UNIVERSIDAD POLITÉCNICA DE MADRID

Escuela Técnica Superior de Ingenieros Informáticos

Departamento de Inteligencia Artificial

Modelling bacterial ecological strategies using an individual-based simulator

Paula Gregorio Godoy

Tesis doctoral presentada para optar al título de Doctor en Inteligencia Artificial

Director:

Dr. Alfonso Rodríguez-Patón Aradas

Co-director:

Dr. Juan Pazos Sierra

Resumen

La cooperación que se establece entre individuos en el mundo microbiano es una conducta social difícil de explicar en el marco de la teoría de la selección natural. Sin embargo, cuando la cooperación es mantenida en el tiempo por un grupo de individuos cercanos, este fenómeno puede dar lugar a lo que se conoce como transición mayor evolutiva, y resulta en la emergencia de un nuevo organismo multicelular. Un organismo se considera multicelular si las células que lo componen no son independientes entre sí sino que forman un grupo que se coordina y comunica, donde su comportamiento final está determinado por la interacción de sus elementos constituyentes.

En esta tesis, se investigan los diferentes pasos que conducen al fenómeno de la multicelularidad. Dichos fenómenos se analizan por medio del simulador basado en individuos, *gro*, una herramienta que permite la simulación de una variedad de experimentos en diferentes contextos ecológicos. *gro* está diseñado para simular el nacimiento y la competición de las bacterias. Si a esto se le suma que también permite la especificación de características individuales, que además pueden ser heredadas, la selección natural resulta un proceso emergente. Gracias al uso de estos simuladores se puede comprender cómo emerge el comportamiento colectivo organizado a partir de interacciones individuales a pequeña escala.

Esta tesis ha sido inspirada por las condiciones que se requieren para que surja dicha multicelularidad. Estas condiciones incluyen el proceso físico-mecánico de la formación de un grupo de células (Capítulo 3), los subsecuentes mecanismos que intervienen en la resolución de conflictos y el mantenimiento de la cooperación entre los individuos constituyentes (Capítulo 4). Finalmente, esta tesis concluye con el estudio de una de las más relevantes consecuencias de la multicelularidad: la división de trabajo (Capítulo 5). Los resultados indican las posibles tendencias evolutivas de cada estrategia, a través de la evolución de las funciones de evaluación de cada una. El objetivo de esta tesis no es el de ofrecer predicciones precisas sino deducir las reglas generales de un sistema basado en la cooperación. Las contribuciones más importantes de esta tesis incluyen la identificación de principios tales como que la competencia celular crea grupos más cohesivos, que la cooperación colectiva requiere del egoísmo individual para prosperar y, finalmente, que la coordinación local entre células puede producir la aparición espontánea de la división de trabajo.

Abstract

Cooperation in the microbial world is a social behaviour difficult to explain in the face of natural selection theory. Yet, when cooperation is consistently maintained within a group, a major transition in individuality can occur: multicellularity. An organism is considered multicellular if the cells that comprise it are not independent of each other but rather form a group that coordinates and communicates, with a resulting behaviour that is determined by the interaction of its fundamental elements.

In this thesis, I have investigated some of the different steps that lead to the emergence of multicellularity. Such processes are simulated by means of the individual-based simulator *gro* throughout a variety of experiments that assume different ecological. *gro* is designed to reproduce birth and competition among the bacterial cells. This added to the specification of the individual traits and the heritability of those traits, natural selection results as an emergent process. Thanks to the individual-based models it is possible to study how apparently organized collective behaviour emerges out of the small-scale interactions between the individuals.

This thesis has been inspired by the conditions that are required for multicellularity to emerge. Briefly, the process of group formation (Chapter 3), the mechanisms for conflict resolution and the maintenance of cooperation (Chapter 4). Finally, this thesis concludes with the study of the most relevant consequence of multicellularity: division of labour (Chapter 5). Results indicate the possible evolutionary tendencies according to the evolution of the fitness of different bacterial ecological strategies. The global aim of this thesis is therefore not to offer specific predictions but rather understand the general properties of a cooperating collective by determining the minimal conditions for the evolutionary transition of multicellularity. The most relevant contributions include the identification of common principles across some of the simulated processes, namely, that cell competition creates more cohesive cell groups, cooperation requires individual selfishness to thrive and specific local cell coordination can result in the emergence of division of labour in homogeneous cell colonies.

Contents

Resumen.....	iii
Abstract.....	iv
Contents	v
List of Figures.....	vii
List of Tables	xv
Chapter 1 Introduction	1
1.1 Social evolution theory in microorganisms.....	2
1.2 Multicellularity and division of labour.....	9
1.3 Ecological and evolutionary feedbacks	11
1.4 Individual-based models.....	13
1.5 Questions addressed in this thesis	17
Chapter 2 Methodology.....	19
2.1 ODD description of the simulator <i>gro</i>	20
Chapter 3 Group formation	27
3.1 Spatial properties of the simulated colony	31
3.2 Group formation and space structuring	39
3.3 Bacterial size and its effect on group formation	51
3.4 Resource limitation and its effect on group formation.....	56
3.5 Size competition in resource limited environments	64
3.6 Conclusions	68
Chapter 4 Individual strategies for the public goods dilemma	71
4.1 Constrained growth in a microcolony of cooperators and defectors.....	77
4.2 Resource limitation enhances cooperation at the colony edge.....	83
4.3 Obligated cross-feeding reduces defector invasion.....	89
4.4 Conjugation of cooperative genes promotes cooperation	95
4.5 Density conditional cooperation is beneficial independently of its costs	107
4.6 Selfish good production promotes cooperation.....	113
4.7 Combined strategies	117
4.8 Conclusions	123

Chapter 5	Division of labour	128
5.1	Why divide labour?	134
5.2	How to divide labour?	147
5.3	Division of labour from coordinated signalling	162
5.3.1	One-signal circuit	162
5.3.2	Two-signals circuit	178
5.3.3	Fitness benefits of emergent division of labour	193
5.3.4	Morphogenetic properties: Repeatability, scalability and resizing	195
5.4	Conclusions	207
Chapter 6	Conclusions	209
6.1	Future lines	212
Appendix		214
References		215

List of Figures

Figure 1.1– Fundamental social interactions.	4
Figure 1.2– Illustrations of different Myxobacteria fruiting bodies from Roland Thaxter, distributed under a CC-BY 2.0 license.	5
Figure 1.3– Steps for the evolution of multicellularity. Cell aggregation and conflict resolution (cooperation) are major requisites to constitute a multicellular organism.	11
Figure 2.1– Process workflow per simulation step in <i>gro</i>	21
Figure 2.2– Different initial cell densities.	23
Figure 3.1– <i>Pseudomonas aeruginosa</i> colonies form complex structures. Images from Scott Chimileski, microbephotography.com.	28
Figure 3.2– Electron micrograph of a cluster of <i>E. coli</i> bacteria magnified 10,000 times. Each individual bacterium is rod-shaped. Image taken by Rocky Mountain Laboratories, NIAID, NIH, distributed under a CC-BY 2.0 license.	31
Figure 3.3– Representation of the cell as a rod-shaped rigid body.	32
Figure 3.4– Distribution of cell length (A) and cell width (B) taken from. Scale bars is 10 μ m. Reprinted with permission from (T. J. Rudge et al. 2013) Copyright 2018 American Chemical Society.	32
Figure 3.5– (A) Fluorescence microscopy image of an <i>E. coli</i> colony provided by Haseloff Lab (courtesy of Dr. Anton Kan). The arrows indicate the image processing (B) and segmentation process (C).	34
Figure 3.6– Alignment of cells in a simulated colony.	35
Figure 3.7– Nutrient distribution modes and resulting colony morphology.	36
Figure 3.8– Radial growth in three different regimes of resource requirements.	38
Figure 3.9– Cells in a bacterial colony are connected by physical contacts that can change with time, very much like a dynamic graph.	39
Figure 3.10– A colony consists on many connected groups. In turn, a group consists on many individual cells.	40
Figure 3.11– Different elements of a physically connected group.	41
Figure 3.12– Dynamical formation of a <i>physically connected group</i>	42
Figure 3.13– Frequency of the different members of a <i>group</i> over time.	43
Figure 3.14– Time evolution of the frequency of the clonal cells and their closest neighbours and image shots at 60'. Simulations were run for various initial number of clonal cells.	46
Figure 3.15– Time shots of the cell arrangement for simulations using different number of strains. The images show how space becomes more structured and patches of clonal cell develop as cells divide. Cells are colour marked but are otherwise equal.	48

Figure 3.16–. Colonies formed from single cells from various strains.....	50
Figure 3.17 - Effect of bacterial size on the colony arrangement and cell alignment. Lighter blue indicates the orientation of the cell is in the same direction as its close neighbours.	52
Figure 3.18 – Shots of the <i>groups</i> with different division lengths.....	53
Figure 3.19–Bacterial sizes and colour code of the different lineages	55
Figure 3.20– Visualization of a group in the centre and in the edge of a colony that is nutrient-limited.	57
Figure 3.21– Frequency of clonal cells and close neighbours in an edge group over time in a colony that is nutrient limited.	58
Figure 3.22–Spatial segregation of five lineages due to nutrient limited growth.....	59
Figure 3.23– Effect of nutrient limitation for different cell sizes and a consumption rate of 0.3 mol/dt.	61
Figure 3.24– Effect of nutrient limitation for different cell sizes and a consumption rate of 0.6 mol/dt.	62
Figure 3.25– Fraction of close neighbours in a group for different consumption rates and cell division lengths.	63
Figure 3.26–Frequency of cells from each strain as a measure of the fitness of the strain size. Competition between sizes is simulated in a pair-wise manner for unrestricted and resource-restricted growth. Colonies are seeded with a single cell from each strain, and colony shots are taken when the population reaches 5000 cells.	66
Figure 3.27–Frequency of cells from each strain as a measure of the fitness of the strain size. Competition between sizes is simulated in a pair-wise manner for unrestricted and resource-restricted growth. Colonies are seeded with an initial well-mixed clump of 100 cells from each strain, and colony shots are taken when the population reaches 5000 cells.	67
Figure 4.1– The public goods dilemma. Adaptation of the selfish mutant that is selected for its higher reproductive rate, a consequence of not producing the common good, results in the extinction of cooperators.....	71
Figure 4.2– Representation of the behaviour of the two phenotypes.	72
Figure 4.3– Diffusion of the concentration of good from the cooperator source.	73
Figure 4.4– Concentration of good over time for three different cells according to their proximity to the source	74
Figure 4.5– Public good coefficient is determined by the availability of good.	75
Figure 4.6– Evolution of the frequency of phenotypes, cooperators and defectors, for high and low-cost production of public goods. Colony shots depict the growth patterns of cooperators (dark blue) and defectors (light blue).	78
Figure 4.7– Growth pattern of two bacterial strains interacting only through mechanical contact.	80
Figure 4.8– Averaged group relatedness, relative and absolute fitness for N=20 simulations at 160'.	81

Figure 4.9– Relative fitness of cooperators for populations constituted with varying cell sizes.	82
Figure 4.10– Evolution of the frequency of phenotypes, cooperators and defectors, for high and low-cost production of public goods in a nutrient limited environment. Colony shots depict the growth patterns of cooperators (dark blue) and defectors (light blue).	85
Figure 4.11– Averaged group relatedness, relative and absolute fitness for N=20 simulations at 160'.	86
Figure 4.12–Relative and absolute fitness for different nutrient consumption rates (N=20 simulations at 160').	87
Figure 4.13– Cross-feeding growth pattern between strains S1 and S2.	90
Figure 4.14–Cross-feeding mediated cooperation between strains S1 and S2. The strain S1 consists of cooperators and defectors that do not secrete any metabolic product.	91
Figure 4.15–Evolution of the frequency of cooperators and defectors of the cross-feeding strain S1. Colony shots depict the growth patterns whit high and low cross-feeding.	92
Figure 4.16– Averaged group relatedness, relative and absolute fitness for N=20 simulations at 160'.	93
Figure 4.17–Frequency of cooperators of the cross-feeding strain S1 in a long-time simulation.	94
Figure 4.18–Cooperative genes are placed in conjugative plasmids. Transmission of the cooperative plasmid converts defectors in new cooperators.	97
Figure 4.19–Time evolution of the frequency of cooperators and defectors when the cooperative trait is placed in a conjugative plasmid. Colony shots depict the growth patterns of cooperators (dark blue) and defectors (light blue) within the same colony.	99
Figure 4.20– Averaged relative and absolute fitness of cooperators and group relatedness for N=20 simulations at 160'.	100
Figure 4.21–Time evolution of the <i>group</i> members when conjugation costs are low.	101
Figure 4.22– A population with cooperative cells (bearing the cooperative plasmid), defectors with plasmid (bearing the defector plasmid) and plasmid-free defectors. Transmission of any plasmid to a plasmid-free defector prevents such defector to be further infected by the other plasmid.	102
Figure 4.23–Time evolution of the frequency of cooperators when the competitive defector plasmid composes 5% of the initial population, for medium and low-cost of conjugation and cooperation.	103
Figure 4.24– Time evolution of the frequency of cooperators when the competitive defector plasmid composes 20% and 33% of the initial population, for medium and low-cost of conjugation and cooperation.	104
Figure 4.25– Averaged relative and absolute fitness of cooperators and group relatedness for N=20 simulations at 160' for different initial fractions of the competitor plasmid.	105

Figure 4.26– Cooperative genes are placed under the regulation of a QS gene. Only when quorum is reached (enough density of QS emitters, yellow signal) cooperators secrete the public good (blue signal).	108
Figure 4.27– Time evolution of the frequency of cooperators and defectors when public good secretion is conditional to the density of cooperators.	110
Figure 4.28– Averaged relative and absolute fitness of cooperators and group relatedness for N=20 simulations at 160'.	111
Figure 4.29– Cooperative cells become a little selfish by reducing the emission and diffusivity of the emitted common good so they become the main beneficiary.	114
Figure 4.30– Time evolution of the frequency of cooperators and defectors when public good secretion is selfish (lower production rate and diffusivity).	115
Figure 4.31– Averaged relative and absolute fitness of cooperators and group relatedness for N=20 simulations at 160'.	116
Figure 4.32– Colony shots for simulations of selfish production of public goods in a resource limited environment. Growth patterns of cooperators (dark blue) and defectors (light blue) are shown for high-cost and low-cost cooperation.	117
Figure 4.33– Average relative and absolute fitness for the simulation of selfish production of public goods in a resource limited environment.	118
Figure 4.34– Average relative and absolute fitness for the simulations of selfish production of public goods when cooperative genes are encoded in a conjugative plasmid.	119
Figure 4.35– Average relative and absolute fitness for the simulations of selfish production of public goods when cooperation is conditional to the local density of cooperators and regulated by QS.	120
Figure 4.36– Conjugative plasmids include cooperative and QS genes. Conjugation spreads the cooperative plasmid (t_1) allowing for the subsequent quorum (t_2) and emission of the public good (t_3).	121
Figure 4.37– Averaged relative and absolute fitness of cooperators and group relatedness for N=20 simulations at 160'.	122
Figure 4.38– Absolute fitness of cooperators of all strategies and conditions for $t=160'$. The colour code of each bar represents the associated cost of cooperation and each strategy (if required) in each case.	127
Figure 5.1– Steps for the evolution of multicellularity. Cell adhesion and cooperation (group formation) are major requisites to constitute a multicellular organism that is constituted of specialized cells that are able to divide labour (group transformation).	129
Figure 5.2– Division of labour in cells performing two tasks.	134
Figure 5.3- Two separate setups to test two different strategies. Simulations can be either in monoculture, for isogenic populations or in co-culture of two types specialized cells.	135
Figure 5.4– The degree of cooperation establishes the length of communication. With the parameters from Table 5-1, a cell can communicate at least with its closest neighbours.	136
Figure 5.5– Molecule reception dictates generation time.	137

Figure 5.6– Emission rates of molecules A and B of the individuals with different degrees of specialization for six different combinations of population 1 and 2.....	138
Figure 5.7– Percentage of cells and their correspondent generation times for the six experiments of Figure 5.6.	139
Figure 5.8– Global fitness of the six simulations with different degrees of specialization ...	140
Figure 5.9– Accelerated emission rates of molecules A and B of the individuals with different degrees of specialization for six different combinations of population 1 and 2.	141
Figure 5.10– Percentage of cells with their correspondent generation times for the six different simulations.	142
Figure 5.11– Global fitness of the six simulations with different degrees of specialization for an accelerating emission rate.	143
Figure 5.12– Global fitness with different degrees of specialization for different emission functions.....	144
Figure 5.13– Global fitness with different degrees of specialization for different emission functions and two communication lengths.	146
Figure 5.14– Four different mechanism to simulate phenotypic variability.....	148
Figure 5.15– Percentage of soma cells over time. Since soma cells do not reproduce, their proportion in a well-mixed population of germ cells becomes diluted over time.	150
Figure 5.16– Population fitness for different initial fractions of soma cells after 3 hours of simulated time. The baseline represents a population consisted only of germ cells.....	151
Figure 5.17– Percentage of soma cells over time for a random and irreversible specialization. Each curve represents the data for a different mutation rate.	152
Figure 5.18– Population fitness for different initial mutation rates after 3 hours of simulated time. The baseline represents a population consisted only of germ cells.	153
Figure 5.19– Percentage of soma cells over time for a random and reversible specialization. Each curve represents the data for a different noise rates.....	154
Figure 5.20– Population fitness for different initial noise rates after 3 hours of simulated time. The baseline represents a population consisted only of germ cells.	155
Figure 5.21– State transitions according to the molecule concentration.	156
Figure 5.22– Individual and multicellular motifs of the consensual and reversible specialization.....	157
Figure 5.23– Percentage of soma cells over time for a consensual and reversible specialization for two different combination of soma threshold/germ threshold.	158
Figure 5.24– Population fitness for different combinations of soma and germ thresholds. ..	159
Figure 5.25– Colony shots of the four simulated model for phenotypic variability.....	161
Figure 5.26– Representation of the one-signal coordinated signalling circuit and its multicellular implications.	163
Figure 5.27– Temporal shots of the cells oscillating between the A state (blue) and the B state (blue). The graph shows the number of cells in each state over time.	164

Figure 5.28– Transitions between state A and state B for two random cells of the population of Figure 5.27.....	165
Figure 5.29– Convergence to the final fraction of A and B cells is independent of the colony size.	166
Figure 5.30– Percentage of cells in state A (blue) and B (green) for different emission rates and two combinations of reception thresholds.....	167
Figure 5.31– Percentage of cells in state A (blue) and B (green) for different emission rates and two combinations of reception thresholds.....	169
Figure 5.32– Temporal shots of a colony undergoing permanent specialization. The graph shows the number of cells in each state over time.	170
Figure 5.33– Transitions between state A and state B for three random cells of the population of Figure 5.32.....	171
Figure 5.34– Spatial patterns emerge as a consequence of cell specialization. The threshold for A controls the width of the blue ring while threshold for B controls the width of the green ring, and thus, the number of blue rings per colony.	173
Figure 5.35– The spatial pattern of cells in each state is influenced by the global shape of the colony. A cells are depicted in green and B cells are blue, to improve contrast.	174
Figure 5.36– Systems behaviour for different combinations of the emission rate, sensitivity thresholds, diffusion and degradation rates. Cells exhibit five different behaviours depending on the model parameter: permanent state A, permanent state B, oscillations between state A and B, transient DOL and permanent DOL.	175
Figure 5.37– Representation of the two-signals coordinated signalling circuit and its multicellular implications.	179
Figure 5.38– Temporal shots of a colony with the two-signal circuit and parameters from Table 5-6. The graph shows the number of cells in each state over time. The resulting global dynamics are oscillation of cells from state A (green) to B (blue).	181
Figure 5.39– Temporal shots of a colony with the two-signal circuit and parameters from Table 5-6. The production time of molecules A and B has a small variability (20 ± 5 minutes). Spatial patterns of cells diving labour in a permanent fashion appear.	182
Figure 5.40– Temporal shots of a colony with the two-signal circuit and parameters from Table 5-6. The production time of molecules A and B has a large variability (20 ± 10 minutes). Spatial patterns of cells diving labour in a permanent fashion appear with many nucleation points.	184
Figure 5.41– Spatial patterns for different combinations of emission rates and reception threshold. Permanent DOL is achieved for almost all the parameter space.	185
Figure 5.42– Resulting spatial patterns for different diffusion rates of molecules A and B.	186
Figure 5.43– Resulting spatial patterns for different degradation rates of molecules A and B.	187
Figure 5.44– Resulting spatial patterns for different degradation rates of molecules A and B.	188

Figure 5.45– Resulting spatial patterns for different combinations of the ratios of <i>Emission rateSensitivity threshold</i> and <i>Diffusion rateDegradation rate</i> . DOL is achieved for all combinations except when diffusion>>degradation.	189
Figure 5.46– Resulting spatial pattern and number of cells in each state when the diffusion and degradation rates ratio of molecule B is two orders of magnitude lower than that of A.	191
Figure 5.47– Resulting spatial pattern and number of cells in each state when the sensitivity threshold of A and B is different.....	191
Figure 5.48– Final number of cells for populations using each signalling motif.	194
Figure 5.49– Lateral inhibition model. Cell in state A inhibits that state in its adjacent neighbour.	195
Figure 5.50– Turing reaction diffusion model. Activator, A, and Inhibitor, I, interact and diffuse leading to spatial patterns of the concentration of the two molecules.	196
Figure 5.51– Both circuits (one and two-signals) have a long-term repression (mediated through the diffusible molecule) and a short-term induction (mediated by the repression of the opposite state).	197
Figure 5.52–Representative ordered spatial patterns of the one and two-signals motif.	198
Figure 5.53– Repeatability of the emergent spatial patterns is measured by the <i>thickness</i> of the green (T_G) and blue (T_B) regions and, in the case of the two-signals pattern, with the number of nucleation points, N	199
Figure 5.54– Size dependency of the emergent spatial patterns is measured by the <i>thickness</i> of the green (T_G) and blue (T_B) regions and, in the case of the two-signals pattern, with the number of nucleation points, N	201
Figure 5.55– Directed patterns created from placing A emitting cells in specific spatial locations for the two-signals motif with no time variability in the molecule production.	202
Figure 5.56– Directed patterns created from placing A emitting cells in specific spatial locations for the two-signals motif with no time variability in the molecule production.	203
Figure 5.57– Pattern scaling of the one-signal motif for two different regimes. When cells growth rate is faster than the signal exchange the pattern cannot form until later stages. If cells growth rate is slower or similar to the temporal scale of the signal, the pattern scales accordingly although it blurs with size.	204
Figure 5.58– Pattern scaling of the two-signal motif for two different regimes. When cells growth rate is faster than the signal exchange the pattern cannot form until later stages. If cells growth rate is slower or similar to the temporal scale of the signal, the pattern scales accordingly although it blurs with size.	205
Figure 5.59– Comparative table of the properties of the patterns created by the two circuits.	206

List of Tables

Table 2-1 General input parameters and their associated models.....	26
Table 3-1 Comparison of validation parameters for simulated and experimental colony sections.....	35
Table 3-2 Fraction of close neighbours in a group at 200'.	47
Table 3-3. Validation parameters for different bacterial division lengths.	52
Table 3-4 Fraction of close neighbours when group size is 200 cells.	54
Table 3-5. Nutrient consumption parameters.....	57
Table 4-1. General parameters for the public goods dilemma simulations	76
Table 4-2. Properties of the simulated QS molecule	109
Table 4-3. Properties of the selfish public good.	114
Table 5-1. Physical parameters of shared molecules that define the degree of cooperation.	136
Table 5-2. Parameters of the simulation for different phenotypic variation mechanism.....	149
Table 5-3. Initial set of parameters for the one-signal circuit simulations.	164
Table 5-4. Set of parameters for the one-signal circuit simulations.	168
Table 5-5. Parameter for the one-signal circuit.	172
Table 5-6. Initial parameters for the two-signal motif simulations.	180
Table 5-7. Model parameters for the fitness simulations of the one-signal and two-signal circuits.....	193

Chapter 1 Introduction

Charles Darwin wrote a book in 1859 entitled “*On the Origin of Species by Means of Natural Selection, or the Preservation of Favoured Races in the Struggle for Life*” that changed the fundamental paradigm in the understanding of natural life and its origins (Darwin 1859). The theory he proposed -the theory of natural selection- provided the first attempt for the explanation of adaptation and speciation in the natural world. Natural selection theory states that individuals within the same species present small variations in morphology and behaviour, and that these variations are heritable (they can be passed on to following generations). Because of competition between individuals, those who outcompete the rest will leave more offspring, and the characteristic that provided such success will also be inherited by their offspring. And so, adaptation happens over the course of generations. If the environment changes, individuals that do best will then be selected, leading to the observable evolutionary change of species.

During the same time that Darwin's book was released for the first time, Gregor Mendel was conducting the first experiments on heritable features which then conformed the foundations of genetics (Abbott & Fairbanks 2016). It was thanks to the ideas of Mendel on heritability that the great mathematician R.A. Fisher was able to generate a more consistent framework for the theory of natural selection. Fisher showed how natural selection could be described by changes in gene frequencies where genes associated with greater individual fitness will increase in frequency (Fisher 1930). So, the updated theory of natural selection shifted focus from individuals to genes, but the fundamental concept remained: natural selection would favour genes that increase an organism's ability to survive and reproduce (i.e. increase the organism fitness).

Natural selection should then lead to a world dominated by organisms interacting mainly by competition. However, cooperation is found at all levels of biological organisation: from genes cooperating to form genomes, to cells cooperating to make multicellular organisms (West et al. 2007).

Cooperative individuals that perform costly behaviours to help others would experience a decrease in their own individual fitness. Behaviours where interactions between organisms are not pure competition thus need another explanation, which can be provided by social evolution theory.

1.1 Social evolution theory in microorganisms

The theory of social evolution, or its associated field sociobiology, was somehow founded by the paper “*The Genetical Evolution of Social Behaviour*” written by the evolutionary biologist W.D. Hamilton (Hamilton 1964). Before describing Hamilton’s theory and explanation for natural cooperative behaviours, which can be applied for almost all kind of organisms, it is worth clarifying that this thesis will be restricted to the context of microorganism. The fundamental reason is that the thesis hypothesis will be tested and simulated using a computational framework that utilizes agents that resemble individual bacteria. Also, microorganisms are particularly social organisms, naturally displaying a variety of social behaviours (Diggle et al. 2007). Scientists have examined in microbes the evolution of social behaviours because they exhibit the fastest generation times of any independent biological organism, allowing the observation of evolutionary change in real time (Dunny et al. 2008). Finally, microbes can be easily manipulated to create synthetically engineered microbes that have been used to control, design, and study a plethora of ecological interactions (Lindemann et al. 2016; Song et al. 2014).

In order to understand the social behaviour that microorganisms present, it is first crucial to describe the manner in which microbes interact. For a social interaction to take place there is the prior requirement of communication. Therefore, social behaviours are intimately linked to the mechanisms of communication (Keller & Surette 2006). Bacteria can engage in interactions through direct and indirect communication. Indirect interactions take place through the shared environment, for instance, by competing for space and common resources (Lloyd & Allen 2015). Direct intentional interactions can also be propagated through the

environment via the exchange (emission and sensing) of diffusible signals. The recognition of such molecules leads to changes in the gene expression of the receptor cell, where gene regulation occurs only for some threshold concentration. Since this mechanism regulates coordinated behaviour in a cell-density-dependent manner, it has been referred to as quorum sensing, and has been naturally found in both gram-positive and gram-negative bacteria (Waters & Bassler 2005). Additional direct interactions include physical mechanical contact (Blanchard et al. 2014), secretion of toxins (Massey et al. 2004) or contact-dependent signalling mechanisms like cell-surface polypeptides (Kruse et al. 2001). Even contact-dependent transmission of DNA molecules could be archived as a bacterial communication mechanism (de la Cruz et al. 2010). Most of these processes will be described in depth in the corresponding sections as they are mentioned here to illustrate the richness of bacterial communication means.

When microorganism interact with each other, they naturally engage in some type of social behaviour. Social behaviours are those that have fitness consequences on individuals other than the actor (called recipients). Social behaviours have been divided into four classes according to the fitness consequences they have for both the actor and the recipient (Figure 1.1). Behaviours can be positive, if they have a positive impact on the actor fitness, or negative, if the effect is detrimental for the actor fitness. Behaviours that provide a positive benefit to actors are selfish if they have a negative effect on recipients, but mutually beneficial if recipients also benefit from it. Behaviours with a cost to the actors are separated into altruism, when recipients benefit, and spite, when recipients also experience a cost. Cooperation can then be defined as any behaviour which provides a benefit to a recipient, so mutualism and altruism are both considered cooperative behaviours (West et al. 2006).

Figure 1.1– Fundamental social interactions.

An example of extreme altruism in bacteria are *Myxobacteria*. In *Myxobacteria*, nutrient starvation causes masses of individual cells to gather and form distinct complex structures (Figure 1.2). When starvation is prolonged, aggregated groups form fruiting bodies in which only a portion of the population develops into reproductive spores. Some cells altruistically lyse, providing nutrients to other cells, others develop into the non-reproductive stalk that elevates the spores, while others differentiate into resistant and reproductive spores (Muñoz-Dorado et al. 2016).

Other cooperative behaviours in microorganism that present difficulties in the eyes of natural selection are collective cell actions driven by public goods production and division of labour. Public goods are goods that are produced by single organisms, but accessible to other organisms as well. Since public goods provide a benefit for the recipients but their production is costly, preserving them is a social, cooperative behaviour which can be exploited by selfish cells, individuals that benefit from the behaviour without contributing to it (Cavaliere et al. 2017a). The public goods dilemma, namely the conflict between group benefit and personal benefit, is best represented by 'The tragedy of the commons', where selection at the individual level leads to a tragedy at the group level. The tragedy of the commons was a metaphor

introduced to describe the dynamics of shared resources (Hardin 1968). It describes shepherds feeding cattle on a common land, each of them benefiting from adding more animals. However, this behavior leads to overexploitation and destruction of the common resource. This metaphor has been applied in evolutionary biology to describe cases where individual selfishness lowers the mean fitness of all individuals in the group (Drescher et al. 2014).

Figure 1.2– Illustrations of different Myxobacteria fruiting bodies from Roland Thaxter, distributed under a CC-BY 2.0 license.

On the other hand, division of labor happens when different individuals of a population engage in different, complementary tasks. It often includes altruistic behaviours, where some individuals reproduce less or not at all. Division of labor happens in some microorganisms, as in the above example of the Myxobacteria, where fruiting bodies allow for more efficient spore dispersal, but also require some cells to die for them to be constructed (Velicer & Vos 2009). Further description of these actions will be provided in the course of this thesis as they conform central topics of chapter 4 (public goods dilemma) and chapter 5 (division of labour).

These actions are difficult to explain by natural selection as they reduce the relative fitness of individuals performing them. Theories of sociobiology are thus concerned with explaining how and under which conditions cooperative behaviours can occur instead of selfish,

competitive behaviours predicted by natural selection. Principal sociobiological theories including inclusive fitness theory, multilevel selection theory and evolutionary game theory will be now described.

Inclusive fitness theory (Kin selection)

Hamilton was the first to offer a formal extension of natural selection to social traits, explaining how apparently contradictory behaviours, in which one individual reduces its fitness to help others can evolve by natural selection. Hamilton proposed a theory, named Kin selection, that argued that, because other members of a population may share one's genes, a gene can also increase its evolutionary success (i.e. fitness) by indirectly promoting the reproduction and survival of other individuals who also carry that gene (Hamilton 1964). The theory is mathematically represented in what is called the Hamilton's Rule. Hamilton's rule states that a cooperative behavior will be selected, if the cost c of the cooperative behavior is less than the product of the benefit b of the trait multiplied by the relatedness r between actor and recipient ($rb - c > 0$). The term relatedness, r , is a comparative coefficient that describes the genetic similarity between actors and recipients, relative to the global similarity of individuals in the population. It is important to note that relatedness refers to the genes encoding the cooperative behavior that is under selection in a cooperative environment.

In summary, inclusive fitness refers to the sum of the two sources affecting the fitness of an individual: individuals gain inclusive fitness directly through their impact on their own reproduction (direct fitness effects) as well as through the reproduction of related individuals (indirect fitness effects), therefore providing an explanation for altruistic behaviors.

Multilevel selection theory

The original natural selection theory treats the individual organism as the basic unit of selection. This means that the variability, even if registered at more fundamental levels, is competing to be selected at the individual level. Yet, the concept of individual can sometimes be an abstract and general notion. It is true that natural selection requires only an entity that

has the properties of variation, reproduction and heritability (Lewontin 1970). So, it has been argued that selection does not necessarily have to apply only to the traditional level of individual organism, but to any level that meets the aforementioned requirements (Okasha 2011).

Multilevel selection theory (MLS) assumes that selection forces can act at many levels of organization, including the *group* level in an ensemble of cells for example. The MLS approach tries to determine at which level is selection stronger (individual or group), and thus, driving the path to evolution. MLS explains cooperative behaviours as behaviours that maximize the relative fitness within the social group, and that they can only be accounted for if natural selection takes place at more than one level of the biological hierarchy. Therefore, selfish individuals might outcompete altruists within groups, but internally, altruistic groups outcompete selfish groups (Wilson 1975; Wilson 2007).

The mathematical framework that supports MLS was derived by George R. Price in an equation named the Price equation (Price 1970). The Price equation describes how a trait or gene changes in frequency from one generation to the next. It is derived from a mapping between individuals in parent and offspring generations and can be applied to any given trait. If it is applied to groups when the parent and offspring entities are social groups, the Price equation provides separated information of the selection force at the group level and the individual level. The relative strength of each component will determine the final direction of evolutionary change.

However, if the equation is used in individuals of the same social group, the price equation simply results in the Hamilton's rule (Damore & Gore 2012). This has been used by the detractors of the MLS theory to state that such approach is just a more complicated way to expose the same principles as the fitness inclusive theory. Detractors claim that MLS does not provide any different predictions from the inclusive fitness theory and that is another way of looking at the same dynamics (West et al. 2007). Therefore, even if many acknowledge that there is some between-group selection, some claim that it cannot be considered a major force in driving evolution (Smith 1976).

In summary, MLS ascribes the existent cooperative behaviour to an increase in the group fitness, that should only be accounted for if there is selection between-groups and not only within groups. Yet, elucidating at what level(s) selection takes place is still currently one of the main debates in evolutionary biology (West et al. 2007; Kramer & Meunier 2016).

Evolutionary game theory

A different theoretical approach that has also been widely used in the study of social behaviours in microorganisms is evolutionary game theory (EGT). EGT is a mathematical framework rather than a sociobiological theory constructed for the explanation of cooperation. Yet, it has been useful in providing explanations for altruism in different contexts (Nowak 2006; Prajapat et al. 2016; Tarnita et al. 2011; Tudge et al. 2016; Axelrod & Hamilton 1981).

Classical game theory is used to describe the interaction of two or more players in which the outcome (payoffs) of each player depends not only on the own decision, but also on the strategies adopted by the co-players (Von Neumann & Morgenstern 2007). In EGT the strategies are not associated to rational choices, but are traits encoded into inherited programs that can be passed to the offspring (for this reason, the terms trait, strategies and behaviors are used in an indistinguishable manner). Briefly, EGT is the study of frequency dependent selection, where the success of a strategy depends on the fraction of individuals in the population that is using that strategy given distinct assigned payoffs to each pair-wise social interaction. When players (individuals or cells) are left to reproduce according to their payoffs, a certain strategy can end up dominating the ensemble. If the strategy becomes stable for evolutionary times and cannot be outcompeted by any other alternative strategies, it is said to be an evolutionary stable strategy (ESS). The extension of the classical framework and the term ESS were both provided by John Maynard Smith and George Price in 1973 in the paper “*The logic of animal conflict*” (Smith & Price 1973). Note that EGT can provide much more than just a justification for the evolution of cooperation since its logic can be used in both inclusive fitness and multilevel selection theories (Foster 2011).

So, the revised theories and frameworks have helped to shade some light into the mystery that is the evolution of cooperation in the context of natural selection. Still to this day, the attempt to unite both terms (cooperation and natural selection) keeps on producing a vast amount of experimental and theoretical work (Fletcher & Doebeli 2009; Foster & Bell 2012; Mitri et al. 2011; Wilson 2007; Hol et al. 2013).

1.2 Multicellularity and division of labour

According to the Kin selection theory followers, individuals are selected to maximize their inclusive fitness and not the overall group fitness (West et al. 2007). However, there can be cases where the interest of individual cells aligns, and thus, group and individual fitness concur as well. In those occasional cases, conflict within those groups had to be reduced to a minimum (Rainey & De Monte 2014). As it was mentioned, group conflict is a consequence of the competition that individuals sharing the same niche experience.

So, for cooperation to succeed as an evolutionary strategy within all the members of a group, the negative effects of conflict among individuals must somehow be either prevented or mitigated (Travisano & Velicer 2004).

Even if rare, the alignment of inclusive and group fitness leads to huge evolutionary consequences, namely, major evolutionary transitions (West et al. 2015; Szathmáry & Smith 1995). A major evolutionary transition occurs when a group of individual entities that was capable of survival and reproduction autonomously before the transition, can only replicate as a greater unit after it. The major transitions in the evolution of life result in the emergence of new kinds of *entities* that follow the hierarchy principles of biological organization (Solé 2016; Fisher 2011). First, replicating molecules turned into populations of molecules, then into chromosomes and then into simple prokaryotes cells. Prokaryote cells lead to eukaryote cells and those could later arrange into colonies of more complex behaviour (Kirk 2005).

As commented, any major transition in evolution requires mechanisms to mitigate the negative effects of competition among individuals within the new cooperative social unit. Even simple evolutionary transitions to primitive forms of cooperation can, in principle, be undermined by the appearance of selfish, competitive individuals. In the transition from fully independent single cells to cells with strong associations, increasingly complex strategies to reduce conflict and avoid exploitation by selfish cells must occur (West et al. 2006). If the major transition happens in a population that involves relatives (*i.e.* clonal groups), phenomena like eusociality in insects or multicellularity arise (Nowak et al. 2010). Clonality can be viewed then, as a mean to mitigate individual competition. It has been suggested that clonality within groups reduces the likelihood that selectively important differences among cells will arise and promotes cooperation among cells to enhance the reproductive output of the group due to Kin selection (West & Cooper 2016).

The evolution of multicellularity is characterized by the integration of lower-level units into higher-level entities and hence is associated with a transition in individuality, from solitary individuals to colonies (De Monte & Rainey 2014). In a broad sense, an organism is considered multicellular if the cells that comprise it are not independent of each other but rather form a group that coordinates and communicates, with a resulting behaviour that is determined by the interaction of its fundamental elements (Shapiro 1998). It has been identified that multicellularity has evolved independently at least once in animals, three times in fungi, six times in algae, and multiple times in bacteria (N. a Lyons & Kolter 2015). This has involved different developmental pathways. Not all organism played the transition to multicellular entities the same way, however, some requirements (like the conflict resolution) are essential to meet the current definition of what constitutes a multicellular organism.

The main evolutionary requirements for multicellularity include the formation of clusters of cells via physical contact or adhesion and the maintenance of cooperation among the individuals (Niklas 2014). Finally, maybe as a consequence or as a requirement, specialization of some cells of the population occurs (Tudge et al. 2013; Ispolatov et al. 2012; West &

Cooper 2016). Cells perform different tasks so that the whole group becomes dependent on all its components (Figure 1.3).

Figure 1.3– Steps for the evolution of multicellularity. Cell aggregation and conflict resolution (cooperation) are major requisites to constitute a multicellular organism.

