood by and for me even when I was difficult to stand.

I would like to start from the former and current PhD students and post-docs I had the privilege to co-advise, as their ideas, investment and energies have been essential for me to shape my own view on the ecology and evolution of collective behaviour. Thomas, Guilhem, Alice, Alice, Orso, Philippe, Shruti, Enrico, Matthieu, Jules, Emil (in chronological order) thank you! Among the undergraduates I interacted with, I am particularly grateful to Hanna and Leonardo for all the discussions on and beyond work.

Writing this memoir – and postponing its writing up to this point – would not have been possible without the support of the Eco-Evolutionary Mathematics team I belong to since I landed in Paris at the end of my PhD. I have learnt a lot in particular from Régis, Sandrine and Eric, not only on science. I greatly appreciate to be surrounded by nice people at the 4th floor of IBENS (as well as in the whole building), and I am particularly grateful to Auguste for regularly reminding me, when I met him in the corridor, that I had an HDR to pass.

My new scientific family at the Max Planck Institute of Evolutionary Biology in Plön is equally important to me, and joining them has been a honour and an invaluable opportunity for further intellectual progress. I thank particularly Arne, Hildegard and Chaitanya for being so patient with my comings and goings, even before the pandemics disrupted them all.

I owe a lot to all the colleagues I have shared my research experiences with. One aspect that I treasure of interdisciplinary research is its shared dimension, which allows every problem to be considered from different points of view. This has let me interact with scholars whose disciplinary and interdisciplinary skills I admire enormously, and who introduced me to concepts and methods that

were essential for my own development. Just to mention the names of those I have exchanged with most, I thank Hugues, Leonardo and Giulio for keeping me connected with physics methods despite my progressive drift towards biology; Lucie, Chris, Olivier for paying attention to my naive views of biology. Clément, Amaury and Mathieu who, a bit like I do, navigate among disciplines, have been the source of uncountable discussions on science and on life, and I appreciate their friendship enormously.

I am equally grateful to all the amazing scientists I have had the occasion to exchange with in France and abroad. Some of them are cited in the memoir for their outstanding influence on my way of thinking, but I hold many more as contributors to the ideas I present in this manuscript. The visiting professors that ENS kindly supported for their month-long stays have been and are still my point of reference for learning how to think.

The role that Francesco and Paul have had for my development as a scientist and as a person can't be overestimated. I am deeply grateful, both for what they gave me and for what they accepted from me.

My dearest friends Barbara, Sara, Lucia and Michela have endured the ups and downs of my self-esteem, and their availability to listen to my rants and enthusiasms has been essential to keep the balance. And I remember Silvia and Claudia, who were close being different, and left too early.

In all the lengthy process of finding my place in an academic world I often found unforgiving, it has been essential to have a marvellous, understanding, eclectic composed – and recomposed – family. Its extension spans very large in my heart, and follows its own rhythms of aggregation and disaggregation. I felt at home every time and anywhere I could sit down and laugh or cry with you - mostly laugh, cook, eat and drink. I miei bambini Irene and Nicolò fill me with joy and pride every day, and remind me the importance of being beyond oneself.

1

Introduction

When I was in high school, my biology teacher Ubaldo Busolin accompanied a few students to attend a seminar on behavioural studies. The speaker showed us pictures of famous politicians at the time, Ronald Reagan and Margaret Thatcher, beside images of primates displaying comparable facial expressions of joy, rage, perplexity. That experience opened my eys on the power of adopting a different point of view on apparently complex behaviours, and opened my eyes on how different systems may follow general rules. It also introduced me to the experiments of Konrad Lorenz and Niko Tinbergen, and to the field of ethology.

This was the starting point of a tortuous path in search of understanding of the way biological units work together. This path has wondered across several areas of education: from biology, to physics, to mathematics, and back to biology. The result is the integration of disparate experiences, questioning and approaches, and fascination for subjects I only regret not knowing deeper.

My ideas on what subject would attract most of my attention was set at the end of my first year as a Biology student at the University of Parma. The professor of Zoology, Davide Csermely, proposed us as an extra reading a paper of Simon Levin on the evolution of cooperation. I realised not only that the subject was very exciting to me, but also that I would not be able to understand the math with the tools available to a biology student.

The same year I had met two fantastic persons, Marilde and Alfredo Porati, who have been mentors and friends as long as they were alive. Their courses of physics were passionating, and their views on life as well. They suggested I could study physics first, learn the mathematical language, then go back to biology. And this is what eventually happened.

My view of collective behaviour thus started off from the physics perspective, as a macroscopic manifestation of some microscopic rules. During my PhD, the macroscopic observable I focused on was synchronised dynamics, notably that of oscillating cells. The challenge was to identify general rules that linked, in a population of many interacting units, the dynamical state of one isolated oscillator to the average state. In particular, I was interested in understanding to what extent statistical physics methods devised to connect microscopic and macroscopic dynamics – that typically assume all interacting units are identical – could be extended to heterogenous populations. In this research, Hugues Chaté has been a master in showing me how one could pass from simple dynamical systems to many coupled degrees of freedom, and then find the way back to low dimensionality. A legacy of this work is my ongoing interest for connecting population-scale measures to individual-level features, and the focus on spatial and time scales involved in observing one and the other.

Just after obtaining my master, I was not sure if and where I would find a place to develop my interests. I had no contact with real biologists, and I felt I was not fit for the academic Italian system, that was strongly organised into disciplinary compartments. The advisors of my master project, Antonio Giorgilli and Sergio Rinaldi, were themselves in loose relation with the physics society, being primarily affiliated with the departments of Mathematics and with the

Engineering School respectively. It is thanks to them and to a series of coincidences that I ended up getting a fellowship to study one more year abroad. I chose to go and join the group of Karl Sigmund, whom I met at a winter school organised by Régis Ferrière in the same building I currently work in. I had been captured by Karl's courses on game theory, in particular on modelling the evolution of cooperation. Together with the charm exerted by Mittle-European literature and culture, working with Karl and with Christoph Hauert – then a post-doc – made of my stay in Vienna a most intense and stimulating period.

The kind of cooperation I found most fascinating is that occurring in groups. How can big achievements, that require the participation of many - possibly diverse - individuals resist to the temptation to free-ride? Can the objective be reached without there be a leader showing the way? The problem is more relevant now then ever, with humanity facing massive global challenges, from climate change to pandemics. It is also important to understand the moral underpinning of political choices - the 'crowds' of Elias Canetti. We considered one aspect of the problem, that can be addressed independently of the choice to cooperate or not: the decision to partake to social interactions. The idea to consider optional participation owes a lot to discussions I had with Francesco d'Ovidio on the way internet protocols avoid the overexploitation of band width. Mathematical models like those we developed may not be that relevant for humans, whose choices and interactions go well beyond our simplistic representation. However, Karl's suggestion that they might be applied to explaining the behaviour of microbes is always in my mind.

Microbial populations were also the biological system that I wanted to describe with the models for synchronisation. In yeast cells, the biochemical mechanisms generating single-cell oscillations were known, and their communication through diffusible metabolites was under study in the group of Preben Sørensen at the University of Copenhagen. The collaboration with Preben and Sune Danø extended beyond my PhD and led me to participate to experiments for the first time. It was a sobering experience, that opened my eyes on the difficulty of connecting abstract mathematical models to reality, but also on the

usefulness of semi-quantitative approaches in doing so.

Even though both the concepts of collective behaviour and of cooperation apply to most microbial populations, there are many flavours of them, and their connection is not always straightforward. While the physics community seems to be mostly interested in the way macroscopic modes emerge from local rules, most evolutionary approaches focus on quantities that were hardly related to microscopic mechanisms (they are more often, nowadays). Fitness and payoff are two central concepts in evolutionary biology and game theory. If they are clearly defined in models, their meaning in terms of levels of description and dynamical processes is all but trivial in natural populations. Uncovering these relationships in biological systems has been the subject of most of my research – including the apparently unrelated side excursions – and this document tries to bring it all together.

An overarching principle that guides my investigations is that qualitative and/or quantitative test of model predictions for a specific biological population should not be hopeless. Following Karl Popper, I think theoretical models should be inter-subjectively falsifiable. Therefore my preferred approaches are those that are based on clear hypotheses on the nature of the units that undergo a collective behaviour, and provide predictions on their deployment on multiple time scales. Models can easily bridge spatial and temporal scales, at least through numerical simulation. However, there are not many biological systems that are amenable to experimentation on multiple scales. Microbial populations are particularly attractive for this reason: one can hope to pin down individual properties, and at the same time observe how cells self-organize on fast time scales, how they are affected by demography on ecological tume scales, how they are shaped by evolution on longer time scales.

Collaborations with evolutionary microbiologists, first and foremost Paul Rainey and Vidya Nanjundiah (thanks to Clement Nizak), motivated me to dare thinking of experiments that could be used to disentangle levels of organisation and time scales. Getting close to the biology of the populations I aimed at modelling opened a Pandora's box of further complexity and revealed

a wealth of additional processes that operate at different scales. At the same time, the advent of single-cell imaging and metagenomic techniques provided ever hardening challenges to models, making it even unclear what expectations are reasonable given the current state of our knowledge: Is competitive exclusion a relevant null model for heterogenous cell assemblies? Is reproductive success predictive of long-term evolutionary dynamics? Do social strategies have the same meaning on ecological and evolutionary time scales? What does one need to count to capture relevant diversity in microbial assemblies?

I will touch upon these themes in this dissertation, and decline them in particular for three classes of heterogenous biological populations that are associated to some forms of collective behaviour: bacteria, 'social' amoebae and microbial communities. Every chapter will provide a brief introduction to a line of questioning and a few key references to put it into context. Moreover, it will provide a list of the publications associated to that theme and the names of the young scientists who have essentially contributed to the research.

Chapter 2 'Game theory and the economics of social behaviour' presents my attempts to concile a very elegant theoretical framework, that of evolutionary game theory, with more down-to-earth details on how groups form in the first place. It starts – from the end, as often – with a reflection on what economics brings into evolutionary biology and what in my opinion it should not.

Chapter 3 'Levels of organisation and evolutionary transitions' introduces a few key ingredients that will be used in the following, notably the levels of description of a biological population and their associated time scales. It also presents a model where these are linked together by evolution of an individual trait in populations that start off with a binary motility phenotype.

Chapter 4 'Phenotypic heterogeneity in aggregative life cycles' tackles implications of phenotypic heterogeneity in microbial populations that form multicellular aggregates. I will first introduce *Dictyostelium discoideum* and the questions it poses to evolutionary biology. Then, I will discuss some ideas on how to assess the effects of different sources of phenotypic variation, and some recent experimental results obtained by Mathieu Forget on frequency-dependent spore

formation efficiency in binary mixes. I will present models designed to test the implications of such experimental observations, and finally touch on their potential evolutionary implications.

Chapter 5 'Community abundance distributions' considers even looser collections of cells, that encompass a much broader genetic diversity: microbial communities. Environmental variations are key elements in determining the dynamics of these communities on the 'fast', ecological time scales. I will discuss how interactions with fluid dynamics shapes diversity distributions of planktonic communities in the global ocean, and how remote sensing can inform us about this process. Then, I will take another perspective and look at a fine-scale characterisation of these communities by genomic methods. Knowledge of the abundances of thousands of species per sample allowed us to address the spatial variability of ecological patterns, in particular species abundance distributions, that underpin common measures of diversity – a key determinant of ecosystem function.

Chapter 6 'Evolution of heterogenous collectives' reports on how communities are increasingly conceived as 'superorganisms' after recognition of their collective functions. Understanding the evolutionary dynamics of communities therefore poses similar problems than that of aggregative multicellular microbes, due to the tension between individual species' competition and community-level functionality. I will discuss the limitations and implications of considering communities as individual entities subjected to Darwinian evolution, and possible ways to improve the action of selection acting on a collective property. I will use a simple two-types community to illustrate the effects of selection for a collective function on cell-level properties. Finally, I will discuss generalisations to complex, multi-species communities.

The last chapter 'Conclusions and perspectives' is dedicated to an outlook on future directions of investigation, and how these are motivated by the results presented previously. A more detailed research project is attached as a separate document.

At the end of every chapter, I listed the names of students I have advised,

the related presentations and publications. A complete list, as well as details on service and funding, is available on the curriculum vitae.

Before starting to discuss the scientific ideas that animated my work, I would like to briefly mention my implication in teaching, mostly at Ecole Normale Supérieure.

I have started in 2009 to support Régis Ferrière in the organisation and teaching of a L3 course on mathematical models for biological sciences, that was mostly a collection of invited seminars by well renown scientists. A couple of years later I became the main organiser of the course and gradually moved to a more traditional structure composed mostly of a series of concatenated lectures and practical exercises, while I kept a few external research seminars. With the help of Mathieu Coppey, who was teaching stochastic modelling, and of a number of other colleagues and students who gave punctual interventions, the course had in my opinion reached a satisfying level of stability and coherence when I handed it over to François Blanquart in 2019. Teaching this interdisciplinary course has been very instructive for learning both how to teach (I had never lectured before) and the functioning of a school that is so different from the institutions I had previously experienced. The interactions with students have been at times exciting and at times frustrating. It revealed the complexity of setting up an interdisciplinary teaching that allowed me to introduce some mathematical concepts at the same time as their domains of applications, and allowed students to try themselves the theoretical tools on subjects of their interest. I keep being convinced that an early introduction to modelling – less to the formal techniques than to the art of phrasing a question in mathematical terms – is fundamental in the education of every scientist. Having suffered myself for the lack of such teaching in a Biology curriculum, I am very happy to witness that now quantitative approaches are increasingly prioritised in training.

Let aside lectures that I have punctually given in other M1 courses at ENS (some of which I temporally organised or co-organised) and in various PhD schools in France and abroad (notably, at KITP, Santa Barbara, USA and ICTP,

Sao Paulo, Brazil), my other major investment in teaching has been the one-week-long module 'Collective Behaviour' for the master IMaLiS of ENS. After a pilot experiment in 2017, this course has evolved towards its current structure, that consists of a combination of morning lectures and, in the afternoon, seminars or lab exercises (coordinated by Sandrine Adiba). The course covers different approaches to collective behaviour, spanning topics from the physics of collective emergent dynamics to evolutionary biology. While the morning lectures are centered on theory and modelling, the afternoons are dedicated to the application of those concepts to controlled experiments, in particular on microbial populations, and lately social insects.

Finally, I wish to acknowledge here that a huge source of inspiration for structuring my teaching have been the interactions with researchers involved in the 'Ecole de formation continue' of CNRS that has been held in the course of a decade on the island of Berder, Morbihan. This school, that was dedicated every year to a different theme at the interface of biology and hard sciences, provided me with my first contacts with French research back in 2003, and involved colleagues who have then been instrumental to my moving to France for a post-doc, in particular Hugues Chaté and François Taddei. I participated to three editions, organised (or co-organised) three and lectured in three. Not only the meeting, that was structured so as to leave a lot of space to discussion, made me start thinking to questions that are still central in my research. It was also the occasion to meet some exceptionally inspiring scientists whose research I greatly admire for its creativity, rigour and depth. In closing this chapter, I remember in particular Yves Couder and Alain Arneodo, recently disappeared at the distance of a few days, who showed me how concepts of an elegant simplicity can teach a lot on the functioning of complex biological systems.

The competition between human beings destroys with cold and diabolic brutality. Under the pressure of this competitive fury we have not only forgotten what is useful to humanity as a whole, but even that which is good and advantageous to the individual.

Konrad Lorenz
Civilized Man's Eight Deadly Sins

2

Game theory and the economics of collective behaviour

A DEFINING FEATURE OF COLLECTIVE BEHAVIOUR is that it results from interactions of individual units that compose an ensemble, such as a population or a group, identifiable from some aggregate property. The composing units, be them players in a group, cells in a body, organisms in a population or a community, belong to the same equivalence class, so that they can be considered similar with respect to some criterion. Equivalence does not mean that they will all be identical, as they could differ with respect to other criteria. Such diversity is the hallmark of biological systems with respect to physical populations. Indeed, if one can think of elementary particles of one type to be all the same, this is almost never true for biological ensembles: communities are composed of different species that belong to a common taxonomi-

cal level and share the same ecological context, animals in groups are different from one another even though they belong to the same species and maybe even the same kindred, cells in a multicellular assemblage (a multicellular organism as well as pluricellular structures like biofilms) can be different both in their genomes and in their intracellular state, and the same is true even within isogenic populations like those that are typically studied in the lab. Without such variation there would be no evolution in the system, and perhaps no beauty either.

Understanding collective behaviour can be challenging in itself even in the absence of variation among individual units. For instance, interaction of oscillating units can give rise to temporal and spatial patterns that have been extensively studied in the framework of synchronisation theory[128,109]. Soft and active matter physics have produced methods to characterise the collective motion of (mostly) identical particles, and these have been applied to animal behaviour[129]. The questioning here is mostly how macroscopic observables, detectable at the level of the population, depend on individual features, for instance the parameters ruling the intrinsic individual dynamics, and on the strength and topology of interactions. I will talk of this type of questioning only in discussing the mechanistic bases of reproductive biases in aggregative multicellular microbes (Chapter 5), but it has largely motivated my research during the PhD[20,22,27,25] and after[26,24,54,23].

On the other hand, evolutionary biology focuses on how different classes of individuals change in frequency over time. In the simplest case, consider a population that is composed of two distinct types of individuals, whose behaviour is transmitted from parent to offspring. When these individuals interact within the population and with their environment, their behaviour determines the probability that they will reproduce. Population demography, driven by the processes of birth and death of individual of each type (typically dependent on some environmental variables as well as on the population composition) will generically drive the system towards an attractor. Evolutionary game theory studies the nature of such attractor as a function of the way traits characteris-

tic of the two types determine their birth and death rates (a.k.a. their fitness). Individual traits, structure and intensity of interactions are all summarised in a single number, the *payoff*, that measures reproductive success of any given type, or *strategy*. In the simple case where individuals interact in pairs in a well-mixed environment, as well in more complicated situations, the temporal dynamics of the frequency of each strategy is described by the so-called *replicator equation*[71]. This equation is mathematically equivalent, for pairwise games, to the Lotka-Volterra equation with competitive interactions, and therefore shares its elegance but also some of its limitations as a model for real systems. One interesting feature of the replicator equation is that it provides a dynamic view on whether equilibrium solutions of the game defined by the matrix including the payoffs (whose entry i, j is the payoff of the interaction of an individual of strategy i with one with strategy j) can be reached when frequency changes are fuelled by natural selection.

This formalism thus naturally connects evolutionary biology to economics, that had developed a wealth of tools to identify optimal equilibrium solutions of games. Conceptual bridges between economics and evolutionary biology were in fact present from the onset, when Charles Darwin found a mechanism for the 'struggle for existence' he observed in the natural world in 'An Essay on the Principle of Population' (1798) by the demographer and economist rev. Thomas Robert Malthus. Despite the often fundamental differences in the support and transmission of behavioural strategies in humans and other organisms, the way economics conceived interactions among individuals kept influencing evolutionary biology – as well as ecology – to the core, to the point that game-theoretical models have widely established as null models for the evolution of both human and non-human behaviour. Suffice to read two highly influential papers written by the ecologist Garrett Hardin. In the first, published in 1960, he proposed the name of 'competitive exclusion principle' for the previously established result that – all else being equal – one (competitively superior) species will generally drive another (competitively inferior) to extinction[63]. Eight years later,

he referred even more explicitly to economics in concluding that by pursuing their own interests, individuals end up undermining any collective entreprise, in what he termed the 'tragedy of the commons'[64]. Interestingly, he also conceived that ecology, and ultimately evolution, could be a source of inspiration for economics: "Natural selection commensurates the incommensurables. ... Man must imitate this process".

The cross-talk between economics and evolutionary biology has been constant, and influenced both disciplines through the widespread use of analogies. The idea that competition – like the struggle for existence – leads to optimal solutions has become a central tenet of neoclassical economics. On the other hand, interactions in the natural world have been widely interpreted in an anthropomorphic perspective, to the point of attributing to cognitively limited organisms, such as microbes, feelings and complex capabilities. More than ten years ago, the philosopher of biology Jean Gayon came up with the idea of exploring the relationship between economics and evolutionary biology. He involved in this project a group including the philosophers of biology Philippe Hunemann and Johannes Martens, the economists Bernard Walliser and Mikael Cozic, the biologist Jean-Baptiste André, and myself. In order to facilitate the dialogue across disciplines, I proposed to identify some terms and concepts that are central to both disciplines, and to compare their meanings and perceived implications. This led to a number of meetings that extended over several years, where single terms would be discussed for the two separate disciplines first, and then in comparison. Under the encouragement of Jean and the subsequent lead by Philippe, Bernard abd Johannes, the synthesis of these discussions has now taken the form of a book that is in press within the Spinger series 'History, Philosophy and Theory of the Life Sciences' under the title 'From evolutionary biology to economics and back: some conceptual transfers'.

One widespread feature of all organisms is that they are seldom found alone. Many animals live in groups, and even unicellular organisms spend most of their lives in close proximity of cells of their same kind (as in bacterial colonies), or

of other strains or species. The boundaries of such groups can be imposed from some external constraint that sets the border between what belongs to the interaction network and who is excluded, such as for the gut microbiome or the stock exchange. In other cases, groups result from the self-organisation of individuals into more or less ephemeral structures, such as swarms, hunting associations, biofilms. Rather than being just one-to-one exchanges, interactions within such groups generally involve a number of individuals, therefore their description requires a generalisation of game-theoretical models based on pairwise payoffs. This has been achieved by extending evolutionary game theory to so-called n-player games[52], that allow to take into account the collective nature of certain fitness-enhancing behaviours. Typically, the size of the group of interacting individuals (or its distribution) are set, so that it is possible to study, for instance, how the evolutionary attractors change when group size is changed.

If the constancy of group size is a reasonable assumption in modelling, for instance, experimental economics studies, where the number of interacting individuals is prescribed, or highly structured societies such as lion prides, in many other cases not only individuals can partake of differently sized groups, but they can – actively or passively – join or leave them. I have long been interested in how group size dynamics influences and is coupled to group-size functionality. In a first study that I conducted before starting my PhD, I wondered what would happen if individuals were given the choice, on top of whether to cooperate or defect in a social setting, also whether to join the group at all. We named individuals whose strategy was to (always) avoid social gatherings loners[66,65]. This term has subsequently been used with slightly different meanings to designate individuals that either have a higher chance to be found alone than others of the same population, or simply individuals that happen to be found outside groups[39,133,116]. As long as loners can persist on their own resources, and if group-related benefits are high enough, then voluntary participation to a game allows the three strategies (loner, cooperate and defect) to coexist in the population in an oscillatory fashion, both in well-mixed and in spatially structured populations. The key element allowing cooperation to surge from its

ashes is that variations in the frequency of loners induces variations in the size of groups of individuals that are effectively interacting. The tragedy of the commons, inevitable within groups, results in players leaving the game and therefore brings the system in a situation where being a cooperator rather than a defector makes a difference to one own's group-related payoff. The applications of this toy model to any real system are problematic in a number of ways. Nevertheless, realisation of the possibly complex interplay between processes involved in group formation and the resulting evolutionary dynamics has been a guiding thread in the models for the eco-evolutionary dynamics of social groups I have subsequently developed.

Group formation in the volunteering model was very simple: a set number of players were drawn at random, and the game was only played by the subset of social players (cooperators and defectors). This way, group composition was random and just reflected the frequencies in the population, bar sampling effects. When we think of groups of organisms, it is difficult to conceive that the individual propensity to join a group is totally disconnected from the role that that individual will have in the game, as these two processes are likely to be influenced by the same class of signals or processes. For instance, if one type of cells is more adhesive than another, it will more likely belong to a cell cluster, and this cell cluster will be more cohesive than a cluster composed of the less adhesive type.

With Thomas Garcia and Guilhem Doulcier, we explored the evolutionary implications of this assumption on two time scales: that of the competition between two variants with fixed differences in adhesion[47], and that of changes in the level of adhesion by successive bouts of pair-wise competition[48]. These models allowed us to explore the potential for the evolutionary success of traits, such as the production of adhesive molecules at the cell surface, that contribute to collective function (cell-cell adhesion is not only essential to keep together multicellular aggregates, but is also involved in development and division of labour) at a cost for the individual cell. As in classic models for social evolu-

tion (e.g. Wilson's trait group model[143]), we considered a scenario where groups form over and again due to some external forcing, and the aggregated phase – where reproduction occurs based on collective-level benefits and individual-level costs – is followed by dispersal in a well-mixed pool (Fig. 2.1).

