



# The coevolution of gene mobility and sociality in bacteria

Tatiana Dimitriu

## ► To cite this version:

Tatiana Dimitriu. The coevolution of gene mobility and sociality in bacteria. Agricultural sciences. Université René Descartes - Paris V, 2014. English. NNT : 2014PA05T005 . tel-00993436

**HAL Id: tel-00993436**

**<https://theses.hal.science/tel-00993436>**

Submitted on 20 May 2014

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

UNIVERSITÉ PARIS DESCARTES  
Ecole Doctorale Interdisciplinaire Européenne Frontières du Vivant

## THESE

pour obtenir le grade de

DOCTEUR

Discipline : Biologie Evolutive

Présentée et soutenue publiquement par

**Tatiana DIMITRIU**

le mercredi 9 avril 2014

# The coevolution of gene mobility and sociality in bacteria

JURY :

Dr. François TADDEI	Université Paris-Descartes	Directeur de thèse
Dr. Ariel LINDNER	Université Paris-Descartes	Codirecteur de thèse
Pr. Bernard GODELLE	Université Montpellier II	Rapporteur
Dr. Ashleigh GRIFFIN	University of Oxford	Rapporteur
Pr. Erick DENAMUR	Université Paris-Diderot	Examineur
Pr. Didier MAZEL	Institut Pasteur	Examineur



# Acknowledgments / Remerciements

First, I would like to thank all the members of my jury, that kindly accepted to give their time to evaluate this work: Bernard Godelle and Ashleigh Griffin, for coming from far away and reading this manuscript, Erick Denamur and Didier Mazel, for being part of this jury. Thank to all of you for your comments and suggestions.

Next, thanks to my supervisors, François and Ariel, that allowed me to do and finish this PhD in the lab with as much freedom as possible. Merci François pour toutes tes idées, intuitions et discussions, thanks to both for advice, thanks Ariel also for advice about the next steps..

Thanks Julien and Dule for the more day-to-day advice, suggestions and science. Julien pour avoir fait découvrir la microbio, et toute sa biodiversité cachée à la petite étudiante de M2, et nos nombreuses discussions par la suite. Dule for sharing your expertise, and being always there to help, plus the teaching hours with M1s that were a welcome change in scientific subjects. And thanks so much for the so detailed and intensive reading and correction of this manuscript in a such short time.

Many thanks to Sam Brown and Didier Mazel, my two PhD tutors, who both saved these years of work in different ways. Sam for giving some structure and coherence to my novice modelling work, and welcoming me so kindly in Edinburgh to work on the general fitness model. Didier without whom the experimental system would have had no chance to work and give nice results, made me realize the far too low transfer, the reason for it and solution to it. Thanks for your help, advice and support over the years.

Thanks to all present and past colleagues in the lab for the discussions, time together, tips and advice. Merci Chantal pour toute ton aide dans les manip et ton soutien; Kevin et Morgane, mes deux stagiaires pleins d'énergie, dont chacun pourrait reconnaître un graphe dans ces pages.. Merci à tous mes amis et collègues de FdV pour m'avoir bien souvent changé les idées le vendredi soir; aux deux Antoine pour les râleries, les pauses communes, et l'aide côté paillasse pour l'un, débogage de mes erreurs de débutante avec matlab et de mes vagues souvenirs de maths pour l'autre..

Merci enfin à ma famille et mes amis, tous ceux qui ont supporté des périodes d'absorption dans des sujets ésotériques que j'avais bien du mal à expliquer clairement ;-). A tous ceux et celles embarqués dans la même galère de manière synchrone: Marianne, Anne-So, Céline.. et pour les soirées jeux qui ont ponctué tout ça. A Titi pour les conseils vocabulaires et bavardages sur mes élucubrations;

Enfin merci coloc pour les si nombreux craquages, coups de blues, questions existentielles, remotivations et bavardages communs des midis, soirs ou week-ends,

Et Guillaume pour tout le reste, son scepticisme récurrent, son oeil de lynx, et pour avoir supporté presque patiemment et exclusivement la plupart des inévitables récriminations qui n'avaient pas leur place ici.

# Foreword

This work focuses on the interplay between two important bacterial traits: cooperation and horizontal gene transfer. Bacterial interactions can be analyzed in terms of social evolution, which has yielded insight into why bacteria present seemingly paradoxical behaviours such as the production of public goods, as well as increased the general understanding of the evolution of cooperation. Horizontal gene transfer is another widespread phenomenon in bacteria, mainly associated with the existence of mobile genetic elements such as plasmids that can move among potentially unrelated bacteria. The mechanisms, ecology and evolutionary consequences of gene mobility have been widely studied, but mobility itself has rarely been studied in the light of social evolution, even though it can alter the behaviour and fitness of interacting cells

Here, I analyze gene mobility while taking into account social interactions, and identify the selective pressures acting on horizontal gene transfer. I investigate the coevolution of gene mobility and sociality in bacteria, showing that both traits strongly influence each other. The coevolution involves not only these two traits, but also at least two types of genetic units: bacterial host chromosomes and mobile elements themselves.

In the introduction, I first present the theory of social evolution, focusing on its application to bacteria (Chapter I). I then describe the mechanisms and ecology of horizontal gene transfer (Chapter II), and finally focus on the coevolution of mobile genetic elements with their hosts (Chapter III).

My work involved both mathematical modelling and experimental work to test the hypotheses generated by modelling. The experimental part required the construction of a synthetic system that I describe first (Chapter 1). I then focus on the consequences of horizontal gene transfer on the maintenance of cooperation (Chapter 2) and on the selection of transfer ability from the host side (Chapter 3), with both chapters presenting the modelling work and experimental tests in parallel. Finally, I address with simulations the interactions that can occur when both transfer and cooperation coevolve (Chapter 4).

# Table of Contents

<b>Foreword</b>	<b>1</b>
<b>Contents</b>	<b>2</b>
<b>List of abbreviations</b>	<b>5</b>
<b>Introduction</b>	<b>6</b>
<b>I Social evolution: between cooperation and conflict</b>	<b>7</b>
I.1 The evolution of social behaviours . . . . .	8
I.2 Conflicts between levels of selection . . . . .	16
I.3 Horizontal transmission and its consequences . . . . .	23
<b>II Mobile genetic elements as bacterial symbionts</b>	<b>29</b>
II.1 Mechanisms of horizontal gene transfer in bacteria . . . . .	29
II.2 Vertical transmission of plasmids . . . . .	35
II.3 Horizontal transmission of plasmids . . . . .	40
<b>III Plasmid-host coevolution</b>	<b>53</b>
III.1 Evolution of plasmids and transmission pathways . . . . .	53
III.2 Plasmids and the transfer of genetic information . . . . .	59
<b>IV PhD project</b>	<b>69</b>
IV.1 Questions . . . . .	69
IV.2 Strategy . . . . .	71
<b>Results</b>	<b>73</b>
<b>Note about the structure of the results</b>	<b>73</b>
<b>1 A synthetic system for plasmid transfer and public good production</b>	<b>75</b>
1.1 Public good production and benefits . . . . .	75
1.2 Plasmid transfer . . . . .	79
1.3 Fluorescent markers for strains and plasmids . . . . .	83
1.4 Combining transfer and cooperative behaviour . . . . .	87
1.5 Final set-up and discussion . . . . .	91

<b>2</b>	<b>Genetic information transfer promotes cooperation</b>	<b>95</b>
2.1	Introduction . . . . .	95
2.2	Design of the study . . . . .	96
2.3	Results . . . . .	98
2.4	Discussion . . . . .	104
2.5	Materials and methods . . . . .	106
<b>3</b>	<b>Indirect selection on chromosomal genes for horizontal transfer</b>	<b>113</b>
3.1	Introduction . . . . .	113
3.2	Selection of chromosomal control of transfer: direct fitness approach . . . .	115
3.3	Selection for transfer in a simple metapopulation . . . . .	122
3.4	Influence of the availability of plasmid-free cells . . . . .	132
3.5	Population structure requirements . . . . .	138
3.6	Discussion . . . . .	142
3.7	Materials and methods . . . . .	147
<b>4</b>	<b>Co-selection of transfer and public good genes</b>	<b>155</b>
4.1	Introduction . . . . .	155
4.2	Co-selection of plasmid transfer and public good production genes . . . .	156
4.3	Co-selection of chromosomal transfer and plasmid cooperation genes . . . .	165
4.4	Discussion . . . . .	174
4.5	Methods . . . . .	178
	<b>General discussion</b>	<b>181</b>
<b>5</b>	<b>General discussion</b>	<b>181</b>
5.1	Conclusions from models and experiments . . . . .	181
5.2	Benefits and drawbacks of the approaches used . . . . .	183
5.3	Possible selection of horizontal transfer in nature . . . . .	188
5.4	Coevolution between plasmids and chromosomes . . . . .	193
5.5	Perspectives . . . . .	196
	<b>Annexes</b>	<b>199</b>
<b>A</b>	<b>Strains and plasmids</b>	<b>199</b>
<b>B</b>	<b>Identification of strains, plasmids and transfer</b>	<b>203</b>
	<b>Bibliography</b>	<b>207</b>
	<b>Abstract</b>	<b>226</b>





# List of abbreviations

BRH: broad host range  
Cm: chloramphenicol  
DNA: deoxyribonucleic acid  
GFP: green fluorescent protein  
GTA: gene transfer agents  
Hfr: high frequency of recombination  
HSL: homoserine lactone  
ICE: integrative and conjugative elements  
Kb: kilo-base pairs  
Kn: kanamycin  
LB: Luria-Bertani broth  
LD: linkage disequilibrium  
Mb: mega-base pairs  
MGE: mobile genetic elements  
MPF: mating pair formation  
NRH: narrow host range  
OD: optical density  
PSK: post segregation killing  
RFP: red fluorescent protein  
RM: restriction-modification  
RNA: ribonucleic acid  
Sp: spectinomycin  
T4CP: type IV coupling protein  
T4SS: type IV secretion system  
Tc: tetracycline  
YFP: yellow fluorescent protein



# Introduction



# Chapter I

## Social evolution: between cooperation and conflict

Evolution by natural selection is based on differential reproduction: biological traits that maximize fitness (the average number of offspring) of entities expressing them will be positively selected. The theory of evolution by natural selection was first introduced by Darwin, and focused primarily on individual organisms [Darwin, 1859]: individuals were thought to maximize their individual fitness in a given environment. This focus on individual fitness suggested that organisms would interact mainly by competition, as embodied in the expression "struggle for existence", and selfish behaviours would be selected. However, organisms often appear to be altruistic, performing costly behaviours that help others, with a decrease in their own individual fitness. One of the most extreme examples is the one of social insects, where sterile workers do not reproduce at all themselves : their individual fitness is zero [Queller and Strassmann, 1998]. Behaviours where interactions between organisms are not pure competition thus need another explanation, which can be provided by social evolution theory.

### Classification of social behaviours

Social evolution is concerned with social behaviours, *ie* behaviours that have fitness consequences on individuals other than the actor (called recipients). Social behaviours have been partitioned into four classes by Hamilton [Hamilton, 1964b, 1970], according to the direct fitness consequences they have for both the actor and the recipient:

Behaviours that provide a direct benefit to actors (positive effect) 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 (negative effect) are separated into altruism, when recipients benefit, and spite, when recipients experience a cost. Cooperation can be defined as any behaviour which provides a benefit to a recipient, and which is selected in actors because of this benefit in recipients (excluding one-way byproduct benefits) [West et al., 2007c]. This broader category includes mutual benefits and altruism, which differ by the effects they have on the actor. Inter-species cooperation is usually called mutualism.

The evolution of both altruism and spite is difficult to explain by natural selection, as those behaviours reduce the relative fitness of individuals performing them, even if they are beneficial for their neighbours or at the scale of a group. Social evolution is thus mainly concerned with explaining how and under which conditions altruistic or spiteful behaviours can occur instead of selfish, competitive behaviours. In this chapter I will address different aspects of social evolution in more detail, using bacterial examples:

Firstly, I will describe generally social behaviours among organisms, as well as the main theory that aims at explaining them, the inclusive fitness theory. Secondly, I will focus on another approach, which concentrates on the occurrence of multiple levels of selection and is appropriate when selection acts in opposing ways among the levels, as between hosts and symbionts. Finally, I will examine the horizontal transmission of symbionts, which is particularly relevant to host-symbiont coevolution towards parasitism or mutualism.

## **I.1 The evolution of social behaviours**

### **I.1.1 The diversity of social behaviours**

Many behaviours can have a social component, affecting organisms other than the one performing them. Social behaviours include the production of public goods, and behaviours that actively target other organisms, either by helping or by harming them. I focus here on behaviours which seem difficult to explain by natural selection: altruism and spite.

#### **Public good production**

Public goods are goods that are produced by given organisms, but accessible to other organisms as well, as opposed to private goods. When these public goods are beneficial but their production or preservation is costly, producing or preserving them is a social, cooperative behaviour which can be exploited by cheaters, individuals that benefit from the behaviour without contributing to it.

Examples of public goods in human societies are environmental goods such as air quality, or human productions such as public libraries. Bacteria also produce a wide range of compounds that are secreted in their common environment and can act as public goods [Crespi, 2001, West et al., 2007c]. Examples of public goods include molecules involved in nutrient acquisition like siderophores which are iron chelators [Griffin et al., 2004], in protection against antibiotics like  $\beta$ -lactamases [Dugatkin et al., 2005] or in competition like toxins or antibiotics [Chao and Levin, 1981]. Structural components like biofilm polysaccharides or biosurfactants can also be considered as public goods [Absalon et al.,

2011, Xavier et al., 2011]. Finally, quorum-sensing is a form of bacterial communication that involves the production of diffusible molecules, auto-inducers, that above a threshold concentration induce various behaviours in recipients. For instance, *Pseudomonas* strains produce different auto-inducers in the N-acyl-homoserine lactone (HSL) family, which control the production of multiple virulence factors (Figure I.1). In the case of quorum-sensing, cooperation can happen at two stages: production of the auto-inducer and response to high auto-inducer concentration by the secretion of other public goods [Diggle et al., 2007b, Sandoz et al., 2007].

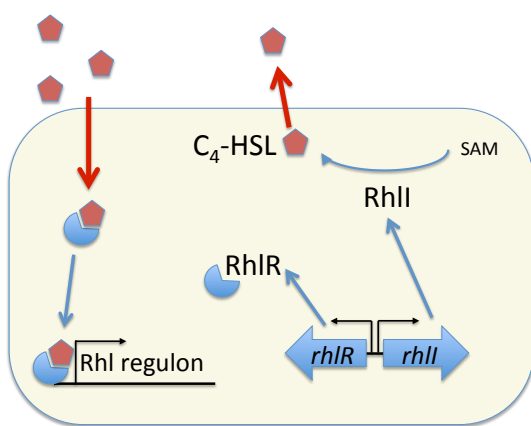


Figure I.1: *Pseudomonas aeruginosa* Rhl quorum-sensing system. The enzyme RhlI produces the autoinducer C<sub>4</sub>-HSL (in red) from SAM (S-adenosylmethionine). C<sub>4</sub>-HSL then diffuses in the extracellular environment and in neighbouring cells (red arrows). Above a threshold concentration, C<sub>4</sub>-HSL binds to the RhlR receptor and induces expression of the Rhl regulon, that includes public good genes like genes for the production of rhamnolipids, surfactants responsible for swarming motility.

Public good production is often linked to pathogen virulence. Virulence is defined as the harm a pathogen causes to its host and it often involves toxin production, or the secretion of molecules into host cells. These molecules can be produced by a few cells only, but have global consequences on the host, behaving as pathogen public goods. For instance, *Salmonella typhimurium* injects virulence factors in the gut tissue of mice to induce global inflammation, which then benefits all individual *Salmonella typhimurium* cells by affecting competitor species [Ackermann et al., 2008].

### Division of labor

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. For instance, sterile social insects only help other individuals to reproduce, and never have offspring of their own. Division of labor also happens in some microorganisms where multicellularity is facultative, as in the cooperative building of fruiting bodies in *Dictyostelium* or *Myxococcus* species. Fruiting bodies allow for more efficient spore dispersal, but also require some cells to die for them to be constructed



[Strassmann et al., 2000, Velicer et al., 2000].

### **Spiteful behaviours**

Spite is defined as an individual harming itself (decreasing its lifetime fitness) in order to harm another. One widespread example of spite in microbes is the production of bacteriocines, proteic toxins that are produced by bacteria in stressful conditions, and kill neighbouring bacteria that are not resistant [Riley and Gordon, 1999].

Bacteriocin production is costly, and bacteriocin release often happens through cell lysis, causing death of the producing cell. Clones of the producing cell, however, are not killed, as the bacteriocin and lysis genes are tightly linked to a gene providing immunity to the bacteriocin. Still, spite presents here the same paradox as altruism: mutants resistant to the bacteriocin, that do not pay the cost of producing and releasing it, will get a direct fitness advantage over bacteriocin producers [Levin et al., 1988].

### **I.1.2 Evolutionary explanations for social behaviours**

Darwin already suggested that altruism in social insects may be explained by effects on kin [Darwin, 1859]. This idea was then taken up by other scientists like Haldane, who famously said that he would give his life "to save two brothers or eight cousins", meaning that as siblings are on average 50% identical by descent, reproductive success through these relatives is equivalent to direct reproductive success through one's own offspring [Haldane, 1932]. However, early formal theories about natural selection mainly neglected interactions between relatives, until Hamilton's contribution [Hamilton, 1964a].

#### **Hamilton's rule**

Hamilton emphasized that the fitness of an individual is the sum of the direct effect of its own genotype and of effects due to its neighbours (neighbour-modulated fitness). When neighbours are related, a gene present in both the actor and the recipient can thus have an indirect effect on the actor via its expression in recipients. Reversely, a given behaviour will influence both the direct fitness of the organism performing it, but also the fitness of neighbours. Its effect on fitness will then be the sum of direct fitness effects, due to offspring of the actor, and indirect fitness effects, due to offspring of related neighbours: this sum is called inclusive fitness.

Effects on indirect fitness are weighted by the genetic relatedness between the actor and the recipient of the behaviour, as not all neighbours have the same probability of bearing the actor's genes responsible for the behaviour. This is summarized in Hamilton's rule,

which states that a behaviour is favored when:

$$b \times r - c > 0$$

with  $c$  being the direct fitness cost for the actor,  $b$  being the fitness benefit to recipients, and  $r$  the genetic relatedness between actors and recipients.  $c$  is thus the impact of a behaviour on direct fitness, and  $b \times r$  its impact on indirect fitness.

Hamilton's rule can be derived rigorously by computing either neighbour-modulated or inclusive fitness [Frank, 1998]. Relatedness then appears as the regression coefficient of the recipient's genotype to the actor's genotype. Relatedness is a relative coefficient: it describes the genetic similarity between actors and recipients, relative to the global similarity of individuals in the population.

Hamilton's rule framework also leads to a classification of evolutionary explanations for cooperation into direct and indirect benefits of a behaviour [West et al., 2007b]. Direct benefits correspond to mutually beneficial behaviours, indirect benefits to altruistic behaviours.

### **Direct benefits of cooperation**

Cooperation might first be beneficial simply because of by-product benefits [Sachs et al., 2004]. This is the case with cross-feeding, when members of two species each feed on by-products of the other. Cooperation can also be beneficial because of feedbacks from recipients, if actors and recipients interact long enough that actors benefit from those feedbacks (partner fidelity feedback, [Sachs et al., 2004]). This mechanism does not necessarily involve kin and can thus happen in inter-species mutualism. The sustained interaction can happen during the life of the actor, or across generations, as in the case of vertical transmission of microbial symbionts (I will focus on symbiont transmission and its consequences in section I.3.1).

Secondly, enforcement mechanisms can maintain cooperation, for instance by making defection from cooperation costly. One enforcement strategy is reciprocity: individuals help preferentially other individuals that helped them in the past [Trivers, 1971]. Another enforcement strategy is the active policing of non-cooperative individuals. For instance, in the legume-rhizobium mutualism soybeans have been shown to penalize rhizobia nodules that do not engage in  $N_2$  fixation, decreasing their reproductive success [Kiers et al., 2003].

## Indirect benefits of cooperation

Indirect benefits of cooperation arise from actors and recipients sharing the alleles responsible for the cooperative behaviour. Shared alleles usually happen between kin, arising from a common ancestor. Because of this, selection because of shared alleles between actors and recipients has been termed kin selection by Maynard Smith [Foster, 2010]. Kin selection can thus be defined as the process by which traits are favoured because of their beneficial effects on the fitness of relatives [West et al., 2007c] (however, in specific cases I will discuss later, shared alleles that lead to indirect fitness benefits do not arise through common ancestry). Hamilton defined two mechanisms that can lead to high genetic relatedness: limited dispersal and kin discrimination [Hamilton, 1964b].

Limited dispersal of individuals leads to the proximity of relatives, so that individuals are likely to interact with relatives purely by chance. Limited dispersal is thought to be important for the evolution of public good production in microorganisms: public goods are released in the medium at a scale where neighbours are likely to be relatives [West et al., 2006]. However, the competition between relatives also has to be taken into account, as it also increases with limited dispersal, and decreases the benefits of cooperation. The relative scale of cooperation and competition will thus determine the effect of limited dispersal: cooperation can still be favored if the scale of competition is larger than the scale of cooperation [Griffin et al., 2004].

Kin discrimination or kin recognition refers to mechanisms that allow cooperators to distinguish relatives from non-relatives, and direct cooperation only toward relatives [West et al., 2007b]. Cooperation can also be restricted to relatives if some mechanism ensures that only relatives can benefit from it, for instance when unrelated individuals cannot benefit from specific public good molecules. This happens in the case of the production of opines, metabolites that can be used only by cells that carry opine catabolism genes, in *Agrobacterium tumefaciens* [Platt et al., 2012]. Cooperative bacteria induce the host plant to produce opines by transferring the T-DNA fragment of their cooperative plasmid to plant cells. The plant then produces opines that diffuse towards bacteria. However, *Agrobacterium* cells without the cooperative plasmid do not bear genes for opine catabolism, and thus do not benefit from it.

The most extreme specificity is observed in greenbeards: genes that specifically direct cooperation towards other individuals that carry the same gene (such as one gene causing formation of a green beard and simultaneously directing cooperation to individuals that

have a green beard [Dawkins, 1976]). Greenbeards are probably rare, but are known in microorganisms, and are often genes coding for cell adhesion proteins, like the *csa* gene in the social amoeba *Dictyostelium discoideum* [Queller, 2003] or the FLO1 gene in the budding yeast *Saccharomyces cerevisiae* that leads to yeast flocculation [Smukalla et al., 2008]. The concept of greenbeards underlines a fundamental feature of inclusive fitness: it considers genetic relatedness at a specific locus, and not a whole-genome genealogical relationship. Those will generally be similar, as usually genes will be shared among organisms because they are genealogical relatives [Gardner and West, 2010], but not always, especially in microbes, as we will see in section I.1.3.

### **The selection of spiteful behaviours**

Spiteful behaviours are defined as having benefits  $b < 0$  for recipients, and costs  $c > 0$  for actors. Following Hamilton's rule, they can be selected when  $b \times r - c > 0$ : relatedness therefore has to be negative. This is possible, as relatedness is a relative coefficient. If a focal individual is more related to some fraction of the population, its relatedness to the rest of the population will be negative [Hamilton, 1970]. For spiteful behaviours to be selected, some kind of discrimination has to take place between positively and negatively related individuals in the population. In the absence of active mechanisms for kin discrimination such as greenbeards, positive relatedness to self can be sufficient, but will be significant only in very small populations where the focal individual is a significant part of the population [Gardner and West, 2004]. In other cases, active discrimination must take place. In the case of bacteriocin production, for instance, the genes for bacteriocin production and immunity to its effect are strongly linked: only individuals that carry the antidote gene associated to the toxin will survive [Riley and Wertz, 2002]. This leads to effective kin discrimination, as kin are not killed by the toxin. Recipients of the behaviour are here non-kin, with negative relatedness to the actor.

### **I.1.3 Social evolution theory in microbes**

It is not surprising that microbes have been frequently used to study social evolution [West et al., 2006, 2007c]. They are particularly social organisms and engage in multiple cooperation and communication behaviours. For instance, 6-10% of *Pseudomonas aeruginosa* genome has been found to be under quorum-sensing control [Schuster et al., 2003]. The reason for such extensive sociality may be that a single bacterial cell generally will not be able to change its environment significantly enough to get benefits from it. Social evolu-

tion theory thus allows us to better understand multiple microbial behaviours (including some with consequences to human health), while microbes simultaneously allow testing general hypotheses of social evolution theory. Finally, they also possess unique properties, most prominently the influence of microbial population structure and horizontal gene transfer on relatedness.

### **Microbes as a testing ground of social evolution theory**

Microbes have been widely used to test social evolution theory, as they are amenable to extensive manipulation. The social characteristics of a given behaviour can be tested, often by using specific mutant strains that do not perform this behaviour. Then, conditions where cooperation is maintained can be explored.

One of the most studied cooperative traits in microbes is the production of siderophores. Experiments have confirmed the importance of relatedness and of the scale of competition [Griffin et al., 2004]; and highlighted additional factors like the regulation of public good production [Kümmerli and Brown, 2010]. Quorum sensing has also been extensively analyzed, demonstrating that both induction and response to quorum sensing can be cooperative [Diggle et al., 2007b].

Finally, synthetically engineered systems can be used in bacteria to decrease unintended interactions with other cell processes [Tanouchi et al., 2012]. For instance, a synthetic system for public good production was constructed by Chuang et al. [2009]. A quorum-sensing signal (HSL) originally present in *Pseudomonas aeruginosa* is introduced into and expressed in engineered *Escherichia coli* cells. Those cells have been further modified to respond to HSL by inducing the expression of antibiotic resistance. As *Escherichia coli* does not naturally produce HSL molecules, the cooperation system should interact minimally with other cell components and enable precise study of parameters governing the selection of cooperation. I will focus on this system later, as we use it for our experiments.