Multicellularity provides several benefits such as the facilitation of colonization of new territories, predation avoidance, resource acquisition, metabolic efficiency, higher resistance to physical and chemical stresses and, finally and more importantly, it provides the opportunity for division of labour (Bonner 1998). Division of labour requires the coexistence of interacting cells that differentiate in order to carry out complementary tasks. Accordingly, differentiation refers to a process where the individual transitions from a state where many functions could be performed, to a configuration where it mainly develops a specific task (Kolter et al. 2015a). The specialization of some individuals of the population to perform different complementary tasks usually implies that the whole group becomes dependent on all its components (Kim et al. 2008a). Extremely, this means that individual cells renounced their capacity to reproduce as independent units and came to reproduce as part of a larger whole (Rueffler et al. 2012).

1.3 Ecological and evolutionary feedbacks

Natural selection theory states that the units of selection (that must be able to variate, reproduce and transmit variations) will compete with each other while adaptation happens over the course of generations. Still, evolution requires adaptation not only to the physical

environment, but also to the social environment as neighbours shape the ecological landscape of a given ecosystem (Levy & Borenstein 2012). Previous sections have focused on the act of cooperation and its consequences (e.g. multicellularity) from an evolutionary perspective, however, the ecology of social organisms has not been discussed. An understanding of the interaction of the population and its environment is critical, as ecology impacts on every component of inclusive fitness.

The interplay between ecological and evolutionary dynamics usually results in changes in the composition of the individuals (i.e. traits, behaviours) of a community over time (Watson et al. 2016). Because of social interactions, certain strategies, such as selfishness and altruism, are selected for or against, resulting in rapid changes (from an evolutionary perspective) in the frequency of the individuals carrying the trait, influencing therefore, the ecology of the global community. These changes in the ecology can then feed-back on the selective advantage of the different traits, leading to an eco-evolutionary feedback (Cavaliere et al. 2017a). For example, the prevalence of a particular trait may be strongly influenced by the existing ecological relationships in the community, and at the same time, it may also dominate the selective pressures on an evolving population within that community. By evolving its ecological relationships, a community may modify its ecological context, and hence, the selection it experiences over subsequent generations (Moreno-Fenoll et al. 2017). For example, in a population of both cooperators and cheaters, cheaters are evolutionary favored because of their higher reproductive rates. However, because of their dependence on cooperators (e.g. on their secreted good), the ecological context will favor cooperators as they are self-sufficient and better adapted to the ecological context. This is an example of how ecology can influence evolutionary trends, and vice versa (Sanchez & Gore 2013).

In summary, explicit consideration of ecological processes can strongly affect the conclusions of a social evolution model. Most importantly, individuals play their behavior on an ecological scale and the survival options of every individual will be affected by its habitat and its access to resources.

1.4 Individual-based models

The complex evolutionary and ecological dynamics of a given group can be mathematically analyzed with EGT, which unlike Kin selection and MLS theories, is a framework that allows to model ecological interactions, namely, social interactions with other individuals in a fixed or changing environment (Liao & Tlsty 2014). Although the dynamics of an EGT model can be studied analytically when the set of strategies is small, EGT equations describe an idealized situation that assumes perfect mixing of populations (*i.e.* no spatial discrimination), deterministic games and infinite population size. So, due to the large number of interactions taking place in microbial communities it is preferable to simulate the dynamics of the community using agent-based model approach (Adami et al. 2016).

Agent or individual-based models (AbMs or IbMs) can be conceived as a mechanistic representation of evolutionary games where the replication, interactions and death of individual agents are explicitly simulated using a system updated by a series of discrete events. These types of models also include the possibility of adding mutations that can introduce novel strategies not yet present in the species, which can be used to simulate random evolution of members of the community (DeAngelis & Mooij 2005). Besides, one fundamental feature of the ecological environment cannot be account for using classical EGT models, this is, the spatial structure and the aggregation of cells, which can be better understood using models that explicitly simulate the dimension of space (Quaghebeur 2017). Partial differential equations and IbMs are the most commonly used approaches to study the spatial properties of cell colonies.

Partial differential equations aim to reveal the general dynamics of a population, following global equations. Alternatively, IbMs derive the global dynamics of the population from local interactions rules described for every single individual. The behaviour of the population then emerges as a result of the interaction between these individuals and the described environment. This feature is of crucial importance because it avoids the definition of the relation between the individual scale and the population level which are required to be

explicitly described by the fitness function of each mathematical model (Breckling et al. 2005).

Most IbMs automatically simulate birth, death and competition among the agents. This added to the specification of the individual traits and the heritability of those traits, natural selection is, then, simply an emergent process in IbMs (Kreft et al. 2000).

A major challenge in the study of complex systems, like microbial communities, is to understand how apparently organized collective behaviour emerges out of the small-scale interactions between the individuals. Complex system research tends to adopt a bottom-up approach, describing types of individuals and environments and then simulating what kind of complex dynamics emerge from the whole system. So, if one is interested in understanding processes such as the evolution of multicellularity, an IbM should be the framework of choice, since such transition is an emergent property that results from a collective behaviour (Hellweger et al. 2016).

In spatially structured environments, like microbial communities, physical cell interactions can generate feedbacks between the emerging spatial structure and the social relation of the individuals. Therefore, IbMs (also named cell-based models in this context) are well suited for simulating multicellular systems such as cell colonies or tissues as more complex behaviours can be achieved. The specific advantages of the IbMs can be summarized as:

- (i) IbMs explicitly simulate the individual allowing to include variability at the individual level.
- (ii) A ruled-based description allows for more realistic assumptions for the model of the individuals than population models.
- (iii) IbMs account for individual adaptive behaviour to their environmental conditions, so the evolution of the whole system arises from the dynamics that govern individuals in their pursuit of optimal fitness.
- (iv) IbMs can simulate phenomena at different temporal and spatial scales. By coupling various model scales, it is easier to explain the complex interactions

underlying the emergent behaviour, analysis that cannot be achieved by examining individual system components or scales in isolation.

- (v) Simple individual rules can give rise to non-explicitly defined large-scale phenomena.

IbMs have been traditionally used in ecology since 1970, and recently have shifted focus into microorganisms, as a powerful way to study biofilms (Hellweger & Bucci 2009). Here, the agent is the idealized microorganism that responds to its local microenvironment, which is continually modified by other agent-cells as they consume and secrete different solutes (Prats et al. 2006). Environmental heterogeneities are tracked by iteratively solving reaction–diffusion equations that describe solute concentration gradients in relation to bulk transport and consumption or secretion within the community. This approach is powerful for exploring questions about biofilm structure and composition, and the results are often supported experimentally (Kreft et al. 2013). There are a number of relevant examples of individual-cell frameworks that are worth naming such as CellModeler (Rudge et al. 2012), BactoSim (Prestes García & Rodríguez-Patón 2015), iDynamics (Lardon et al. 2011) and *gro* (Gutiérrez et al. 2017; Jang et al. 2012). All of them have been developed with the purpose of describing and representing the behaviour of bacteria, but present differences in characteristics such as the scope of the simulations, the levels of abstraction allowed or their capacity to be extended. A complete review in those frameworks can be found in (Gorochowski 2016).

However, despite having been long exploited to model ecological systems, the use of IbMs in the simulation of multicellular related phenomena has been greatly underused (DeAngelis & Grimm 2014). Some possible explanations include the complexity of developing robust, verified IbMs, a process that is costly and time-consuming (Zomorodi & Segré 2016). IbMs usually require a large number of parameters, where variation in their numerical values can strongly influence the resulting predictions. Because the precise value of some parameters should be extracted by individual measurements of the cell processes, IbMs are often fed with interpolated data from global phenomena. This constraints the prediction capacity of the simulator and requires extensive scanning of the parameters space, similarly to a sensitivity

analysis in equation-based models (Prestes García & Rodríguez-Patón 2016). Large-scale microbial simulations where one typically deals with a considerably large number of individual cells, are also a challenge to overcome. The problem becomes even more complicated if spatial heterogeneity is considered, especially in two- or three-dimensional discretized space. Due to these limitations, IbM studies that include spatial heterogeneity have been so far only applied to small-scale (from micrometers to centimeters) environments (van der Vaart et al. 2015).

IbMs can be critical in addressing a variety of ecologically and evolutionary relevant questions such as quantifying the impact of inter-species interactions and environmental factors on the emergence of cooperation, coexistence of cooperators and cheaters and the evolutionary transitions from individual cells to multicellular communities (DeAngelis & Mooij 2005).

1.5 Questions addressed in this thesis

The aim of the introduction was to provide some context on the questions and methodologies that this thesis will cover. As a guideline, this thesis has been inspired by the conditions that are required for multicellularity to emerge. Briefly, the process of group formation and the mechanisms for conflict resolution and the maintenance of cooperation. Such processes will be simulated by means of the individual-based simulator *gro*, whose logic and inner assumptions will be further detailed in the methodology chapter. Despite treating with evolutionary problems, the time scale of simulations restricts the results to the ecological context. However, because of the discussed feedback between ecology and evolution, relevant evolutionary trends can be withdrawn. In this thesis, simulations will assume different ecological contexts, and results will indicate the possible evolutionary tendencies according to their fitness evolution during the simulated experiments.

Specifically, I plan to determine the general necessary conditions for a genetic trait (in the form of a strategy or behaviour) to adapt (*i.e.* increase its frequency) in the context of competition or cooperation with other traits. The global aim of this thesis is not to offer specific concrete predictions but rather understand the essential general properties of a cooperating collective system using an Artificial Life approach, and so, determine the minimal conditions for the evolutionary transition of multicellularity (Taylor & Jefferson 1993; Bedau 2003).

The thesis is organized in six chapters with an introductory chapter (**Chapter 1**).

In **Chapter 2**, I describe the global methodology I have intended to follow throughout this thesis. Specially, I describe the operations and main biological functions of the cell-based simulator *gro* following the ODD protocol.

Chapter 3 studies the effect of spatial constraints in growth and division and how they affect the process of group formation. Environmental and physical conditions such as resource limitation and cell size are varied to study their impact on the stability of contacts within a

group. Since *gro* does not allow the simulation of cell adhesion, I look for the scenarios that emulate the relevant features of the process of group formation. For that, I define a new metric that follows the number and duration of contacts of a given clonal cell line. I argue that this metric is crucial because it describes group stability and it can help determine whether cooperation could be maintained in more complex configurations (as in chapter 4).

In **Chapter 4**, I test several mechanisms and environmental conditions that have been reported to prevent defectors from overtaking a population of cooperators in a public good sharing scenario. An analysis of the partial and global fitness of each simulation mechanisms is followed by a transversal comparison that provides a qualitative understanding of the relevance of each mechanism in a controlled environment.

Chapter 5 covers simulations related to division of labour and the rules required for its emergence in a population of equal individuals. For that, first, I study the required benefit that division of labour must provide in order to be an evolutionary stable strategy. Then, I simulate possible sources of variability in a homogeneous population, a process which is also required for division of labour. Finally, a novel cell-based model that allows for the emergence of division of labour without explicit fitness is presented.

Finally, most relevant conclusions and future lines are reviewed in **Chapter 6**.

Chapter 2 Methodology

This thesis makes use of the individual-based model (IbM) of *gro* (Jang et al. 2012). As commented in the introduction, IbMs explicitly treat each individual cell as a discrete independent entity that interacts with other individuals and with its surroundings. They allow the simulation of spatially heterogeneous environment and permit the explicit modelling of individual variability (e.g. growth rate, substrate uptake, secreting substances, genetic mutations, etc.) which is central to the study of the emergence of division of labour and multicellularity.

Population behaviour is a consequence of the local interactions of the individuals and the selective pressure exerted by the environment. Given a specific configuration (*i.e.* some specific ways for the agents to interact) and environmental constraints, the strategy that would succeed in most individuals of the population is, arguably, the one that would maximise their reproductive success. Throughout this thesis, the simulation of different individual strategies will test their fitness benefits. For that, the reproductive success of the individuals carrying each trait (having a fixed strategy) will be used as an indicator. Then, a comparison of the reproductive success (which translates as the number of individuals) of each strategy follows.

The simulation workflow (Define-Simulate-Analyse) can be detailed as:

1. Define an initial configuration of the simulation in terms of:
 - a. Environmental conditions.
 - b. Internal conditions: define the strategy to analyse for each individual.
2. Run several simulations enough for statistical significance.
3. Analyse data outputs in relation to the initial configuration: look for the most successful strategy given the initial conditions and constraints in terms of the reproductive success. Comparison can be performed within the same configuration and across different configurations.

2.1 ODD description of the simulator *gro*

Here, a description of the extended version of the *gro* simulator is provided using the ODD (Overview, Design concepts and Detail) protocol. The ODD protocol provides a standard template for model specification (Grimm et al. 2010). An extensive explanation of the underlying algorithms of the model is described and justified in detail elsewhere (Gutiérrez et al. 2017), and empirical tests have demonstrated the framework's ability to make accurate predictions for real biological systems.

1. Purpose

The purpose of this model is to simulate the growth of bacterial colonies consisted of rod-shaped cells in a two-dimensional surface. Originally, it was conceived to prototype engineered behavior in bacterial cells (Pascalie et al. 2016).

2. State Variables and Scales

The model is a two-dimensional representation of a biofilm and comprises two entities: bacterial cells and its environment. The two-dimensional environment represents an infinite petri-dish discretized in equally-sized square units. The grid serves as a scaffold for the spatial localization of each cell. It also handles the diffusion of signaling and nutrient molecules. Bacterial cells are represented as rod-shape rigid bodies. The main state variable, that characterizes each cell, is an array of the protein expression levels associated to the bacterial genetic material. Proteins are digital abstractions that can only be in two states: expressed (1) and non-expressed (0). The behavior (*e.g.* growth rate, emission rate, nutrient uptake rate...) of every cell is then determined by the cell state, which is in turn, defined by the protein array. In this thesis, only the behavior of cells is considered, and therefore, the protein level is obviated. Proteins are required for the control of the cell behavior but are used only as a mean to simulate the desired behavior. Other important state variables include the cell position in the grid, its orientation and its length.

Time is discretized with a constant time step, $\Delta t = 0.1s$. Starting from a single bacterium, colonies of 10^5 bacteria are reached in 9 min in a standard computer. Since grid cells have a size of $3 \times 3 \mu m$ in an infinite domain, and almost 10^6 cells can be reached, colonies diameter ranges from the hundredths of microns to few millimeters. Bacterial cells have fixed diameters ($\sim 1 \mu m$) and variable lengths (2-6 μm).

3. Process Overview and Scheduling

At every discrete time step, every cell in a random order updates its state, as it is shown in Figure 2.1. The order of execution of this process is shuffled to avoid any bias due to a purely sequential execution. After updating each cells state, the environment is resolved. This way, the physical contacts emerging from cell growth are computed and the nutrient space is also updated according to the previous consumption from cells. Finally, the signals grid is updated following diffusion and degradation of the remaining molecules.

Figure 2.1– Process workflow per simulation step in *gro*.

4. Design Concepts

- **Emergence:** spatial patterns of bacteria emerge from local interactions. Specific cell arrangement emerges from local intercellular competition for space while cells are growing. Collective behaviors, such as division of labour (chapter 5) emerge as a consequence of bacterial communication.

- **Adaptation:** individual cells exhibit fixed traits or behaviors, but inherent competition derives in the adaptation of the fittest. Besides, cells can adapt their position and growth according to the local availability of space and nutrients.
- **Fitness:** when the final output is the number of individuals performing a given strategy, their fitness is considered implicitly as an emergent property of the competition between individuals. Furthermore, some simulations explicitly impose a fitness penalty depending on the strategy (*e.g.* produce certain molecule or conjugate a plasmid).
- **Sensing:** the simulated cells are able to sense and absorb the nutrient concentration that is placed in the cell-grid where they are located. If given the instructions, cells can also sense and absorb extracellular diffusible molecules such as quorum sensing molecules, essential metabolites or toxins. In addition, microorganisms detect overlap with neighbouring cells due to cell growth.
- **Interaction:** bacterial cells interact with their nearby individuals for nutrient access, growth, molecule exchange and plasmid transfer.
- **Stochasticity:** there are numerous sources of stochasticity, including: (i) the initial cell locations are randomly chosen within a particular region, (ii) initial cell lengths are randomly chosen around an average value, following an uniform distribution, (iii) the cell division length is chosen randomly around an average division size, following an uniform distribution, (iv) upon cell division, daughter cell sizes are chosen stochastically around equally sized daughter cells, (v) also upon cell division, daughter cells are oriented in a random direction, (vi) the order in which agents are updated during a single global time step is made randomly, (vii) the amount of nutrients that cell consume is randomly chosen around an average value, (viii) the search for a potential recipient cell when transferring a plasmid via conjugation, and the specific moment of transfer is randomly calculated, and finally, (ix) the digital protein expression that controls the behavior of each agent has an associated expression and degradation time and variability.

Also, the probability of a malfunctioning promoter can also be included in the experiment description.

- **Observation:** at each time step the state variables for bacteria are recorded and can be saved into an external file when specified.

5. Input data

A list of the most relevant parameters for the specification of a simulation is in Table 2-1.

6. Initialization

The specific behavior of each cell type (strain or behavior) must be specified in a standardized format. Once rules and parameters are established, the initialization of the experiment is required. The number and type of cells, the spatial coordinates (x,y) on the plate and a dispersion area can be specified. This allows to place randomly each cell type in a particular surface area, creating various initial seed densities (Figure 2.2).

Figure 2.2– Different initial cell densities.

7. Submodels

a. Signaling model.

Signals in *gro* are implemented through a set of grids (one for each signal) that store the signal concentration at each grid location. At each time step of the simulation, diffusion and degradation are applied to update the concentrations of the signals over the whole grid. The dynamic-sized grid automatically resizes whenever it is needed. Signals are not initialized in the environment but rather, they are created by every

individual cell. Cells that are programmed to secrete a certain signal, will do so every simulation time step. The signal is initially placed in the grid cell under the secreting cell. The signal diffuses and degrades in the grid until another cell absorbs it. Cells can be programmed to absorb certain signalling molecules that later trigger the expression of certain inner proteins. This process emulates the process of quorum sensing where reading cells check for a given amount of signal concentration and consume it from the environment. If the amount of absorbed signal is over a given threshold set by the user, it triggers the expression of a binary protein. This protein then integrates into the rest of the protein specification.

b. **Growth models**

i. Unlimited exponential growth model

gro imposes an exponential growth for each bacterium described by:

$$\frac{dL}{dt} = k_{basal}L$$

where L is the length of the bacterium and k_{basal} is the basal growth rate, which is an input parameter required in the simulation description.

ii. Nutrient limited growth model

The growth rate of a given bacterium can be further modulated by external factors such as nutrient availability, waste accumulation or even pressure. So, to simulate a more realistic limitation of growth, a Monod-based growth model is implemented in *gro*. The model modulates the basal growth rate, k_{basal} , as a function of nutrient availability near the cell, following the equation:

$$k = k_{basal} \cdot N = k_{basal} \cdot \frac{S}{S + \kappa}$$

where S is the amount of nutrients available in the cell surroundings and κ is the half saturation constant (which is also an input parameter). The consumption rate of the nutrients, their initial amount and their distribution in the environment are input model parameters that have to be specified in the simulation description. Nutrients are

consumed by the cells from a grid that stores the amount of nutrient at their specific given location. The grid is dynamic in size, extending its size whenever a bacterium reaches any border of the grid.

iii. Cross-feeding/Toxins model

The absorption of specific signals can also modulate the growth of individual cells. These dynamics seek to simulate cross-feeding and toxin warfare mechanisms, respectively. Cells that are programmed for the absorption, consume a given amount of molecule from the environment (if possible). The effect of absorbing the molecule affects the growth rate of the bacterium in two different ways: either it allows its normal growth (cross-feeding) or it prevents it (toxins). If the interaction type is set to positive (+), then bacteria would need to absorb a given amount of signal to grow normally: the amount of absorbed signal directly modulates the bacterial growth rate (simulating a cross-feeding interaction between the producer cell and the receptor cell). On the other hand, if the interaction type is set to negative (-), bacteria are still forced to absorb, but now the signal effect is detrimental, slowing down the growth of the cell (simulating the obligated effect of a toxin in the environment). The effect of such toxin or required metabolite is encapsulated in a coefficient, C_r , that is calculated according to the molecule availability following the function:

In summary, the growth rate of each bacterium at each time step can be calculated as:

$$k = k_{basal} \cdot N \cdot C_r$$

Where k_{basal} is the basal growth rate, N is the monod-based nutrients coefficient and C_r is the coefficient that reflects on the effect of toxins or cross-feeding molecules in the environment.

c. Conjugation model

The conjugation method is programmed as an action triggered by the expression of the proteins involved in the conjugation process (*i.e.* conjugation proteins). When the conjugation proteins are expressed in a given cell, there is a probability, p that the conjugative plasmid transfers to a random neighbour bacterium. This probability is given by the specified conjugation frequency, γ , and the generation time of the cell, G , in this way:

$$p = \frac{\gamma \Delta t}{G}$$

The conjugation probability is calculated for every conjugative plasmid. The probability value is density-independent, arising only from the inherent frequency rate of the plasmid in the host. This rate is given by the average conjugation frequency per life cycle γ and by the bacterial performance (in form of its instant growth rate). Once the conjugation probability p is calculated, a uniform distributed variable decides if the plasmid gets transmitted. If so, a second random variable defines which of the neighbours receives the plasmid.

<i>Parameter</i>	<i>Unit</i>
Growth model	
Basal growth rate	$\mu\text{m}/\Delta t$
Division length mean and standard deviation	μm
Nutrient model	
Initial concentration	Molecules/grid cell
Consumption rate	Molecules/grid cell $\cdot \Delta t$
Signals model	
Diffusion rate	Molecules $\cdot$ grid cell/ Δt
Degradation rate	Molecules/ Δt
Conjugation model	
Conjugation frequency	Events/generation time

Table 2-1 General input parameters and their associated models.

Chapter 3 Group formation

It has been argued that one of the first requirements for the emergence of multicellularity is the formation of clusters of adhered cells (Niklas 2014; Abedin & King 2010; Gumbiner 1996). The process of cell aggregation can arise through two distinctive means. Either cells stay together after division and thus create clumps of clonal individuals, or cells come together, independently of their genotype, and stick to each other (Tarnita et al. 2013; Bonner 1998; Koschwanez et al. 2011). Wherever the path, the process of group formation is fundamental for the subsequent steps that lead to the emergence of multicellularity. Cooperation among the individual cells requires some sort of communication: cells interact through the exchange of specific molecules. Truthful communication arises when both the transmitter and the receiver can exchange messages for enough time to allow cooperation to evolve into its extreme version: specialisation (Asfahl & Schuster 2016; Cavaliere et al. 2017b). This is what the attachment of cells permits; it provides stability for cooperation to be reciprocal and consistent. However, how exactly is the process of adhesion, from the mechanical and molecular standpoint, can occur on many different ways: aggregation within an extracellular matrix, filamentation by incomplete cell division, or direct contact mediated by transmembrane proteins (Claessen et al. 2014; Lyons & Kolter 2015). However, cell coordination is also possible without adhered group formation, as in the quorum sensing systems found throughout the microbial world.

Group formation is bounded by the environment it is embedded. Spatial constraints will dictate the shape and global morphology of the aggregate (Lloyd & Allen 2015). In the case of microorganisms growing on a plate, how nutrients are dispensed from the plate surface and gravitational effects produce that some colonies form planar, almost two-dimensional structures. The mechanical interactions of growing cells shape the organization and disposal of individual cells that, despite the spatial constraints, can form complex assemblies.

In Figure 3.1, two shots of *Pseudomonas aeruginosa* colonies are shown as an example of this phenomenon.

Figure 3.1– *Pseudomonas aeruginosa* colonies form complex structures. Images from Scott Chimileski, microbephotography.com

In nature, microorganisms are commonly found in surface-associated communities known as biofilms (Flemming et al. 2016). Biofilms are microbial communities surrounded by an extracellular polymeric substance (EPS) matrix that confers cell-cell and cell-surface adhesion properties (Hall-Stoodley et al. 2004). Biofilms have been described as multicellular entities that conform small-scale ecosystems in which a variety of cells, sometimes from different species, cooperate to obtain resources, such as food and space to grow (Kolter et al. 2015b). Biofilm formation is commonly viewed as a cooperative enterprise, where bacteria work together for a common goal. Apart from the secretion of the EPS, some communities of cells can manufacture large volumes of lubricants that allow a colony to swarm over soft surfaces, and some can produce iron-scavenging molecules which are low-molecular-mass compounds that bind and transport iron, a frequently limiting nutrient in some environments such as the human host. The physical aggregation of cells in these complex structures provides protection for cells living at the centre of the structure from attack by predators and it also provides a controlled environment for their residents in the face of fluctuating and harsh

conditions facilitating the development of complex interactions between individual cells (Dunny et al. 2008). The proximity in which cells reside within the biofilm and their social nature provides cells with the opportunity to specialize in a particular task, such as acquiring food or acting as a reservoir of genetic material from which to regrow the entire structure. Cells divide the labor of maintaining the colony and, sometimes, differentiate into specialized forms for their function. For example, some cells of the bacterium *Bacillus subtilis* secrete extracellular matrix and anchor in place, while others release spores that allow them to colonize different places and survive difficult conditions (Lyons & Kolter 2017).

Either in naturally existing biofilms or in laboratory plates, cells are constrained in space and influence each other in a distance-dependent manner. A fundamental factor for the survival of a microbial cell is the nature and number of its neighbour cells. These neighboring cells ultimately determine the signaling molecules or toxins it will perceive, and, most fundamentally, its access to resources. Therefore, the most basic and fundamental way of interacting is through direct contact. Mechanical contact is what will then determine the relationships that are established within a group of cells. Spatial constrained growth limits the number of neighbours, a factor that will also be influenced by the geometric shape of the bacteria in question. This, in turn, influences the ability of the group to form stable contacts where cooperation could arise.

Objectives of this chapter:

Here, I present a systematic survey on the spatial structures of simulated bacterial communities. Various initial environmental conditions are tested for colony simulation and analysis. I look for the scenarios that emulate the reported process of group formation. I specifically investigate the impact of bacterial shape and resource limitations in the stability of physical contacts among cells.

In the first section, *Properties of the simulated colony*, I describe and validate against experimental data some of the physical properties of the simulated colonies *in gro*. I

investigate the arrangement and packaging of cells by quantifying the average number of neighbours, the global density and the relative alignment of cells.

In the second section, *Group formation and space structuring*, I simulate growing colonies and analyze its structure with a proposed metric: *physically connected group*. The *group* is given by the physical contacts that a single random bacterium and its descendants establish throughout their lifetime. The metric measures the number of individuals in a cluster and their nature, differentiating between clonal cells and close not related contacts. A statistical investigation of the resulting patterns follows, demonstrating the metric usefulness as stability indicator.

The third section, *Bacterial size and its effect on group formation*, analyses the effect of cell size in the number and type of contacts defined by a *physically connected group*. It studies how shape influences the global spatial structure and considers shape as a mechanism to maintain stable connexions.

In the fourth section, *Resource limitation and its effect on group formation*, colony simulations enforce a scenario where resources are highly limited. Space structuring and group metrics are analysed accordingly.

Finally, the section entitled *Size competition in resource limited environments* compares and measures the fitness advantage that size provides in scenarios where nutrients are in shortage.

3.1 Spatial properties of the simulated colony

The *gro* framework considers microbial communities as collections of elongating cells on a hard surface. Cells grow and divide, causing the colony to expand in space. *gro* includes a self-implemented physics engine, called CellEngine. CellEngine is a two-dimensional physics engine used for real-time calculation of rigid body dynamics. It has been specially developed to resolve the collisions of rod-shaped bacteria growing in non-inertial systems (Figure 3.2). A more detailed description of *gro* can be found in the Methodology (chapter 2) and in a recent publication (Gutiérrez et al. 2017). Here, I present some empirical tests that demonstrate the framework’s ability to make accurate representations of real biological systems.

Figure 3.2– Electron micrograph of a cluster of *E. coli* bacteria magnified 10,000 times. Each individual bacterium is rod-shaped. Image taken by Rocky Mountain Laboratories, NIAID, NIH, distributed under a CC-BY 2.0 license.

There are some biophysical aspects of a colony that can be quantified. These aspects reflect some general features that most surface-growing colonies have in common. The fashion individual cells grow, divide and push each other when competing for space, creates the emergent structure of a bacterial colony. In this section, I compare three global properties of an actual *E. coli* colony with the properties of a simulated one. For that, we have analysed some microscopy images of *E. coli* bacterial colonies taken by our colleagues at the Haseloff

Lab, in the University of Cambridge (Figure 3.5). Since the computation of growth and division with *gro* is now calculated with an algorithm developed in our laboratory, I want to ensure that the algorithm does not introduce any artefacts and that the physics of growth is realistic enough to represent actual colonies. The possible discrepancies that arise as a consequence of the computational representation of the cell are going to be identified and reduced, if possible.

Figure 3.3– Representation of the cell as a rod-shaped rigid body.

Simulated cells have a constant width of 1 μm and a variable length (Figure 3.3). In these simulations, I use the length and width that was derived by the Haseloff Lab in (T. J. Rudge et al. 2013), as they are the providers of the microscopy image that was used for the validation of the simulated colonies (Figure 3.4). Cells begin with a length of approximately 2 μm and divide once they have reached a length of $4 \pm 0.5 \mu\text{m}$, following a uniform distribution. The division length is calculated at the birth of every daughter cell. When a cell reaches its division length it generates two daughter cells that split the total volume of the mother cell. There is a variability in the partition that ranges from 40 to 60% of the volume of the mother.

Figure 3.4– Distribution of cell length (A) and cell width (B) taken from. Scale bars is 10 μm . Reprinted with permission from (T. J. Rudge et al. 2013) Copyright 2018 American Chemical Society.

The experimental procedure used to obtain Figure 3.5 forces bacteria to grow in 2D since the colony forms between agar pads. So, bacteria arrange almost as if the colony was two-dimensional, like it is in the simulated scenario. The arrangement of the individual bacteria was analysed by image processing techniques. Specifically, we subdivided each colony image into square images and process them with the OpenCV-Python libraries. This process was performed for the whole colony leading to a total of 20 images. Each image was processed, and individual bacteria were detected by segmentation (Figure 3.5). The segmented image was then incorporated into the simulator, for the calculation of its physical properties. With the extracted data, three different parameters were computed. These values summarize important structural properties of the colony.

First, the average number of neighbours is calculated by identifying the number of neighbour cells (which are in direct contact) of each cell. This was done for all 20 images. Cells at the edge of the image were ignored, as the total number of neighbours might not be revealed in the shot. The average number of neighbours is a consequence of the geometry of the cell, the average cell size and the packaging of the cells in space. It is important because it reveals the capacity of the simulator to distribute and group growing cells.

Second, the density of the colony is calculated as follows:

$$D = \frac{\sum \text{Bacterium Area}}{\text{Colony Area}}$$

The colony density returns a value between $[0, 1]$, and it informs about how well bacteria are packed in the simulation.

Finally, the relative direction of growth of each bacterium was measured and related to that of its neighbours. I refer to this parameter as the alignment of the bacterium, A_b , and is calculated as:

$$A_b = \frac{1}{N} \sum_{i=1}^N \vec{l}_b \cdot \vec{l}_i$$

where l_b is the direction of growth of the bacteria whose alignment is been calculated, and N represent the number of neighbours of each cell. Its value is comprised between $[0, 1]$, with 0 meaning that the cell is growing in a perpendicular direction in relation to its neighbours, and 1 meaning that the cell growth axis is in the same direction as its neighbours (Figure 3.6). This parameter informs about the inner order of the colony.

Figure 3.5– (A) Fluorescence microscopy image of an *E. coli* colony provided by Haseloff Lab (courtesy of Dr. Anton Kan). The arrows indicate the image processing (B) and segmentation process (C).

Results from both simulated and experimental colonies are shown in Table 3-1. Data corresponds to the mean value and standard deviation from the processed 20 images. The global alignment refers to the average of the individual alignment of all bacteria in the section.

	Experimental	Simulated
Average number of neighbours []	5.3 ± 0.5	5.5 ± 0.2
Colony density []	0.8 ± 0.1	0.9 ± 0.1
Average alignment []	0.8 ± 0.2	0.8 ± 0.2

Table 3-1 Comparison of validation parameters for simulated and experimental colony sections.

In Figure 3.6 the computation of the alignment is shown directly in each bacterium with a colour code being deep blue completely unaligned and light blue completely aligned. Simulated cells sustain a high local nematic order, which has said to be induced by nearby cells, cell division and the elongated shape of the cells (Cho et al. 2007; Volfson et al. 2008; Sheats et al. 2017). The histogram of experimental alignment depicts the same distribution (data not shown).

Figure 3.6– Alignment of cells in a simulated colony.

Results shown in Table 3-1 indicate a fair agreement between simulated and experimental colonies. It is worth mentioning that the agreement in the number of neighbours and global alignment demonstrates that the packaging of cells is performed quite realistically considering that cells are simulated as rigid-bodies and cannot exhibit the flexibility actual bacteria have. The inferior value in the experimental colony density is most certainly due to problems in the segmentation process that was unable to detect certain regions of cells (as it can be observed in Figure 3.5), causing that value to artificially drop.

The physical arrangement of cells simulated by *gro* is rather accurate. As mentioned in the introduction, a critical factor in the global colony shape that imposes a high competition among bacteria, is to grow under a limited resource scenario. To simulate a more realistic limitation of growth in the simulator *gro*, we implemented a library that offers a built-in Monod-based growth dynamic. It modulates the growth rate of any cell as a function of resource availability in the vicinity following a Monod equation. Depending on the spatial resource distribution, individual growth creates different colony morphologies. In Figure 3.7, the resulting shapes of each colonies are depicted along with the corresponding initial distribution of resources.

Figure 3.7– Nutrient distribution modes and resulting colony morphology.

The library offers the possibility to distribute nutrients in three fashions: homogenous, gradient and random. The graphs of Figure 3.7 represent a transversal section of the distribution of nutrients on a plate, from the centre outwards. The homogeneous distribution places an equal amount of nutrients, while the gradient distribution reproduces a radial extinction of nutrients. This mode can be useful for long-time simulations where the degradation of nutrients, or the inability to access them, occurs. The random distribution represents space inhomogeneity of any kind, and the variability can be also controlled.

The light blue cells of the colony shots in Figure 3.7 are no longer acquiring any nutrient and, therefore, they are not growing. Dark blue cells are growing cells placed above food reservoir.

In the same way that I validated the physics engine, the effect of the resource limitation in the colony growth was confirmed by a simple but well-illustrated phenomenon. In a single colony, a key parameter is the radial growth rate. This is, the velocity of expansion of the colony or the speed of the front of the colony. If growth is unlimited, bacteria within the colony divide leading to a global exponential growth. This unlimited growth translates in space as an exponential radial growth rate: the area of the colony πR^2 changes exponentially, and thus the radius R would change as the square root of the area, which is exponential as well.

When bacteria in the inner regions of the colony stop growing after some time due several reasons such as nutrient depletion, waste accumulation or build-up pressure, the radial growth rate decreases and becomes linear for long time periods (Fujikawa & Morozumi 2005). As only the external bacteria divide, the expansion is now due to the outer ring of non-exhausted bacteria. In the simulator, we would observe linear radial growth for long times when the number of bacteria that do not grow is much larger than the number of bacteria that still grow. This phenomenon is observed and shown in Figure 3.8. In the figure, three colonies are simulated under different conditions of resource limitation. The left image shows a colony without any growth limitation. The correspondent graph illustrates the exponential nature of growth. When bacteria are forced to acquire nutrients for division, the slope of the curve decreases, but it remains exponential. This corresponds to the middle image. The super-limited case shows the experimental behaviour of bacterial colonies, as the radial growth becomes linear with time (Su et al. 2012; Pipe & Grimson 2008; Farrell et al. 2013).

Figure 3.8– Radial growth in three different regimes of resource requirements.

3.2 Group formation and space structuring

Once it has been established that the physics engine of the *gro* framework is able to simulate realistic colonies, simulations in this section study how the collective structure of cell groups arises from the activity of many individual cells. For the systematic study of the spatial constraints and their impact on the cells arrangement in subsequent sections, here, a new metric is introduced, defined and characterized. The metric is based on the concept of group as a set of cells that are genetically related as well as the cells that are in contact with them. The concept is referred to as *physically connected group*. The ability of this proposed abstraction to identify the important emergent property of space structuring will be discussed.

Figure 3.9– Cells in a bacterial colony are connected by physical contacts that can change with time, very much like a dynamic graph.

Individual cells in a colony are in direct contact with their neighbours. Similar to a dynamic graph (Figure 3.9), the identity of these neighbours changes with time even if the average number of neighbours remains constant. Theoretically, in a colony, every cell influences the position and orientation of every other cell, but the strength of the interaction drops with the distance (or number of cells) separating those two. In an effective way, each cell can influence only a finite number of close cells, therefore, in practice, there is actually a limited number of cells that will influence each other. Eventually, these cells will end up influencing some other

cells. A *physically connected group* is defined as follows: given a bacterium b_i , there is a fixed number of bacteria with whom b_i will interact via physical contact during certain simulation time. The number of cells that interact via physical contact of any duration with b_i and its descendants, represent the *group*. So, basically, a *group* is consisted of 1) clonal cells of b_i , and, 2) any neighbour cell of a clonal cell of b_i . The duration of the physical contact defines the spatial stability of the *group* and the colony. The evolution of this number is a metric of the colony inner structuring and contact stability.

Figure 3.10– A colony consists on many connected groups. In turn, a group consists on many individual cells.

There are as many *physically connected groups* as bacteria in a colony. Bacteria (and our physics algorithm) tend to minimize the space among cells, which leads in these simulations to a constant density of cells per μm^2 . Because of this constant density, the general assumption is that a group defined by a single bacterium b_i can be extrapolated to the rest of bacteria in the same colony. This assumption makes sense for all bacteria, except for bacteria in and near the border of the colony. The *group* represents the intermediate step from individuals to colonies (Figure 3.10). It is the optimal level at which to study the effects that a single cell creates over the whole colony. By studying the dynamics of *group* formation and their shape, information that can be hidden in the dynamics of the whole colony appears.

Spatial and contact stability are of crucial importance when studying the impact of the spatial constraints in the process of cell communication and cooperation. The more stable the cell contacts, the more plausible is that cell communication could derive in cell cooperation.

The definition of *group* helps to find the conditions that benefit contact stability, at least, long enough for communication to be reliable and reciprocal.

A visual representation of a *physically connected group* and the different natures of its members is represented in Figure 3.11.

Figure 3.11– Different elements of a physically connected group.

As it can be seen in the figure, the members of the *group* are divided into three types:

- Clonal cells (daughter cells of primordial b_i)
- Close/stable neighbours.
- Far/transient neighbours.

Contact types are calculated according to the duration of the physical connexion with any of the clonal cells. The b_i cell establishes a lineage of cells (depicted in deep blue in Figure 3.11).

The daughter cells of the primordial cell b_i are marked from birth. Any contact a clonal cell has, will also be part of the *group*. Close and far neighbours correlate with the contact duration they have with any of the clonal cells. Basically, the spatial stability (stable or transient neighbour) is linked with the proximity of the contact (close or far neighbour).

A close stable contact would begin as a transient connexion. If the contact remains for at least a generation time, the contact type is then considered close (since it is in direct contact) and stable. Monitoring the formation of the *group* over time reports about the spatial stability of its elements. Temporal shots of a *group* formation are shown in Figure 3.12.

Figure 3.12– Dynamical formation of a *physically connected group*.

Initially, bacterium b_i divides and starts to generate transient contacts with the surrounding cells. After some time, some of the initial transient contacts become more stable. Because of the competition for space which comes as a result of growth and cell division, the *group* shape deforms and many of the transient contacts end up far away from the lineage defined by b_i .