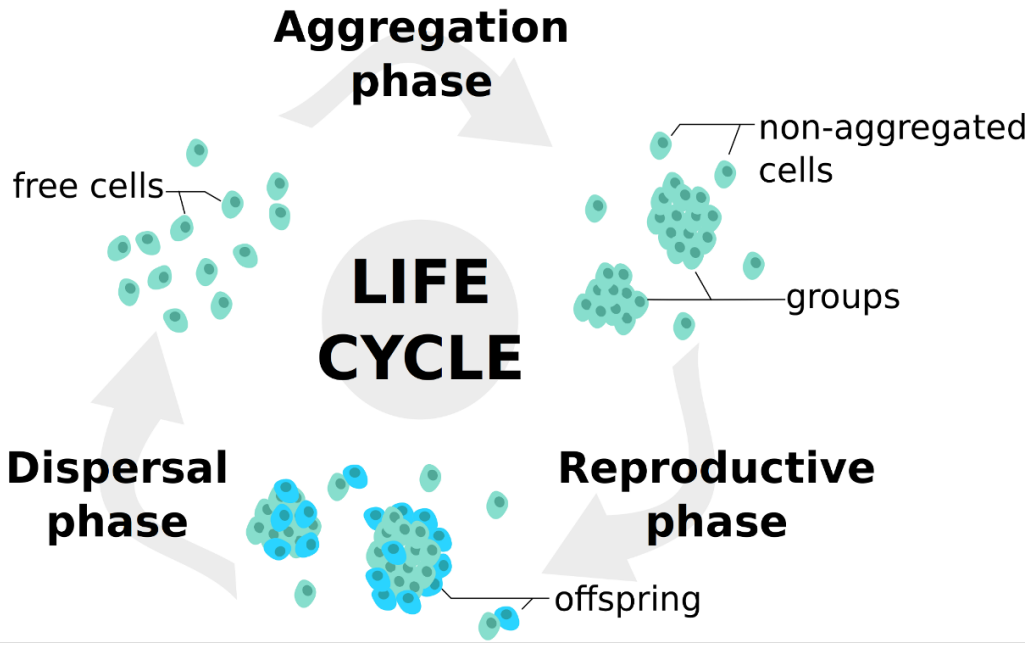

Figure 2.1: In aggregative life cycles, selection acts on cell behaviour in the multicellular stage, that is the consequence both of the costs of individual investment and the benefits induced by collective contributions. Possible additional differences among groups at this stage are superseded by dispersal in a common pool as single-cell propagules, that are free to group again later independently of the group they previously belonged to. The essential features of such life cycle are captured by one-shot multiplayer games where players assort in groups repeatedly. Graphics by Guilhem Doulcier.

Unlike the trait group and the volunteering model, group formation here does not occur by random sampling in the population, but is biased by the same trait that concurs in providing collective functions: adhesion. Individuals are still chosen at random to have the chance to join a group, but the probability they will eventually do so depends on adhesiveness of both the focal cell and of the group (represented by a randomly chosen 'recruiter'). Once groups are formed, payoffs are attributed according to a public goods game (a standard for

experimental economics), where individual payoff, or fitness, increases linearly in the average contribution and decrease linearly in the individual cost[121]. Here, costs and investments derive exclusively from the level of adhesiveness, so that more sticky types provide a more valuable contribution to the collective, but pay a higher cost independent of the fact that they end up or not in a group.

In Garcia & De Monte we showed that even if adhesion-induced interactions are not assortative (the expected proportion of stickier individuals among those a focal individual attaches to does not depend on its type), more adhesive types will increase in frequency in the population. More generally, this mechanism works also if assortment is not neutral, even when it is negative (as long as it is not too much so). This is partially (but not exclusively) motivated by the fact that more adhesive types will be found in smaller proportion in the fraction of individuals that remain alone, and therefore will not access the possible collective benefits. The way group formation was depicted in this work was highly simplified in order to allow to solve the problem analytically. I will talk about more realistic models, that consider individual motility underpinning encounters between individuals, in Chapter 4. One noticeable conclusion of this study is that in order for the more adhesive type to invade the population, it must be initially present in a sufficient proportion: only a few individuals would not be able to assort sufficiently to offset the cost of increased adhesiveness.

The aggregation process depicted by this model was highly simplified so as to only retain the key features of group self-organisation. Some more realism can be introduced if we focus on more specific aggregation processes, such as those occurring in microbial populations that crawl on surfaces. In Garcia et al., we considered how differential adhesion introduces biases the aggregation of self-propelled particles, thus letting the frequency of particles of one type inside and outside the aggregates to deviate from the frequency of that same type in the binary mix. We could thus explore not only the conditions for the success of stickier particles (that confirmed the previous more abstract results), but also the patterns associated to different evolutionary stable strategies. One feature of this model is that the ecological time scale of aggregation (after which groups

are dissolved and particles dispersed) mediates between fast, transient ecology and long-term evolutionary changes. The importance of such parameter will be discussed in several of the following chapters, and in particular in Chapter 4, where I will present a slightly more realistic version of the self-propelled particle model, which accounts for quantifiable features of cell motility.

I will end this chapter discussing an extension of the previous models that describes adhesiveness as a continuous, and not a binary trait as previously[48]. This model allows to ask the question of what evolutionary trajectories can lead to the establishment of a sufficient level of adhesion, so that the benefits of collective behaviour overcome the cost to the individual. Clearly, if producing some kind of glue is costly, but its effects for mediating adhesion are scarce, then selection will not favour adhesion purely on the ground that it sustains collective function. The framework of adaptive dynamics allows to determine how the continuous trait 'adhesiveness' changes under the successive appearance of mutant individuals whose trait is only slightly different from the resident type. Given the previously mentioned result that a quorum of similarly adhesive types must be present for adhesion to become beneficial, one could expect that single mutant individuals could never drive an increase in the trait. However, costs are also proportional to the investment in glue production, so two infinitesimal changes need to be compared. Our calculations revealed that a small increase in adhesion in a handful of individuals can ratchet up adhesiveness and lead to more on more cohesive collectives. This is however possible only if there is a certain amount of adhesiveness to start with. In applications to the real world, and notably to the evolutionary emergence of multicellular organisation, one can think that some kind of glue was originally produced for reasons that were unrelated to collective function, for instance to attach to surfaces or to be able to exert traction on a substrate. In the moment cells were adhesive enough for this to provide also a collective advantage, then it would have become both the cause and the consequence of the evolution of more and more cohesive aggregates.

What these models do not explain is how did the cycles of aggregation and dispersal initiated. The next chapter is dedicated to the formalisation and modelling of the establishment of such life cycles.

Supervision:

Thomas Garcia (PhD 2014): co-supervised with Edith Perrier and Leonardo Gregory-Brunnet

Selected invited presentations:

Plenary speaker at the XXXIV Dynamics Days Europe 2014, Bayreuth, Germany (September 2014)
Invited speaker at the Aquavit symposium, Max Planck Institute of Evolutionary Biology, Plön, Germany (May 2015)
Invited speaker at the symposium in honour of Karl Sigmund, Vienna, Austria (October 2016)
Invited speaker at the conference Emergence and Evolution of Biological Complexity, Bangalore, India, (February 2017)
Keynote speaker at Learning, Evolution and Games, Lucca, Italy (June 2022)

Related publications:

Thomas Garcia and Silvia De Monte (2013)
Group formation and the evolution of sociality
Evolution, 67(1), 131–141

Garcia T., Gregory-Brunnet L. and De Monte S. (2014)
Differential adhesion between moving particles as a mechanism for the evolution of social groups
PLoS Comp. Biol. 10(2): e1003482

Thomas Garcia*, Guilhem Doulcier* and Silvia De Monte (2015)
The evolution of adhesiveness as a social adaptation
eLife, 4, e08595

André Jean-Baptiste*, Cozic Mikael*, De Monte Silvia*, Gayon Jean*, Huneman Philippe*, Martens Johannes*, Walliser Bernard* (in press)
From evolutionary biology to economics and back: some conceptual transfers
Ed. Spinger
Chapters: 'Crisis', 'Cycles', 'Diversity', 'Equilibrium',
'Heredity/Transmission', 'Time scales'

...time was not passing...it was turning in a circle...

Gabriel García Márquez
One hundred years of solitude

3

Levels of organisation and evolutionary transitions

Life is hierarchically organised, so that societies are made up of organisms whose organs are composed of cells, where organelles and nuclei contain genes and other sequences capable of self-replication. Each of these levels has originated in the course of life evolution based on the progressive complexification of lower levels of organisation. Although the nature of the first replicators is still uncertain, as are the processes that led to their emergence from inanimate matter, it is clear that the process of self-replication is a core ingredient for evolution by natural selection. Once natural selection acts on a population of self-replicating units, it can shape their features towards more adaptive and more complex solutions.

In Chapter 6 I will discuss more extensively some of the evolutionary pro-

cesses involved in transitioning from one level of organisation, composed of units that self-replicate, to another level having the same property. Such 'major evolutionary transitions' have concentrated the attention of biologists, modellers and philosophers in the past decades[98,51,103]. Here, I will focus on how a unit with defined spatiotemporal boundaries comes to be established. In particular, even though some principles could extend to other levels of organisation, I will consider the transition from unicellular to multicellular life style, thus on the first appearance of aggregates of cells with a given collective function. Having occurred numerous times independently along the tree of life, this has been deemed a 'minor' major evolutionary transition[58]. General organising principles rather than serendipity are thus expected to underlie the emergence of multicellular structures[58,137]

One first path to the emergence of multicellular aggregates is that cells remain attached after division, and the collective is composed of a clonal lineage[84,108]. The fact that clonal expansion by cell reproduction is a feasible, and arguably the most important, evolutionary route to multicellularity has been demonstrated by experimental evolution[113,62] and through the reconstruction of the phylogeny of volvocine algae[69] and choanoflagellates[13]. Among many others, clonal growth has the great advantage of reducing the level of conflict experienced within the multicellular body, as all its composing cells share the same genetic material. When all cells are essentially the same, they have a priori the same expectation of reproductive success, even though they contingently manifest different behavioural strategies (e.g. cooperators and defectors). This opens the door to the establishment of division of labour, even to the extent to which one of the cell types dies, like in somatic tissues. Clonal growth – even more when it is combined with single-cell bottlenecks, another potent means of purging deleterious mutations – drastically reduces within-group genetic heterogeneity that can be inherited across generations of the collective, multicellular stage. This is a key ingredient to ensure the evolutionary stability of multicellular function. Even if such teleological explanation is widely accepted as satisfactory in an evolutionary biology perspective, open questions still remain on the

mechanisms that created a diversity of life cycles by variations in the modes of propagule dispersal and the regulation of behaviour in time[124].

Beside life cycles encompassing clonal growth, in many important circumstances pluricellular aggregates are formed by cells that do not share the same genome, and are sometimes not even of the same species or the same clade. Biofilms range from single-strain colonies to complex communities where bacteria coexist with protists. The microbiota comprises a variety of unicelllular organisms that share a same environment (e.g. the gut or the skin) and its boundaries are essentially set by the spatial extension of the host. Formation of functional multicellular aggregates has been observed not only in bacteria (e.g. of the genus *Myxococcus*[101]), but also among close unicellular relatives of Metazoans, such as *Capsaspora owczarzaki*[119]. Moreover, it has been recently shown that mutations can stabilize a stress-dependent multicellular aggregative stage in the unicellular alga *Chlamydomonas reinhardtii*. A surprising organism that I will present in more detail in Chapter 4, *Dictyostelium*, builds its multicellular differentiated fruiting bodies by assembling cells that belong to close or sometimes even distantly related strains. Analogously, other unicellular eukaryotes as well as prokaryotes possess life cycles that contain a stage where disparate cells join to form fruiting bodies. These collective assemblies, that can reach a considerable level of internal structure by cell differentiation, improve the dispersal efficiency of a subset of cells that enter dormancy, the spores. As was pointed out[44,105,102], aggregative life cycles can offer the advantage of providing rapid access to a number of functions that are distributed among cells belonging to different strains or species. Aggregation is likely to limit the complexity of the multicellular body, because it risks of disrupting the synchrony of temporally organised developmental processes. Nonetheless, it allows to promptly exploit the diversity of functions that are available in the populations inhabiting a given environment at a given time. Moreover, it can produce large multicellular collectives on a time scale that is substantially faster than that required to grow them from a handful of cells, and is less limited by the availability of nutrients

that sustain such growth. Since a large size can in many cases bear advantages *per se* to a group of cells, for instance in reducing grazing by predators or protection against stressors, aggregative multicellularity appears in principle to be as viable as clonal multicellularity in providing support for the evolution of collective adaptations, and has been suggested to play a possible role even in the emergence of clonal multicellularity[104].

What makes aggregative life cycles particularly puzzling from the evolutionary point of view is that they make it much harder to avoid exploitation by cell subpopulations – commonly called 'cheaters' – which divert resources necessary for group function towards increasing their own reproduction. Without a way to distinguish between self and non-self, thus basically to enforce assortment by some kind of policing, it is unavoidable that multicellular bodies formed through aggregation will experience strong genetic conflicts over the advantages ensued from collective living. As a consequence of such conflicts, the demise of collective function is expected. For instance, as discussed in Chapter 2, an aggregate might contain cells that produce less of the glue that keeps the body together. The energy that is spared can be invested in higher division rate, which would result in their share of the population to increase over time, and in the consequent extinction of glue producers. In extreme cases, one type might completely forego reproduction in favour of another type. Such kind of suicidal altruist strategy, where one fraction of the aggregating cells dies and favours, by so doing, the remaining cells, occurs in *Dictyostelium* and myxobacteria[101], but also in other microbial species that do not have an explicit aggregative phase[1]. The question is then why a strategy resulting in self-sacrifice does not go extinct, and with it the benefits is produces.

The narrative I have just exposed matches well the formalism of evolutionary game theory that was discussed in the previous chapter: the tragedy of the commons seems unavoidable in aggregative life cycles. However, it simplifies the biological reality in many ways. Game-theoretical models – among others those discussed in Chapter 2 – have integrated more realistic hypotheses on the way strategic decisions are taken, groups are formed or reproductive success is deter-

mined. However, the vast majority of these models is based on two key assumptions: 1) that interactions within multicellular aggregates only occur within a pre-existent temporal framework, whereby cells form groups over and again, and 2) that a cell's reproductive value is exclusively assessed as a consequence of its social behaviour within groups.

The second hypothesis disregards both the process of group formation, and the possible growth that cells may undergo in isolation. The impact of neglecting the solitary phase of an aggregative life cycle and base all evolutionary predictions on the performance in the collective social phase was pointed out for *Dictyostelium*[39,133,116]. Nizak, Tarnita and colleagues showed that, once the multicellular cycle of *Dictyostelium discoideum* is triggered by starvation, 'loner' cells systematically remain outside aggregates. These cells are able to start dividing as soon as nutrients are made available without excessive delay. Through mathematical models, those authors showed that bet-hedging between solitary and group living is the optimal solution in unpredictably varying environments, where committing to the developmental cycle may cause cells to forego unexpected feeding opportunities. In their first models, these authors considered that the probability of aggregating was a genetically determined trait[39,133], but later they refined the model by introducing a mechanistic description of how the decision to aggregate is taken, which accounts among other things for the experimentally observed density-dependence of the fraction of loners[116]. They made the point that trade-offs between strategies – each defined by the probability of cell aggregation – were sufficient to explain coexistence of multiple types in a (meta)population. Indeed, even if different types of cells had the same chances of reproducing in the social phase, the fact that some were found more often in aggregates resulted in a positive bias in spore production. Such advantage of more aggregative cells, that should drive the system towards complete aggregation, is contrasted by the short-term advantages of solitary cells. Key to the maintenance of strategies with different aggregation propensity is the uncertainty induced by stochastic environmental fluctuations. As for the

cycles of aggregation-dispersal in game theoretical models, these cycles are exogenous and imposed on the system. However, the developmental cycle, whose time scale is endogenous, is considered to be fixed. Whereas exogenously varying conditions are unavoidable in any natural system (though perhaps with a lower degree of randomness and independence in different patches than what assumed in simple models), it is not clear what sets the endogenous scale that defines the collective phase of the aggregative life cycle. Presumably, ancestral populations with a unicellular life style did not possess a proto-developmental program that just needed to be triggered to give rise to the social phase of the life cycle. Therefore, whereas the previously discussed models account for the maintenance of diversity in aggregative microbes, they do not explain how such life 'cycles' have emerged in the first place.

Other models addressed more specifically the question of the evolution of a life cycle, that is a succession including a dispersal phase and a multicellular, often differentiated, phase. The adaptive value of life cycles characterised by different modes of group formation was investigated by modelling exogenous periodic variations that impose alternating selective pressures on a trait affecting aggregation[124]. Periodic environmental variations have also been used as a means to change selective pressures in experiments by Paul Rainey and colleagues[8,62]. In these works, however, they were used as convenient proxies of variations that, in more natural conditions, are expected to derive from the ecology of the system itself. The idea that the feedbacks between a microbial population and its environment can be the source of selective pressures that favour alternatively 'cheaters' and 'cooperators' was formalised by Rainey & Kerr. They pointed out that when cells can stochastically switch between alternative phenotypes (by mutation or phenotypic variation), a system where multicellular aggregates grow by clonal expansion and die proportionally to the number of cheaters they contain can entertain cycles that are reminiscent of life cycles. Indeed, the initial growth advantage provided by the function of a collective composed of cooperators is eroded by the emergence of cheater mutants. These

mutants, unable to produce the public good that confers collective advantages – in their case a glue keeping cells attached to each other and to a physical support – can on the other hand act as 'propagules' by setting themselves apart from the collective and initiating a new lineage of cooperative cells.

Environmental feedbacks on reproductive success lead in general to the coupling of the equations describing population ecology and evolution to an external, environmental variable[86]. This coupling opens up new dynamical regimes that were not achievable for pair-wise games, and notably they allow eco-evolutionary cycles[140,134]. In particular, the relevance of eco-evolutionary cycles to systems exhibiting collective behaviour was stressed for blooms of toxic algae[38].

If theoretical approaches coupling the specific ecology of a microbial system to its evolution appear relevant also to conceptually frame the life cycle of *Dicyostelium*, there are a number of questions that needed to be made clear: what are the individual properties that bear fitness consequences in aggregative microbes? When do they make cycles possible? How does the eco-evolutionary dynamics change when they evolve?

In Miele & De Monte, we developed an eco-evolutionary model to explore the onset of aggregative life cycles, starting from pre-existing populations with different phenotypic traits. We chose those traits based on some commonalties between different organisms that have aggregative life cycles, and more specifically on *Dicyostelium*. Our starting point was the observation that the same trait may bear implications for reproductive success both when cells are within an aggregate, as most often considered, and when they are alone. One such trait, that I already discussed, is adhesion, that influences both group formation – thus the probability of remaining outside the groups – and group cohesion. Another such trait is cell motility. Motility allows single-cells to prey, and also fuels the collective displacement of aggregates towards richer food patches. In doing so, it increases fitness both when cells are isolated and when they belong to a multicellular body. At the same time, motility enhances cell dispersal, which opposes assortment by mixing cells with different features. Motility might

therefore be counter-selected if increasing relatedness was the primary means of sustaining collective performance.

We made the hypothesis that motility underpins a trade-off between solitary and collective performance. Cells that move too slowly are neither efficient in feeding (because their prey might escape them), nor at aggregating within a set timeframe. However, when they are within a multicellular aggregate containing also faster cells, they will generally sort to the back of a moving group and be propelled by faster cells. The latter, on the other hand, may spend most of their energy in helping slower cells aggregate, reach better food patches, or, in the case of *Dicyostelium*, forming a stalk that increases their dispersal. The high costs involved in providing such functions to the group can however be compensated when cells are alone, because faster cells can outcompete slower ones in the search for local food.

We formalised this hypothesis in an eco-evolutionary model describing the frequencies of slow and fast cells, their total number, and the resource they feed upon. For simplicity, we assumed that slow cells would only reproduce in the multicelllular stage, and fast cells in the solitary phase, and die otherwise (this clear-cut distinction can be relaxed without altering qualitatively the results). The most important parameters of this system are the growth rate of fast cells (in isolation) λ_F and the level of exploitation of fast by slow cells λ_S, that sets the reproductive success of the latter in the collective phase. We found that this model has two possible qualitatively different dynamics (Figure 3.1). For small exploitation and small single-cell fitness (black region), fast and slow cells coexist in equilibrium (or, for even lower values, fast cells drive slow cells to extinction). However, for higher values of these parameters, slow and fast cells undergo oscillations coupled to those of the resource, whose period depends on the system parameters. These oscillations remind those observed in aggregative life cycles: aggregation occurs when resources grow scarce (therefore, cells – mostly fast ones – are not occupied chasing their food); fast cells are exploited within groups (as per the tragedy of the commons); the prevalence of slow cells causes a decreased resource consumption, paving the way for a successive boom of fast

solitary cells and the consequent overconsumption of resources.

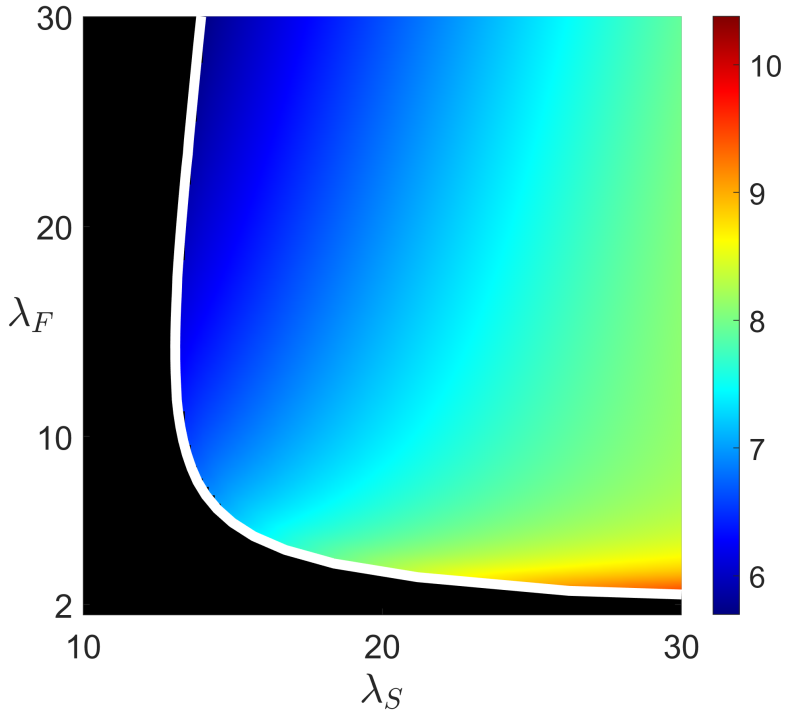

Figure 3.1: Eco-evolutionary cycles of the model introduced in Miele & De Monte have two qualitatively different dynamics. The heatmap represents the period of the 'life-like' cycle as a function of the main system parameters: the solitary growth rate of fast cells and the level of exploitation of fast cells by slow cells within a multicellular aggregate.

The question remains, however, as to where in this phase space we should expect a biological system to be. The advantages of individual motility may manifest already in a purely unicellular life style, so that λ_F could be nonvanishing from the start. There is however no reason to believe that exploitation should be high when collective benefits are still marginal. For this reason, we decided to explore what would be the effect of selection on the exploitation level λ_S.

In order to do this, we had to go beyond the classical game-theoretical description, where the replicator equation describes the dynamics of the frequencies of different, fixed, strategy types (e.g. cooperators and defectors). Even by adding other variables, as in the eco-evolutionary formulation, strategies belong to a finite set that is initially assigned. A solution to this well-known limitation

consists in considering a strategy as defined by a parameter that varies continuously, and that affects interactions between cells and their environment. As long as mutations in the parameter occur sufficiently seldom with respect to the time scale of eco-evolutionary dynamics, then the framework of adaptive dynamics[33] allows to study how the system behaviour changes in the long term. The evolutionary dynamics is here determined by successive invasions of more adapted mutants.