### **Origins of population structure in microbes**

Because of microbial specificities, relatedness will often arise through mechanisms that differ from what happens in social animals, which have long been the main focus of social evolution theory. Kin discrimination based on genetic or environmental cues is thought to be rare in microbes, as mechanisms for kin recognition often require cognitive abilities [Strassmann et al., 2011]. Discrimination will more often be due to specificity, with specific public good molecules that only the producing strain can use, or greenbeard mechanisms.

Except for this, population structure in microbes will mostly arise through limited dispersal. Limited dispersal is extreme in the case of clonal growth of non-motile individuals, leading 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 [Hallatschek et al., 2007]. Simulations of biofilm formation have shown that this can favor cooperative phenotypes [Nadell et al., 2010]. It was also shown experimentally in yeast that clonal growth favors cells that cooperatively produce the exoenzyme invertase (providing extracellular sucrose digestion), as successful sectors of producer cells arise. In less extreme cases than growth on solid surfaces, medium viscosity can impact both individual motility and public good diffusion, favoring cooperation [Kümmerli et al., 2009a]. In the absence of viscosity, a metapopulation composed of multiple patches will also be structured if subpopulations differ in genotype frequencies: this is besides a way to study cooperation dynamics experimentally, by culturing bacteria in patch-structured metapopulations, where inclusive fitness and multilevel selection parameters can be easily estimated [Chuang et al., 2009, Dumas and Kümmerli, 2012, Datta et al., 2013]. Moreover, differences in frequencies can arise naturally in the absence of complex mechanisms, simply through stochastic variations in founder populations [Chuang et al., 2009].

### **Origin of gene and relatedness changes in bacteria**

Relatedness in bacteria arises mainly through limited dispersal, leading to spatial structuring of populations. However, a striking particularity of bacteria is the widespread horizontal gene transfer that I will describe in Chapter II: genetic changes arise not only by mutation, but also by gene transfer from other organisms, via mobile genetic elements on which natural selection also acts [Levin and Bergstrom, 2000]. This further implies that different genes do not share the same relatives, and relatedness can differ among genes in the same organism [West et al., 2007c]. Particularly, genes responsible for social behaviours can spread in a population [Smith, 2001], modifying relatedness and indirect fitness effects at the social locus [Nogueira et al., 2009]. The consequences of these relatedness changes are not obvious, and make us wonder about the selective pressures acting on mobile but also non-mobile genome loci affecting social behaviours. Part of my work will focus on better understanding and quantifying the importance of relatedness in the social evolution of microbes.

We have analyzed here situations where all organisms that engage in cooperative or

competitive behaviours were of the same kind. However, interactions can happen between entities present at different levels: this will happen in symbioses where one partner hosts the other, like intracellular parasites or mutualists (endosymbionts), or when different individuals cooperate in a way that leads to a higher-level entity.

## **I.2 Conflicts between levels of selection**

### **I.2.1 Levels of selection: definition and examples**

#### **Definition**

Lewontin pointed out that natural selection is a very general mechanism, which applies to any entity that has the properties of variation, reproduction and heritability [Lewontin, 1970]. Entities under selection must present phenotypic variation, with different phenotypes leading to differences in fitness. Finally, the characters of parents and offspring must be correlated (heritability). With these properties natural selection will act, leading to the selection of fittest entities and evolutionary change.

Thus, natural selection does not apply only to the traditional level of individual organisms, but can also apply to within-individual entities, like genes, or higher entities, like populations. Levels of selection are levels of the biological hierarchy at which natural selection acts [Okasha, 2010].

#### **Selection within cells**

Variants of a molecule can be characterized by differential reproduction and heritable characters within individuals. The most common cases are DNA sequences that differ in fitness through diverse mechanisms leading to differential reproduction, as DNA has obviously both properties of variation (through mutations) and heritability (through replication). DNA sequences that multiply more effectively than others accumulate within cells, a process called molecular drive. Multiplication can be due to sequences leading to biased replication, like repeated sequences, or to specific functions like the transposition mechanisms of transposable elements. Sequences with no known functions other than self-replication have been called selfish DNA [Orgel and Crick, 1980].

Selection also happens among other molecules that can vary in copy number within the cell. This is the case with bacterial plasmids, which are DNA molecules that are independent from the chromosome, and are present in the cytoplasm in varying numbers. Variants with higher copy number, for instance with an altered control of replication

[Paulsson, 2002], will have a higher chance of being transmitted to daughter cells during cell division, leading to intra-cellular selection on plasmid copy number. Similarly, eucaryotic organelles can also proliferate more or less within cells depending on their genotype. For instance, mitochondria with genome deletions can be selected for in *Saccharomyces cerevisiae* cells [Taylor et al., 2002].

### **Selection within multicellular organisms**

For multicellular organisms, selection can also happen at the cellular level. Cancers are a clear example of cellular selection, with cell variants that differ in their reproductive success, leading to proliferation within an organism [Merlo et al., 2006]. Cellular selection also takes place in regular development, as with the acquired immunity of vertebrates, where lymphocytes with receptors that recognize specific antigens will either be activated, leading to multiplication, or killed.

### **Selection among groups**

Finally, selection can also happen among populations or groups of organisms. The topic of group selection has been widely discussed, the debate focusing partly on the existence and definition of group selection and partly on the possibility for group selection to be significant compared to within-group selection. For instance division into isolated groups and variation among groups has been seen as essential for group selection, which represents quite restrictive conditions in nature [Maynard Smith, 1976]: group selection will be favored over within-group selection only if migration between groups is extremely low, maintaining differences between groups [Leigh Jr, 2010b]. Thus, group selection has been thought as unlikely to be significant as a major evolutionary force [Maynard Smith, 1976].

Another less restrictive type of group selection was termed "trait-group selection" or "new group selection" [West et al., 2006, Wilson, 1975]. This theory considers cases where interactions take place in groups only at specific stages of the life cycle of individuals (for instance non-dispersal stages), and selection can act on those "trait-groups". This version of group selection can actually be mathematically reconciled with a kin selection approach [West et al., 2006, Queller, 1992].

What can be seen as a clear example of group selection [Lewontin, 1970] is the coevolution of rabbits and myxoma, viruses that cause myxomatosis, after introduction of the virus in Australia. Viruses were first strongly lethal, but became less virulent after coevolution. Those viruses could not be transmitted from dead rabbits, leading to very strong between-group selection for decreased virulence. In this case, selection at the group level



is clearly visible, because groups are easily distinguished (each of them within a host).

## **I.2.2 Diverging interests between levels of selection**

When selection happens at different levels, the direction of selection can differ among levels, leading to conflicts. Selection acting within organisms can harm the organism; selection between organisms within groups can lead to costs at the group level.

### **Conflicts within organisms**

Molecules that proliferate within cells, or cells that do so within multicellular organisms, often lead to costs at the level of the organism. Uncontrolled cell division causes cancers; accumulation of repeated DNA sequences can give rise to diseases, like Huntington's disease, a neurodegenerative disorder arising from the accumulation of repeats in the sequence of a protein-coding gene [Rubinsztein et al., 1994]. Transposition can cause chromosome breakage and reduced fertility, and insertions are often lethal [Burt and Trivers, 2006]. Increases in plasmid copy number are also costly for the cell [Harrison et al., 2012], leading to a conflict between plasmids and cells [Paulsson, 2002]: plasmids variants with increased replication have a higher chance of being present in daughter cells, so are selected within cells, but cells containing these plasmids are counter-selected.

### **The tragedy of the commons**

Diverging selective pressures at the individual and group levels is another way of analyzing cooperation: 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 herders grazing cattle on a common land, each of them benefiting from adding more animals. However, this behaviour leads to over-exploitation 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 [Rankin et al., 2007]. Examples of individual selfishness can be the overexploitation of external resources, but also the act of not contributing to a resource, like not producing a public good, or engaging in active competition (instead of restraining from it).

### **Evolution of virulence**

The case of host-parasite coevolution is itself a case of tragedy of the commons: each individual parasite has a short-term interest in maximum exploitation of the host, leading

to competition among pathogens within hosts. Overexploitation of host resources can lead to death of the host, and collapse of the population of parasites. Prudent exploitation is then a form of cooperation, and increased relatedness should thus lead to decreased virulence [Frank, 1996]. For instance, restricted migration (leading to high relatedness) was shown to select for prudent host exploitation in an experimental evolution study between bacteriophages and their hosts [Kerr et al., 2006].

However, real cases are often more complex, as virulence itself can be due to cooperative traits [Brown, 1999]: numerous virulence factors are public goods. For instance, the cooperative production of siderophores is involved in *Pseudomonas aeruginosa* virulence in an insect host [Harrison et al., 2006]. So, high relatedness can also favor increased virulence, and the effects of social conflicts on virulence will depend on the type of social interactions involved in virulence [Brown et al., 2002, Buckling and Brockhurst, 2008].

### I.2.3 The outcome of conflicts: multilevel selection

The outcome of conflicts between levels of selection can be studied by a multilevel selection approach, which analyzes the strength of selection at different levels. Mathematically, the multilevel selection approach uses the Price equation, allowing us to quantify selection at different levels. This approach can then be applied to cooperative behaviours, particularly when they happen in group-structured populations.

#### The Price equation

The Price equation [Price, 1972] describes how the average value of a character changes 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 character. With the character value being  $z$ , its change from individual to offspring  $\Delta z$  and the number of offspring  $w$ , the mean change from individual to offspring in the population  $\Delta \bar{z}$  follows Price equation (where  $\bar{w}$  is the population mean fitness):

$$\Delta \bar{z} = \text{cov}(w/\bar{w}, z) + E((w/\bar{w})\Delta z) \quad (\text{I.1})$$

The first component  $\text{cov}(w/\bar{w}, z)$ , or change due to selection, is the covariance between character value and reproductive success. The second component  $E((w/\bar{w})\Delta z)$ , or change due to transmission, is the expectation of changes from individual to offspring, weighted by individual number of offspring: it describes how, in average, offspring differ from their parents [Gardner, 2008].

Price equation can simply be applied to individuals: in this case, it describes natural selection at the individual level. However, it can also be applied to groups when the parent and offspring entities are social groups. Indexing groups with  $i \in N$ , this yields:

$$\Delta \bar{z} = \text{cov}_N(w_i/\bar{w}, z_i) + E_N((w_i/\bar{w})\Delta z_i) \quad (\text{I.2})$$

The first term corresponds to between-group selection, and the second term to within-group selection. The proportion of individuals in a population structured in groups can thus be followed by taking into account changes within and among groups (Figure I.2).

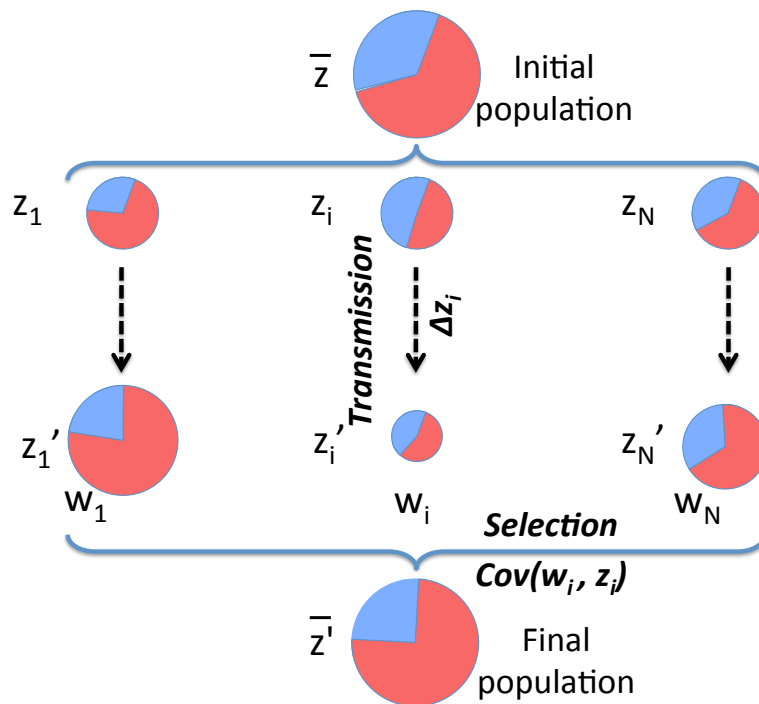


Figure I.2: **Price equation components and multilevel selection.** A population with the initial proportion of  $\bar{z}$  red individuals is distributed in  $N$  subpopulations (of equal size, for the sake of simplicity), each with initial proportion  $z_i$  of red individuals. In each subpopulation  $i$ , this proportion changes to  $z'_i$  (changes due to transmission), and the subpopulation size is multiplied by the factor  $w_i$  (success of subpopulation  $i$ ). When subpopulations are mixed, selection occurs among them (changes in  $z$  depend on  $z_i$  correlation with  $w_i$ , or  $\text{cov}(w_i, z_i)$ ), leading to a final population with the proportion  $\bar{z}'$  of red individuals.

This equation allows us to separate selection in among- and within-groups components. The relative strength of these components will determine the final direction of evolutionary change. Conflicts between levels of selection are cases where these two components have opposite signs. Moreover, Price equation further highlights that group selection will happen whenever there is a correlation between the genetic composition of a group and

group success.

### **Group selection and kin selection**

The validity of group selection and kin selection approaches and the relationship between them has been the subject of numerous debates [Leigh Jr, 2010b]. Still, Hamilton’s rule can be rigourously derived using Price equation [Hamilton, 1970], as well as group selection results. It can also be shown that group selection and kin selection are mathematically equivalent [Lehmann et al., 2007]. Both approaches essentially rely on the fact that interactions occur within groups, and individuals are more related within groups than among groups. In a multilevel selection framework, relatedness can be thought of as the fraction of the total weighting of selection that happens at the group level, or as the fraction of the total variance that is among groups [Frank, 2012]. To maintain cooperation, benefits have to be redirected to cooperators, which in a group-structured population is equivalent to saying that some groups are enriched in cooperators.

Group selection and kin selection are thus two ways of studying the same processes [West et al., 2006]. The multilevel selection approach is by nature well suited to analyze selection by looking at groups. However, it is less easy to apply when populations are not structured in clearly defined groups. Still, when groups can be defined, cooperation will very often increase group size, affecting the covariance component in Price equation (the  $cov_N(w_i/\bar{w}, z_i)$  term in Equation I.2). In the cases described below, both the multilevel and the kin selection approaches can be used.

### **Cooperation with effects on group size**

Cooperation can increase group fitness by increasing local carrying capacity. This is very often the case with microbial public goods [West et al., 2007c, Platt and Bever, 2009], for instance siderophores or secreted exoenzymes, that lead to higher local density. Even without clearly defined groups, a local increase in bacterial density thanks to public goods can favor cooperators.

With clearly defined groups, effects on group size and their consequences on the maintenance of cooperation can be easily followed. This has been done in experiments with the synthetic system we have mentioned, and that we will also use [Chuang et al., 2009]: the authors designed a population structure with subpopulations differing in their initial proportion of public good producers, providing positive relatedness between producers at the metapopulation scale. Population growth was dependent on the proportion of producers, providing the covariance between cooperative ability and reproductive success needed to

select for cooperation at the group level.

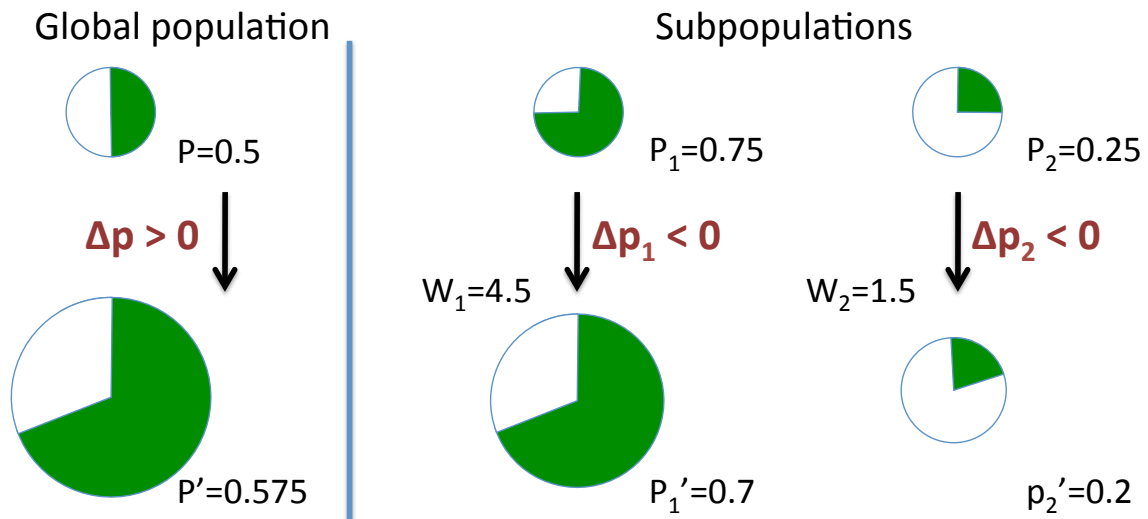


Figure I.3: **Principle of Simpson's paradox** (adapted from [Chuang et al., 2009]). In two groups with given proportions of P element, the proportion of P decreases within each group but increases overall. The right side shows the change in two subpopulations represented by pie charts of P (green) slices; the initial and final subpopulations with respectively  $p_i$  and  $p'_i$  proportions of P are connected by black arrows. The left side shows the corresponding composition of the initial and final global population formed by these two subpopulations, with respectively  $p$  and  $p'$  proportions. Simpson's paradox is that P decreases in each subpopulation ( $\Delta p_i < 0$  for every group  $i$ ) but increases overall ( $\Delta p > 0$ ), because the global proportion of P is a group size-weighted average that differs from the nonweighted average.

This population structure and benefits of cooperation lead to an instance of Simpson's paradox (Figure I.3), as interpreted by the authors to refer to a well-known statistical paradox: despite a decrease in producers frequency in all populations (because public good production is costly), producers increase in frequency at the global scale when all populations are pooled. Here, pooling corresponds to a global competition among groups, where populations with more producers contribute more to fitness.

In conclusion, selection can act at different levels, which often leads to conflicts when, for example, fitness benefits at an individual scale are costly at a group scale. A further complication is that transmission of genes or individuals can happen not only by multiplication within groups, but also by migration between groups. This is particularly the case of symbioses where groups of symbionts are living within a host: migration of symbionts between groups is equivalent to horizontal transmission from the perspective of the host.

## I.3 Horizontal transmission and its consequences

### I.3.1 Symbionts with vertical or horizontal transmission

Symbiosis refers to a close and long-term interaction between species. It sometimes exclusively describes only the beneficial associations, but usually also includes parasitism, while beneficial symbioses are called mutualism [Herre et al., 1999]. Symbioses are often asymmetric [Leigh Jr, 2010a], with one partner living within the cells or the body of the other partner, or host. Mobile genetic elements, that I will describe extensively in Chapter II, can be seen as symbionts living in association to host genes, if symbiosis is seen in a broad sense as an association between genetically dissimilar units [Smith, 2007].

#### Costs and benefits of symbiosis

Symbionts vary in their costs and benefits to the host. One extreme example is harmful parasites, that can lead to host disease and death. Other symbionts may provide benefits, like nutrients or antibiotics [Leigh Jr, 2010a, Sachs et al., 2011]. The net effect of symbionts on their host depends on the environment. For instance, bacterial plasmids bear antibiotic resistance genes, that will benefit their host only in the presence of antibiotics.

Symbionts usually benefit from the host, that provides protection or nutrients, but vary in their dependance to the host [Herre et al., 1999]. Some symbionts have a free-living phase and can survive in the environment; other are completely dependant on their host, like obligate endosymbionts. The dependance of hosts and symbionts is related to the ability of symbionts to undergo horizontal transmission between hosts.

#### Vertical and horizontal transmission

Vertical transmission is the transmission of symbionts from parent to offspring. It characterises obligate endosymbionts like *Buchnera* bacteria in insects. Horizontal transmission is the transmission of symbionts between different hosts, that are not necessarily related (Figure I.4). Horizontal transmission happens for infectious parasites, but also for mutualistic symbionts like *Rhizobium*. Sometimes, horizontal transmission is in fact analogous to vertical transmission, when it happens at a local scale and leads to the re-infection of host offspring.

Horizontal transmission also characterises selfish genetic elements like plasmids. In bacteria, mobile genetic elements are sequences that have the ability to move between bacteria (horizontal gene transfer). More generally, selfish genetic elements can be con-

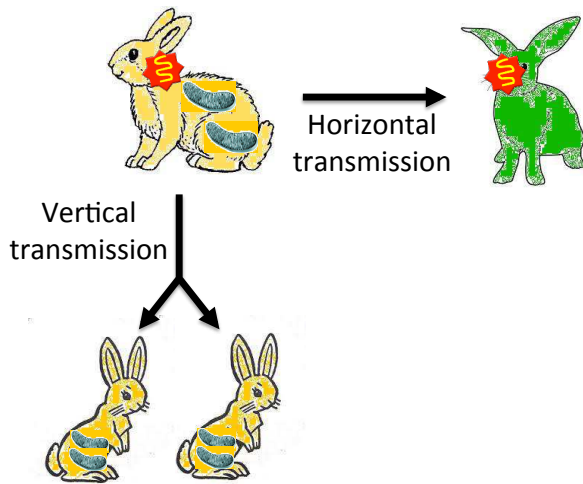


Figure I.4: **Vertical and horizontal transmission.** Vertical transmission happens from parent to offspring (yellow individuals), horizontal transmission happens between individuals that are not parent and offspring (yellow to green here). Examples represented are mitochondria (with vertical transmission, in blue) and the myxoma virus causing myxomatosis (with horizontal transmission, in red) in rabbits.

sidered as having horizontal transmission from the gene's perspective [Smith, 2007], even if they are transmitted from parent to offspring, like segregation distorters during meiosis in Eukaryotes [Burt and Trivers, 2006], as their transmission differs from the transmission of host genes: they effectively invade new host gene lineages during sexual reproduction.

### Consequences on symbiosis

With vertical transmission, both hosts and symbionts benefit from host reproduction by increased reproductive success [Herre et al., 1999]. This effect of vertical transmission can be seen as an example of partner-fidelity feedback, one of the mechanisms ensuring the maintenance of cooperation between species cited in section I.1.2 [Sachs et al., 2004]. Strict vertical transmission has led to some of the major evolutionary transitions, after which symbionts cannot replicate independently anymore, but only as a part of a larger whole [Szathmary and Smith, 1995]. This is the case of some eucaryotic organelles, such as mitochondria, that arised from formerly free-living, symbiotic bacteria.

With horizontal transmission, symbionts gain access to new hosts, decoupling their fitness from initial hosts [Herre et al., 1999]. Selection can then favor symbiont traits that are deleterious for initial hosts. Particularly, symbiont traits increasing horizontal transmission can be selected as they lead to increased reproductive success for the symbiont. These traits can be costly for the host. Host-symbiont conflicts about reproduction can also happen: symbionts with maternal transmission in sexual species, as mitochondria or *Wolbachia* bacteria for instance will benefit from biased sex ratios towards females. However, symbionts generally benefit as well from not harming the host, by increased vertical transmission or increased period of interaction.

### **I.3.2 A trade-off between vertical and horizontal transmission**

#### **The transmission-virulence trade-off**

The transmission-virulence trade-off is a general hypothesis about host-parasite interactions, including both long-term interactions with vertical transfer and transitory infections. Virulence is the harm a parasite does to its host, either by increasing mortality, or by non-lethal effects that reduce fecundity [Alizon et al., 2009]. The trade-off hypothesis relies on the idea that there is a genetic trade-off between the rate of parasite horizontal transmission between hosts and the duration of the infection, because increasing transmission increases virulence [Anderson and May, 1982, Ewald, 1987]. Decreasing the duration of the infection then decreases the probability of further transmission. This can lead to an evolution of optimal virulence, maximizing global transmission, which has been shown in some studies. For instance, intermediate virulence maximises transmission in *Daphnia* parasites [Jensen et al., 2006]. Still, the trade-off may not always exist, as for instance virulence can be non-adaptive, *ie* not linked to increased horizontal transmission (for example, meningitis is caused by the rare infection of the meningeal space by bacteria that are not likely transmitted further from this space [Frank, 1996]).

For parasites with vertical transmission, the transmission-virulence trade-off becomes a vertical transmission - horizontal transmission trade-off, as increased virulence also reduces vertical transmission of the parasites [Lipsitch et al., 1996]. If increasing horizontal transmission is costly for host fitness, it will also decrease the part of parasite fitness due to vertical transmission.

#### **Evolution of horizontal transmission**

The evolution of horizontal transmission should depend on its benefits relative to accrued costs to the initial hosts, which decrease vertical transmission. Benefits from investing in horizontal transmission will mainly depend on the availability of new hosts [Anderson and May, 1982]: if many hosts are available, then horizontal transmission will contribute more to fitness than vertical transmission.

This hypothesis has been confirmed by some experimental studies. For instance, filamentous bacteriophages are bacterial viruses that can be transmitted by both vertical and horizontal transmission, which ratio can be controlled experimentally. Bull and coauthors showed that evolution with more opportunities for vertical or horizontal transmission leads to respectively lower or higher virulence [Bull et al., 1991, Messenger et al., 1999].



They further confirmed that low virulence was adaptive in the absence of available hosts for horizontal transmission, and high virulence was adaptive in the absence of vertical transmission, increasing horizontal transmission.

We focused here on the trade-off between symbiont horizontal transmission and host reproductive success. However, there are cases where symbiont transmission can actually benefit the host, through indirect effects.