A graph with the relative frequencies of each of the different members of the group is shown in Figure 3.13. In this particular simulation, bacteria have a division length of $4 \pm 0.5 \mu\text{m}$ and

a generation time of roughly 20 minutes. The first four shots of Figure 3.12 correspond to the behaviour that is observed in Figure 3.13 during the first 80'. This early stage is when the group develops, and the relative frequency of the three types changes dramatically. During this formation stage, many contacts which were initially transient become stable (notice that all initial contacts are transient). When the variability in the relative frequency of cell types is reduced, values converge and tend to stabilize around a constant value. This is when all transient contacts become stable contacts and space becomes structured.

Given the average generation time and the moment of stabilization, at least four generations were required for the space to stabilize and structure in this simulation.

Figure 3.13– Frequency of the different members of a *group* over time.

Evidently, the process of group formation is highly random and the data shown in Figure 3.13 correspond to a single run. In each run, the shape of the group changes. Still, the relative frequencies of the members and the discrimination of the two stages (formation and stability of the *group*) remain fairly constant. For $N=20$ simulations, the proportion of members returns the following coefficients for 160 minutes of simulation, yielding the composition:

$$\text{physically connected group} = (0.10 \pm 0.01) * [\text{clonal}] + (0.20 \pm 0.05) * [\text{stable contacts}] + (0.70 \pm 0.05) * [\text{transient contacts}]$$

Most of the group is composed by transient contacts that were at some point in close proximity with any of the clonal cells but that then separated and kept on diving. It is important to notice that any cell that had any contact with the clonal set holds its state as a contact. However, the characteristic of close contact only holds if the cell is near. If a cell in close proximity is pushed away from the clonal set it will automatically transform into a transient contact and its offspring will be marked as well.

As it was mentioned in the introduction of the chapter, groups can be formed by two means. Either cells divide and stay together after division or they come together with other cells and create some sort of permanent bond. The *group* metric defined here, distinguish and includes those two possibilities in its definition. Since transient contacts provide little information about the *group* stability other than the transition from transient to permanent, from now on they will be removed from the *group* definition. So, if transient contacts are put aside, the proportion of clonal cells and their close contacts is a fine metric of the percentage of clonal cells and non-clonal cells in a given *group*, which now can be reinterpreted as:

$$\frac{[\text{clonal cells}]}{\text{group size}} + \frac{[\text{permanent contacts}]}{\text{group size}} = \% [\text{Staying together}] + \% [\text{Coming together}]$$

And the fraction of cells of each category becomes:

$$\text{Connected group} = (0.35 \pm 0.05) * [\text{Staying together}] + (0.65 \pm 0.05) * [\text{Coming together}]$$

The phenomenon of space structuring is what has been suggested to promote staying together (i.e. clumps of clonal cells) over coming together (Tarnita et al. 2013). Because of the spatial constraints, cells after division would stay close to each other and be more inclined to remain with those of its kind. This clonal clusters can be generated passively, without active aggregation among clones (Nadell et al. 2010). However, data from these simulations indicates that, proportionally within a group, there are more non-clonal stable contacts than clonal cells. So, daughter cells will reach far away from its mother if there is no other kind of social interactions holding them together. Probably because of the radial nature of the

expansive growth in colonies, a clonal lineage would most certainly form elongated shapes. Cells are not inclined to stay together in a big clump because of the axial growth of rod-shaped cells (Nuñez et al. 2017; J. Rudge et al. 2013). The *group* defined by the lineage of a single cell would have an elongated profile most certainly due to their individual shape (this conclusion is further study in the next section). However, spatial structuring happens when the *group* starts to widen instead of expanding along the direction of growth. This is when the group thickens and the isolation of clonal cells in the inner regions of the *group* takes place (Figure 3.12).

The ratio of clonal and neighbouring cells hints the profile of the group. Because neighbouring cells represent the perimeter of the clonal cells, the ratio will inform if the cluster is more elongated (i.e. % neighbours > % clonal) or circular (i.e. % neighbours < % clonal). This phenomenon can be observed in Figure 3.14. If the group is defined by only one cell, the initial tendency of the cell to divide and grow along the longitudinal axis is transmitted and the global shape of the group will result elongated. On the other hand, if the seed of the group is already globular, such as the case where a group of cells are initially defining the clonal group, then, the group shape will result more circular. The middle and bottom images of Figure 3.14 are examples of this peculiarity, where an originally circular group of cells is marked as clonal cells in the beginning of the simulation. The resulting group shape stays more circular as more initial cells are conforming the initial set. Also, the relative frequency of clonal cells and close contact cells is reversed evidencing the ability of the proposed metric to inform about the individual cell arrangement and the global geometry of the group. Notice that the relative frequency of cells in each case stabilizes for different times. The perpendicular discontinuous line in the graphs, indicates roughly the moment of stabilization. That stabilization is a consequence of space structuring. The stabilization appears sooner (2 generation times and 1 generation time for the 50 initial cells and 100 initial cells, respectively) for the middle and bottom situations because the initial conditions are altered to recreate such stabilization. I have argued that space structuring occurs when the expanding group starts to grow in various directions (the group widens) and the number of contacts

becomes constant consequently. Once the shape is formed (formation stage), clonal cells and close contacts keep on dividing in a rhythmic way and the general form holds (stable stage).

Figure 3.14– Time evolution of the frequency of the clonal cells and their closest neighbours and image shots at 60'. Simulations were run for various initial number of clonal cells.

The phenomenon of space structuring is further analyzed in two different settings. Instead of considering and following only the growth and shape of one lineage, simulations involving different lineages were run. Lineages or strains only differ in a random colour code that it is assigned to visually follow the process of structuring. Still, all cells have the same generation time and a division length of $4 \pm 0.5 \mu\text{m}$. In the first set of simulations, 100 cells of each strain are randomly placed in high density conditions in the simulation environment. Zoomed temporal shots of the evolution and formation of clonal patches (same colour patches) are shown in Figure 3.15. The image reflects how an initially disordered mixture of strains can become highly structured such that the final community contains large single-genotype patches that span many cell lengths.

In each of the simulations ($N=20$) a random cell was used as a probe to measure the fraction of cells of its same lineage and its close contacts. The fraction of close contacts for the different number of strains is shown in Table 3-2 under the well-mixed column. The fraction of close contact is invariant of the number of strains, as expected. Also, the value is the same as in the previous simulations.

	Well-mixed (100 cells/strain)	1 cell/strain
5 strains	0.65 ± 0.05	0.16 ± 0.05
4 strains	0.64 ± 0.05	0.15 ± 0.06
3 strains	0.65 ± 0.05	0.11 ± 0.03
2 strains	0.65 ± 0.05	0.05 ± 0.02

Table 3-2 Fraction of close neighbours in a group at 200'.

Figure 3.15– Time shots of the cell arrangement for simulations using different number of strains. The images show how space becomes more structured and patches of clonal cell develop as cells divide. Cells are colour marked but are otherwise equal.

The next simulations consider a different initial environment. The effect of the group elongation can be due not only to the axial growth of cells but also it is highly influenced by the growth of the rest of individuals in the colony. If single cells from different strains are placed near, in an otherwise empty fictional plate, the structure they form is rather different (Figure 3.16). Growing lineages clash with each other competing for space in the inner regions of the colony and expand freely colonizing the outer regions. The relevance of these simulations is to show how the constraints that cells impose on each other impact on contact stability.

When single cells are left to grow with less spatial competition, the outer regions of the colony, where expansion takes place, are predominantly conformed by a specific cell lineage. The shape of these *groups* is far more globular than in the previous case, as it reflected in the data in Table 3-2 (1 cell/strain column). The fraction of neighbours drops significantly, reflecting that when colonizing (without the individual competition for space in an already crowded environment) the members of a *group* are mostly clonal cells.

However, unlike the previous case, this scenario is sensitive to the number of strains. Initial conditions have a strong impact in the later development of the colony structure, especially when simulations are initialized with very few agents. Simulations of colonies with five initial cells present a more disorganized inner region than the ones with less initial number of cells. The fraction of close neighbours in a group reflects such disorganization.

Previous theoretical and experimental work has revealed that initially diverse and well-mixed microbial populations will commonly lose their diversity when growing in dense groups such as bacterial colonies, leading to large patches of single genotypes. This phenomenon has been observed *in silico*, *in vitro* and *in vivo*, and causes clonal patchiness as a function of cell division (Hallatschek et al. 2007; Hol et al. 2013; Smith et al. 2017b; Müller et al. 2014). Population structure in microbes will mostly arise through limited dispersal (population viscosity or structure) which can generate high degrees of relatedness between interacting

individuals because it will tend to keep relatives together, allowing cooperation to be directed indiscriminately toward all neighbours.

Figure 3.16—. Colonies formed from single cells from various strains.

Limited dispersal is extreme in the case of growth of non-motile individuals, as in the presented case, which leads to the formation of microbial colonies. Furthermore, even with initial mixing, subsequent growth on solid surfaces leads to spatial segregation of genotypes by sector formation (Figure 3.15).

As a summary, simulations have revealed that the concept of *physically connected group* exists, and that it serves well as a metric of space structuring, which is fundamental for the stability that simulated cell adhesion would have provided. I have defined and characterized the metric of *connected group* as a readout for visualizing the mechanical and spatial stability in a simulator that is not capable of emulating actual adhesion between cells.

3.3 Bacterial size and its effect on group formation

So far, I have revised the general properties of the simulated bacteria in *gro* and described the natural formation of *groups* and how to quantify them. The concept of *physically connected group* has been defined and some important properties have already arisen. Every simulated colony displays two well-differentiated stages, prior to spatial structuring when the space is disrupted easily and right after, when the change in neighbours begins to decrease. This is all reflected in the evolution and formation of the *group*. I have mentioned in section 3.1 that the average number of neighbours is a consequence of the cell geometry, if cells are properly packed. This number remains constant with time, but, the specific neighbour will probably change. Because I am interested in finding the mechanisms that prevent the neighbourhood disruption and maintain cells tighter together as if they were mechanically stuck, the transient contacts quantity will not be as important as the stable and close contacts which will be the parameter to maximize.

The most important property at the individual level, or at least the one that could have a more profound effect on the colony arrangement, is bacterial shape (Smith et al. 2017; Kennard et al. 2015; Young 2007). Cells exhibit different sizes, even within the same strains. Most individual-based models do not allow cell shape and length to be altered.

The simulation framework *gro* allows to modulate cell length alongside cell division, physical interactions, and metabolic interactions via nutrient consumption. Recent studies have suggested that rod-shaped cells can drive collective behaviors in microbial groups because of their tendency to align their orientations with nearby cells and surfaces (Sheats et al. 2017; Farrell et al. 2017; Volfson et al. 2008). Aligned cells are also subject to buckling interactions, which fold neighboring cell groups into one another to form fractal-like interdigitations (Rudge et al. 2013).

In this section, the effect of cell length in the formation and structure of the *group* will be analysed. Results will determine which cell length generates more stable and permanent connexions.

Five different average division lengths were simulated. Simulations of colonies of 50000 individuals were grown for five different average cell division lengths (N=20 for each division length). As expected, cell size has a strong impact on bacterial arrangement. Our physical engine, CellEngine, was optimized to pack and distribute growing cells of roughly 4 μm or less, and therefore, rigid body dynamics fails to achieve order in colonies with large-size bacteria. Bacteria of smallest sizes are packed more compactly, as it can be seen in the physical parameters of Table 3-3.

	2 μm	3 μm	4 μm	5 μm	6 μm
Avg. Neighbourhood []	6.2 ± 0.1	6.0 ± 0.1	5.5 ± 0.2	5.5 ± 0.3	5.5 ± 0.3
Density []	0.95 ± 0.05	0.94 ± 0.05	0.9 ± 0.1	0.88 ± 0.15	0.85 ± 0.15
Alignment []	0.77 ± 0.20	0.81 ± 0.20	0.8 ± 0.2	0.75 ± 0.25	0.7 ± 0.3

Table 3-3. Validation parameters for different bacterial division lengths.

The global alignment, the average number of neighbours and the global density drop for long bacteria because of their rigid nature (cells are modelled as rigid bodies). Contacts among cells are reduced due to the physical engine inability to pack such large bodies.

Figure 3.17 - Effect of bacterial size on the colony arrangement and cell alignment. Lighter blue indicates the orientation of the cell is in the same direction as its close neighbours.

The *group* shape and stability are also highly influenced by cell size. Image shots of the group forming for different cell division lengths are shown in Figure 3.18. Simulations start with 100 cells in high density. Bacteria b_i is placed randomly in the initial plantation and it is marked with a dark blue colour. Descendants from b_i are marked in the same way. Close, stable contacts are marked in light blue. The shots of Figure 3.18 are taken at different times so the group size is equal. Cells have different division lengths but the same initial length of $2 \pm 0.5 \mu\text{m}$, excepting cells that divide at $2 \pm 0.2 \mu\text{m}$ that have an initial length of $1.5 \pm 0.2 \mu\text{m}$. The growth rate is constant and equal in all simulations leading to different generation times. This way, $2\mu\text{m}$ cells divide much faster ($\sim 20'$) than $6 \mu\text{m}$ ($\sim 75'$).

Figure 3.18 – Shots of the *groups* with different division lengths.

Clonal cells remain connected forming a single domain for small lengths but begin to disconnect as the cell length is increased. For division lengths of $5.0 \pm 0.2 \mu\text{m}$ and $6.0 \pm 0.2 \mu\text{m}$, cell arrangement changes so dramatically that clonal cells do not longer form a physically connected *group*. Instead, every cell is surrounded by not related cells increasing the contact surface of the clonal cell group. The average fraction of close neighbours for $N = 20$

simulations is shown in Table 3-4. Larger division lengths have higher fraction of neighbours in their groups because of the disconnection effect. Groups of 2 μm cells are more globularly shaped because the axial growth phenomenon is less pronounced. Therefore, small-length groups tend to be more clonally connected.

	100 cells initial cells	1 initial cell
2 μm	0.55 ± 0.01	0.11 ± 0.05
3 μm	0.60 ± 0.03	0.15 ± 0.06
4 μm	0.68 ± 0.03	0.18 ± 0.03
5 μm	0.69 ± 0.05	0.30 ± 0.02
6 μm	0.72 ± 0.05	0.42 ± 0.02

Table 3-4 Fraction of close neighbours when group size is 200 cells.

The next scenario considers again cells from different strains and the mixing and colliding of sectors of clonal cells. Five previously marked cells are placed in the simulated plate at distance of 6 μm forming a symmetrical cross. In the centre of the cross a blue marked cell is placed. The blue strain is the reference strain to which the *group* is measured. The fraction of neighbouring cells is included in Table 3-4 (1 initial cell column). The blue lineage in Figure 3.19 is constrained by the growth of the other four lineages. This spatial constriction should maintain blue cells restricted to the inner region of the colony, however, blue cells from colonies consisting of 5 and 6 μm bacteria tend to disperse regardless of the spatial restriction of the other lineages.

Figure 3.19—Bacterial sizes and colour code of the different lineages

To summarize this section, cell length has proved to be a crucial attribute regarding *group* formation. Simulations have revealed that, on average, cells with smaller sizes form *groups* that have a lower fraction of neighbours, which means that the group could be rather globular and consisted primarily of clonal cells.

Simulations with higher variability in division lengths were run as well (data not shown). Results with higher variability agree with the presented observations: the larger the cells, the less connected are the groups. This observed phenomenon should be partially attributed to the inability of the physical engine to model the growth and arrangement of large cells as it was appointed by the drop in the colony density and cell alignment (from the validation parameters), a process which has not been reported experimentally (Cho et al. 2007).

3.4 Resource limitation and its effect on group formation

The way cells arrange in a colony is a consequence of the competition for space that occurs while cells are growing and dividing. This competition is manifested in the shape and size of each cell that ultimately creates a short scale evolutionary game at each spatial location. Cells that were randomly granted a specific size in a specific moment and that were able to outcompete any other close neighbour, could then establish in a spatial location to keep on growing and dividing. This effect has been reportedly to be magnified when the resources for growth are limited: the competition for space that shapes the connected groups becomes fiercer when cells are also competing for resources (Nadell et al. 2010). This competition acts almost as an adhesive force within the group, segregating clonal strains in a more pronounced way. Even when different strains or species are initially well mixed, populations that grow towards a limited nutrient pool often experience strong spatial constraints as some cell lineages are cut off by the actively growing front. This process, referred to as spatial genetic drift, induces population subdivision into monoclonal sectors (*i.e.* positive assortment) and has been documented in agar colonies of various strains (Van Dyken et al. 2013).

Cells in natural systems utilize mechanisms for displacing neighbours that are in direct competition. Antibiotic secretion and direct injection of toxins into adjacent cells are some examples of mechanism that increase the basal level of competition (Baquero et al. 2011). Also, size and metabolism can be selective traits that benefit cells in specific situations. For example, fast-growing organisms will consume nutrients before their neighbours take over their space. Even within the same group, cells are simultaneously sharing resources and competing for them. The success or failure of all these interactions depends strongly on precise details of the physical circumstances in which the colonies are growing.

The resource dependent growth feature is used in this section to simulate and study the commented phenomenon. Simulations take place in an environment of initially fixed substrate concentration where cells consume at a fixed rate as well.

The chosen initial amount of nutrients and the consumption rate are shown in Table 3-5. Different cell consumption rates will be simulated far along in this section.

Initial nutrient molecules per square grid	50
Consumption rate (dt^{-1})	0.30

Table 3-5. Nutrient consumption parameters.

Even if nutrients are evenly spread out in the fictional plate they affect the spatial arrangement of cells in an anisotropic manner. Cells are bound to consume nutrients to grow, and so, nutrients are first depleted from the colony centre (where cells are placed initially). When nutrients are exhausted, the corresponding cell above would stop growing. Only cells at the edge of the colony can continue to expand as they colonize rich nutrient regions. The metric of the *physically connected group* is now measured closer to the edge of the colony because it provides longer time measurements, as the centre stops growing much sooner, due to the anisotropy in space of nutrient consumption (Figure 3.20)

Figure 3.20– Visualization of a group in the centre and in the edge of a colony that is nutrient-limited.

The frequency of clonal cells and close neighbours of a *group* for a given representative run is shown in Figure 3.21. The sample group is taken relatively close to the edge of the colony, in the expansive region.

Figure 3.21– Frequency of clonal cells and close neighbours in an edge group over time in a colony that is nutrient limited.

The initial stage of group formation is similar to the canonical case presented in section 3.2. However, lineage segregation at the colony surface promotes the growth of a random clonal cell which, unlike any of the close neighbouring cells, ends up colonizing the sector. The ratio of clonal cells and close neighbours switches, resulting in a *group* that is consisted mostly of clonal cells. This experiment proves that the initial assumption (that was also supported by many empirical and theoretical works) that resource limitation promotes colonization, and thus segregation at the colony edge, can be also simulated in this framework.

This initial proof of concept is further studied in conditions of initially well-mixed cells. Simulations were run for different bacterial lengths and five different bacterial lineages (only distinctive by colour). Representative shots of the colonies with different cell lengths are

shown in Figure 3.22. Colony shots are taken at different times, so they consist on the same number of cells. Remember that larger sizes have larger generation times because the growth rate of all cells is the same. Besides, all cells are set with the same consumption rate and were initially placed in the centre of the simulated plate.

Figure 3.22–Spatial segregation of five lineages due to nutrient limited growth.

The segregation effect is produced in the 2, 3 and 4 μm colonies but fails to happen in larger size colonies. Intriguingly, despite sharing the same growth rate, the size effect plays a role in the segregation phenomena. The general conclusion is that larger sizes feel less competition than smaller sizes, for the same environment.

To better understand this effect, let's imagine the grid where virtual nutrient molecules are place at the beginning of every simulation. The grid cell has a length of 3 μm , which leads to a total surface of 9 μm^2 . This means that smaller cells will have to share the fixed amount of nutrients that are place in every grid-cell among many more cells than the large cells. As an example, in this scenario, a grid-cell is occupied, on average, by 4.5 two-micron cells whereas

only 1.5 six-micron cells fit in the same space. Therefore, two-micron cells will have one third less ($4.5/1.5$) of the available molecules per individual cell to consume, in compare to six-micron cells. Since the absorption rate is the same for all cells, nutrient depletion will happen faster in small size populations. So, the actual available nutrient molecule per cell is what sets the degree of competition, and therefore, the degree of segregation. Even when the consumption rate and the yield of that absorbed nutrient (growth rate) are the same for all cases, it is the individual cell and not the total mass (*i.e.* the yield of nutrient into cell mass) which is creating the competition.

Next simulations were initialized by placing five different marked cells in the simulated plate (Figure 3.23). When the consumption rate is the initial value of 0.3 molecules per time step, the segregation effect is, again here, more noticeable for small length populations. The high competition that small bacteria experience in this setting leads to the distinctive pinwheel pattern. When several lineages are expanding into new territory, whichever strain finds itself at the frontier of the population will grow preferentially, creating the pinwheel pattern shown here. As cells divide and expand out into the plate, they quickly segregate with well-defined sections of colour. The pinwheel effect diminishes together with the effective competition in the different populations. 5 and 6 μm colonies barely feel the selective pressure and their growth pattern resembles that of an unlimited growth scenario (Figure 3.19).

nutrient consumption rate= 0.3 molecules/dt

Figure 3.23– Effect of nutrient limitation for different cell sizes and a consumption rate of 0.3 mol/dt.

On the other hand, when the consumption rate is increased, large cell populations transform into the pinwheel pattern as well (Figure 3.24). Notice that an increment in the basal absorption rate is analogous to a decline in the initial amount of resources. This increment in the absorption rate is enough to create a stronger competition that segregates the lineages right from the centre of the colony. Accordingly, small sizes drain their resources so quickly that the colony presents the star-like shape typical of very low nutrient plates (Tokita et al. 2009; Cooper et al. 1968).

Figure 3.24— Effect of nutrient limitation for different cell sizes and a consumption rate of 0.6 mol/dt.

Finally, to properly quantify this phenomenon and study more explicitly how does the spatial segregation impact the attributes of a *group*, the average fraction of close neighbours in a *group* is represented as a function of cell length and the consumption rate in Figure 3.25. Each point in the surface plot represents the average value after $N=20$ simulations. The experiment set is an initial well-mixed case of cells with different sizes and different consumption rates. The figure above the graph represents visually the effect of increasing the consumption rate for the initial well-mixed simulation. The three-dimensional graph resembles the slope of a mountain where the highest point represents the highest fraction of close neighbours. Remember that such quantity is complementary to the clonal members of a *group* which translates graphically as the segregation. So, greater segregation implies that groups will be consisted primarily on clonal related cells. As previous visual results pointed out, cells segregate more when their nutrient consumption rate is higher, and their division length is shorter.

This section has provided further proof that nutrient dependent growth is a strong force that shapes the spatial arrangement of bacterial cells growing on a plate. Resource limitation increments the clonality of a group only when bacteria start to compete for food and segregation becomes more abrupt. Besides, the observed dependence of cell competition with cell size should be tested as well experimentally. That would provide important information on the nutrient uptake process and the dynamics of growth. Experimental results could confirm whether the assumptions made here (namely, that the growth rate is independent of size and that the same amount of nutrients yields the same mass creation) are correct.

Figure 3.25– Fraction of close neighbours in a group for different consumption rates and cell division lengths.

3.5 Size competition in resource limited environments

This final section will not deal with the process of *group* formation and contact stability but rather follows the effect discovered in the previous section. Individual shape has proven to be a relevant factor in the arrangement of bacterial colonies. Size and cell morphology can be competitive advantages that play an important role as evolutionary strategies. Several studies suggest that, by influencing biomechanical interactions between microbes, shape may have far-reaching consequences for the properties and prospects of a cell within a community (Young 2007; Nuñez et al. 2017; Smith et al. 2017a; Bell & Mooers 1997; Gounand et al. 2016). Similar to the simulations and experiments of Smith et al. (Smith et al. 2017), the simulations presented here compare the fitness advantages of two different strains that present varying cell sizes. Simulations are run for two distinctive initial configurations. First, simulations are initialized with only one cell of each strain (Figure 3.26). With that initial setting, two environmental conditions are implemented and compared: either growth is unlimited, or it depends on the absorption of nutrients. Varying cell sizes are simulated in a pair-wise manner ranging from small (division length of 3 μm), medium (division length of 4 μm) to large (division length of 6 μm). Each pair-wise competition is run several times to achieve statistical significance.

When cells are able to grow without the requirement of nutrient uptake, smaller sizes show a clear fitness advantage (Figure 3.26). For example, if an initial small cell is left to grow along with another large cell, the proportion of smaller cells after 3 hours is four times greater than the proportion of large cells. That trend is less accused if the competition is between small and medium sizes or medium and large sizes, which incidentally present the same fitness advantage respectively. This advantage turns around when growth is dependent on nutrient acquisition. In that scenario, cells are imposed to absorb 0.3 molecules of nutrient every simulation step in order to grow normally. Here, the effect of the previous section prevails and favours larger sizes. Because large cells are able to drain the entire reservoir of nutrients

almost by themselves, without competing with each other, they expand more rapidly than smaller cells.

Unrestricted growth

Resource-restricted growth

Figure 3.26—Frequency of cells from each strain as a measure of the fitness of the strain size. Competition between sizes is simulated in a pair-wise manner for unrestricted and resource-restricted growth. Colonies are seeded with a single cell from each strain, and colony shots are taken when the population reaches 5000 cells.

Surprisingly, while a smaller size is advantageous when coexisting with a larger size strain in the unlimited growth scenario, a greater volume becomes beneficial when the competition for resources is fiercer. Notice that the advantage of size does not increase with the size difference as the relative frequency of cells for the small-large competition is almost the same as in the small-medium and medium-large. Size also provides a crucial advantage at the edge, as it facilitates the reaching of further regions that can still harbour food.

This set up did not provide direct cell-cell competition, as only the cells in the interface were in direct contact. So, this configuration informed about the advantage of each strain almost in isolation, as cells from each lineage divided practically as independent processes. The second configuration is that of an initially well-mixed plate with 100 cells from each strain. Here, direct competition can be better tested. Results are shown in Figure 3.27. The main difference lies in the obtained relative fitness. In both scenarios, the difference in the fraction of cells is much smaller than in Figure 3.26. So, when in direct competition (either for space or for resources), the reproductive advantage of each strategy is slightly reduced.

The relevance of the presented results derives from the yet unidentified consequences of the variety of shapes and sizes in microbial communities and their corresponding evolutionary and ecological significances. Simulations predict that size can have strong effects on cells survival and reproduction. It was observed how long rod-shaped cells are better at colonizing the expanding edges of a colony, especially if resources are scarce, whereas smaller cells have greater reproductive success when the environmental conditions are more favourable.

Figure 3.27—Frequency of cells from each strain as a measure of the fitness of the strain size. Competition between sizes is simulated in a pair-wise manner for unrestricted and resource-restricted growth. Colonies are seeded with an initial well-mixed clump of 100 cells from each strain, and colony shots are taken when the population reaches 5000 cells.

3.6 Conclusions

This chapter has investigated the effect of the physical spatial constraints in the process of *group* formation in two-dimensional bacterial colonies constituted of rod-shape cells.

The conclusions withdrawn from each section are now enlisted:

- The global physical properties of the simulated bacterial colonies with *gro*, such as cell arrangement, colony density and the average number of neighbours of a given cell, lie within the same range of those of actual *E. coli* colonies (Table 3-1). Therefore, the physical algorithm that simulates cell growth and cell distribution due to growth-related pressures, offers an accurate approximation making *gro* a valid framework for the simulation of bacterial structural phenomena.
- The proposed concept of *physically connected group* provides a useful standard for determining contact stability. Since social interactions take place among sets of individuals that are small compared to the total population, *groups* do not require discrete boundaries because the important feature is that social interactions are local, compared to the size of the total population. The information that the proposed metric provides can be further enlisted as:
 - The dynamical process of *group* formation presents two separate and distinct phases: first, a variable phase where the group is forming and contacts among neighbouring cells are transient, and, then, a stable phase where the group scales while individual cells divide, maintaining most of the contacts. The appearance of the latter phase represents the structuring of space that is a measure of *group* stability (Figure 3.13).
 - The members of a *group* can be categorized in two sets: clonal cells (that define the group itself) and non-clonal cells that surround the clonal cluster. Thus, a *group* can be described following its own definition of %Clonal cells + %Non-clonal cells = 100%, and so, the identity of a group can be referenced by one of those fractions (e.g. groups formed by cells of 3 μm have an average of 60% of non-clonal neighbours.).

- The fractions of close neighbours of a *group* (and, evidently its complementary value of clonal cells) can also inform about the general group shape. If the fraction of close neighbours is lower than 0.5, the general group shape will be globular, while if it is higher the group shape will be more filamentous (Figure 3.14).
- The shape of a *group* determines the spatial structuring phenomena. Globular groups will experience earlier the stable phase of space structuring. On the other hand, elongated groups require more cell divisions to reach the structured state (Figure 3.14).
- Group shape, and subsequently spatial structuring and stability, depends as well on the initial number of cells. The lower the initial number of cells the more compact and stable is the emergent group. This is a consequence of the decrease in spatial competition that simulated cells experience when growing surrounded by other growing cells (Figure 3.16).
- Bacterial shape (*i.e.* length) has a profound impact in the contact stability of the group. This conclusion is obtained through two different sets of readouts.
 - Global readouts such as cell arrangement, colony density and the average number of neighbours are influenced by the cell average length. Large cells (5 and 6 μm) present less order effects (lower alignment). Density and average number of neighbours also drop for large sizes because of the rigid body model used by the physical engine (Table 3-3).
 - Derived readouts such as the properties of the *group* are also highly influenced by the average cell length. Large cells (5 and 6 μm) cannot maintain connected *groups* of clonal cells and present therefore low stability. On the other hand, cells with smaller sizes form *groups* consisting primarily of clonal cells (Figure 3.18).
- Cell competition via nutrient acquisition acts as an adhesion force that promotes *groups* consisting on clonal cells. The degree of the competition is tuned through cell size, with smaller sizes reacting more strongly to the resource competition and forming segregated sectors in the colony (Figure 3.25).

- Different cell size lineage competition promotes smaller sizes when competition is only at the spatial level. However, if competition via nutrient acquisition is included, larger sizes result as the best ecological strategy (Figure 3.26 and Figure 3.27).

Some of these conclusions will be exploited in the next chapter, that studies the effect of different mechanism and strategies to help maintain a population of cooperative cells that is at risk of be taken by cheaters. The concept of *group* will be useful then, and it will be extended to differentiate into cooperative and cheater neighbors. Some models of biofilm growth have predicted that the observed spatial structuring promotes cooperative behavior (Kreft et al. 2000; Xavier et al. 2009; Nadell et al. 2010; Leinweber et al. 2017; Nadell et al. 2016). The spatial arrangement of cells within biofilms strongly influences the relative fitness benefits of the different phenotypes (mainly, cooperators and defectors). The arrangement of cells in space is therefore crucial to whether competitive or cooperative interactions are advantageous in a given environmental context. In summary, these spatial relationships may be paramount for understanding the evolution of cellular cooperation. When different cell lineages are segregated in space, those expressing cooperative phenotypes are more likely to benefit others of their own kind. On the other hand, when different cell lineages are mixed together, cells that exploit the resources of others can thrive.

Chapter 4

Individual strategies for the public goods dilemma

Bacteria are social beings. Because of their high reproductive rate, unlike other living organisms, they do not disperse after division but rather stay together and form communities (Li & Tian 2012). Within these communities, bacteria often exhibit coordinated behavior that is directed to the common goal of the survival of the colony (Crespi 2001; West et al. 2007). In those cases, bacteria rely on the costly production of certain molecules that provide shared benefits for the community. Examples of these secreted molecules range from informational signals like quorum sensing, biofilm polymers, iron-scavenging siderophores or digestive enzymes (Diggle et al. 2007; Allen et al. 2016; Boyle et al. 2013). These secreted molecules are referred to as public goods as they provide a global benefit for the commons.

Figure 4.1– The public goods dilemma. Adaptation of the selfish mutant that is selected for its higher reproductive rate, a consequence of not producing the common good, results in the extinction of cooperators.

When bacteria collaborate through public good sharing, there is always the risk of the appearance of cheating cells, which exploit the public goods without providing any contribution to them. These cheaters or defectors are cells that gain a relative fitness advantage arising from not investing the costs associated with public good production (Ghoul et al. 2014; Czárán & Hoekstra 2009; Stewart et al. 2016; Travisano & Velicer 2004). This can lead to the conquest of defectors and the elimination of the cooperative phenotype (Figure 4.1).

For example, the fruiting body morphogenesis and sporulation of the *Myxococcus xanthus* can be disrupted when defectors overtake the population even to the extent of driving it to its extinction (Fiegna & Velicer 2003).

Despite this puzzling dilemma, cooperative and collective behaviour is observed across many biological entities (Claessen et al. 2014). The accepted hypothesis is that cooperators can only outcompete defectors if any net costs of cooperation are sufficiently offset by increased rates of interaction with fellow cooperators (Allen et al. 2016). In this sense, several mechanisms and environmental effects have been suggested to maintain and promote cooperation (Asfahl & Schuster 2017; Ghoul & Mitri 2016; Mitri & Foster 2013). In this chapter, such different scenarios will be simulated to elucidate their effect on maintaining and promoting cooperation in a situation where defectors have already infested a population of cooperators.

Objectives of this chapter:

In each section, relevant simulations will be testing a specific strategy or condition and its success to maintain the relative frequency of cooperators in a public goods scenario. However, all simulations will use the parameters described in this introduction unless specified otherwise in the corresponding section. The fundamental difference among sections will be the strategy or condition to analyse.

Simulations consider two distinctive phenotypes: cooperators and defectors. Both phenotypes present fixed strategies; they are obligated to follow their strategy and are unable to alter their given status as cooperators or defectors. Cooperation is based on the production of a diffusible public good which is secreted into the environment. Cooperators and defectors are in constant need of a small amount of the public good in order to grow (Figure 4.2).

Figure 4.2– Representation of the behaviour of the two phenotypes.

The production rate, uptake, diffusion and degradation constant of the public good define how far the secreted public good travels, and, therefore its public availability. Realistic values of the aforementioned constants were chosen to define the environment against which the comparison of the different mechanisms will be performed. The specific values are:

Diffusion rate of good	$0.1 \frac{\text{mol} \cdot \text{cell_size}}{dt}$
Degradation rate of good	$0.3 \frac{\text{mol}}{dt}$
Cooperators emission rate of good	$20 \frac{\text{mol}}{dt}$
Absorption rate of good	$3 \frac{\text{mol}}{dt}$

So, cooperators secrete 20 molecules per simulation step into the environment. The secretion is deposited into the environmental grid where it diffuses and degrades until another cell absorbs about 3 molecules each time step. Given the diffusion and degradation rates, the profile of the public good concentration decays as follows:

Figure 4.3– Diffusion of the concentration of good from the cooperator source.

A cell grid away from the cooperator source, the good concentration drops to 10% of its value. Remember that the grid length is roughly half of an adult cell (2 μm). If there was no absorption and only one emission source (a cooperator) the time evolution of the public good concentration for the grid cells #1 (source cell), #2 (1 grid cell from source) and #3 (2 grids away from source) in a single dimension would be:

Figure 4.4— Concentration of good over time for three different cells according to their proximity to the source

So, basically, cooperators are positioned above a large reservoir of public good (that they themselves produce) and have plenty of resources to thrive. Proximal first neighbours that would be located around cell #2 can access a much lower amount of good. Still, the concentration of good is still above the required uptake rate of every cell. The actual profile of public good would ultimately depend on the position of every cooperator and every defector and thus, will depend on the initial random disposition of cells. The presented simplification attempts to justify the realism of the selected values, or at least serve as a characterization of the simulation framework.

The reproductive rate of each cell is a measure of its individual fitness. This reproductive or growth rate depends on various factors. One important factor is the bacterial access and subsequent absorption of the public good. Every simulation step, cells are imposed with the instruction of absorbing a fixed quantity of good. The absorbed good relates to the growth rate of the cell through the public good coefficient. The public good coefficient is calculated according to the absorption of public good at each simulation time step following the function in Figure 4.5. If the available amount of good is lower than the required absorption rate, the public good coefficient will transmit this implication to the bacterium growth rate.

Figure 4.5– Public good coefficient is determined by the availability of good.

Another important factor affecting the individual growth rate is the cost of cooperation. Cooperators have a penalty in their growth rates for producing the public good. The cost coefficient reflects such penalty and is considered only in the calculation of the growth rate of cooperators. In summary, the growth rate of each bacterium at each time step is calculated as:

$$growth_{rate} = Basal\ growth\ rate * Public\ goods\ coefficient * Cost\ coefficient$$

Simulations are initialised with 500 cooperators and 500 defectors. The initial density is chosen to simulate an initial mixing of strains that starts with direct cell-cell contact (high density). Therefore, each strategy will be evaluated with regard to its power to promote cooperation with no initial premises of the spatial structuring (well-mixed initial set-up).

Conclusions of the efficiency of each strategy are drawn from the comparison of two metrics: the relative and the absolute fitness of cooperators. The absolute fitness is directly drawn from the number of cooperators at the final simulation time while the relative fitness refers to the fraction of cooperators, also at the final simulation time.

The cost coefficients are reduced to two values (high and low) because the collective fitness is a sum of the individual fitness which is, in turn, dictated by the individual growth rate of each cell. Therefore, intermediate costs would lead to intermediate fitness and can be extrapolated from the presented data. The effect of the initial density of cells and the initial

frequency of cooperators can be determinant in the evolution of the cooperative phenotype as it has been proposed before in other works (Shigaki et al. 2013). However, the scope of this chapter is to analyse the impact of different strategies in a common scenario for that particular set of parameters. Results should not be taken out of the context of this thesis as it is only a generalized way to qualitatively study the effect of each strategy in a controlled framework.

A summary with the specific parameters and initial conditions common to all simulations in this chapter is in Table 4-1.

Basal growth rate	$0.0347 \frac{\mu m}{dt}$
Low cost coefficient	0.9
High cost coefficient	0.5
Diffusion coefficient of good	$0.1 \frac{mol \cdot cell_size}{dt}$
Degradation rate of good	$0.3 \frac{mol}{dt}$
Cooperators emission rate of good	$20 \frac{mol}{dt}$
Absorption rate of good	$3 \frac{mol}{dt}$
Initial number of cells	10^3
Initial frequency of cooperators	0.5

Table 4-1. General parameters for the public goods dilemma simulations

4.1 Constrained growth in a microcolony of cooperators and defectors

Theory and experiments investigating bacterial colonies on agar have shown that constrained movement can promote the coexistence of defectors and cooperators because cooperators can be more easily surrounded by other related cooperators (Xu & Van Dyken 2017; Kim et al. 2008b; Harcombe et al. 2014; Momeni et al. 2013). An experimental example of the cooperation-stabilizing effect of spatial structure was performed in the social amoeba *Dictyostelium discoideum* (Strassmann 2015). The fruiting body production showed the importance of spatial positioning of related, cooperating individuals. It was found that just millimeters of separation between genetically distinct foraging cells was sufficient to produce clonal fruiting bodies, thereby ensuring the cooperative task of sporulation was shared among relatives. Theoretical work has also supported this theory. Nadell *et al.* used a two-dimensional individual-based model to uncover how cooperative and defector cells can spontaneously segregate from each other in space as the size of a biofilm colony expands (Nadell et al. 2010). They found that the spatial segregation allows cooperative cells to preferentially interact with other cooperative cells, thereby avoiding exploitation by defectors and favoring the maintenance of cooperation.