We showed that selection always acts as to increase the level of exploitation within groups, that is what would expect from the 'tragedy of the commons' scenario. Therefore, slow cells gain more and more advantages from the motility of fast cells. However, such increase of λ_S drives the system across the boundary where eco-evolutionary life-like cycles start occurring. The intuition for this observation is that, while slow cells become increasingly exploitative within groups, temporal compartmentalisation of conflicts limits their detrimental effect when averaging on the time scale of the cycle. By a combination of analytical results and numerical simulations, we showed that selection towards harsher exploitation continues also in the oscillating region, causing the amplitude of the life-like cycles, thus the total population size excursions, to increase. This system would thus naturally tend towards small population bottlenecks, where variability within a population could be maintained due to stochastic drift. Another interesting mechanism that might favour the coexistence of multiple types is that evolution slows down when exploitation becomes larger, since increasingly exploitative strategies face diminishing returns. The assumption that the time scales of parameter evolution are longer than those of the eco-evolutionary cycles may hence brake down, opening the door to more complicated coexistence scenarios.

Higher levels of organisation are typically associated to a time scale, that determines the temporal extension of the collective structure. In our model, this time scale emerges from the evolutionary dynamics and is tightly linked to the ecology of the composing populations, as well as to their strategic choices. In fact, one realises that the cellular level and the life cycle come to be associated

to distinct time scales. If one looks at the instantaneous growth rates, then the time scale associated to the fastest or average growth of any cell type is shorter than the cycle duration. This indicates that along the emergent life-like cycle, cells duplicate several times, as in multicellular bodies. Moreover, direct comparison of 'fitness' in terms of the growth rates at the two levels is not meaningful without specifying the time scale on which population expansion is assessed, a point that is not often considered in multi-level selection approaches (but see Bourrat).

Although this model does not explicitly account for the process of group formation as those discussed in Chapter 2, its hypotheses reflect some features of populations of motile particles: the tendency to spontaneously form aggregates, and the increased aggregation efficiency at high density. Active matter models, where cells are described as self-propellled particles, offer a more realistic description of the aggregation dynamics of *Dicyostelium*, as will be discussed in more detail in Chapter 4.

The model presented in this chapter is a first attempt at conceptualising how higher levels of organisation may have acquired their characteristic temporal structure, where collectives last longer than their composing units, and occur over and again. The solution we propose does not rely on externally imposed environmental variations, but fluctuating selection derives from the coupled ecological and evolutionary dynamics of the system (in the absence of the evolutionary component, the consumer-resource ecology is always at equilibrium). The main disparity with respect to actual aggregative life cycles is that here the motility phenotype is heritable across cell generations, whereas cell populations are able to produce a variety of phenotypes that are only partially heritable, if at all. As a consequence, their frequency can change also independently of the genotype. In our model, that only describes phenotypes, this would lead to an additional term in the frequency equation. This term could potentially alter the dynamics of the system, but it will not change its (generic) behaviour as long as it is small enough.

A general form of heterogeneity occurring in isogenic (or almost isogenic) cell populations is phenotypic multistability. The ability of cells to switch between alternative phenotypes has been shown to be a route to the establishment of life cycles and the concomitant evolution of division of labour in *P. fluorescens*[62]. In this experiment, a genetic switch allowed cell populations to steadily produce two phenoypes, one of which was adapted to collective life and the other to the solitary stage. Physiological or epigenetic switches were also obtained as a result of alternating selective pressures that did not involve explicitly a collective phase (though one of the two phenotypes is typically associated to biofilms)[8]. To what extent phenotypic diversity is a consequence of group living, having evolved in formerly uniform multicellular aggregates because of the advantages provided by division of labour, or phenotypic diversity was the driver of the emergence of multicellular life cycles is still unclear. The typical scenario that is theoretically hypothesised and corresponds to some experimental observations is that first groups form, then they differentiate. Alternatively, one could imagine that the capacity of producing multiple phenotypes drove the emergence of multicellular structures as we know them, and division of labour was obtained by creating an appropriate local environment that would evoke previously existing diversity.

The interplay of genetic and epigenetic (or physiological) sources of phenotypic diversity with the ecological and evolutionary dynamics started to be studied only recently, and I was involved in some interesting experimental studies addressing the mechanisms of environment-dependent bacterial phenotype switches[114,31]. I will not detail these works here, but go back to them in my Research Project, as including phenotypic variation at different time scales is one of the most exciting lines of research I intend to pursue in the future.

Supervision:

Leonardo Miele (MSc in Physics of Complex Systems 2017)

Selected invited presentations:

Invited speaker at the Workshop 'Evolution of life cycles', Plön, Germany (October 2018)
Invited speaker at the Discussion Meeting on Conflict and Cooperation in Cellular Populations, NCBS, Bangalore, INDIA (February 2020)

Related publications:

Miele L., De Monte S. (2021)
Aggregative cycles evolve as a solution to conflicts in social investment
PLoS Computational Biology 17(1): e1008617

All demands for justice and all theories of equality ultimately derive their energy from the actual experience of equality familiar to anyone who has been part of a crowd.

Elias Canetti
Crowds and Power

4

Phenotypic heterogeneity in aggregative life cycles

A NECESSARY CONDITION FOR EVOLUTION BY NATURAL SELECTION is that (cell) populations subjected to selection are phenotypically heterogenous. The main source of phenotypic variation are genetic differences, whose heritability sustains directional changes over multiple generations. Most evolutionary models are focused on this type of heritable variation, and on its effects on reproductive value, or fitness. Theoretical frameworks like game theory and adaptive dynamics (see Chapter 2 and 3) assume for simplicity that the phenotype, upon which selection acts, is perfectly inherited, thus implying that the phenotype is univocally determined by an invariable (or slowly variable) genotype.

Other non-genetic mechanisms, from phenotypic plasticity to epigenetic inher-

itance, however allow cells with the same genotype to have different phenotypes, that can be maintained over variable time scales. Phenotypic variation that is not due to genetic changes is increasingly viewed as key to facilitate evolution by providing pre-adaptations and increasing evolvability[107]. In models addressing the evolution of cooperative behaviour, however, the complexity of describing the map between individual and collective behaviour has so far largely obscured the exploration of the role of non-genetic sources of variability.

This chapter is dedicated to studying the effects of phenotype variability of different origin in the self-organisation of multicellular aggregates, and to exploring its potential evolutionary implications. Even though some of the conclusions are hopefully valid in general for cellular populations – in particular for cell populations that undergo an aggregation stage – models will be developed having in mind the biology of the 'social amoeba' *Dictyostelium discoideum* (Dicty).

Dicty is mostly found in the form of soil-dwelling individual cells that feed on bacteria. When cells run out of nutrients and are sufficiently dense, they aggregate in multicellular 'bodies' containing hundreds of thousands cells. Remarkably, these multicellular structures display a wealth of collective properties[79]. Their aggregation is one of the best described processes of cellular self-organisation. Cells that were formerly moving in isolation become increasingly sensitive to cyclic AMP. This metabolite is produced by starved cells, and sensitive cells not only chemotax towards higher cAMP concentrations, but also emit more cAMP themselves. In this way, they relay a local chemical signal on the scale of the whole population. Looking at aggregation from the bird's eye, as can be achieved in dark-field images from a low-magnification microscope, one sees density waves akin to those observed in nonlinear chemical reactions, that span from target waves to spiral waves, depending on the experimental conditions. Models of excitable media[74] have been extensively used to describe the propagation properties of such waves, indicating that under general conditions one can expect that self-organisation leads to structure the space of aggregation in domains. Cells belonging to one of those domains form streams and most of

them converge towards a unique center to form a tightly packed 3-D mound. In the streams, cells attach head-to-tail, so that in the mound they turn all in the same direction producing a whirling, round aggregate. A tip starts to emerge from the mound's upper part, and to move outwards bringing with it the rest of the cells, that come to be organised in a slug. The slug is highly tactic and provides a very efficient way for cells to move coherently towards more intense light, higher temperature, and higher humidity, corresponding to the conditions encountered at the soil surface. During the time the slug migrates, other chemical signals become important. Such morphogens structure the slug in proto-differentiated regions that will later turn into different tissues. Among those tissues, a majority of cells ends up forming a spore ball that is lifted above the ground by other cells, mostly those composing a stalk, that will eventually die. Spores instead develop into resistance forms that can forego feeding for a long time, and are dispersed by arthropods picking up the spore mass. A small fraction of 'loner' cells that did not join the aggregates, on the other hand, does not commit to multicellular differentiation and thus to social behaviour, but is still able to divide if nutrients are supplied[39,133].

Let aside the recent work, that was more extensively discussed in Chapter 3, on such 'loner' cells, most attention in evolutionary biology has been focused on the two terminal developmental fates that imply drastically different contributions to the following generation: the stalk and the spores. In particular, the reproductive value of one Dicty strain with respect to another is traditionally measured in terms of the fraction of cells each of those strains contributes to a chimera's spore pool[127]. Typically, the same amount of cells of two genetically different populations are let co-aggregate and spores collected from all the fruiting bodies that were formed. The strain that forms the largest fraction of spores is called 'cheater', because its genotype frequency will be increased in the next generation of solitary cells, which emerges after spores exclosure. The other genotype is called a cooperator. As we pointed out in Forget et al. (commented in Van Cleve), this definition views genotypes as players that engage in a binary game, and treats a strain as an 'organism', or unit of selection. Coop-

erators pay a cost because they are underrepresented in the successive generation, either because of their reduced frequency in the spores alone, or of their possible consequent increased frequency in the stalk. As discussed in Chapter 2, the expectation is that over multiple rounds of interactions between the same cheater and cooperator strain, the former will exclude the latter. Hence, cooperation should be eliminated over evolutionary time scales, and its existence in nature is a paradox. The solution of this 'paradox of cooperation' led several authors to invoke, for Dicty as well as for other aggregative microbes like *Myxococcus*, similar concepts as those used to explain cooperation in humans and in social insects[141,126]. If these studies indicated that the theory of sociobiology – initially formulated to understand social insects' sterile castae[61]– could be potentially applied to understand the ultimate causes of cooperation in populations with different structures, a multiplicity of proximate causes has been proposed to explain how such structure emerges. Some confusion derives from the fact that these multiple mechanisms do not all have the same status in terms of their evolutionary predictions, thus are not equivalent ultimate causes[43].

In order to identify the most appropriate null model for the evolution of sociality in heterogenous cell populations, it is useful to view behaviour at the level of the individual cell, rather than of the strain or population. In this perspective, players that are engaged in social interactions are characterised by their own phenotype (e.g. the probability of becoming a spore), and not by the genotype of the strain they belong to (e.g. the fraction of spores produced by a given strain). From this perspective, it is easier to conceive that the phenotype of any cell is determined not only by the genes it carries, but also from its environment. A key feature of collective behaviour is that such social context is primarily influenced by other cells and is in general highly dynamic. Therefore, the notion of adaptation to the environment needs to be modulated so as to encompass processes that occur on times scales faster than a single life cycle. It is often assumed that defining social behaviour from the point of view of the strain and of the single cell is equivalent: cheating as defined by spore bias is identified with a higher probability (determined a priori by the genotype) of

a cell to become a spore, and a consequent lower probability of becoming stalk and die. Equivalently, one could say that the behavioural strategy of one cell is the genetically encoded probability of cooperating (in game-theoretical terms, this would be a mixed strategy).

Even in this simple case, where the (probabilistic) behaviour is genetically encoded, the existence of different levels of organisation, and thus of social groups inside which interactions take place, can create a mismatch between the definition of cheating at the level of the individual and of the population composed of individuals of that type. For instance, individuals could interact pairwise via a Prisoner's dilemma, but when population structure is taken into account, the resulting effective game played by different populations can be a stag-hunt game, where the predicted outcome is the coexistence of multiple strategies[106,135].

The outcome of selection for collective properties of the multicellular stage becomes even more uncertain when one factors in the observation that neither the probability of becoming a spore, nor the structure of the population, are defined exclusively by a cell's genotype, but are a product of many factors, including the previous history of the cells, the timing of starvation, and of properties of the other population, including cell type and number. We have reviewed these various factors in Forget et al..

In the course of the ongoing thesis of Mathieu Forget, we have explored the possible role of different sources of phenotypic diversity in determining spore bias. In particular, we focused on the biases that can be induced in the course of aggregation by the fact that cells are, at the moment of starvation, in different states of growth. We called *chronochimerae* the mixes of two isogenic cell populations that were harvested at different times along their demographic cycle (e.g. in early, mid and late exponential phase). Since they are genetically identical, we can exclude that biases observed in the spores are due to kin discrimination, which is considered to be a prominent factor in determining fitness differences among strains. Co-aggregation of the two subpopulations in chronochimerae allowed us to evaluate the extent to which reproductive success is affected by phenotypic variation occurring at a scale that may occur in

nature, where environmental microstructure is likely to make cells more or less starved at the moment of aggregation (that is triggered when only a small fraction of cells start emitting the common cAMP signal). Moreover, differences in demographic phase at the moment of harvesting provide a control parameter that can be continuously tuned, thus affecting phenotypes in a gradual way.

The results of Mathieu Forget, that I do not detail here because they are the subject of a manuscript in preparation, demonstrated that spore bias (the frequency of spores belonging to one subpopulation) is affected by the time of harvesting. These results complete a previous observation that spore bias could be caused by differential duration of the starvation phase, and extend it in several ways. First of all, they demonstrate that the bias is frequency-dependent, as for most genetic chimerae: in general, the relative selective advantage of one type decreases when the ratio of cell types is highly skewed, and is maximal for more balanced mixes. Second, the bias can change direction depending on the difference in demographic timing. This physiologically-induced bias is quantitively compatible with biases observed in some (but not the most extreme) genetic chimeras[15], indicating that it can potentially interfere with selection acting on genetic causes. In order for this to happen, however, it is necessary that genetic differences do not completely override those induced by demography. In order to test if this was the case, Mathieu repeated the same experiments using two different strains, which were derived by Sandrine Adiba in an experiment where selection was applied on the ability to adhere to a surface[2]. These strains produce very strong biases in chimeras with the ancestor, even though we do not expect them to be different in terms of their specific recognition systems, because they are the result of a (relatively short) evolutionary trajectory driven by selection on isolated cells. Although in this case demographic variation was unable to completely revert the sign of the bias, it caused it to changed quantitatively in accordance to what observed in isogenic populations. Even more strikingly, when Mathieu repeated the same experiments by mixing well-characterised strains, including the most famous 'cheating' strain ChtA, he showed that the bias could be completely reversed, so that a cheater would

transform into a cooperator when it is mixed with a culture of a 'cooperator' strain that is sufficiently more advanced in the demographic cycle. Not only the definition of cheating is relative, but there are intermediate situations where the frequency-dependent bias changes sign, indicating bistability of the evolutionary outcome of repeated interactions.

The observation that genetic and plastic phenotypic changes are additive (to a certain extent at least) supports the idea that unavoidable variation in the timing of aggregation, a parameter that is not under the control of any single cell, can compete with selection on genetically determined traits. Similar to previously proposed 'lottery-like' mechanisms that interfere with the causal relationship between the cell genotype and its developmental fate (like dependance from the cell cycle phase[59] and from local density[116]), differences in growth state can potentially underpin the apparent lack of strong directional selection for increasingly cheating genetic variants.

This work paves the way to the systematic exploration of the interplay between genetic and plastic variability in determining biases of evolutionary relevance for *Dictyostelium discoideum*. The observation that the demographic age at starvation reproducibly affects social behaviour provides a means to tune spore bias continuously – unlike what is achieved by mutagenesis – and without affecting the genetic architecture. In complement to molecular approaches, our work thus opens a route to explore the coupling of short-term processes, involved in aggregation and development, with the evolutionary dynamics.

If non-genetic biases in reproductive fate can be of similar magnitude of those induced by genetic differences, defining a priori what is cheating behaviour becomes tricky, especially when cells are genetically close. Either cheating is defined just based on the outcome of a single aggregation – so that bias is not necessarily a meaningful predictor of the long-term evolutionary dynamics – or it is based *a posteriori* on evolutionary success, so that it can be difficult to connect it to individual cell behaviour within a single life cycle. These different scales are bridged if spore bias is univocally determined by the genotype, and is

independent of the context. Otherwise, as it seems likely to be the case for *Dictyostelium*, definitions of cheating based on proximate and ultimate causes may not overlap, making the use of measures of social behaviour questionable.

One possible way forward in defining meaningful proxies for the selective value of social behaviour in Dicty is to take into account not only the fraction of spores that one subpopulation produces, but also the total number of spores produced by a chimera. As eco-evolutionary models extend the replicator equation by encompassing demography on top of frequency dynamics (see Chapter 3), the definition of reproductive success in Dicty can be expanded beyond the measure of the relative success in spore formation. Indeed, different subpopulations can differ not only in the fraction of spores they produce, but also in their number compared to isogenic aggregations. As was discussed in Buttery et al. and Adiba et al., taking the latter information into account can alter, even qualitatively, the classification of social behaviour. Moreover, in a metacommunity-structured population, variation in the number (not only the fraction) of spores produced allows the coexistence of types that have different levels of investment in the social phase[133].

Going beyond the simple definition of cheating based on spore bias, however, requires a deeper understanding of how cells get partitioned in different aggregates. Beside the already mentioned fraction of isolated cells, additional meaningful population statistics are the average aggregates' composition (fraction of cells of the two types within one aggregate) and the degree of segregation (variation in composition, detecting the tendency of cells to co-aggregate with those of the same type). In limit cases when the distribution of aggregate composition is simple, these quantities are in simple relationship with the usual measures. For instance when all aggregates have the same composition, and all cells behave the same in the multicellular stage, the average composition corresponds exactly to the fraction of spores. This can nonetheless be different from the fraction of cells of the two types in the initial mixture, if they differ in the proportion of loners they produce. When like aggregates with like, on the opposite, spore bias will be the average of the spores formed by each type in isolation,

provided that this is estimated by matching the initial cell number to their initial frequency. One can expect that, in general, none of these extreme cases will occur, so the relation between spore production in isolation and in chimeras will be nontrivially affected by factors that alter group formation as much as by biases induced in the multicellular development.

As was discussed in Chapter 2, cell physical properties, such as adhesion, can produce, during group formation, assortment between cells belonging to different populations. We therefore decided to explore whether the changes in spore bias observed in chronochimeras correlated with changes in single-cell behaviour of the composing populations. First, we considered whether qualitative changes in spore bias could be traced back to early stages of the multicellular cycle by determining whether biases were already present at the moment of aggregation. Indeed, biases opposite to those observed in the spores were retrieved when looking at the fraction of isolated cells. This indicated that the fraction of cells that remain outside aggregates might influence what population gets overrepresented in the spores, as also postulated in models including a loner strategy[39,133,116].

We then looked for possible candidate cell phenotypes that could provoke such biases during aggregation. Mathieu examined two properties: the ability of cells to attach to a substrate (that has implications for their social behaviour[2]) and their motility. Whereas adhesion to a surface increases monotonously during the demographic cycle preliminary observations indicate that cells (or at least a fraction of them) are most motile in the mid- to late exponential phase. Previous works moreover suggested that, before aggregation in a uniform population, cell motility is bimodally distributed, and that such differences may have bearings in terminal cell fate determination[55]. Moreover, recent observations found that this fate is also affected by levels of ATP before aggregation[70], which may be related to cell motility.

Based on these observations and on the possible evolutionary implications discussed in Chapter 3, we decided to explore the possible biases induced by

motility differences in virtual cell populations. Rather than a model for the specific aggregation process described above, however, this has to be conceived as a means for establishing null expectations on the biases that can be achieved in the absence of any sophisticate form of cell-cell interaction, e.g. signalling, specific recognition, gene regulation, that we know are important for Dicty and beyond.

We modelled cells as self-propelled particles endowed with an intrinsic capacity of propulsion and with the possibility of interacting with each other on a short range (less than a cell's radius), in order to mimic cell adhesion. Such an individual-based model had been proposed by Szabó et al. as relevant to describe the collective motion of fish keratocytes, that have amoeboidal motility akin to that of Dicty. In this simple model, each cell's phenotype is defined by a handful of parameters: a scalar velocity v, the time scale of persistence of motion τ (precisely, the resistance to variations of the particle's direction of motion upon perturbation by an external force), the cell radius R_{eq}, which is the position where the attractive adhesive forces (acting between cells within a distance $R_0 < 2R_{eq}$) and the short-range repulsive forces equilibrate, and the intensity of these forces. Moreover, an angular noise term ensures that the system cannot attain unrealistically ordered configurations. In the framework of active matter physics, such a model and others with different choices of the single-cell and interaction properties, have been shown to give rise to a wealth of collective behaviours, ranging from swarming to motility-induced phase separation[112]. Understanding how changes in the microscopic, particle-level, properties can lead to phase transitions, that is changes in the macroscopic or collective behaviour of populations of particles has been addressed with numerical simulations but also with analytical tools drawn from soft matter physics. Such powerful theoretical approaches let researcher focus more on the asymptotic, statistically stationary, collective states, than on transients. For a biological system like Dicty, however, transients are important, because once mounds are formed, their cells proceed along a developmental path that drastically changes their individual properties. Active matter physics models can however be used to explore the

transient dynamics leading an initially uniformly distributed population of cells to form aggregates within a finite time.

In collaboration with Leonardo Gregory Brunnet (Universidade Federal do Rio Grande do Sul, Bazil), we have been studying the aggregation of the previously described Szabo's particles both in uniform and heterogenous populations. In particular, we focused on variations of the scalar speed of particle motion. We numerically simulated populations large enough to produce multiple aggregates (illustrated by the snapshot reproduced in Figure 4.1), similar to what is observed in the experiments. First, we characterised the phase diagram of this system at finite time. We found that, at sufficiently small (but not excessively small) densities, particles are found to aggregate efficiently for intermediate values of motility, whereas there are two regions where most particles are found in isolation. For large velocities, the transition out of the condensed phase corresponds to a liquid-gas transition enabled by an increase in particle speed. For small velocities, particles that are found outside would instead join aggregates if one waited long enough, and the transition between a gas and a liquid phase is only a finite-time effect. However, such finite-time effect could be relevant for a biological system, whose timing of aggregation is constrained by the existence of a developmental program with a set time scale. It could as well have been important at the onset of multicellular organisation, where the selective value of a collective structure may depend on a varying environmental variable, as discussed in Chapter 3. We therefore decided to focus on the finite-time result of aggregation rather than on the asymptotic state of the system (though for large values of the velocity, the two are almost equivalent).

Using the phase diagram for a homogeneous population as a reference, we then studied the effect of velocity differences on the composition of the aggregates and on the partition of cells between the aggregated phase and the 'gas' of solitary particles. We placed ourselves in a region of the parameter space where initially randomly and uniformly distributed particles aggregated in several groups that did not display directional collective movement, sometimes leaving a fraction of the particles moving in isolation. In order for the model to represent

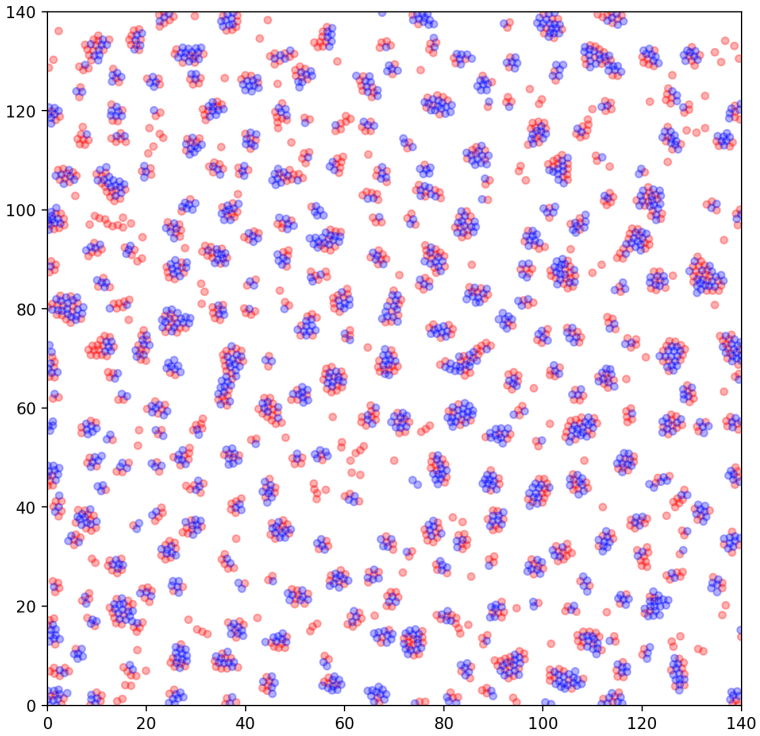

Figure 4.1: Snapshot of a numerical simulation for a population of self-propelled particles. The particles undergo, when isolated, a persistent random walk, and interact within a finite range by elastic interactions, as proposed by Szabó et al. to model ameboid cells. In the simulations Mathieu Forget realised, particles are chosen to have differential speed (here, red particles are slower than blue ones). When initialised in uniform random positions, the bias in their representation within and outside clusters depends on the degree of heterogeneity chiefly through the degree of condensation of each particle type in isolation, and can be qualitatively understood based on the phase diagram of homogenous populations.