### I.3.3 Benefits of horizontal transmission for symbiosis

Symbiont transmission can be beneficial to the host if other populations in the community are taken into account. Indeed, populations in a community interact in multiple and often antagonistic ways. The transmission of symbionts can affect these interactions in a way that favors initial hosts (summarized in Figure I.5).

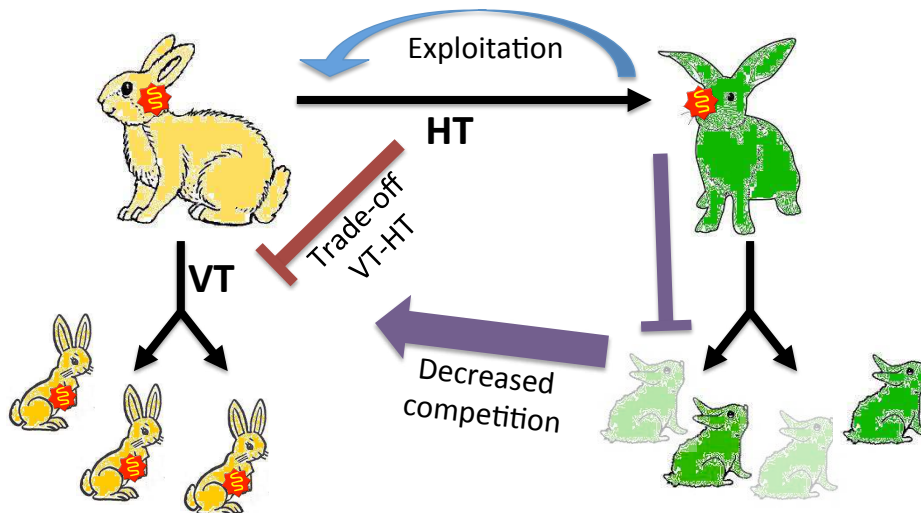


Figure I.5: **Consequences of horizontal transmission on host fitness.** Potential consequences of symbiont horizontal transmission (HT) to a non-related individual (green) for both the host (yellow) and the symbiont (red) vertical transmission (VT). HT negatively affects the host because of its costs (trade-off, red arrow); but can be indirectly beneficial (blue arrows) because it frees up resources from the non-related individual, or because it helps the host lineage in competition with the non-related lineage.

#### Apparent competition via parasites

Populations of different species or ecotypes are often in competition. Competition happens first of all when populations use the same resources (exploitation competition, by resource consumption). But competition can also be due to direct active behaviours that interfere with the other population (interference competition, for instance by aggression). Finally, another mechanism of competition is apparent competition [Price et al., 1986],

which happens indirectly through common parasites or predators: the presence of one population increases the density of these enemies, which then increase the harm to the other population. Apparent competition is a form of interference competition, where interference happens through shared predators or parasites.

In this context, horizontal transmission can be seen as a mechanism of interference competition, where hosts benefit from increased horizontal transmission of their parasites to competitors (Figure I.5, violet arrows). An advantage of that strategy compared to interference competition not involving parasites is that parasites can amplify on competitors, and have a long-lasting effect. Recently, examples of interference competition have been found in microbes and their symbionts. For instance, virus carriers in *Escherichia coli* were found to outcompete sensitive cells thanks to virus spread [Brown et al., 2006], initial hosts being protected from lysis by bearing the virus in a lysogenic state. Amplification of the virus on sensitive cells was essential for this effect.

In an evolutionary context, interference competition is analogous to spite, so its maintenance is again difficult to explain, with individuals that do not perform a spiteful behaviour still benefiting from the harm done on competitors. Spite is even more difficult to explain than altruism, as it requires negative relatedness between actor and recipients (see section I.1.2). Spite through parasites could be easier to maintain, as amplification in recipients reduces costs for actors [Dionisio, 2007], and actors can often be protected by immunity, allowing a form of kin discrimination.

### **Manipulation via horizontal transmission**

Transmission of parasites can also be beneficial because they induce a specific reaction in recipients that is beneficial to initial bearers of the parasite (Figure I.5, blue arrow). For instance, herbivore insects are vectors of plant pathogenic viruses [Belliere et al., 2004] that inhibit the plant response against insects. Insects can also secrete symbiotic bacteria [Chung et al., 2013], which fake a bacterial attack and inhibit the response to herbivory. A more specific case of manipulation by horizontally transmitted symbionts is the case of polydnaviruses [Strand and Burke, 2012]. Those viruses are transmitted vertically in insect parasitoids, and virions are produced only when larvae are injected into the host insect. Virions then infect cells of this host insect, where their expression inhibits the immune response and interferes with growth, allowing the development of larvae. In a similar way, *Agrobacterium tumefaciens* bacteria transfer a specific DNA sequence to plant cells they parasitize, the T-DNA. This segment bears genes with plant-specific ex-

pression that are involved in the production of specific metabolites that the bacteria can use [Otten et al., 1992] (see Figure III.2). In those last cases, symbionts are horizontally transmitted at a short timescale, but do not actually reproduce in the secondary host. They benefit only from vertical transmission in their original host, with no direct increase in fitness through horizontal transmission.

In conclusion, individuals affect their neighbours in multiple ways. A major factor controlling the evolution of social interactions towards cooperation or conflict is the occurrence of shared genes, or relatedness, between interacting individuals. Shared genes usually arise from common descent, but horizontal gene transfer in bacteria modifies this heritability pattern. Mobile genetic elements have two sides: they are responsible for horizontal gene transfer, but are also bacterial symbionts with patterns of transmission that natural selection will act on. In the following chapters, I focus on the properties of these peculiar bacterial symbionts.

## Chapter II

# Mobile genetic elements as bacterial symbionts

Mobile genetic elements (MGEs) are defined as segments of DNA that encode proteins mediating the movement of their own DNA within genomes or between cells [Frost et al., 2005]. Mobile genetic elements are not specific to prokaryotes. However, they are strongly affected by a key feature of prokaryotes, reproduction through clonal multiplication. Thus, mobile elements do not spread in populations through sexual reproduction like in Eucaryotes, but through other mechanisms, decoupled from reproduction. Archaeal mobile elements seem to be generally similar to bacterial ones, but are little known [Smillie et al., 2010]. For this reason, I focus on bacterial mobile elements.

Here, I will first describe the different mechanisms by which genes can spread between bacterial cells, and then the maintenance of mobile elements in bacterial populations by both vertical and horizontal transmission. I will focus on a class of mobile elements most relevant to my work, plasmids.

## II.1 Mechanisms of horizontal gene transfer in bacteria

Horizontal gene transfer between bacteria can happen in three ways: transfer of naked extracellular DNA (transformation), direct transfer through cell-cell contact (conjugation) or transfer through bacterial viruses (transduction).

### II.1.1 Natural transformation: transfer of naked DNA

#### Discovery of horizontal transfer

The discovery of horizontal gene transfer is actually linked to the discovery of the nature of genetic material. First, Griffith's experiment showed that an agent, so-called "transforming principle" originating from virulent pneumococci transformed another non-virulent strain into a virulent one [Griffith, 1928]. This transforming principle was then shown to be DNA [Avery et al., 1944]. DNA was thus able to be taken up by cells unrelated to the cells it originated from, demonstrating the existence of horizontal gene transfer. By analogy to the horizontal transmission of symbionts, horizontal gene transfer refers to the transmission of genetic information between organisms that does not happen through

reproduction, but between potentially unrelated organisms.

### **Mechanisms of natural transformation**

Transformation is one mechanism of horizontal transfer where DNA is acquired from the extracellular environment. Transformation can be enforced experimentally through physical or chemical treatments, but natural transformation is an active uptake of extracellular DNA [Johnsborg et al., 2007], that takes place in some bacterial species. Natural transformation involves the binding of DNA to specific receptors on the cellular surface, then uptake of single-strand DNA. Finally, non-replicating DNA molecules will be maintained in the cell only if they are integrated by homologous recombination.

All these processes happen in so-called competent bacteria: competence, the ability to take up extracellular DNA, has to be induced by the cell. Natural transformation is well described in some model species like *Bacillus subtilis*, but it is not known whether it happens in many other species. Competence may exist, but require specific conditions that have not yet been tested in the lab. For instance, *Escherichia coli* can be naturally competent after growth under static conditions [Sun et al., 2006]. On the contrary, competence may have been selected only in a few bacterial species.

### **The origin of transformed DNA**

Transformed DNA molecules are not always random. In some cases, DNA uptake can be sequence-specific, like in *Neisseria* species: efficient uptake occurs only for molecules containing a specific sequence around 10bp in length, which is enriched in the genome of the corresponding species. In other cases, the DNA present in the environment of the cell can be non-random: its release from neighbouring bacteria can be active. In streptococci, induction of competence leads to the lysis of a subpopulation of cells [Steinmoen et al., 2002]. This has been interpreted as a way of ensuring transformation of homologous DNA when there is no sequence specificity in DNA uptake. Finally, another active mechanism for ensuring that some DNA sequences get imported is the active secretion of DNA *via* a type IV secretion system in *Neisseria gonorrhoeae* [Hamilton et al., 2005].

## **II.1.2 Transduction: transfer through bacterial viruses**

A second mechanism for horizontal transfer was discovered in *Salmonella* [Zinder and Lederberg, 1952]. Here, gene transfer involved a "filtrable agent" adsorbed by bacterial cells. This process was defined as transduction: the transfer of genetic information mediated by bacteriophages, or bacterial viruses.

## Life cycles of bacteriophages

Bacteriophages consist of a DNA or RNA genome, protected by a proteic capsid. They attach to the surface of the cell and inject their genome inside the cytoplasm. Lytic phages immediately replicate using the cellular machinery, and then lyse the cell when virions are liberated. Temperate phages do not immediately lyse the cell: they go through an additional phase, called lysogeny, where the viral genome integrates in the host genome as a prophage.

## Gene transfer *via* transduction

Bacteriophages can transfer parts of the host genome in different ways [Canchaya et al., 2003]. In specialized transduction, the chromosomal DNA adjacent to the phage integration site is packaged with phage DNA because of imprecise excision of the prophage. In generalized transduction, any part of the host genome can be accidentally packaged in the capsule instead of phage DNA. Transduction can involve up to 100kb of DNA, and is limited by the content of DNA present in the capsule.

Transduction is also part of a peculiar form of genetic exchange, known in only a few bacterial species, that involves gene transfer agents (GTA) [Lang et al., 2012]. GTAs are phage-like entities that, unlike phages, always contain a random part from the cell's genome but are too small to package all the genes predicted to encode GTA particles themselves: GTA genes do not spread through viral infection, contrary to phages.

Finally, prophages themselves can carry genes that provide a new phenotype to host cells, a process called lysogenic conversion. For instance, a filamentous phage is responsible for the production of cholera toxin in *Vibrio cholerae* [Waldor and Mekalanos, 1986]. Gene transfer via bacteriophages is widespread on an evolutionary timescale: many bacteria contain multiple prophages, that can represent up to 16% of chromosomal DNA, and are often involved in host virulence, being expressed during lysogeny and independently from the rest of the prophage genome [Canchaya et al., 2003].

On a shorter timescale, the frequency of horizontal gene transfer by transduction is linked to phage dynamics. Both generalized transduction and GTAs involve rare events of gene packaging from the whole genome. Specialized transduction and lysogenic conversion concern specific genes, but happen at much higher rates, closer to phage infection rates.

### II.1.3 Conjugation: transfer through cell-cell contact

#### Discovery of conjugation

This third mechanism for horizontal transfer was discovered by Tatum and Lederberg by studying auxotrophic mutants in *Escherichia coli*. They showed that genetic recombination was happening in mixed cultures, and that this phenomenon needed cell-cell contact between bacteria [Tatum and Lederberg, 1947]. Conjugation was then defined as a process by which DNA is transferred from a donor to a recipient cell during cell-cell contact.

Moreover, this phenomenon did not take place in all the strains that were studied and was asymmetrical: transfer happened from "donor" to "recipient" cells. Additional studies revealed that conjugation was driven by the presence of an "infective factor", then called the F factor, present in donor cells [Hayes, 1953]. Finally, it was shown that this factor was a DNA molecule, independent from the host chromosome, that transferred between cells [Marmur et al., 1961]: the F plasmid. The same mechanism of conjugation was also involved in the transfer of a molecule conferring drug resistance [Watanabe, 1963], indicating the existence of multiple kinds of plasmids.

#### Overview of the conjugation steps

Conjugation involves several steps. Donor and recipient cells are brought into close contact by the process of mating-pair formation, through specific organelles called conjugative pili [Arutyunov and Frost, 2013], which are formed by a type IV secretion system (T4SS). In parallel, DNA is associated to a protein complex called the relaxosome and transported to the mating pore. The nicked DNA strand (T-strand) is then transferred to the recipient, where it is recircularized. Finally, both strands are replicated.

#### Mating-pair formation

Mating-pair formation involves the assembly of pili by a type IV secretion system (T4SS) [Arutyunov and Frost, 2013]. Pili are constructed by polymerization of pilin in a helical array, and can measure from 1 to 20  $\mu\text{m}$  in length and 8 to 11 nm in diameter for F plasmid. Pili mediate the recognition of recipient cells.

Pili involved in conjugation are classified according to their morphology [Bradley et al., 1980]. Rigid pili are not able to retract or stabilize mating pairs, so they are associated with transfer in solid environments only. Thick flexible pili, including F plasmid pili, can retract by depolymerization, which brings together donors and recipient cells and stabilizes the mating pair. Unlike rigid pili, thick flexible pili allow efficient transfer in

liquid medium. DNA can transfer through extended pili [Babic et al., 2008], but transfers more efficiently after pilus retraction.

### **DNA mobilization**

DNA mobilization is the processing of DNA by transfer and replication proteins. A key protein is the relaxase, that binds specifically to a cognate *oriT* sequence. In association to accessory proteins, the relaxase forms a complex with DNA called the relaxosome [Alvarez-Martinez and Christie, 2009]. This complex is then recognized by the coupling protein (T4CP), which brings it to the mating pore. The relaxase cleaves the *oriT* sequence at the nicking site (*nic*), then stays associated with the 5' end of the cleaved DNA strand. Single-strand DNA associated to the relaxase is transferred to the recipient cell, where the relaxase ensures plasmid recircularization. *oriT* is therefore a specific sequence that is needed for a DNA molecule to be recognized by the relaxase and mobilized. Moreover, as *oriT* is the only determinant of transfer in *cis*, transfer proteins encoded on another genetic element can mobilize in *trans* any sequence bearing *oriT*. However, the recognition by the relaxase is highly precise: just a few base pairs differences between sequences can prevent recognition [Harley and Schildbach, 2003].

Mobilization concerns both plasmid or chromosomal molecules. Plasmids are independent DNA molecules, but can be integrated in the bacterial chromosome, leading to mobilization from the integrated *oriT* and transfer of parts of the chromosome. This is what happens in Hfr (High frequency of recombination) cells in *Escherichia coli*, where recombination of loci on the chromosome is due to the insertion of F plasmid, initiating conjugation from the chromosome. Other chromosomal sequences, called Integrated and Conjugative Elements (ICE), bear their own *oriT*. They undergo an additional step of excision to create a circular intermediate, from which mobilization occurs.

Finally, two other potentially widespread mechanisms of gene transfer have still been discovered recently. DNA transfer within and between species was shown to happen through nanotubes bridging neighbouring cells [Dubey and Ben-Yehuda, 2011]. Similarly, DNA transfer through membrane vesicles has also been observed [Yaron et al., 2000].

#### **II.1.4 Mobile elements: gene vectors and symbionts**

Except in the case of natural transformation, horizontal gene transfer is mainly associated to the spread of mobile elements: these elements actively transfer themselves and mobilize occasionally other sequences, by either transduction or conjugation.



## Diversity of mobile elements

The most well-known mobile elements are bacteriophages and plasmids, but many other elements exist. Bacteriophages are characterized by an extracellular phase where DNA is encapsulated. However, some prophages like the P1 prophage are also maintained as plasmids in the cytoplasm during lysogeny. Conjugation can be mediated by a huge diversity of mobile elements, whose boundaries are often not precisely defined [Osborn and Böttner, 2002], and include plasmids (extrachromosomal elements), but also genomic islands or conjugative transposons (integrative chromosomal elements), which are widespread but less studied than plasmids [Wozniak and Waldor, 2010]. In comparison to plasmids, these do not encode maintenance functions, but additionally encode excision functions.

We will here mainly focus on plasmids, that carry more diverse genes than the small cargo regions of phages, and have been shown to be responsible for most of the transfer events between bacteria [Halary et al., 2009].

## Plasmid definition and prevalence

The word plasmid was first introduced as a "generic term for any extra-chromosomal hereditary determinant" [Lederberg, 1952], which included mitochondria or viruses, by analogy to endosymbioses. Nowadays, a plasmid is typically defined as "a stable, self-replicating entity, which is smaller than the cellular chromosome and which usually does not contain genes required for essential cellular functions" [Frost et al., 2005]. The criterion of not bearing essential genes is often difficult to test, as it strongly depends on the ecological context, with some genes being essential only in specific conditions.

Although the prevalence of plasmids in bacteria is insufficiently quantified, we know it varies greatly. In *Escherichia coli*, cells carry on average between 0.45 and 2.6 plasmids per cell depending on isolates [Souza et al., 1999]. However, in the extreme case of *Borrelia* species, up to 21 plasmids can coexist in some strains [Casjens et al., 2000].

## The structure of plasmids

Plasmids are generally double-stranded circular DNA molecules, but exceptions exist, like the linear plasmids of *Streptomyces* [Frost et al., 2005]. Plasmid sizes are hugely diverse and range from 0.85 kb to 2 Mb for some megaplasmids [Slater et al., 2008a]. Plasmids are generally organized in functional genetic modules [Toussaint and Merlin, 2002]: replication, copy number control and maintenance modules ensure vertical transfer, conjugation or mobilization modules ensure horizontal transfer, and adaptation modules,

consisting of so-called accessory genes, increase the prevalence of plasmid-bearing cells in the population by providing benefits to the cell in some conditions (Figure II.1).

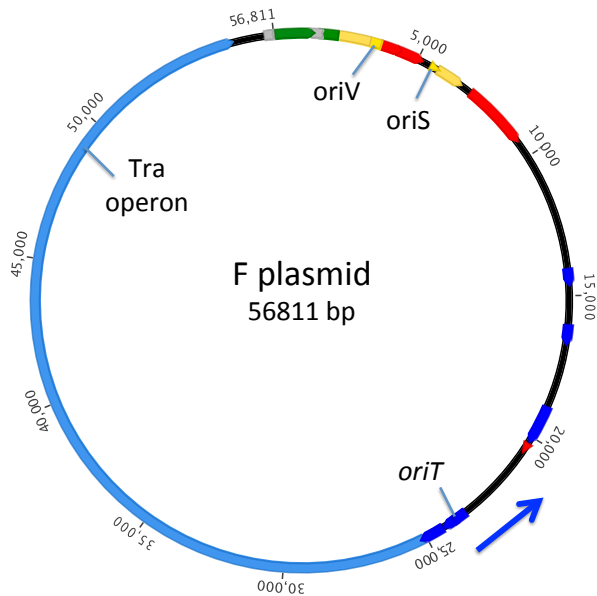


Figure II.1: **Plasmid structure: the example of F.** The F plasmid shown is pOX38::Tc [Anthony et al., 1994]. Genes involved in conjugation are represented in blue. The ones located at the right of *oriT* ensure plasmid maintenance in recipients after conjugation. The blue arrow indicates the region that first enters the recipient cell. Yellow indicates replication and copy number control genes. Maintenance genes are represented in red, accessory genes are represented in green. *oriT* and origins of replication (*oriV* and *oriS*) are indicated. Sequences in gray are insertion sequences.

All these modules will play a role in the maintenance of plasmids in host populations, by affecting either vertical transmission or horizontal transmission of plasmids.

## II.2 Vertical transmission of plasmids

Vertical transmission of plasmids depends on two parameters: the maintenance of plasmids during host cell division and the net effect of plasmids on host fitness. Plasmid effect on host fitness determines if plasmid-free cells have a selective advantage compared to plasmid-bearing cells.

### II.2.1 Plasmid maintenance mechanisms

For plasmids to be maintained in both daughter cells after division, several maintenance mechanisms are needed (Figure II.2). First, plasmid replication has to be coupled to the growth of the cell. They further have to be accurately distributed between daughter cells at division, which involves partitioning (ensuring that each daughter cell gets at least one plasmid copy) and multimer resolution systems (that counteract plasmid multimerization leading to unequal sharing of plasmids between daughter cells). If any of these mechanisms fails, plasmid-free cells may arise. Those will usually not reproduce, because of the existence of post-segregational killing mechanisms that kill plasmid-free cells.

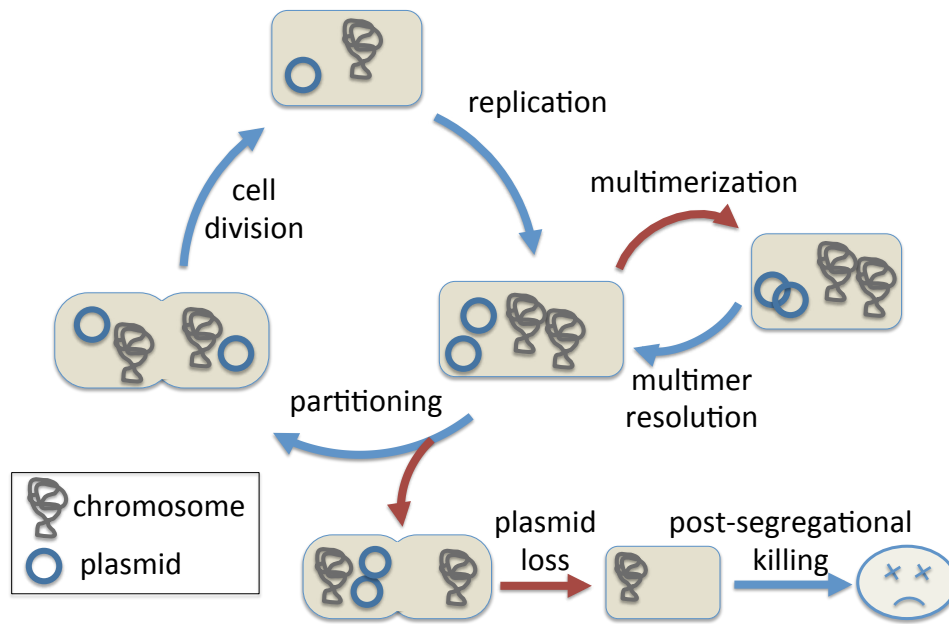


Figure II.2: **Plasmid maintenance mechanisms.** Mechanisms enhancing plasmid maintenance in bacterial lineages are indicated with blue arrows, factors that decrease it or failures of maintenance mechanisms are indicated with red arrows.

### Plasmid replication and copy number

Circular plasmids replicate by rolling circle, theta-type or strand-displacement mechanisms [del Solar et al., 1998]. Generally one essential plasmid region contains the loci associated to replication: the origin(s) of replication (called *ori*, such as *oriV* and *oriS* for F plasmid), often a gene coding for a protein involved in the initiation of replication, termed Rep protein, and additional genes involved in the regulation of replication. Replication itself then involves both plasmid-encoded and host-encoded proteins.

Each plasmid is characterized by a copy number, the number of plasmid copies per cell under specific growth conditions, which is usually regulated at the initiation stage by negative regulatory circuits [del Solar et al., 1998]. Negative regulators reduce the probability of replication when the number of plasmid copies in each cell increases. The reduction is mainly achieved by regulation of the expression of Rep proteins. The regulation of the copy number lowers the burden on the host cell by preventing excessively high numbers of plasmids per cells. It also reduces the variance in copy number, decreasing the probability that plasmid-free cells arise during bacterial division [Paulsson, 2002].

This negative control of copy number has consequences on plasmid competition within cells. Two plasmids with identical replicons cannot stably coexist within a cell lineage, as they share the same regulatory mechanism: the copy number will be regulated for the sum of both plasmids, effectively decreasing the copy number of each of them. This

leads to loss of one of both plasmids, a phenomenon called plasmid incompatibility. For this reason, the classification of plasmids was historically based on incompatibility groups (Inc): groups of incompatible plasmids, for which we can deduce that they share the same replication mechanisms.

Plasmids can be separated in two categories based on their copy number. High-copy number plasmids (typically around 20 copies per cell, but up to 200 copies in some cases [Smillie et al., 2010]) can rely on random partitioning, as cells not bearing any plasmid will arise very rarely by chance. Low-copy number plasmids (up to 5-10 copies per cell [Thomas, 2000]) have a high chance of being lost randomly at division in absence of partitioning mechanisms.

### **Plasmid partitioning and addiction**

Plasmids, and mainly low-copy plasmids, have multiple mechanisms that favor their equal distribution to both daughter cells. Homologous recombination between identical plasmids first leads to multimerization of plasmids. This phenomenon causes a decrease in the effective number of molecules segregating at division, increasing the chance of plasmid loss. Many plasmids bear a multimer resolution system that counters multimerization, increasing their effective copy number [Bahl et al., 2009]. Low copy number plasmids further bear active partitioning systems, consisting in an ATPase associated to a DNA-binding protein. These proteins bind on a specific plasmid site, functioning as centromere-like elements. The ATPase forms filaments and ensures migration of plasmids during division, so that both daughter cells receive at least one plasmid copy [Bahl et al., 2009].