It was shown in chapter 3 that when bacteria are forced to grow and divide in a two-dimensional surface, space structuring appears as an emergent property. Limited dispersal (*i.e.* population viscosity) generates high degrees of relatedness between interacting individuals because it will tend to keep relatives together, allowing cooperation to be directed indiscriminately toward all neighbors. Limited dispersal is extreme in the case of growth of non-motile individuals, as in the presented case, which leads to the formation of microbial colonies. Furthermore, it was shown how, even with initial mixing, bacterial growth on solid surfaces leads to spatial segregation of genotypes by sector formation. So, in this section, the strategy of spatial constrained growth will be investigated as a feasible strategy to help sustain the initially disperse population of cooperators, when defectors have already infested their space.

Simulations were run for the parameters of Table 4-1. The evolution of the frequency of cooperators is shown for a single run in Figure 4.6. Colony shots were taken for the final simulation time $t=160'$.

Figure 4.6– Evolution of the frequency of phenotypes, cooperators and defectors, for high and low-cost production of public goods. Colony shots depict the growth patterns of cooperators (dark blue) and defectors (light blue).

Clustering of cells due to limited dispersal helps sustain the frequency of cooperators when the cost of cooperation is low (cost coefficient=0.9). However, when cooperation is costly (cost coefficient=0.5), the frequency of cooperators declines in time. Figure 4.6 provides a visual example of the pattering that arises from an initial well-mixed group of cooperators and defectors. Cell arrangement is modulated by the space structuring and the availability of the public good.

The total simulation time allows for an average of 14 generations of defectors (if they have access to the public good). The imposed costs on the growth rate of cooperators returns an average of 12 and 7 generations for low and high-cost, respectively. Another determinant factor for the elected simulation time was the final equilibrium of the ratio of defectors and cooperators. Defectors cannot overtake the entire population as they need the public good produced by cooperators. In circumstances where death is included, the fitness of the population oscillates due to exploitations by defectors, followed by a recovery in which cooperators start to proliferate. Again, this produces that defectors start to thrive, and exploitation happens again. When death is not modelled (as in these simulations), the fraction of defectors converges to an equilibrium value which results as the minimal number of cooperators required to produce the public good and sustain exploitative defectors. This proportion is sustained and constant. For the specific properties of the public good this converges to an 80% of defectors when cooperation costs are high. The average fraction of defectors for the final simulation time (160') is 76% which is close to the equilibrium value (Figure 4.8). Therefore, the simulation time was chosen to permit that the selective pressure and the differential advantage of each phenotype is appropriately reflected in their maximal achievable proportions.

When cooperation is not costly, the average generation time of cooperators and defectors is very similar, and the pattering effect is mostly due to the availability of the public good. The diffusion and degradation rates of the public good, as well as its production rate, are sufficiently high for cells to access it without incurring in further competition. That is why

the pattering of the low-cost simulations resembles that of the control case, where there is no other interaction but the mechanical contact (Figure 4.7).

Figure 4.7– Growth pattern of two bacterial strains interacting only through mechanical contact.

To further investigate how social interactions affect the local arrangement of cells, the metric of *physically connected group* is imported from the previous chapter. In this context, a group is defined by a random cooperative bacterium. The proportion of clonal cells informs about the clustering effect that was studied also in the previous chapter. The proportion of close stable neighbors can be divided into whether they are cooperators (not from the same lineage than the reference cell) or defectors. The relatedness within each group is calculated in the following way:

$$\text{Group relatedness} = \frac{\text{clonal cells} + \text{close cooperators}}{\text{group size}}$$

The group relatedness is a measure of how well distributed the clusters of cooperators are. When the group relatedness correlates with the relative fitness of cooperators, that means cooperators are homogeneously dispersed throughout the colony, and there are no spatial inhomogeneities as the group represents a random sample from the colony. The group relatedness is ultimately a way to quantify the spatial arrangement, from the point of view of the cooperators, and reflects how social interactions affect the arrangement of cells. Generally, the relatedness is a measure of the statistical association between cooperators, and high

relatedness effectively means that recipients of cooperation benefits are likely to be cooperators themselves.

The averaged group relatedness, the relative and absolute fitness of cooperators for $N=20$ simulations at 160' is shown in Figure 4.8. On average, cooperation is only maintained when its costs are low. Restricted growth provides the minimal required benefit for cooperators so they can co-exist with defectors without a competitive disadvantage. However, the higher the cost, the more difficult is to offset the relative advantage of defectors. Restricted growth, and the subsequent assortment of cooperators is not enough to maintain cooperation if cooperation is high-cost. The absolute fitness is given by the number of cooperative cells normalized by the number of cells in a monoculture of cooperators when cooperation is costly (which has an absolute fitness of 1). The absolute fitness measures the average time between divisions and the general reproductive success of cooperators. As observed in the figure, the absolute fitness in co-culture is lowered in half independently of the cost of cooperation with respect to the monoculture counterpart.

Figure 4.8– Averaged group relatedness, relative and absolute fitness for $N=20$ simulations at 160'.

These results demonstrate that even if restricted growth causes cooperators and cheaters to cluster with their respective progeny, this positive assortment can only do so much.

The effect of cell size in cell arrangement and *group* formation was simulated and analysed in the previous chapter. Simulations proved that larger sizes were more competitive when resource acquisition was at stake, but smaller sizes were less spatially disruptive in the process of *group* formation. To elucidate if there is indeed a correlation between size and promotion of cooperation via spatial assortment, different bacterial division lengths were simulated in this same environment. Cells had the same generation time independently of their size, so results avoid the bias of being run for the same simulation time. Costs of cooperation were assumed to be high. Results shown in Figure 4.9 corroborate the initial hypothesis: smaller cells create more stable clusters where cooperators insulate themselves from defectors exploitation.

Figure 4.9– Relative fitness of cooperators for populations constituted with varying cell sizes.

Extreme cases (2μm and 6μm) reflect a difference of 0.2 in their relative fitness (i.e. fraction of cooperators) that can only be accounted for the spatial inability of larger cells to remain closer to their kin.

In summary, simulations have revealed that restricted growth helps sustain cooperation only if the relative advantage of defectors is low. Besides, as it was elucidated in chapter 3, the effect of spatial structuring is enhanced for small cell sizes that form more spatially stable groups. Group stabilization correlates with higher frequencies of cooperators because clustering and positive assortment are enabled by spatial stability.

4.2 Resource limitation enhances cooperation at the colony edge

Previously, it was studied how restricted growth on surfaces can generate explicit spatial dynamics by repeated clonal cell division that form high-density multicellular assemblages. Cell division origins a mechanical force that randomly displaces nearby cells causing daughter cells to remain near their mothers, which creates positive spatial assortment of genotypes. In turn, this positive assortment promotes that the public goods produced by cooperators are excessively received by similar genotypes, providing a means of avoiding exploitation by defectors (Hallatschek et al. 2007; Nadell et al. 2010; Nadell et al. 2016; Momeni et al. 2013; Tarnita et al. 2011; West et al. 2006; West et al. 2007).

In the introduction of this thesis, two theories that attempt to explain the evolution of cooperation were discussed. First, kin selection, in which generosity toward family members aids in the survival of one's genes, and then, group selection, in which a group that cooperates is more successful than one that does not. However, in the absence of kin recognition, cooperative benefits do not discriminate across the population, and benefits are shared with individuals who are spatially nearby (West et al. 2007). In the previous section it was concluded that achieving great assortment highly depends on the competition for space that emerges because of the differential fitness rates. In this context, defectors were able to exploit resources despite the spatial restrictions as the cooperators assortment was too low to properly insulate them.

A scenario which has been proved to enhance positive assortment is resource limitation. A great computational work by Nadell *et. al* showed how the substrate concentration shaped the arrangement of cells that led to strong segregation of lineages. As it was proved in the previous chapter, the lower the resources the higher the competition and thus, the segregation. They further showed how this lineage segregation was beneficial for the maintenance of cooperation (Nadell et al. 2010). It can be postulated that active lineage segregation is then, a passive way to achieve positive assortment.

However, the way resource limitation generates segregation is spatially dependent. Natural existing colonies tend to exhaust resources in the center and expand radially through their borders. During colony expansion, cells at the front experience the highest competition. Yet, cells that make it to the edge can colonize new territory and thus expand without competition from other cells (Lloyd & Allen 2015). Several studies have demonstrated this theory showing that cooperation is promoted at the expanding boundaries of a population. Van Dyken *et al.* proved in an experiment using the budding yeast *Saccharomyces cerevisiae*, that despite losing to defectors in nonexpanding conditions, cooperators increase in frequency in spatially expanding populations. They concluded that cooperation was promoted by increasing positive assortment at population frontiers (Van Dyken *et al.* 2013). In the same year, another experimental work concluded that cooperation is strongly enriched on the front of expanding populations because cooperators can outrun an invading wave of defectors under certain conditions (Datta *et al.* 2013). Theoretical predictions also support this hypothesis. A spatial model of a mixed population of cooperators and defectors shown that cooperators are favoured at the edge of an expanding colony, and under certain conditions, they can spread into new territories faster than they are invaded by defectors (Korolev 2013).

Simulations in this section are also based on the parameters from Table 4-1. However, here, cells are forced to absorb a particular amount of nutrients in order to grow. The intention is to facilitate the segregation of cooperators and defectors, especially at the edge of the colony. Using the same parameters that were explored in the previous chapter, the number of nutrient molecules that was initially placed in each cell grid was 50. In simulations of Figure 4.10 and Figure 4.11, cells are forced to absorb 0.3 molecules of nutrient each time step.

Representative results of the colony patterns and the relative fraction of cooperators is shown in Figure 4.10. The characteristic lineage segregation is more noticeable when the costs of cooperation are low because the relative frequency of each phenotype is more balanced, and so, the expansion at the edge allows the process of colonization for both cooperators and defectors. However, given the defectors requirement of the public good (produce by cooperators), the presence of the defectors subpopulations on the edge are much smaller than

those of cooperators, that are self-sufficient. Therefore, positive assortment is facilitated at the edge, as a consequence of colonization to obtain resources, which substantially increments the relative fitness of cooperators

Figure 4.10– Evolution of the frequency of phenotypes, cooperators and defectors, for high and low-cost production of public goods in a nutrient limited environment. Colony shots depict the growth patterns of cooperators (dark blue) and defectors (light blue).

Average results for $N = 20$ simulations are summarized in Figure 4.11. Results from the previous section are also shown to facilitate comparison. Substantial increments in the relative fitness of cooperators were obtained. Segregation in the high-cost scenario created clusters of cooperators in the colony edge that led to an average of 1.4 increment in the frequency of cooperators, in contrast with simulations with no nutrient dependent growth. Group relatedness was measured at the colony edge since the hypothesis was that it would increase in that location. Error bars are larger for the resource limited case because of the spatial inhomogeneity that highly depends on the initial conditions of each run. Although the initial thesis is correct, and segregation promotes cooperation over cheating, the absolute fitness of cooperators is significantly reduced in this environment. So, there are more cooperators than defectors (especially in the edge), but, overall, there are far few cells than in the previous experiment. The global population is strongly reduced, especially for the high-cost scenario, hinting that population collapse is a probable risk in such a harsh environment.

Figure 4.11– Averaged group relatedness, relative and absolute fitness for $N=20$ simulations at 160'.

To investigate the possible trade-off of that risk, namely, the promotion of cooperation at the edge and the collapse of the population, different nutrient absorption rates were simulated. The total amount of resources was kept constant for all simulations, and only the required intake of the cells was varied.

Figure 4.12—Relative and absolute fitness for different nutrient consumption rates (N=20 simulations at 160°).

Figure 4.12 shows how the relative fitness of cooperators increases when the consumption rate of individual cells rises from 0.1 to 0.4. The increment in consumption is equivalent to a reduction in substrate concentration which is the source of spatial segregation.

For consumption rates higher than 0.4 the relative fitness slightly decreases because the whole population is unable to expand and colonize due to the lack of resources. This effect is reflected in the absolute fitness which decreases as the consumption rate increases, reaching values that indicate that the population hardly expands for that nutrient intake cases (0.5 and 0.6).

As a summary, simulations in this section have revealed that resource limitation enhances cooperation at the edge but reduces global fitness. While resource-limited growth produces an increment in the assortment of cooperators, particularly at the colony edge, it that can be globally impoverishing to the extent of suffocating the whole colony.

4.3 Obligated cross-feeding reduces defector invasion

This chapter aim is to investigate the individual or ecological strategies that sustain cooperators in the face of exploitation by defectors. Another phenomenon that has been confirmed to stabilize the coexistence of both cooperators and defectors is the interaction with a foreign strain. As it was commented in the introduction of this thesis, bacterial social interactions (beneficial or detrimental) are created as a consequence of residing in a common environment (Nadell et al. 2016). In the case of metabolite exchange, strong inter-species dependencies arise as the species that share a common environment evolve together (Germerodt et al. 2016). Different strains can engage in the exchange of essential metabolites leading to altruistic (one way) or mutualistic (bidirectional) interactions. These metabolites-mediated interactions are named syntrophic interactions or cross-feeding. It is the objective of this section to corroborate if cross-feeding of metabolic by-products is a key mediator of positive interactions, as it has been noted before (Mitri et al. 2011; Estrela & Brown 2013; Momeni et al. 2013; Brileya et al. 2013; Müller et al. 2014).

Simulations will try to recreate the experiments conducted by Momeni *et al.*, that experimentally produced two species of yeast that mutually cooperate, each providing a metabolite that was essential to the other. They observed how these species interacted with each other and a third species of yeast that cheated by consuming one of the metabolites without releasing any metabolite of its own. Experiments revealed that the two species that cooperated self-organized into mixed clusters, with the cheating species being excluded from these clusters (Momeni et al. 2013).

First, cross-feeding of two different strains was simulated as it is represented in Figure 4.13. Two strains, S1 and S2, were modelled to secrete a certain metabolite into the environment. Each strain was modelled to require the secreted metabolite of the other one, but not its own. To make things simpler, the physical properties of the metabolites (diffusion and degradation) were the same than those used to model the public good (Table 4-1). Cells from each strain required the same amount of metabolite to survive (3 molecules per simulation step).

Figure 4.13– Cross-feeding growth pattern between strains S1 and S2.

Cells were modelled to produce the metabolite at two distinctive rates. When cells secreted a high number of molecules (20 molecules per simulation step), the metabolite environmental build-up allowed cells to easily absorb their requirements and spatially accommodate into strain-like clusters. Therefore, the obligation is moderate, and the inter-species interaction is considered to be *low* (Low cross-feeding in Figure 4.13).

On the other hand, if the production rate is lowered in half (10 molecules per simulation step), there is no effective accumulation of metabolites and cells are mutually obligated to grow in a highly mixed fashion to maximize the uptake of the counterpart product. Here, the interaction is much stronger than the previous case, and is considered to be *high* (High cross-feeding in Figure 4.13). So, effectively, the only change in the high and low cross-feeding scenarios is the number of molecules each strain produces.

Once a cross-feeding obligated cooperation is established between lineages S1 and S2, simulations investigate the prospects of this association if a fraction of one of the strains starts to defect (Figure 4.14).

Figure 4.14—Cross-feeding mediated cooperation between strains S1 and S2. The strain S1 consists of cooperators and defectors that do not secrete any metabolic product.

In order to remain within the same range of parameters, the costs of metabolite production will be those of the previous simulations. But in this case, the low and high production of metabolite will have according costs. Low production of metabolite, which creates high cross-feeding, will have a lower imposed cost (low-cost penalty from Table 4-1). Accordingly, high producing cells, a condition that generates high cross-feeding, will be imposed with a higher-cost (high-cost penalty from Table 4-1).

Simulations were initialized with 500 cells of strain S2 and 500 of S1. Half of S1 cells are defectors that consume the secreted metabolite from S2, but do not produce the metabolite that S2 requires. Evolution of the frequency of cooperators and defectors of the cross-feeding strain S1 for a single representative run is shown in Figure 4.15, and agree with the averaged results from Figure 4.16, which indicate that strong inter-species interaction (high cross-feeding) works as an insulator against defectors. Low cross-feeding, on the other hand, is more permeable and provides little benefit in comparison with the high-cost cooperation data from section 4.1. Notice that the lower costs are applied in the high cross-feeding system. Lower production enforces cooperation providing an interaction that is stronger than the metabolic advantage of defectors.

Figure 4.15—Evolution of the frequency of cooperators and defectors of the cross-feeding strain S1. Colony shots depict the growth patterns with high and low cross-feeding.

Figure 4.16– Averaged group relatedness, relative and absolute fitness for N=20 simulations at 160'.

The absolute fitness of cooperators is severely affected, even more than in a resource limited environment. This is a consequence of the strong interdependence, which limits the global fitness of the colony. Also, the group relatedness metric has little to say in this scenario, as the mechanism whereby defectors are pushed out is not increased kin interaction (i.e. group relatedness) but increased interaction with a third part. That is why the relative fitness does not correlate with group relatedness as it normally does.

To investigate if the high cross-feeding simulations were more successful because of the associated lower costs, extra simulations were run in high cross-feeding and high-cost production conditions. The question is if the interaction between strains is stronger than the fitness advantage of defectors.

Figure 4.17–Frequency of cooperators of the cross-feeding strain S1 in a long-time simulation

Figure 4.17 shows how even with high-costs, the interaction between strains is able to rescue cooperation. However, the initial well-mixed distribution of strains requires some time for the system to start. Local patches of cooperators from both strains start to proliferate at long times, expanding in defectors-free regions.

As a summary, simulations have revealed that strong obligated cross-feeding is able to rescue cooperation independently of its costs. Mutually cross-feeding cells will grow well towards each other, forming large areas of contact. As a result, defectors are squeezed out, as they do not facilitate the growth of the strain that they depend on. However, this cross-feeding dependency only provides a benefit at the expense of the global colony fitness.

4.4 Conjugation of cooperative genes promotes cooperation

It has been discussed how positive assortment can emerge because of spatial structuring and nutrient limited growth, and how it helps insulate cooperative phenotypes from exploitative cells. When positive assortment of cells occurs, the benefits of the public goods are directed preferentially to the same organisms that are producing them (i.e. cooperators). In this section, another mechanism that bias the partner association in cooperative interactions is simulated and analysed as a promoter for cooperation. The mechanism is the active transfer of cooperative genes (that encode for the production of the public good) to neighbouring cells via conjugation. With the proposed strategy, the clustering and positive assortment of cells arises independently of the spatial constraints.

The transfer of genes between cells is often denoted as Horizontal Gene Transfer (HGT) because it is complementary to the vertical transmission of genes that occurs naturally in cell division (Frost et al. 2005). In general, HGT between bacteria can happen in three ways: transfer of naked extracellular DNA (transformation), transfer through bacterial viruses (transduction) or direct transfer through cell-cell contact (conjugation) (Norman et al. 2009). Hence, conjugation is a biological process by which DNA is transferred from a donor to a recipient cell during cell-cell contact (de la Cruz et al. 2010). Conjugation involves several steps. First, donor and recipient cells are brought into close contact by the process of mating-pair formation through specific organelles called conjugative pili, which are formed by a type IV secretion system (T4SS). Then, when the mating-pair is connected through the secretion system, the donor cell produces a complex formed by the single-stranded DNA and other transports proteins that enable the transfer to the recipient cell, where it is recircularized and replicated (Arutyunov & Frost 2013; Smillie et al. 2010).

It has been observed that specific genes that code for cooperative traits such as the production of a public good are often located on mobile elements, like conjugative plasmids (Dimitriu et al. 2015). A bioinformatics analysis of the *E. coli* genome revealed that cooperative genes are over-represented on plasmids (Nogueira et al. 2009). An example of these genes are the ones

coding for β -lactamases which are secreted proteins that degrade antibiotics extracellularly, and are also predominantly located on plasmids (Touchon et al. 2012). Another fundamental collective and cooperative function that has proved to be associated with plasmid conjugation is biofilm formation (Molin & Tolker-Nielsen 2003; Ghigo 2001).

These evidences suggest a link between the genes involved in social interactions and their mobility, therefore indicating that a natural strategy of bacteria is to grant specific traits to unrelated cells through horizontal transmission. The following testing hypothesis is that the enforcement of cooperation in defectors cells can be achieved through conjugation. If the public good genes are encoded on a mobile plasmid, then, conjugation could convert defector cells into cooperators and, thus, enforce cooperation by infection.

Some previous efforts in modelling have attempted to test this hypothesis. Mc Ginty *et. al* used a population genetic approach to model the evolution of mobile plasmids that encoded genes involved in the production of public goods. They found that due to the increase in relatedness, conjugation of public goods bearing plasmids increased the stability of cooperation via kin selection (Mc Ginty et al. 2013). These results followed preceding theoretical models that also argued that transmission of public goods bearing plasmids reinforces the population fitness benefit in a structured environment, as it increases relatedness on a local scale (Rankin et al. 2011; Nogueira et al. 2009). In another relevant study, which has very much inspired the simulations proposed here in this section, Dimitriu *et. al* experimentally shown that conjugation helps maintain public good production in the face of defectors. They proved that conjugation spreads genes to neighbouring cells, increasing positive assortment and directing a higher proportion of public good benefits to cooperators (Dimitriu et al. 2014).

Considering these evidences, simulations in this section will analyse if conjugation is indeed a powerful mechanism for the maintenance of cooperation, as models and experiments indicate. Simulations will also provide a means to quantify the effect of conjugation in comparison with the other simulated strategies.

Similar to past simulations, half of the initial population will consist on cooperators that contribute to the production of a public good that is required for the appropriate growth of all cells. The other half will consist on defectors that do not contribute to the public good production and, thus, have a competitive fitness advantage. However, here, the cooperative gene (which for the sake of clarity will be represented as a single gene) is placed in a self-conjugative plasmid that is able to transfer to neighbouring cells.

Figure 4.18–Cooperative genes are placed in conjugative plasmids. Transmission of the cooperative plasmid converts defectors in new cooperators.

As indicated in Figure 4.18, the conjugative plasmid is able to transfer without the need of other accessory proteins (green gene). Within that same plasmid the gene in charge of the synthesis of the diffusible public good is placed (blue gene). The dynamics of the mechanism is shown in the right side of the figure. Cooperators cells will infect neighbour defector cells and convert them into cooperators.

gro has been extended to incorporate the logic of plasmid conjugation (Gutiérrez et al. 2017). Basically, each plasmid is modelled with a specific associated transfer rate. That transfer rate is normalized by the generation time of each cell, so eventually, plasmid transfer is simulated as a frequency-dependent process. But, the effective transfer or conjugation rate will depend on the generation time of the bacterial host. Recipient cells are selected randomly from the pool of close neighbours. An average and realistic value for the conjugation rate (or conjugation frequency) is 1 transfers per generation time. So, a plasmid will be able to

conjugate once during its host lifetime, in a random moment of its host cell cycle (Fox et al. 2008; Stalder & Top 2016; Zhong et al. 2012; Andrup & Andersen 1999; del Campo et al. 2012).

In this case, because there are two operations -cooperate and conjugate- additional costs are included. Simulations model scenarios were cooperation and conjugation are either high or low-cost. This results in various combinations for the final growth rate:

$$growth_{rate} = Basal\ growth\ rate * Public\ goods\ coefficient * Cost\ Conjugation * Cost\ Cooperation$$

Three different combinations were used to model cooperators growth rate. The two first combinations occur when the cost of conjugation and cooperation are equal: both conjugation and cooperation are either high-cost or low-cost. The third possibility is that one operation is high-cost while the other is low-cost. Because the effective growth rate does not discriminate which operation is imposing the cost, there is only one crossed combination, which is labelled as medium-cost. The final cost coefficients are:

$$High\text{-}cost\ coefficient=0.25 \quad Medium\text{-}cost\ coefficient=0.45 \quad Low\text{-}cost\ coefficient=0.81$$

Notice that the low-cost coefficient here is higher than the canonical low-cost (which was 0.9) and that the penalty for conjugation imposes a higher burden in general.

In Figure 4.19 the time evolution of the frequency of cooperators and defectors for the three different cost (individual growth rates of cooperators) is shown. As it can be observed in the graph and in the averaged bar graph of Figure 4.20, results show that even with higher costs the relative and absolute fitness of cooperators is increased. If these results are compared with the simplest strategy of section 4.1 (merely growing in a spatially structured environment), the relative fitness of cooperators (i.e. relative frequency of cooperators) is 1.25 times higher when cooperation is high-cost in both scenarios, even though the cost in this strategy is doubled. In the case of low-cost cooperation and conjugation, the relative fitness increases 1.8 times and returns a population of 95% of cooperators for $t=160'$. However, the absolute fitness of cooperators (i.e. normalized number of cooperators) is differentially lower: 1.4 and 1.2 times lower for the high-cost and low-cost scenario respectively.

Colony shots show how cooperators form clusters with varying size depending on the cost of the strategy. These clusters are formed by division due to restricted growth (as in section 4.1) and principally because of conjugation. This is reflected as well in the group relatedness metric of Figure 4.20. Group relatedness increases for all costs, reaching an averaged value of 0.95 when cooperation and conjugation are low-cost. Group relatedness for high-costs increases from 0.3 to 0.7 (2.3 times higher). This is a direct consequence of the plasmid transfer that transforms close-by defectors in cooperators within the same group.

Figure 4.19–Time evolution of the frequency of cooperators and defectors when the cooperative trait is placed in a conjugative plasmid. Colony shots depict the growth patterns of cooperators (dark blue) and defectors (light blue) within the same colony.

Figure 4.20– Averaged relative and absolute fitness of cooperators and group relatedness for N=20 simulations at 160’.

A representative run of the evolution of the *group* is shown in Figure 4.21. While the fraction of clonal cells remains constant, the fraction of cooperators increases as the fraction of defectors decreases. Surprisingly, the *group* relatedness (Figure 4.20) general shape does not correlate with the relative fitness. High and medium-costs return almost the same averaged relatedness, despite presenting different relative and absolute fitness. This is most certainly a consequence of the frequency-dependent process of conjugation. Even if cells are simulated with the same conjugation frequency, since it is relative to the cell growth rate, high and medium-cost scenarios are limited by the transfer events cells experience during the simulation time. Conjugation events are so rare that, even when increasing the group relatedness, the system does not differentiate below the limit of the medium case.

In summary, conjugation has proven to be a good strategy for the maintenance of cooperation if both operations (conjugation and cooperation) are selected together and co-evolve to minimize costs. In that case, when costs are in the low-medium range, conjugation of cooperative plasmids helps sustain and even increases the fraction of cooperators.

Through the increase of genetic similarity between adjacent cells, conjugation is able to construct a social niche in which the cooperative phenotypes are selectively advantageous.

Figure 4.21–Time evolution of the *group* members when conjugation costs are low.

Cooperative plasmid transfer competing with a non-cooperative conjugative plasmid.

In this subsection, simulations study the competition between producer and non-producer conjugative plasmids in a structured population. This situation can arise if the cooperative gene that is encoded in the conjugative plasmid gets lost by deletion, a very likely scenario (Merhej et al. 2009; Haft et al. 2009; Garcillán-Barcia et al. 2011; Johnson & Nolan 2009; Keen 2012). So, the question is if conjugation of plasmid bearing cooperative genes can still maintain cooperation if a fraction of those plasmids lose the cooperation gene, and thus, become mutated defectors plasmids. Defector plasmids would still be able to conjugate and spread across the population but with no effect to the cooperators cause. This situation would seem little threatening if it was not by the incompatibility effect. Assuming that both plasmids, cooperative and defector, are identical apart from the cooperative gene(s), this results in an incompatibility problem. Plasmid incompatibility is caused by sharing the same genes that control plasmid copy inside their host, which results in mutual inhibition of replication (Novick 1987; Smillie et al. 2010; Garcillán-Barcia & de la Cruz 2008; Gooding-townsend et al. 2015).

Therefore, incompatible plasmids cannot be stably maintained within a host cell. It has been seen how this effect can lead to the rapid displacement of one of the plasmids (Gooding-townsend et al. 2015). Most certainly, the plasmids that would be displaced would be the one that spreads more slowly. So, the hypothesis is that if defector plasmids have the same or higher transfer rate compared to cooperative plasmids, cooperation would again be threatened by the spread of cheaters, now at the plasmid level.

The logic of this situation is translated into the simulated population as it is represented in Figure 4.22. Initially, cooperative cells (bearing the cooperative plasmid), defectors with plasmid (bearing the defector plasmid) and plasmid-free defectors, are placed together in the simulated plate. The expected dynamics is that cooperators and plasmid defectors will engage in a competitive arms race. Both strains will attempt to infect plasmid-free defectors and turn them into one of their own, removing the possibility of conjugation for the other strain.

Figure 4.22– A population with cooperative cells (bearing the cooperative plasmid), defectors with plasmid (bearing the defector plasmid) and plasmid-free defectors. Transmission of any plasmid to a plasmid-free defector prevents such defector to be further infected by the other plasmid.

The parameters of the simulation remain the same as in the previous experiment. Both plasmids conjugated at a frequency of 1 events per cell life time. However, since it was already proved that high-cost conjugation and cooperation did not rescue cooperation in the former

experiment, that condition has been obviated here. The growth rate of defectors remains the same (found in Table 4.1) while the growth rate of cooperators will depend again on the imposed costs:

Medium-cost coefficient=0.45 Low-cost coefficient=0.81

The growth rate of plasmid defector cells will suffer the penalty costs only of conjugation, with their according costs:

Medium-cost coefficient=0.5 Low-cost coefficient=0.9

Notice that the costs used for defectors correspond to the high and low costs initially proposed (medium costs in this case correspond to high-costs in the former case). Still, plasmid defector cells are conferred with the advantage of avoiding the costs of cooperation and thus, will have faster reproduction rates. Simulations explore this scenario for different initial fractions of the defector plasmid. Figure 4.23 shows the time evolution of cooperators if the initial fraction of defectors plasmids is 5%. Specifically, simulations were initialized with 500 cooperators, 500 plasmid-free defectors and 50 plasmid defector cells.

Figure 4.23–Time evolution of the frequency of cooperators when the competitive defector plasmid composes 5% of the initial population, for medium and low-cost of conjugation and cooperation.

As observed, the effect of the competitor plasmid is harmless in such low densities. Average results are shown in the bar graph of Figure 4.25. The relative fitness of cooperators remains practically unaltered in this scenario. However, further initial addition of competitor plasmids reveals an interesting effect.

Figure 4.24 shows evolution of the frequency of cooperators and defectors (including plasmid-free and plasmid-bearing defectors) for two representative runs. Simulations with an initial 20% of defector plasmids were initialized with 500 cooperators, 500 plasmid-free defectors and 250 plasmid defector cells, while simulations with an initial 33% were initialized with the same number of cells of each strain (500).

Figure 4.24– Time evolution of the frequency of cooperators when the competitive defector plasmid composes 20% and 33% of the initial population, for medium and low-cost of conjugation and cooperation.

Average results are summarized in Figure 4.25. The final frequency of cooperators (*i.e.* the relative fitness of cooperators) drops for higher fractions of the competitor plasmid. Still, the effect of the competition of the defector plasmid does not imply a failure for the strategy of conjugation. The fraction of cooperators is able to rise above the 50% in all scenarios, regardless of the cost or number of defector plasmids.

Figure 4.25– Averaged relative and absolute fitness of cooperators and group relatedness for N=20 simulations at 160' for different initial fractions of the competitor plasmid.

In both final setups (20% and 33% of defector plasmids), the initial fraction of cooperators (40% and 33%, respectively) was able to overcome the relative frequency of defectors.

This is rather surprising considering that cooperators have lower reproductive rates, and consequently, less conjugative events (remember that the frequency of conjugation is normalized by the host generation time). This effect is due to the dependence that all cells have of the public good produced by cooperators. The regions where defectors plasmids are not near cooperative cells, their host cells cannot reproduce, lowering or stopping altogether their ability to conjugate.

In a work carried out by Dimitriu et al., they presented experimental and modelling evidence of a very similar set up (Dimitriu et al. 2014). They wanted to measure if genetic transfer of cooperative genes could sustain cooperation. They measured the proportion of producers in a configuration where producers, producers with the ability to transfer the cooperative plasmid, non-producers and non-producers with a competitor plasmid were mixed in different proportions. They found that cooperation could not be maintained in the absence of transfer and that transfer is not sufficient to maintain cooperation in a well-mixed population, if both producer and non-producer alleles are mobile. However, similar to the results here presented, they found that when the population is already structured, there was an increase in the global proportion of cooperative producers as transfer reversed the direction of the selective pressure to favour them.

So, to summarize this section, the transmission of good production genes via conjugation has proved to be a successful strategy for the promotion of cooperation as long as the costs in conjugation are moderate. Simulations showed that even for medium costs, which result in an increase of almost double the reproductive rate of non-producers, transmission could maintain and even increase the population of cooperators.

4.5 Density conditional cooperation is beneficial independently of its costs

Quorum sensing (QS) is a form of cell–cell communication observed across diverse bacterial species (Keller & Surette 2006). It involves the secretion of diffusible signals into the environment that, when sensed by other cells, regulate gene expression in response to the signal concentration. In this process, cells would only trigger gene expression until a critical threshold of signal concentration (and consequently population density) is reached. Microorganisms use QS mechanisms to sense the existence of neighbouring species, measure their cell densities, and coordinate their group behaviors to motivate social traits of the microbial population accordingly (Waters & Bassler 2005). It has also been proved to control other behaviors such as virulence (Rutherford & Bassler 2012) and biofilm formation (Nadell et al. 2008).

Many of the QS-regulated cells exhibit signs to be subjected to the public good dilemma, since nonresponsive QS mutants function as social cheats that not produce and not respond to the QS associated signals (Diggle et al. 2007). However, this could be used as an advantage for QS sensitive cells if the production the public good was coupled to the secreting and sensing of QS signals. As a matter of fact, it has been suggested that QS has evolved in part to restrict secretion investments when they are inefficient (Cornforth et al. 2014). The producers' strategy could be to first, sense the density of producers and then, only if enough density of producers is reached, secrete the public goods. This strategy would limit the cooperative investments to social environment enriched for cooperators. It would also allow cells to cooperate when it is effective, thus saving resources under low density conditions. So, the hypothesis is that QS regulation of public goods production will limit the ability of defectors to socially exploit cooperators in a high-density environment (Allen et al. 2016).

Several computational works have supported this hypothesis (Cavaliere & Poyatos 2013; Czárán & Hoekstra 2009; Schluter et al. 2016), and there is experimental evidence as well

(Drescher et al. 2014). Therefore, the task here is to re-evaluate to what extent this strategy is beneficial, especially in comparison with those already simulated, such as conjugation of cooperative genes.

Simulations in this section will analyse if density conditional cooperation is indeed a powerful mechanism for the maintenance of cooperation (as models and experiments indicate). Simulations will also provide a means to quantify the effect of this strategy in comparison with the other simulated strategies.

Similar to past simulations, half of the initial population will consist on cooperators that contribute to the production of a public good that is required for the appropriate growth of all cells. The other half will consist on defectors that do not contribute to the public good production and, thus, have a competitive fitness advantage because of their higher reproductive rate. However, here, the cooperative gene (which for the sake of clarity will be represented as a single gene) is regulated under a QS promoter (Figure 4.26).

Figure 4.26– Cooperative genes are placed under the regulation of a QS gene. Only when quorum is reached (enough density of QS emitters, yellow signal) cooperators secrete the public good (blue signal).

Secretion of public good will be restricted to cooperative cells that are surrounded by enough cooperative cells that are emitting and sensing QS in order to be recognized. Once and only when a quorum is reached, cooperators would start to produce the public good.

The physical properties of the QS molecule are those so groups of cooperators need a range of 30 to 50 cells to reach the required concentration threshold. Higher cell densities cannot be achieved in the initial setup as only 500 cooperative cells are randomly placed in the simulated plate. If the quorum threshold were to be higher, cells cannot reach quorum (there is a low probability of having an initial cluster of more than 50 cells) and they cannot start the emission of the public good. This results in the suspension of QS and therefore the system is unable to start. The number of initial cells was kept the same for all sections in order to have a fair comparison, so the minimum density corresponds to 30 cells if they are forming a compact cluster. Used values of the QS molecule are found in Table 4-2.

Diffusion rate	$0.2 \frac{\text{mol} \cdot \text{cell_size}}{dt}$
Degradation rate	$0.5 \frac{\text{mol}}{dt}$
Cooperators emission rate	$1 \frac{\text{mol}}{dt}$
Cooperators threshold for quorum	$8 \frac{\text{mol}}{dt}$

Table 4-2. Properties of the simulated QS molecule

In this case, because there are two operations –secretion of public good and QS- additional costs are included. Simulations model scenarios were cooperation and QS emission are either high or low-cost. This results in various combinations for the final growth rate:

$$growth_{rate} = \text{Basal growth rate} * \text{Public goods coefficient} * \text{Cost QS} * \text{Cost Cooperation}$$

Three different combinations were used to model cooperators growth rate. The two first combinations occur when the cost of QS and cooperation are equal: both QS emission and cooperation are either high-cost or low-cost. The third possibility is that one operation is high-cost while the other is low-cost. Because the effective growth rate does not discriminate which operation is imposing the cost, there is only one crossed combination, which is labelled as medium-cost. The final cost coefficients are:

$$\text{High-cost coefficient}=0.25 \quad \text{Medium-cost coefficient}=0.45 \quad \text{Low-cost coefficient}=0.81$$

Notice that the low-cost coefficient here is higher than the canonical low-cost (which was 0.9) and that the penalty for QS emission imposes a higher burden in general.

In Figure 4.27 the time evolution of the frequency of cooperators and defectors for the three different costs is shown. As it can be observed in the graph and in the averaged bar graph of Figure 4.28, results show that independently of the costs, the relative fitness of cooperators is increased.

Figure 4.27– Time evolution of the frequency of cooperators and defectors when public good secretion is conditional to the density of cooperators.

The representative run that is shown in Figure 4.27 depicts how the relative frequency of cooperators (*i.e.* relative fitness) starts at 0.5 in the beginning of the simulation and rapidly grows until it reaches a fraction of around 0.65, then growth slows down until it finally reaches a fraction of 0.7 at $t=160'$. This dynamic is given by the underlying rules of the mechanism and is shown in every run in a fashion that is independent of the costs of public good and QS secretion. It is during that early stage that the initial groups of cooperators start to produce the good, which is first consumed by them, to finally be reached by the defectors. Colony shots of Figure 4.27 also depict the group formation of cooperative cells where defectors are left to grow in the interface of such clusters.

Figure 4.28— Averaged relative and absolute fitness of cooperators and group relatedness for N=20 simulations at 160'.

The simulated mechanism presents an important trade-off between the reproductive advantage that defectors have over cooperators and the global requirement of public good.

Since defectors have restricted access to the public good in this scenario specially, two problematics present. The defectors advantage over cooperators diminish with costs (reduces the reproductive advantage) but also diminish if the public good is produced in less quantity, as it happens when the costs are high (higher costs imply less producers cells, and so less leftovers for defectors). This trade-off is reflected in the group relatedness of each cost scenario, which is shown in Figure 4.28.

In summary, density conditional cooperation (cooperate only when surrounded by cooperators) is a strong protection mechanism against exploitation by defectors. The mechanism is cost independent for the relative fitness although the absolute fitness of cooperators strongly resents the costs effect on the number of cells.

4.6 Selfish good production promotes cooperation

The final studied mechanism for the promotion and maintenance of cooperation is the selfish production of public good by cooperators. When the secreted public good is costly to produce, the evolutionary fate of cooperation is determined by whether or not the shared good benefits cooperative cells rather than defectors. This will depend on how far the secreted public good travels, which is a factor determined by its production, its transport by advection and diffusion, its uptake and its decay. When the spatial scale of public goods sharing is similar to the spatial scale to that of the cooperative group, public goods dilemmas can be resolved. Therefore, clonal clustering tends to promote the evolution of public good production, as long as the public good does not rapidly diffuse throughout the system and is available for other cell types rather than their producers. So, the tested hypothesis is that cooperation can also be stabilized if a small fraction of the public good is retained by the producing cell (Figure 4.29). This can be achieved by two means. Either the cell internally retains part of the required good or it reduces its production and diffusivity so the main beneficiary is itself, or, at most, its closest neighbors.