Dicty aggregation, we wanted indeed to avoid regions of the parameter space where particles would form groups just for steric reasons (due to their finite extension, as for Motility-Induced Phase Transtions[41]) or groups displayed collective movement[60]. A first observation we could draw is that all the aggregates had a similar composition, reflecting the lack of initial segregation between particles. Therefore, all aggregation-induced biases observed in the groups were due to the different composition of the pool of solitary cells. Contrary to the models for loners, where joining a group is the result of a strategic decision (either dic-

tated by a fixed probability of aggregating[39,133], or dependent on signalling[116]), here the fact of partaking or not an aggregate is a 'passive' outcome of differences in the particle physical properties.

As a consequence of their physical underpinning, biases in aggregates' composition depend on the properties of the two populations of cells in the mix, notably velocity and adhesion parameters of the two types, but also their relative frequency. We therefore addressed the nature of the possible compositional biases and how they depend on the differences between the co-aggregating particles and on their frequency in the population. This study, that is detailed in a manuscript in preparation, showed that frequency dependence of the bias is the rule rather than the exception, as soon as particles of different speed are let co-aggregate. In principle, such biases can belong to four qualitative classes (Figure 4.2): either the focal particle type is always overrepresented in the groups, or the other is, or the sign of the bias depends on its frequency, that is the fraction of particles of that type that were initially present in the mix. In the latter case, the two simplest scenarios are that the bias is positive when the frequency of the focal type is small and negative otherwise, and vice-versa. This distinction is important because these four different scenarii give rise to different predictions on the evolutionary success of the focal population across succesive cycles of aggregation and dispersal. If the properties of the two subpopulations do not change in time, in the previously listed four scenarii the focal type would be predicted, respectively, to fixate (driving the other type to extinction, as in the case of the tragedy of the commons), to go extinct, to coexist with the second type or to be alternatively fixed or excluded, depending ont its initial frequency (bistability).

It turns out that co-aggregation of particles with differential motility does not appear to be able, at least for densities low enough to be in a realistic range for Dicty cells, to produce all these possible classes of frequency-dependent biases. Indeed, in our simulations the focal type was. always observed to have a positive bias at low frequency and a negative one at high frequency, that is the condition required for the two types to coexist in the long run. This scenario oc-

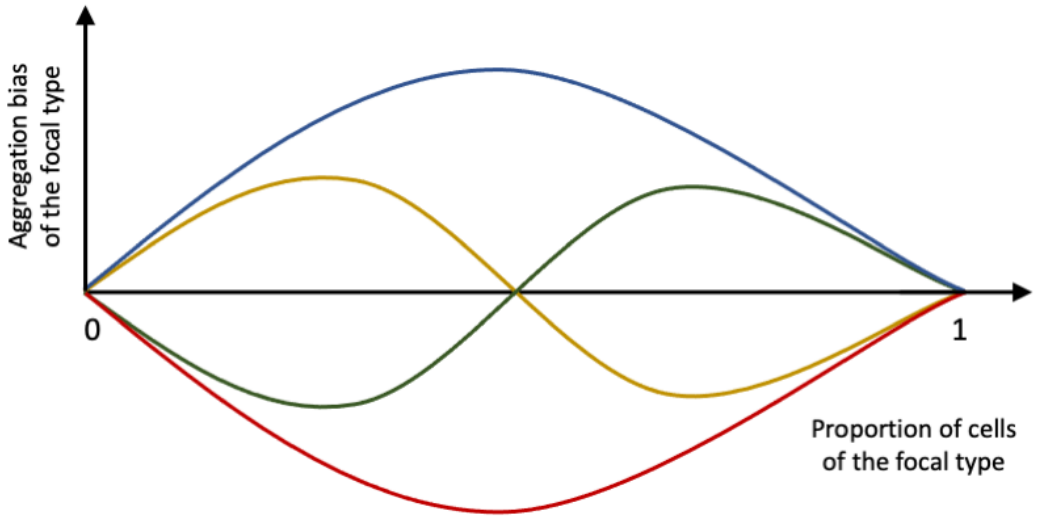

Figure 4.2: Illustration of the four simplest forms that can take frequency-dependent biases, such as those induced by differences in the efficiency with which cells aggregate. If aggregate composition is also reflected in the proportion of spores produced, i.e. if there are no biases introduced by development (as one might expect for primitive multicellular organisms), then these biases correspond to four possible evolutionary outcomes: the focal type is selected (blue curve); it is driven to extinction (red curve); one or the other of these outcomes occurs, depending on the initial frequency of the focal type (green curve); both types coexist (yellow curve). Of these four scenarios, only the first three are produced starting from an individual-based model of aggregation of self-propelled particles.

curs when both particle types aggregate rather efficiently on their own, as is the case for most Dicy strains, and in particular for chronochimerae.

This observation can be explained by considering that group formation occurs under the competing processes of delayed aggregation of slower particles and evaporation of faster ones. Asymptotically, the slowest type will be more represented in the aggregates because faster particles evaporate more. Faster particles can however be transiently more frequent within aggregates because they need less time to join them. Knowledge of the fraction of particles that aggregate in a homogenous population is a good indicator of what kind of bias will manifest when different populations are let co-aggregate, and a way to predict

the outcome of a complex aggregation process.

The predictive power of such macroscopic observable extends to simulations run for different values of particle adhesion. Moreover, it can be used to understand why the bistability scenario is never observed. A bias that would be negative at low frequency and positive at high frequency would require that the total aggregated fraction was a concave function of particle speed, whereas the opposite is true: slow particles tend to join aggregates with some delay, and fast particles to evaporate.

In our experiments on isogenic chronochimerae, we never observed biases that changed sign with varying the frequency. However, we could evidence it in some genetically heterogeneous chronochimerae. This kind of bias was moreover observed in experiments on chimerae of different natural strains[90] and of different species of the genus *Dyctiostelium*[118]. Although the authors provide a different explanation for their observations, in the former study[90], interestingly, the bias was always such that the rare type would invest more in the stalk – thus would be underrepresented in the spores – and the reverse when common, corresponding to the bistability scenario. On the other hand, the second study[118] revealed a strain of *Dictyostelium giganteum* that displayed positive bias when rare and negative when more common against at least two other strains, and even more complicated nonlinear frequency-dependence whereby the sign of the bias changed twice. Our studies suggest that, in order to understand the possible role of aggregation biases in shaping the composition of groups – which will inevitably affect successive biases established during multicellular development – it is necessary to look at the whole life cycle, and not only at its endpoint (as also claimed in Tarnita et al.). Moreover, we have shown that cell properties might play an essential role in constraining the possible biases achieved during aggregation, thus affecting the evolutionary trajectory not only of extant aggregative microbes, but also, potentially, of the fist aggregative multicellular life forms.

Supervision:

Mathieu Forget (PhD 2022), co-supervised with Sandrine Adiba

Selected invited presentations:

Keynote speaker at the conference Learning, Evolution and Games, Lucca, Italy (June 2022)

Related publications:

Forget M., Adiba S. and De Monte S. (2021)
Social strategies of D. discoideum across scales of organization
Peer Community Journal, Volume 1, article no. e58.

Adiba S., Forget M. and De Monte S. (2022)
Evolving social behaviour through selection of single-cell adhesion in Dictyostelium discoideum
bioRxiv, in revision

Forget M., Adiba S.*, De Monte S.* (2022)
Ephemeral cheating: timing of aggregation modulates social behaviour in Dictyostelium discoideum
in revision

All the physical and chemical laws that are known to play an important part in the life of organisms are of this statistical kind.

Erwin Shrödinger
What is life?

5

Community abundance distributions

Interactions among cells affect collective behaviour also when they are not embedded in the physical structure of a multicellular body, as we have considered in the last chapter. Inside or outside well-recognizable multicellular organisms, single cells compete for nutrients, exchange signals, predate or parasitise other cells, in short undergo an ecological dynamics that is fundamentally dictated by their genotype through the phenotype. Even when we neglect phenotypic plasticity, genetically-based phenotypic heterogeneity is the rule in most microbial ecosystems. Think of cellular assemblages such as biofilms, mutualistic or symbiotic associations and communities (groups of organisms interacting with one another and with the environment in a specific region). They are all associated to collective functions, such as stress resistance (achieved for instance through the production of an extracellular matrix), exploitation of resources, or ecological functions like ecosystem productivity and

resilience, that rely on the interplay of different – typically heterogeneous – component cells. Letting, for the time being, aside the question of how do such collective functions evolve (that will be the subject of the following chapter), I will focus here on how diverse microbial communities assemble and what might be the signatures of underlying ecological processes, both those common to all communities and those that manifest only in specific ones.

As for the case of aggregation in Chapter 4, this chapter addresses processes that occur on an ecological time scale, therefore where the types of cells that are interacting are thought not to change (as they would along an evolutionary trajectory). This oversimplification neglects the possible coupling between ecology and evolution (as discussed in Chapter 3) that would be expected when evolution is sufficiently fast to generate new types on the time scale of the observations. Though certainly important for microbial populations, evolution adds another layer of complexity and will be discussed in Chapter 6. The work I'll present here inscribes instead in the broad field of ecology that aims to uncover general patterns and processes in ecological systems[12,87,91].

Macroecology and biogeography have strived for decades to identify the governing principles beyond the spatial distribution of organisms, rooting them in their relationship with one another and to their environment. The idea is that ecosystems similar in terms of abiotic and biotic interactions should be associated to similar values of some community-level statistics, for instance species diversity, turnover, productivity and so on. These regularities would in turn order into a number of equivalence classes the multiplicity of biomes that are observed on Earth. Even more importantly, if these classes were to be associated to some underlying mechanism, for instance metabolism or species interactions, then one might hope to formulate predictions on how a given ecosystem transitions from one to another class, or how it is maintained where it is. This program of research has been applied traditionally to communities of macroorganisms, for instance birds or plants, where observational data were made available by systematically measuring all organisms of any given type (typically, species) present in a given area at a given time. Particularly apt to the purpose have been commu-

nities that contained a large number of species – for instance, certain fish and tree communities – because they provided sufficient power to compare distributions.

The most widely available data on ecological communities are counts of individuals belonging to the same species, without distinguishing thus among stages of life cycle stage, nor considering other sources of phenotypic variation. These observations provide, among others, information on biodiversity, species distribution ranges and turnover, and are essential in the identification of different biomes. As a consequence, patterns that have been extensively studied are those that appear when one looks at species abundances. These abundances are classically ordered independent of the identity of the species, for instance by rank in the rank-abundance plots (RADs) or by number of individuals, pooling together all species with abundance in a certain interval, as in Species-Abundance Distributions (SADs)[99,95]. Ecologically important indexes are directly related to these representations: species richness - the number of different species present in a community - is the support of a RAD; indexes such as Shannon diversity, Simpson index and species evenness are directly calculated from these distributions; while other estimators, e.g. Chao's for species richness, are based on the assumption of their regularity.

A major obstacle to assessing the generality of the principles highlighted for species-rich communities, typically observed in tropical ecosystems, was that the vast majority of non-tropical biomes did not support a sufficiently high number of species, or even of organisms within species, to make statistical comparisons meaningful. For microorganisms, this limitation used to be even more severe, because the classification of single cells in taxonomic types required at least a very expert eye, and was often impossible. Ramon Margalef's amazing dataset of diatom abundances included some thousands of species, but nothing of the kind was available for bacteria, that display a much smaller morphological diversity. The game changed suddenly 15 years ago when shotgun sequencing was first applied to environmental samples[117], revealing a previously unforeseen genetic diversity of marine plankton bacteria. Following this first study,

a plethora of other investigations looked at microbial communities through the lens of ever more powerful genomic methods, all confirming the enormous diversity of genes present in the environment. Though the best way to identify microbial 'species' – or ecologically significant groups of organisms – is in itself an open and very debated subject[36,88,53,115], it was natural to apply to those new datasets the approaches previously developed for macroorganisms[34]. This effort is still ongoing, and the state of the art gets continuously reassessed as newer and more extensive datasets become available.

I approached this field via my interest in microbial plankton communities. Cells of these communities are relatively well-mixed, on the scale they are sampled at, by small-scale turbulence, and therefore share some features with lab cellular cultures of the type I had previously studied[23] – at least more than microbial populations in other, more structured, environments. Like in *Dictyostelium*, cells in plankton communities can potentially share public goods through diffusion in the water, and they certainly interact through soluble compounds, other than competing for common resources. Microbial communities are hence increasingly conceived as highly integrated ensembles of cells, and viewed as a superior form of 'organism', as I will discuss in more depth in the next Chapter 6. On the other hand, plankton communities are not contained in visible boundaries that ensure their spatiotemporal coherence, and - contrary to most organisms - are constantly subjected to mixing with neighbouring communities and are strongly influenced by environmental variation. Their position at the boundary of what we typically conceive as coherent biological units made them particularly attractive to my eyes as an ideal system to address the interplay of ecological dynamics and environmental forcing in shaping collective function.

Beside their relevance for addressing theoretical questions, plankton microbial communities also happened to offer the opportunity of developing collaborations in my local work environment. Chris Bowler and his team at IBENS (particularly Lucie Zinger) were indeed heavily involved in the *Tara oceans* expedition, which was set to explore and characterise marine plankton biodiversity

across the world oceans[78,10]. The data from the metabarcoding analysis of the nearly 400 samples from over 120 locations, spanning all oceanic basins except the Artic ocean, were starting to be produced, and to reveal the extraordinary genomic diversity of marine plankton. Clustering of single reads into Operational Taxonomic Units (OTUs) produced counts of microbial 'species' that could be compared to species distributions observed in macroorganisms[130,30]. Under Chris' direction, Shruti Malviya was looking in particular into how protists (unicellular eukaryotes) diversity varies from one to another place in the global ocean. She was finding that diversity changed surprisingly little from location to location, the main signal of variation between samples being the presence of geographical features, such as barriers that separated different oceanic regions, or upwellling currents. The latitudinal gradient of Shannon diversity commonly observed in macroorganisms (later evidenced for some groups of protists[75]) was not evident[30], despite the important role that temperature played in community assembly. The curiosity for the reason of this apparent mismatch between micro and macroorganism diversity led me to look into the source of the ecological indexes that Shruti had been analysing, that is the distribution of OTU abundances.

Before I address the adherence of plankton communities to known macroecological patterns, I'd like to briefly present a part of my work – realised in collaboration with oceanographers Francesco d'Ovidio (LOCEAN, Paris) and colleagues – that provided me with a picture of the physical scales and processes that are relevant for marine plankton populations. Aforementioned metagenomic approaches are a great tool to precisely caracterize the local composition of plankton communities, to the cost, however, of a huge effort in sequencing and data analysis. They are therefore less adapted to the observation of communities on large spatial scales than other methods which are less taxonomically precise, but have higher spatiotemporal resolution, such as flow cytometry or remote sensing[19]. It was by combining different datasets of satellite-based observations, numerical simulations and mathematical methods from nonlin-

ear dynamics that we addressed the question of how ocean circulation sets the boundaries of coherent plankton assemblages[37]. Our analysis revealed that oceanic fronts and eddies that structure the ocean surface at the sub-mesoscale (tens to hundreds of Kms) separate communities characterised by their dominant phytoplankton type (identified by analysing water-leaving radiance with the algorithm Physat developed by Severine Alvain) (Figure 5.1, right panel). The calculation of the Finite-Size Lyapunov Exponents - a method originally introduced to characterise the chaotic behaviour of non-autonomous dynamical systems - and Lagrangian simulations informed by altimetry-derived surface currents, moreover, allowed us to estimate that water mass coherence was maintained – in the highly energetic regions where plankton blooms are most powerful and diverse, at least – for time scales of a few weeks, that is similar to the duration of a bloom. This time elapsed, such 'fluid-dynamical niches' are destroyed by the complex dynamics of ocean currents, as coherent patches elongate in narrow filaments that are eventually dispersed by small-scale turbulence. A bird's eye view of the seascape obtained by Lagrangian computation of the FSLEs (Figure 5.1, left panel) reveals how complex and dynamic the environment experienced by planktonic cells can be, as soon as they leave the coherent water mass that constitutes their 'niche".

These considerations led us to conceive that, if plankton communities are chiefly influenced by the quality of their environment, the dynamics of the water mass they belong to imposes a sort of a spatiotemporal 'life cycle', whereby a given community thrives in a local patch before being dispersed by currents on scales much larger than what possibly achieved by diffusion. As a consequence, one might expect that local diversity is enhanced when two conditions are met: a landscape-level heterogeneity of communities (e.g. Figure 5.1, right panel) and a sufficiently high mixing induced by surface currents. Having shown that these two features are related, we proposed that a proxy for local diversity is an index of sub-mesoscale seascape heterogeneity. This hypothesis was confirmed during the thesis of Alice Soccodato (co-directed with Francesco d'Ovidio) by analysing the diversity of functional groups in the MIT Darwin model, where the ecology

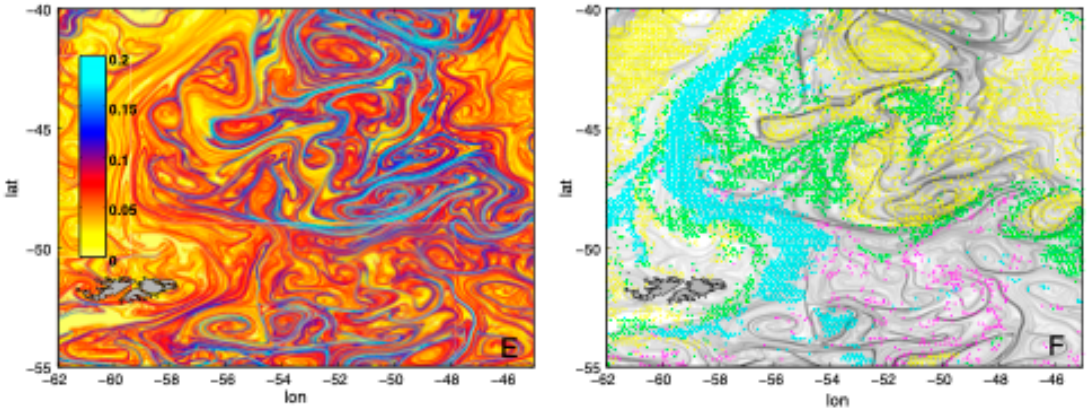

Figure 5.1: Seascape of the ocean surface off the Patagonian coast[37]. Left Panel: Finite-Size Lyapunov Exponents of the ocean surface velocity field computed from remote-sensing of altimetry. The higher the exponent, the more intense the stretching along the ridge, and the larger the barrier that the front represents to circulation between adjacent water masses. Right panel: the manifolds identified by the largest FSLEs are superposed to the spatial distribution of phytoplankton dominant types, identified through the ocean colour spectrum via the PHYSAT algorithm (cyan, coccolitophores; green, diatoms; yellow, small eukaryotes and cyanobacteria; pink, phaeocystis).

of a virtual plankton community is simulated in the environment set by a global circulation model[42,123]. Its application to remote-sensing data[29] showed that the seascape diversity index displayed patterns consistent with previous observations, both in terms of the geographical distribution of diversity and of the relationship between sea surface temperature and diversity. They nonetheless showed some features that differed from the majority of observations for other organisms: a diversity peak at intermediate latitudes, that had been predicted by the Darwin model and is a signature of the ocean circulation structure; and a similarly unimodal trend in the relation between temperature and diversity, as predicted by some evolutionary models based on the metabolic theory of biodiversity[125]. These patterns suggested that, given their embedding in a complex fluid dynamics environment, plankton populations may display ecological patterns that vary with respect to those of macroorganisms, but variations could be also expected relative to microorganisms that live in environments with different spatiotemporal variability, as for instance the gut microbiome.

The opportunity to pursue this idea was provided by the *Tara oceans* dataset and by the hiring of a brilliant post-doc, Enrico Ser-Giacomi (currently post-doc in Mick Follow's group at MIT), who worked with me for one year. By looking at the distribution of relative OTU abundance sample-by sample, we realised that the small variability observed in the Shannon diversity index reflected a strong uniformity in the RADs and SADs. In particular, it stroke us that most of the sample-to-sample variation was concentrated in the distribution of abundant species, whereas the abundance of rare ones seemed to decay at a similar pace. Plotting these distributions in double logarithmic scale, the decay revealed to follow a power law (Figure 5.2 a). The challenge now was to make this observation quantitative, and in particular to determine whether the inevitable variations in the power-law exponent – which would greatly influence diversity estimators like Shannon or evennes indexes – could bring some information on the nature of the community. Having realised that dominant species greatly biased Shannon diversity in blooming stations (the presence of a small number of very abundant species would drastically reduce this index, without that implying necessarily that the background community was less diverse), we hoped that geographical patterns of rare species might be more robust than those identified considering the whole community.

The larger part of research on ecological patterns, including in microbes, had focused on finding which among a large number of theoretically predicted distributions provided the best qualitative match for empirical distributions, and more and more sophisticated statistical tools were being developed in order to tell apart ambiguous cases. Their application to the increasingly available large microbial datasets was however inconclusive on which model would provide the best explanation for the observations[68]. I see two reasons (at least) why this program is unlikely to succeed on its own. First, not only some of the theoretical distributions share so many important features that they are probably indistinguishable given the sampling and sequencing errors, but they sometimes converge, for some parameter values, to identical distributions. Second, and even more substantially, I see no reason that one single distribution should fit the

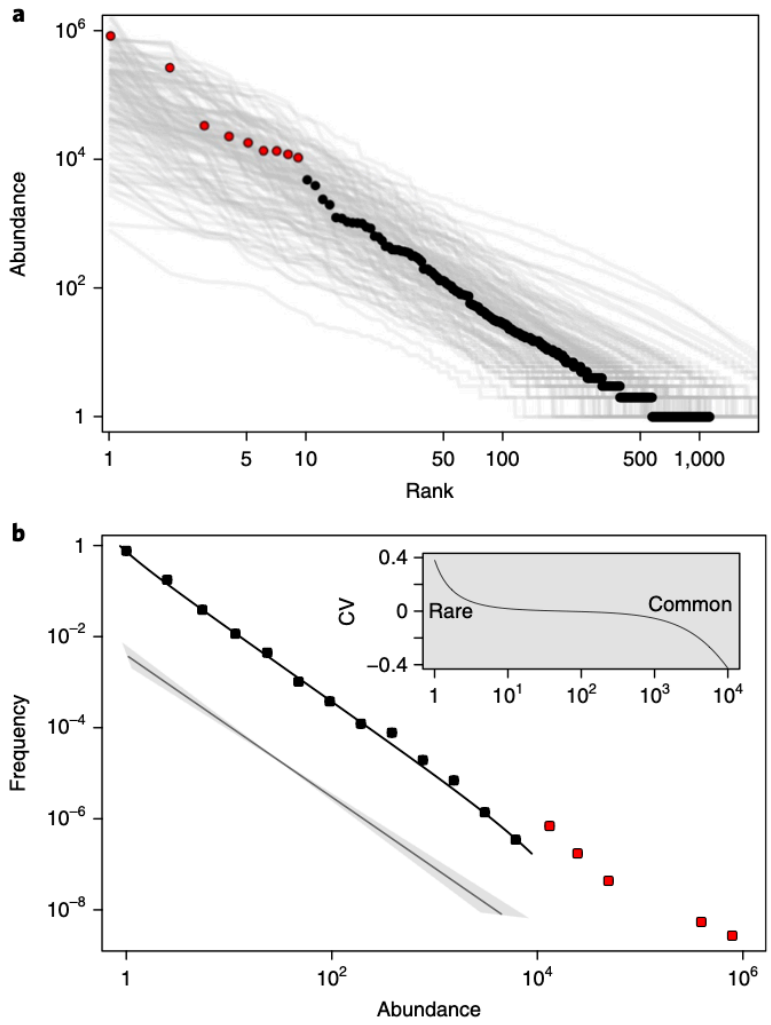

Figure 5.2: Abundance distribution of OTUs for marine plankton protists, obtained from the *Tara oceans* expedition. a) Rank-abundance plots for all samples (grey lines) reveal that the abundance decay for non-dominant species consistently follows a power-law. b) Emmpirical SAD for the community highlighted in panel a. The frequency of non-dominant species (black symbols) displays a different regime from dominant species (red symbols). The distribution (black curve) obtained by adaptively fitting a model to the non-dominant component of the community has a leading associated power-law trend. Its exponent λ displays a limited range of variability across physical location and cell size. Illustrated below in grey are the average exponent and its variation across all samples. The inset displays the coefficient of variation (CV) between the best-fitting distribution and its associated power-law, showing that the model can capture deviations from the power-law scaling for both abundant and rare species. The picture is reproduced from Ser-Giacomi et al..

distributions of all species of communities open to immigration. As convincingly demonstrated by Ann Magurran and Peter Handerson, open communities will 'import' all the time species that are not endemic, so that species abundances will result from the superposition of distributions corresponding to different and largely independent ecological processes[92].