Despite these mechanisms, some plasmid-free cells may still arise. Additional mechanisms, called addiction systems or post-segregational killing (PSK) systems, ensure the elimination of plasmid-free cells after division. PSK systems are based on two genes, encoding a protein toxin and a corresponding antitoxin that is either an antisense RNA or a protein, and prevents toxin function (Figure II.3). The key feature of these systems is that the toxin is more stable than the antitoxin: after plasmid loss, the expression of both genes stops, and the toxin remains present longer than the antitoxin, killing the cell or stopping its growth. PSK systems ensure that plasmid-free cells do not outcompete plasmid-bearing cells. However, they do not always bring a direct advantage to a plasmid, but mainly need to be understood in terms of competition between plasmids [Cooper and Heinemann, 2005]: when two plasmids with and without PSK compete within the same cell, plasmids with PSK systems have a selective advantage against plasmids without

them, as PSK will bias plasmid segregation in favor of plasmids carrying them.

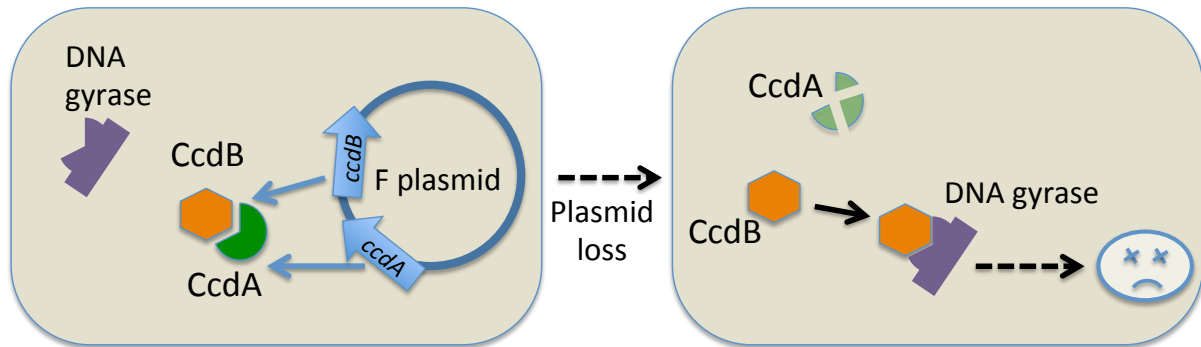


Figure II.3: **PSK system in the F plasmid.** When F is present in a cell, both *ccdA* and *ccdB* genes are expressed. CcdA protein forms a complex with CcdB protein, preventing the diffusion of CcdB monomers. When F is lost upon division, CcdA is rapidly degraded by proteases. CcdB protein targets the DNA gyrase and inhibits it, leading to cell death.

### Plasmid loss

Despite these maintenance mechanisms, plasmid segregation (loss after cell division) will still occur at low rates. Actual rates of segregation are difficult to measure, as plasmid-free cells spread rapidly when plasmids impose a cost on the host [Lenski and Bouma, 1987]: differential growth effects are much stronger than actual loss. Loss rates measured by high-throughput methods and minimal times of growth are actually extremely low: R1 plasmid with a non-functional partition system was estimated to have a loss rate as low as  $10^{-3}$  per cell per generation, and loss of R1 with a functional partition system was not measurable [Lau et al., 2013].

When partitioning and addiction systems are functional, the major force acting on the vertical maintenance of plasmids will thus be their effects on host fitness.

### II.2.2 Plasmid effects on host fitness

First, plasmid maintenance in cells leads to unavoidable costs for the host. Then, accessory genes can be either beneficial or costly to the host depending on the environment. The net effect of plasmids on host fitness will thus be strongly variable.

#### Costs of bearing plasmids

Plasmid presence in a cell has been shown to lead to numerous costs. This has been called the "metabolic burden" of plasmids, and can lead to increased lag phase or decrease of the exponential growth rate of cells in culture [Diaz Ricci and Hernández, 2000]. First, plasmid replication and maintenance are costly. Both plasmid copy number and plasmid

size have an influence on this cost. However, the main factor in plasmid cost seems to be the expression of genes carried by the plasmid. Gene expression is metabolically more costly than replication, and leads to a competition in the cell for limiting ribosomes and RNA polymerase [Bailey et al., 1986], which can strongly outweigh any costs of plasmid replication [Bentley et al., 2009]. Finally, PSK systems also impose a cost on the host, as plasmid-free cells that arise do not reproduce [Cooper and Heinemann, 2005].

Some costs of plasmids are increased by plasmid mobility between unrelated hosts, as plasmid genes may be maladaptive in the context of the new host [Baltrus, 2013]. For instance, proteins may misfold, or interact with cellular networks and metabolism in a detrimental way. Codon usage will also differ for genes originating from distant hosts, increasing the cost of gene expression. A high cost of protein expression can also be explained by recent horizontal transfer, with non-optimal expression of new proteins. Interestingly, mechanisms are known that specifically reduce these costs of plasmid gene expression: H-NS like proteins are global regulators of gene expression that have a preference for AT-rich sequences, which corresponds usually to horizontally acquired genes. H-NS-like protein binding to plasmid genes suppresses the expression of these genes [Lucchini et al., 2006]. Moreover, some plasmids actively diminish their cost to the cell by encoding H-NS like proteins themselves, thus supplementing host H-NS proteins [Doyle et al., 2007]. Finally, in the case of transferable plasmids, horizontal transfer mechanisms are also costly to the host cell. I describe these costs in greater detail below.

### **Costs of plasmid transfer**

Plasmid conjugation is thought to be costly to the cell, first because of metabolic costs. Costs of conjugation are mainly observed for plasmids with derepressed conjugation (as we will see in section II.3.1, conjugation is often repressed, reducing strongly the expression of transfer genes). Derepressed plasmids strongly express transfer genes, including pilin synthesis and T4SS assembly genes.

In experimental plasmid-host coevolution, a correlation between plasmid conjugation rate and plasmid cost to the cell has often been found [Turner et al., 1998]. Derepressed plasmids generally decrease the growth rate of the cell: for instance, the R1-16 derepressed plasmid increases the generation time of *Escherichia coli* cells by 14%, while R1 plasmid, which conjugation is repressed, increases it only by 6% [Haft et al., 2009]. Finally, the synthesis of the type IV secretion system affects the integrity of the cell envelope, and elicits extracytoplasmic stress responses [Zahrl et al., 2006].

In natural environments, a further cost linked to plasmid transfer is due to the existence of male-specific phages. These phages target specifically bacterial cells that bear conjugative plasmids, because their entry in the cell is based on specific binding to conjugative pili. For instance, M13 and R17 phages recognize specifically domains of the F pilin subunit [Manchak and Frost, 2002], and can infect only *Escherichia coli* cells that bear F plasmid. The costs imposed by bacteriophages will vary depending on bacteriophage abundance and their accessibility to bacterial cells, and also on the type of bacteriophage (some male-specific phages, including M13, are filamentous phages that are liberated from the cell by secretion, so only reduce the cell growth rate but do not lyse the cell).

### **Benefits of bearing plasmids**

Benefits of bearing plasmids are hugely variable, and depend on plasmid accessory genes. Some plasmids, called cryptic plasmids, do not bear any genes known to be involved in other functions than maintenance or transfer.

The earliest described and most studied function of plasmid genes benefiting the cell is antibiotic resistance [Watanabe, 1963]. So-called R-plasmids often bear multiple resistance genes, and are responsible for the transfer of multiple antibiotic resistance between distantly related species. Plasmids can also provide resistance to heavy metals, or specific metabolic abilities like enzymes degrading specific substrates. Some plasmids confer virulence properties to their hosts, by bearing toxin production genes, like the S-endotoxin in *Bacillus thuringiensis* [González et al., 1982] or the Ti plasmid of *Agrobacterium tumefaciens* that is responsible for pathogenicity [Otten et al., 1992]. Plasmids are also involved in the mutualistic behaviour of bacteria, as in the case of nodulation in *Rhizobium*, where nodulation and N<sub>2</sub> fixation genes are present on symbiotic plasmids [Sachs et al., 2010], or in bacterial warfare, as in the case of bacteriocines, bacteriocidal proteins that kill bacteria closely related to the producing strain [Riley and Wertz, 2002].

The reason why plasmids bear some specific classes of accessory genes is a crucial evolutionary question, that I will address further in section III.2.1 and in my work.

## **II.3 Horizontal transmission of plasmids**

Horizontal transmission of plasmids happens mainly by conjugation, and plasmids can be classified according to their transfer properties: self-transmissible or conjugative plasmids carry themselves both mobilization and mating-pair-formation genes. Mobilizable

plasmids are plasmids that use other mating-pair formation systems present in the cell (usually of a co-infecting self-transmissible plasmid), but carry mobilization genes (at least a relaxase recognizing their own *oriT*, sometimes also the type IV coupling protein). Finally, non-mobilizable plasmids do not spread by conjugation at all. In sequenced plasmids of proteobacteria, the frequency of conjugative plasmids was estimated as 28%, and another 24% were classified as mobilizable, based on the presence of relaxase or T4SS sequences [Smillie et al., 2010]. The rate of plasmid transmission by conjugation is highly variable, as it depends not only on plasmid and chromosomal factors, but also on donor and recipient cells, as well as environmental conditions.

### II.3.1 Regulation of transfer in donor cells

Plasmid transfer by the donor cell is mainly regulated by changes in the expression of transfer proteins. The expression of the transfer apparatus imposes a burden on the cell, diminished by regulation. The regulation of gene expression involves both plasmid and chromosomal genes. Here I will mainly focus on F-like plasmids, as the F plasmid is one of the most studied conjugative plasmids, and the one we will use in our experiments.

#### Self-regulation of transfer by plasmids

Conjugative plasmids control their own transfer, by regulating the expression of transfer proteins [Zatyka and Thomas, 1998]. In the case of F, the transfer region consists in approximately 40 genes, coding for pilin and pilus assembly proteins, DNA metabolism and regulation of transfer. The transfer region is adjacent to the origin of transfer that contains the nicking site *nic* (Figure II.4). It contains two monocistronic operons, encoding *traM* and *traJ*, followed by the 30-kb long *tra* operon whose first gene product is *traY*. For transfer to happen, the *tra* operon needs to be expressed, yielding T4SS assembly and mating-pair formation. In parallel, the relaxosome is assembled with the relaxase *TraI*. Nicking by the relaxase is performed after receiving a signal, through *TraM*, that a functional mating pair has been formed.

In F-like plasmids, the expression of the *tra* operon is usually repressed, so that only 0.1% of the bacteria actually transfer plasmids to recipients. F plasmid is a derepressed plasmid: all cells actively transfer plasmids. This is due to the interruption of *finO* by an IS3 insertion sequence [Cheah and Skurray, 1986], that inactivates the repression mechanisms. Repression is based on the fertility inhibition system *FinOP*, that consists in two genes, *FinO* and *FinP*. *FinP* encodes the antisense RNA *FinP*, that is transcribed



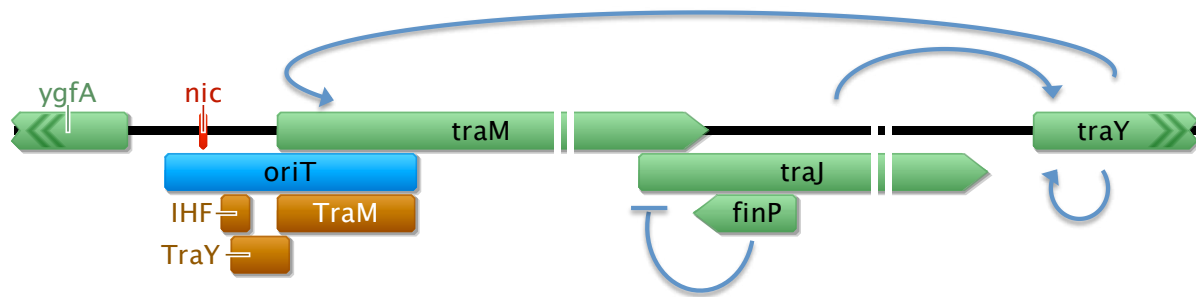


Figure II.4: **Details of *oriT* and adjacent genes map, and key elements of transfer initiation.** *OriT* is in blue, with known protein binding sites indicated in brown. Adjacent genes are represented in green (*traM* and *traJ* are truncated). The nicking site *nic* is shown in red. Regulatory interactions are indicated by blue arrows.

constitutively, and binds to the ribosome binding site for *traJ*, preventing *traJ* translation. FinO protein stabilises FinP.

Activation of the transfer operon happens with a positive feedback loop involving TraM, TraJ and TraY (Figure II.4): TraJ activates the transcription of the *tra* operon. TraY, the first protein expressed from the *tra* operon, also activates itself the transcription of the *tra* operon, and *traM* expression.

This mechanism of fertility inhibition is thought to be responsible for the transient derepression known for F-like plasmids : new transconjugants are characterized by high frequency transfer, leading to an "epidemic spread" of plasmids in a population of naive cells [Lundquist and Levin, 1986]. This could happen because after transfer, FinO and FinP are not immediately present in new plasmid-bearing cells, so there is a huge transient expression of the *tra* operon.

F-like plasmids are thus characterized by a tight control of *tra* gene expression, leading to only a small part of the population expressing transfer functions. Other systems will behave differently: for instance, IncP plasmids have a low but constant level of pili expression [Zatyka and Thomas, 1998].

### Control of gene transfer by the host

Chromosomal genes are also involved in the regulation of transfer [Frost and Koraimann, 2010]. *Escherichia coli* mutants where pili production and transfer are inhibited have been discovered early [Beutin and Achtman, 1979]. Most of those genes are transcription factors, acting as global regulators of gene expression. IHF is a nucleoid-associated protein involved in relaxosome assembly and nicking by the relaxase, and also in *tra* gene expression. Another transfer control protein that was discovered early is ArcA: ArcA is a

response regulator of a two-component system involved in sensing the redox state of the cells, and also activates transcription from the *tra* operon promoter [Strohmaier et al., 1998].

Numerous nucleoid-associated proteins have been shown to interact with the expression of *tra* genes, often at the level of *traJ* expression (Figure 1.6). As mentioned previously, H-NS proteins globally repress the expression of AT-rich sequences, including the *tra* operon. The interaction of plasmid and chromosomal components is complex: some plasmids encode H-NS-like proteins; and TraJ actually seems to activate *tra* gene expression by inhibiting the silencing effect of H-NS [Will and Frost, 2006]. This regulation by host proteins is highly variable. Related plasmids are often not regulated in the same way by the proteins from the same host [Serna et al., 2010].

Many additional host factors play a role in plasmid transfer and are involved in sensing environmental changes. For instance, *cpxAR* is a two-component system, involved in the response to extracytoplasmic stress. Cpx stands for "conjugative plasmid expression", as *cpxA* was discovered because of mutations decreasing transfer ability. Factors such as CpxAR will thus allow a regulation of plasmid transfer by environmental conditions.

### **Regulation by environmental conditions**

Regulatory systems may also modulate plasmid transfer depending on various kinds of environmental signals. For instance, ArcA activates transfer of the *Salmonella typhimurium* virulence plasmid under microaerobiosis conditions, which mimic conditions of the mouse intestine [Serna et al., 2010]. This activation could be an adaptation allowing transfer in high density environments, where recipients are potentially abundant. Transfer is often regulated by the availability of nutrients: F transfer is upregulated in the presence of glucose for instance, and repressed in stationary phase [Frost and Manchak, 1998]. Temperature is also known to regulate transfer: R27, a resistance plasmid in *Salmonella typhi*, transfers optimally at 25°C, where repression by H-NS proteins is absent. At 37°C, H-NS proteins repress transfer [Forns et al., 2005]. Finally, transfer can be influenced by social cues: in *Agrobacterium tumefaciens* or *Rhizobium* species, transfer is regulated by quorum sensing [Zatyka and Thomas, 1998], that indicates high density of cells.

In conclusion, plasmid conjugation can be modulated by numerous factors. Some of these factors signal changes in the environment that probably are favorable to transfer, like the availability of recipients. Some factors are plasmid-encoded, while others are encoded by the host chromosome: the host has thus the ability to control plasmid transfer.

However, this does not imply that the observed regulation mechanisms necessarily benefit either the host or the plasmid. Indeed, the activation of a host protein could for instance be the result of exploitation by the plasmid, that diverts the function of this protein to its own interest.

### II.3.2 Regulation of transfer in recipient cells

The expression of transfer genes is regulated in the donor cell, but effective transfer also involves the recipient cell. First, mating-pair formation relies on contacts between donor and recipients, that depend on the recipient surface properties. Then, plasmid entry and maintenance in the recipient cell depend not only on its capacity to replicate and be maintained in the recipient cell, but also on recipient defense mechanisms. Finally, the regulation of transfer in the donor cell can itself depend on signalling involving recipients.

#### Chromosomal factors modifying plasmid entry

Recipient genes involved in plasmid entry have been discovered early, with reports of *Escherichia coli* mutants with reduced F recipient ability [Reiner, 1974]. Those mutants were sensitive to multiple stresses, suggesting membrane alterations. It was then shown that most of these strains were mutated in either OmpA or lipopolysaccharide (LPS) synthesis genes, each of them reducing 100-fold the cell recipient ability [Manoil and Rosenbusch, 1982]. Both OmpA and LPS are major cell envelope components. They are specifically recognized by TraN and TraG proteins of F plasmid during mating-pair formation [Anthony et al., 1999]. This recognition of cell-surface receptors will contribute to the host range specificity of plasmids (the range of hosts in which a plasmid is transmitted and maintained): broad host range plasmids can form mating pairs with various cell types, allowing transfer to distant species. The most extreme case is the transfer of T-DNA from *Agrobacterium tumefaciens* to eucaryotic cells [Bundock et al., 1995].

Thus, cell surface receptors control plasmid host range. However, inside a given species, mutations in these genes are strongly deleterious to the cell, reducing the possibility for variation in recipient ability in natural conditions. A high-throughput screening of *Escherichia coli* mutants, conducted in solid medium, concluded that no non-essential recipient genes had an essential role in conjugation [Pérez-Mendoza and de la Cruz, 2009]. LPS genes are actually not essential for mating-pair-formation in solid conditions. Still, a few genes, like the *nhaA* gene, a membrane antiport protein, affect recipient ability also in solid conditions. Other recipient factors actively involved in plasmid entry seem to be

rare, as the transport of DNA is essentially controlled by the donor cell.

Still, a mechanism actively limiting plasmid entry has been discovered very recently [Ho et al., 2013]. It is based on T6SS secretion systems, that are generally involved in antagonistic behaviours between bacteria: they act by delivering toxic effectors to neighboring cells. In *Pseudomonas aeruginosa* cells, one of these T6SS is activated in response to attack from neighbouring T6SS<sup>+</sup> cells. It was shown that the same T6SS acts against the conjugation of RP4 plasmid, in response to mating-pair formation. This strongly reduces plasmid conjugation rate (86% decrease compared to T6SS<sup>-</sup> cells), and imposes a strong cost on RP4-bearing cells, as donor cells are killed in the process.

### **Exclusion of plasmid entry by plasmid genes**

Actually, the most efficient system reducing recipient ability is not encoded on the bacterial chromosome, but on plasmids themselves: it was discovered early that F-bearing cells are not efficient recipients for conjugation [Lederberg et al., 1952]. This phenomenon is plasmid-specific: a plasmid prevents the entry of a similar plasmid. Part of this is due to plasmid incompatibility which I already described, but the main part of it, called entry exclusion, is an active barrier to the physical transfer of DNA between cells carrying closely related plasmids [Garcillán-Barcia and de la Cruz, 2008].

Entry exclusion can happen at two steps: mating aggregates stabilization (surface exclusion) and DNA entry in the cell (entry exclusion in the strict sense). In F plasmid, both surface and entry exclusion take place. Surface exclusion is due to TraT, an outer membrane protein that inhibits the stabilization of mating aggregates, decreasing transfer 10-fold. Entry exclusion is due to TraS, an inner membrane protein, that decreases transfer up to 1000-fold. TraS seems to inhibit transfer by interacting with the TraG protein present in the inner membrane of donor cells [Audette et al., 2007]. Entry exclusion has been interpreted as a mechanism that diminishes the competition between related plasmids [Garcillán-Barcia and de la Cruz, 2008]. Simulations demonstrated that a plasmid with entry exclusion can gain an advantage in competition against a plasmid without entry exclusion [van der Hoeven, 1985]. Still, entry exclusion does not completely prevent plasmid transfer. For instance, entry exclusion is reduced in stationary phase for the F plasmid, leading to the phenomenon of "F<sup>-</sup> phenocopies": F-bearing cells in stationary phase are efficient recipients, behaving similarly to F-free cells [Frost and Manchak, 1998].

## Defense mechanisms against horizontal transfer

While chromosomal mechanisms controlling plasmid entry into the cell are rare, there are numerous systems that actively defend the cell against the further establishment of plasmids (or, more generally, horizontally transferred DNA) (Figure II.5).

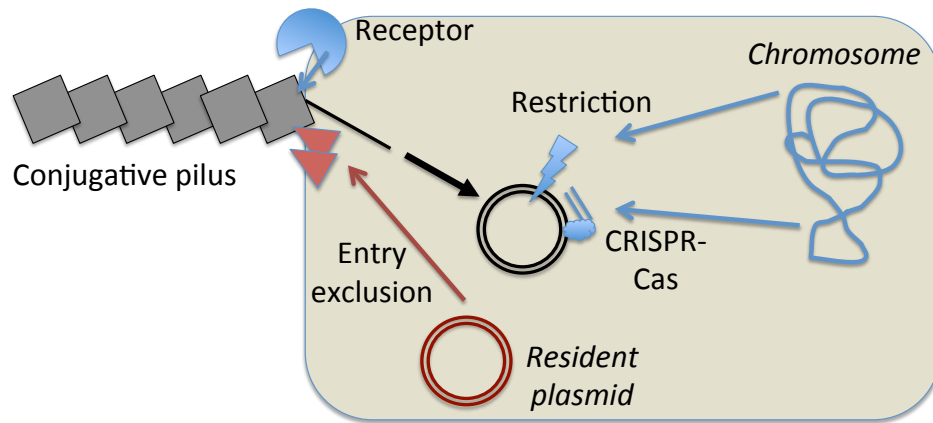


Figure II.5: **Summary of the main mechanisms controlling cell recipient ability.** The incoming plasmid and pilus are represented in black, defense mechanisms encoded on the chromosome in blue, and defense mechanisms encoded on a resident plasmid in red. Cell receptors and entry exclusion act at the plasmid entry stage; restriction and CRISPR-cas systems target plasmid DNA after circularization and replication.

Restriction-modification (RM) systems are composed of a restriction enzyme, an endonuclease that cleaves DNA after recognition of a specific sequence (the restriction site), and a modification enzyme, that protects the restriction site, hindering its recognition by the restriction enzyme, usually by methylation. RM systems decrease the efficiency of horizontal transfer when the transferred DNA contains restriction sites that are not protected by modification.

Plasmid transfer can be strongly decreased by the presence of RM systems in recipients: adding three restriction sites to the RK2 plasmid decreased its transfer efficiency by a factor of  $10^5$  [Guiney, 1984]. However, RM is not an efficient protection at an evolutionary timescale: a plasmid or phage that escapes restriction (if by chance modification happens faster than restriction) is then protected by methylation and can invade new hosts sharing the same RM system. Plasmids are also relatively protected against restriction, compared to phages, as they enter the cell through conjugation in a single-strand form that is not recognized by restriction enzymes. Still, RM systems seem to impose an evolutionary pressure against plasmids, as there are fewer restriction sites on plasmids than expected by chance [Guiney, 1984]. Finally, many plasmids also encode anti-restriction genes, located

in the region of the plasmid that is first transferred. They are rapidly expressed in the recipient cell, where their product protects plasmid DNA from restriction by binding to the DNA or by modulating RM activity [Tock and Dryden, 2005].

Another recently discovered defense mechanism against plasmids is the CRISPR-cas system. CRISPR stands for "Clustered, Regularly Interspaced Short Palindromic Repeat". CRISPRs are short repeats that are separated by spacers, sequences that match bacteriophage or plasmid sequences [Marraffini and Sontheimer, 2010]. The spacers specify the targets of CRISPR. They are transcribed, processed by Cas (CRISPR-associated) proteins, and then serve as guides for recognition of the target DNA. This recognition then leads to resistance against phages or plasmids that bear this target DNA, by mechanisms that are still poorly known [Barrangou et al., 2007, Marraffini and Sontheimer, 2008].

A key feature of CRISPRs is that spacers can be rapidly acquired upon infection by a phage or a plasmid, representing a form of acquired immunity. CRISPR presence and spacer content is thus highly variable among bacterial species and strains. For instance, F plasmid is not targeted by CRISPRs in the classical *Escherichia coli* laboratory strain K12, but a corresponding spacer is present in other strains [Westra et al., 2013].

Overall, a huge diversity in the effective recipient ability of bacteria is caused not only by a passive absence of cell recognition or plasmid entry, but also via active defense mechanisms against horizontal transfer.

### **The case of pheromone plasmids**

A peculiar case of plasmid transfer control is the case of sex-pheromone plasmids in *Enterococcus faecalis*, where the control of transfer involves both plasmid donor and recipient cells [Dunny, 2012]. In the model plasmid of this group, pCF10, a clumping reaction that leads to very efficient plasmid transfer can be observed specifically when donor cells and recipient cells (not bearing the plasmid) are mixed together. Clumping is due to the secretion of small peptides, the sex pheromones, by recipients. Donors then sense the presence of the pheromone and in response synthesize a surface adhesin, the aggregation substance that is responsible for clumping. The striking feature of this system is that sex pheromone genes are encoded on the chromosome, and pheromone synthesis is repressed by plasmids present in the cell (Figure II.6).