The second scenario has been observed experimentally. Experiments on agar plates have shown that siderophore secretion by *P. aeruginosa* is more strongly favoured as agar concentration is increased, because this decreases public good diffusivity and limits the receipt of cooperative help to neighbouring cells (Kümmerli et al. 2009). The first scenario has been considered by a theoretical model that show how by keeping some fraction of the produced good for personal use cooperating cells coexist with cheaters (Prajapat et al. 2016). Their results suggest that individuals exhibiting a little selfishness while still contributing to the population are best suited to resist cheater invasion. Another theoretical study also showed that cooperation is lost beyond a critical diffusivity of public goods. Furthermore, for the case of unequal diffusivity of public goods, the species with slower-diffusing public goods will dominate the co-culture and destroy cooperation by driving the other species to extinction (Menon & Korolev 2015).

Simulations in this section will analyse if the selfish production of public goods (Figure 4.29) is indeed a powerful mechanism for the maintenance of cooperation (as models and experiments indicate).

Figure 4.29–Cooperative cells become a little selfish by reducing the emission and diffusivity of the emitted common good so they become the main beneficiary.

Similar to past simulations, half of the initial population will consist on cooperators that contribute to the production of a public good that is required for the appropriate growth of all cells. The other half will consist on defectors that do not contribute to the public good production and, thus, have a competitive fitness advantage because of their higher reproductive rate. However, here, the public good properties are changed from past simulations. Now, the emission rate is more than three times smaller, the diffusion coefficient is half and the degradation rate is almost double than the previous case (generous scenario, parameters from Table 4-1). The new properties of the diffusible good are enlisted in Table 4-3.

Diffusion coefficient of good	$0.05 \frac{mol \cdot cell_size}{dt}$
Degradation rate of good	$0.5 \frac{mol}{dt}$
Cooperators emission rate of good	$6 \frac{mol}{dt}$
Absorption rate of good	$3 \frac{mol}{dt}$

Table 4-3. Properties of the selfish public good.

The associated cost coefficients remain as in previous sections: 0.5 penalty for high-cost and 0.9 penalty for low-cost secretion. In Figure 4.30 the time evolution of the frequency of cooperators and defectors for the different costs is shown for a single representative run.

Figure 4.30– Time evolution of the frequency of cooperators and defectors when public good secretion is selfish (lower production rate and diffusivity).

This strategy not only maintains cooperation but it consistently increases the frequency of cooperators for both the high and the low-cost scenario. Average results shown in Figure 4.31 agree with the runs of Figure 4.30. Cell arrangement can be observed as well in the colony shots. The reproductive advantage of defectors is overcome by cooperators that cluster to meet the more demanding requirements of public good of this scenario. This mechanism significantly reduces the advantage of defectors that survive by spreading over the colony

attempting to maximize their access to the short-lived public good. Even if costs of production return similar relative fitness of cooperators, their effect is observable both in the group relatedness and the absolute fitness of cooperators. High-costs cooperation is still globally impoverishing although it does not reach the critical level of other simulated strategies. On the plus side, the low-cost strategy is able to return the best absolute fitness of cooperators apart from the control case of the first section.

Figure 4.31– Averaged relative and absolute fitness of cooperators and group relatedness for N=20 simulations at 160’.

In this section it was proved how the preferential access to resources promotes cooperation especially if the scale of diffusion of the cooperative secreted compounds is similar to the natural clustering of cooperative cells due to division. This was achieved by lowering the production of the secreted good while at the same time reducing its diffusivity and its half-life. So far, this strategy has delivered the best results in the promotion of cooperation offering simultaneously the greatest relative and absolute fitness of cooperators for the final simulation time.

4.7 Combined strategies

After simulating and analysing each of the proposed strategies in an artificially controlled environment, here, simulations combining some of the strategies will study the impact and the possible synergetic effect they can have in maintaining cooperation. As the more promising results were obtained for the selfish strategy, where the production and transport of the public good preserved it closer to the producer, this strategy will be studied in combination with resource limitation, conjugation and QS-triggered cooperation. Finally, both conjugation of the cooperative genes and QS-triggered cooperation will show how these natural bacterial mechanisms provide the perfect recipe for the promotion of cooperation.

Selfishness and resource limitation

The first combination of strategies is the selfish production of public goods in a scarce resource environment. Simulations were run using the same parameters as in section 4.6 but here cells were forced to consume nutrients for growth. The individual consumption rate of nutrients was set to 0.3 molecules/dt. Representative colony shots of the two scenarios (high and low-cost cooperation) are shown in Figure 4.32.

Figure 4.32– Colony shots for simulations of selfish production of public goods in a resource limited environment. Growth patterns of cooperators (dark blue) and defectors (light blue) are shown for high-cost and low-cost cooperation.

Data from the simulations is summarized in Figure 4.33. The frequency of cooperators (i.e. relative fitness) and the normalized total number of cooperators (i.e. absolute fitness) are shown for these actual simulations (selfishness + resource limitation) in addition to the results provided by the individual strategies. The combined effect of both conditions does not deliver a better fitness than the case of the selfish strategy alone. Growth in a competitive environment where nutrients are scarce results as an imposition that only lowers the relative and absolute fitness of the cooperators. The effects of a more disruptive space structuring are irrelevant when compared to the limited production of goods that the selfish strategy utilizes. If this scenario is observed from another angle, where resource limitation is not a strategy but rather an imposed consequence of the environment, the fundamental conclusion is that selfishness remains being a valid strategy for cooperation not to be outcompeted. Actually, in a low-cost scenario, the frequency of cooperators is the same independently of the nutrient limitation.

Figure 4.33—Average relative and absolute fitness for the simulation of selfish production of public goods in a resource limited environment.

Selfishness and Conjugation

The second combination of strategies is the selfish production of public goods by cooperators that can transfer the cooperative gene, which is encoded in a conjugative plasmid. Simulations were run with the parameters defined in section 4.6 (public good production and transfer properties) and section 4.4 (conjugation frequency and derived costs from conjugation). Results, which are depicted in Figure 4.34, show that conjugation of the cooperative genes and selfish production of public goods indeed returns a higher frequency of cooperators when used in combination. Even if the improvement is differential when compared with the selfish strategy alone, it does provide a larger advantage for the conjugation strategy if the expenses of cooperation are high. Notice that the low-cost scenario does not reflect significant differences neither in the relative nor in the absolute fitness. This implies that the cost of cooperation is more fundamentally determinant for the prospect of the cooperators than any proposed strategy.

Figure 4.34— Average relative and absolute fitness for the simulations of selfish production of public goods when cooperative genes are encoded in a conjugative plasmid.

Selfishness and QS

The third combination of strategies is the selfish production of public goods by conditional cooperators that utilize QS to measure the density of cooperators. Simulations were run with the parameters defined in section 4.6 (public good production and transfer properties) and section 4.5 (properties of the QS molecule). Results, which are depicted in Figure 4.35, show that there is no additional benefit in using this two strategies together when costs are high as the resulting relative fitness is slightly lower than that of the QS alone and significantly lower than the selfish strategy alone. Only when costs are low does the QS strategy benefits form a selfish production of public good. However, in bot costs scenarios, the absolute fitness is reduced when compared to each strategy separately.

Figure 4.35– Average relative and absolute fitness for the simulations of selfish production of public goods when cooperation is conditional to the local density of cooperators and regulated by QS.

Conjugation and QS

Final simulations consider the plausible combination of both density conditional cooperation regulated by QS and conjugation of cooperative genes. Both mechanisms are complementary and represent opposite strategies of communication in the bacterial world. On one hand, emission and reception of QS represent a fast mechanism for communicating across a large spatial scale. On the other hand, conjugation allows the transmission of larger amounts of information. However, the process of transmission is slow and is confined to individuals close to the message source. Here, I propose the combination of both mechanisms in a sequential manner.

First, if the initial density of cooperative cells is below the threshold required for quorum, the mechanism of conjugation would improve the group relatedness by transmitting the cooperative genes and the QS density sensor (Figure 4.36). That process would result in larger clusters of cooperators that could then reach a consensus and start producing the public good.

Figure 4.36—Conjugative plasmids include cooperative and QS genes. Conjugation spreads the cooperative plasmid (t_1) allowing for the subsequent quorum (t_2) and emission of the public good (t_3).

Simulations were run with the parameters defined in section 4.4 (conjugation frequency and derived costs from conjugation) and section 4.5 (properties of the QS molecule and derived costs for QS production). The combination of both strategies plus the costs of public good production result in the highest penalty for the reproductive rate of cooperators in all the

presented simulations. In these simulations, only the combinations for high-cost and low-cost of all processes was studied.

Results of the relative and absolute fitness of cooperators as well as the group relatedness are shown in Figure 4.37. Despite the large costs that the combination of all the mechanism imply, results indicate how the relative fitness of cooperators is not only increased but it basically reaches its maximum value. When using both mechanism, defectors are almost extinguished. This is possible because cooperation is conditional to high densities of cooperators, which drastically reduces the advantage and number of defectors. The remaining defectors are at the same time converted to the cooperative phenotype by the transmission of the implicated genes.

Figure 4.37– Averaged relative and absolute fitness of cooperators and group relatedness for N=20 simulations at 160’.

To my knowledge, the reported process has not yet been experimentally or theoretically investigated even though it would probably take place in natural bacterial colonies. It would be interesting to find out if the use of both mechanisms does indeed impose an unbearable metabolic burden, and so, bacteria evolve to primarily use one or the other.

4.8 Conclusions

This chapter has investigated the effect of different environmental conditions and mechanisms for cooperative producers to avoid being exploited by non-producing cheats in a public good sharing colony. Simulated colonies were composed of an equal initial mix of cooperators and defectors, which competed using different strategies. Each individual cell has a fitness that depends on its strategy and on the strategy of the individuals with whom it interacts. The objective of each strategy was to increase or, at least, maintain the relative frequency of cooperators throughout the simulated experiment. Individuals that used more successful strategies had higher chances to propagate and thus, of increasing their frequency in the community. Specific conclusions withdrawn from each section are now enlisted:

- It was shown in chapter 3 how restricted growth on a two-dimensional surface amplifies the positive assortment of cells by keeping clonal cells together after division. However, this effect is not strong enough to prevent exploitation by defectors if the costs of good production impose a metabolic burden greater than 10%. Coexistence of both phenotypes is possible if costs lie within that 10%, otherwise defectors outnumber cooperators at a constant rate over the course of 10 generations. The effect of spatial structuring is enhanced for small cell sizes that form more spatially stable groups. Cell size correlates with higher frequencies of cooperators because clustering and positive assortment are enabled by spatial stability. Downsizing cell length from 4 μm to 2 μm increases 1.3 the relative fitness of high-costs producers but, still, is insufficient to rescue the cooperative phenotype.
- Resource restriction increases the natural positive assortment of cells and enhances cooperation by segregating cells at the colony edge (where colonization takes place). However, a highly nutrient restricted environment can diminish the absolute fitness of cooperators (the number of divisions) to the extent of suffocating the whole colony.
- If a species engages in metabolic cross-feeding with another species, the stronger the interaction the more difficult it would be for a mutant defector to take advantage. Cross-

feeding cells will grow towards each other, forming large areas of contact. As a result, defectors are squeezed out. This interaction is able to rescue cooperation independently of its costs but, similar to the resource restricted environment, at the expense of the global colony fitness.

- Through the increase of genetic similarity between adjacent cells, the transmission of cooperative genes via conjugation is able to sustain and even increase the relative fraction of cooperators. Despite being a strategy that is susceptible of failure due to plasmid loss or gene deletion, competition with a defector plasmid can be overcome even at high frequencies of infectivity. Conjugation of cooperative plasmids has proven to be a good strategy for the maintenance of cooperation if both operations (conjugation and cooperation) are selected together and co-evolve to minimize costs. Still, co-existence of both phenotypes is achieved even when defectors have twice the reproductive rate of cooperators. Besides, the absolute fitness of cooperators is not as severely affected as in the previous conditions.
- If the secretion of the common good is conditional to the local density of cooperators, the cooperative phenotype is able to avoid exploitation independently of the associated cost of good and QS production. Cooperators can thrive in large groups whereas defectors are left to survive in the interface of such regions. This dynamic does not allow to eliminate defectors but permits the coexistence of both, with cooperators being the advantageous phenotype. A final fraction of 70% of cooperators is reached independently of the cost of the strategy. High-cost simulations showed that, even if the strategy is cost-invariant, it can impose a high burden as the final number of cooperators lies far beyond the control case.
- If the diffusivity of the public good is reduced so that the spatial scale of the good is similar to the scale of clonal clusters, then the molecule benefits are mostly directed towards cooperators rather than defectors. This strategy promotes cooperation to the point of extinguishing cheaters. Again, this is achieved regardless of the assumed costs of cooperation. Modulation of signal diffusivity and half-life provides a great advantage at virtually no cost. Since this strategy has proved to be so powerful and simple, simulations

using the new diffusion and degradation constant were conducted, leading to the following results:

- A selfish production of public goods does not benefit from the lineage segregation that a scarce resource environment provides. Growth in a competitive environment where nutrients are scarce results as an imposition that only lowers the relative and absolute fitness of the cooperators.
- Conjugation of cooperative plasmids when the production of goods is restricted increases the success of the conjugation strategy. This combination results as the most beneficial strategy for cooperation promotion of all those presented here.
- Density conditional cooperation does not benefit from the selfish production of goods unless the associated costs are low.
- The combination of both conjugation of cooperative genes and QS controlled cooperation is able to offset the large costs they impose on cooperators and leads to defectors population to almost complete extinction.

Figure 4.35 depicts a summary with the relative fitness of cooperators for all experiments performed in this chapter. Notice that the strategies have been ordered according to their contribution to the fitness of cooperative producers. The colour of each bar represents the imposed costs of cooperation for each particular situation. Three distinctive regions can be identified, which are indicated by grey dotted lines. The left region includes the strategies that returned a relative fitness of cooperators lower than 0.4. This means that those strategies were not good enough for the maintenance of cooperation under the assumption taken in this chapter. Notice how those strategies are all using high-cost mechanisms which could not offset the benefits of defectors.

Figure 4.35—Relative fitness of cooperators of all strategies and conditions for $t=160'$. The colour code of each bar represents the associated cost of cooperation and the strategy in each case.

The middle region reflects strategies which were able to maintain the relative fraction of cooperators between 0.5 and 0.7. Such strategies deliver a regime where coexistence with exploitative defectors is possible and the public good dilemma is resolved. Strategies include various costs for the involved mechanism which indicate their robustness if applied in more realistic contexts.

Finally, the right region includes those strategies that were able to almost eliminate the defector phenotype. As expected, that region is mostly consisted on low-cost mechanisms. However, selfishness and the combination of QS and conjugation belong to that category for both high and low-costs.

Yet, the metric of the relative fitness is not complete and needs to be complemented with the absolute fitness values. This is given in Figure 4.36 where the absolute fitness of cooperators

is represented for all the simulated cases following the order of Figure 4.35. The absolute fitness value indicates the final number of cooperative cells for the final simulation time which is a measure of the number of divisions and the global viability of the population. The reference value that normalizes the metric is the number of cells in a monoculture of cooperators in a high-cost production regime. Values below 0.5 of absolute fitness indicate that after 160' populations were unable to reach half of the number of cooperators. As expected, almost all strategies that were assuming low-costs are above or around 1. Selfishness, conjugation of cooperative genes and their combination when the costs are low result as the strategies that maximize at the same time the relative and the absolute fitness of cooperators.

Figure 4.38–Absolute fitness of cooperators of all strategies and conditions for $t=160'$. The colour code of each bar represents the associated cost of cooperation and each strategy (if required) in each case.

Chapter 5 Division of labour

Multicellularity is considered one of the major transitions in evolution, a concept that encompasses the phenomena in which natural selection transforms autonomously replicative genes/cells/individuals into new higher-level structures. In such structures, individuals are no longer independent of each other but rather form a group that coordinates and communicates with a resulting behaviour that is determined by the interaction of its fundamental elements (Szathmáry & Smith 1995). The evolution of life on Earth has involved approximately eight major transitions, where a group composed of initially independent and self-replicant individuals form a more complex life form that can only replicate as a group (Biernaskie & West 2015). One of the major transitions is multicellularity. It has been identified that multicellularity has evolved independently more than 20 times, at least once in animals, three times in fungi, six times in algae, and multiple times in bacteria (Lyons & Kolter 2015). The major challenge is to explain why single cells should join and become mutually dependent, in a way that leads to a more complex life form that can only replicate as a whole. Explaining the transition from single cells to multicellular organisms is still a major challenge for evolutionary biology (De Monte & Rainey 2014).

Multicellularity has involved different evolutionary pathways, leading to the belief that there is not a single explanation for its origins. However, some requirements have been identified across many species and are thought to be essential to meet the current definition of what constitutes a multicellular organism (Niklas 2014). The evolutionary requirements for multicellularity include (i) the formation of clusters of cells via physical contact or adhesion to form a new evolutionary unit, and (ii) communication among the individuals leading to coordinated activity (Figure 5.1).

There are some essential disadvantages to multicellularity. For example, it is much costly to adhere and communicate with your neighbours than to live a solitary life. Employing strategies that assure communication and encourage cooperation to reduce the possible

conflict within the group does not come for free (De Monte & Rainey 2014). Yet, these pitfalls can be greatly compensated by the benefits that multicellularity provides. It has been proved to improve: colonization of new territories, predation avoidance, resource acquisition, metabolic efficiency, higher resistance to physical and chemical stresses and, finally and more importantly, it implies cell differentiation and division of labour (Bonner 1998).

Figure 5.1– Steps for the evolution of multicellularity. Cell adhesion and cooperation (group formation) are major requisites to constitute a multicellular organism that is constituted of specialized cells that are able to divide labour (group transformation).

Division of labour requires the coexistence of interacting specialized cells that carry out complementary tasks. Accordingly, specialization refers to a process where the individual transitions from a state where in principle many functions could be performed, to a configuration where it mainly develops a specific task. The specialization of some individuals of the population usually implies that they become dependent on the rest of the population. This means that individual cells renounced their capacity to reproduce as independent units and came to reproduce as part of a larger whole (Rueffler et al. 2012).

Several types of microbes have evolved some division of labour among its colony members and could be considered multicellular organisms (Shapiro 1998; Dunny et al. 2008). Therefore, division of labour is not exclusive of multicellularity but can emerge as well in groups of cells such as bacterial populations. Bacterial manifestations of division of labour take different forms which range from undifferentiated chains to morphologically differentiated structures (Lyons & Kolter 2015). Some well-known examples include myxobacteria and cyanobacteria. Filamentous cyanobacteria are able to carry out two incompatible tasks, nitrogen fixation and photosynthesis, in two spatially separated and

specialized cells: the carbon-fixating vegetative cells and nitrogen-fixating heterocysts. In contrast, myxobacteria present an aggregative and motility-dependent multicellularity.

Myxococcus xanthus, the best studied specie of myxobacteria, hunt socially via the coordinated secretion of lytic enzymes that digest bacterial and fungal prey. During starvation, they undergo a dramatic transition where individual cells migrate together to create fruiting bodies for the spreading of the spores. Around 50% of the colony cells die during this transition in a process called programmed cell death (PCD), in order to supply nutrients for the fruiting body cells. Another example of PCD is the virulent strain *Salmonella typhimurium* when invading the gut tissue. Its altruistic members die while preparing the ground for the infection of the remaining part of the population (Claessen et al. 2014).

Recently, some experimental models have been developed to study the first steps in the evolution of multicellularity. In a recent paper involving yeast cell populations, Ratcliff *et al.* found that snowflakes structures developed division of labour with some cells undertaking PCD (Ratcliff et al. 2012). They used gravity to select for primitive multicellularity, by selecting the faster sedimenting cells that lead to the appearance of stable aggregates. After repeated culture transfers, the connected cluster of cells, so called snowflake phenotypes, appeared in a predictable way. They demonstrated that larger groups presented larger rates of apoptosis, proving that PCD is a viable evolutionary strategy that increases the fecundity of the snowflake phenotype. In another experiment, Kim *et al.* created mutants of the specie *Pseudomonas fluorescens* that showed collaboration and division of labour between the mutant and the parent strain. They worked together to gain new territory by diving the labour through the expression of a wetting polymer that facilitates the advance of the colony, which was later executed by the parent strain (Kim et al. 2008b).

On top of the experiments that recreate multicellularity and division of labour, there has been many theoretical models that have been developed to answer some of the fundamental questions of these processes. The complexity of such processes accentuates the importance of mathematical modelling in providing more precise conclusions based on generalization from

data and empirical work. Much modelling effort has focused on the fitness advantages of undifferentiated cell clusters, the benefits of cell specialization and the conditions for the evolutionary stability of division of labour (Yamagishi et al. 2016).

A special case that has drawn much interest is the study of the canonical division of labour in bacteria. Canonical division of labour refers to the emergence of two distinct phenotypes within the same strain, namely, a reproductive type often called *germ* and a vegetative type called *soma*. In this context, somatic cells perform specific, non-reproductive tasks for the whole group. Several models have addressed questions that arise from this scenario: how beneficial is to sacrifice a part of the group in the reproductive effort?, and if so, under which circumstances?

In a model provided by Willensdorfer, he defined a population of soma and germ cells and calculated the benefits that soma cells conveyed, providing that germ cells contribute less to the group benefit than somatic cells (Willensdorfer 2009). The optimality of the result was calculated by defining colony fitness as the product of the cost of soma, the cost of colony size, and the benefit of group living. Based on this function, he predicts when somatic cells will evolve and what the proportion of somatic cells will be. He concludes that division of labour is favoured by many different benefit functions, but that it is only maintained if the unicellular organism is not able to receive large benefits without the division of labour. He further shows that the fraction of somatic cells will generally be much smaller than the reproductive germ cells even if somatic cells contribute much more to the colony fitness than reproductive cells.

A different method was used by Gavrillets to model the same phenomenon (Gavrillets 2010). The model specifically studies the trade-off between the investment on viability (soma cells) and the investment on fertility (germ cells). He models the evolution of a colony with these two phenotypes for different trade-offs and proves that total division of labour increases the fitness of the colony. Similar conclusions were derived in the model of Goldsby *et al.*, who were able to simulate the emergence of reproductive division of labour between soma and

germ cells (Goldsby et al. 2014). They used what they referred to as “digital evolution”, an approach where cells are modelled as digital entities that run programs which are subjected to evolution. Regardless of the specific modelling techniques (mathematical equations describing the system fitness or a more complex computational framework), these models evidence that the necessary condition for sustained division of labour is the existence of strong trade-offs between the somatic and reproductive functions.

Another interesting scenario is the modelling of the division of labour of two incompatible tasks, as with the example of cyanobacteria where individual cells need to perform two incompatible metabolic processes such as oxygenic photosynthesis and anoxic nitrogen fixation. Ispolatov *et al.* developed a quantitative model that shows that a multicellular form emerges from genetically identical cells when poorly compatible physiological processes are divided through the specialization of spatially separated cells (Ispolatov et al. 2012). For that, they assumed a fitness cost to producing both metabolic products simultaneously. Their model also enabled the formation of transient aggregates of up to two cells that later dissociate into single cells. Cells modify their investment into each task based on the paired group state and the global fitness with the objective of increasing its individual cell fitness. Finally, they tracked the evolution of the time a cell spends in an aggregate state as a representative function of the evolution of multicellularity. Their result claims to be proof that division of labour is possible in a homogeneous population, without the pre-requisite of phenotypical heterogeneity. The spontaneous emergence of division of labour occurs if it provides a fitness advantage to the individual cell in the aggregated state, and if the regulatory mechanism that allows cells to specialize is already present in the individual cell.

Objectives of this chapter:

Specific conditions for the emergence and further maintenance of division of labour include:

- (i) For its emergence: individuals should exhibit phenotypic variation and exhibit a cooperative behaviour (Kolter et al. 2015b).

- (ii) For its maintenance: division of labour should provide a benefit for all individuals (West & Cooper 2016).

So, in this chapter, simulations will attempt to address how these conditions must be enforced.

In the first section, *Why divide labour?*, the second condition will be put to the test. Simulations of populations with different degrees of specialization will compare different reward functions. The testing hypothesis is that division of labour is more beneficial if the fitness return for investment in a given task is accelerating (West & Cooper 2016).

In the following section, *How to divide labour?*, the first condition will be examined. In a context where division of labour is programmed to emerge in the form of soma and germ cells, different sources of phenotypic variability will be simulated. The objective is to find which source of phenotypic variation provides a more efficient division of labour (*i.e.* higher global fitness) and determine the optimal fraction of sterile cells in each case. Noisy permanent and transient sources for variation are simulated. Finally, a novel mechanism based on the emission and sensing of diffusible molecules is proposed as phenotypic variability generator. The mechanism is able to produce differentiation, providing the optimal fraction of soma and germ cells.

To finish, in *Division of labour from coordinated signalling* the novel regulatory mechanism that is proposed in the previous section is further characterized. The genetic network is generalized to generate spontaneous differentiation in an initially equal-cell population without the requirement of any fitness reward. Spatial patterns are obtained as a consequence of specialization, hinting that the proposed network could be useful as well in a morphogenetic context.

5.1 Why divide labour?

In this section, I want to test how large the benefit of division of labour should be for it to be a proper evolutionary stable strategy. I am interested in which reward function produces an increase in the global fitness of the colony (i.e. number of cells after certain simulation time). The main hypothesis under which the following simulations are motivated is that the fitness benefit of the division of labour has to be beneficial at the group level, and that it should be an accelerating benefit (West & Cooper 2016). An accelerating fitness return is recreated when the task becomes more efficient as more effort is put into it, or equivalently, when there is a penalty for the execution of incompatible tasks in the same location (Ispolatov et al. 2012).

Figure 5.2– Division of labour in cells performing two tasks.

I have devised a simple experiment to test when, given different benefit/cost ratios, is more beneficial to divide labour. Let us imagine that there is a population of cells that performs two independent tasks (Figure 5.2). Both tasks involve the production and emission of a certain diffusible molecule, namely, molecule A (associated to task A) and molecule B (associated to task B). All cells are required both molecules to properly grow and divide. Each cell has a constant and finite set of resources to spend in the production of molecules. The imposition is that if any cell invests a proportion of resources (X) into task A, then, the remaining proportion ($1-X$) would be invested into task B.

Now, the question is if the fitness of the population would increase if part of the population specializes in carrying out a single task and the remaining group of cells takes up the other required function. Because both tasks involve the emission of the product into the shared environment, it is possible that the whole group benefits from the labour without carrying the

cost of production. In other words, the molecules are equally shared, but the cost of metabolite production is individual.

By comparing the number of individuals of a non-specialized community with the number of individuals of a specialized one, it is possible to determine which strategy provides a higher fitness (Figure 5.3). Given two strategies (*e.g.* divide or not divide labour) one can compare the final fitness outcome (*i.e.* number of individuals in the population) and derive robust conclusions from the process. So, the following simulations explore the conditions under which division of labour leads to optimized system output, as measured by the reproductive success of the whole colony.

Figure 5.3- Two separate setups to test two different strategies. Simulations can be either in monoculture, for isogenic populations or in co-culture of two types specialized cells.

The assumption is that the efficiency of this strategy (*i.e.* divide labour) is going to depend on the imposed reward/cost ratio of each task and the degree of cooperation. Initially in the following simulations, the degree of cooperation will remain fixed, and it is given by the physical properties of the produced molecules:

Diffusion rate	$0.2 \frac{\text{mol} \cdot \text{cell_size}}{dt}$
Degradation rate	$0.3 \frac{\text{mol}}{dt}$

Table 5-1. Physical parameters of shared molecules that define the degree of cooperation.

Cells would be emitting molecule A and molecule B with the degradation and diffusion properties specified in Table 5-1. The degree of cooperation is given by those arbitrary values because they define a specific length of communication. This is, provided that a certain cell is emitting, for instance, one molecule A per simulation step into the shared environment, then, how many individuals can receive the signal?

Figure 5.4— The degree of cooperation establishes the length of communication. With the parameters from Table 5-1, a cell can communicate at least with its closest neighbours.

In Figure 5.4, it is shown how the blue cell is emitting 1 molecule per time step and the reception threshold is set to 1 molecule per time step. Cells that sense above the reception threshold are depicted in black. In this worst-case scenario, cells can communicate with their closest neighbours. The length of communication is further modulated by the emission rate and the reception threshold of each cell, which also modulates the cell individual fitness. This is, the division time of the cell will depend on the amount of signal that it reads from the environment. The more signal a cell can receive, the shorter the division time, and thus, the higher its reproductive fitness (Figure 5.5)

Figure 5.5– Molecule reception dictates generation time.

Simulations compare the overall fitness of the population after 4 hours of simulated experimental time. Each simulation consists on a colony composed by two subpopulations randomly distributed in space. Each subpopulation has the same initial number of cells (1000 initial cells per population, 2000 in total). The key matter is that the level of specialization of each individual within the subpopulation varies, ranging from subpopulations of non-specialized cells to subpopulations that only undertake one of the tasks. Simulation were run for the six different specialization scenarios (Figure 5.6).

The first simulation considers two equal subpopulations that do not divide labour, where their individuals simultaneously produce molecules A and B. Each individual has an emission rate of 5 molecules of each type per time step. Therefore, at each time step, 2000 cells are emitting 5 molecules of A and 5 molecules of B. The next experiment considers a case where the individuals are slightly specialized. In this case, population 1 is consisted of 1000 cells that have an emission rate of 6 molecules of A per time step while the emission rate of molecule B is set to 4. On the other hand, the remaining subpopulation, population 2, consists of 1000 cells that emits 4 molecules of A and 6 molecules of B per time step. Notice that population 2 is always complementary to the distribution of tasks of population 1, in all simulated scenarios. Overall, in all six simulations, the global emission of each molecule is set to be 10 molecules per time step. Succeeding simulations consider individuals that are increasingly more specialized. The final experiment simulates the extreme case where subpopulation 1 consists only on cells that emit molecule A, and the cells of subpopulation 2 emit molecule

B. Again, the global emission rate is kept constant and set to 10 molecules per time step. So, if the reception thresholds define the rewards through the reproductive rate, the emission rate defines the cost of producing molecules A and B. Since the resources for production are fixed at 10 molecules per time step, the cost is constant and equal for all cells. In other words, the number of emitted molecules only depends on the amount of resources dedicated to each task, being this relation linear (*i.e.* constant cost function).

Statistical significance was achieved for N= 10 runs per scenario with negligible variance. The large initial number of cells removes the global stochasticity that could arise, proving that the global fitness as it is defined (number of cells after 4 hours of experiment time) is a reliable and robust readout.

Figure 5.6– Emission rates of molecules A and B of the individuals with different degrees of specialization for six different combinations of population 1 and 2.

In Figure 5.7 the fraction of cells with their corresponding generation times is depicted. In the first experiment (non-specialized cells) most cells (~60%) are able to divide roughly every 60'. That specific distribution provides a global fitness that will be used as a reference value. The global fitness of the rest of the experiments is normalized with that reference control value (Figure 5.8). Every distribution of individual fitness determines the global fitness. The distribution of the percentage of cells and their division time changes dramatically from the non-specialized experiment (top left graph in Figure 5.7) to the completely specialized (bottom right graph in Figure 5.7).

Figure 5.7– Percentage of cells and their correspondent generation times for the six experiments of Figure 5.6.

The lack of benefit for specialization is a consequence of the molecule diffusion. Cells that produce both molecules are not required to *wait* for the already-diffused-and-degraded molecule that is coming from its neighbours. Since the intangible cost of diffusion, associated to process of specialization, is not compensated in any way, there is no advantage for cooperation and division of labour.

Figure 5.8 shows the global fitness of the six simulations with various degrees of specialization. As commented, the final number of cells in the non-specialized experiment is used as a reference value for normalization. The average global fitness declines as cell specialize. Notice that total specialization returns a global fitness which is almost half of the non-specialized case. This proves that a constant cost function does not promote division of labour over non-division.

Figure 5.8— Global fitness of the six simulations with different degrees of specialization

As mentioned in the beginning of the section, the working hypothesis is that the rewards received from the performance of a given task have to be accelerating for division of labour

to be more beneficial than non-specialization. To test this hypothesis, the emission rate of each molecule will be now a function of its specialization. Instead of modifying the rewards provided by the reception threshold (which will remain the same), in the following experiments, the emission rate assigned to each scenario will grow exponentially (Figure 5.9). The total cost of each individual remains equal and constant, but, the cost per molecule is now reduced in proportion to its specialization.

Figure 5.9– Accelerated emission rates of molecules A and B of the individuals with different degrees of specialization for six different combinations of population 1 and 2.

This emission function represents the idea that the task becomes more efficient as more effort is put into it. The mastery in the production of a molecule is rewarded by increasing the rate of the emitted signal for the same invested resource.

In order for the non-specialized case to be the same as in the previous example, the new emission rate is related to the previous one as:

$$New\ emission_{rate} = \frac{(old\ emission_{rate})^2}{5}$$

That results in the emission of 20 molecules per simulation step for the fully specialized cell.

Now, the distributions of the percentage of cells in each time slot are shifted in a different manner (Figure 5.10). Cells that divide labour completely (20/0) and almost completely (16,2/0,2) provide a higher fitness than the non-specialized case.

Figure 5.10— Percentage of cells with their correspondent generation times for the six different simulations.

These results are a consequence of the trade-off that exists between the molecule loss due to diffusion and the higher efficiency that division of labour provides. Figure 5.11 shows how the global fitness initially decreases for cases with partial specialization where the reward in signal emission does not compensate still the loss of emission. As specialization is more extreme, the global fitness increases. Not complete specialization (16.2:0.2) already returns a higher fitness than the non-specialized case (global fitness of 1,1). Finally, fully specialization achieves an average fitness 1.3 higher than the non-specialized case proving that accelerating benefits, in the form of more efficient molecule production, are required for division of labour to be a successful strategy.

Figure 5.11– Global fitness of the six simulations with different degrees of specialization for an accelerating emission rate.

These simulations are now extended to different emission functions (reward functions) other than the linear reward (Figure 5.6) and the accelerated quadratic reward (Figure 5.9). A cubic

emission function ($\sim \text{linear_emission}^3$) and the square root function ($\sim \text{linear_emission}^{1/2}$) are simulated for different degrees of specialization (Figure 5.12). The cubic emission function corresponds to a scenario where specialization is highly rewarded, while the square root function mimics a scenario where specialization is punished. Constant ($\sim \text{linear_emission}$) and quadratic ($\sim \text{linear_emission}^2$) emission function were already discussed. Results corresponding to the same emission function were connected with dotted lines for an easier visualization. The small variation in the results lead to the substitution of error bars for big dots for the same reason.

Notice that for slight specialization the constant and the square root function returns higher fitness than the quadratic function. It is interesting that the square root function, which implies decelerated benefits or, in other words, less emission in the specialized case than in the non-specialized, returns the same fitness tendency than the linear function.

Figure 5.12– Global fitness with different degrees of specialization for different emission functions.

Results from Figure 5.12 conclude again that only accelerated benefits (quadratic and cubic) are able to increase the global fitness of a population that has divided labour. If the biological function A does not become more efficient as it is decoupled from some other function, B, the emission functions could be represented as the linear and square root demonstrating in those cases, that division of labour is not a beneficial strategy.

It is important to notice that the benefits of division of labour can emerge only if individuals remain spatially close to each other. Proximity is a requisite for cell-to-cell interactions because it is necessary for the exchange of molecules through diffusion. The effect that the degree of cooperation has in promoting division of labour is considered next. The degree of cooperation, which is represented as the length of communication, is changed in the following simulations to include a shorter communication length and a longer communication length.

Figure 5.13 depicts the global fitness of populations with the six different degrees of specialization in two distinctive contexts, for the four emission functions.

First, results in the short communication range are shown in the top of the figure. That communication length (diffusion=0.1 mol/dt; degradation=0.5 mol/dt) simulates a signal that can only be sensed by the cell that is emitting it and occasionally by some close neighbours. Because the half-life of that molecule is so small, the accelerated benefits of the quadratic function cannot compensate the loss of signal, and division of labour results as a poorer strategy than non-specialization. The general fitness of specialized populations is lowered as they rely on a more long-live communication system. It is only the cubic function that is able to rescue specialized populations by producing much more signal.

On the other hand, when the range of communication is widened (diffusion=0.5 mol/dt; degradation=0.1 mol/dt) so that second neighbours from the source are able to sense the emitted molecule, the system becomes so saturated that accelerated emission functions do no longer provide a competitive advantage. The half-life of the molecules is so high that there is no loss in information when sharing, and the specialized and non-specialized populations become almost undifferentiable for all emission functions.

Figure 5.13– Global fitness with different degrees of specialization for different emission functions and two communication lengths.

As a summary, simulations have revealed that division of labour is an evolutionary stable strategy only if: (i) the transport properties of the shared compounds remain in the local proximity of the producer and, (ii) the benefits for task specialization are accelerating (i.e. increase faster than a linear function).

5.2 How to divide labour?

In the previous section, I have proved that an accelerating benefit is fundamental for division of labour to outcompete a non-divided strategy. For its emergence, division of labour is said to require individuals to exhibit phenotypic variation within the group (Kolter et al. 2015b). In this section, I want to test how division of labour can emerge by focusing on this condition. Individuals have to exhibit phenotypic variation, this is, it is required that not all individuals follow the same strategy. Here, I will simulate different mechanisms that create differences among the individuals of the same colony. Simulations consider the canonical case of division of labour, where there are only two possible phenotypes or strategies. A certain fraction of cells becomes soma (non-reproductive, helper phenotype) and the remaining cells continue as germ (reproductive phenotype). So, what rules or mechanisms should decide which cells become soma to maximize the global fitness? And, what proportion of individuals should carry out the different tasks?

Even in homogeneous environments and isogenic populations, bacteria can exhibit cell-to-cell phenotypic variability. This is referred to as individuality (Davidson & Surette 2008). Phenotypic heterogeneity allows colonies to adapt and survive to sudden changes in the environment and is central in the emergence of division of labour.

The fundamental idea is that phenotypic heterogeneity can provide such benefits because, heterogeneous colonies have more diverse strategies to face the environment and thus are better adapted for evolution (Ackermann 2015). The most important internal molecular mechanisms that give rise to phenotypic variation, can be summarized as:

- Genetic differences or mutations.
- Noise in gene expression.
- Cell-cell interactions mediated by diffusible molecules or physical contact between individual cells.

Let us imagine a population of germ cells and somatic cells, where soma cells give up the chance to reproduce to help germ cells reproduce. So, we have two types of cells: germ cells, G, and somatic cells, S.

In a population consisted only of germ cells, the fitness of the population would be proportional to the number of cells it has reached during the experiment time. This fitness is given by the generation time of the germ cells. I assume that germ cells can reproduce on their own, without the help of the somatic cells. Somatic cells are sterile cells that help increase the growth of the germ cells by secreting a molecule that germ cells absorb. The proportion of somatic cells is increased in following different rules or mechanisms. The questions that this section will attempt to answer are: what is the best mechanism for phenotypic variability to divide labour in term of the population fitness? And, within each mechanism: what is the optimal proportion of sterile/reproductive cells that maximizes the global fitness of the colony? In this section, the methodology is as follows: four different mechanism, with their corresponding rules and dynamics, will be simulated in independent experiments. Finally, the global fitness of each strategy will be compared.