The solution to these problems appeared to us twofold. First of all, we should find a way to find an objective criterion to separate the dominant from the rare component of the community. Contrary to Magurran and Handerson, we could not connect OTUs to their range, as most of planktonic species are poorly ecologically characterised, if at all. We thus thought of using the regularity of the power-law decay as providing an ansatz for what part of the community we could consider 'rare' (instead of setting arbitrary thresholds on maximal abundance and/or occurrence, as commonly done). To avoid possible criticisms on the reliability of identified patterns, we also needed to be confident that a power law really provided a good fit to the empirical distributions. We solved these issues by fitting the observations with a model that embedded the two main qualitative types of abundance decay: a power law and an exponential. We could thus identify which of the two trends was dominating in the data, while at the meantime realising a quantitative and not just qualitative comparison among samples. Moreover, we could investigate possible correlations of the quantitatively fitted parameters with environmental measures.

Among the numerous classical models that had been proposed to explain SADs[96,99], the few that produced (often as a limit case) power laws were predicting decay exponents of 1, whereas we did not want to constrain *a priori* the decay exponent. Inspired by the review by Azaele et al., Enrico decided to consider a variation of Hubble's neutral model that maintained species exchangeability, but introduced a density-dependence on the stochastic birth and death rates (a model we later realised had already been studied by He[67]). In the limit of large number of species, that Enrico derived, the ensuing distribution was the product of a power-law and of an exponential. Interestingly, the power-law exponent was no longer fixed but depended on the parameters weighting the

non-linear components of the demographic rates. Moreover, the distribution deviated from a power-law not only at high abundances, where it predicted an exponential cutoff, but also at very small abundances, as illustrated in the in inset of Figure 5.2 b. We adapted the algorithm by Clauset et al.[18] in order to devise an adaptive fit of the empirical data: starting from the 100 rarest OTUs, we would progressively enlarge the window where the model was fitted, until the significance of the fit was lower than a very strict threshold, that corresponded to almost perfect match between the data and the theoretical curve (published as supplementary data of Ser-Giacomi et al.). The procedures for fitting the distribution and assessing its statistical significance involves direct comparisons of the empirical and theoretical distributions, and not of the binned abundance plots. Subsequent analysis of the same data without taking this precaution has indeed led to identify a higher degree of variability among samples[94], which I believe is the consequence of the less accurate fitting protocol.

The results of the fit were quite surprising to us. First of all, the model was flexible enough to fit the vast majority - but not all - of the samples. The fit was overly dominated by the power-law behaviour, which extended over about 4 decades, making us confident that its detection was not an artifact of the data representation. Secondly, OTUs that were identified as belonging to the 'rare' component of the community constituted the near totality (more than 99%) of the OTUs present in each sample, indicating that the larger part of the OTUs composing the community obeys the same scaling everywhere in the ocean (Figure 5.3). We discovered that, contrary to our initial expectations, the exponent of the power law decay varied very little from sample to sample (less than 8% overall), and had a value (1.53 ± 0.08) statistically different from the exponent 1 predicted by models. Our exponent, on the other hand, seemed to be compatible with that identified for the time distribution of abundances in models with intermittent demography[40].

Given that plankton protist communities vary considerably from one place to the other both in composition and in abundance, it was highly unexpected to find that the vast majority of OTUs would follow quantitatively the same law,

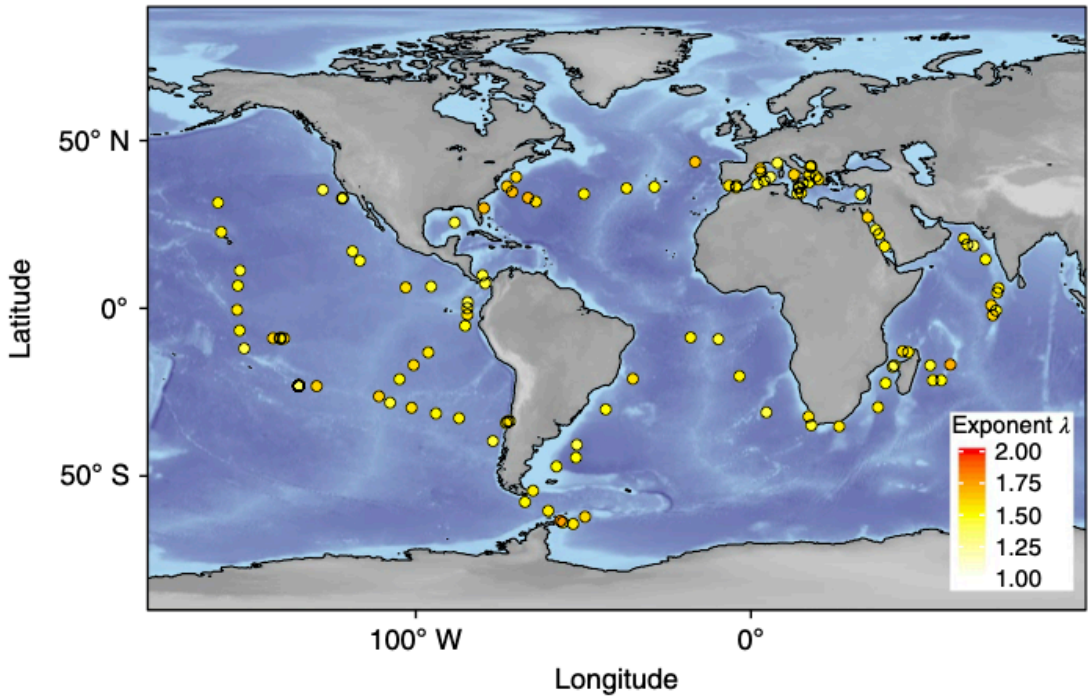

Figure 5.3: Exponent of the power-law decay of the 'rare' component of the community in the over 100 different sampling stations of the *Tara oceans* expedition, showing the very small extent of its variation despite the variability of oceanic regions, environmental conditions and biomes.

and with a precision that is hardly encountered in ecological patterns. A natural question that this observation poses is what are the ubiquitous ecological processes that produce such a pattern. A second, related question is what kind of information could be gathered by looking only at the small fraction of dominant OTUs – which nonetheless made up more than half of the total number of reads.

In order to address the latter question we looked at another ecological pattern: species turnover across communities. The amount of compositional change among communities is typically measured by the Jaccard or the Bray-Curtis indexes, which compare the types present based on their presence/absence or abundance, respectively. The higher those indexes, the more the communities are divergent, a feature that in macroorganismal communities typically reflects

geographical distance. By computing these metrics on our dataset, we realised that the correlation between the ecological distance computed with or without considering the tail of rare species was extremely high ($r > 0.98$). This indicates that dominant species contain the largest part of the biogeographical information on how communities change from place to place. A corollary of this observation is that deep sequencing of environmental samples should not substantially improve the characterisation of microbial biogeography, which could explain why low-resolution methods like flow cytometry and remote sensing work so well in identifying large-scale patterns. The weak relationship between the rare component of communities and environmentally-driven macroecological patterns was confirmed by examining correlations of the fitted parameters (notably the power-law exponent) with the metadata of the *Tara oceans* dataset[16]. These covered a number of environmental parameters, be they abiotic (e.g. latitude, SST, nutrient and oxigen concentration, transport-related indexes) or reflecting the biotic state of the ecosystem (e.g. total chlorophyll concentration). Perhaps unsurprisingly given the small variation of the exponent, no consistent correlation was found. This suggests that the ecological processes involved with the assemblage of the 'rare biosphere' occur on a larger spatiotemporal scale than the idiosyncratic conditions a given community experiences, while such conditions – on the contrary – determine what are the locally dominant species.

If the vast majority of OTUs does not play a role in establishing macroecological patterns, and is not strongly influenced by the local environmental context, then what processes are relevant to establish their abundance? This question is very much at the heart of the ongoing research in collaboration with Emil Mallemin (post-doc in co-direction with Arne Traulsen, MPI Evolutionary Biology), Giulio Biroli (Dept. of Physics of ENS) and other colleagues of a working group on ecological dynamics (Ada Altieri, MSC lab, University of Paris; Mathieu Barbier, CIRAD, Montpellier; Lucie Zinger, IBENS; Jules Fraboul, Dept. of Physics of ENS; Giulia Lorenzana, Dept. of Physics of ENS).

The functional form for the abundance distribution used in Ser-Giacomi et al.

was derived from a model that, like Hubble's neutral models, is species-symmetric. This means that species are assumed to be no different from one another, apart for a density dependence that is only important when abundances are very low. These assumptions seem to patently contradict what we know about microbial populations. On the one hand, their growth rate – which is the most common measure of fitness – can be widely different in a given environmental context, as evidenced by the fact that planktonic blooms are such massive events to be visible even from space. On the other hand, microbial populations are often so large that very seldom they would be reduced, by dilution or predation, to just a handful of individuals. Moreover, the exact origin of the nonlinear terms in the birth and death rates is unclear. They are often interpreted as reflecting immigration and emigration events, but the fact that these two opposite fluxes to and from a given environment should obey the same scaling is in my mind non trivial, especially when they are driven by passive transport with a water mass.

Our explanation for this model to work seemingly very well in fitting the empirical observations for plankton is that all cell populations in the ocean obey dynamical regimes that are effectively equivalent (something different could occur for instance in gut microbiome communities, that decay exponentially[76]). Every species would experience every now and then a blooming event, which is essential to increase drastically its abundance, but limited in time and space. Time series analysis indeed suggests that plankton communities commonly face an important OTU turnover[93], on top of repeatable seasonal patterns[45,50]. Once the bloom is over, species would disperse on large spatial scales due to mixing by ocean currents, as their fluid dynamical niche gets elongated into filaments[37]. Therefore, most of the time a species would experience environmental conditions that are not optimal for its growth. The sheer number of cells generated during a bloom (possibly combined with an Allee effect reducing predation on rare species, or a decrease in mortality due to viruses) would however allow species to hang around for a long time. In microbial ecology, the idea that most of the observed microbial diversity is due to species that are not actively growing is commonly called the seed-bank hypothesis[81,49]. Dynamically, the exis-

tence of a seed bank would correspond to the presence of alternating periods of rest and of growth, as simulated by a model with intermittent population dynamics that predict power-law abundance distributions in the time series of a single species[40]. If single-species models including a simple population dynamic and external perturbations seem to go a long way in accounting for patterns in microbial abundance (see also Grilli, where the population dynamics was subjected to external stochastic perturbations) and could be considered more realistic for describing large populations than stocastic birth-death models, it remains similarly difficult for these models to justify the complete neglect of all ecological interactions with other species.

A number of recent and ongoing works, including some I am pursuing, is addressing the degree to which the ecology of species-rich ecosystems as microbial ones is reproduced by models that do not appear to reflect the complexity and diversity of interactions. My working hypothesis is that there are regimes where the diversity of interactions, coupled with the high number of species, reduces the influence of the rest of the community to an effective term. In particular, I am interested in non-equilibrium regimes where the population dynamics is, by virtue of biological interactions or through environmental forcing, intermittent, as this seems to reflect more closely the observations on plankton communities.

A first step in this direction is the identification of a model adapted to describe the main features of species-rich communities without being too complicated for mathematical treatment. Many authors use the so-called generalised Lotka-Volterra equations to describe the continuous-time dynamics of the (continuous) abundances N_i of species $i \in \{1, S\}$, one formulation of which is the following:

$$\frac{\mathrm{d}N_i}{\mathrm{d}t} = \frac{N_i}{K_i}\left(K_i - N_i - \sum_{j \neq i} \alpha_{ij} N_j\right). \tag{5.1}$$

Here, the constant K_i represent the carrying capacity of species i in isolation and quantifies the weight of nonlinear effects of intra-specific competition on net growth rate. For simplicity, all species are commonly assumed to have the

same maximum growth rate, so that this parameter is rescaled out of the equations. The dynamics of every specie is influenced by the abundance of any other species j through the interaction strength $\alpha_{i,j}$. Although this model does not describe explicitly the mechanism of interaction, as do for instance models that account for exchanged resources (see Chapter 6), it can still be used as a tool to study how population dynamics is influenced by the existence of a multitude of weighted interactions. In a seminal paper, Robert May[97] studied the linear stability of an ecological equilibrium in the case when the interaction matrix $\alpha_{i,j}$ is a random matrix, concluding that diverse, species-rich ecosystems should not be found in a coexistence equilibrium. Many criticisms have been moved to the relevance of his conclusions for real ecosystems, and the debate is still going on. However, the idea of using random interactions as a null model to study the properties of large communities has given rise to a large body of work in the so-called 'disordered' approach. In a spectrum of models for ecosystem dynamics, these studies are positioned to the opposite limit of the other prominent – and arguably wrong – model for ecosystem dynamics, that predicts competitive exclusion based on generalisations of a two-species competitive Lotka-Volterra model[63]. The study of disordered generalised Lotka-Volterra equations witnessed a burst in recent years as methods from statistical physics started to be applied to investigate the dynamical regimes and statistical properties of complex ecosystems[80,14,7,3,32]. These studies thus offer the opportunity of using powerful analytical tools to disentangle the many different sources of complexity of models for species-rich communities.

In an ongoing study by Emil Mallemin in collaboration with Giulio Biroli and Arne Traulsen, that I mention in my Research Project, we decided to investigate the relative capacity of neutral and Lotka-Volterra model to capture statistical features of microbial communities, keeping particularly in mind plankton protists. Indeed, if these models are known to share many common features, it is still unclear what exactly should be measured in order to tell them apart. And if this is impossible, then what causes models that are based on radically different assumptions to be effectively equivalent? The aforementioned work

group on ecological dynamics has been extensively discussing these points, notably thanks to the work of Jules Fraboul, whom I co-advise with Giulio Biroli, and of Giulia Lorenzana, co-advised by Giulio Biroli and Ada Altieri, and a collective manuscript on the subject is in preparation.

Stepping aside from the dynamic complexity that can manifest in large ecosystems with random species interactions, the next chapter will deal with a related question. Indeed, a common and well-taken point of criticism of disordered approaches is that there is no reason (apart from the presumption of ignorance on the side of the modeller) that the interaction matrix should have a random form. In particular, we might expect that evolution of interactions among species would always modify its entries, and therefore break some of the regularities upon which rely analytical treatments. Competitive interactions are critical in this respect, as discussed in Chapters 2 and 3, since they underpin conflicts that can drastically reduce species diversity, exactly like the tragedy of the commons (the Lotka-Volterra and replicator equations are indeed mathematically equivalent[71]).

We will see that selection of a system of interacting species for its collective function can, given the ecological conditions that make the community a Darwinian individual, push interactions in a different direction than predicted by game-theoretical arguments, and moreover imprint a non-random structure on species-rich communities.

Supervisions:

Alice Soccodato (PhD 2014), co-supervised with Francesco d'Ovidio
Shruti Malviya (PhD 2018), supervised by Chris Bowler
Enrico Ser-Giacomi (Post-doc 2017)
Emil Mallemin (Post-doc 2021-2024), co-advised by Arne Traulsen

Selected invited presentations:

Invited speaker at the Workshop Microorganisms in Turbulent Flows, Leiden, The Netherlands (February 2016)
Invited speaker at the Meeting Regular Patterns in Biology: Causes and Consequences, Princeton, USA (April 2018)
Invited speaker at the Meeting on 'Biodiversity, Epigenetics and the Environment', Collège de France, Paris, France (April 2019)

Related publications:

d'Ovidio F.*, De Monte S.*, Alvain S., Dandoneau Y., Lévy M. (2010)
Fluid dynamical niches of phytoplankton types
PNAS 107: 18366.

De Monte S., Soccodato A., Alvain S. and d'Ovidio F. (2013)
Can we detect oceanic biodiversity hotspots from space?
ISME J 7(10): 2054–2056.

Soccodato A.*, d'Ovidio F*., Lévy M., Jahn O., Follows M.J., De Monte S. (2016)
Estimating planktonic diversity through spatial dominance patterns in a model ocean
Marine genomics 29: 9-17

Ser-Giacomi E., Zinger L.*, Malviya S.*, De Vargas C., Karsenti E., Bowler C., De Monte S. (2018)

Ubiquitous abundance distribution of non-dominant plankton across the world's ocean

Nature Ecology and Evolution, 2 :1243–1249.

Mark J. Costello et al. (2017)

Chapter *Methods for the Study of Marine Biodiversity*

in *The GEO Handbook on Biodiversity Observation Networks*, Springer.

... life is in fact a hierarchy of processes (e.g., metabolic, developmental, ecological, evolutionary) and any abstraction of an ontology of fixed entities must do some violence to this dynamic reality.

John Dupré & Maureen O'Malley

6

Evolution of heterogenous collectives

It is an evidence for everybody that collective functions evolve. Nonetheless, the mechanisms by which they do and the most appropriate level for describing their dynamics are still the subject of intense investigation. If one thinks of collective behaviour as the emergence of group-level properties, it is easy to conceive that variation in these properties will lead to a consequent variation in the collective behaviour. The role of self-organisation in revealing the collective effects of selection acting on individuals has been stressed as perhaps the most important in establishing the features of new levels of organisation[137]. The starting blocks here would be traits of solitary individual units that come to assume a meaning when such units hapen to be found within a collective. Particularly important in this perspective are processes that allow collectives to emerge as units, thus offering the context for individual-level traits to become meaningful[28,110]. Cells that possess certain properties in their soli-

tary living can for instance find themselves interacting with one another as they grow crowded in a limited space. This can happen as a passive consequence of some kind of externally imposed compartmentalisation, or derive from ecological conditions that are controlled by the cells themselves[9]. As I discussed in Chapter 3, key in defining the spatiotemporal extension of a collective is the life cycle it is embedded in, that determines two key aspects of collective living: what is the social context of cells embedded in a multicellular organism, and how does such multicellular phase reproduce. A lot of exciting work in experimental evolution has shown that it is relatively easy to harness functions of unicellular ancestors to evolve life cycles that bear great resemblance with the salient features of multicellular life cycles, including reproduction by single-cell bottlenecks and division of labour[113,62]. Mutations that prevent yeast cells to detach at division and that allow bacterial cells to switch their phenotype between sessile and planktonic forms are recruited when a selective pressure is imposed on the collective phase. Life cycles ensue as the temporal arrangement offered when the constraints imposed by these 'unicellular pathologies' are taken into account[111].

The collectives that derive from reprogramming cellular phenotypes end up being largely clonal, so that selection for collective properties boils down to selection on the genes that control those phenotypes. It is less obvious, instead, that a selective pressure on collective function would menage to act on the properties of heterogeneous collectives. The main reason for this is that groups formed over and again by distinct types would need to withstand intra-group competition, making it exceedingly easy for one type to exploit the others at its own advantage. Akin to the tragedy of the commons, such conflicts are expected to destabilize collectives on short (if different cells compete within the multicellular body, as for cancer) and long time scales. Despite the seemingly insurmontable challenge, interactions of diverse cellular types within groups are widespread and often so well integrated that organismal identity is attributed to the assembly - on top of the composing units. Symbiotic associations, where the component, co-localized cellular lineages are locked into coexistence by reciprocal dependencies that make life as an independent organism disadvantageous, are a

prime example of such stable, heterogenous collectives. Microbial communities in the environment, moreover, appear to be maintained over time in all their complexity, as discussed in Chapter 5, despite composing cells being not even physically attached. A growing number of researchers conceive such communities as units at a supracellular level of organisation, identified by the interactions of species with one another and with other organisms. Ecological communities are for instance viewed as 'superorganisms' whose coherence derives from spatial self-organisation[83] and whose diversity underpins ecosystem-level functions. Similarly, the microbiota is increasingly seen as an integrated component of larger collective assemblages, the 'meta-organism' or 'holobiont', that include also host species on other levels of biological organisation[6,122]. If the advantages produced by group living and cross-feeding are intuitively evident, the question of the evolution of interactions poses in different terms. First, how feasible is a breakdown of collective properties due to the invasion of 'cheater' types? Second, how does collective-level selective pressure translate into evolution of cell-level interactions, so as to produce collective adaptation? I will come back to the first question in my Research Project, and concentrate for the time being on the second, starting from a simple two-species community and moving on to multi-species communities.

In order for collectives to be considered as full-fledged Darwinian individuals that evolve by natural selection, they must satisfy some conditions on their spatiotemporal structure and coherence. The most famous phrasing of such conditions was proposed by Richard Lewontin and applies to any entity that is capable of reproduction[82]. Necessary for a population of these units to undergo evolutionary changes is that there are phenotypic variation, differential fitness and heritability of fitness. These three conditions are often summarised in the expression 'heritable variance in fitness'. Although this formalisation of Darwin's ideas was hugely influential in pointing out where to look for signatures of natural selection in action, it involves concepts that are problematic to apply to microbial collectives. First and foremost, it is not always clear what measure of

fitness should be adopted (as fitness is the aggregate result of different demographic processes[35]). This is particularly critical in situations where it is not even evident to discern reproduction events, such that parental collectives can be clearly assigned to offspring collectives (thus reproductive output evaluated). Secondly, and related to the previous point, it is even more difficult to conceive what exactly heritable fitness means. In population genetics, heritability is calculated as a population-level mean, and thus independently on the mechanisms that enact inheritance of phenotypic traits in a given situation.

In De Monte & Rainey, we have argued that Lewontin's conditions can be rephrased in less restrictive (or more operational) terms, that allow to address the action of natural selection in the proximity of mayor evolutionary transitions - first and foremost that to multicellularity - where assessing fitness may be problematic. We proposed that three generalised conditions are variation, recurrence and genealogy. All these properties can be defined, starting from the identity of their composing units, for any collective, including heterogeneous assemblies, aggregates that exchange only part of their units, and in general loose collections of reproducing units. Think for instance to the microbiota, a collection of cells of different identities, that is open to the environment but at the same time constrained in its spatiotemporal extension by the host, and as a consequence at least partially vertically transmitted. The advantage of extending the scope of natural selection is that it becomes clear that complete integration into a new level of organisation is not necessary for evolution to start acting – even though with little efficiency – on collective properties at the collective level. This point of view makes less mysterious the emergence of new levels of organisation, that was the missing link between the two formulations of multi-level selection theory (so-called MLS1 and MLS2). A second point this generalisation allowed to stress is the importance of life cycles in the recurrent structuring of collectives. As I argued in Chapter 3, the emergence of a collective time scale is essential in order to recognise collectives as such. There are many different ways this can occur, which opens the way to the existence of a variety of forms of collective genealogy (for instance, aggregative vs clonal mul-

ticelullarity).