The *Enterococcus faecalis* chromosome actually encodes numerous pheromone genes, and each of them induces conjugation of a specific pheromone-plasmid family. Each plasmid encodes a corresponding inhibitor peptide, that represses the synthesis of the

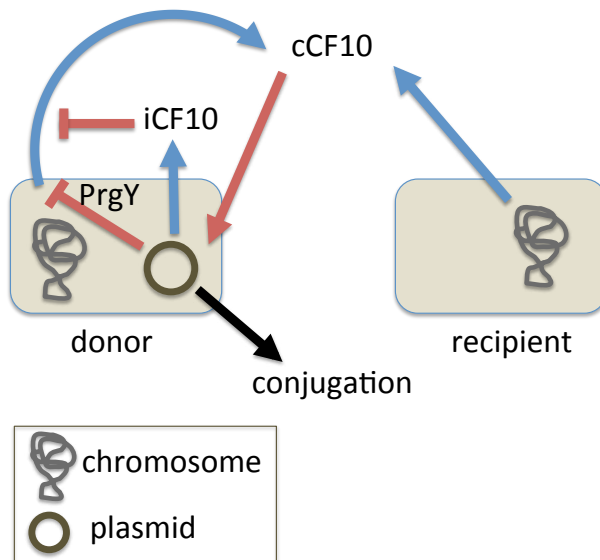


Figure II.6: **Regulation of pCF10 plasmid conjugation** (modified from [Dunny, 2012]). The plasmid-free recipient cell secretes the pheromone cCF10, encoded by a chromosomal gene; the donor cell contains the pCF10 plasmid, which bears genes coding for a conjugation system synthesized in response to the presence of cCF10 in the medium. The plasmid also codes for production of two negative regulators, PrgY and iCF10, which prevent self-induction of donors by endogenous cCF10. PrgY reduces the amount of cCF10 secreted by the donor cell, whereas iCF10 acts as a competitive inhibitor of cCF10.

pheromone in cells bearing this plasmid. The evolution of this system is scarcely known and deserves more investigation, but clearly involves a strong coevolution between plasmids and hosts. The existence of pheromones could be interpreted as an active mechanism for receiving plasmids in recipient cells, where receiving a plasmid would be beneficial (which implies that recurrent plasmid loss occurs in recipients). However, chromosomal pheromones could also be host proteins exploited as cues by selfish plasmids. This behaviour is, as far as we know, restricted to the species *Enterococcus faecalis*.

### II.3.3 Plasmid maintenance through horizontal transfer?

Plasmids bear numerous modules that favor their vertical and horizontal transmission. However, they also impose a cost on their hosts, reducing vertical transmission, and horizontal transmission can be limited for multiple reasons, including host resistance mechanisms. Thus, a key question is how and under which conditions are plasmids actually maintained in bacterial populations. Put differently, are plasmids parasitic or mutualistic symbionts? Are they maintained mainly because of infectious transfer, or thanks to the benefits they bring to their hosts?

#### Population dynamics of bacterial plasmids

Population dynamics of plasmids has been extensively studied with mathematical modelling, beginning with a seminal paper from Stewart and Levin [Stewart and Levin, 1977]. The model described the growth of bacteria bearing a conjugative plasmid in a chemostat, and showed that at equilibrium, a plasmid is maintained if the following relation holds:

$$\gamma N > \alpha \rho + \tau$$

where  $\gamma$  is the rate of plasmid transfer,  $N$  the equilibrium recipient density,  $\alpha$  the negative effect of the plasmid on host growth rate,  $\rho$  the rate of cell turn-over in the chemostat and  $\tau$  the plasmid segregation rate. Globally, this means that plasmid infectious transmission ( $\gamma N$ ) has to overcome plasmid loss through segregation and spreading of plasmid-free cells. Infectious transmission will be a function of plasmid transfer rate and cell density: high cell densities are favorable to transfer. This result also formalized the fact that a plasmid with little or no transfer ability will need to bring some advantage to its host to overcome segregation. Reversely, a plasmid with a given cost will need some minimal transfer ability to be maintained in the population as a parasitic element.

Plasmid rate of transfer is thus a key parameter of plasmid maintenance in populations. To be able to compare transfer among studies, where the cell densities, donors to recipients ratio and conjugation time can vary, the intrinsic plasmid transfer rate coefficient  $\gamma$  has been introduced [Levin et al., 1979]. Transfer is assumed to depend on random encounter between donor and recipient cells, following a mass-action law.  $\gamma$  is then defined as the proportionality constant that gives the number of new transconjugants over a given period of time, as a function of the densities of plasmid-bearing (donor) and plasmid-free cells (recipient) cells:

$$\dot{T} = \gamma \times D \times R$$

with  $T$  (transconjugants),  $D$  (donors) and  $R$  (recipients) expressed in cells  $\text{mL}^{-1}$ .  $\dot{T}$  is the change in the number of transconjugants in a unit of time,  $\gamma$  is thus expressed in  $\text{mL cell}^{-1}\text{h}^{-1}$ .  $\gamma$  can be estimated either with a short encounter of donors and recipients, or by sampling at the end of growth, which is called the end-point method [Simonsen et al., 1990] and relies on multiple assumptions, including a random distribution of cells, and both equal transfer and equal growth rate for all cells in the population.

The intrinsic plasmid transfer rate coefficient is widely used when experiments are coupled to modelling of plasmid transfer. However, experiments aimed at estimating the actual invasion of populations often measure more intuitive parameters, like the proportion of transconjugants compared to recipients ( $T/R$ ), or compared to donors ( $T/D$ ).

### **Diversity of transfer rates among plasmids and strains**

In the laboratory, transfer rates are usually estimated by following antibiotic resistance markers located on plasmids. With this method, R1 plasmid transfer rate measured among natural isolates of *Escherichia coli* was shown to vary between  $10^{-18}$  and  $10^{-11}$   $\text{mL cell}^{-1}\text{h}^{-1}$  [Gordon, 1992]. Other parameters concerning plasmid population dynamics were



roughly estimated, under the growth conditions of *Escherichia coli* in the gut [Simonsen, 1991]. With those estimations, R1 transfer rate should be at least  $10^{-10}$  cells mL<sup>-1</sup> in order for R1 to be maintained as parasitic DNA. It was thus concluded that the maintenance of plasmids as parasites is hugely unlikely.

However, both plasmid cost and bacterial densities are poorly known in reality, so lower transfer rates could be sufficient to maintain plasmids. Moreover, it has been shown that for the same plasmid R1, "amplifier strains" with high transfer rates can have a key role in plasmid maintenance, even if present in low frequencies [Dionisio et al., 2002]. Transfer ability was shown to vary over eight orders of magnitude among *Escherichia* species, and adding an amplifier strain allowed plasmid invasion of recipients even when starting from an initial donor strain with low transfer ability. The existence of amplifier strains with high donor or recipient ability is itself a striking fact that I will focus on later, but in any case it favors plasmid maintenance. Another feature favoring plasmid maintenance is the phenomenon of transitory derepression shared by many plasmids [Lundquist and Levin, 1986], which can allow epidemic spread of plasmids in a new population of cells. Finally, plasmid transfer will be strongly dependent on the environmental conditions, that will influence bacterial densities and the regulation of transfer.

### **Plasmid horizontal transfer *in situ***

In natural environments, many parameters influence both the amount of horizontal transfer and the actual spread of horizontally transferred plasmids, that additionally takes selection of transconjugants into account [Slater et al., 2008a]. Horizontal transfer is measured by following plasmid-encoded phenotypes after the introduction of donor strains in a given environment. One extensively used technique is to mark a plasmid with a fluorescent marker, allowing detection by fluorescence microscopy: plasmids marked with the green fluorescent protein (GFP) allow the detection of transconjugants with more sensitivity compared to antibiotic resistance markers, because this method removes the need for growth on selective agar plates [Hausner and Wuertz, 1999].

Studies of horizontal transfer *in situ* have generally shown the existence of "hot spots" for horizontal transfer: environments where the total number of gene-transfer events is high [Elsas and Bailey, 2002], usually soil or plant surfaces, or surfaces in aquatic environments, which are characterized by high nutrient concentrations. Transfer is favored by a high intrinsic conjugation rate, but also by high densities of both donor and recipient cells, and by a local spatial structure that favors interactions between cells. In environ-

ments with low cell densities, such as bulk soil, observed transfer frequencies are generally below  $10^{-5}$  transconjugants per donor cell. In conjugation hotspots, they can go up to  $10^{-1}$  transconjugants per donor cell.

Many studies have concluded that the kinetics of transfer is profoundly different in natural environments compared to *in vitro* studies: for instance, a plasmid was shown to transfer more efficiently in the gut than what could be measured *in vitro* [Netherwood et al., 1999]. A *Bacillus thuringiensis* plasmid that encodes genes for production of Bt toxin was found to invade up to 80% recipient cells in host insect larvae, despite low conjugation frequency in culture or in soil [Vilas-Bôas et al., 1998]. The regulation of transfer probably often involves unknown signals or cues, that indicate environmental conditions adequate for plasmid transfer, and are absent in laboratory conditions.

Finally, plasmid transfer can be favored in biofilms: biofilms are communities of bacterial cells associated with surfaces, generally enclosed in a extracellular matrix. They are stable environments characterized by a high density of cells, with high metabolic activity and numerous cell-cell contacts. Correspondingly, the frequency of conjugation has often been shown to be higher in biofilms compared to liquid environments [Molin and Tolker-Nielsen, 2003]. Conjugation can however be limited by the accessibility of recipients in biofilms [Reisner et al., 2012]. Strikingly, plasmid conjugation itself has been shown to promote biofilm formation [Ghigo, 2001], because of aggregation driven by conjugative pili. This can lead to a positive feedback, favoring both biofilm formation and horizontal transfer [Madsen et al., 2012]. The dynamics of transfer in biofilms is still not well understood, and would benefit from modelling studies that take into account the aggregation driven by plasmids.

### **Conclusion: plasmid parasitic maintenance through horizontal transfer?**

Rates of plasmid horizontal transmission are hugely variable across strains, plasmids and environments. Horizontal transfer was thought to be too low to allow maintenance of parasitic plasmids, but it can actually be greatly increased by mechanisms like transitory derepression in recipients, or the existence of a few amplifier strains. In some cases, horizontal transfer is indeed essential for plasmid maintenance in populations: for instance, an *Escherichia coli* plasmid was shown to decline in frequency in the gastrointestinal tract in the absence of conjugation. Conjugation allowed this plasmid to be stably maintained [Bahl et al., 2007]. Strikingly, numerous plasmids are not transmissible by conjugation or mobilization [Smillie et al., 2010]. They may transfer horizontally at low frequen-

cies through other mechanisms, like exchange through nanotubes or co-integration within other mobile elements. Still, absence of high transfer rates in these plasmids suggests that they do not behave as parasites and have very low net costs to their host cells.

When transfer is necessary for maintenance, transfer seems to be indeed a parasitic feature of conjugative plasmids. However, plasmid hosts can potentially control plasmid horizontal transfer, as numerous chromosomal genes are involved in plasmid conjugation and its regulation, and the same plasmid has hugely variable transfer rates, both among strains and species [Dionisio et al., 2002]. This should allow coevolution between plasmids and hosts to take place.

In a coevolutionary context, the existence of strains with very high transfer ability requires an explanation. High transfer ability could be due to two reasons: either the host is being manipulated by parasitic plasmids and high transfer ability has not actually been selected for, or plasmid transfer actually benefits the host, which could explain the selection of chromosomal genes favoring transfer. Plasmid transfer thus needs to be studied in an evolutionary way with specific attention to the following questions: (1) what are the selective pressures acting on it? and (2) how are they influenced by ecological conditions and plasmid-host co-evolution?

General knowledge about the evolution of cooperation, conflict and horizontal transmission suggests that both the fitness effect of plasmids on the host and the horizontal transfer of plasmids will be strongly affected by plasmid-host coevolution. The focus of my work will be the evolution of transfer ability, specifically viewed from the perspective of the host. In the next chapter, I describe what is currently known about plasmid-host coevolution, particularly concerning the evolution of transfer.

## Chapter III

# Plasmid-host coevolution

Plasmids have various rates of horizontal transmission and can bring both costs and benefits to their bacterial host. The net effect of all these factors can be an increase or decrease of the host fitness in presence of plasmids, leading to either mutualism or parasitism. Additionally, the evolution of the bacteria-plasmid symbiosis is affected by the level at which selection acts (within or among bacterial hosts), as well as by the horizontal transmission rate. The resulting bacteria/plasmid coevolution can classically be analyzed in terms of horizontal *vs* vertical transmission patterns described in Chapter I, seeing plasmids as bacterial symbionts. However, because plasmids transmit genetic information among hosts, the bacteria-plasmid symbiosis also has striking peculiarities we must consider. For example, coevolution also involves changes in the content of genes located on plasmids *vs* chromosomes. In this chapter, I will summarize the current knowledge of the forces shaping the coevolution between plasmids and their bacterial hosts.

## III.1 Evolution of plasmids and transmission pathways

I first focus on the classical patterns of horizontal *vs* vertical transmission, excluding social effects of symbiont transfer.

### III.1.1 Plasmid-host coevolution with vertical transmission

Pure vertical transmission of symbionts is thought to be rare, and lead to an alignment of interests between hosts and symbionts, as symbionts that do not provide a net benefit to their hosts will lead to the extinction of their host lineage [Ebert, 2013]. Concerning plasmids, while horizontal transmission can be absent or strongly reduced, vertical transmission is always widespread. For many plasmids, no conjugation mechanism (self-transfer or mobilization sequences) is known [Smillie et al., 2010]. These plasmids may still be transmitted horizontally through other mechanisms, like transformation, but probably at much lower rates. Lysogenic phages also undergo mainly vertical transmission (sometimes as plasmids, as in the case of P1 prophage).

### **Alignment of interests in the absence of transfer**

Plasmid-host coevolution in the absence of transfer has been extensively studied in experimental coevolution studies. These studies were motivated by the issue of antibiotic resistance, trying to predict if antibiotic resistant plasmids disappear in the absence of antibiotics selective pressure. Experiments usually start with a parasitic plasmid imposing a cost on its host, and have repeatedly found a decrease in plasmid costs to the host, with sometimes even evolution towards mutualism, where the evolved plasmid brings benefits to its host. Genes responsible for the change in plasmid effects can be located on the chromosome, [Bouma and Lenski, 1988], the plasmid [Modi et al., 1991] or both [Dahlberg and Chao, 2003, Dionisio et al., 2005].

Evolution can happen extremely rapidly: indeed, many studies did not even impose a positive selection pressure on plasmid maintenance, so adaptations were selected before plasmid loss occurred, even for plasmids lacking stabilizing mechanisms like PSK [Harrison and Brockhurst, 2012]. Adaptations can involve deletions of plasmid genes that are not beneficial in the conditions of the experiment: for instance, a large deletion including a tetracycline resistance gene was selected in the absence of tetracycline [Modi et al., 1991], removing the cost of resistance gene expression.

In the case of conjugative plasmids, a decrease in plasmid conjugation rate can also be observed [Dahlberg and Chao, 2003]. Here, horizontal transmission was theoretically possible, but not favored, as all cells initially carry the plasmid, effectively reducing transfer rates due to plasmid exclusion. Reducing costly horizontal transmission favored plasmid vertical transmission by favoring the host.

### **Within-host plasmid competition**

Despite the alignment of interests between the plasmid and its host due to vertical transmission, within-host competition is still possible if plasmid variants coexist in the cell. Competition is thought to be limited, as vertical transmission leads to successive bottlenecks in plasmid populations [Paulsson, 2002], decreasing the chances for long-term coexistence of variants. Moreover, selection pressures acting among cells are stronger than within cells: cell population size is generally higher than plasmid population size within cells, allowing more efficient selection. Additionally, heritability at the group level (groups of plasmids within cells) is very high, with reduced migration between groups (even with high rates of horizontal transfer by conjugation, vertical transfer is still more frequent at the cell level).

In the context of within-host competition, plasmid horizontal transmission through conjugation has two opposing effects. When conjugation happens to a plasmid-free cell, it creates a bottleneck [Paulsson, 2002], as only one plasmid molecule is transferred during conjugation, increasing relatedness between plasmids in each newly infected cell. However, when conjugation happens to a cell bearing a competing plasmid (which is rarer because of entry exclusion mechanisms), it increases the within-cell competition between plasmids. But horizontal transmission will also influence plasmid-host coevolution in other ways.

### **III.1.2 Evolution of plasmid horizontal transmission**

Horizontal transmission generally leads to a decoupling of hosts and symbiont interests, which can favor the evolution of parasitism. In the case of conjugation, horizontal transfer indeed can decouple plasmid and host interests, but also has other consequences on the host cell that I will mention briefly, linked to biofilm formation.

#### **A trade-off between horizontal and vertical transfer**

As we already saw, horizontal transmission by conjugation is costly to the host ([Turner et al., 1998]), which can decrease vertical transmission, materializing the trade-off between vertical and horizontal transmission described in section I.3.2. The selection of either vertical or horizontal transmission is then thought to depend on the accessibility of recipient, plasmid-free cells: in their absence, vertical transmission will be favored; with many plasmid-free cells, vertical transmission will decline, being a less substantial part of plasmid fitness.

When plasmid-free cells are present in a limited amount, vertical transmission is favored in the long-term: a competition experiment between R1 plasmid and its derepressed mutant R1-16 (with 1000-fold higher conjugation rate) showed first a faster invasion by R1-16, but then an increase in R1 frequency, as R1-16-bearing cells declined in frequency compared to R1-bearing cells, and no plasmid-free cells were left [Haft et al., 2009]. When the cost of horizontal transfer is high, transfer can also be counter-selected: for instance, the exposure to male-specific phages was shown to select for the loss of plasmid conjugative ability [Jalasvuori et al., 2011]. Enforcing various rates of plasmid horizontal transmission is complicated, as no direct selection of horizontal transmission is possible (there is no extracellular phase that can simply be selected for, like in the case of phages), and vertical transmission is always present. When the tradeoff between vertical and horizontal transfer was first confirmed [Turner et al., 1998], the authors did not observe an

effect on recipient availability on the evolution of horizontal transfer. This may be due to the fact that the effective proportion of plasmid-free cells is constantly changing during the experiment and is not always high. Indeed, mathematical models have shown that increasing the rate of horizontal transmission can actually lead to more vertical transmission at equilibrium, as the whole population is rapidly invaded [Lipsitch et al., 1995, 1996]. Predicting the evolution of either vertical or horizontal transmission based on the assumed availability of recipient cells is thus difficult.

Moreover, the relationship between plasmid horizontal transfer and virulence to the host is complicated by competition of plasmids within cells, which is also enhanced with opportunities for horizontal transmission. Particularly, when the density of plasmid-bearing cells is high, superinfection (secondary infection of a cell that already bears a plasmid) can evolve between plasmids and bypass entry exclusion mechanisms [Smith, 2011]. The mechanism of superinfection is not known, but superinfection has been shown to be highly deleterious to the host, leading to a tragedy of the commons with strongly reduced final host density [Smith, 2012].

Horizontal transmission thus decouples plasmid and host interests, leading to accrued costs to the host and indirectly allowing increased competition between plasmids.

### **Direct benefits of plasmid conjugation**

Despite its costs to the host, plasmid horizontal transmission can also benefit the host directly in some conditions. Indeed, plasmid conjugation very often enhances the formation of biofilms by host cells. Derepressed plasmids, that strongly express the conjugation machinery, were first shown to allow biofilm formation in an *Escherichia coli* strain unable to form biofilms in the absence of plasmids [Ghigo, 2001]. At least in the case of F plasmid, this is due to non-specific contacts, with conjugative pili acting as adhesion factors to both other cells and abiotic surfaces. Plasmids also influence biofilm formation in other ways. They often bear other genes promoting aggregation, as for instance clumping factors [Luo et al., 2005], or are involved in the regulation of chromosomal genes. For instance, R1 plasmid induces a stress response and decreases motility [Yang et al., 2008].

Moreover, conjugation by itself is necessary for the spread of biofilm-related genes, as it amplifies their expression. When plasmid-free recipients are present, epidemic spread takes place with transitory derepression of transfer. In many cases, synergistic biofilm formation is then observed when different natural strains are mixed, following conjugative plasmid transfer [Reisner et al., 2006].

In a chemostat-like environment, where sticking to surfaces is advantageous, biofilm formation will represent a strong benefit of conjugation for the host cell. Biofilms may also provide protection to cells, including protection against male-specific phages that target conjugative pili, as phages have less access to cells within biofilms [May et al., 2010]. As we saw in the previous chapter, biofilms themselves favor horizontal transmission through a high proximity between cells. Increasing biofilm formation could then be seen as a way to increase both vertical and horizontal transmission of plasmid genes.

Horizontal transmission strongly influences the coevolution between plasmids and their hosts, and leads to both costs and benefits to the host. Still, we have analyzed here coevolution in terms of the symbiosis between fixed host and plasmid entities, considering only one type of host. As horizontal transfer can happen between unrelated hosts, another consequence of horizontal transmission is plasmid host range.

### III.1.3 Horizontal transfer and plasmid host range

#### Variation in plasmid host ranges

Plasmid host range is defined as the diversity of host species in which a plasmid can be maintained [Bahl et al., 2009], and is hugely variable across plasmids. A narrow-host-range (NHR) plasmid is present in only a few bacterial species or genera, like IncF plasmids that are confined to *Enterobacteriaceae*. On the other hand, a broad-host-range (BHR) plasmid can be maintained in very different host species. IncP plasmids, for instance, are able to be transferred to and replicate in most Gram-negative species. BHR plasmids often carry specific adaptive traits, like multiple replication proteins used in different hosts [Adamczyk and Jagura-Burdzy, 2003].

Host range is generally thought to be limited by the plasmid ability to replicate and be maintained in distant hosts, rather than by the conjugation range itself. Indeed, conjugation can happen between very distant organisms: the most striking example is that of *Agrobacterium tumefaciens* T-DNA, that can be transferred by conjugation to plant cells. Even the NHR F plasmid can be transferred to *Pseudomonas* bacteria, but is not maintained after conjugation [Guiney, 1984]. Overall, the horizontal transmission is generally a successful process but does not always imply long-term maintenance of plasmids in recipients.



### Evolution of plasmid host range

Plasmid host range can evolve and change with only a few mutations, often concerning replication proteins, and either expand [De Gelder et al., 2008] or shift the host range [Sota et al., 2010]. The evolution of host range can be seen as the evolution of generalist (transferring between distinct hosts) or specialist plasmids (remaining in one host), with a probable trade-off between the maintenance in one host and the ability to replicate in others [Sota et al., 2010].

Plasmids that transfer frequently among distinct hosts will experience lower selective pressures for mutualism towards each of these hosts. Moreover, being able to infect different hosts can also benefit plasmids: for instance, the coexistence of hosts with varying transfer rates was shown to favor plasmid maintenance [Dionisio et al., 2002, Dionisio, 2005]. "Amplifier strains" provide efficient horizontal transmission, but suffer a high cost because of the presence of male-specific phages, which leads to a high risk of both strain and plasmid extinction when in isolation. Non-amplifier strains do not suffer such cost, so ensure plasmid long-term maintenance in populations despite low horizontal transmission.

In some cases, horizontal transmission happens for mobile elements that are in reality extremely specialized towards one host. This is the case of polydnviruses in insect parasitoids [Strand and Burke, 2012] or *Agrobacterium* plasmid T-DNA [Otten et al., 1992], both described in section I.3.3, that are transmitted to hosts where they are not stably maintained (host insects and plants, respectively): transfer here benefits the primary host in the short term but is not actually horizontal transmission in the evolutionary sense, as mobile elements are not transmitted again from secondary hosts.

Long-term coevolution of plasmids and hosts thus also involves the issue of specialization. Specialization can also be analyzed at the level of genes, asking why a given gene would be located on a plasmid or on the chromosome, which I will do in sections below.

#### III.1.4 The paradox of plasmid beneficial genes

Many plasmid genes are strongly beneficial to the host, at least in some environments. A major question is thus to understand why genes are located on mobile elements, and not on the chromosome [Eberhard, 1989, 1990]. Indeed, recombination between plasmids and chromosomes happens frequently. Genes have thus access to both plasmids and chromosomes on an evolutionary timescale. From the perspective of the host, beneficial traits should be integrated into the bacterial chromosome, which would remove the cost

of plasmid maintenance [Bergstrom et al., 2000]: cells with the beneficial gene integrated into the chromosome would outcompete cells where this gene is on the plasmid. Except for the genes that are involved in plasmid maintenance or infectious spread, the presence of additional genes on plasmids must be explained by specific properties of these elements.

First, one specificity of plasmids, compared to the chromosome, is the increased gene dosage. This can be selected when high gene expression is beneficial for the cell: for instance, high concentration of an antibiotic can select, in the short term, for increased copy number rather than rare promoter mutations [Thomas, 2004].

Secondly, another particular feature of plasmids is exactly their ability to transfer horizontally. Besides the direct advantage of horizontal transmission for the plasmid as a parasite, we must wonder whether the horizontal transmission could benefit either the host, or some accessory genes, and in turn explain the location of those genes on plasmids.

## III.2 Plasmids and the transfer of genetic information

The pool of genes that transfer horizontally by association to mobile genetic elements is not random: different sorts of genes that are preferentially transferred can be identified, in addition to the genes involved in plasmid maintenance and transfer.

### III.2.1 Classes of horizontally transferred genes

#### **Accessory genes: benefits in specific environments**

A common feature of many plasmid genes is that they bring optional benefits to the host in specific conditions or environments. This includes genes involved in antibiotic resistance, in the degradation of specific substrates, or genes responsible for the host specificity of pathogenic or mutualistic bacteria [Eberhard, 1989]. Accessory genes are thus often involved in local adaptation of bacteria. Benefits of carrying plasmids have indeed been shown to vary at short spatial scales and with environmental conditions in the case of a plasmid coding for mercury resistance [Slater et al., 2008b]. This observation has led to the vision of mobile genetic elements constituting a "communal pool" of genes [Norman et al., 2009] that are available, when needed, to procaryotic organisms in the environment. In the same way, plasmids were considered as "agents of open source evolution" [Frost et al., 2005]. Such plasmid role indeed has major consequences for bacterial evolution, but does not explain why accessory genes are specifically found on mobile elements.