Figure 5.14– Four different mechanism to simulate phenotypic variability.

In Figure 5.14, the four mechanisms and their implications are shown. Soma cells (blue) sacrifice their reproductive resources and invest them into the production of a substance that promotes growth in germ cells (green). Germ cells that are close enough to the soma cells are able to exploit the growing substance and become super-growing cells (dark green cells). Mechanisms will be further described with their correspondent results.

Starting with an isogenic population of germ cells, simulations will differ only in the nature of the mechanism that incorporates the soma phenotype. The physical properties of the growth substance emitted by soma and the sensitivity of the germ cells to that substance (sensitivity threshold) are the same in all simulations and correspond to the values in Table 5-2.

Diffusion rate of growth molecule	$0.05 \frac{\text{mol} \cdot \text{cell}_{\text{size}}}{dt}$
Degradation rate of growth molecule	$0.5 \frac{\text{mol}}{dt}$
Emission rate of soma cells	$20 \frac{\text{mol}}{dt}$
Sensitivity threshold for super growth	$3 \frac{\text{mol}}{dt}$
Generation time of germ cells	40 min.
Generation time of super-growing germ cells	22 min.

Table 5-2. Parameters of the simulation for different phenotypic variation mechanism.

The values of diffusion, degradation, emission and sensitivity of the growth substance return a specific benefit value that the soma cell provide. Given the constraints of the spatial structure, those values represent a short-life diffusive molecule that can only be profitable for germ cells in right physical proximity with the secreting soma cells. I have assumed these values as they represent the more restrictive conditions of cooperation. Since the simulations consider the explicit position of cells in space, the overall differences in the dynamics of soma, germ and super-growing germ cells can only be accounted for by the simulated sources of variability. Note that, this comparative methodology can only be performed with an

individual-based model since the combination of the individual interactions and the spatial restrictions are responsible for the global system behaviour.

Initially specialized

The first group of simulations assumes an initially specialized scenario that works as a control experiment against which the rest of the simulations will be compared. It establishes the baseline benefit of having a set of soma cells in a germ population. The assumption is that cells are already specialized in the two phenotypes that follow the rules described above. The effect of different proportions of initial soma cells in the population was explored.

Figure 5.15– Percentage of soma cells over time. Since soma cells do not reproduce, their proportion in a well-mixed population of germ cells becomes diluted over time.

Figure 5.15 shows how the initial proportion of soma cells gets diluted due over time for nine different simulations. After three hours of simulated time, almost all simulations return an equal final fraction of soma cells, regardless of the initial proportion. The dynamics of the dilution process reflect the trade-off that arises from having higher proportions of soma (which translates as a higher proportion of super-growing germ cells) but at the same time not contributing to the population reproduction capacity. It is a matter of finding a balance

between having many sterile workers and a queen bee, or many queens that do not provide that many offspring.

The final population fitness after 180 minutes of simulation time, is shown in Figure 5.16. The 0% bar corresponds to a population that was conformed solely of germ cells and provides the reference value. Simulations were performed for $N=20$ times with error bars showing little deviations from the average value. The trade-off is again reflected in the graph that shows that the optimal balance is achieved when the proportion of initial soma cells is close to half the total initial population (45%). Notice that any fraction of soma cells improves the global fitness of the control case, except when only the 15% of initial bacteria are germ cells.

Figure 5.16– Population fitness for different initial fractions of soma cells after 3 hours of simulated time.

The baseline represents a population consisted only of germ cells.

Random and irreversible specialization

The next group of simulations recreates division of labour emergent from a random and irreversible mechanism that would turn germ cells into soma. These rules can be interpreted as random mutations that transform the reproductive organism into the sterile phenotype. Mutations occur randomly over time with every individual having the same mutation probability. Once the mutation has occurred, the individual will remain in the somatic phenotype. That trait could be inheritable as well, however, since soma cells do not reproduce it is not vertically transmitted.

gro allows the specification of individual behaviour with certain probabilities. This means that is possible to define the initial state of a cell and assign it a probability to transit to another defined state. That probability, although defined individually, refers to the final proportion of individuals that would experience such transition in a given time. So, if all initial cells are assigned a probability of, for instance, 0.5 in 3 hours, this mutation rate would return a final population of 50% cells in the initial state, and 50% cells in the mutated state after 3 hours. However, the internal model that calculates these fractions assumes that *mutated* cells divide as well. So, the initial mutation rates that were used in these simulations do not return a final population that coincides with that assigned value, because soma cells do not reproduce.

Figure 5.17– Percentage of soma cells over time for a random and irreversible specialization. Each curve represents the data for a different mutation rate.

In Figure 5.17 the percentage of soma cells over time for the random and irreversible specialization is shown. Each curve corresponds to a mutation rate. The mutation rates that were utilized range from 0.05 to 0.95 of cells transitioning in 3 hours. Because of the dilution problem, the curve that should correspond to an 85% of emergent soma cells, stabilizes at 40%. Also, due to dilution, the fraction of soma cells converges to a specific value even though the mutation rate provides a constant flow of soma cells. The equilibrium is achieved when the mutation rate equals the division rate of the germ cells. Darker highlighted curves are the ones that return the highest fitness (Figure 5.18).

Figure 5.18– Population fitness for different initial mutation rates after 3 hours of simulated time. The baseline represents a population consisted only of germ cells.

Simulations with 0.55 and 0.65 initial mutation rates provide the best population fitness. This corresponds to fraction of 20% of soma cells, as highlighted in Figure 5.17. Results continue

to exhibit the commented trade-off, but with this mechanism the achieved global fitness is much higher than the previous case (from 1.5 to almost 2 times higher). The more plausible reason is that the mutation mechanism provides a constant fraction of somatic cells, which could be arguably better than the diminishing function shown in Figure 5.15.

Random and reversible specialization

The next group of simulations recreates division of labour emergent from a random and reversible mechanism. These rules can be interpreted as changes in the phenotype arising from internal genetic noise. In these simulations, initial germ cells can transition to a soma phenotype with a random probability. However, that transition is not permanent and will be undone after some time. The time that soma cells remain in that state is 40 ± 10 minutes, which is also the average generation time of germ cells. The dynamics of the soma cells resemble that of the mutation curves for early time periods (Figure 5.19). Then, the fraction of soma cells decays because initially transformed soma cells start to return to the germ state. Higher noise rates (>0.55) decay much faster because the dilution effect is also greater in those cases. The dark curve corresponds to the best global fitness, shown in Figure 5.20.

Figure 5.19– Percentage of soma cells over time for a random and reversible specialization. Each curve represents the data for a different noise rates.

The fitness histogram is shifted and now peaks at a rate of 0.75 (Figure 5.20). Even though the higher fitness value is only slightly better than the best case of the mutation mechanism, here, various noise rates return fitness higher than 1.8. With this mechanism, the perpetual trade-off between super-growing cells and non-growing cells is resolved over time. The fraction of soma cells does not forever decay (like the initially specialized case) or remain constant (like the mutation case) but has two temporarily distinguishable strategies. First, for high noise rates, a saturation of transient soma cells provides a growth boost in the remaining germ cells. Later, the percentage of soma decays slowly enough to release enough cells into the reproductive state again. This combined strategy, has proven to be successful in the improvement of the colony fitness.

Figure 5.20– Population fitness for different initial noise rates after 3 hours of simulated time. The baseline represents a population consisted only of germ cells.

Consensual and reversible specialization

Finally, the last group of simulations recreates a transient division of labour emergent from a consensual, cooperative mechanism. The idea is that cells self-organize through the exchange of a certain molecule that dictates when cell should transition to the soma state, and vice versa.

If cells interact and coordinate phenotypes at a local level, this can ensure a precise and appropriate ratio of the different phenotypes, even in small groups. All cells regulate according to the circuit in Figure 5.22. The substance secreted by soma cells acts at the same time as a regulator to switch between the germ and soma state. Each state, soma and germ, has an associated sensitivity threshold (Figure 5.21). When a soma cell senses that the concentration of signal is higher than its threshold, then the cell becomes a germ cell. On the other hand, if a germ cell senses that the signal concentration is below its threshold, the cell would turn into a soma cell.

Figure 5.21– State transitions according to the molecule concentration.

The population begins at the germ state and the lack of signal produces that most cells turn into the soma state. After some time, the accumulation of signal will transform the cells back to the germ state until a balance between the proportion of soma and germ cells is reached.

Figure 5.22 (left) shows the unicellular inner states of germ G and soma S and how they relate. Following the described rules, state G and S are mutually exclusive, hence the double repression depicted in the figure. This represents the premise that cells can only be in one of the two states. The absence of signal reinforces the repression of S over G, and thus, cells turn into the S state.

On the other hand, when signal is in excess, the repression branch from S to G weakens, and thus, cells would turn into the G state.

At the multicellular level (Figure 5.22, right), cells that are in state S will influence their neighbours (and themselves) by promoting the G state through the emission of the regulatory molecule. On the other hand, cells in the G state will passively promote the S state in their vicinity by not emitting the regulatory signal.

Figure 5.22– Individual and multicellular motifs of the consensual and reversible specialization.

The parameters of cooperation, namely the diffusion and degradation of the molecule produced by soma cells, are fixed. All mechanisms have shared the production of the same type of molecule (Table 5-2). However, here, two new modulable parameters appear. Cells require two extra parameters that inform them when (how much signal concentration) to transition from the germ to the soma state, and vice versa. The amount of signal required for the transition is the sensitivity threshold. The sensitivity threshold for the transition into soma and the threshold for the transition back to germ do not necessarily have to be equal (Figure 5.21). Simulations were run for different combinations of the both. Two relevant examples of how the proportion of soma cells changes over time are shown in Figure 5.23.

The values 20/10 refer to the sensitivity threshold for soma (mol/dt) and the sensitivity threshold for germ (mol/dt), respectively. This means that soma cells, in this example, would require more than 20 molecules every simulation time step to initiate the transition to the germ state (*i.e.* soma sensitivity threshold). Germ cells would require more than 10 molecules per

time step to remain in the germ state. Sensing lower concentration would result in germ cells becoming soma (*i.e.* germ sensitivity threshold).

The percentage of soma cells initially oscillates until it reaches an equilibrium. Most combinations of thresholds stabilize at a percentage that lies between 3%-7% which is by far the smallest percentage of soma cells that any mechanism has provided. Despite the small fraction of germ cells that this mechanism sacrifices, the overall fitness results are much greater than any of the other simulated mechanisms in this section. The general fitness achieved with consensual specialization (Figure 5.24) can be up to 5 times higher than the second-best mechanism (random and reversible specialization, Figure 5.20).

Figure 5.23— Percentage of soma cells over time for a consensual and reversible specialization for two different combination of soma threshold/germ threshold.

A proximate explanation for the outperformance of this mechanism, in compare to the rest of sources of phenotypic variability, is that the number of generated soma cells is much lower than the number of super-growing cells. This mechanism is able to obtain the optimal fraction of the trade-off between sterile production and global growth by adapting the number and spatial location of soma cells as a function of the current needs.

Figure 5.24– Population fitness for different combinations of soma and germ thresholds.

The best results within this mechanism, according to the fitness they provide, are obtained when at least one of the two following conditions is met: (i) sensitivity thresholds are high (compared to the average signal concentration in the environment) and/or (ii) the soma threshold is equal or larger than the germ threshold.

The first condition implies that both soma and germ cells are more resistant to change its state (they require more signal to build up). This minimizes avoidable fluctuations between both states while still behaving in a responsive manner. Since the sensitivity threshold to transition to the super-growing state is much lower than both the soma and the germ threshold (remember it is fixed at 3 molecules per time step), the population of germ cells (which will be resistant to transit to the soma state) will be almost completely consisted of super growing cells.

The second condition points out that if both thresholds cannot be high, at least, a higher soma threshold ensures that soma cells only turn into germ cells if there is a high enough signal concentration in the environment, thus, benefiting the rest of the germ cells that would most likely be super growing cells.

As a final remark, it is important to consider a major difference in this last mechanism compared to the two previous random specialization models. The constant fraction of soma cells achieved with this mechanism does not imply that the same individuals have to carry the burden of being soma from the beginning. The burden of transiently being a non-reproductive, helper phenotype is virtually shared by all cells. The mechanism assures that the fraction of soma cells remains constant but does not lock the state of each individual cell. This reversibility requirement, which is also found in the random and reversible specialization, is not the fundamental difference. Because of cell self-organization, a given cell can become soma for some time and then come back to being germ in a consensus way that is dictated by the needs of the local cells. Besides, soma cells are not randomly generated but emerge in specific spatial locations. The emergence of soma cells is spatially more efficient, leading to a lower number of soma cells that are strategically located.

Shots of the cell arrangements of the four simulated mechanisms are shown in Figure 5.25. The consensual and reversible specialization (bottom, right) provides the minimal fraction of soma cells (blue cells) that, at the same time, create the maximum fraction of super-growing cells (dark green cells). Soma cells are homogeneously distributed in the population providing the highest efficiency. Random and reversible specialization (bottom, left), on the other hand, is unable to transform most of the regular germ cells (light green) into super-growing cells. Since the emergence of soma cells is random, much effort is lost in having several soma cells being spatially close. The mechanism of random permanent specialization (top, right) offers the same fitness benefits than the reversible specialization, but with a different strategy. The mechanism of random and irreversible specialization generates such a high proportion of soma cells that even if not spread homogeneously, are enough to transform germ cells into super-growing cells.

Figure 5.25– Colony shots of the four simulated model for phenotypic variability.

In these simulations, it was evidenced that a mechanism that relies on coordination (via the secreting and sensing of a certain signal) is more efficient as a difference generator than random differentiation (whether it is permanent or transient). This was proved to be true for the case of canonical division of labour, where cells differentiate into non-reproductive or reproductive phenotypes. The trade-off between reproduction and viability is better resolved with a consensual mechanism because it maximizes the distribution of signal by coordinating the emergence of soma cells through time and space.

5.3 Division of labour from coordinated signalling

In the previous section the information derived hinted that phenotypic variability may not be such a fundamental requirement for division of labour. While it has been widely proved that phenotypic variability can lead to division of labour (DOL), it was shown how DOL can also emerge as a consequence of cell communication. This happens through the spontaneous symmetry breaking of an initially homogeneous and indistinguishable group of cells (Ispolatov et al. 2012).

In this final section, the genetic network that generated consensual and reversible specialization is further characterized. In the following, two genetic circuits will be explored as difference generators in an initially homogeneous population. Both circuits are based on intercellular signal regulation, the first uses only one signal (it is the same circuit as the consensual and reversible specialization mechanism) while the second regulates the specialization of the two tasks using two different signals. The exploration of the circuit properties will be conducted mainly without any cost/fitness functions in stationary non-growing populations. Finally, the fitness benefits of these circuits are studied in a more general case as well as their attributes as morphogenetic networks.

5.3.1 One-signal circuit

Simulations will explore the properties of the circuit of Figure 5.22 in a static, non-growing population of cells. The circuit imposes a series of rules that every individual in the colony must follow. In this general case, cells can be either in the A state or in the B state, but not both. By exploring the different physical parameters of the circuit, the potential for it to be a difference generator will be analysed. DOL can be transient if cells switch between states perpetually, or it can be permanent, if cells end up in only one of the two states resulting in a static solution.

The mechanism is the same as the one explained in the previous section 5.2 and is shown again in Figure 5.26. The substance secreted by cells in the A state acts as a regulator to switch between the A and B states. Each state has an associated sensitivity threshold equivalent to the one presented in Figure 5.21. If an A cell senses that the concentration of signal is higher than its threshold, then the cell transits to the B state. On the other hand, if a B cell senses that the signal concentration is below its threshold, the cell would turn into the A state.

Simulated cells begin in the B state and the lack of signal automatically produces all cells to turn into the A state. Then, the accumulation of signal will convert some cells back to the B state until a balance between the proportion of A and B cells is reached.

The multicellular motif (right, Figure 5.26) is also analogous to the previous scenario. Cells influence each other indirectly (hence the arrows). Cells in A will influence their neighbours (and themselves) by promoting the B state through the emission of the regulatory molecule. On the other hand, cells in the B state will passively promote the A state in their vicinity by not emitting the regulatory signal.

Figure 5.26– Representation of the one-signal coordinated signalling circuit and its multicellular implications.

The behaviour of the whole system will depend on the model parameters, namely, the diffusion and degradation rates of the coordination signal, the signal emission rate and the sensitivity threshold of both states. The parameters for the first exploration are directly inherited from the simulations of the previous section. Specifically, the set of parameters that provided the highest fitness return are listed in Table 5-3. Temporal sequential shots of the

colony behaviour are shown in the top part of Figure 5.27. Blue cells represent cells in the A state while green cells represent cells in the B state.

Diffusion rate of molecule	$0.05 \frac{\text{mol} \cdot \text{cell_size}}{dt}$
Degradation rate of molecule	$0.5 \frac{\text{mol}}{dt}$
Emission rate of A cells	$20 \frac{\text{mol}}{dt}$
Sensitivity threshold for B	$40 \frac{\text{mol}}{dt}$
Sensitivity threshold for A	$40 \frac{\text{mol}}{dt}$

Table 5-3. Initial set of parameters for the one-signal circuit simulations.

Figure 5.27– Temporal shots of the cells oscillating between the A state (blue) and the B state (blue). The graph shows the number of cells in each state over time.

Figure 5.27 also depicts the number of cells in each state over time. Cells oscillate between the two states, at first synchronously and then increasingly asynchronously. The initial synchrony is visually observed in the shots, as all cells are in one of the two states. As the signal diffuses over the colony, desynchronization starts to happen first at the edge, and then later spreads all over the colony. Eventually, the number of cells in each state converges into a certain proportion. Specialization is achieved in a transient manner with 60% of the population performing task A and the remaining 40% performing task B on average. Despite this convergence, the specific individuals that carry out each task changes with time. The internal transitions of two cells from this same experiment are shown in Figure 5.28.

Figure 5.28– Transitions between state A and state B for two random cells of the population of Figure 5.27.

In order to check if this process is size independent, this is, if the number of individuals affects the final collective effect, simulations for different fixed population sizes were run and are shown in Figure 5.29. Despite minor differences in the amplitude of the oscillations, all colonies converged roughly to a 60/40 division of states. Proving, therefore, that the mechanism is size independent. This makes sense since cell interactions are local and so, cells do not distinguish between large or small colonies.

Figure 5.29– Convergence to the final fraction of A and B cells is independent of the colony size.

A small scan over the values of the emission rates for two different threshold combinations was performed (Figure 5.30).

Figure 5.30– Percentage of cells in state A (blue) and B (green) for different emission rates and two combinations of reception thresholds.

The percentage of cells in state A are depicted in blue, next to the remaining fraction which are cells in the B state (green). Notice that the sum of cells in each state is 100, which indicated that cells are consistently in only those states (they cannot be in neither or in both states at the same time). The circuit is able to differentiate cells into the two given states at a fraction that

can be tuned by both the emission rate and the sensitivity thresholds. Practically all sorts of distributions can be achieved. Cells tend to spend more time in the A state as the emission rate decreases. If the reception threshold for A is much higher than for B (bottom row), cells will remain in state B much longer for high emission rates.

Further explorations in the parameter space were conducted. As mentioned earlier, the systems' behaviour depends on the diffusion and degradation rates of the coordination signal, the signal emission rate and the sensitivity threshold of both states. The effect of the variance of the thresholds and the emission rate was able to tune the average value of cells in each state. Now, different diffusion and degradation rates are tested. The new values represent a more cooperative signal, with a higher diffusion and lower degradation rates, respectively. New parameters are summarized in Table 5-4.

Diffusion rate of molecule	$0.2 \frac{\text{mol} \cdot \text{cell_size}}{dt}$
Degradation rate of molecule	$0.3 \frac{\text{mol}}{dt}$
Emission rate of A cells	$[20, 10, 5, 2] \frac{\text{mol}}{dt}$
Sensitivity threshold for B	$[40, 10] \frac{\text{mol}}{dt}$
Sensitivity threshold for A	$[40, 30] \frac{\text{mol}}{dt}$

Table 5-4. Set of parameters for the one-signal circuit simulations.

Results of this simulation are shown in Figure 5.31. A higher cooperative signal tilts the balance in favour of cells being in state B (the no emissive state). This is because a higher diffusion and lower degradation rates cause more signal to accumulate and reduce the requirement of cells to transition to the A state. Again, the interplay among the involved parameters provides a mean for tuning the population into the desirable fraction of A and B cells. A given task may be equally shared by cells in a community or it may be reserved for only just a tiny fraction. Whatever the case, the proposed genetic network is able to transiently differentiate cells into any given proportion.

Figure 5.31– Percentage of cells in state A (blue) and B (green) for different emission rates and two combinations of reception thresholds.

While almost all simulations present a transient DOL (cells transit between states maintaining the general proportion constant), simulations with an emission rate of 2 mol/dt and sensitivity thresholds of $A_th/B_th = 30/10$ offers a permanent DOL with cells rapidly fixing in one of the two states permanently.

Temporal sequential shots of the colony behaviour and number of cells in each state over time for this particular permanent DOL case, are shown in Figure 5.32. Blue cells represent cells in the A state while green cells represent cells in the B state.

Figure 5.32– Temporal shots of a colony undergoing permanent specialization. The graph shows the number of cells in each state over time.

After a few global oscillations the system stabilizes, and permanent DOL emerges, with cells stuck in the state dictated by the states of their neighbours. This is due to signal coordination among neighbouring cells. Because of the asynchrony among neighbouring cells, there is a chance that a cell meets its requirements for molecule A by producing it or by being near a producer cell, losing the need to change its state in the future. Cells do not engage in this behaviour at the same time, but rather, the coordination is performed in small groups that communicate with each other in a domino-like fashion.

The internal states of three random cells of this simulation are shown in Figure 5.33. All cells end up in a given state and remain there indefinitely. However, each cell has a specific differentiation time, as cell 1 stabilizes in state B after one oscillation while cell 2 requires two periods.

Figure 5.33– Transitions between state A and state B for three random cells of the population of Figure 5.32.

In order to continue the exploration of the parameter space of this motif, the diffusion and degradation rates are again modified to emulate an even more durable and diffusible molecule. Meanwhile, the emission rate of soma cells continuous with the same value. A large combinatorial study of different sensitivity thresholds is performed. The new parameters are summarized in Table 5-5.

Diffusion rate of molecule	$0.5 \frac{\text{mol} \cdot \text{cell_size}}{dt}$
Degradation rate of molecule	$0.1 \frac{\text{mol}}{dt}$
Emission rate of A cells	$2 \frac{\text{mol}}{dt}$
Sensitivity threshold for B	$[20, 10, 5, 2, 0] \frac{\text{mol}}{dt}$
Sensitivity threshold for A	$[60, 50, 40, 30] \frac{\text{mol}}{dt}$

Table 5-5. Parameter for the one-signal circuit.

Surprising spatial patterns appear for almost the entire range of parameters (Figure 5.34). In these simulations, the internal motif was able to achieve multicellular consensus and permanent DOL. As in the previous case, cells fixate on a certain state that accommodates the local need of the regulatory signal. Nevertheless, unlike simulations from Figure 5.32, which returned a scattered distribution of cells of both states, here, cells organize in space creating ordered patterns.

Cells fixate in the A state (blue cells) in the border of the colony because the signal concentration is lower in that region, activating the sensitivity threshold of the A state. Notice that the sensitivity threshold of A ranges from 60 to 30 molecules per time step, while the production rate is as low as 2 molecules per time step. Evidently, the shortage of signal in the colony edge triggers the expression of more signal (i.e. activating the A state). The fixation of cells in the edge in the A state induce the B state in their inner neighbours. This fixation of the B state will trigger, on a later stage, a second ring of cells in the A state. This cascades until the centre of the colony, thus, creating the ring-like pattern.

Figure 5.34– Spatial patterns emerge as a consequence of cell specialization. The threshold for A controls the width of the blue ring while threshold for B controls the width of the green ring, and thus, the number of blue rings per colony.

The depicted range of B threshold (0-20) corresponds to the values that exhibited permanent and ordered DOL. The only exception is the image for the 30/20 case that was not able to stabilize neither in time nor in space. Notice that the green region is where the signal concentration is higher than the B threshold, and the blue region is where the signal concentration is lower than the A threshold. In a way, the B threshold acts as signal sink. For example, when the B threshold is 0, all regions are above that signal concentration as long as the outer ring of cells consists of A cells that produce the molecule. As the B threshold increases, the regions where the signal concentration is not high enough are supplied by the

emergence of A cells. The higher the threshold, therefore, the higher the requirement of A cells.

Figure 5.35– The spatial pattern of cells in each state is influenced by the global shape of the colony. A cells are depicted in green and B cells are blue, to improve contrast.

It has been observed how the pattern is due to the colony shape where the molecule is diffused from the edge, and that pattern is transmitted radially from the edge to the centre of the colony. This means that if cells were initially arranged in different ways, the emergent pattern will still follow the contour shape of the colony, as it can be seen in Figure 5.35. The threshold combination is 50/2 for the examples of Figure 5.35.

A more exhaustive exploration of the parameter space follows, so that the requirements for the different collective behaviours that this circuit provides can be understood (Figure 5.36). However, the parameter space of this model is excessively broad to perform a thorough exploration of each of the parameters. Instead, related parameters were combined to reduce the search in the parameter space. First, the emission rate of cells in the A state has been conserved as it is. Then, the sensitivity thresholds of each state are considered in conjunction as the ratio A_{th}/B_{th} , where A_{th} refers to the A sensitivity threshold, and, B_{th} to the B threshold. For the sake of simplicity, B_{th} was kept constant at a value of 10 molecules per time step while A_{th} ranged from 5 to 60 molecules per time step. Finally, diffusion and degradation rates are also combined as shown in Figure 5.36.

Figure 5.36– Systems behaviour for different combinations of the emission rate, sensitivity thresholds, diffusion and degradation rates. Cells exhibit five different behaviours depending on the model parameter: permanent state A, permanent state B, oscillations between state A and B, transient DOL and permanent DOL.

The circuit results show five different regimes where DOL emerges for a specific set of parameters. In the z-axis, each of the depicted surfaces represents a different relation of the sensitivity thresholds. Remember that A_{th} dictates the maximum amount of A signal that can be sensed by cells before turning into the B state. On the other hand, B_{th} dictates the minimum concentration of signal to become a B cell. When $A_{th}/B_{th}=0.5$, the sensitivity threshold for A is set to 5 molecules per time step, which is half the sensitivity threshold for B. In this scenario, A cells would turn to B cells much sooner than B cells would turn to A cells, and thus, the system is unable to stabilize. Cells would predominantly oscillate between states A and B, indefinitely. Both thresholds are required to be, at least, equal for transient division of

labour to emerge. Ath has to be substantially higher than Bth (Ath/Bth=3, Ath/Bth=6) so cells can divide labour efficiently and permanently.

Other parameters greatly influence the behaviour of cells as well. The emission rate and the ratio of diffusion and degradation (horizontal axis) determine the spatial scale of the signal, and thus, its concentration on the environment. Low emission rates (0.1 molecules per time step) do not allow the system to start, and so cells remain forever in the A state producing signal. Only when the diffusion rate is high enough (and the degradation rate is low enough) low emitting cells can transition to the B state. Still, the signal scale is spatially too short and to poor, and cells end up oscillating between the two states at the same time (*i.e.* oscillations). The opposite happens when the emission rates are much higher than the sensitivity thresholds (100 molecules per time step) and the degradation rates are low (0.01 molecules per time step). The initial production of signal by all cells that turn into the A state saturate the colony, and so, cells are then, permanently stuck in the B state for as long as the simulations took place (300 minutes). It can be observed that only short-range signals when $Ath/Bth \geq 1$ are able to deliver permanent or transient DOL.

The signal temporal scale also determines the system behaviour. If production is high (higher than the sensitivity thresholds) but the molecule lifetime is short, the resulting reach of the signal will also be local. However, a rapidly changing signal does not give enough time for the system to communicate and organise. Therefore, permanent DOL requires a communication of cells that is local and temporarily stable. On the contrary, if the signal concentration changes too quickly the system will oscillate or, in the best-case scenario, reach a transient DOL.

It was shown how DOL can be achieved transiently, where cells do not permanently specialize but still coordinate to provide global fractions of cells that could be performing two different tasks. This type of transient DOL could be interesting for multicellular organism or bacterial colonies that need to adapt to sudden or temporary changes in the environment. For example, as in the previous section, giving up the ability to reproduce momentarily in favour of helping

others could be more beneficial in evolutionary terms than a permanent specialization. Transient DOL also allows for the compartmentalization of different processes in a delocalized manner, which can also be more efficient in certain environments (Afroz et al. 2014; Kalmar et al. 2009). Depending on the processes that cells would like to complete, sensitivity thresholds could adapt to provide different proportions for the DOL. Also, the physical nature of the coordinating molecule can be tuned to achieve diverse ratios of A and B cells (Dulla & Lindow 2008). A high cost metabolic process may only be carried out by a small fraction of cells, as in the example of soma and germ cells (Kirk 2001). From the results here presented, the basic requirements for a **transient DOL** can be enlisted as:

- (i) $A_{th}/B_{th} \approx 1$
- (ii) Molecule properties should define a short-range spatial scale.

On the other hand, permanent DOL offers the opportunity for total specialization as a result of collective agreement. Because of this, spatial patterns of cells of different nature can emerge. This type of DOL could then be useful when tasks have to be spatially arranged or are incompatible and need to be carried out in separated locations (Zhang et al. 2006). The internal logic could also be found in morphogenetic or developmental genetic networks, as it serves as well as a pattern generator (Payne et al. 2013; Davies 2017; Schaerli et al. 2014). From the results here presented, the basic requirements for a **permanent DOL** can be enlisted as:

- (i) $A_{th}/B_{th} \gg 1$
- (ii) Molecule properties should define at the same time:
 - a. Short-range spatial scale.
 - b. Long-range temporal scale.

5.3.2 Two-signals circuit

In this section, the one-signal motif is extended by adding a second regulatory molecule. The internal and multicellular motif of the proposed circuit is shown in Figure 5.37. The circuit imposes a series of rules that every individual in the colony must follow. The general functioning rules of this extended version are:

- Cells have two possible states, A or B. Both states are mutually exclusive.
- Cells in A emit the diffusive molecule A (green) whereas cells in B emit molecule B (blue).
- Cells work to have both molecules A and B in their surroundings. Therefore, when individual cells sense that the concentration of the molecule they are not producing is below its sensitivity threshold and the one they produce is abundant, they immediately change to the opposite state (*e.g.* if a cell in A senses that the concentration of molecule B is below its threshold while the molecule A is above its threshold, the cell would transit to state B).
- When the concentration of both A and B molecules is above the cell sensitivity threshold, cells will remain in its original state.
- When the concentration of both A and B molecules is below the cell sensitivity threshold, cells will remain in its original state.
- By convention, the system is initialized in the A state.

Figure 5.37– Representation of the two-signals coordinated signalling circuit and its multicellular implications.

These logic rules can be better exemplified with the following analogy. Imagine that the community of cells is a small village. Villagers need both water and bread to live, they organise so that there is always enough water and bread for everybody. At first, everybody goes to the well to get water, and then, some will start baking bread, since there is none yet. The easiest way for the villagers is to share their product with their immediate neighbours. Therefore, if a villager is currently baking bread he will need to get water from his neighbour. Yet, his neighbour has just starting baking bread because his neighbour across the street is collecting water. Then, the first villager will switch his task to water collection and stop making bread, as there is an excess of bread in his surroundings. With time, all villagers will get organised without an explicit general command from the mayor (*i.e.* transient and permanent DOL).

Simulations will explore the properties of the proposed motif in a static, non-growing population of cells. Model parameters in this new motif include the diffusion and degradation rates of molecules A and B, the cells emission rate of molecules A and B, and their respective sensitivity thresholds. Additionally, the production time of molecules is included as a parameter in this model. To start the circuit characterization and its ability to provide any kind of DOL, simulations will consider the symmetric case where molecules A and B share the

same attributes. Also, the sensitivity threshold and the emission rate of cells will also be the same for both A and B molecules. A random set of parameters was chosen following the symmetry requirements and is shown in Table 5-6.

Diffusion rate of molecules	$0.2 \frac{\text{mol} \cdot \text{cell_size}}{dt}$
Degradation rate of molecules	$0.3 \frac{\text{mol}}{dt}$
Emission rate of cells	$10 \frac{\text{mol}}{dt}$
Sensitivity threshold of cells	$10 \frac{\text{mol}}{dt}$
Production time of molecules	20 minutes

Table 5-6. Initial parameters for the two-signal motif simulations.

For this set of parameters, cells oscillate between the state A and B (Figure 5.38). Cells start by simultaneously entering the state A where they express the molecule A (represented in green in the figure). When enough molecule A has built up and the sensitivity threshold has been reached, cells commute to state B where they express molecule B (represented in blue). Because of the diffusive nature of the molecule, these transitions occur in the edge of the colony and spreads towards the centre.

Interestingly, if the initial production of molecules is not synchronized, a different phenomenon is observed. The process of molecule production could be indeed a noisier mechanism. In the following simulations, all parameters remain the same except for the production times of molecules A and B that now present a small variability (20 ± 5 minutes). The result is that cells do not longer oscillate but, after some time, get locked in a given state (Figure 5.39). Dark cells shown are cells in state A that did not start the emission of the associated molecule, because of the variability in the production time. Cells get locked in a given state due to signal coordination among neighbouring cells. Because of the time variability in producing the molecule, there is a chance that a cell meets its requirements for molecule A and B by producing, for example, molecule A and getting molecule B from its surrounding neighbours. This was not possible in the previous simulation because cells were

highly synchronized, and thus, all neighbours were roughly performing the same tasks at the same time.

Figure 5.38— Temporal shots of a colony with the two-signal circuit and parameters from Table 5-6. The graph shows the number of cells in each state over time. The resulting global dynamics are oscillation of cells from state A (green) to B (blue).

In the example shown in Figure 5.39, cells engage in molecule sharing and divide labour in a permanent way. Since cells whose state has been locked influence other cells, the process spreads in the colony creating a very characteristic spatial pattern. The fraction of cells that engage in the process grows linear with time. Initially, cells start to oscillate between the state A and B to meet their molecule requirements. Randomly, a peak of one of the molecules somewhere in space induces that a cell fixes its state. In turn, this cell will induce the fixation of the opposite state in its neighbours, giving birth to a nucleation point. When the nucleation of the pattern starts, it starts to recruit cells that equally split the task of producing both

molecules. Cells that are still not recruited continue to oscillate until they are influenced by their close neighbours. The oscillation in the number of cells starts to lose amplitude as the pattern stabilises, and finally, it reaches a stable equilibrium where half the population is performing task A and the other half is performing task B, all of them in a precise spatial location.

Figure 5.39— Temporal shots of a colony with the two-signal circuit and parameters from Table 5-6. The production time of molecules A and B has a small variability (20 ± 5 minutes). Spatial patterns of cells diving labour in a permanent fashion appear.

The plethora of spatial patterns can be further characterized depending mainly on: 1) number of nucleation points, and 2) the *thickness* of the pattern. Nucleation points are those cells that locally arrange and lock their state independently of the rest of the colony. They are responsible for the propagation of the pattern and cannot be directed as they are a consequence of an emergent self-coordinating process. The *thickness* of the pattern refers to the average number of cells that lie between two equal-performing task regions. In other words, it refers to the thickness of the resulting stripes. In the example of Figure 5.39, the *thickness* of the pattern is quite thin, for both states A and B, as it involves only two cells on average.

So, the first requirement for this motif to provide DOL is that neighbouring cells need some kind of variability in their states. If the system is purely symmetrical and the production time of molecules is synchronized, cells will constantly change their state almost in unison. It would be like having only one cell permanently switching from task A to task B. This was not a requirement in the one-signal circuit because it was already asymmetrical by definition.

The following simulation proves the importance of the initial desynchronization. The standard deviation of the production time is now half the value of the production time (20 ± 10 minutes). This leads to permanent DOL, with different spatial arrangement of cells (Figure 5.40).

Because cells interact with different-state neighbours due to larger variability, more nucleation points are initially formed. The path to equilibrium, or pattern formation, is much faster in this simulation than it was for the one of Figure 5.39. This is because the spatial solution is more constrained due to the many nucleation points. Here, cells are being recruited from five different fronts, which speeds up the process of permanent DOL emergence. Still, notice how the thickness of the pattern remains the same. Cells adapt to meet the five fixed solution points (nucleation points).

Figure 5.40— Temporal shots of a colony with the two-signal circuit and parameters from Table 5-6. The production time of molecules A and B has a large variability (20 ± 10 minutes). Spatial patterns of cells diving labour in a permanent fashion appear with many nucleation points.

The characterization of the motif is followed by varying the emission rate and the sensitivity threshold. Diffusion and degradation rates are those from Table 5-6. The production time of molecules is set to be 20 ± 5 minutes. Remember that cells have the same values of those parameters for state A and state B.

The characterization is performed in order to find the regions where this motif provides any kind of DOL. Simulations were repeated a total of $N=20$ for each combination of parameters. Figure 5.41 summarizes the obtained results with representative colony shots.

Figure 5.41– Spatial patterns for different combinations of emission rates and reception threshold.

Permanent DOL is achieved for almost all the parameter space.

The parameter space has three main regions, namely, when the emission rate is lower, equal or higher than the sensitivity threshold. DOL is achieved in all regions with different spatial and temporal solutions. The spatial solutions (spatial patterns) are affected by the parameter values of these three regions. First, the diagonal represents the region for which the emission rate is equal to the sensitivity threshold (20/20, 10/10 and 5/5). Patterns in the diagonal have an equal overall *thickness* of 2.0 ± 0.1 cells for both A and B regions. On average, they also presented similar number of nucleation points, $N=2.5 \pm 1.5$.

Below the principal diagonal of Figure 5.44, is the region where the emission rate is higher than the sensitivity threshold. Here, the pattern thickens, and the number of nucleation points increases for the same time variability. Positions (10/20) and (3/10) depict a *thickness* of 3.0 ± 0.5 cells and an average number of nucleation points of $N = 5 \pm 2$. In the extreme case of (3/20), there is an excess in signal concentration in the environment, and the aforementioned effect gets amplified, resulting in a *thickness* of 4.0 ± 0.5 cells and an average number of nucleation points of $N = 9 \pm 3$.

Above the diagonal is the region where the emission rate is lower than the sensitivity threshold. Positions (20/10) and (10/3) present the same pattern. The resulting pattern is commonly referred to as “salt & pepper”. It can be considered the thinnest pattern possible (one bacteria) with multiple nucleation points. In the extreme case where the emission is much lower than the reception, (3/20), no stable spatial pattern emerges, only transient DOL. Cells do transition from state A to state B but the pattern cannot be maintained due to the low concentration of signal that implies too fast transitions.

Now, simulations will study the effect of the diffusion rate on the pattern shape (Figure 5.42).

Figure 5.42– Resulting spatial patterns for different diffusion rates of molecules A and B.

Permanent DOL emerges for a wide range of diffusion rates. The pattern thickens as the diffusion rate of molecules increases. This proves that the vicinity is defined by the signal diffusivity (*i.e.* the communication length). The pattern mimics the scale of the diffusive signal. While higher diffusion rates lead to the thickening of the pattern, the main thickener agent is the degradation rate, as it is shown next.

Figure 5.43– Resulting spatial patterns for different degradation rates of molecules A and B.

In Figure 5.43 different degradation rates were tested for the parameters of Table 5-6. When the degradation rate is lower than the reference value ($0.3 \frac{mol}{dt}$), the resulting pattern thickens substantially. Lower degradation rates imply that the signal is spread far from the signal source (long communication range). Besides, the sum of molecules in the space (i.e. the combination of signals from different cells) produces a spatial arrangement which is unordered in space, although it continuously returns an equal fraction of cells in state A and state B. In the extreme of higher degradation rate ($0.5 \frac{mol}{dt}$), the obtained pattern resembles that of a low diffusion rate. In this case, the molecule concentration required for the state to switch can only be reached by the closest neighbours of the source cell, before degrading. This describes a poorly communicated colony that turns into a salt & pepper permanent DOL solution.