Not all forms of genealogy are however equally conducive to support heritability of a collective character. Indeed, any collective property that relies on specific relations among units will be highly affected if the composition can be significantly altered by migration or horizontal transfer of parts of the collective. A particularly efficient way to harness selection on collective properties is to impose from the exterior a 'scaffold' that ensures collectives to be compartmentalised and genealogically arranged. Serial transfer protocols, such as those commonly implemented in experimental evolution, are an instance of such scaffolding. On its own, serial dilution through a single-unit bottleneck has been shown to be able to drive the evolution of unit-level properties that are commonly attributed to the existence of higher levels of selection, such as reduced growth rate at the lower level and increased decoupling between the reproductive efficiency of single units and collectives[9]. Moreover, it is expected to mimic the spatiotemporal arrangement that could derive from ecology in a metapopulation. But what will be the evolutionary dynamics of a heterogeneous ecological community scaffolded by serial transfer?

The difficulty of describing such process stems from the need to take into account two levels at the same time. Indeed, birth and death occur both for units (on the ecological time scale between tranfers, or within a 'collective generation'), and for collectives (on the time scale of collective generations). Figure 6.1 represents graphically a typical serial transfer selection protocol, where communities composed by units (individuals) endowed with a proper ecology (based on intra- and inter-specific interactions) get assessed at the end of an externally imposed period. This time scale defines the extent of a 'collective generation' and determines when the processes of collective death (due to selection) and birth (due to transfer, possibly with migration) take place. If it is clear that within-collective ecology determines – perhaps in complex ways – the phenotypes of the collective that are accessible to selection, it is far less evident how selection for such emergent properties might drive the evolution of the composing units. For instance, as argued in Chang et al. and De Monte, collective-level properties

could change simply due to the rearrangement in the ecological structure of the community – for instance, by different combinations of species out of a regional pool, or because of differences in initial composition – without there being any change in the individual-level parameters.

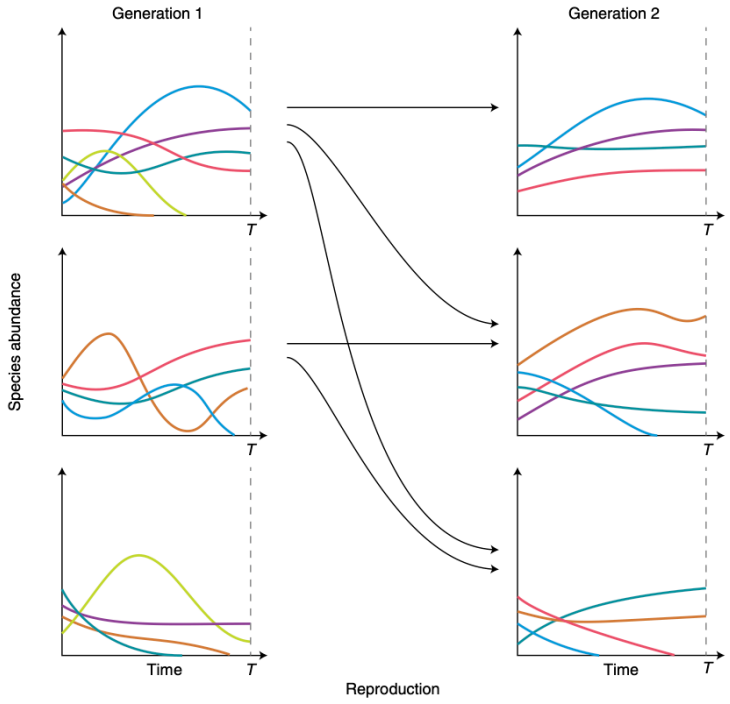

Figure 6.1: Schematic representation of a process of serial transfer of a complex community. A population of communities is initialised by sampling from a same regional pool of species, giving rise to communities that differ in composition both qualitatively (the type of species) and quantitatively (the number of individuals for each species). After the duration T of a collective generation, communities are simultaneously assessed for some property (say total biomass), and only the communities with highest score get propagated. A new generation of collectives is started by sampling from one or more of the selected communities (there are many different ways to implement this, as discussed in Chang et al.). By itself, this process can produce changes in the collective function that is under selection, and this also in the absence of any change in the parameters that underpin the ecology of the community, although open-ended evolution is not possible in this system unless evolution is allowed to occur also at the scale of the individual interactions. Image reproduced from De Monte.

That selection applied to the collective level should be able to steer lower-level traits towards those that produce the target outcome seems rather intu-

itive, but the big question is under what conditions, and with what efficiency. A long debate on the relevance of group selection in shaping behaviour within groups has raged for decades in evolutionary biology, and was 'solved' by the proposal that individual-level and group selection are equivalent when one is able to correctly account for population structure in computing inclusive fitness. Apart from its philosophical implications, this statement however does not help in assessing when selection operating on collectives can yield the desired results. Letting aside more abstract considerations, a number of studies have addressed the action of group-level selection in specific contexts, both in experimental and modelled populations. These efforts have been in particular focused on microbial communities because of their fast evolution, the ease in harvesting them, and the potentially high impact of collective-level functions of microbial assemblages, for instance in applications to bioremediation or production of specific compounds[85,4]. As pointed out in Sánchez et al., despite the high expectations based on the previous intuitive arguments, artificial selection of microbial communities witnessed a variable and in no case spectacular success, calling for a deeper understanding of the eco-evolutionary processes that drive observed community-level changes.

A first way to get useful insights on the ecological processes guiding community evolution is to focus on the simplest possible assemblies, those composed of only two different species. Several theoretical works have addressed this situation in the recent years, exploring under what conditions evolution could produce a collective outcome that is costly for the individual units[145,139,144]. A key parameter in determining the efficiency of selection turned out to be the degree of heritability of community composition. Indeed, although the existence of a collective genealogy produces material overlap of collectives across generations (in terms of the composing units), heritability can be limited both by lateral exchanges[139] and by ecological constraints[144]. Clearly, if the collective phenotype is not heritable, natural selection cannot operate at the collective level.

During the PhD thesis of Guilhem Doulcier (co-directed by Paul Rainey and Amaury Lambert), we have addressed whether and how heredity of a collective phenotype can evolve in a two-species community (illustrated in Fig. 6.2).

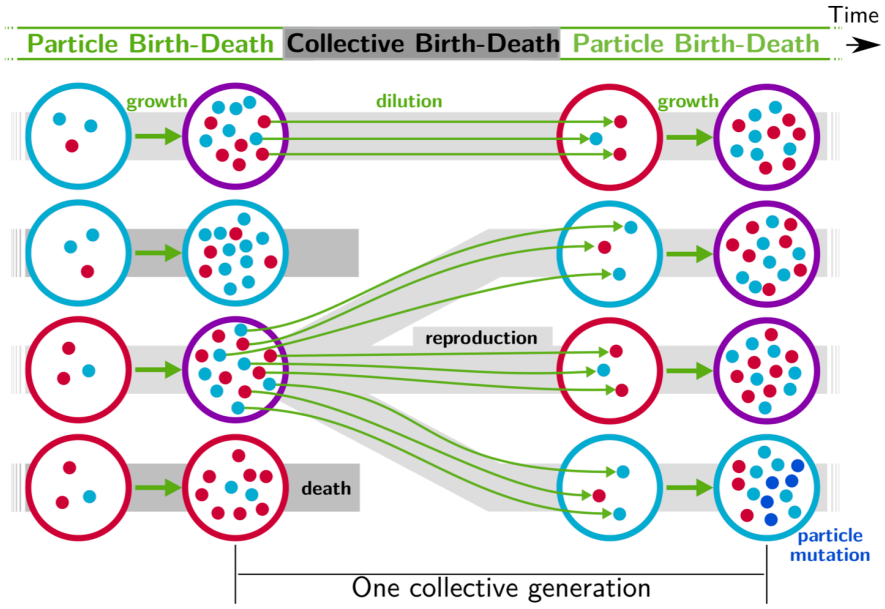

Figure 6.2: Schematic description of the model for the evolution of a two-species community under selection for balanced community composition. Each species is made up of at most two populations that carry the same colour: a resident and a mutant type. The dynamics of the model can be looked at on two time scales. Within the duration T of a collective generation, the composition changes continuously in time $t \in 0, T$ due to the ecological interactions (fixed in the newborn collective based on succes in the parental collective and mutation). Between collective generations, the community composition (quantified by collective colour) follows a discrete-time dynamics. Selection acting at the 'adult' stage ($t = T$) keeps only a fixed proportion of collectives that are closest to the target composition. Newborn collective of the following collective generation (in a population of fixed size) is seeded by sampling B particles from one of the selected collectives. The bottleneck size B determines the amount of stochastic variation in the composition of newborn collectives with respect to that of their parents.

Initially, collectives were composed of two types of units or 'particles', marked by different colours (red and blue). Each type was attributed a set of six (different) parameters determining its birth and death rates on the basis of density-dependent intra- and inter-specific interactions in a well-mixed environment. Notably, inter-specific parameters were chosen to be competitive, and maximal

growth rates different. At the end of a collective generation of duration T, a sample of B particules was propagated in order to found a new generation of collectives, mimicking the dilution operated in serial transfer protocols. Newborn collectives were then allowed to 'grow' to an adult stage by deploying their ecological dynamics for the duration T. The collective phenotype we were interested in was the composition of the adult community. Selection was let operate at the collective level by discarding a fraction of communities whose composition was farthest away from the target balanced one (associated to a purple colour of the community).

In the absence of selection, the ecological dynamics within collectives favoured one of the two types over and again. For short enough generation times, the fastest-growing type would for instance prevail. After a handful of generations, the population of collectives was composed by monocromatic communities, indicating that the collective function is lost if collective-level selection is neutral, as expected. Application of selection at the collective level, on the other hand, allows to maintain a population of purple collectives through a mechanism that was called 'stochastic corrector' by Grey et al.. Stochastic correction relies on fluctuations in initial composition due to sampling that, given a large enough population of collectives and a small enough bottleneck, will end up producing some – maybe just a few – 'good' collectives (good in the sense that they will have, at the adult stage, a colour close to the target). Not only this mechanism has a low yield, but it does not allow collective function to be conserved outside the selection treatment: as soon as the selective pressure is discontinued, the system is back to the neutral dynamics, and purple collectives go extinct. If this was all that artificial selection had to offer, it would not be surprising that it is of scarce practical use.

If particles are allowed to evolve, however, the picture changes completely, albeit differences emerge on a much longer time scale. Particle-level parameters change progressively in a highly predictable way. First, within-collective selection drives an increase in maximal growth rates, that tend to align with one another (if they are exactly identical, collectives will maintain a constant colour

if collective generations do not last too long). But as faster and faster growth is selected, density-dependent interactions grow in importance, and thus start to be 'seen' by natural selection thanks to their phenotypic effect. In the long run, interactions among particles will be selected, such that the two populations attain at time T the same population size. Notice that this does not exclude large excursions in colour within a collective generation, where cells modify their growth rates in time, akin to a developmental cycle. We called 'developmental corrector' the mechanism that, through selection of specific ecological parameters, ensures that adult collective phenotype is reliably inherited despite the possible fluctuations in the newborn community composition, and gives rise to a population of collectives that are all close to the optimal colour. Such increased collective inheritance – obtained through canalisation or convergence towards a collective state – is manifest when selection is discontinued. Indeed, the average colour of the population is very well maintained in the population of collectives on the time scale of hundreds of collective generations. It changes very slowly (on evolutionary time scales) only due to neutral drift of the underlying particle-level parameters.

We used a continuous approximation of the within-collective ecological dynamics (the two-species Lotka-Volterra competitive equations) and a hypothesis of time-scale separation (new mutations occur once a previous mutant has invaded the population of particles of its colour) to make sense of the previous observations, which were obtained through a stochastic, individual-based simulation. This semi-analytical approach allowed us to point out the key role of the 'growth function' G that maps the newborn composition into the adult composition. Analogous to the Waddington's epigenetic landscape, this function depends on particle-level parameters (in developmental terms, these could be set by the genes) and determines which final state collective colour will converge to. The evolutionary stable equilibrium is simply the stable fixed point of the map $x(\tau + 1) = G(x(\tau))$, and the speed of convergence to this equilibrium is determined by the derivative of G. A study of the bifurcations that are possible give this function allowed us to understand how different system parameters,

most importantly those that can be controlled by the experimenter (see Fig. 6.3), affect the qualitative outcome of collective-level selection. The most notable conclusion is that growth functions that are associated to high inheritance of the collective phenotype are always selected, though with different efficiency depending on the parameters. Typically, evolution of interactions is facilitated by a decoupling in the (faster) individual particle growth rate from the (slower) duration of a collective generation. If instead of evolving freely, the ecological parameters obey some constraint (for instance, particle growth rate is bounded), high inheritance evolves nonetheless, though along different evolutionary trajectories.

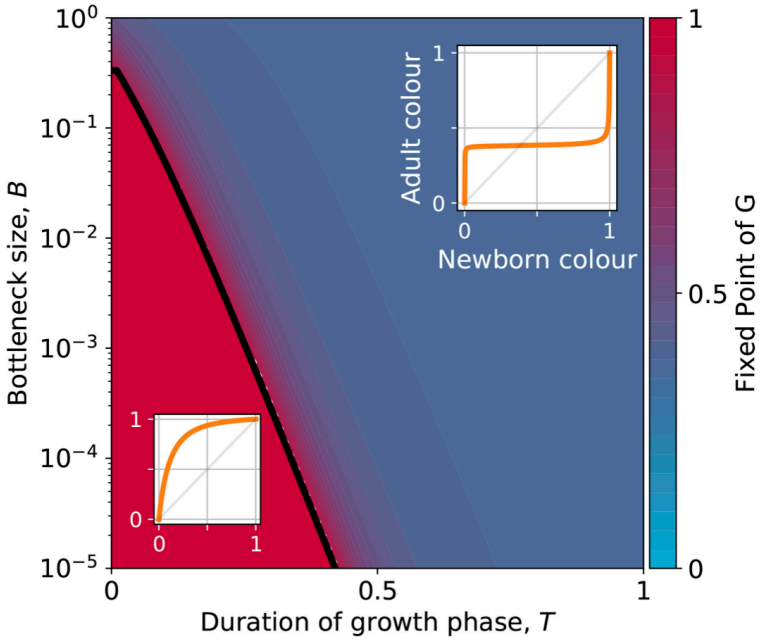

Figure 6.3: Dependence of the fixed point of the growth function G (adult collective colour) on the collective-level parameters B (bottleneck size at reproduction) and T (duration of a collective generation). For any given set of particle-level parameters, macroscopic parameters such that interactions can evolve before selection drives collectives to extinction are those whereby newborn collectives have sufficiently large populations, and grow for a sufficiently long time. Once these conditions are met, the function G will change over evolutionary times so that its value for a specific choice of B and T matches the target colour. The insets present qualitatively different shapes of the growth function in the regions where the evolutionary stable equilibrium is monochromatic (bottom left) and purple (top right).

Important to the study I will discuss next is the observation that the evolution of interactions modifies the ecological relationship between species. If interspecific interactions are constrained to be only competitive, then they generically evolve towards a state of maximal asymmetry: in the course of a collective generation, one species grows freely, and the other (that with weakest intraspecific interactions) is slaved to the dynamics of the former. Asymmetry is common in long-established symbiotic interactions, such as for instance between organells and the eukaryotic cell. Our model suggest it may be the inevitable outcome of optimising the robustness of the collective developmental process by concentrating its control on only one of the composing partners. Our study stresses moreover that the need to reach a determined state at the moment of reproduction does not require this state to be stable for the whole duration of a collective association, thus opening the way to the deployment of different developmental paths. Finally, ecological interactions may not be bound to remain qualitatively the same: competitive interactions may turn mutualistic in the course of time. If this is allowed to occur, evolution for collective function drives the system towards increasingly mutualistic states, until the Lotka-Volterra system (that is not bounded by higher nonlinearities) diverges. Despite such unrealistic outcome, the observation that selection at the collective level might prevent strongly competitive species to take over the entire community aligns with the group selection view of collective evolution and with the idea that costly collective behaviour can be favoured[145,139].

Let me now turn to more complex communities. If ecosystems composed of a handful of species have been essential for understanding the ecological dynamics induced by interactions between and among species and how these take effect, they constitute extreme simplifications of real communities. In particular, microbial communities are commonly composed of hundreds or thousands of co-existing different OTUs, many of which can only be characterised by their genomic signature. It is therefore unclear what null model should these communities be compared to. Competitive exclusion clearly does not appear to be a

particularly relevant scenario of reference, despite it being still invoked as the expected outcome of ecological interactions, for instance in the so-called "paradox of the plankton"[73]. The contradiction between the prediction of simple competitive models and the observation of a large number of co-existing species led to the formulation of alternative null models based on a radically different assumption, where all species are considered identical with respect to their fitness. The most famous formulation of such species-symmetric models is Stephen Hubbell's neutral model, that explains ecological patterns as consequences of individual processes of birth, death, mutation and dispersal[72]. Assuming species equivalence allows to draw conclusions on the distribution of species abundances for communities with a very large number of species. The comparison of these predictions with empirical distributions in general fare pretty well, but it is still not completely clear why would the assumption of neutrality hold in any real system (I discussed this point also in Chapter 5). An alternative starting point is to consider that species demographic parameters are influenced by the presence of other species in the community, therefore fitnesses depend on community composition. In particular, as suggests the previously discussed two-species model, the existence of interactions seems to be key in allowing collective function to evolve by collective-level selection. A simple way to model communities with a large number of species that are not equivalent in terms of their interactions is to generalise two-species Lotka-Volterra equations to multiple species. In equation 5.1 introduced in Chapter 5, interactions boil down to density-dependent terms weighted by constant coefficients $\alpha_{i,j}$, whose values are constant for any focal pair of species. Though many criticisms have been moved against the realism of such a formulation (including the impossibility to neglect resources and higher-order interactions), this system offers the opportunity to further analytical investigations and to make use of exact results (like, in force of their structural equivalence, with the replicator equation used in evolutionary game theory[71]).

With Jules Fraboul (whose PhD is ongoing under the joint direction by Giulio Biroli and myself), we have therefore decided to explore the evolutionary dy-

namics of species-rich communities described by generalised, disordered Lotka-Volterra equations. Choosing random entries for the interaction coefficients corresponded to constraining them as little as possible. If on the one hand such lack of constrains is certainly not to be expected in real communities, it allows to explore more generally how collective-level selection shapes the statistical properties of the interactions in complex heterogeneous communities. Our model has a nested structure similar to what previously discussed for the two-species consortium, with two notable differences in the way reproduction with mutation is implemented. First, we wanted to avoid as much as possible transient ecological regimes, for both theoretical and practical reasons. Indeed, as shown by numerical simulation in Chang et al. and commented in De Monte, when selection is applied on a feature that reflects a transient ecological arrangement of species, evolutionary outcomes can be suboptimal and are in general much harder to foresee. Moreover, if the community is always close to a steady state, perturbation methods can be applied to obtain both an analytical description and numerically efficient simulations. The second difference consists in the way mutations of species-level parameters are implemented. In principle, mutations could occur any time during the duration of a collective generation, in a process that we do not want to describe explicitly. When looking at the composition of an adult community, therefore after a growth phase of duration T, one might expect to face a spectrum of mutants belonging to different species. We therefore considered that, as far as their effect on collective evolution is concerned, mutations impact the whole interactions matrix and not just one entry at the time. This choice reflects the spirit of collective evolution being based on collective-level mutations – that thus transcend the identity of individual species. We wanted however that such mutations did not impose an a priori bias on the collective phenotype. This invariance in expectation of the mutated interactions is achieved when mutations are chosen as appropriate Gaussian perturbations of the interaction matrix. The evolutionary dynamics of the community can thus be conceived as a time-discrete transition between equilibrium states associated to the interaction matrices that get successively selected. If the

mutation process maintains the statistical structure of the interaction matrix, selection applied to the collective level singles out specific realisations of the mutated matrices, and in this way gets potentially imprinted onto their structure.

We concentrated on selection for total community abundance (the study of the response to other choices for the selective optimum is under way). This choice reflects, on the one side, artificial selection for microbial communities of higher productivity, that is deemed to be of primary importance in ensuring community stability[77]. On the other side, contrary to the previously discussed two-species model, it does not prescribe a specific target community composition, thus leaving the door open to multiple solutions that are selectively equivalent, for instance when biomass is distributed over all species and when it is concentrated in just a few. Finally, the dependence of total abundance on the interactions' statistics is known for random matrices[14], allowing us to foresee in which direction selection should push the initial community.

As it has been observed in previous numerical studies[142,17], selection imposed at the collective level always improves the target function. If interactions remained perfectly random along the evolutionary trajectory, it is expected that an increase in total abundance would be associated to a decrease in the average interaction strength and an increase in interaction variance. Numerical simulations reveal that the evolutionary trajectory satisfies these qualitative expectations: as the total abundance of the community increases (eventually diverging in finite time), interactions become more and more mutualistic (whereas they were, at the beginning, chiefly competitive). Quantitatively, however, their statistics deviate more and more from the prediction of the disordered approach, indicating that some structure emerges in the interaction matrix. Figure 6.4 illustrates the change that the interaction matrix undergoes along an evolutionary trajectory, indicating that the community gets restructured so that species with more permissive initial intra-specific interactions tend to become, in the end, more mutualistic.

The change in inter-specific interactions is connected to the emergence of an isolated negative eigenvalue (Fig. 6.5) of the interaction matrix, that remains

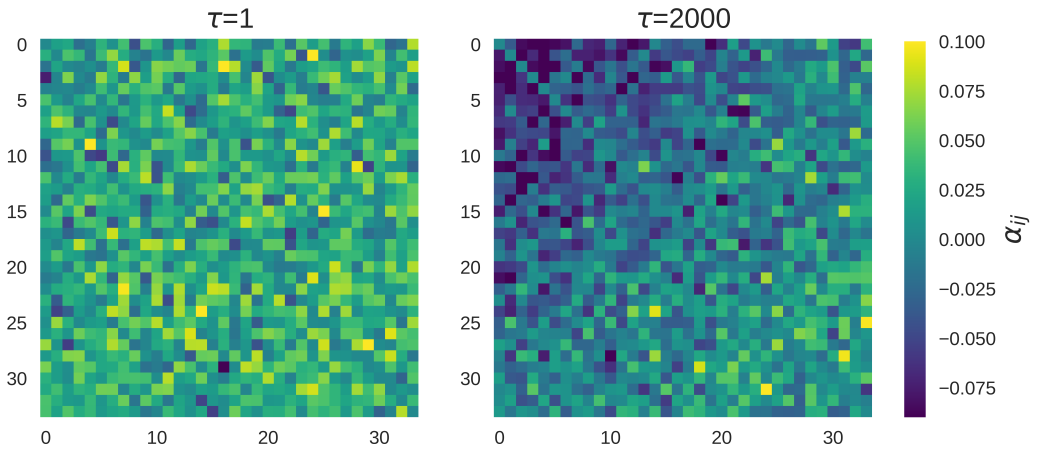

Figure 6.4: Ancestral and evolved interaction matrices in a numerical simulation of collective-level selection for high total abundance in communities described by the generalised Lotka-Volterra model (equation 5.1). Displayed are only the species that are not extinct at collective generation 2000. The initial matrix has randomly assigned entries (with zero symmetric correlation), and correspondingly no pattern is visible, set aside the diagonal that is null by definition (intra-specific interactions are taken into account by the vector K_i of the carrying capacities). The entries of the evolved matrix have smaller values, indicating that interactions have become more mutualistic. Moreover, its structure is no longer purely random, as evidenced particularly when species are ranked by decreasing carrying capacity: species that become more mutualistic tend to be those that, by virtue of their largest carrying capacity, are expected to contribute the most to collective function. The correlation is however not perfect, and our analytical approach allowed us to identify the origin of such correlations in the interplay, in structuring inter-specific interactions, of the target function, the vector of carrying capacities, and the linear stability of the ecological equilibrium.

otherwise disordered. This property allowed us to describe the evolution of interactions as successive addition of order-one perturbations to a disordered background. Interestingly, the eigenvector associated to the isolated eigenvalue, which is very strongly correlated to the species abundance distribution, can vary a lot among realisations of the evolutionary trajectory. This indicates that, if the statistical impact of selection on the interactions is predictable, it might be extremely difficult to foresee exactly what solution the community will find. Even a perfect knowledge of the initial community would not be sufficient to determine which community composition will eventually be selected.