At the gene level, one argument for the presence of accessory genes on mobile elements

is that mobility would allow them to persist without continued positive selection in a given local population or a given genotype [Eberhard, 1990, Bergstrom et al., 2000]. In a structured environment with only sporadic selection for accessory genes, they would be regularly lost from populations that are not under this selection (for instance, presence of antibiotics). However, if other populations are under selection at this point, the genes will be maintained and will be able to re-infect initial populations

Sporadic selection can explain the observation of accessory, transiently beneficial genes being located on mobile elements. However, it does not imply that the donor host directly benefits from transferring accessory genes: receiving them is indeed beneficial, but favoring competitors by providing them with a beneficial allele is not.

### **Genes involved in social interactions**

Plasmids and other mobile elements also frequently bear genes involved in social interactions between bacteria [Rankin et al., 2010]. Indeed, many emblematic social genes are known to be mobile. Antibiotic resistance is in many cases a social trait, when it involves degradation or inactivation of the antibiotic present in the environment: neighbouring bacteria that do not participate in inactivation can still benefit from it. The most prevalent example of social antibiotic resistance is  $\beta$ -lactamases, secreted proteins that degrade antibiotics extracellularly, and are predominantly located on plasmids [Barlow and Hall, 2002]. Spiteful interactions also involve mobile genes: bacteriocines are a widespread example, and are very often located on plasmids [Riley and Wertz, 2002].

Many genes involved in the interaction between symbiotic bacteria and their hosts are also located on mobile elements. For instance, the cholera toxin is located on a filamentous phage that is characterized by high rates of transduction in the gut [Waldor and Mekalanos, 1986]. *Bacillus thuringiensis* endotoxin gene, that causes host death, is plasmidic, with high rates of conjugation *in vivo* [Vilas-Bôas et al., 1998]. For mutualistic interactions, the trend is the same: in *Synorhizobium meliloti*, transferable plasmids bear all the genes involved in nodulation and mutualistic nitrogen fixation for the host plant [Blanca-Ordóñez et al., 2010]. Interactions between bacteria and hosts can usually be considered as social at the bacterial level, because the behaviour of a few cells influences the response of the host to many bacteria, not only the ones performing this behaviour.

The pool of genes present on mobile elements can moreover be analyzed quantitatively. The rigorous proof that some behaviour is social requires experimental work, but an indirect indication of the social property of proteins is given by their localization: proteins

are more likely to be involved in social interactions if they are secreted or present at the surface of cells, where they can be used by, or interact with neighbouring cells. A bioinformatic analysis of *Escherichia* genomes has revealed that genes predicted to encode secreted proteins are overrepresented on plasmids or mobile regions of the chromosome compared to non-mobile regions of the chromosome [Nogueira et al., 2009].

Genes involved in social interactions thus seem to have a strong evolutionary relationship with mobile genetic elements. Here, contrary to the case of local adaptation genes, gene transfer may have a beneficial effect on the donor cell, as it may modify the social behaviour of the neighboring, recipient cell.

### **Mobilization of chromosomal loci at a low frequency**

Finally, the existence of mobility mechanisms is in some cases linked to the random mobilization of chromosomal genes at a low level. This is sometimes due to the rare event of plasmid integration into the chromosome, as in the case of Hfr cells in *Escherichia coli*, or to the errors in packaging during bacteriophage production (see section II.1). However, some gene transfer mechanisms actually seem to have no other consequences than chromosomal gene mobilization. Gene transfer agents (GTA), for instance, are involved only in the transfer of random fragments of the chromosome [Lang et al., 2012] (see section II.1.2). As genes coding for GTA production cannot be packaged in a single GTA particle, this mechanism does not seem to be selected by allowing infectious spread of GTAs, but is more likely the consequence of actual gene exchange between host chromosomes.

We can conclude that gene transfer has the potential to benefit the host, in more or less direct ways, depending on the transferred genes: either because of the spread of beneficial genes, or by effects on bacterial sociality. I describe both hypotheses in more detail below.

## **III.2.2 Gene transfer benefits through obtaining novel genes**

Novel genes could benefit the host in two ways: first directly, if the host receives a beneficial allele through horizontal transfer, or more indirectly, if sexual recombination is beneficial in a given environment.

### **Fitness effects of receiving plasmids**

Recipient ability itself does not seem to be costly as reducing it requires active defense mechanisms that are probably costly (see section II.3.2). Still, recipient ability can present

a strong cost in very specific conditions, as recipients can die when donor cells are in excess, which has been called lethal zygotis [Skurray and Reeves, 1973]. Lethal zygotis is not well understood and is probably linked to membrane alterations. Cell death due to plasmid reception is also reduced when entry exclusion genes are expressed [Garcillán-Barcia and de la Cruz, 2008], but could still represent a significant cost if transfer is abundant.

More indirectly, receiving a plasmid can be either beneficial or costly for a bacterium, as determined by the plasmid net effect on fitness. If the plasmid is costly for the cell, defense mechanisms can be selected for. For instance, CRISPR-mediated immunity may be maintained in the presence of costly plasmids [Levin, 2010]. Conversely, when genes present on the plasmid benefit the host, defense mechanisms can be selected against. When artificial selection of conjugation events is applied, deletions in the corresponding CRISPR loci targeting the beneficial plasmid are found in the chromosome [Jiang et al., 2013]. More generally, a negative correlation has been observed between the presence of CRISPRs and multidrug resistance in *Enterococci* [Palmer and Gilmore, 2010].

The cell recipient ability, when it has short-term benefits for the focal cell, can thus be selected. Gene recombination, or sex, may also be selected in more indirect ways, which I detail below.

### **Horizontal transfer and the evolution of sex**

Sex can be defined in various ways in bacteria: it has been thought to encompass any DNA inheritance from a source other than the parent cell [Narra and Ochman, 2006], including all mechanisms for horizontal gene transfer. However, other definitions insist on the fact that the mechanisms considered should additionally be selected for due to the benefits of genetic exchange themselves, by opposition to side-effects of mobile elements infectious spread [Redfield, 2001]. Studies about the evolution of sex in bacteria generally focus on chromosomal recombination [Levin, 1988], which can still be a side effect of plasmid integration (the case of Hfr cells), or involve phage-like mechanisms (the case of GTAs).

Many theoretical studies have analyzed in which conditions gene recombination can evolve [Otto and Lenormand, 2002]. One effect of genetic recombination is to reduce linkage disequilibrium (non-random genetic association between alleles at different loci). This allows beneficial mutations that arise in different lineages to be associated in the same genome, reducing the phenomenon of clonal interference. Recombination can thus generate variability, by creating genotypes with new allele combinations. This is thought to be beneficial in changing environments, where variability allows faster adaptation [Becks and

Agrawal, 2012]. Benefits of sex have indeed been experimentally identified in bacteria. For instance, a naturally competent *Helicobacter pylori* strain adapted faster to a new environment than its non-competent mutant, generating a fitness advantage for competence [Baltrus et al., 2007]. Recombination mediated by F plasmid conjugation also increased the speed of adaptation in *Escherichia coli* populations [Cooper, 2007]. Benefits of sex can thus be an indirect benefit of the spread of mobile genetic elements.

Horizontal transfer benefits here not only mobile elements themselves, but also host bacteria. Still, benefits of sex do not explain why some genes are transferred at much higher frequencies than others, by association to mobile elements, particularly genes involved in social interactions. Here, it seems that gene transfer could benefit not only recipient cells, but also donors, as it can alter the behaviour of their neighbours.

### III.2.3 Gene transfer benefits through bacterial sociality

Transferring mobile elements can act on social interactions in two ways. Mobile elements can generally modify the fitness of recipient cells, which can participate to competition between bacterial strains. Moreover, genes controlling specific social behaviours can be expressed in recipients, with consequences for neighbouring donor or recipient cells.

#### Gene transfer and bacterial warfare

Mobile elements can have a deleterious effect on recipients, participating in bacterial warfare. Such effect has mainly been shown with bacteriophages characterized by lysis and lysogenization states. Bacteria that bear the lysogenic phage in an integrated form are protected from lysis. When a small fraction of the population undergoes lysis, the abundant produced virions can have a strong impact on susceptible competitor cells. Strains of *Salmonella enterica* bear multiple functional prophages, which were shown to confer a strong fitness advantage to strains that carry them in coculture with susceptible strains [Bossi et al., 2003]. In an invasive context, biological weapons are more efficient than toxins, as they are amplified on recipient cells [Brown et al., 2006]. However, the advantage of bearing prophages is only transient, as phages can become lysogenic in recipient cells, protecting them from further lysis [Gama et al., 2013].

When mobile elements act like biological weapons, their infectious transfer clearly benefits the initial host, at least in the short term. However, plasmids have not yet been shown to act as weapons. Plasmids would probably have a lower impact on the fitness

of recipients than phages, as they do not lead to cell lysis. Still, plasmids with a higher cost in recipients compared to donors could bring a competitive advantage to donor cells. One potential observation of bacterial competition through plasmids is the fact that the transmission of IncN plasmids by conjugation to *Klebsiella* recipients leads to the death of the recipients [Rodríguez et al., 1995]. However, the mechanisms of death and the selective pressures acting on it are not known and deserve more investigation.

### **Transfer of altruistic traits and bacterial cooperation**

Transfer of genes coding for altruistic behaviours to recipients could be another benefit of transfer for donor cells. As we have seen, the maintenance of public good production, like other cooperative behaviours, is threatened by the spread of cheaters, which benefit from the public good without producing it. It has thus been proposed that plasmid localization of many public good genes can be explained by the fact that horizontal gene transfer helps maintain cooperation. Two different explanations have been proposed:

(A) The first hypothesis relies on the enforcement of cooperation in previously non-producing cells by horizontal transfer [Smith, 2001]. A public good gene, encoded on a transferable plasmid, converts non-producer cells into producers after plasmid conjugation (Figure III.1, A and B). It should be noted that "enforcement" refers here to the change in genotype of the recipient cell following cooperation gene transfer, and not to its classical use in social evolution, that describes mechanisms favoring cooperation through manipulating costs and benefits of cooperative behaviours ([West et al., 2007b], see section I.1.2).

This hypothesis assumes that non-producing cells do not bear the corresponding plasmid. Plasmid transfer simply enforces cooperation through infectious spread, which compensates for the competitive disadvantage of the producing allele. However, cheater plasmids bearing a non-producing allele are not considered in this model, even though these plasmids will likely appear easily, for instance by deletion of the production allele from a producing plasmid. If cheater plasmids have the same, or higher, transfer rate compared to cooperative plasmids (Figure III.1-C), the infectious advantage of cooperative alleles disappears and cooperation is again threatened by the spread of cheaters, now at the plasmid level [Mc Ginty et al., 2011].

The consideration of cheater plasmids suggests that enforcement of cooperation by infectious plasmid spread is not a stable strategy, and does not explain the observed localization of public good genes on mobile elements [Nogueira et al., 2009].

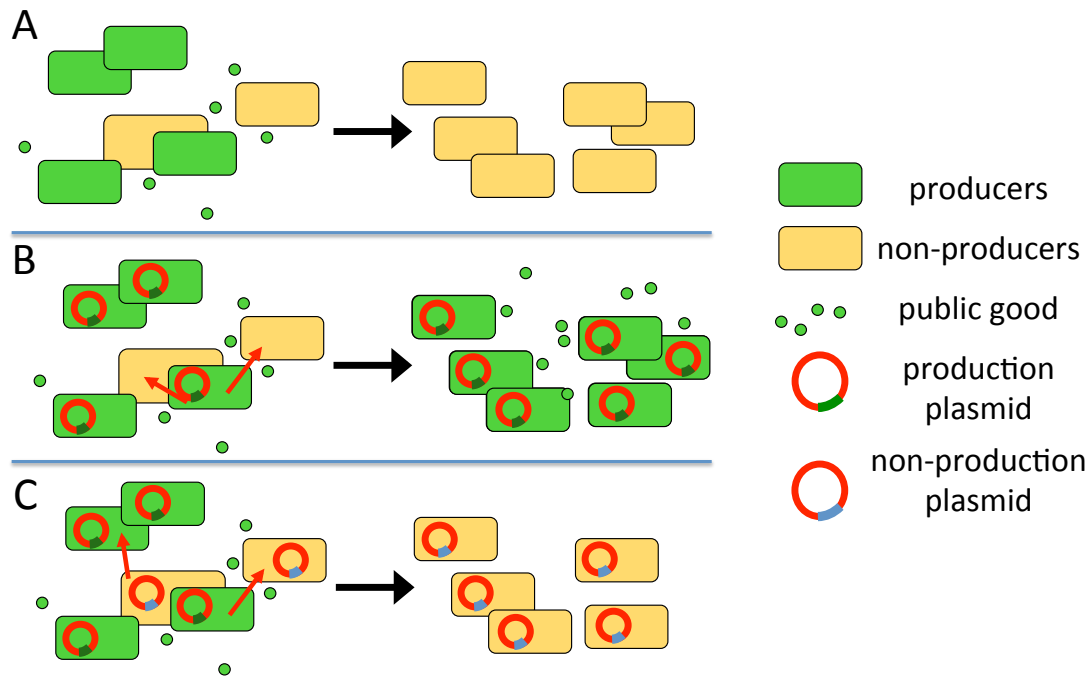


Figure III.1: **Enforcement hypothesis for the maintenance of cooperation**, adapted from [Smith, 2001]. **A:** in the absence of plasmids, producers (green) are outcompeted by non-producers (yellow). **B:** when the production allele (dark green) is located in a plasmid (red circle), it can be transferred to non-producers (red arrows), converting them to producers: producers are maintained. **C:** when both the cooperative (dark green) and cheater (blue) alleles are able to transfer horizontally (red arrows), benefits of infectious transfer cancel each other, and producers are outcompeted again.

(B) The second explanation focuses on social evolution theory: most explanations for the maintenance of cooperation rely on the existence of positive genetic relatedness between actors and recipients of the cooperative behaviour. Relatedness should thus be analyzed in the context of horizontal gene transfer. In the model we just described [Smith, 2001], there could be no positive relatedness, as the population was well-mixed. However, the second model studied a patch-structured population, where patches can differ in producer proportions, providing relatedness at the meta-population level [Nogueira et al., 2009].

The model follows relatedness, as well as the effect of horizontal transfer. When horizontal transfer is included in the model, it increases relatedness at the meta-population level, as it increases the probability that two individuals in a subpopulation carry the same allele by spreading this allele within the subpopulation. When any allele spreads by infectious transfer within patches of a metapopulation, this has the net effect of increasing the total variation in allele frequencies among subpopulations, which is equivalent to increasing relatedness. This is a general result, and does not specifically focus on cooperative or cheating alleles. Still, as increasing relatedness favors cooperation, horizontal



transfer should generally promote cooperation in a structured population.

More detailed population genetics models have recently been analyzed, including both population structure and producing/non-producing alleles, and confirming that both infectious transmission and indirect effects on relatedness could favor the maintenance of cooperative alleles by the two mechanisms described above [Mc Ginty et al., 2013].

### **Indirect selection of transfer through public good benefits**

The goal of the two models just described was mainly to explain why cooperation genes are located on plasmids, and to identify mechanisms that can maintain cooperation. However, the maintenance of cooperation can itself be interpreted as an indirect benefit of horizontal transfer: when cooperation is beneficial, this could allow transfer to be selected. This could be a form of second-order selection, via the effect of transfer on cooperation.

Transferring plasmids to another population, or another species, can also be beneficial if it makes recipients cooperate, even if the donor cell is not involved in the cooperative behaviour itself. This can be seen as a case of manipulation via plasmid transfer. An extreme example is the transfer of plasmid T-DNA to plant cells in *Agrobacterium tumefaciens*, where the transferred T-DNA induces the formation of tumours in plants, and manipulates host cells to produce molecules that benefit bacterial growth (Figure III.2). A segment of the virulence pTi plasmid, called the T-DNA, is excised and undergoes conjugation to cells of the host, where it is integrated into the genome and expressed [Otten et al., 1992]. The T-DNA segment bears genes coding for the production of opines, specialized molecules that then can be used for growth by bacteria, and also growth factors that enhance tumour formation [Platt et al., 2012].

Interestingly, bacteria do not contribute to the "cooperative" behaviour here, as the T-DNA genes are expressed only in host cells. However, transferring the T-DNA is itself a cooperative behaviour, as neighbouring cells that do not transfer it could still benefit from opines secreted by the plant.

The induction of cooperative behaviours in recipient cells, either from the same population (promoting cooperation) or in another strain or species (manipulation) seems to represent an indirect benefit of plasmid conjugation through gene transfer. This could provide a selective pressure for the evolution of horizontal transfer, from both chromosomal and plasmid genes sides.

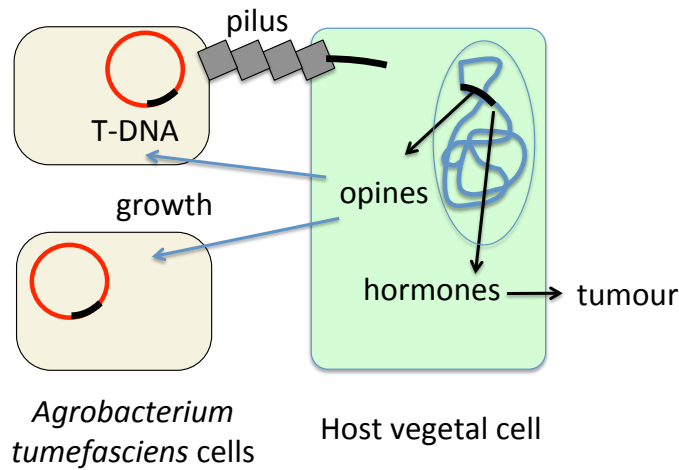


Figure III.2: **Consequences of T-DNA transfer by *Agrobacterium tumefaciens*.** The T-DNA is transferred through a conjugative pilus to host plant cells, where it integrates into the nuclear genome. The cell then synthesizes both plant hormones, that lead to tumour formation, and opines, that are used for bacterial growth.

In conclusion, the coevolution between plasmids and their bacterial hosts is strongly linked to the occurrence of horizontal transfer, which impacts the evolution of plasmids into parasites or mutualists. Moreover, selective pressures also act at the gene level to determine if genes are located on mobile elements or on the chromosome. Mobile elements leading to genetic exchange among hosts can provide benefits of sex and of access to a common gene pool to their hosts, including the possibility of changing the genotype of the host's social partners. The interactions between plasmids and their hosts create a challenge I will address in the next chapters: understanding and quantifying the selective pressures governing the selection of plasmid horizontal transfer exerted by the host.



## Chapter IV

# PhD project

In this work, I focus on the link between the selection of horizontal gene transfer and cooperation in bacteria. The observed bias in the pool of genes present on mobile elements [Nogueira et al., 2009] suggests an evolutionary link between horizontal transfer and cooperation. Modelling work has further shown that plasmid transfer could favor the maintenance of cooperative behaviours [Smith, 2001, Nogueira et al., 2009, Mc Ginty et al., 2013]. Moreover, the strong variability in transfer abilities in bacteria, on both donor and recipient sides, is controlled by plasmid as well as chromosomal loci. Because of this dual control, both will be able to respond to selective pressures acting on transfer rates. To understand the evolution of plasmid transfer, selective pressures thus have to be understood for both plasmids and chromosomes.

From the perspective of the host, transferring plasmids could be beneficial for various reasons. One of them is the maintenance of cooperation; another is apparent competition via parasites, in an inter-species context (where one bacterial species or strain would differ significantly, in terms of coevolution with the plasmid, from a competing strain). Here, I focus on an intra-species case, in the sense that I do not analyze interactions where donors and recipients strongly differ in plasmid cost, plasmid gene expression or other traits, but only in plasmid presence or plasmid transfer rate.

### IV.1 Questions

I first focus on the hypothesis that plasmid transfer promotes the maintenance of cooperation; in a second step, I analyze in which conditions transfer could be selected through benefits to the host cell.

#### IV.1.1 Can plasmid transfer promote cooperation?

This question has been addressed with numerous mathematical models that I have presented in section III.2.3 [Smith, 2001, Nogueira et al., 2009, Mc Ginty et al., 2011, 2013]. These models provide insight into the ways in which plasmid transfer could promote cooperation: either by direct enforcement of cooperation in recipient cells by transfer of

cooperation genes, or by an indirect effect on population structure, changing the relatedness of horizontally transferred genes. However, these models lack experimental validation. Particularly, they do not identify the rates of transfer that would be necessary to provide the expected effect on cooperation. Moreover, some key constraints influencing the process of conjugation in natural systems, like the physical ones, could be missed.

I thus aim here at testing the two hypotheses about the link between cooperation and conjugation experimentally. Besides experiments, I will also use models to help understanding population dynamics. By coupling the two research methods, I will test the following two hypotheses (Chapter 2): (1) can transfer simply promote cooperation by enforcement? and (2) can the indirect effect of transfer on relatedness provide for a more stable maintenance of cooperation than (1), and under which conditions?

### **IV.1.2 Can cooperation promote plasmid transfer?**

The second part of this work is to study if plasmid transfer can be selected because of benefits to the bacterial host. This novel hypothesis has not been investigated before, either in experimental or in modelling studies. We wish to understand how and when plasmid transfer could be beneficial through its effect on the maintenance of cooperation, and if this could in return select for transfer itself, despite potential transfer costs.

Plasmid horizontal transfer is obviously directly beneficial from the perspective of the plasmid: it helps the plasmid to gain access to new hosts (unless the trade-off with vertical transmission strongly disfavors horizontal transmission). Our hypothesis thus has to be addressed separately in the cases of plasmid and chromosomal genes. Can chromosome genes promoting transfer be favored at all? Can plasmid genes promoting transfer be more strongly selected in the context of cooperation? Finally, what happens when selection acts simultaneously on cooperation and on gene transfer? I will first focus on the selection of chromosome genes controlling the transfer of a given plasmid, distinguishing between donor and recipient abilities (Chapter 3). I will then analyze the simultaneous selective pressures acting on transfer and cooperation genes from both plasmid and chromosomal sides (Chapter 4).

## IV.2 Strategy

### IV.2.1 Experimental approach

To answer these questions, I will separately manipulate both cooperation and conjugation in our system. We chose a synthetic approach, where we independently control public good production and plasmid transfer. The synthetic system I adopted and modified is described in Chapter 1: it is a combination of public good production genes from *Pseudomonas aeruginosa*, and the transfer control system of *Escherichia coli* F plasmid.

This synthetic system is not a truly natural system, but a combination of well-studied natural components. The main benefit of the synthetic approach is the reduction in the risk of unwanted interactions: we can then study the effect of precisely the factors we are interested in. Indeed, unwanted interactions could be especially prevalent in natural systems: the maintenance of cooperation is influenced by many factors, particularly by the existence of complex regulatory mechanisms in cooperative strains [Kümmerli et al., 2009b, Xavier et al., 2011]. Plasmid transfer is also strongly regulated by host and environmental factors [Frost and Koraimann, 2010]. A given cooperative behaviour and plasmid transfer could thus be regulated by the same signal in a natural system. This could be due to independent reasons (the signal indicating environmental conditions that are favorable to both transfer and cooperation), or precisely to previous co-evolution of transfer and cooperation, but distinguishing between both cases in laboratory conditions would be difficult. Concerning horizontal transfer, naturally occurring plasmids carry genes involved in many traits, which could lead to side-effects unrelated to the hypothesis we want to test. Moreover, adding fluorescent markers to the strains and plasmids used allows us to precisely follow both transfer and competitors. With this system, I compete strains or plasmids that differ in public good production or transfer ability, but not in other traits except fluorescence.

Finally, the experimental set-up involves either well-mixed populations, or a simple meta-population that consists of patches of well-mixed populations with varying proportions of competitors, providing non-random population structure and following the approach of [Chuang et al., 2009].

### IV.2.2 Modeling approach

First, I created a detailed model of the population dynamics of bacterial strains, plasmids and transfer (Chapters 2, 3 and 4). I used numerical solutions of differential equation

systems that closely follow the experiments, and are parametrized based on experimental estimations. Using such modeling approach allows for a deeper understanding of the system dynamics and the effect of transfer on it. We are also able to vary the transfer rates and other properties of transfer in a more flexible way than in the experiments, and avoid or control for unwanted side-effects that still are unavoidable in experiments.

In addition to this detailed model, I constructed a more general one capturing the selective pressures acting on chromosomal genes controlling transfer ability, in order to broadly analyze the conditions that allow for selection of transfer and its interaction with the evolution of cooperation (Chapter 3), which have not yet been described.

# Results





# Note about the structure of the results

My work has involved both mathematical modelling and experiments. The experimental synthetic system is generally common to all experiments. I first describe it in Chapter 1, and detail the different parts of the system I have constructed and the experimental protocol that I developed.

The three last chapters (2, 3 and 4), presenting both experimental and modelling results, have been written as independent chapters, as I am in the process of submitting them for publication as separate articles. For this reason, some explanations are repeated in all of them, namely:

- Chapter 2 and Chapter 3 contain experimental results, and as such each of them describes some of the associated strains, plasmids and protocols already presented in Chapter 1.

- In Chapters 2, 3 and 4, mathematical simulations that closely follow the experimental system are presented. Although presented independently, all the simulations are closely related, and share much of the nomenclature.

Finally, additional experimental details are presented in Annexes: strains and plasmid construction (Appendix A), identification of strains, plasmids and transfer by plating and flow cytometry techniques (Appendix B).