The characterization of the motif is followed by a full screening over different diffusion and degradation rates (Figure 5.44). Emission and threshold rates are those from Table 5-6. The production time of molecules is set to be 20 ± 5 minutes. Remember that cells have the same values of those parameters for state A and state B. The characterization is performed in order to find the regions where this motif provides any kind of DOL. Simulations were repeated a total of $N=20$ for each combination of parameters. Figure 5.44 summarizes the obtained results with representative colony shots.

Figure 5.44— Resulting spatial patterns for different degradation rates of molecules A and B.

DOL is achieved for any combination of diffusion and degradation rates. As seen before, if the degradation rate is too high, the spatial pattern is so thin that it is difficult to achieve permanent DOL. In the (0,5/0,5) case, only half the simulations delivered the presented permanent DOL while the other half showed the arranged oscillation typical of transient DOL. The spatial pattern becomes unordered for small degradation rates. This happens when no nucleation points or stripes can be identified.

Finally, the diffusion and degradation rates, as well as, the emission and sensitivity rates are combined to provide a full screening of the model parameters. The production time of

molecules is set at 20 ± 5 minutes. Simulations were run for $N=20$ times for each combination, and representative colony shots are depicted in Figure 5.45.

Figure 5.45– Resulting spatial patterns for different combinations of the ratios of $\frac{\text{Emission rate}}{\text{Sensitivity threshold}}$ and $\frac{\text{Diffusion rate}}{\text{Degradation rate}}$. DOL is achieved for all combinations except when diffusion $\gg$ degradation.

Three different dynamics are obtained. First, DOL is achieved for all combinations except when the diffusion rate ($1 \frac{\text{mol} \cdot \text{cellsize}}{\text{dt}}$) of both molecules is much higher than the degradation rate ($0.01 \frac{\text{mol}}{\text{dt}}$). In these cases (left column), the signal spatial scale is larger than the colony. Cells begin in state A (green) and the production of A saturates the environment. When cells

turn to state B (blue) the same happens. Since cells are instructed to remain in the same state when both molecules are above their sensitivity threshold, they remain in the B state for as long the simulation lasted (300 minutes).

In the remaining parameter space DOL is achieved. Transient DOL emerges when spatial scale of the signal is short and emission is lower than the sensitivity threshold. It was shown how when the emission is lower than the sensitivity threshold, cells form the salt & pepper pattern as it represents the smaller vicinity arrangement. In these cases, cells can only interact with their immediate neighbours because the regulating signal is both spatial and temporary short.

Permanent DOL with ordered patterns appears when the spatial scale of the signal is short, but it is temporarily more stable. This happens when the emission rate is higher than the sensitivity threshold. As seen in Figure 5.45, the larger the difference between those values, the thicker the pattern. Finally, unordered patterns of cells undergoing permanent DOL are found for medium ranges of diffusion and degradation rates.

Independently of the pattern, permanent DOL in these past simulations always provided an equal fraction of cells in state A and B. This is because the assigned parameters were considered the same for both molecules. Still, the fraction of cells in each state can be tuned to create any distribution.

In Figure 5.46, the diffusion and degradation ratio of molecule A is two orders of magnitude higher than the one of B. The emission and sensitivity threshold values were $10 \frac{mol}{dt}$ for both states. Because the lifetime of signal B was shorter, cells were forced to stay in the B state to balance the environmental concentration of both molecules. The resulting behaviour is a permanent DOL with a much higher fraction of cells in state B (95%).

Figure 5.46– Resulting spatial pattern and number of cells in each state when the diffusion and degradation rates ratio of molecule B is two orders of magnitude lower than that of A.

In Figure 5.47, the effect of different sensitivity thresholds is observed. Diffusion and degradation rates were both set to $0,1 \frac{\text{mol}}{\text{dt}}$ and the emission rate was set to $20 \frac{\text{mol}}{\text{dt}}$. The resulting behaviour is a permanent DOL with a higher fraction of cells in state A (63%). A lower sensitivity threshold represents a cell that is more reactive to the signal concentration. That means that the concentration of B is $20/3$ times lower than the concentration of A in the colony.

Figure 5.47– Resulting spatial pattern and number of cells in each state when the sensitivity threshold of A and B is different.

As a summary, this section has demonstrated how the proposed circuit is able to create DOL in a variety of conditions, starting from a homogenous group of cells. The two-signal circuit has proved to be more robust than the one-signal circuit. This means that the extended version is able to provide DOL for a larger parameter region. Not only that, but it also provides a permanent DOL solution for almost all conditions. The only fundamental condition of this logic is that it requires an initial perturbation of the cell states. This is easily achieved with slightly noisier production times of the regulating molecules. Besides, some permanent DOL solutions presented the formation of spatial patterns of specialised cells. These patterns could be categorized through two important features: the number of nucleation points and the motif *thickness*. The final arrangement of cells depends on factors such as the diffusion and degradation rates of the shared molecules, the cell emission rates and cell susceptibility to transition to the opposite state (sensitivity threshold) and, more fundamentally, the time variability in the production of the diffusible molecules.

5.3.3 Fitness benefits of emergent division of labour

So far, both proposed circuits have been characterized and their functioning regions have been identified. The one-signal circuit was able to provide DOL for a smaller region of the parameter space in compare to the two-signal circuit. Here, in this section, instead of simulating non-growing populations of cells, simulations will consider growing cells that divide labour according to the rules of the two proposed circuits. The objective is to test which of the circuits delivers a higher fitness.

The assumptions of the simulations of this section are the following. First, it will be assumed that a cell that is in a productive state (*i.e.* that is producing a molecule) will have an imposed cost that will translate into its individual growth rate. Finally, the secreted molecule(s) will be required by all cells to grow at their maximum growth rates. Model parameters were chosen so both circuits deliver an equal fraction of cells in each state (50% in A and 50% in B). The specific parameters for both models are shown in Table 5-7.

	One-signal	Two-signals
Diffusion rate of molecules	$0.2 \frac{\text{mol} \cdot \text{cell_size}}{dt}$	$0.2 \frac{\text{mol} \cdot \text{cell_size}}{dt}$
Degradation rate of molecules	$0.3 \frac{\text{mol}}{dt}$	$0.3 \frac{\text{mol}}{dt}$
Emission rate of cells	$10 \frac{\text{mol}}{dt}$	$20 \frac{\text{mol}}{dt}$
Sensitivity threshold of A	$40 \frac{\text{mol}}{dt}$	$10 \frac{\text{mol}}{dt}$
Sensitivity threshold of B	$40 \frac{\text{mol}}{dt}$	$10 \frac{\text{mol}}{dt}$
Production time of molecule(s)	20 ± 5 minutes	20 ± 5 minutes
Generation time of producers	50 ± 10 minutes	
Generation time of non-producers	30 ± 10 minutes	

Table 5-7. Model parameters for the fitness simulations of the one-signal and two-signal circuits.

The costs of emission are translated as a larger generation time for the producers. The final imposition was that cells from both models were required to absorb a certain amount of molecule(s) to grow. There was no benefit for absorption but rather a penalty if the minimal

amount was not met. Every simulation step, cells are imposed with the instruction of absorbing a fixed quantity of molecule. The molecule coefficient is calculated according to the absorption of molecule at each simulation time step following the function:

If the available molecule is lower than the required absorption rate, the molecule coefficient will transmit this implication to the bacterium growth rate. The growth rate of each bacterium at each time step is then, calculated as:

$$\text{growth}_{rate} = \text{Basal growth rate} * \text{molecule coefficient}$$

The basal growth rate relates to the generation time as, $\text{growth}_{rate} = \frac{\ln(2)}{\text{generation time (minutes)}}$

The results of simulations with each circuit are summarized in Figure 5.48. Simulations were run for 300 minutes a total of $N=20$ times, with small variability in the final output of cells. Evidently, cells with the one-signal circuit had less costs to carry, as they only produce signal in one of the states. As a result, simulations with the one-signal circuit delivered, on average, twice as many cells as the simulations with the two-signal circuit, for the assumed costs. If the assumption that producing any molecule has an associated cost is correct, as it points to reason, then, wherever the costs, the one-signal circuit will always provide a better fitness. So, even if the two-signal circuit is more robust and requires less parameter tuning, it is also less beneficial in evolutionary terms.

Figure 5.48– Final number of cells for populations using each signalling motif.

5.3.4 Morphogenetic properties: Repeatability, scalability and resizing

In this section, I investigate the potential of the proposed genetic circuits as not only difference generators for DOL, but also their ability to create spatial patterns interesting for morphogenesis. Morphogenesis is the biological process that causes an organism to develop its shape (Davies 2013). This process usually requires the patterning of cells via their spatial distribution during the embryonic development of an organism (Davies 2013). As it turns out, there are common mechanisms found in both DOL and morphogenesis, as seen, for example, in the patterning of nitrogen-fixing cells in cyanobacteria (Zhang et al. 2006). Both morphogenesis and DOL require some sort of specialization or differentiation among cells, but morphogenesis also involves this specialization to be ordered in space (Davies 2017). Some of the mechanisms involved in morphogenesis occur at the cellular level, with cell adhesion, proliferation and motility as mechanical forces that shape the final morphology of an organism (Gumbiner 1996). Still, many morphogenetic processes occur at the genetic and molecular level, where the interaction of gene products and molecules (here named morphogens) are the cause of pattern formation (Kondo & Miura 2010; Hironaka & Morishita 2012).

Figure 5.49– Lateral inhibition model. Cell in state A inhibits that state in its adjacent neighbour.

While many morphogenetic models based their spatial structure on pre-existing information, some interesting models of pattern formation are able to create patterns *de novo* in homogenous fields of cells with no existing cues (Koch & Meinhardt 1994). An important mechanism for creating a difference among initially homogeneous cells is lateral inhibition,

a type of cell-cell interaction whereby a cell that adopts a particular state inhibits its immediate neighbours from adopting the same state (Munteanu & Solé 2008; Collier et al. 1996). This mutual inhibition mechanism between adjacent cells is usually mediated by Delta–Notch signalling system, as it happens in the *Drosophila* neurogenic ectoderm (Campos-Ortega 1995). This mechanism creates a fine-grained pattern of cells in opposite states (Figure 5.49). Another major example of self-regulated pattern formation is the reaction-diffusion model proposed by Turing as the basis for morphogenesis (Turing 1952). The model is based on two diffusible molecules that interact in a certain manner, leading to distinguishable ordered regions of each molecule (Figure 5.50).

Figure 5.50– Turing reaction diffusion model. Activator, A, and Inhibitor, I, interact and diffuse leading to spatial patterns of the concentration of the two molecules.

In the reaction-diffusion model, an activator molecule activates its own production and that of the other molecule, which acts as an inhibitor of the activator. Where activator levels are high, self-induction assures a high concentration of the activator. If the activator concentration is below a certain threshold, the simultaneous activation of the inhibitor will locally deplete the activator molecule. These interactions (short self-activation and long self-repression), together with the physical properties of the diffusible molecules, produce a spaced-out pattern of some regions dominated by the activator and others by the inhibitor. Turing systems are

thought to underlie many different examples of developmental patterning (Maini 2004), including finger formation (Raspopovic et al. 2014) and hair follicle spacing (Sick et al. 2006).

These two important models of morphogenesis are mentioned here as they have served as inspiration for the proposed self-coordinating circuits. Notice that, in general, both the one-signal and the two-signal circuits are also self-regulating and do not need prior cues for the formation of spatial patterns.

First, both proposed circuits employ some sort of lateral inhibition. In these circuits, where differentiation between two exclusive states is emergent, cells would not inhibit its state on neighbouring cells but instead, activate the opposite mutually exclusive state. This interaction is not contact mediated as in lateral inhibition systems, but rather is mediated through the regulatory diffusive molecules.

On the other hand, the basis of the Turing reaction-diffusive model (as seen from the activator perspective) is that two incoherent pathways are initiated at different speeds. A short-term induction is required (the activator directly self-activates), as well as, a long-term repression (the activator self-represses through the activation of its inhibitor). Both proposed circuits utilize this same logic (Figure 5.51).

Figure 5.51– Both circuits (one and two-signals) have a long-term repression (mediated through the diffusive molecule) and a short-term induction (mediated by the repression of the opposite state).

In the one-signal circuit, as seen from the A state, the state self-induces by repressing its inhibitor state, B, within a cell. On the other hand, incoherently, A also self-represses through the emission of the diffusible molecule, which, in turn, weakens the A repression over B.

In the two-signal circuit, this logic applies from both the point of view of A and B states. This could be the reason for its robustness, since the Turing logical rules takes place symmetrically and simultaneously.

In the following, several characteristics of spatial patterns will be investigated in the proposed circuits. These characteristics include the repeatability of patterns, their dependence with size and shape, and their ability to scale as the colony grows. For the sake of simplicity, a single representative pattern was chosen from each motif. The patterns are depicted in Figure 5.52.

Figure 5.52—Representative ordered spatial patterns of the one and two-signals motif.

The one-signal pattern is created with the parameters of Table 5-5, with $A_{th}/B_{th}=60/2$. The two-signal pattern is created with parameters of Table 5-6 and a production time of 20 ± 5 minutes. These patterns are selected as they represented the simplest ordered patterns that result from each motif.

Repeatability

Repeatability refers to the reliability of the emergent spatial pattern in every simulation. Since there are many random inherent processes, a consistent pattern should be the one that maintains its shape, regardless of the noise of initial conditions in a non-growing cell population. The magnitudes of the *thickness* and the number of nucleation points will be used as reference to determine the pattern variation between runs. Nucleation points only apply in the two-signal patterns. Figure 5.53 shows three different runs from the $N=20$ runs that were performed to obtain the average values of the reference magnitudes.

Figure 5.53– Repeatability of the emergent spatial patterns is measured by the *thickness* of the green (T_G) and blue (T_B) regions and, in the case of the two-signals pattern, with the number of nucleation points, N .

The one-signal patterns (Figure 5.53, top row) showed high repeatability with an average *thickness* of 10 green cells in the outer ring and 2 blue cells in the inner ring. The errors shown in Figure 5.53 represent the variation from the average values between runs. The pattern was consistent 100% of the runs, with small variations due to the different initial arrangement of cells.

The two-signal patterns (Figure 5.53, bottom row) showed less repeatability due to the noise in the nucleation points. The appearance of nucleation points is an intrinsic random process and cannot be directed: nor its position neither its number. Still the *thickness* of each coloured stripe presented high repeatability among runs.

Size dependency

Size dependency refers to the dependence of the spatial pattern with the total body size, which in this case is the number of cells. The change in the patterning magnitudes with size as the group of cells grow is called scaling and it will be tested later. Here, cells will not grow. Simulation will be initialised with three different number of cells, namely, 10^2 , 10^3 and 10^4 cells. Each simulation will be run for $N=20$ times, and the average value of the reference magnitudes will be used as indicator of pattern variation between the different initial conditions.

Results are shown in Figure 5.54. The one-signal patterns (top row) present equal *thickness* of the green rings. The blue *thickness* is also conserved for the 10^3 and 10^4 simulations although there is not enough space for a ring in the 10^2 colony to compare with. In general, the ring-like motif scales as expected for the different initial sizes revealing that the pattern is size invariant.

The two-signal patterns (Figure 5.54, bottom row) showed the same phenomena as in the repeatability simulations. Again, the relative *thickness* is conserved across the various initial number of cells. The pattern general design changes due to the randomness (in number and location) of the nucleation points. As commented, nucleation points are a statistical phenomenon that ultimately depends on the number of cells.

Figure 5.54– Size dependency of the emergent spatial patterns is measured by the *thickness* of the green (T_G) and blue (T_B) regions and, in the case of the two-signals pattern, with the number of nucleation points, N .

Therefore, the final arrangement of cells is strongly size dependent for the two-signal pattern. This could be improved if the noise variability in the production time was altogether removed. This, of course, would not return any spatial pattern and only temporal oscillations between both states. In this case, informational cues could be useful to control the number and location of the nucleation points. For example, Figure 5.55 shows the emergent patterns of the noise free two-signal circuit when intentional informational cues were included.

A strategically placed cell that basally emits the A molecule was able to create the stripe-like pattern (left image). The position in where the cells was placed is marked with a yellow dot.

Figure 5.55– Directed patterns created from placing A emitting cells in specific spatial locations for the two-signals motif with no time variability in the molecule production.

The same simulation was performed for four fixed nucleating cells (right image). This could be done for as many nucleating cells and positions as desired, providing a means for this motif to be totally repeatable and even size-independent. Notice that the *thickness* of the pattern is invariant of the number of nucleation points.

Shape dependency

Similar to the size dependency, the shape dependency refers to the correlation between the general shape of the colony and the imprinted pattern. This phenomenon was shown already in Figure 5.35 for the one-signal circuit but is here extended to the two-signal pattern as well. Simulation were performed $N=20$ for two different arrangement of cells. Again, the reference magnitudes shown in Figure 5.56 refer to the average thickness for each of the shapes.

Even if variation between runs is low (high repeatability), the one-signal patterns are strongly influenced by the general contour of the colony. While the *thickness* of the blue region is conserved, the green inner region adapts to the spatial solution and varies strongly between shapes.

Figure 5.56– Directed patterns created from placing A emitting cells in specific spatial locations for the two-signals motif with no time variability in the molecule production.

On the other hand, the two-signal pattern showed that despite the low repeatability in the nucleation points, its average value was not affected by the colony general shape. The *thickness* is also conserved, proving that this motif is shape-independent.

Scaling

Finally, scaling refers to the adaptation of the pattern with size. A scale invariant pattern is able to maintain a constant ratio between the emergent pattern and the changing colony size (Cao et al. 2016). So far, it was studied the properties of pattern formation in a static non-growing population of cells. The question to whether the spatial pattern can be maintained in a population of growing and diving cells has yet to be answered.

Figure 5.57– Pattern scaling of the one-signal motif for two different regimes. When cells growth rate is faster than the signal exchange the pattern cannot form until later stages. If cells growth rate is slower or similar to the temporal scale of the signal, the pattern scales accordingly although it blurs with size.

In Figure 5.57, representative shots of different temporal stages of the one-signal pattern are shown for two conditions. First, if the generation time of the cells is short (*i.e.* they divide fast) compared to the speed of the signal exchange, the resulting pattern forms as the colony expands (top row). As seen in section 5.3.1, the signal temporal scale is defined by the ratio between the diffusion and degradation rate of the signal. If the molecule lifetime is greater than the cells generation time, the pattern is not able to scale in this regime since growth overtakes the circuit capacity to adapt. On the other hand, if the growth rate of cells is larger

than the temporal scale of the signal, the pattern emerges during the early stages. Then, the pattern scales well as it is able to maintain the relative *thickness* of the green rings. Blue cells appear in the right location for the scale to be maintained although the pattern they formed is blurred and discontinuous.

cell growth rate > signal temporal scale

cell growth rate < signal temporal scale

Figure 5.58— Pattern scaling of the two-signal motif for two different regimes. When cells growth rate is faster than the signal exchange the pattern cannot form until later stages. If cells growth rate is slower or similar to the temporal scale of the signal, the pattern scales accordingly although it blurs with size.

Figure 5.58 depicts also both regimes for the two-signal patterns. When cells divide faster than the signal lifetime, reference magnitudes cannot be measured as the pattern only forms for long time periods. In addition, the number of nucleation points grows as cell divide and does not stop increasing until the whole pattern is formed. The emergent pattern maintains the thickness of past simulation even at the cost of the blurriness of the pattern. The more complex the pattern (i.e. more nucleation points) the more difficult it is to self-rescale. However, when the growth process is set at a slower pace, the pattern is formed at the first stage. Then, as the colony grows, the pattern rescales accordingly. The thickness is conserved throughout the entire simulation time and the number of nucleation points remains constant.

Results regarding the scaling of patterns showed two differentiated regimes in both circuits. The signal temporal scale has to be larger than the growth rate. Only then, the circuit can sense, react and adapt to the colony size.

As a summary, both one-signal and two-signal motifs are able to create *de novo* patterns without initial spatial cues or pre-patterning of any kind. Results of both patterns are summarized in the following table (Figure 5.59).

One-signal patterns	Two-signals patterns
High repeatability.	Pattern <i>thickness</i> is repeatable. Nucleation points are random and the overall pattern has low repeatability.
Size-independent.	Pattern <i>thickness</i> is size independent but nucleation points are highly dependent on the initial size of the population. The pattern is size-dependent.
Pattern thickness is shape-dependent.	Shape-independent.

Patterns are scale-invariant as long as the temporal scale of the signal (s) is longer than the growth rate.

Figure 5.59– Comparative table of the properties of the patterns created by the two circuits.

Remember that the repeatability and size dependency of the two-signal patterns can be corrected with the initial disposition of pre-patterned cells. Both motifs could work as pattern generators interesting for morphogenesis (Day & Lawrence 2000), synthetic biology (Basu et

al. 2005; Payne et al. 2013) or even artificial intelligence (Mamei et al. 2006; Mamei et al. 2004).

5.4 Conclusions

Division of labour (DOL) has been defined as an evolutionary process that occurs when cooperating individuals specialize to carry out specific tasks. It requires two fundamental conditions: (i) individuals carry out different tasks (phenotypic variation) and, (ii) the division of tasks provides an inclusive fitness benefit to all of the individuals involved (adaptation) (West & Cooper 2016; Claessen et al. 2014). This chapter has studied the conditions for the emergence and maintenance of DOL of two different tasks in a variety of scenarios. Most relevant conclusions can be enlisted as:

- The first hypothesis that was tested was that for DOL to be an evolutionary successful strategy it should provide a benefit for all individuals (West & Cooper 2016). DOL was implemented through the tasks of production and absorption of two different diffusible molecules. Simulations of populations with different degrees of specialization were compared for different reward functions. Results indicate that DOL would be an evolutionary stable strategy only if: (i) the diffusible molecules remain in the local proximity of the producer and, (ii) the benefits for task specialization are accelerating. Both conclusions, although hypothesized, had never been theoretically or experimentally tested before. Observed results serve as the first prove of the fitness requisites for DOL in a real simulated environment.
- The second tested hypothesis was that for DOL to emerge individuals should exhibit phenotypic variation (Kolter et al. 2015b). DOL was implement through two distinctive phenotypes, namely, germ (reproductive phenotype) and soma (helper, non-reproductive phenotype). Simulations model the emergence of these two phenotypes through different sources of internal variability: mutations (random and irreversible), noise (random and reversible), and coordination through signal emission/reception (not random and

reversible). Results showed that reversible and coordinated DOL was more efficient as a difference generator. The initial hypothesis was proved wrong, as heritable changes are not always required for DOL to emerge. The main result is a proof of the principle that DOL can emerge in genetically homogeneous populations via spontaneous breaking of symmetry by the proposed regulatory mechanism.

- The novel mechanism proposed for consensual and reversible DOL (**one-signal circuit**) was further characterized proving that transient and permanent DOL can emerge as a self-coordinating process among alike cells without the requirement of any fitness reward. The proposed internal cell regulation has demonstrated to create interesting collective behaviour for certain regions of the parameter space. The characterization of the model in terms of its parameters defines specific conditions for transient and permanent DOL:

Transient DOL

- $Ath/Bth \approx 1$
- Molecule sharing should be spatially local.

Permanent DOL

- $Ath/Bth \gg 1$
- Molecule sharing should be spatially local and temporarily long-lived.

- The proposed circuit was extended to add a second regulatory molecule (**two-signal circuit**). Characterization of this novel motif has demonstrated its ability to permanently generate DOL in a variety of conditions, only with the requisite of a small source of variation among cells. This is achieved through small variability in the production time of the regulating molecules. In general, the two-signal circuit has proved to be more robust as DOL generator than the one-signal motif in simulations of non-growing cells.
- A comparison of the fitness benefits of each circuit in a cooperative scenario showed that the two-signal circuit is less beneficial in evolutionary terms. This follows under the assumption that molecule producing is metabolically more costly than non-production.
- In conditions where the reproductive rate of cells is slower than the half-life of the molecules, both circuits are able to create *de novo patterns* that can scale with size. The one-signal pattern shows a high repeatability together with great size invariance. Finally, the two-signal circuit shows great *thickness* repeatability and shape invariance.

Chapter 6 Conclusions

Understanding the evolution of cooperation is still nowadays one of the major evolutionary questions. The importance of such question relies on a key consequence derived from cooperation. When cooperation is maintained and promoted in a group, a major transition in individuality can occur where groups of individual cooperators transcend to a new self-replicating entity: multicellularity. The main evolutionary requirements for multicellularity include the formation of groups of cells and the maintenance of cooperation among the individuals of the group. In this thesis, I have tried to investigate in isolation the different steps that lead to the emergence of multicellularity, as well as study one of the most important benefits of multicellularity: division of labour.

Chapter 3 investigated the effect of the physical spatial constraints in the process of *group* formation in two-dimensional bacterial colonies constituted of rod-shape cells. The proposed metric of *physically connected group* was able to measure the spatial effects of cell size and resource acquisition in the stability of the *groups*. The general conclusion is that spatial competition improves *group* stability by increasing the segregation of cells into clonal *groups*. This effect is further amplified if competition for resources is high. The counterintuitive inference is that individual cell competition (for space and resources) promotes the assortment of groups into clonal clusters. This can in turn influence the relative fitness benefits of the groups, which will primarily consist on a fixed phenotype (mainly, cooperators and defectors). Therefore, stable groups of cooperators can be maintained if the individual cells that comprise them are embedded in a high competitive environment. In addition to this mayor conclusion, it was proved how small and homogenous size cells formed colonies of groups that showed more cohesion. Differences in sizes could act as exploitative strategies that destabilize the group. Differentially larger sizes would have an evolutionary advantage when competition for resources is heavy.

In **Chapter 4** simulations represented a population based on the consumption of a public good, where cells themselves were responsible for the production of the diffusible good. Simulations studied different simple strategies and environmental conditions for maintaining and promoting cooperation in a scenario where non-producers (defectors) had already infested a population of producers (cooperators). Results showed that the first and foremost important factor for cooperators not to be displaced by defectors is the production cost of the public good. If the costs of public good production (*i.e.* cooperation) are sufficiently low, there is no reproductive advantage in exploitation. The strategy that resulted in the highest promotion of the cooperative phenotype was that of lowering the production rate and diffusivity of the good. This strategy showed protection for producers almost independently of its costs without the need of other more complicated mechanisms such as conjugation or QS.

Chapter 5 studied the conditions for the emergence and maintenance of division of labour of two different tasks in a variety of scenarios. Results indicate that dividing the task of emitting two different molecules, in different cells within the same group, is more evolutionary beneficial if the group is small and the benefits for task specialization are, at least, exponential with the degree of specialization. The most important contribution of this chapter is the proposal of two different regulatory mechanisms that provided evidence that division of labour can emerge in genetically homogeneous populations via the coordination of cells through the emission and sensing of diffusible molecules. The proposed regulatory mechanisms are original and lead to interesting spatial patterns that are unprecedented to this date.

Understanding natural systems represents a great challenge due to the complexity that governs them. This is why theoretical and computational models often end up recreating biologically inspired artificial systems. Still, it is possible to make complex artificial systems to simulate natural phenomena. The aim of this thesis has not been to recreate in detail microbial or cellular systems, but rather, derive general principles from individual rules of behaviour and interaction in a controlled artificial environment. Throughout this thesis, it has been possible to identify common principles underlying the processes here discussed: competition creates

more cohesive cell groups, cooperation requires individual selfishness to thrive or local cell coordination can result in division of labour.

This thesis would not have been possible without the considerable previous work of developing and extending the general features of the simulator *gro* (Gutiérrez et al. 2017). The individual based model approach, and more specifically the use of *gro*, has allowed to propose experiments in different biological contexts that resulted in the emergence of many high level phenomena. *gro* was able to simulate simple and general interaction rules of real ecosystems creating artificial ecosystems with patterns similar to what are observed in nature. It was proved how these simulations can be helpful to understand theoretical questions regarding evolutionary process. The major and more general contribution of this thesis has been demonstrating the viability and versatility of IbMs as methodological tools for the investigation of natural collective phenomena. In particular, *gro* has proved to be a useful platform to investigate several broad ecological questions, as well as long-term evolutionary patterns and processes in biology and ecology.

6.1 Future lines

Many of the assumptions and scenarios proposed on this thesis could benefit from a more exhaustive exploration using the same methodology, but that regrettably are beyond the scope of this thesis.

To begin with, it would be interesting to cross-validate important results and conclusions with other similar frameworks. There are a series of related individual-cell models that were built to simulate with high detail some specific problems. For example, simulators like iDynoMiCS (Lardon et al. 2011) utilize chemical equations that allow for more realistic simulations of the diffusion of chemical fields such as nutrients or signals. Other frameworks, like CellModeller (Rudge et al. 2012) have a more precise physical engine that permits simulations of 3D environments and also cell adhesion. Replicating the experiments presented in this thesis with those frameworks could inform, for example, about the relevancy of the effect of the third dimension in bacterial groups. The cell-based Chaste software (Mirams et al. 2013) can also account for changes in cell–cell adhesion and it is ideal for dense tissue-like environments. Experiments on DOL and pattern formation utilizing this software could improve the biological relevancy of the conclusions presented in this thesis.

As mentioned in the beginning of Chapter 4, the effect of the initial density of cells and the initial frequency of cooperators can be also determinant in the evolution of the cooperative phenotype in the scenario represented in that Chapter. It would be interesting to conduct follow-up simulations to analyse the impact of such factors together with the proposed strategies to better understand their impact on the obtained results. These follow-up simulations could also include cell death, which would require including further interesting assumptions such as how long for a non-growing cell to be presumed dead and, how realistic is to remove it from the simulated plate to leave space for the other cells.

On the theoretical front, I want to highlight two of the many proposed experiments in this thesis that could benefit from a more rigorous and quantitative study. First, conditions established in section 5.1 favouring DOL, could be more rigorously established. Similar to the

work performed in (Tsoi et al. 2018), where they analysed different metabolic pathways to derive a general criterion for outperforming DOL, the different degrees of specialization could be measured as function of their proposed readout, the maximization of the overall productivity. It is worth mentioning that, although initial assumptions and models greatly differ, resulting conclusions offer the same principles.

The second proposal that could benefit of a more analytical study is the intracellular motifs described in section 5.3. The proposed dynamics could be described into partial differential equations where a more comprehensive sensitivity analysis of the model parameters could be carried out. Furthermore, transferring the genetic motifs into a synthetic network to experimentally study its potential as a pattern generator could allow the development of high-level principles of biological self-organization that underlie embryogenesis in general (Isalan et al. 2005; Boehm et al. 2018; Tayar et al. 2017).

Finally, other results that should be experimentally tested include the observed dependency of cell competition with cell size (as it was proposed in section 3.5) and the model assumption that growth rate is independent of size so the same amount of nutrients yields the same mass creation (as it was proposed in section 3.4). Such experiments would provide important information on the cellular nutrient uptake process and the dynamics of growth.

Appendix

The specifications of the simulations presented in this thesis and videos can be found [here](#).

References

- Abbott, S. & Fairbanks, D.J., 2016. Experiments on Plant Hybrids by Gregor Mendel. *Genetics*, 204(2), pp.407–422. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/27729492>
- Abedin, M. & King, N., 2010. Diverse evolutionary paths to cell adhesion. *Trends in cell biology*, 20(12), pp.734–42. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/20817460>
- Ackermann, M., 2015. A functional perspective on phenotypic heterogeneity in microorganisms. *Nature Publishing Group*, 13(8), pp.497–508. Available at: <http://dx.doi.org/10.1038/nrmicro3491%5Cnpapers2://publication/doi/10.1038/nrmicro3491>.
- Adami, C., Schossau, J. & Hintze, A., 2016. Evolutionary game theory using agent-based methods. *Physics of Life Reviews*, 19, pp.1–26. Available at: <https://www.sciencedirect.com/science/article/pii/S1571064516300884>.
- Afroz, T. et al., 2014. Bacterial sugar utilization gives rise to distinct single-cell behaviours. *Molecular microbiology*, 93(6), pp.1093–1103. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/24976172>.
- Allen, R.C. et al., 2016. Quorum sensing protects bacterial co-operation from exploitation by cheats. *The ISME journal*, 10 (January), pp.1–11. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/26744811>.
- Andrup, L. & Andersen, K., 1999. A comparison of the kinetics of plasmid transfer in the conjugation systems encoded by the F plasmid from *Escherichia coli* and plasmid pCF10 from *Enterococcus faecalis*. *Microbiology*, 145(8), pp.2001–2009.
- Arutyunov, D. & Frost, L.S., 2013. F conjugation: Back to the beginning. *Plasmid*, 70(1), pp.18–32. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/23632276>
- Asfahl, K.L. & Schuster, M., 2017. Social interactions in bacterial cell-cell signaling. *FEMS Microbiology Reviews*, 41(1), pp.92–107.
- Axelrod, R. & Hamilton, W.D., 1981. The Evolution of Cooperation. *Science, New Series*, 211(4489), pp.1390–1396. Available at: <http://links.jstor.org/sici?sici=0036-8075%2819810327%293%3A211%3A4489%3C1390%3ATEOC%3E2.0.CO%3B2-6>.
- Baquero, F., Coque, T.M. & de la Cruz, F., 2011. Ecology and evolution as targets: the need for novel eco-evo drugs and strategies to fight antibiotic resistance. *Antimicrobial agents and chemotherapy*, 55(8), pp.3649–60. Available at: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3147629&tool=pmcentrez>

&rendertype=abstract.

- Basu, S. et al., 2005. A synthetic multicellular system for programmed pattern formation. *Nature*, 434(7037), pp.1130–4. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/15858574>.
- Bedau, M.A., 2003. Artificial life: Organization, adaptation and complexity from the bottom up. *Trends in Cognitive Sciences*, 7(11), pp.505–512. Available at: <http://people.reed.edu/~mab/publications/papers/BedauTICS03.pdf>.
- Bell, G. & Mooers, A.O., 1997. Size and complexity among multicellular organisms. *Biological Journal of the Linnean Society*, 60(3), pp.345–363. Available at: <https://academic.oup.com/biolinnean/article-lookup/doi/10.1111/j.1095-8312.1997.tb01500.x>.
- Biernaskie, J.M. & West, S.A., 2015. Cooperation, clumping and the evolution of multicellularity. *Proceedings of the Royal Society B: Biological Sciences*, 282(1813), p.20151075. Available at: <http://rspb.royalsocietypublishing.org/lookup/doi/10.1098/rspb.2015.1075>.
- Blanchard, A.E., Celik, V. & Lu, T., 2014. Extinction, coexistence, and localized patterns of a bacterial population with contact-dependent inhibition. *BMC systems biology*, 8(1), p.23. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/24576330>.
- Boehm, C.R., Grant, P.K. & Haseloff, J., 2018. Programmed hierarchical patterning of bacterial populations. *Nature Communications*, 9(1), p.776. Available at: <http://www.nature.com/articles/s41467-018-03069-3>.
- Bonner, J.T., 1998. The origins of multicellularity. *Integrative Biology: Issues, News, and Reviews*, 1(1), pp.27–36. Available at: <http://doi.wiley.com/10.1002/%28SICI%291520-6602%281998%291%3A1%3C27%3A%3AAID-INBI4%3E3.0.CO%3B2-6>.
- Boyle, K.E. et al., 2013. Exploiting social evolution in biofilms. *Current Opinion in Microbiology*, 16(2), pp.207–212.
- Breckling, B. et al., 2005. Emergent properties in individual-based ecological models - Introducing case studies in an ecosystem research context. *Ecological Modelling*, 186(4), pp.376–388.
- del Campo, I. et al., 2012. Determination of conjugation rates on solid surfaces. *Plasmid*, 67(2), pp.174–182. Available at: <http://linkinghub.elsevier.com/retrieve/pii/S0147619X12000121>.
- Campos-Ortega, J.A., 1995. Genetic mechanisms of early neurogenesis in *Drosophila melanogaster*. *Molecular Neurobiology*, 10(2–3), pp.75–89. Available at: <http://link.springer.com/10.1007/BF02740668>.
- Cao, Y. et al., 2016. Collective Space-Sensing Coordinates Pattern Scaling in Engineered Bacteria. *Cell*, 165(3), pp.620–630. Available at:

- <http://www.ncbi.nlm.nih.gov/pubmed/27104979>.
- Cavaliere, M. et al., 2017a. Cooperation in microbial communities and their biotechnological applications. *Environmental Microbiology*, 19(8), pp.2949–2963.
- Cavaliere, M. & Poyatos, J.F., 2013. Plasticity facilitates sustainable growth in the commons. *Journal of the Royal Society, Interface*, 10(81), p.20121006. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/23365195>.
- Cho, H. et al., 2007. Self-organization in high-density bacterial colonies: Efficient crowd control. *PLoS Biology*, 5(11), pp.2614–2623.
- Claessen, D. et al., 2014. Bacterial solutions to multicellularity: a tale of biofilms, filaments and fruiting bodies. *Nature reviews. Microbiology*, 12(2), pp.115–24. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/24384602>.
- Collier, J.R. et al., 1996. Pattern formation by lateral inhibition with feedback: a mathematical model of delta-notch intercellular signalling. *Journal of theoretical biology*, 183(4), pp.429–46. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/9015458>.
- Cooper, a L., Dean, a C. & Hinshelwood, C., 1968. Factors affecting the growth of bacterial colonies on agar plates. *Proceedings of the Royal Society of London. Series B, Containing papers of a Biological character. Royal Society (Great Britain)*, 171(23), pp.175–199.
- Cornforth, D.M. et al., 2014. Combinatorial quorum sensing allows bacteria to resolve their social and physical environment. *Proceedings of the National Academy of Sciences*, pp.1–5. Available at: <http://www.pnas.org/cgi/doi/10.1073/pnas.1319175111>.
- Crespi, B.J., 2001. The evolution of social behavior in microorganisms. *Trends in ecology & evolution*, 16(4), pp.178–183.
- Czárán, T. & Hoekstra, R.F., 2009. Microbial communication, cooperation and cheating: Quorum sensing drives the evolution of cooperation in bacteria. *PLoS ONE*, 4(8).
- Damore, J.A. & Gore, J., 2012. Understanding microbial cooperation. *Journal of Theoretical Biology*, 299, pp.31–41. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/21419783>.
- Darwin, C., 1859. *On the origin of species by means of natural selection, or, the preservation of favoured races in the struggle for life*, London : John Murray,. Available at: <https://www.biodiversitylibrary.org/bibliography/68064>.
- Datta, M. Sen et al., 2013. Range expansion promotes cooperation in an experimental microbial metapopulation. *Proceedings of the National Academy of Sciences of the United States of America*, 110(18), pp.7354–9. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/23569263>.
- Davidson, C.J. & Surette, M.G., 2008. Individuality in bacteria. *Annu.Rev.Genet.*, 42(0066–4197 (Print)), pp.253–268. Available at: [c:%5CKARSTEN%5CPDFs%5CGrundlagen-](http://www.ncbi.nlm.nih.gov/pubmed/18688111)

- PDFs%5CGrund-2008%5CDavidson - Surette-Individuality in bacteria.pdf.
- Davies, J., 2017. Using synthetic biology to explore principles of development. *Development*, 144(7), pp.1146–1158. Available at: <http://dev.biologists.org/lookup/doi/10.1242/dev.144196>.
- Davies, J.A., 2013. *Mechanisms of morphogenesis*, Elsevier Academic Press. Available at: <https://www.sciencedirect.com/science/book/9780123910622>.
- Davies, J.A. & Davies, J.A., 2013. Chapter 2 – Key Principles of Morphogenesis. In *Mechanisms of Morphogenesis*. pp. 7–16.
- Day, S.J. & Lawrence, P.A., 2000. Measuring dimensions: the regulation of size and shape. *Development (Cambridge, England)*, 127(14), pp.2977–87. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/10862736>.
- DeAngelis, D.L. & Grimm, V., 2014. Individual-based models in ecology after four decades. *F1000prime reports*, 6(June), p.39.
- DeAngelis, D.L. & Mooij, W.M., 2005. Individual-Based Modeling of Ecological and Evolutionary Processes. *Annual Review of Ecology, Evolution, and Systematics*, 36(2005), pp.147–168. Available at: <http://arjournals.annualreviews.org/doi/abs/10.1146/annurev.ecolsys.36.102003.152644>.
- Diggle, S.P. et al., 2007. Cooperation and conflict in quorum-sensing bacterial populations. *Nature*, 450(7168), pp.411–414. Available at: <http://www.nature.com/doi/10.1038/nature06279>.
- Dimitriu, T. et al., 2014. Genetic information transfer promotes cooperation in bacteria. *Proceedings of the National Academy of Sciences*, pp.1–6. Available at: <http://www.pnas.org/cgi/doi/10.1073/pnas.1406840111>.
- Dimitriu, T. et al., 2015. Mobile genetic elements are involved in bacterial sociality. *Mobile genetic elements*, 5(1), pp.7–11. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/26435881>.
- Drescher, K. et al., 2014. Solutions to the public goods dilemma in bacterial biofilms. *Current Biology*, 24(1), pp.50–55. Available at: <https://www.sciencedirect.com/science/article/pii/S0960982213012712>.
- Dulla, G. & Lindow, S.E., 2008. Quorum size of *Pseudomonas syringae* is small and dictated by water availability on the leaf surface. *Proceedings of the National Academy of Sciences of the United States of America*, 105(8), pp.3082–7. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/11248098>.
- Dunny, G.M., Brickman, T.J. & Dworkin, M., 2008. Multicellular behavior in bacteria: communication, cooperation, competition and cheating. *BioEssays : news and reviews in molecular, cellular and developmental biology*, 30(4), pp.296–8. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/18348154>.