An analytical approach allowed us to understand the origin of the isolated

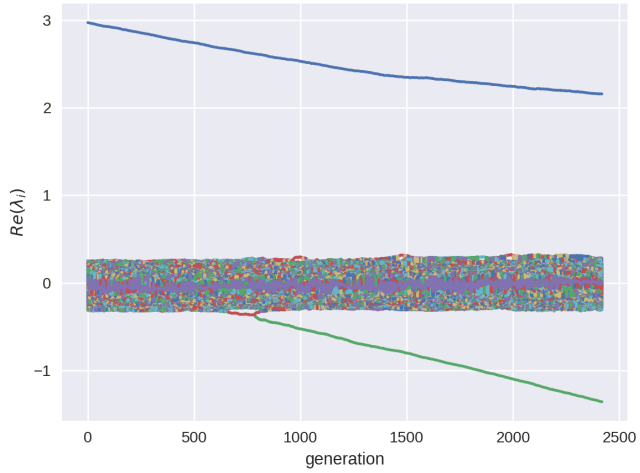

Figure 6.5: Real part of the eigenvalue spectrum along a simulated evolutionary trajectory. Selection for increased total community abundance drives the inter-species interactions to become more mutualistic (and diverse). This manifests in an order-one perturbation of the interactions matrix, evident in the emergence of an isolated negative eigenvalue. The dynamics of the divergence and its scaling with the system parameters can be reproduced by an approximation of the equations for the evolution of the interaction matrix, obtained when its variance is small.

eigenvalue and its relation with the type and strength of the selective pressure, the number of species in the community, the amplitude of the mutational step and the number of collectives subjected to selection. Although we were able to solve exactly the discrete-time stochastic dynamical system only in a limit case, the validity of this description was demonstrated by the correct prediction of how the speed of evolutionary change scales with the system's parameters. For instance, it turns out that massively increasing the number of communities that get assessed in parallel – as promised by new millifluidic technologies that allow to scan thousands of milliliter-scale microbial reactors – is not expected to provide an enormous gain in terms of selection efficiency, at least as long as mutations remain normally distributed. The formula for the evolution of the interaction matrix quantifies how the equilibrium abundances and their elasticity with respect to changes in the carrying capacity interplay with the target of selection to determine what kind of evolutionary change is expected to occur more promptly.

There is much more to explore on models of this kind, both making the ecology more realistic, and pursuing the theoretical explanation of their general properties. Particularly interesting for me is the perspective of comparing the predictions with experimental observations, both for environmental communities and experimental ones. For instance, one could imagine that microbial communities subjected to different ecological regimes, thus diffrent 'scaffolds' should reflect more or less closely the predictions of the model. Inferring from data the interaction coefficients of microbial communities with a large number of species is the object of ongoing active research, so that I expect methods to improve quickly in the coming years. Alternatively, interactions can be directly estimated by a high-throughput screening of pair-wise or multi-species competition experiments. Recent huge efforts have been put in realising these measures, suggesting that communities of the gut microbiome seem to display a prevalence of mutualistic interactions[138], as would be also expected in artificially evolved communities. On the other hand, scaffolding in a structure that ensures the prevalence of vertical transmission is less likely to be expected for marine plankton communities. Although direct measures are in this case impossible due to the prevalence of unculturable species, we can exploit the predicted species abundance distribution for comparison with the observations presented in Chapter 5. These two distributions do not qualitatively match, as the empirical distributions possess a heavy tail of rare species, whereas the model predicts truncated Gaussian distributions, that decay exponentially. Whether this reflects the weakness of community-level selection or a different effect of mutations on the variation of the ecosystem properties, is still to be established. Recent models for the metabolic activity in natural microbial communities suggest that they organize along a gradient of resource usage, from highly competitive to highly cross-feeding communities, and that these features reflect on the stability of the community to environmental disturbance and invasion by new species[89]. It would be interesting to explore whether this corresponds to different degrees of collective-level selection. Some guidelines on how I intend to develop these ideas are given in my Research Project.

Supervisions:

Guilhem Doulcier (PhD 2019), co-supervised with Paul Rainey and Amaury Lambert

Jules Fraboul (PhD ongoing), co-supervised with Giulio Biroli

Selected invited presentations:

Invited speaker at the conference Ecologie et Biologie Evolutive, Modèles Déterministes et Stochastiques, Toulouse, France (October 2017)
Invited speaker at the Workshop Mathematical Approaches to Cell-Cell Communication and Collective Behaviours, Banff, Canada (July 2018)
Discussion leader at the GRC Microbial Population Biology on 'The Structure, Ecology and Evolution of Interactions Within Microbial Populations', Andover, New Hampshire, USA (July 2019)
Invited speaker at the Workshop Evolution of Interacting Populations, Plön, Germany (September 2019)
Keynote speaker at the conference Mathematical Models in Ecology and Evolution conference, Reading, UK (July 2022)

Related publications:

Doulcier, G., Lambert, A., De Monte, S.*, Rainey, P. B.* (2020)
Eco-evolutionary dynamics of nested Darwinian populations and the emergence of community-level heredity
eLife, 9:e53433

Fraboul J., Biroli. G., De Monte, S. (2021)
Artificial selection of communities drives the emergence of structured interactions
bioRxiv

7

Conclusions and perspectives

What we call a biological population is characteristically composed of units that are homogenous enough to be recognised as belonging to the same equivalence class (lineage, species, organism), but different enough to support evolution by natural selection. The tension between uniformity and diversity lies at the core of mathematical representations of biological systems across scales. On the one hand, focusing on the shared features of the units composing a collective allows to more easily connect properties of lower and higher levels of organisation. On the other hand, if diversity among units has in many cases a constructive role (complementarity and redundancy increase a system's robustness and scope of action) it can also generate disruptive conflicts.

The decision of what diversity to take into account, and what to neglect or to describe by effective terms, is in my view strictly coupled to the dynamics

of any biological system. A unit at a given level of biological organisation has a characteristic extension in space, but also in time. Addressing the changes that occur at that level requires that its boundaries are sufficiently well established. If one thinks naturally to a finite spatial extension, I think that temporal boundedness is at least as important. Such temporal limitation is provided by the existence of 'cycles' such that the unit occurs over and again (together with a number of other units of the same kind). It is precisely re-occurrence that allows to detect evolution at a given level of organisation, be it a cell or a multicellular body or an ecosystem. One can for instance compare the size of cells before they divide, the dimension of bones at adulthood, and a climax community. Their variation on long time scales may provide indications on the selective processes that drive their adaptation, but often relies on synchronic observations rather than on a diachronic charcaterisation of the whole life cycle.

Form this point of view, it is clear that our capacity of detecting evolution in microbial collectives hinges upon the possibility to determine at what point in time they should be compared. The models I have discussed in Chapter 6 showed that the choice of the moment when selection is applied is critical to community evolution, and that transient ecological dynamics can be as important as the asymptotic regimes in determining its course. Similarly, in Chapter 4 I have argued that the evolution of sociality in aggregative organisms can be biased by altering the moment when collective development begins. When we start to leave the well-regulated structure of paradigmatic life cycles, it becomes essential to know how and when collectives start to be recognized as such, and what is their internal dynamics. In other words, the coupling of ecology and evolution is in microbial communities as tight as development is coupled to multicellular adaptation. In loose collections of cells, however, the problem that genes are not the only determinants of collective properties displays much more acutely than in metazoans. The influence of the environment, be it endogenous to the collective or exogenously imposed, reduces and constrains the degree to which relevant properties are remembered along the evolutionary trajectory.

Among the many questions that such increased complexity entails, I am par-

ticularly drawn to inquire how lower-level phenotypes and time scales become integrated in those of the higher level. How does the system remember its identity? When can lower-level temporal variability be 'folded up' into spatial diversity within the higher level? What changes matter and what do not in the course of evolution?

In the research project that is provided as a separate document I illustrate some directions in which I intend to pursue my studies in order to contribute to answering those questions. Beyond the choice of specific subjects, that are given there in some detail, I would like to conclude this memoir by a reflexion on the method I adopt in my research. Indeed, from the picture that emerges from the curent document it perhaps transpires that I have often found my approach sitting uncomfortably at the boundary of several disciplines. I am often wondering if the question I am asking requires a step back from disciplinary details, or I am just totally missing the point that someone more qualified than me would get. It was François Taddei who first drew my attention to the inevitable trade-off that interdisciplinary research has to navigate, that between knowing everything of nothing and knowing nothing of everything. This was certainly a well-taken point, but was not enough to discourage me to let curiosity be my guide – something I try to instil to the young researchers I interact with. I hope that this memoir reveals the thread that was behind the variety of different themes I addressed in my works. As much as I think interdisciplinarity is my cup of tea (also if it is not necessarily the most efficient way to proceed through a career), I am also fully aware that it is only the interaction with colleagues who had more specific and excellent expertise that allowed me to connect dots. And the diversity of experiences, temperaments, and personal choices is what I consider the single most important feature of scientific – just as well as microbial – collectives.

References

[1] Ackermann, M., Stecher, B., Freed, N. E., Songhet, P., Hardt, W.-D., & Doebeli, M. (2008). Self-destructive cooperation mediated by phenotypic noise. *Nature*, 454(7207), 987–990.

[2] Adiba, S., Forget, M., & Monte, S. D. (2022). *Evolving social behaviour through selection of single-cell adhesion in Dictyostelium discoideum*. Technical report, bioRxiv. Section: New Results Type: article.

[3] Altieri, A., Roy, F., Cammarota, C., & Biroli, G. (2021). Properties of Equilibria and Glassy Phases of the Random Lotka-Volterra Model with Demographic Noise. *Physical Review Letters*, 126(25), 258301. Publisher: American Physical Society.

[4] Arias-Sánchez, F. I., Vessman, B., & Mitri, S. (2019). Artificially selecting microbial communities: If we can breed dogs, why not microbiomes? *PLOS Biology*, 17(8), e3000356.

[5] Azaele, S., Suweis, S., Grilli, J., Volkov, I., Banavar, J. R., & Maritan, A. (2016). Statistical mechanics of ecological systems: Neutral theory and beyond. *Reviews of Modern Physics*, 88(3), 035003. Publisher: American Physical Society.

[6] Bang, C., Dagan, T., Deines, P., Dubilier, N., Duschl, W. J., Fraune, S., Hentschel, U., Hirt, H., Hülter, N., Lachnit, T., Picazo, D., Pita, L., Pogoreutz, C., Rädecker, N., Saad, M. M., Schmitz, R. A., Schulenburg, H., Voolstra, C. R., Weiland-Bräuer, N., Ziegler, M., & Bosch, T. C. G. (2018). Metaorganisms in extreme environments: do microbes play a role in organismal adaptation? *Zoology*, 127, 1–19.

[7] Barbier, M., Arnoldi, J.-F., Bunin, G., & Loreau, M. (2018). Generic assembly patterns in complex ecological communities. *Proceedings of the*

National Academy of Sciences, 115(9), 2156–2161. tex.ids: barbierGenericAssemblyPatterns2018a, barbierSupportingInformationSI Publisher: National Academy of Sciences Section: Biological Sciences.

[8] Beaumont, H. J. E., Gallie, J., Kost, C., Ferguson, G. C., & Rainey, P. B. (2009). Experimental evolution of bet hedging. *Nature*, 462(7269), 90–93. Number: 7269 Publisher: Nature Publishing Group.

[9] Black, A. J., Bourrat, P., & Rainey, P. B. (2020). Ecological scaffolding and the evolution of individuality. *Nature Ecology & Evolution*, 4(3), 426–436. Number: 3 Publisher: Nature Publishing Group.

[10] Bork, P., Bowler, C., de Vargas, C., Gorsky, G., Karsenti, E., & Wincker, P. (2015). Tara Oceans studies plankton at planetary scale. *Science*, 348, 873.

[11] Bourrat, P. (2015). Levels of Selection Are Artefacts of Different Fitness Temporal Measures. *Ratio*, 28(1), 40–50.

[12] Brown, J. H. (1999). Macroecology: Progress and Prospect. *Oikos*, 87(1), 3–14. Publisher: [Nordic Society Oikos, Wiley].

[13] Brunet, T. & King, N. (2017). The Origin of Animal Multicellularity and Cell Differentiation. *Developmental Cell*, 43(2), 124–140. Publisher: Elsevier.

[14] Bunin, G. (2017). Ecological communities with Lotka-Volterra dynamics. *Physical Review E*, 95(4), 042414.

[15] Buttery, N. J., Rozen, D. E., Wolf, J. B., & Thompson, C. R. (2009). Quantification of Social Behavior in D. discoideum Reveals Complex Fixed and Facultative Strategies. *Current Biology*, 19(16), 1373–1377.

[16] Chaffron, S., Guidi, L., d'Ovidio, F., Speich, S., Audic, S., De Monte, S., Iudicone, D., Picheral, M., Pesant, S., Tara Oceans Consortium, C., & Tara Oceans Expedition, P. (2014). Environmental context of selected samples from the Tara Oceans Expedition (2009-2013). Publisher: PANGAEA Type: dataset.

[17] Chang, C.-Y., Vila, J. C. C., Bender, M., Li, R., Mankowski, M. C., Bassette, M., Borden, J., Golfier, S., Sanchez, P. G. L., Waymack,

R., Zhu, X., Diaz-Colunga, J., Estrela, S., Rebolleda-Gomez, M., & Sanchez, A. (2021). Engineering complex communities by directed evolution. *Nature Ecology & Evolution*, 5(7), 1011–1023. Bandiera_abtest: a Cg_type: Nature Research Journals Number: 7 Primary_atype: Research Publisher: Nature Publishing Group Subject_term: Community ecology;Evolutionary theory Subject_term_id: community-ecology;evolutionary-theory.

[18] Clauset, A., Shalizi, C. R., & Newman, M. E. J. (2009). Power-Law Distributions in Empirical Data. *SIAM Review*.

[19] Costello, M. J., Basher, Z., McLeod, L., Asaad, I., Claus, S., Vandepitte, L., Yasuhara, M., Gislason, H., Edwards, M., Appeltans, W., Enevoldsen, H., Edgar, G. J., Miloslavich, P., De Monte, S., Pinto, I. S., Obura, D., & Bates, A. E. (2017). Methods for the Study of Marine Biodiversity. In M. Walters & R. J. Scholes (Eds.), *The GEO Handbook on Biodiversity Observation Networks* (pp. 129–163). Cham: Springer International Publishing.

[20] Danø, S., Hynne, F., Monte, S. D., d'Ovidio, F., Sørensen, P. G., & Westerhoff, H. (2002). Synchronization of glycolytic oscillations in a yeast cell population. *Faraday Discussions*, 120(0), 261–275. Publisher: The Royal Society of Chemistry.

[21] De Monte, S. (2021). Ecological recipes for selecting community function. *Nature Ecology & Evolution*, 5(7), 894–895. Bandiera_abtest: a Cg_type: Nature Research Journals Number: 7 Primary_atype: News & Views Publisher: Nature Publishing Group Subject_term: Evolutionary theory;Experimental evolution;Microbial ecology Subject_term_id: evolutionary-theory;experimental-evolution;microbial-ecology.

[22] De Monte, S. & d'Ovidio, F. (2002). Dynamics of order parameters for globally coupled oscillators. *EPL (Europhysics Letters)*, 58(1), 21. Publisher: IOP Publishing.

[23] De Monte, S., D'Ovidio, F., Danø, S., & Sørensen, P. G. (2007). Dynamical quorum sensing: Population density encoded in cellular dynamics. *Proceedings of the National Academy of Sciences of the United States of America*, 104(47), 18377–18381.

[24] De Monte, S., d'Ovidio, F., Mosekilde, E., & Chaté, H. (2006). Low-Dimensional Chaos in Populations of Strongly-Coupled Noisy Maps. *Progress of Theoretical Physics Supplement*, 161, 27–42.

[25] De Monte, S., d'Ovidio, F., Chaté, H., & Mosekilde, E. (2004). Noise-Induced Macroscopic Bifurcations in Globally Coupled Chaotic Units. *Physical Review Letters*, 92(25), 254101. Publisher: American Physical Society.

[26] De Monte, S., d'Ovidio, F., Chaté, H., & Mosekilde, E. (2005). Effects of microscopic disorder on the collective dynamics of globally coupled maps. *Physica D: Nonlinear Phenomena*, 205(1), 25–40.

[27] De Monte, S., d'Ovidio, F., & Mosekilde, E. (2003). Coherent Regimes of Globally Coupled Dynamical Systems. *Physical Review Letters*, 90(5), 054102. Publisher: American Physical Society.

[28] De Monte, S. & Rainey, P. B. (2014). Nascent multicellular life and the emergence of individuality. *Journal of Biosciences*.

[29] De Monte, S., Soccodato, A., Alvain, S., & D'Ovidio, F. (2013). Can we detect oceanic biodiversity hotspots from space? *The ISME Journal*, 7(10), 2054–2056.

[30] de Vargas, C., Audic, S., Henry, N., Decelle, J., Mahé, F., Logares, R., Lara, E., Berney, C., Le Bescot, N., Probert, I., Carmichael, M., Poulain, J., Romac, S., Colin, S., Aury, J.-M., Bittner, L., Chaffron, S., Dunthorn, M., Engelen, S., Flegontova, O., Guidi, L., Horák, A., Jaillon, O., Lima-Mendez, G., Lukeš, J., Malviya, S., Morard, R., Mulot, M., Scalco, E., Siano, R., Vincent, F., Zingone, A., Dimier, C., Picheral, M., Searson, S., Kandels-Lewis, S., Acinas, S. G., Bork, P., Bowler, C., Gorsky, G., Grimsley, N., Hingamp, P., Iudicone, D., Not, F., Ogata, H., Pesant, S., Raes, J., Sieracki, M. E., Speich, S., Stemmann, L., Sunagawa, S., Weissenbach, J., Wincker, P., & Karsenti, E. (2015). Ocean plankton. Eukaryotic plankton diversity in the sunlit ocean. *Science (New York, N.Y.)*, 348(6237), 1261605.

[31] Demarre, G., Prudent, V., Schenk, H., Rousseau, E., Bringer, M.-A., Barnich, N., Nhieu, G. T. V., Rimsky, S., Monte, S. D., & Espéli, O. (2019). The Crohn's disease-associated Escherichia coli strain LF82 relies on SOS and stringent responses to survive, multiply and tolerate antibiotics within

macrophages. *PLOS Pathogens*, 15(11), e1008123. Publisher: Public Library of Science.

[32] Denk, J. & Hallatschek, O. (2021). *Self-consistent migration puts tight constraints on the spatio-temporal organization of species-rich metacommunities*. Technical report. Company: Cold Spring Harbor Laboratory Distributor: Cold Spring Harbor Laboratory Label: Cold Spring Harbor Laboratory Section: New Results Type: article.

[33] Dercole, F. & Rinaldi, S. (2008). *Analysis of Evolutionary Processes: The Adaptive Dynamics Approach and Its Applications*. Princeton University Press.

[34] Dickey, J. R., Swenie, R. A., Turner, S. C., Winfrey, C. C., Yaffar, D., Padukone, A., Beals, K. K., Sheldon, K. S., & Kivlin, S. N. (2021). The Utility of Macroecological Rules for Microbial Biogeography. *Frontiers in Ecology and Evolution*, 9, 196.

[35] Doebeli, M., Ispolatov, Y., & Simon, B. (2017). Towards a mechanistic foundation of evolutionary theory. *eLife*, 6.

[36] Doolittle, W. F. & Zhaxybayeva, O. (2010). Metagenomics and the Units of Biological Organization. *BioScience*, 60(2), 102–112.

[37] D'Ovidio, F., De Monte, S., Alvain, S., Dandonneau, Y., & Lévy, M. (2010). Fluid dynamical niches of phytoplankton types. *Proceedings of the National Academy of Sciences of the United States of America*, 107(43), 18366–70.

[38] Driscoll, W. W., Hackett, J. D., & Ferrière, R. (2016). Eco-evolutionary feedbacks between private and public goods: evidence from toxic algal blooms. *Ecology Letters*, 19(1), 81–97.

[39] Dubravcic, D., van Baalen, M., & Nizak, C. (2014). An evolutionarily significant unicellular strategy in response to starvation stress in Dictyostelium social amoebae. *F1000Research*, 3.

[40] Ferriere, R. & Cazelles, B. (1999). Universal power laws govern intermittent rarity in communities of interacting species. *Ecology*, 80(5), 1505–1521.

[41] Fily, Y. & Marchetti, M. C. (2012). Athermal Phase Separation of Self-Propelled Particles with No Alignment. *Physical Review Letters*, 108(23), 235702.

[42] Follows, M. J., Dutkiewicz, S., Grant, S., & Chisholm, S. W. (2007). Emergent Biogeography of Microbial Communities in a Model Ocean. *Science*. Publisher: American Association for the Advancement of Science.

[43] Forget, M., Adiba, S., & De Monte, S. (2021). Social conflicts in *Dictyostelium discoideum* : a matter of scales. *Peer Community Journal*, 1.

[44] Foster, K. R., Fortunato, A., Strassmann, J. E., & Queller, D. C. (2002). The costs and benefits of being a chimera. *Proceedings. Biological sciences / The Royal Society*, 269(1507), 2357–62.

[45] Fuhrman, J. A., Hewson, I., Schwalbach, M. S., Steele, J. A., Brown, M. V., & Naeem, S. (2006). Annually reoccurring bacterial communities are predictable from ocean conditions. *Proceedings of the National Academy of Sciences*, 103(35), 13104–13109. Publisher: National Academy of Sciences Section: Biological Sciences.

[46] Garcia, T., Brunnet, L. G., & De Monte, S. (2014). Differential adhesion between moving particles as a mechanism for the evolution of social groups. *PLoS computational biology*, 10(2), e1003482.

[47] Garcia, T. & De Monte, S. (2013). Group formation and the evolution of sociality. *Evolution*, 67(1), 131–141.

[48] Garcia, T., Doulcier, G., & De Monte, S. (2015). The evolution of adhesiveness as a social adaptation. *eLife*, 4, e08595.

[49] Gibbons, S. M., Caporaso, J. G., Pirrung, M., Field, D., Knight, R., & Gilbert, J. A. (2013). Evidence for a persistent microbial seed bank throughout the global ocean. *Proceedings of the National Academy of Sciences of the United States of America*, 110(12), 4651–5.

[50] Gilbert, J. A., Steele, J. A., Caporaso, J. G., Steinbrück, L., Reeder, J., Temperton, B., Huse, S., McHardy, A. C., Knight, R., Joint, I., Somerfield, P., Fuhrman, J. A., & Field, D. (2012). Defining seasonal marine microbial community dynamics. *The ISME journal*, 6(2), 298–308.

[51] Godfrey-Smith, P. (2009). *Darwinian Populations and Natural Selection*. Oxford, New York: Oxford University Press.

[52] Gokhale, C. & Traulsen, A. (2014). Evolutionary Multiplayer Games. *Dynamic Games and Applications*, 4.

[53] Goldford, J. E., Lu, N., Bajić, D., Estrela, S., Tikhonov, M., Sanchez-Gorostiaga, A., Segrè, D., Mehta, P., & Sanchez, A. (2018). Emergent simplicity in microbial community assembly. *Science*, 361(6401), 469–474. Publisher: American Association for the Advancement of Science Section: Research Article.

[54] Gomes Da Silva, I., De Monte, S., d'Ovidio, F., Toral, R., & Mirasso, C. R. (2006). Coherent regimes of mutually coupled Chua's circuits. *Physical Review E*, 73(3), 036203. Publisher: American Physical Society.

[55] Goury-Sistla, P., Nanjundiah, V., & Pande, G. (2012). Bimodal distribution of motility and cell fate in Dictyostelium discoideum. *International Journal of Developmental Biology*, 56(4), 263–272.

[56] Grey, D., Hutson, V., & Szathmáry, E. (1995). A re-exam ination of the stochastic corrector model. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, 262(1363), 29–35. Publisher: Royal Society.

[57] Grilli, J. (2020). Macroecological laws describe variation and diversity in microbial communities. *Nature Communications*, 11(1), 4743. Bandiera_abtest: a Cc_license_type: cc_by Cg_type: Nature Research Journals Number: 1 Primary_atype: Research Publisher: Nature Publishing Group Subject_term: Biodiversity;Community ecology;Ecological modelling;Macroecology;Microbial ecology Subject_term_id: biodiversity;community-ecology;ecological-modelling;macroecology;microbial-ecology.