## Chapter 1

# A synthetic system for plasmid transfer and public good production

To study the selection of both cooperation and horizontal plasmid transfer, we need a genetic system with the following components:

- (1) production and non-production alleles controlling the production of a public good
- (2) transfer and no-transfer alleles controlling the transfer of the production and non-production alleles. Transfer ability genes can be on the transferred plasmids (plasmid control of transfer), or on the chromosome of the initial plasmid-bearing strains (chromosomal control of transfer). Chromosomal genes can also control recipient ability.
- (3) a way to distinguish strains and plasmids, evaluate fitness and follow transfer.

To create such a system, I started from an existing synthetic system, constructed to study public good production [Chuang et al., 2009]. I added plasmid transfer and fluorescent markers to follow multiple genotypes and transfer, and obtained the synthetic system I used in all my experiments, which I describe in this chapter.

## 1.1 Public good production and benefits

### 1.1.1 Description of the system

The published bacterial cooperation system [Chuang et al., 2009] consists of two strains of *Escherichia coli*, called producers (P) and non-producers (NP) (Figure 1.1). Producers constitutively express the RhlI enzyme, which produces the quorum-sensing auto-inducer C<sub>4</sub>-HSL (N-butyryl-L-homoserine lactone). C<sub>4</sub>-HSL is a membrane-permeable molecule that diffuses in the medium, and is able to induce gene expression in cells expressing the *rhlR* gene. Both P and NP cell types express RhlR, and also bear a chloramphenicol-resistance gene under control of the  $P_{rhl*}$  promoter, localized on the chromosome. C<sub>4</sub>-HSL binds to RhlR and induces resistance to low concentrations of the antibiotic chloramphenicol (Cm) in both P and NP cells. P cells can be identified by fluorescence, because the public good production allele is linked to the GFP coding sequence, and both are located on a plasmid that we call P<sup>+</sup>GFP. NP cells also bear a control plasmid, representing the

non-production allele, that we call  $P^-$ . Both plasmids share the same backbone with SC101\* origin of replication and a kanamycin resistance marker, and are generally referred to as pZS plasmids (arising from the pZ vector system [Lutz and Bujard, 1997]).

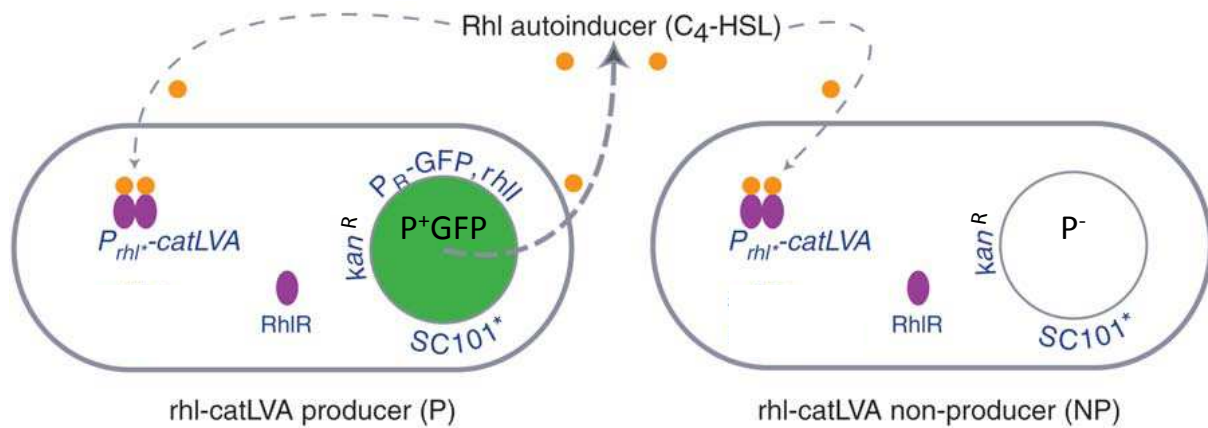


Figure 1.1: **Original producer and non-producer strains** (adapted from [Chuang et al., 2009]). The RhlI enzyme synthesizes the Rhl autoinducer  $C_4$ -HSL, which by binding to the RhlR receptor induces expression of the *catLVA* sequence, coding for Cm resistance, from the  $P_{rhl*}$  promoter. In P cells, the *rhlI* gene is linked to GFP and localized on  $P^+$ GFP plasmid; NP cells bear a control  $P^-$  plasmid. Plasmids have SC101\* replication origin and are maintained with kanamycin resistance genes ( $kan^R$ ).

### 1.1.2 Properties of the system

#### Expected properties of a cooperation system

This system has all the general characteristics of a public good system:

First, the public good production is beneficial at the population level (Figure 1.2-A): the growth of a culture in the presence of Cm is strongly positively correlated with the proportion of producers in the culture.

Secondly, producing the public good and GFP has a cost: producers decline in frequency within populations. This cost corresponds to a 1.04 to 1.05 increase in NP frequency per generation, in the absence of Cm [Chuang et al., 2009].

Finally, Chuang et al. showed that cooperation can be maintained in this system by controlling population structure: they initiate a metapopulation that consists in 10 subpopulations with increasing proportions of P. The correlation between the growth of subpopulations and producer proportions within subpopulations leads to a global increase in producer proportion when pooling subpopulations, despite the local decrease in each population (Figure 1.2-B). This result can be analyzed either in terms of multilevel selec-

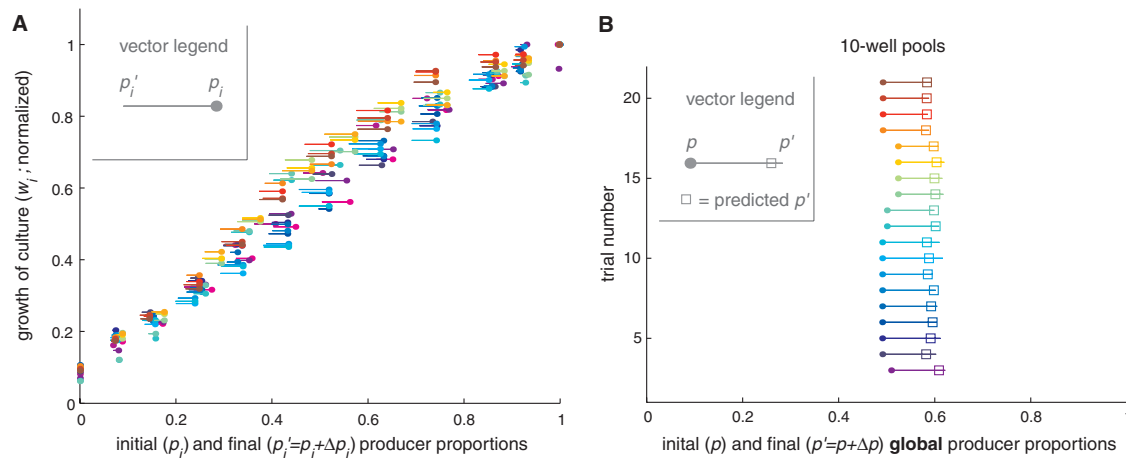


Figure 1.2: **Benefits and costs of public good production** from [Chuang et al., 2009]. Subpopulations containing producers and nonproducers mixed in 12 different proportions were first grown in the absence of Cm and then diluted 100-fold into Cm at time 0. **A:** Normalized growth of subpopulations in the presence of Cm as a function of initial producer proportion. Dots indicate initial producer proportion  $p_i$  in each subpopulation  $i$ ; the end of each line indicate final producer proportion  $p'_i$ . **B:** Initial ( $p$ ) and final ( $p'$ ) proportions of producers after pooling all non-pure populations represented in A, squares being predicted  $p'$  values calculated by using the data from individual subpopulations of (A). Each color represents an independent trial.

tion - individual costs and group benefits, leading to Simpson's paradox [Chuang et al., 2009] (see section I.2.3) - or in terms of Hamilton's rule - relatedness arising from the distribution of initial proportions within groups [Chuang et al., 2010].

### Pre-induction necessary to obtain public good benefits

To enable the use of the Chuang et al. system in our work, we further characterized its behavior. A key constraint is that a long pre-incubation with producers of  $C_4$ -HSL is necessary to induce Cm resistance in non-producers before adding Cm (pre-induction step): if producers and non-producers are mixed only when confronted to the antibiotic, producing  $C_4$ -HSL presents a direct advantage instead of a cost. Indeed, producers have already been confronted to  $C_4$ -HSL and expressed the Cm-resistance enzyme. Non-producers have not, and cannot directly grow as rapidly as producers in the presence of Cm.

We tested the influence of pre-induction time on P relative fitness when mixtures of P and NP were launched in the presence of Cm. 4 hours of pre-induction are necessary to observe a net cost to producers (Figure 1.3). This result is likely a combination of two effects: production of  $C_4$ -HSL by producers depending on P density, then induction of gene expression in non-producers.

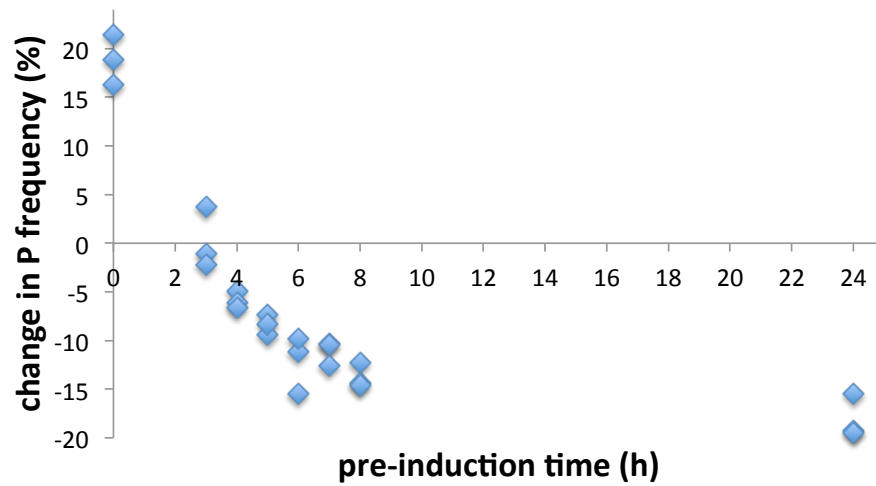


Figure 1.3: **Effect of pre-induction time on P vs NP fitness.** Stationary phase cultures of P and NP were mixed equally with a 100-fold total dilution during a varying pre-induction time. They were then diluted again by 100-fold and  $6.25\text{ng}\cdot\mu\text{L}^{-1}$  Cm was added. The change in frequency of P after 12h of growth in the presence of Cm, at  $37^\circ\text{C}$ , is shown as a function of pre-induction time.

### Transient benefits of cooperation

Another feature of the system is that cooperation acts on the growth rate and not on the final yield: when the concentration of Cm used is low (such as the  $6.25\text{ ng}\cdot\mu\text{L}^{-1}$  we use), cells manage to grow slowly even in the absence of  $\text{C}_4\text{-HSL}$ . Thus, the benefits of cooperation are transient: after enough time, all cultures will reach the same final density, so the correlation between producers proportion and culture growth (that leads to Simpson's paradox) will be lost. We tested this effect by analyzing the kinetics of the producer ratio.

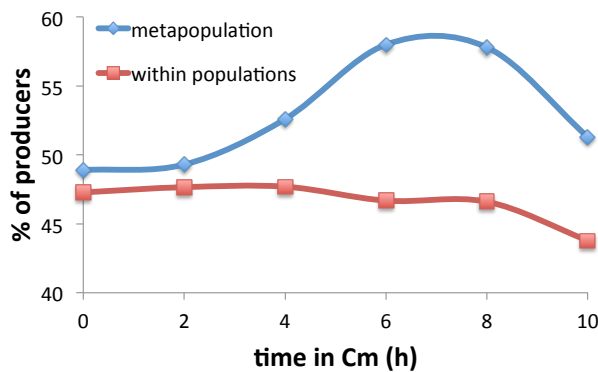


Figure 1.4: **Transient benefits of cooperation.** The frequency of producers at the metapopulation scale (10-well pools, blue) and within subpopulations (red), with the same population structure as in [Chuang et al., 2009] (see Figure 1.2), is shown as a function of time in the presence of Cm at  $37^\circ\text{C}$ , after 12h of pre-induction.

We confirm that the selection of cooperators is indeed transient, peaking around 7h after the dilution in Cm-containing medium at  $37^\circ\text{C}$  (Figure 1.4). After this, subpopulations with fewer producers manage to reach the same densities as populations with more producers, and the benefit of production disappears.

## 1.2 Plasmid transfer

### 1.2.1 Design of the system

As the cooperation system is based on growth in a metapopulation, with well-mixed subpopulations that ensure efficient diffusion of the public good within them, the transfer system should work efficiently in liquid. We chose to work with the F plasmid, one of the best-studied *Escherichia coli* plasmids [Frost et al., 1994], that transfers by conjugation at high rates in liquid cultures.

#### Mobilization of *oriT* in *trans*

F conjugation machinery recognizes a specific sequence, the origin of transfer (*oriT*), where it initiates transfer. *oriT* is normally localized on the F plasmid itself, but another plasmid bearing *oriT* can be mobilized by the F machinery [Johnson et al., 1981] (see section II.1.3). Based on this, we decided to insert *oriT* in the two pZS plasmid types ( $P^+$  and  $P^-$  described above, additionally containing fluorescence genes I will detail in section 1.3: either GFP or YFP), creating respectively  $P^+ oriT$  and  $P^- oriT$ , and provide the conjugative machinery separately, with a helper F plasmid (Figure 1.5). F plasmid should also be unable to transfer itself, as F transfer to recipients could affect the system and the outcome of competitions. This can be achieved by mutating F plasmid own *oriT* sequence (Figure 1.5).

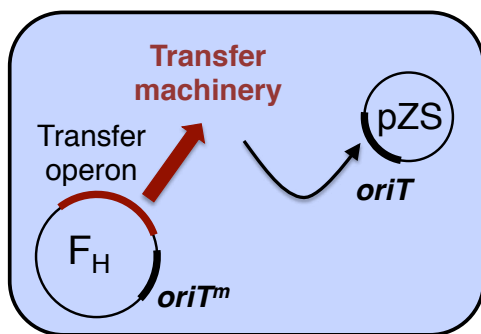


Figure 1.5: **Design of the system: mobilization of pZS plasmids by a mutant F plasmid.**  $F_H$  plasmid is a helper F plasmid bearing *oriT<sup>m</sup>*, a mutant *oriT* not recognized by the F conjugation machinery coded by F transfer operon. The transfer machinery acts only in *trans*, and mobilizes the pZS plasmids bearing the wild-type *oriT* (black arrow).

#### Mimicking plasmid or chromosomal control of donor ability

To mimic plasmid control of transfer (Figure 1.6-A), we compete pZS plasmids with or without *oriT*, in experiments where all cells bear a F plasmid. Plasmids with *oriT* ( $T^+$ ) can be transferred, plasmids without *oriT* ( $T^-$ ) cannot. For recipients to first receive transferred plasmids, entry exclusion needs to be inactivated by a deletion in *traS* gene. We call the F plasmid  $F_{HR}$  (helper plasmid with recipient ability).



Chromosomal genes known to regulate plasmid transfer are mainly global regulators [Frost and Koraimann, 2010] (see section II.3.1), so mutating them would affect not only transfer but also other functions of the cell. We mimic a chromosomal control of donor ability without modifying the chromosome. To do so, we use one of two experimental setups that will be used in Chapter 3: in the first one (Figure 1.6-B), donor cells can mobilize pZS plasmids to both donor and non-donor cells (transfer can happen within kin, see Chapter 3), as none of the cells express entry exclusion genes. Transfer ability is controlled by the presence of  $F_{HR}$  plasmid in the donor cells. In the second setup (Figure 1.6-C), transfer can happen only to a separate recipient strain (transfer does not happen towards kin, see Chapter 3), as both donor and non-donor strains bear the non-mobile F helper plasmid that expresses functional entry exclusion, called  $F_{HE}$ . Donor strains bear a pZS plasmid with  $oriT$ , so can mobilize it to recipients. Recipients are a separate strain unable to transfer the plasmid, as they do not bear F plasmid.

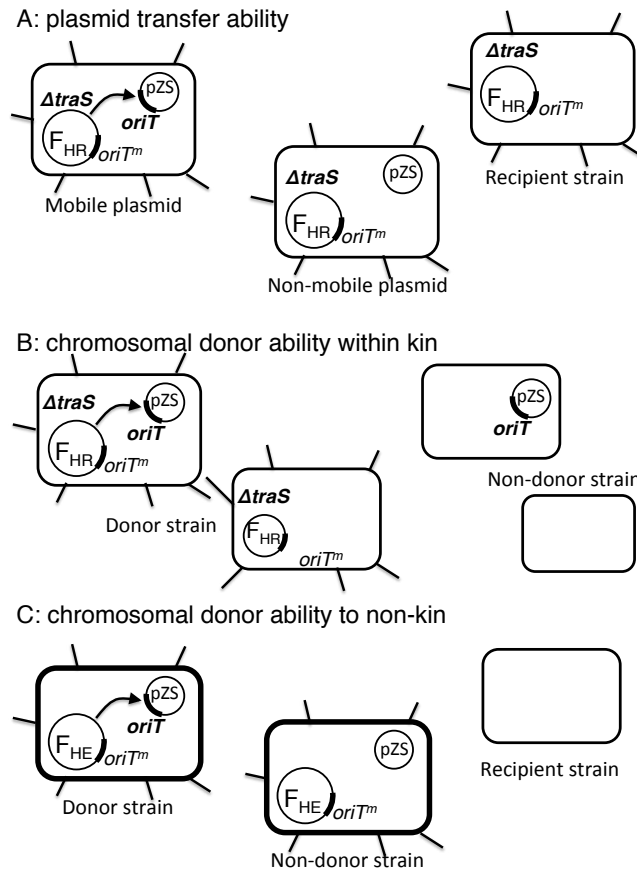


Figure 1.6: **Control of donor ability.** Black arrows represent mobilization of pZS plasmids by F conjugation machinery; thick black cell membranes represent functional entry exclusion; pili represent the presence of F conjugation machinery in the cell. **A: Plasmid control of transfer.** Plasmid transfer is controlled by  $oriT$  presence on the pZS plasmid.  $traS$  mutation suppresses entry exclusion and allows transfer to recipients bearing  $F_{HR}$ . **B: Chromosomal donor ability within kin.** Donor ability is controlled by  $F_{HR}$  presence in the cell, all pZS plasmids bear  $oriT$ . **C: Chromosomal donor ability to a separate recipient.** All transfer happens from the initial donor strain bearing  $F_{HE}$  plasmid and a pZS plasmid with  $oriT$  to recipients.

### Mimicking chromosomal control of recipient ability

To control recipient ability, we use again F entry exclusion (see section II.3.2). Recipients bear  $F_{HR}$  with deficient entry exclusion, non-recipients (with strongly reduced recipient ability) bear  $F_{HE}$  (Figure 1.7).

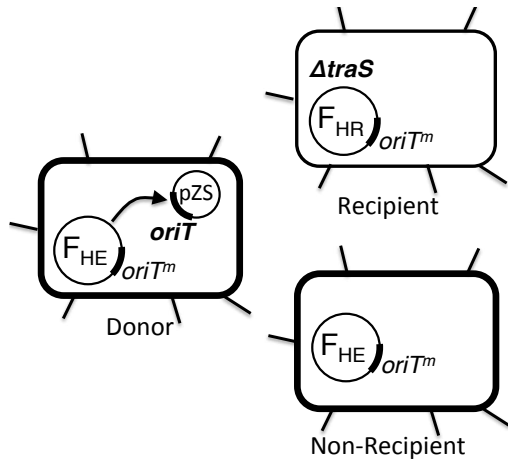


Figure 1.7: **Control of recipient ability.** Black arrows represent mobilization of pZS plasmids by F conjugation machinery; thick black cell membranes represent functional entry exclusion inhibiting plasmid entry into the cell; pili represent the presence of F conjugation machinery in the cell. Donors and non-recipients bear  $F_{HE}$ , recipients bear  $F_{HR}$ . Donors additionally bear a pZS-*oriT* plasmid, which can be transmitted mainly to recipients.

### 1.2.2 Mobilization of plasmids

The F plasmid should mobilize plasmids bearing *oriT*, without transferring itself. To this end, I designed F plasmid mutants with different versions of *oriT*, and tested both their ability to mobilize pZS-*oriT* plasmids and their own transfer rate.

#### Mutant *oriTs*

A mutant of F with deletion of the whole *oriT*, F-*oriT::frt*, was found to mobilize other plasmids with low efficiency (see Figure 1.9 left). The deletion of *oriT* probably alters expression of the nearby transfer operon (see Figure II.4). Thus, I then constructed mutant *oriTs* that are close to the wild-type sequence, but should be poorly recognized by the transfer machinery (Figure 1.8). First, the origin of transfer of R100 plasmid

F-oriT	C	A	C	C	A	C	A	C	C	C	C	A	C	G	C
F-oriT-R100	C	A	C	C	A	C	A	C	A	C	T	A	C	G	C
F-oriT-7	C	A	C	C	T	C	A	C	C	C	C	A	C	G	C
F-oriT-78	C	A	C	C	T	C	A	G	C	C	C	A	C	G	C

Figure 1.8: **Sequences of F *oriT* mutants tested for mobilization efficiency.** The red line indicates the nicking site; deviations from the wild-type F *oriT* sequence are highlighted in mutant sequences.

(from the same family that F) differs from F-*oriT* by only 2 nucleotides near the nicking

site, which was shown to lead to 1000-fold reduction in transfer by F machinery [Harley and Schildbach, 2003]. The other sequences F-*oriT*-7 and F-*oriT*-78 were also shown to decrease the F-relaxase binding affinity in structural studies [Stern and Schildbach, 2001].

### Mobilization efficiency

All mutant F plasmids except the one with *oriT* deletion mobilized pZS-*oriT* plasmids at a high rate in liquid (Figure 1.9). Moreover, their self-transfer ability was strongly decreased compared to the wild-type F plasmid. We chose to use the F-*oriT*-78 mutant as F<sub>HE</sub>: its self-transfer is near the detection limit (the measured transfer rate corresponds to only a few colonies recovered after 30-min mating at high densities) and 1000-fold smaller than its mobilization efficiency for pZS-*oriT* plasmids.

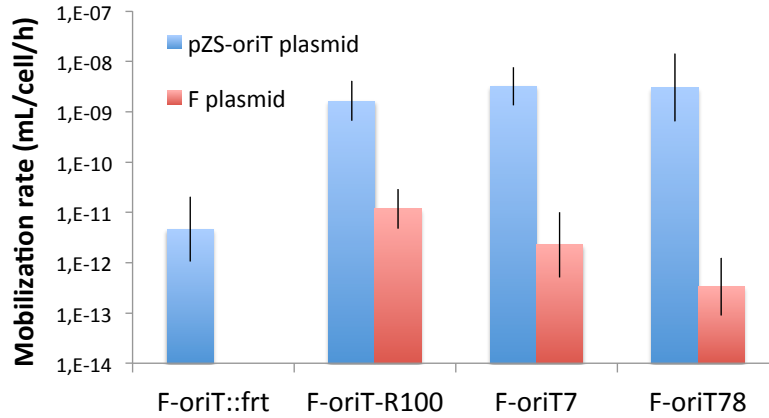


Figure 1.9: **Mobilization efficiency by F mutants** to the control recipient strain MG Nal<sup>R</sup> Rif<sup>R</sup>. Results are a geometric mean of at least 7 replicates. Results obtained for the two pZS-*oriT* plasmids (P<sup>+</sup>GFP*oriT* and P<sup>-</sup>*oriT*) were pooled for each F mutant, as they did not differ significantly from each other. Geometric standard deviations are shown. F-*oriT*::*frit* self-transfer was never detected.

Interestingly, self-transfer of F<sub>HE</sub>, already very low when in association to pZS-*oriT* plasmids, is completely undetectable when in association to pZS plasmids, *ie* in the absence of the wild-type *oriT* sequence in the cell (data not shown). The association of the relaxase to the wild-type sequence probably favors the transfer of mutated sequences.

### 1.2.3 Entry exclusion in recipients

#### Entry exclusion caused by F<sub>HE</sub>

The mobilization rate is strongly reduced when recipients bear F<sub>HE</sub> compared to recipients without a F plasmid (Figure 1.10). This reduction leads to only a few percent of plasmid invasion overnight, compared to almost full invasion in its absence.

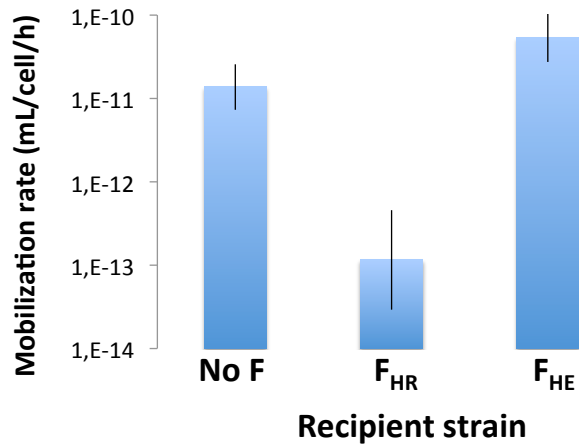


Figure 1.10: **Entry exclusion caused by F plasmids.** The mobilization efficiency of  $P^{+}YFP-oriT$  plasmid to the control recipient strain MG  $Nal^R$   $Rif^R$  was measured with  $F_{HE}$  in donor cells and with or without  $F_{HE}$  or  $F_{HR}$  in recipient cells. Results are the geometric mean of 3 replicates, standard deviations are shown.