- Van Dyken, J.D. et al., 2013. Spatial population expansion promotes the evolution of cooperation in an experimental prisoner's dilemma. *Current Biology*, 23(10), pp.919–923.
- Estrela, S. & Brown, S.P., 2013. Metabolic and Demographic Feedbacks Shape the Emergent Spatial Structure and Function of Microbial Communities S. Allesina, ed. *PLoS Computational Biology*, 9(12), p.e1003398. Available at: <http://dx.plos.org/10.1371/journal.pcbi.1003398>.
- Farrell, F.D. et al., 2017. Mechanical interactions in bacterial colonies and the surfing probability of beneficial mutations. *Journal of the Royal Society, Interface*, 14(131). Available at: <http://www.ncbi.nlm.nih.gov/pubmed/28592660>.
- Farrell, F.D.C. et al., 2013. Mechanically driven growth of quasi-two-dimensional microbial colonies. *Physical Review Letters*, 111(Octobre), pp.1–5.
- Fiegna, F. & Velicer, G.J., 2003. Competitive fates of bacterial social parasites: persistence and self-induced extinction of *Myxococcus xanthus* cheaters. *Proceedings of the Royal Society B: Biological Sciences*, 270(1523), pp.1527–1534. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/12965020>.
- Fisher, R.A., 1930. *The genetical theory of natural selection.*, Oxford: Clarendon Press. Available at: <https://www.biodiversitylibrary.org/bibliography/27468>.
- Fisher, R.M., 2011. The Major Transitions in Evolution Revisited. Available at: <http://mitpress.universitypressscholarship.com/view/10.7551/mitpress/9780262015240.001.0001/upso-9780262015240>.
- Flemming, H.C. et al., 2016. Biofilms: An emergent form of bacterial life. *Nature Reviews Microbiology*, 14(9).
- Fletcher, J.A. & Doebeli, M., 2009. A simple and general explanation for the evolution of altruism. *Proceedings of the Royal Society of London*, 276(September 2008), pp.13–19.
- Foster, K.R., 2011. The sociobiology of molecular systems. *Nature Reviews Genetics*, 12(3), pp.193–203. Available at: <http://dx.doi.org/10.1038/nrg2903>.
- Foster, K.R. & Bell, T., 2012. Competition, not cooperation, dominates interactions among culturable microbial species. *Current biology : CB*, 22(19), pp.1845–50. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/22959348>.
- Fox, R.E. et al., 2008. Spatial structure and nutrients promote invasion of IncP-1 plasmids in bacterial populations. *The ISME journal*, 2(10), pp.1024–39. Available at: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2605096&tool=pmcentrez&rendertype=abstract>.
- Frost, L.S. et al., 2005. Mobile genetic elements: the agents of open source evolution. *Nat.Rev.Microbiol.*, 3(1740–1526 (Print)), pp.722–732. Available at: [c:%5CKarsten%5CPDFs%5CBakteriologie-PDFs%5CBakt-2005%5CFrost et al.- Mobile genetic elements- the agents of open source evolution.pdf](http://www.ncbi.nlm.nih.gov/pmc/articles/PMC1526174/).

- Fujikawa, H. & Morozumi, S., 2005. Modeling Surface Growth of *Escherichia coli* on Agar Plates. *Applied and environmental microbiology*, 71(12), pp.7920–26. Available at: <http://aem.asm.org/content/71/12/7920>.
- Garcillán-Barcia, M.P., Alvarado, A. & de la Cruz, F., 2011. Identification of bacterial plasmids based on mobility and plasmid population biology. *FEMS microbiology reviews*, 35(5), pp.936–56. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/21711366>.
- Garcillán-Barcia, M.P. & de la Cruz, F., 2008. Why is entry exclusion an essential feature of conjugative plasmids? *Plasmid*, 60(1), pp.1–18.
- Gavrilets, S., 2010. Rapid Transition towards the Division of Labor via Evolution of Developmental Plasticity C. T. Bergstrom, ed. *PLoS Computational Biology*, 6(6), p.e1000805. Available at: <http://dx.plos.org/10.1371/journal.pcbi.1000805>.
- Germerodt, S. et al., 2016. Pervasive Selection for Cooperative Cross-feeding in Bacterial Communities. *PLoS Computational Biology*, pp.1–21.
- Ghigo, J.M., 2001. Natural conjugative plasmids induce bacterial biofilm development. *Nature*, 412(6845), pp.442–5. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/11473319>.
- Ghoul, M. et al., 2014. An experimental test of whether cheating is context dependent. *Journal of Evolutionary Biology*, 27(3), pp.551–556. Available at: <http://doi.wiley.com/10.1111/jeb.12319>.
- Ghoul, M. & Mitri, S., 2016. Special Series: Microbial Communities The Ecology and Evolution of Microbial Competition. *Trends in Microbiology*, pp.1–13. Available at: <http://dx.doi.org/10.1016/j.tim.2016.06.011>.
- Goldsby, H.J. et al., 2014. The Evolutionary Origin of Somatic Cells under the Dirty Work Hypothesis L. Keller, ed. *PLoS Biology*, 12(5), p.e1001858. Available at: <http://dx.plos.org/10.1371/journal.pbio.1001858>.
- Gooding-townsend, R., Holder, S.T.E.N. & Ingalls, B., 2015. Displacement of Bacterial Plasmids by Engineered Unilateral Incompatibility. , (April), pp.19–21.
- Gorochoowski, T.E., 2016. Agent-based modelling in synthetic biology. *Essays In Biochemistry*, 60(4), pp.325–336.
- Gounand, I. et al., 2016. Size evolution in microorganisms masks trade-offs predicted by the growth rate hypothesis. *Proceedings. Biological sciences*, 283(1845), p.20162272. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/28003453>.
- Grimm, V. et al., 2010. The ODD protocol: A review and first update. *Ecological Modelling*, 221(23), pp.2760–2768. Available at: <http://dx.doi.org/10.1016/j.ecolmodel.2010.08.019>.
- Gumbiner, B.M., 1996. Cell adhesion: The molecular basis of tissue architecture and morphogenesis. *Cell*, 84(3), pp.345–357. Available at:

- <http://www.ncbi.nlm.nih.gov/pubmed/8608588>
- Gutiérrez, M. et al., 2017. A New Improved and Extended Version of the Multicell Bacterial Simulator gro. *ACS Synthetic Biology*, 6(8), pp.1496–1508. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/28438021>
- H. Koschwanez, J., R. Foster, K. & W. Murray, A., 2011. Sucrose Utilization in Budding Yeast as a Model for the Origin of Undifferentiated Multicellularity L. Keller, ed. *PLoS Biology*, 9(8), p.e1001122. Available at: <http://dx.plos.org/10.1371/journal.pbio.1001122>
- Haft, R.J.F., Mittler, J.E. & Traxler, B., 2009. Competition favours reduced cost of plasmids to host bacteria. *The ISME journal*, 3(7), pp.761–9. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/19340086>
- Hall-Stoodley, L., Costerton, J.W. & Stoodley, P., 2004. Bacterial biofilms: from the natural environment to infectious diseases. *Nature reviews. Microbiology*, 2(2), pp.95–108. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/15040259>
- Hallatschek, O. et al., 2007. Genetic drift at expanding frontiers promotes gene segregation. *Proceedings of the National Academy of Sciences of the United States of America*, 104(50), pp.19926–30. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/18056799>
- Hamilton, W.D., 1964. The genetical evolution of social behaviour. *Journal of Theoretical Biology*, 7(1), pp.1–16. Available at: <http://www.sciencedirect.com/science/article/pii/0022519364900384>
- Harcombe, W.R. et al., 2014. Metabolic resource allocation in individual microbes determines ecosystem interactions and spatial dynamics. *Cell Reports*, 7, pp.1104–1115.
- Hellweger, F.L. et al., 2016. Advancing microbial sciences by individual-based modelling. *Nature Reviews Microbiology*, 14, pp.461–471.
- Hellweger, F.L. & Bucci, V., 2009. A bunch of tiny individuals-Individual-based modeling for microbes. *Ecological Modelling*, 220(1), pp.8–22.
- Hironaka, K. & Morishita, Y., 2012. Encoding and decoding of positional information in morphogen-dependent patterning. *Current opinion in genetics & development*, 22(6), pp.553–61. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/23200115>
- Hol, F.J.H. et al., 2013. Spatial Structure Facilitates Cooperation in a Social Dilemma: Empirical Evidence from a Bacterial Community. *PLoS ONE*, 8(10), p.e77042. Available at: <http://dx.plos.org/10.1371/journal.pone.0077042>.
- Isalan, M., Lemerle, C. & Serrano, L., 2005. Engineering gene networks to emulate Drosophila embryonic pattern formation. *PLoS biology*, 3(3), p.e64. Available at: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=1044831&tool=pmcentrez&rendertype=abstract>
- Ispolatov, I., Ackermann, M. & Doebeli, M., 2012. Division of labour and the evolution of

- multicellularity. *Proceedings of the Royal Society B: Biological Sciences*, 279(1734), pp.1768–1776.
- Jang, S.S. et al., 2012. Specification and simulation of synthetic multicelled behaviors. *ACS Synthetic Biology*, 1(8), pp.365–374.
- Johnson, T.J. & Nolan, L.K., 2009. Pathogenomics of the virulence plasmids of *Escherichia coli*. *Microbiology and molecular biology reviews : MMBR*, 73(4), pp.750–74.
Available at:
<http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2786578&tool=pmcentrez&rendertype=abstract>
- Kalmar, T. et al., 2009. Regulated Fluctuations in Nanog Expression Mediate Cell Fate Decisions in Embryonic Stem Cells M. A. Goodell, ed. *PLoS Biology*, 7(7), p.e1000149. Available at: <http://dx.plos.org/10.1371/journal.pbio.1000149>
- Keen, E.C., 2012. Paradigms of pathogenesis: targeting the mobile genetic elements of disease. *Frontiers in cellular and infection microbiology*, 2(December), p.161.
Available at:
<http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3522046&tool=pmcentrez&rendertype=abstract>
- Keller, L. & Surette, M.G., 2006. Communication in bacteria: an ecological and evolutionary perspective. *Nature reviews. Microbiology*, 4(4), pp.249–58. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/16501584>
- Kennard, A.S. et al., 2015. Individuality and universality in the growth-division laws of single *E. coli* cells. *arXiv*.
- Kim, W., Levy, S.B. & Foster, K.R., 2008a. Rapid radiation in bacteria leads to a division of labour. *Nature Communications*, 354(1388), pp.1395–1405. Available at: <http://dx.doi.org/10.1038/ncomms10508>.
- Kim, W., Levy, S.B. & Foster, K.R., 2008b. Rapid radiation in bacteria leads to a division of labour. *Nature Communications*, 354(1388), pp.1395–1405. Available at: <http://dx.doi.org/10.1038/ncomms10508>.
- Kirk, D.L., 2005. A twelve-step program for evolving multicellularity and a division of labor. *BioEssays*, 27(3), pp.299–310.
- Kirk, D.L., 2001. Germ–Soma Differentiation in *Volvox*. *Developmental Biology*, 238(2), pp.213–223. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/11784005>
- Koch, a. J. & Meinhardt, H., 1994. Biological pattern formation: From basic mechanisms to complex structures. *Reviews of Modern Physics*, 66(4), pp.1481–1507.
- Kolter, R., Vlamakis, H. & van Gestel, J., 2015b. Division of Labor in Biofilms: the Ecology of Cell Differentiation. *Microbiology Spectrum*, 3(2). Available at: <http://www.asmscience.org/content/journal/microbiolspec/10.1128/microbiolspec.MB-0002-2014>

- Kondo, S. & Miura, T., 2010. Reaction-diffusion model as a framework for understanding biological pattern formation. *Science (New York, N.Y.)*, 329(5999), pp.1616–20. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/20929839>
- Kramer, J. & Meunier, J., 2016. Kin and multilevel selection in social evolution: a never-ending controversy? *F1000Research*, 5. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/27158472>
- Kreft, J.-U. et al., 2013. Mighty small: Observing and modeling individual microbes becomes big science. *Proceedings of the National Academy of Sciences of the United States of America*, 110(45), pp.18027–8.
- Kreft, J.U., Booth, G. & Wimpenny, J.W.T., 2000. Applications of individual-based modelling in microbial ecology. *Microbial biosystems: New frontiers (Proceedings of the 8th international symposium on microbial ecology)*, pp.917–923.
- Kruse, T. et al., 2001. C-signal: a cell surface-associated morphogen that induces and coordinates multicellular fruiting body morphogenesis and sporulation in *Myxococcus xanthus*. *Molecular microbiology*, 40(1), pp.156–68. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/11298283>
- Kümmerli, R. et al., 2009. Viscous medium promotes cooperation in the pathogenic bacterium *Pseudomonas aeruginosa*. *Proceedings. Biological sciences*, 276(1672), pp.3531–8. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/19605393>
- de la Cruz, F. et al., 2010. Conjugative DNA metabolism in Gram-negative bacteria. *FEMS microbiology reviews*, 34(1), pp.18–40.
- Lardon, L.A. et al., 2011. iDyNoMiCS: next-generation individual-based modelling of biofilms. *Environmental microbiology*, 13(9), pp.2416–34.
- Leinweber, A., Fredrik Inglis, R. & Kümmerli, R., 2017. Cheating fosters species co-existence in well-mixed bacterial communities. *ISME Journal*, 11(5), pp.1179–1188.
- Levy, R. & Borenstein, E., 2012. Reverse Ecology: from systems to environments and back. *Advances in experimental medicine and biology*, 751, pp.329–345.
- Lewontin, R.C., 1970. The Units of Selection. *Annual Review of Ecology and Systematics*, 1(1), pp.1–18. Available at: <http://www.annualreviews.org/doi/10.1146/annurev.es.01.110170.000245>
- Li, Y.-H. & Tian, X., 2012. Quorum sensing and bacterial social interactions in biofilms. *Sensors (Basel, Switzerland)*, 12(3), pp.2519–38. Available at: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3376616&tool=pmcentrez&rendertype=abstract>
- Liao, D. & Tlsty, T.D., 2014. Evolutionary game theory for physical and biological scientists. I. Training and validating population dynamics equations. *Interface focus*, 4(4), p.20140037. Available at: <http://www.scopus.com/inward/record.url?eid=2-s2.0-84903164215&partnerID=tZOtx3y1>

- Lindemann, S.R. et al., 2016. Engineering microbial consortia for controllable outputs. *The ISME Journal (in press)*, pp.1–8. Available at: <http://dx.doi.org/10.1038/ismej.2016.26>.
- Lloyd, D.P. & Allen, R.J., 2015. Competition for space during bacterial colonization of a surface. *Journal of The Royal Society Interface*, 12(110), p.20150608. Available at: <http://rsif.royalsocietypublishing.org/lookup/doi/10.1098/rsif.2015.0608>.
- Lyons, N.A. & Kolter, R., 2017. *Bacillus subtilis* Protects Public Goods by Extending Kin Discrimination to Closely Related Species. *mBio*, 8(4), pp.e00723-17. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/28679746>
- Lyons, N.A. & Kolter, R., 2015. On the evolution of bacterial multicellularity. *Current opinion in microbiology*, 24, pp.21–8. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/25597443>
- Lyons, N. a & Kolter, R., 2015. On the evolution of bacterial multicellularity. *Current Opinion in Microbiology*, 24, pp.21–28. Available at: <http://dx.doi.org/10.1016/j.mib.2014.12.007>.
- Maini, P., 2004. The impact of Turing’s work on pattern formation in biology. *Mathematics Today*, August, pp.140–141. Available at: <http://eprints.maths.ox.ac.uk/563>.
- Mamei, M., Tolksdorf, R. & Zambonelli, F., 2006. Case studies for self-organization in computer science. *Journal of Systems Architecture*, 52(8–9), pp.443–460. Available at: <https://www.sciencedirect.com/science/article/pii/S1383762106000166>
- Mamei, M., Vasirani, M. & Zambonelli, F., 2004. Experiments of morphogenesis in swarms of simple mobile robots. *Applied Artificial Intelligence*, 18(9–10), pp.903–919. Available at: <http://www.tandfonline.com/doi/abs/10.1080/08839510490509081>
- Massey, R.C., Buckling, A. & French-Constant, R., 2004. Interference competition and parasite virulence. *Proceedings. Biological sciences / The Royal Society*, 271(1541), pp.785–8. Available at: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=1691666&tool=pmcentrez&rendertype=abstract>
- Mc Ginty, S.É. et al., 2013. The interplay between relatedness and horizontal gene transfer drives the evolution of plasmid-carried public goods. *Proceedings. Biological sciences*, 280(1761), p.20130400. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/23760639>
- Menon, R. & Korolev, K.S., 2015. Public Good Diffusion Limits Microbial Mutualism. *Physical Review Letters*, 114(16), p.168102. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/25955075>
- Merhej, V. et al., 2009. Massive comparative genomic analysis reveals convergent evolution of specialized bacteria. *Biology Direct*, 4(1), p.13. Available at: <http://biologydirect.biomedcentral.com/articles/10.1186/1745-6150-4-13>
- Mirams, G.R. et al., 2013. Chaste: An Open Source C++ Library for Computational Physiology and Biology A. Prlic, ed. *PLoS Computational Biology*, 9(3), p.e1002970.

- Available at: <http://dx.plos.org/10.1371/journal.pcbi.1002970>
- Mitri, S. & Foster, K.R., 2013. The genotypic view of social interactions in microbial communities. *Annual review of genetics*, 47, pp.247–73. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/24016192>
- Mitri, S., Xavier, J.B. & Foster, K.R., 2011. Social evolution in multispecies biofilms. *Proceedings of the National Academy of Sciences of the United States of America*, 108 Suppl, pp.10839–46. Available at: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3131810&tool=pmcentrez&rendertype=abstract>
- Molin, S. & Tolker-Nielsen, T., 2003. Gene transfer occurs with enhanced efficiency in biofilms and induces enhanced stabilisation of the biofilm structure. *Current Opinion in Biotechnology*, 14(3), pp.255–261. Available at: <http://linkinghub.elsevier.com/retrieve/pii/S0958166903000363>
- Momeni, B., Briley, K.A., et al., 2013. Strong inter-population cooperation leads to partner intermixing in microbial communities. *eLife*, 2013(2), pp.1–23.
- Momeni, B., Waite, A.J. & Shou, W., 2013. Spatial self-organization favors heterotypic cooperation over cheating. *eLife*, 2013(2), pp.1–18.
- De Monte, S. & Rainey, P.B., 2014. Nascent multicellular life and the emergence of individuality. *Journal of Biosciences*, 39(2), pp.237–248. Available at: <http://link.springer.com/10.1007/s12038-014-9420-5>
- Moreno-Fenoll, C. et al., 2017. Eco-evolutionary feedbacks can rescue cooperation in microbial populations. *Scientific Reports*, 7(January), pp.1–9. Available at: <http://dx.doi.org/10.1038/srep42561>.
- Müller, M.J.I. et al., 2014. Genetic drift opposes mutualism during spatial population expansion. *Proceedings of the National Academy of Sciences of the United States of America*, 111(3), pp.1037–42. Available at: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3903240&tool=pmcentrez&rendertype=abstract>.
- Munteanu, A. & Solé, R., 2008. Neutrality and robustness in evo-devo: emergence of lateral inhibition. *PLoS computational biology*, 4(11). Available at: <http://dx.plos.org/10.1371/journal.pcbi.1000226.g012>
- Muñoz-Dorado, J. et al., 2016. Myxobacteria: Moving, Killing, Feeding, and Surviving Together. *Frontiers in microbiology*, 7, p.781. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/27303375>
- Nadell, C.D. et al., 2008. The evolution of quorum sensing in bacterial biofilms. *PLoS biology*, 6(1), p.e14. Available at: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2214811&tool=pmcentrez&rendertype=abstract>

- Nadell, C.D., Drescher, K. & Foster, K.R., 2016. Spatial structure, cooperation and competition in biofilms. *Nature Reviews Microbiology*, 14(9), pp.589–600. Available at: <http://dx.doi.org/10.1038/nrmicro.2016.84>.
- Nadell, C.D., Foster, K.R. & Xavier, J.B., 2010. Emergence of spatial structure in cell groups and the evolution of cooperation. *PLoS computational biology*, 6(3), p.e1000716. Available at: <http://journals.plos.org/ploscompbiol/article?id=10.1371/journal.pcbi.1000716>
- Von Neumann, J. & Morgenstern, O., 2007. *Theory of games and economic behavior*, Princeton University Press. Available at: <https://press.princeton.edu/titles/7802.html>
- Niklas, K.J., 2014. The evolutionary-developmental origins of multicellularity. *American Journal of Botany*, 101(1), pp.6–25. Available at: <http://www.amjbot.org/cgi/doi/10.3732/ajb.1300314>.
- Nogueira, T. et al., 2009. Horizontal gene transfer of the secretome drives the evolution of bacterial cooperation and virulence. *Current biology : CB*, 19(20), pp.1683–91. Available at: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2773837&tool=pmcentrez&rendertype=abstract>
- Norman, A., Hansen, L.H. & Sørensen, S.J., 2009. Conjugative plasmids: vessels of the communal gene pool. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences*, 364(1527), pp.2275–89. Available at: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2873005&tool=pmcentrez&rendertype=abstract>
- Novick, R.P., 1987. Plasmid incompatibility. *Microbiological reviews*, 51(4), pp.381–95. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/3325793>
- Nowak, M.A., Tarnita, C.E. & Wilson, E.O., 2010. The evolution of eusociality. *Nature*, 466(7310), pp.1057–62. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/20740005>
- Nowak, M. a, 2006. Five rules for the evolution of cooperation. *Science*, 314(5805), pp.1560–1563.
- Núñez, I.N. et al., 2017. Artificial Symmetry-Breaking for Morphogenetic Engineering Bacterial Colonies. *ACS Synthetic Biology*, 6(2), pp.256–265.
- Okasha, S., 2011. Précis of Evolution and the Levels of Selection. *Philosophy and Phenomenological Research*, 82(1), pp.212–220. Available at: <http://doi.wiley.com/10.1111/j.1933-1592.2010.00470.x>
- Pascalie, J. et al., 2016. Developmental Design of Synthetic Bacterial Architectures by Morphogenetic Engineering. *ACS Synthetic Biology*, p.acssynbio.5b00246. Available at: <http://pubs.acs.org/doi/abs/10.1021/acssynbio.5b00246>.
- Payne, S. et al., 2013. Temporal control of self-organized pattern formation without morphogen gradients in bacteria. *Molecular systems biology*, 9(697), p.697. Available

- at:
<http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3817405&tool=pmcentrez&rendertype=abstract>
- Pipe, L.Z. & Grimson, M.J., 2008. Spatial-temporal modelling of bacterial colony growth on solid media. *Molecular BioSystems*, 4(3), p.192. Available at: <http://xlink.rsc.org/?DOI=b708241j>.
- Prajapat, M.K. et al., 2016. Analysis of a strategy for cooperating cells to survive the presence of cheaters. *Mol. BioSyst.*, 12(11), pp.3338–3346. Available at: <http://xlink.rsc.org/?DOI=C6MB00427J>.
- Prats, C. et al., 2006. Individual-based modelling of bacterial cultures to study the microscopic causes of the lag phase. *Journal of Theoretical Biology*, 241(4), pp.939–953.
- Prestes García, A. & Rodríguez-Patón, A., 2016. Sensitivity analysis of Repast computational ecology models with R/Repast. *Ecology and Evolution*, 6(24), pp.8811–8831. Available at: <http://doi.wiley.com/10.1002/ece3.2580>
- Price, G.R., 1970. Selection and Covariance. *Nature*, 227(5257), pp.520–521. Available at: <http://www.nature.com/doifinder/10.1038/227520a0>.
- Quaghebeur, W., 2017. Individual-Based Modelling Of Biodiversity In Microbial Communities. , pp.2016–2017. Available at: http://lib.ugent.be/fulltxt/RUG01/002/352/373/RUG01-002352373_2017_0001_AC.pdf.
- 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), pp.599–620. Available at: <http://www.annualreviews.org/doi/10.1146/annurev-ecolsys-120213-091740>.
- Rankin, D.J. et al., 2011. Bacterial cooperation controlled by mobile elements: kin selection and infectivity are part of the same process. *Heredity*, 107(3), pp.279–281. Available at: <http://www.nature.com/doifinder/10.1038/hdy.2011.59>
- Raspopovic, J. et al., 2014. Digit patterning is controlled by a Bmp-Sox9-Wnt Turing network modulated by morphogen gradients. *Science*, 345(6196), pp.566–570. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/25082703>
- Ratcliff, W.C. et al., 2012. Experimental evolution of multicellularity. *Proceedings of the National Academy of Sciences*, 109, pp.1595–1600.
- Rudge, J. et al., 2013. Cell polarity-driven instability generates self-organized, fractal patterning of cell layers. *ACS Synthetic Biology*, 2(12), pp.705–714.
- Rudge, T.J. et al., 2012. Computational modeling of synthetic microbial biofilms. *ACS synthetic biology*, 1(8), pp.345–52. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/23651288>.

- Rueffler, C., Hermisson, J. & Wagner, G.P., 2012. Evolution of functional specialization and division of labor. *Proceedings of the National Academy of Sciences*, 109(6), pp.E326–E335. Available at: <http://www.pnas.org/cgi/doi/10.1073/pnas.1110521109>
- Rutherford, S.T. & Bassler, B.L., 2012. Bacterial Quorum Sensing: Its Role in Virulence and Possibilities for Its Control. *Cold Spring Harbor Perspectives in Medicine*, 2(11), pp.a012427–a012427. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/23125205>
- Sanchez, A. & Gore, J., 2013. Feedback between Population and Evolutionary Dynamics Determines the Fate of Social Microbial Populations S. P. Ellner, ed. *PLoS Biology*, 11(4), p.e1001547. Available at: <http://dx.plos.org/10.1371/journal.pbio.1001547>
- Schaerli, Y. et al., 2014. A unified design space of synthetic stripe-forming networks. *Nature communications*, 5(May), p.4905. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/25247316>
- Schluter, J. et al., 2016. The Evolution of Quorum Sensing as a Mechanism to Infer Kinship J. C. W. Locke, ed. *PLOS Computational Biology*, 12(4), p.e1004848. Available at: <http://dx.plos.org/10.1371/journal.pcbi.1004848>
- Shapiro, J. a, 1998. Thinking about bacterial populations as multicellular organisms. *Annual review of microbiology*, 52, pp.81–104.
- Sheats, J. et al., 2017. Role of growth rate on the orientational alignment of *Escherichia coli* in a slit. *Royal Society Open Science*, 4(6), p.170463. Available at: <http://rsos.royalsocietypublishing.org/lookup/doi/10.1098/rsos.170463>
- Shigaki, K. et al., 2013. Effect of Initial Fraction of Cooperators on Cooperative Behavior in Evolutionary Prisoner's Dilemma Game Y. Moreno, ed. *PLoS ONE*, 8(11), p.e76942. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/24244270>
- Sick, S. et al., 2006. WNT and DKK Determine Hair Follicle Spacing Through a Reaction-Diffusion Mechanism. *Science*, 314(5804), pp.1447–1450. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/17082421>
- Smillie, C. et al., 2010. Mobility of plasmids. *Microbiology and molecular biology reviews : MMBR*, 74(3), pp.434–52. Available at: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2937521&tool=pmcentrez&rendertype=abstract>
- Smith, J.M., 1976. Group Selection. *The Quarterly Review of Biology*, 51(2), p.277. Available at: <http://www.journals.uchicago.edu/doi/10.1086/409311>
- Smith, J.M. & Price, G.R., 1973. The Logic of Animal Conflict. *Nature*, 246(5427), pp.15–18. Available at: <http://dx.doi.org/10.1038/246015a0>
- Smith, W.P.J. et al., 2017a. Cell morphology drives spatial patterning in microbial communities. *Proceedings of the National Academy of Sciences*, 114(3), pp.E280–E286. Available at: <http://www.pnas.org/lookup/doi/10.1073/pnas.1613007114>
- Smith, W.P.J. et al., 2017. Cell morphology drives spatial patterning in microbial

- communities. *Proceedings of the National Academy of Sciences of the United States of America*, 114(3), pp.E280–E286. Available at:
<http://www.ncbi.nlm.nih.gov/pubmed/28039436>
- Solé, R., 2016. The major synthetic evolutionary transitions. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences*, 371(1701). Available at:
<http://www.ncbi.nlm.nih.gov/pubmed/27431528>
- Song, H. et al., 2014. Synthetic microbial consortia: from systematic analysis to construction and applications. *Chemical Society reviews*. Available at:
<http://www.ncbi.nlm.nih.gov/pubmed/25017039>
- Stalder, T. & Top, E., 2016. Plasmid transfer in biofilms: a perspective on limitations and opportunities. *npj Biofilms and Microbiomes*, 2(August), p.16022. Available at:
<http://www.nature.com/doifinder/10.1038/bmt.2016.231>.
- Stewart, A.J., Parsons, T.L. & Plotkin, J.B., 2016. Evolutionary consequences of behavioral diversity. *Proceedings of the National Academy of Sciences*, pp.E7003–E7009. Available at: <http://arxiv.org/abs/1606.01401>.
- Strassmann, J.E., 2015. Kin Discrimination in Dictyostelium Social Amoebae. Available at:
https://strassmannandquellerlab.files.wordpress.com/2016/07/strassmann-2016-journal_of_eukaryotic_microbiology.pdf
- Su, P.-T. et al., 2012. Bacterial colony from two-dimensional division to three-dimensional development. *PloS one*, 7(11), p.e48098. Available at:
<http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3498271&tool=pmcentrez&rendertype=abstract>
- Szathmáry, E. & Smith, J.M., 1995. The major evolutionary transitions. *Nature*, 374(6519), pp.227–232. Available at: <http://www.nature.com/doifinder/10.1038/374227a0>
- Tarnita, C.E., Taubes, C.H. & Nowak, M. a, 2013. Evolutionary construction by staying together and coming together. *J Theor Biol*, 320(C), pp.10–22. Available at:
<http://dx.doi.org/10.1016/j.jtbi.2012.11.022>.
- Tarnita, C.E., Wage, N. & Nowak, M. a, 2011. Multiple strategies in structured populations. *Proceedings of the National Academy of Sciences of the United States of America*, 108(6), pp.2334–7. Available at:
<http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3038739&tool=pmcentrez&rendertype=abstract>
- Tayar, A.M. et al., 2017. Synchrony and pattern formation of coupled genetic oscillators on a chip of artificial cells. *Proceedings of the National Academy of Sciences of the United States of America*, 114(44), pp.11609–11614. Available at:
<http://www.ncbi.nlm.nih.gov/pubmed/29078346>
- Taylor, C. & Jefferson, D., 1993. Artificial Life as a Tool for Biological Inquiry. *Artificial Life*, 1(1_2), pp.1–13. Available at:
http://dx.doi.org/10.1162/artl.1993.1.1_2.1%5Cnhttp://www.mitpressjournals.org/doi/p

- df/10.1162/artl.1993.1.1_2.1%5Cnhttp://www.mitpressjournals.org/toc/artl/1/1_2.
- Tokita, R. et al., 2009. Pattern Formation of Bacterial Colonies by *Escherichia coli*. *Journal of the Physical Society of Japan*, 78(7), p.74005. Available at: <http://journals.jps.jp/doi/abs/10.1143/JPSJ.78.074005>.
- Touchon, M. et al., 2012. Antibiotic resistance plasmids spread among natural isolates of *Escherichia coli* in spite of CRISPR elements. *Microbiology (Reading, England)*, 158(Pt 12), pp.2997–3004. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/23059972>.
- Travisano, M. & Velicer, G.J., 2004. Strategies of microbial cheater control. *Trends in Microbiology*, 12(2), pp.72–78.
- Tsoi, R. et al., 2018. Metabolic division of labor in microbial systems. *Proceedings of the National Academy of Sciences of the United States of America*, p.201716888. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/29463749>
- Tudge, S., Watson, R. & Brede, M., 2013. Cooperation and the Division of Labour. In *Advances in Artificial Life, ECAL 2013*. MIT Press, pp. 1–8. Available at: <http://mitpress.mit.edu/sites/default/files/titles/content/ecal13/978-0-262-31709-2-ch001.pdf>
- Tudge, S.J., Watson, R.A. & Brede, M., 2016. Game theoretic treatments for the differentiation of functional roles in the transition to multicellularity. *Journal of Theoretical Biology*, 395, pp.161–173. Available at: <http://www.sciencedirect.com/science/article/pii/S0022519316000825?via%3Dihub>.
- Turing, A.M., 1952. The Chemical Basis of Morphogenesis. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 237(641), pp.37–72. Available at: <http://rstb.royalsocietypublishing.org/cgi/doi/10.1098/rstb.1952.0012>
- van der Vaart, E. et al., 2015. Calibration and evaluation of individual-based models using Approximate Bayesian Computation. *Ecological Modelling*, 312(2015), pp.182–190. Available at: <http://www.sciencedirect.com/science/article/pii/S0304380015002173>.
- Velicer, G.J. & Vos, M., 2009. Sociobiology of the Myxobacteria. *Annual Review of Microbiology*, 63(1), pp.599–623. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/19575567>
- Volfson, D. et al., 2008. Biomechanical ordering of dense cell populations. *Proceedings of the National Academy of Sciences of the United States of America*, 105(40), pp.15346–51. Available at: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2563119&tool=pmcentrez&rendertype=abstract>.
- Waters, C.M. & Bassler, B.L., 2005. Quorum sensing: cell-to-cell communication in bacteria. *Annual review of cell and developmental biology*, 21, pp.319–46. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/16212498>

- Watson, R.A. et al., 2016. Evolutionary Connectionism: Algorithmic Principles Underlying the Evolution of Biological Organisation in Evo-Devo, Evo-Eco and Evolutionary Transitions. *Evolutionary Biology*, 43(4), pp.553–581. Available at: <http://link.springer.com/10.1007/s11692-015-9358-z>
- West, S.A. et al., 2015. Major evolutionary transitions in individuality. *Proceedings of the National Academy of Sciences*, 112(33), pp.10112–10119.
- West, S.A. et al., 2006. Social evolution theory for microorganisms. , 4(August), pp.597–607.
- West, S.A., Diggle, S.P., et al., 2007. The Social Lives of Microbes. *Annual Review of Ecology, Evolution, and Systematics*, 38(1), pp.53–77. Available at: <http://www.annualreviews.org/doi/10.1146/annurev.ecolsys.38.091206.095740>.
- West, S.A. & Cooper, G.A., 2016. Division of labour in microorganisms: an evolutionary perspective. *Nat Rev Micro*, 14(11), pp.716–723. Available at: <http://dx.doi.org/10.1038/nrmicro.2016.111%5Cnhttp://10.1038/nrmicro.2016.111%5Cnhttp://www.nature.com/nrmicro/journal/v14/n11/abs/nrmicro.2016.111.html#supplementary-information>.
- West, S.A., Griffin, A.S. & Gardner, A., 2007. Evolutionary Explanations for Cooperation. *Current Biology*, 17(16), pp.661–672.
- West, S.A., Griffin, A.S. & Gardner, A., 2007. Social semantics: Altruism, cooperation, mutualism, strong reciprocity and group selection. *Journal of Evolutionary Biology*, 20(2), pp.415–432. Available at: <http://doi.wiley.com/10.1111/j.1420-9101.2006.01258.x>.
- Willensdorfer, M., 2009. On the evolution of bacterial multicellularity. *Evolution*, 63(2), pp.306–323. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/19154376>.
- Wilson, D.S., 1975. A theory of group selection. *Proceedings of the National Academy of Sciences of the United States of America*, 72(1), pp.143–6. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/1054490>.
- Wilson, D.S., 2007. Rethinking the Theoretical Foundation of Sociobiology. *The Quarterly Review of Biology*, 82(4), pp.327–348. Available at: <http://www.journals.uchicago.edu/doi/10.1086/522809> .
- Xavier, J.B., Martinez-Garcia, E. & Foster, K.R., 2009. Social evolution of spatial patterns in bacterial biofilms: when conflict drives disorder. *The American naturalist*, 174(1), pp.1–12. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/19456262>.
- Xu, S. & Van Dyken, J.D., 2017. Microbial expansion-collision dynamics promote cooperation and coexistence on surfaces. *Evolution*, pp.1–49. Available at: <http://doi.wiley.com/10.1111/evo.13393>.
- Yamagishi, J.F., Saito, N. & Kaneko, K., 2016. Symbiotic Cell Differentiation and Cooperative Growth in Multicellular Aggregates. *PLoS Computational Biology*,

12(10), pp.1–17.

Young, K.D., 2007. Bacterial morphology: why have different shapes? *Current Opinion in Microbiology*, 10(6), pp.596–600.

Zhang, C.-C. et al., 2006. Heterocyst differentiation and pattern formation in cyanobacteria: a chorus of signals. *Molecular Microbiology*, 59(2), pp.367–375. Available at: <http://doi.wiley.com/10.1111/j.1365-2958.2005.04979.x>

Zhong, X. et al., 2012. On the meaning and estimation of plasmid transfer rates for surface-associated and well-mixed bacterial populations. *Journal of theoretical biology*, 294, pp.144–52. Available at: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3346278&tool=pmcentrez&rendertype=abstract>

Zomorodi, A.R. & Segré, D., 2016. Synthetic Ecology of Microbes: Mathematical Models and Applications. *Journal of Molecular Biology*, 428(5), pp.837–861. Available at: <http://dx.doi.org/10.1016/j.jmb.2015.10.019>.