[58] Grosberg, R. K. & Strathmann, R. R. (2007). The evolution of multicellularity: a minor major transition? *Annual Review of Ecology, Evolution, and Systematics*, 38(1), 621–654.

[59] Gruenheit, N., Parkinson, K., Brimson, C. A., Kuwana, S., Johnson, E. J., Nagayama, K., Llewellyn, J., Salvidge, W. M., Stewart, B., Keller, T., van Zon, W., Cotter, S. L., & Thompson, C. R. L. (2018). Cell Cycle Heterogeneity Can Generate Robust Cell Type Proportioning. *Developmental Cell*, 47(4), 494–508.e4.

[60] Grégoire, G. & Chaté, H. (2004). Onset of Collective and Cohesive Motion. *Physical Review Letters*, 92(2), 025702. Publisher: American Physical Society.

[61] Hamilton, W. D. (1964). The genetical evolution of social behaviour. I. *Journal of Theoretical Biology*, 7(1), 1–16.

[62] Hammerschmidt, K., Rose, C. J., Kerr, B., & Rainey, P. B. (2014). Life cycles, fitness decoupling and the evolution of multicellularity. *Nature*, 515(7525), 75–79. Bandiera_abtest: a Cg_type: Nature Research Journals Number: 7525 Primary_atype: Research Publisher: Nature Publishing Group Subject_term: Experimental evolution Subject_term_id: experimental-evolution.

[63] Hardin, G. (1960). The Competitive Exclusion Principle. *Science*, 131(3409), 1292–1297. Publisher: American Association for the Advancement of Science Section: Articles.

[64] Hardin, G. (1968). The Tragedy of the Commons. *Science*, 162(3859), 1243–1248. Publisher: American Association for the Advancement of Science.

[65] Hauert, C., De monte, S., Hofbauer, J., & Sigmund, K. (2002a). Replicator Dynamics for Optional Public Good Games. *Journal of Theoretical Biology*, 218(2), 187–194.

[66] Hauert, C., Monte, S. D., Hofbauer, J., & Sigmund, K. (2002b). Volunteering as Red Queen Mechanism for Cooperation in Public Goods Games. *Science*, 296(5570), 1129–1132. Publisher: American Association for the Advancement of Science Section: Report.

[67] He, F. (2005). Deriving a neutral model of species abundance from fundamental mechanisms of population dynamics. *Functional Ecology*, 19(1), 187–193.

[68] Hendershot, J. N., Read, Q. D., Henning, J. A., Sanders, N. J., & Classen, A. T. (2017). Consistently inconsistent drivers of microbial diversity and abundance at macroecological scales. *Ecology*, 98(7), 1757–1763. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/ecy.1829.

[69] Herron, M. D. & Michod, R. E. (2008). Evolution of Complexity in the Volvocine Algae: Transitions in Individuality Through Darwin's Eye. *Evolution*, 62(2), 436–451. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1558-5646.2007.00304.x.

[70] Hiraoka, H., Nakano, T., Kuwana, S., Fukuzawa, M., Hirano, Y., Ueda, M., Haraguchi, T., & Hiraoka, Y. (2020). Intracellular ATP levels influence cell fates in Dictyostelium discoideum differentiation. *Genes to Cells*, 25(5), 312–326.

[71] Hofbauer, J. & Sigmund, K. (1998). Evolutionary Games and Population Dynamics by Josef Hofbauer.

[72] Hubbell, S. P. (2001). *The Unified Neutral Theory of Biodiversity and Biogeography*. Princeton University Press.

[73] Hutchinson, G. E. (1961). The Paradox of the Plankton. *The American Naturalist*, 95(882), 137–145. Publisher: The University of Chicago Press.

[74] Huyan, C., Golden, A., Zhu, X., Mehta, P., & Sgro, A. E. (2021). *Robust coordination of collective oscillatory signaling requires single-cell excitability and fold-change detection*. Technical report. Company: Cold Spring Harbor Laboratory Distributor: Cold Spring Harbor Laboratory Label: Cold Spring Harbor Laboratory Section: New Results Type: article.

[75] Ibarbalz, F. M., Henry, N., Brandão, M. C., Martini, S., Busseni, G., Byrne, H., Coelho, L. P., Endo, H., Gasol, J. M., Gregory, A. C., Mahé, F., Rigonato, J., Royo-Llonch, M., Salazar, G., Sanz-Sáez, I., Scalco, E., Soviadan, D., Zayed, A. A., Zingone, A., Labadie, K., Ferland, J., Marec, C., Kandels, S., Picheral, M., Dimier, C., Poulain, J., Pisarev, S., Carmichael, M., Pesant, S., Acinas, S. G., Babin, M., Bork, P., Boss, E., Bowler, C., Cochrane, G., Vargas, C. d., Follows, M., Gorsky, G., Grimsley, N., Guidi, L., Hingamp, P., Iudicone, D., Jaillon, O., Kandels, S., Karp-Boss, L., Karsenti, E., Not, F., Ogata, H., Pesant, S., Poulton, N., Raes, J., Sardet, C., Speich, S., Stemmann, L., Sullivan, M. B., Sunagawa, S., Wincker, P., Babin, M., Boss, E., Iudicone, D., Jaillon, O., Acinas, S. G., Ogata, H., Pelletier, E., Stemmann, L., Sullivan, M. B., Sunagawa, S., Bopp, L., Vargas, C. d., Karp-Boss, L., Wincker, P., Lombard, F., Bowler, C., & Zinger, L. (2019). Global Trends in Marine Plankton Diversity across Kingdoms of Life. *Cell*, 179(5), 1084–1097.e21. Publisher: Elsevier.

[76] Jeraldo, P., Sipos, M., Chia, N., Brulc, J. M., Dhillon, A. S., Konkel, M. E., Larson, C. L., Nelson, K. E., Qu, A., Schook, L. B., Yang, F., White, B. A., & Goldenfeld, N. (2012). Quantification of the relative roles of niche and neutral processes in structuring gastrointestinal microbiomes. *Proceedings of the National Academy of Sciences*, 109(25), 9692–9698.

[77] Jones, M. L., Rivett, D. W., Pascual-García, A., & Bell, T. (2021). Relationships between community composition, productivity and invasion resistance in semi-natural bacterial microcosms. *eLife*, 10, e71811. Publisher: eLife Sciences Publications, Ltd.

[78] Karsenti, E., Acinas, S. G., Bork, P., Bowler, C., Vargas, C. D., Raes, J., Sullivan, M., Arendt, D., Benzoni, F., Claverie, J.-M., Follows, M., Gorsky, G., Hingamp, P., Iudicone, D., Jaillon, O., Kandels-Lewis, S., Krzic, U., Not, F., Ogata, H., Pesant, S., Reynaud, E. G., Sardet, C., Sieracki, M. E., Speich, S., Velayoudon, D., Weissenbach, J., Wincker, P., & Consortium, t. T. O. (2011). A Holistic Approach to Marine Eco-Systems Biology. *PLOS Biology*, 9(10), e1001177. Publisher: Public Library of Science.

[79] Kessin, R. H. (2001). *Dictyostelium: evolution, cell biology, and the development of multicellularity*. Cambridge University Press, 1 edition.

[80] Kessler, D. A. & Shnerb, N. M. (2015). Generalized model of island biodiversity. *Physical Review E*, 91(4), 042705. Publisher: American Physical Society.

[81] Lennon, J. T. & Jones, S. E. (2011). Microbial seed banks: the ecological and evolutionary implications of dormancy. *Nature Reviews Microbiology*, 9(2), 119–130. Bandiera_abtest: a Cg_type: Nature Research Journals Number: 2 Primary_atype: Reviews Publisher: Nature Publishing Group Subject_term: Environmental microbiology;Evolution;Metagenomics;Microbial ecology Subject_term_id: environmental-microbiology;evolution;metagenomics;microbial-ecology.

[82] Lewontin, R. C. (1970). The Units of Selection. *Annual Review of Ecology and Systematics*, 1(1), 1–18. _eprint: https://doi.org/10.1146/annurev.es.01.110170.000245.

[83] Liautaud, K., van Nes, E. H., Barbier, M., Scheffer, M., & Loreau, M. (2019). Superorganisms or loose collections of species? A unifying theory of community patterns along environmental gradients. *Ecology Letters*.

[84] Libby, E. & Rainey, P. B. (2013). A conceptual framework for the evolutionary origins of multicellularity. *Physical Biology*, 10(3), 035001. Publisher: IOP Publishing.

[85] Lindemann, S. R., Bernstein, H. C., Song, H.-S., Fredrickson, J. K., Fields, M. W., Shou, W., Johnson, D. R., & Beliaev, A. S. (2016). Engineering microbial consortia for controllable outputs. *The ISME Journal*, 10(9), 2077–2084. Bandiera_abtest: a Cc_license_type: cc_y Cg_type: Nature Research Journals Number: 9 Primary_atype: Reviews Publisher: Nature Publishing Group Subject_term: Biological techniques Subject_term_id: biological-techniques.

[86] Lion, S. (2018). Theoretical Approaches in Evolutionary Ecology: Environmental Feedback as a Unifying Perspective. *The American Naturalist*, 191(1), 21–44. Publisher: The University of Chicago Press.

[87] Longhurst, A. R., Ed. (2007). *Ecological Geography of the Sea (Second Edition)*. Burlington: Academic Press.

[88] Louca, S., Parfrey, L. W., & Doebeli, M. (2016). Decoupling function and taxonomy in the global ocean microbiome. *Science*, 353(6305), 1272–1277.

[89] Machado, D., Maistrenko, O. M., Andrejev, S., Kim, Y., Bork, P., Patil, K. R., & Patil, K. R. (2021). Polarization of microbial communities between competitive and cooperative metabolism. *Nature Ecology & Evolution*, 5(2), 195–203. Bandiera_abtest: a Cg_type: Nature Research Journals Number: 2 Primary_atype: Research Publisher: Nature Publishing Group Subject_term: Computational biology and bioinformatics;Ecology;Microbiology;Systems biology Subject_term_id: computational-biology-and-bioinformatics;ecology;microbiology;systems-biology.

[90] Madgwick, P. G., Stewart, B., Belcher, L. J., Thompson, C. R. L., & Wolf, J. B. (2018). Strategic investment explains patterns of cooperation and cheating in a microbe. *Proceedings of the National Academy of Sciences*, 115(21), E4823–E4832.

[91] Magurran, A. E. (2004). *Measuring Biological Diversity*. Wiley. Google-Books-ID: tUqzLSUzXxcC.

[92] Magurran, A. E. & Henderson, P. A. (2003). Explaining the excess of rare species in natural species abundance distributions. *Nature*, 422(6933), 714–716.

[93] Martin-Platero, A. M., Cleary, B., Kauffman, K., Preheim, S. P., McGillicuddy, D. J., Alm, E. J., & Polz, M. F. (2018). High resolution time series reveals cohesive but short-lived communities in coastal plankton. *Nature Communications*, 9(1), 266. Bandiera_abtest: a Cc_license_type: cc_by Cg_type: Nature Research Journals Number: 1 Primary_atype: Research Publisher: Nature Publishing Group Subject_term: Marine biology;Microbial ecology Subject_term_id: marine-biology;microbial-ecology.

[94] Martín, P. V., Buček, A., Bourguignon, T., & Pigolotti, S. (2020). Ocean currents promote rare species diversity in protists. *Science Advances*. Publisher: American Association for the Advancement of Science.

[95] Matthews, T. J. & Whittaker, R. J. (2014). Fitting and comparing competing models of the species abundance distribution: assessment and prospect. *Frontiers of Biogeography*, 6(2).

[96] Matthews, T. J. & Whittaker, R. J. (2015). REVIEW: On the species abundance distribution in applied ecology and biodiversity management. *Journal of Applied Ecology*, 52(2), 443–454. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/1365-2664.12380.

[97] May, R. M. (1972). Will a Large Complex System be Stable? *Nature*, 238(5364), 413–414. Bandiera_abtest: a Cg_type: Nature Research Journals Number: 5364 Primary_atype: Research Publisher: Nature Publishing Group.

[98] Maynard Smith, J. & Szathmary, E. (1997). *The Major Transitions in Evolution*. Oxford: Oxford University Press.

[99] McGill, B. J., Etienne, R. S., Gray, J. S., Alonso, D., Anderson, M. J., Benecha, H. K., Dornelas, M., Enquist, B. J., Green, J. L., He, F., Hurlbert, A. H., Magurran, A. E., Marquet, P. A., Maurer, B. A., Ostling,

A., Soykan, C. U., Ugland, K. I., & White, E. P. (2007). Species abundance distributions: moving beyond single prediction theories to integration within an ecological framework. *Ecology Letters*, 10(10), 995–1015.

[100] Miele, L. & De Monte, S. (2021). Aggregative cycles evolve as a solution to conflicts in social investment. *PLOS Computational Biology*, 17(1), e1008617. Publisher: Public Library of Science.

[101] Muñoz-Dorado, J., Marcos-Torres, F. J., García-Bravo, E., Moraleda-Muñoz, A., & Pérez, J. (2016). Myxobacteria: Moving, Killing, Feeding, and Surviving Together. *Frontiers in Microbiology*, 7, 781.

[102] Márquez-Zacarías, P., Conlin, P. L., Tong, K., Pentz, J. T., & Ratcliff, W. C. (2021). Why have aggregative multicellular organisms stayed simple? *Current Genetics*.

[103] Okasha, S. (2006). *Evolution and the Levels of Selection*. Oxford: Oxford University Press.

[104] Olson, B. J. (2013). From brief encounters to lifelong unions. *eLife*, 2, e01893. Publisher: eLife Sciences Publications, Ltd.

[105] Pentz, J. T., Márquez-Zacarías, P., Bozdag, G. O., Burnetti, A., Yunker, P. J., Libby, E., & Ratcliff, W. C. (2020). Ecological Advantages and Evolutionary Limitations of Aggregative Multicellular Development. *Current Biology*, 30(21), 4155–4164.e6.

[106] Peña, J., Lehmann, L., & Nöldeke, G. (2014). Gains from switching and evolutionary stability in multi-player matrix games. *Journal of Theoretical Biology*, 346, 23–33.

[107] Pfennig, D. W., Ed. (2021). *Phenotypic Plasticity & Evolution: Causes, Consequences, Controversies*. Boca Raton: CRC Press.

[108] Pichugin, Y. & Traulsen, A. (2020). Evolution of multicellular life cycles under costly fragmentation. *PLOS Computational Biology*, 16(11), e1008406. Publisher: Public Library of Science.

[109] Pikovsky, A., Rosenblum, M., & Kurths, J. (2001). *Synchronization: A Universal Concept in Nonlinear Sciences*. Cambridge Nonlinear Science Series. Cambridge: Cambridge University Press.

[110] Rainey, P. B. & De Monte, S. (2014). Resolving Conflicts During the Evolutionary Transition to Multicellular Life. *Annual Review of Ecology, Evolution, and Systematics*, 45(1), 599–620.

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

[112] Ramaswamy, S. (2010). The Mechanics and Statistics of Active Matter. *Annual Review of Condensed Matter Physics*, 1(1), 323–345. _eprint: https://doi.org/10.1146/annurev-conmatphys-070909-104101.

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

[114] Remigi, P., Ferguson, G. C., McConnell, E., De Monte, S., Rogers, D. W., & Rainey, P. B. (2019). Ribosome Provisioning Activates a Bistable Switch Coupled to Fast Exit from Stationary Phase. *Molecular Biology and Evolution*, 36(5), 1056–1070. Publisher: Oxford Academic.

[115] Rodríguez-Ezpeleta, N., Zinger, L., Kinziger, A., Bik, H. M., Bonin, A., Coissac, E., Emerson, B. C., Lopes, C. M., Pelletier, T. A., Taberlet, P., & Narum, S. (2021). Biodiversity monitoring using environmental DNA. *Molecular Ecology Resources*, 21(5), 1405–1409. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/1755-0998.13399.

[116] Rossine, F. W., Martinez-Garcia, R., Sgro, A. E., Gregor, T., & Tarnita, C. E. (2020). Eco-evolutionary significance of "loners". *PLOS Biology*, 18(3), e3000642.

[117] Rusch, D. B., Halpern, A. L., Sutton, G., Heidelberg, K. B., Williamson, S., Yooseph, S., Wu, D., Eisen, J. A., Hoffman, J. M., Remington, K., Beeson, K., Tran, B., Smith, H., Baden-Tillson, H., Stewart, C., Thorpe, J., Freeman, J., Andrews-Pfannkoch, C., Venter, J. E., Li, K., Kravitz, S., Heidelberg, J. F., Utterback, T., Rogers, Y.-H., Falcón, L. I., Souza, V., Bonilla-Rosso, G., Eguiarte, L. E., Karl, D. M., Sathyendranath, S., Platt, T., Bermingham, E., Gallardo, V., Tamayo-Castillo, G., Ferrari, M. R., Strausberg, R. L., Nealson, K., Friedman, R., Frazier, M., & Venter, J. C. (2007). The Sorcerer II Global Ocean Sampling Expedition: Northwest

Atlantic through Eastern Tropical Pacific. *PLOS Biology*, 5(3), e77. Publisher: Public Library of Science.

[118] Sathe, S. & Nanjundiah, V. (2018). Complex interactions underpin social behaviour in Dictyostelium giganteum. *Behavioral Ecology and Sociobiology*, 72(10), 167.

[119] Sebé-Pedrós, A., Irimia, M., del Campo, J., Parra-Acero, H., Russ, C., Nusbaum, C., Blencowe, B. J., & Ruiz-Trillo, I. (2013). Regulated aggregative multicellularity in a close unicellular relative of metazoa. *eLife*, 2, e01287. Publisher: eLife Sciences Publications, Ltd.

[120] Ser-Giacomi, E., Zinger, L., Malviya, S., De Vargas, C., Karsenti, E., Bowler, C., & De Monte, S. (2018). Ubiquitous abundance distribution of non-dominant plankton across the global ocean. *Nature Ecology & Evolution*.

[121] Sigmund, K. (2010). *The Calculus of Selfishness*. Princeton University Press. Publication Title: The Calculus of Selfishness.

[122] Simon, J.-C., Marchesi, J. R., Mougel, C., & Selosse, M.-A. (2019). Host-microbiota interactions: from holobiont theory to analysis. *Microbiome*, 7(1), 5.

[123] Soccodato, A., d'Ovidio, F., Lévy, M., Jahn, O., Follows, M. J., & De Monte, S. (2016). Estimating planktonic diversity through spatial dominance patterns in a model ocean. *Marine Genomics*, 29, 9–17.

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

[125] Stegen, J. C., Ferriere, R., & Enquist, B. J. (2012). Evolving ecological networks and the emergence of biodiversity patterns across temperature gradients. *Proceedings. Biological Sciences*, 279(1731), 1051–1060.

[126] Strassmann, J. E. & Queller, D. C. (2011). Evolution of cooperation and control of cheating in a social microbe. *Proceedings of the National Academy of Sciences*, 108(Supplement 2), 10855–10862.

[127] Strassmann, J. E., Zhu, Y., & Queller, D. C. (2000). Altruism and social cheating in the social amoeba Dictyostelium discoideum. *Nature*, 408(6815), 965–967.

[128] Strogatz, S. H. (2012). *Sync: How Order Emerges from Chaos In the Universe, Nature, and Daily Life*. Hachette Books. Google-Books-ID: ZQeZAAAAQBAJ.

[129] Sumpter, D. J. T. (2010). *Collective Animal Behavior*. Princeton University Press.

[130] Sunagawa, S., Coelho, L. P., Chaffron, S., Kultima, J. R., Labadie, K., Salazar, G., Djahanschiri, B., Zeller, G., Mende, D. R., Alberti, A., Cornejo-Castillo, F. M., Costea, P. I., Cruaud, C., D'Ovidio, F., Engelen, S., Ferrera, I., Gasol, J. M., Guidi, L., Hildebrand, F., Kokoszka, F., Lepoivre, C., Lima-Mendez, G., Poulain, J., Poulos, B. T., Royo-Llonch, M., Sarmento, H., Vieira-Silva, S., Dimier, C., Picheral, M., Searson, S., Kandels-Lewis, S., Bowler, C., de Vargas, C., Gorsky, G., Grimsley, N., Hingamp, P., Iudicone, D., Jaillon, O., Not, F., Ogata, H., Pesant, S., Speich, S., Stemmann, L., Sullivan, M. B., Weissenbach, J., Wincker, P., Karsenti, E., Raes, J., Acinas, S. G., Bork, P., Boss, E., Bowler, C., Follows, M., Karp-Boss, L., Krzic, U., Reynaud, E. G., Sardet, C., Sieracki, M., & Velayoudon, D. (2015). Structure and function of the global ocean microbiome. *Science*, 348(6237), 1261359–1261359.

[131] Szabó, B., Szöllösi, G. J., Gönci, B., Jurányi, Z., Selmeczi, D., & Vicsek, T. (2006). Phase transition in the collective migration of tissue cells: Experiment and model. *Physical Review E*, 74(6).

[132] Sánchez, ., Vila, J. C. C., Chang, C.-Y., Diaz-Colunga, J., Estrela, S., & Rebolleda-Gomez, M. (2021). Directed Evolution of Microbial Communities. *Annual Review of Biophysics*, 50, 323–341.

[133] Tarnita, C. E., Washburne, A., Martinez-Garcia, R., Sgro, A. E., & Levin, S. A. (2015). Fitness tradeoffs between spores and nonaggregating cells can explain the coexistence of diverse genotypes in cellular slime molds. *Proceedings of the National Academy of Sciences*, 112(9), 2776–2781.

[134] Tilman, A. R., Plotkin, J. B., & Akçay, E. (2020). Evolutionary games with environmental feedbacks. *Nature Communications*, 11(1), 915. Number: 1 Publisher: Nature Publishing Group.

[135] Van Cleve, J. (2017). Stags, Hawks, and Doves: Social Evolution Theory and Individual Variation in Cooperation. *Integrative and Comparative Biology*, 57(3), 566–579. Publisher: Oxford Academic.

[136] Van Cleve, J. (2021). The cell-level perspective in social conflicts in Dictyostelium discoideum. *Peer Community in Evolutionary Biology*, 1, 100122. Company: Peer Community in Evolutionary Biology Distributor: Peer Community in Evolutionary Biology Institution: Peer Community in Evolutionary Biology Label: Peer Community in Evolutionary Biology Publisher: Peer Community In.

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

[138] Venturelli, O. S., Carr, A. V., Fisher, G., Hsu, R. H., Lau, R., Bowen, B. P., Hromada, S., Northen, T., & Arkin, A. P. (2018). Deciphering microbial interactions in synthetic human gut microbiome communities. *Molecular Systems Biology*, 14(6), e8157. Publisher: John Wiley & Sons, Ltd.

[139] Vliet, S. v. & Doebeli, M. (2019). The role of multilevel selection in host microbiome evolution. *Proceedings of the National Academy of Sciences*, 116(41), 20591–20597.

[140] Weitz, J. S., Eksin, C., Paarporn, K., Brown, S. P., & Ratcliff, W. C. (2016). An oscillating tragedy of the commons in replicator dynamics with game-environment feedback. *Proceedings of the National Academy of Sciences*, 113(47), E7518–E7525.

[141] West, S. A., Griffin, A. S., Gardner, A., & Diggle, S. P. (2006). Social evolution theory for microorganisms. *Nature Reviews Microbiology*, 4(8), 597–607.

[142] Williams, H. T. P. & Lenton, T. M. (2007). Artificial selection of simulated microbial ecosystems. *Proceedings of the National Academy of Sciences*, 104(21), 8918–8923.

[143] Wilson, D. S. (1975). A theory of group selection. *Proceedings of the National Academy of Sciences*, 72(1), 143–146.

[144] Xie, L. & Shou, W. (2021). Steering ecological-evolutionary dynamics to improve artificial selection of microbial communities. *Nature Communications*, 12(1), 6799.

[145] Xie, L., Yuan, A. E., & Shou, W. (2019). Simulations reveal challenges to artificial community selection and possible strategies for success. *PLOS Biology*, 17(6), e3000295.

This thesis was typeset using LaTeX, originally developed by Leslie Lamport and based on Donald Knuth's TeX. The template by Jordan Suchow has been released under the permissive MIT (X11) license, and can be found online at github.com/suchow/Dissertate.