This result implies first that transfer between cells containing F plasmids will be strongly reduced if plasmids bear functional entry exclusion systems. This allows us mimic chromosomal control of transfer to non-kin, as described before (Figure 1.6-B). Reversely, entry exclusion will hinder plasmid control of transfer (Figure 1.6-A), so entry exclusion genes need to be mutated in this set-up.

#### Entry exclusion with F *traS* mutant

With *traS* deletion, the recipient ability of cells bearing  $F_H$  plasmid is restaured (Figure 1.10). In all experiments, we will use  $F-oriT-8$  plasmid with *traS* deletion as the  $F_{HR}$  plasmid (helper with high recipient ability). Using this construct, we will be able to study the selection of recipient ability by competing strains with  $F_{HR}$  and  $F_{HE}$ , and mimic plasmid control of transfer by putting  $F_{HR}$  in all cells.

## 1.3 Fluorescent markers for strains and plasmids

### 1.3.1 Design of the system

Our goal is to follow at least two different competitors simultaneously in order to evaluate fitness, and to follow transfer. To this end, we use a combination of fluorescent proteins to distinguish strains and plasmids, that we follow with flow cytometry (see Appendix B for a description of the flow cytometry analysis).

#### Choice of fluorescences

To mark the strains, we use either fluorescence genes located on non-transmissible plasmids, or a chromosomal marker. To mark plasmids and follow their potential transfer to recipients, we use fluorescence genes located on plasmids themselves.

The exact combination of fluorescence markers used depends on the competition under

study. Generally, one strain is marked with the red fluorescence protein (RFP) gene, localized either on the chromosome or on a non-mobilizable plasmid (pSB plasmid). RFP identifies either one of the two competitor strains, or distinguishes the initial donors from recipients. Plasmids of interest (pZS plasmids, with or without *oriT*) are marked with distinguishable green fluorescent protein (GFP) or yellow fluorescent protein (YFP) genes.

I present here the setup we used the most (chapter 2 and chapter 3 section 3.3.2), where RFP on the non-mobile pSB plasmid identifies recipients (R). Transfer results in cells that are marked with two fluorescent proteins (Figure 1.11). The set-up implicitly

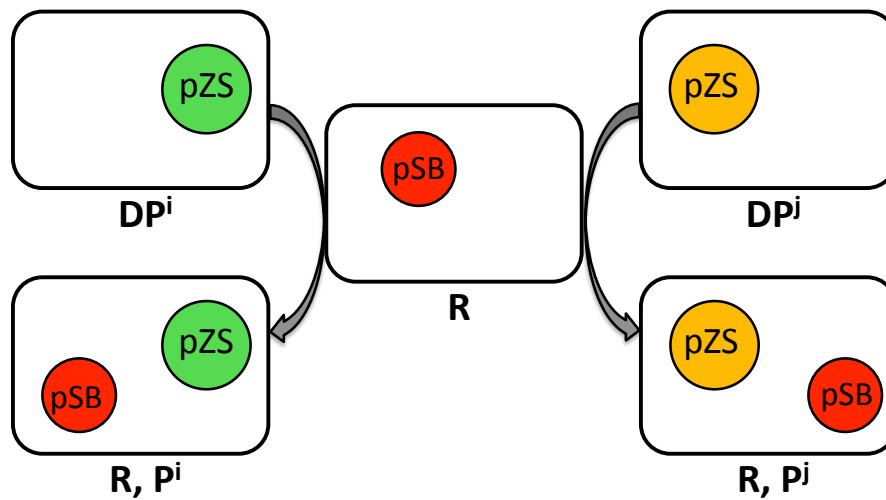


Figure 1.11: **Exemple set-up to follow two plasmids and strains.** Recipients (R) are distinguished from initial plasmid-bearers (D) with the RFP expressed by pSB plasmid; two plasmids of interest  $P^i$  and  $P^j$  are pZS plasmids that can be identified with GFP and YFP. They can be mobilized to recipients (if they bear *oriT*, not shown here), which leads to cells marked by two fluorescences.

assumes that transfer happens only from initial plasmid-bearers to recipients. This will mostly be the case when mimicking chromosomal control of donor ability to non-kin, thanks to the entry exclusion. However, with plasmid control of transfer or chromosomal control of donor ability to kin (where entry exclusion is absent), who receives the plasmid will mostly depend on strain frequencies: conjugation is governed by mass-action laws, so transfer will mostly happen to recipients if recipients are the majority of cells. We can quantify the process since the double GFP+YFP cells are apparent in flow cytometry.

### Maintenance of markers

Chromosomal markers for the strains are stably maintained, but plasmid markers are not, because of plasmid loss. The ideal way to maintain them would be to have specific antibiotic resistance markers for each plasmid. However, when tested with ampicilline

resistance, the antibiotic marker lead to a huge delay in growth. This is probably due to the high cost of bearing multiple plasmids and antibiotic resistance genes. We thus work only with the initial antibiotic marker, kanamycin. pZS plasmids and the pSB plasmid are both kanamycin-resistant, but have different replication origins to avoid plasmid incompatibility (Figure 1.11). Initially, all plasmids are maintained with kanamycin. As transfer happens, one of the plasmids could possibly be lost in recipients, but these events should be rare, as experiments are done on a time-scale of few generations. We confirmed this in our experiments: we do not observe any decrease of RFP fluorescence for populations of recipient cells that received pZS plasmids (data not shown).

### Final combinations of genotypes and their phenotypes

I cloned the genes for YFP and GFP in association with all pZS plasmids, with or without *oriT* (see plasmid maps, Appendix A). Transfer from donors is controlled by the combination of *oriT* sequence on the pZS plasmid, and the presence of a F helper plasmid in the cell; recipient ability is controlled by the *traS* allele present on F (table 1.12).

Property	Phenotype	Genotype (pZS plasmid)	Genotype (host)
public good production	P <sup>+</sup>	<i>rhlI</i>	-
	P <sup>-</sup>	-	-
strain donor ability - to kin  - to a separate strain	D <sup>+</sup>	<i>oriT</i>	$F_{HR}$
	D <sup>-</sup>	<i>oriT</i>	-
	D <sup>+</sup>	<i>oriT</i>	$F_{HE}$
strain recipient ability	D <sup>-</sup>	-	$F_{HE}$
	R <sup>+</sup>	-	$F_{HR}$
plasmid autonomous transfer	R <sup>-</sup>	-	$F_{HE}$
	T <sup>+</sup>	<i>oriT</i>	$F_{HE}$
fluorescence	T <sup>-</sup>	-	$F_{HE}$
	green	<i>GFP</i>	-
	yellow	<i>YFP</i>	-
	red	-	<i>RFP</i>

Figure 1.12: **Summary of the relevant strain and plasmid phenotypes and genotypes controlling them.** Genotypes include genes on pZS plasmids, and "host" non-mobile genes that can be localized either on the chromosome or on a non-mobile plasmid.

### 1.3.2 Combining transfer and fluorescence

At this stage, we have all the components needed to study the interplay between transfer and cooperation. However, we have not yet tested how they behave when combined. The following paragraphs describe the issues that arose when combining all the components of the system, and their resolution.

#### The issue of temperature

We first needed to choose the temperature that would optimize plasmid transfer and facilitate the best flow cytometry measurements. GFP and YFP are well distinguished by flow cytometry at 30°C. However, populations of cells expressing GFP and YFP overlap up to 10% at 37°C, because of the appearance of low-fluorescent cells in both populations, that cannot be properly classified. This lead us to use 30°C for measurements of the final system (with transfer, public good production and fluorescences). Unfortunately, transfer efficiency is strongly reduced at low temperature (all our initial measures of transfer were done at 37°C). Transfer efficiency decreases linearly with temperature, and is up to a 10-fold less at 30°C than at 37°C. This is confirmed by studies that have shown that F piliation depends strongly on temperature [Novotny and Lavin, 1971, Sowa et al., 1983]. To resolve these issues, we settled on a compromise protocol: the growth during the pre-induction step is at 35°C, where transfer happens quite efficiently, while all other steps (including pre-culture of individual strains and growth in the presence of Cm) are done at 30°C (see the final protocol in Figure 1.17).

#### Flow cytometry measurements

When testing the generic design shown in Figure 1.11, recipients are clearly distinguished from initial plasmid-bearers with flow cytometry (Figure 1.13). Moreover, this distinction is stable over time as transfer happens, confirming that red fluorescence is not lost from recipient cells on the timescale of our experiments, despite the reduced selection for plasmid maintenance when two kanamycin-resistant plasmids are present in the cell.

GFP and YFP are also clearly distinguished in initial plasmid-bearing cells, without any overlap between the fluorescences. The distinction is less accurate in recipient cells after transfer (Figure 1.13), with up to 2.5% overlap between YFP- and GFP-positive populations (measured by comparing separately grown populations). This overlap is mainly due to cells with lower fluorescence intensity, that may have received a plasmid recently.

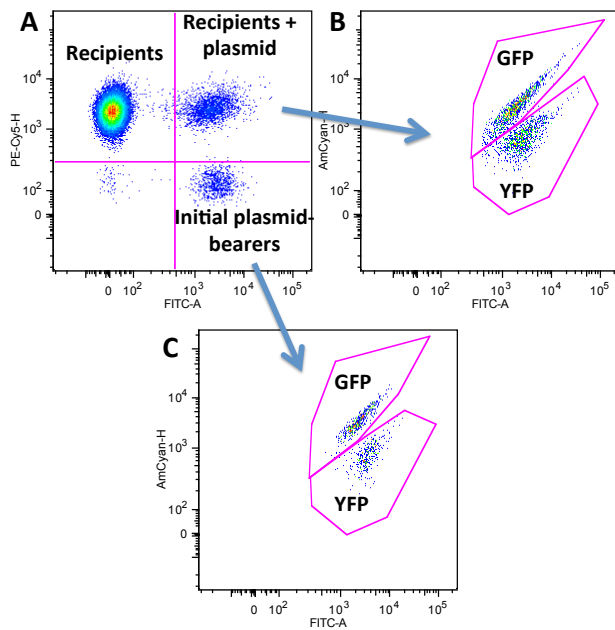


Figure 1.13: **An example of the identification of cells and transfer with flow cytometry.** Initial plasmid bearers and recipients with or without plasmids are first distinguished (A) based on RFP fluorescence (Pe-Cy5-H axis) and mixed GFP- and YFP-fluorescence (FITC-A axis), then GFP and YFP expressing cells are separated with specific filters (AmCyan-H *vs* FITC-A) for both recipients bearing plasmids (B) and initial plasmid-bearers (C). The flow cytometry analysis is described in Appendix B.

## Double transfer events

When two types of mobilizable plasmids are present, transfer with both GFP and YFP plasmids may also happen to the same recipient cell. In the absence of entry exclusion, it may further happen between initial plasmid-bearing cells. A key parameter controlling transfer partners is the initial cell frequency, because plasmid transfer is governed by a mass-action law: if initial plasmid-bearing cells are low in frequency, we only observe cells with either GFP or YFP. When they are high in frequency, the distinction between GFP and YFP populations is blurred in initial as well as in secondary bearers, which corresponds to the appearance of cells containing both fluorescent proteins (Figure 1.14).

## 1.4 Combining transfer and cooperative behaviour

In the previous section I described the system that makes it possible to follow transfer of two plasmids to recipient cells, and identify them accurately. Before starting experiments, we must also examine the interactions between the transfer and the cooperation system.

### 1.4.1 Transfer affects the response to public goods

After dilution in the presence of Cm, we observed the appearance of a population of donor cells with low green fluorescence when mixing donor strains (with both  $F_{HE}$  plasmid and  $pZS-oriT$  plasmids) with recipients (data not shown). Those cells are not clearly marked by propidium iodide or other markers of death, but they do not grow on LB-agar plates. This low-fluorescent population appears only in the presence of recipients, and is strongly



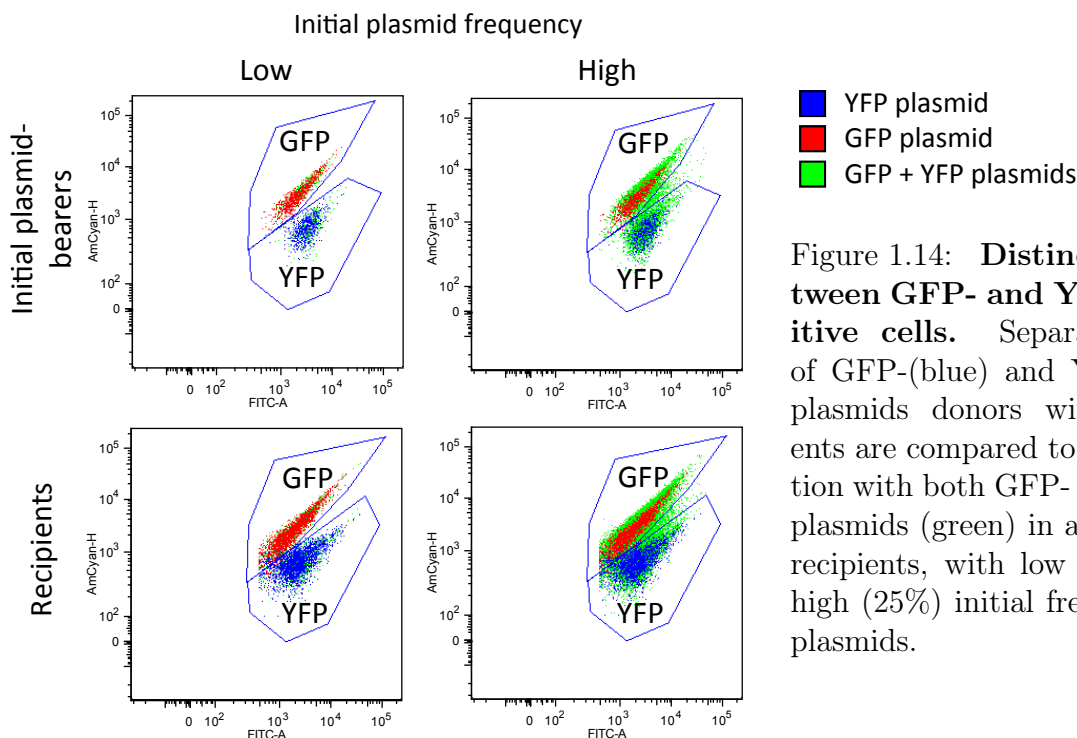


Figure 1.14: **Distinction between GFP- and YFP- positive cells.** Separate mixes of GFP-(blue) and YFP-(red) plasmids donors with recipients are compared to a population with both GFP- and YFP-plasmids (green) in a mix with recipients, with low (2.5%) or high (25%) initial frequency of plasmids.

reduced at 30°C compared to 37°C, suggesting a link with effective transfer. It moreover affects more non-producers than producers, and adding C<sub>4</sub>-HSL to the medium further reduces this population. Finally, even with sufficient pre-induction time, P<sup>-</sup> donors decrease in frequency in competition with P<sup>+</sup> donors when adding Cm. All these observations suggest an altered permeability to C<sub>4</sub>-HSL or a less efficient induction of gene expression by C<sub>4</sub>-HSL, in some way linked to plasmid transfer.

We did not investigate further this behaviour, but searched for conditions that reduce its effect on our experiments. The final protocol we settled on includes an additional dilution step at 30°C, after transfer but before dilution with Cm (see the final protocol in Figure 1.17). This step almost completely suppresses the low-fluorescent population, probably because it decreases the expression of the transfer operon before cells are exposed to Cm.

## 1.4.2 Modified public good benefits

### Frequency-dependance of benefits

Studying the interplay between transfer and cooperation implies that the initial proportion of producer cells will be reduced as recipients are added to the mix. We found the relative fitness of producers in the presence of Cm to be frequency-dependant at low frequencies of producers (Figure 1.15):

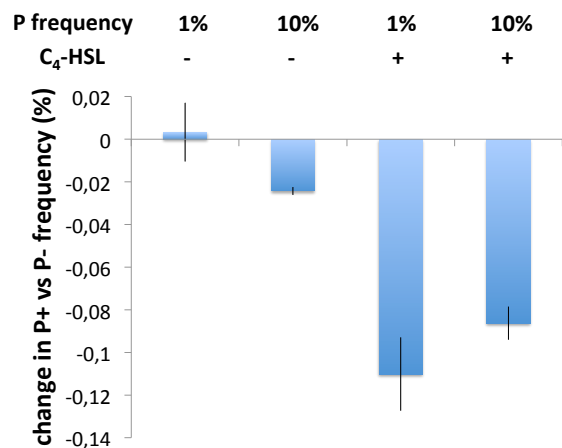


Figure 1.15: **Change in producers frequency** compared to non-producers after 13h of growth in the presence of Cm, depending on both their initial global frequency and 0.75 $\mu$ M C<sub>4</sub>-HSL presence from the pre-induction stage. Producers, non-producers and recipients all beared F<sub>HE</sub> plasmid. Results are mean of 3 replicates, standard deviations are shown.

When initially present in a frequency of 1%, producers appear to have an advantage over non-producers (with huge variation across replicates) instead of a cost. When C<sub>4</sub>-HSL is added, it restores the apparent cost of producers. Although the percent change in frequency between populations with and without C<sub>4</sub>-HSL is higher for the initial 10% frequency of P<sup>+</sup>, we expect that this difference arises from an increased number of generations when more public good is present. The frequency-dependance suggests a differential benefit of low public-good concentrations for producers, which could be explained by incomplete diffusion or intake of C<sub>4</sub>-HSL.

### Poor induction of recipients

We observed that recipients bearing pSB plasmid grow poorly in the presence of Cm. This is due to an inefficient response to C<sub>4</sub>-HSL, and not a lower basal growth in the absence of C<sub>4</sub>-HSL: when C<sub>4</sub>-HSL is added exogenously, recipients are outcompeted by initial plasmid bearers in the presence of Cm. The observed effect was unexpected: the only thing that makes these cells different is that the *RFP* gene is localized on a plasmid with higher copy number, which seems to affect the response to C<sub>4</sub>-HSL in an unknown way. We did not observe poor growth in presence of Cm for recipients marked with the chromosomal RFP marker.

Such growth difference could alter the conclusions of our experiments when recipient and non-recipient strains are competed, or when plasmids are competed (as poor growth of recipients will also decrease the plasmid fitness gain from transfer to recipients). Competitions concerning recipient ability (Chapter 3) were done with chromosomally marked recipients, avoiding this issue.

## Public good benefits with transfer

An implicit expectation for our system is that plasmids should have the same behaviour in recipients as in the initial plasmid-bearers: the *rhII* gene should be expressed in recipients and promote growth in the presence of Cm as well as it does in donors. To confirm the uniform plasmid behavior across strains, I measured growth in the presence of Cm for mixes of producers and recipients, comparing producers with or without transfer ability.

After 13 h of growth at 30°C, transfer increases overall population growth when comparing populations with or without transfer, started with equal frequencies of initial plasmid-bearing cells (Figure 1.16-A). The maximal benefit is observed for 30% of initial producers. Moreover, if growth is plotted as a function of the total percentage of producers before dilution in Cm (including recipients that received P<sup>+</sup>*oriT* plasmid), the curves become roughly similar for both strains (Figure 1.16-B). These results suggest that recipients with P<sup>+</sup>*oriT* plasmid are effective at promoting growth: *rhII* is expressed and active.

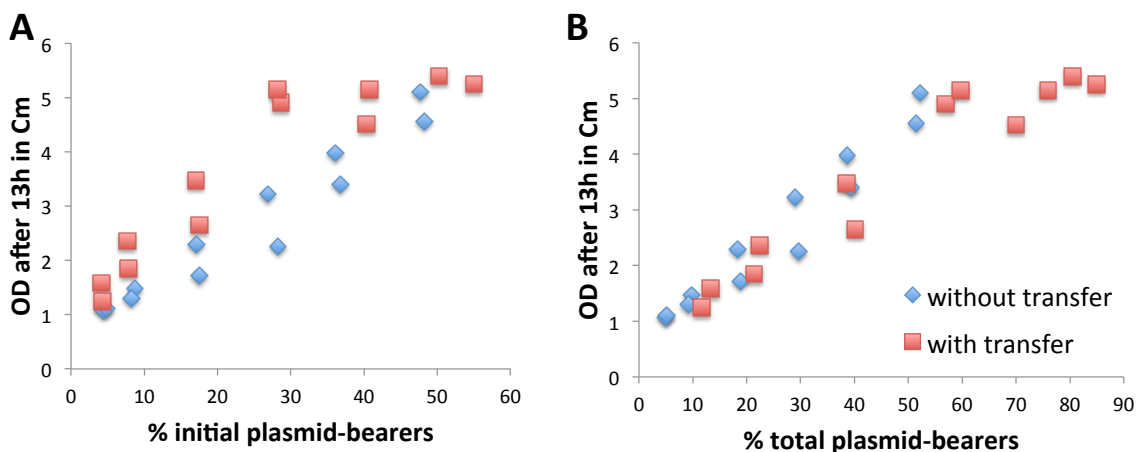


Figure 1.16: **Effect of the transfer of the public good production gene on population growth.** Mixes of strains bearing  $F_{HE}$  and P<sup>+</sup>GFP (no transfer) or P<sup>+</sup>GFP-*oriT* (transfer) with recipients were diluted in the presence of Cm after transfer and pre-induction. The optical density at 600nm (OD<sub>600</sub>) was measured after 13h of growth at 30°C, and expressed as a function of the initial frequency of initial plasmid-bearers (excluding recipients, A) or total plasmid-bearers (including recipients, B).

## 1.5 Final set-up and discussion

### 1.5.1 The experimental system

#### Strains and plasmids

We have constructed multiple bacterial strains with a combination of production/non-production, transfer/non-transfer and fluorescent alleles (see Figure 1.12). Depending on the set-up (see Figure 1.6), those alleles can or cannot be transferred to recipients. We can thus directly address questions about the interplay between transfer and cooperation. Additionally, we can use this system to study selection of transfer in the absence of public-good cooperation (see Chapter 3 section 3.3.1).

#### Possible experiments using this system

We can distinguish two competitors ( $D^+/D^-$ ,  $R^+/R^-$ ,  $T^+/T^-$  or  $P^+/P^-$ ), and the transfer of two plasmids to recipients (Figure 1.11). We can thus do pairwise competitions in the presence of recipients. The choice of markers (GFP and YFP) will be done as a function of desired costs: the net apparent effect of production as well as transfer ability should be costly. Surprisingly, transfer ability is actually costly in this set-up (see next chapters), even when F plasmid is present in all cells: this means that here the transfer itself is a significant part of transfer cost, and not only the pili production.

Transfer seems to be efficient enough for us to observe its consequences in our experiments. Indeed, the transfer rates achieved in the experiments are high: instantaneous transfer rates in liquid are more than  $10^{-9}$  mL cell<sup>-1</sup>h<sup>-1</sup>, which is in the range of derepressed plasmids like F itself [Simonsen, 1991]. Moreover, the transfer of production genes increases benefits of cooperation, as needed (higher growth rate in the presence of Cm).

#### Precision and robustness of the system

As shown in section 1.3.2, the precision of flow cytometry measurements is good in the absence of transfer (no overlap between populations). With transfer, the accuracy of identifying genotypes declines, as populations of recipients with YFP- or GFP-expressing plasmids overlap up to 2.5%. Unfortunately, this hinders very precise measurements of plasmid fitness differences. Still, we are mainly interested in measures of among-population selection, which remain possible: these measures rely on the distinction between among- and within-population changes, independantly from the precise boundary between populations.

As transfer is very efficient, cells with both GFP- and YFP-expressing plasmids may appear when plasmid frequencies are high. This complicates the set-up, and goes against simplifying assumptions from the models (no transfer between plasmid donors, recipients are either producers or non-producers). However, as shown in section 1.3.2, this issue can be limited by setting low initial plasmid frequencies.

## 1.5.2 The experimental protocol

### The generic protocol

First, we mix cells with a 10-fold dilution of cultures previously grown to stationary phase at 30°C ( $t_0$ ). Cultures then grow at 35°C up to an optical density (OD) of 3. Transfer mainly happens at this step. We then dilute cultures 10-fold at 30°C and let cells grow to stationary phase (OD around 5,  $t_1$ ). Finally, we dilute cultures 100-fold at 30°C in the presence of Cm. Cultures grow until  $t_2$ , at which point the experiment is stopped.

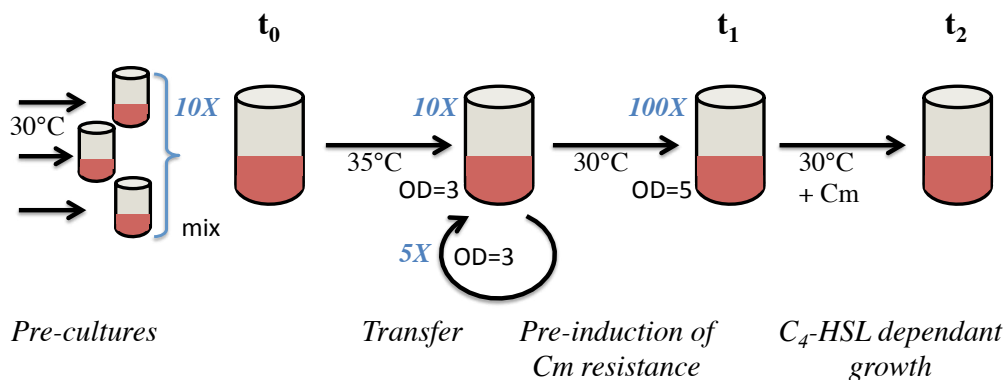


Figure 1.17: **Generic protocol merging transfer and cooperation.** Black arrows show growth conditions, with growth temperature indicated. Blue numbers indicate dilution, performed at the OD (600nm) indicated in black. The 5-fold dilution step is optional and amplifies transfer.

### Protocol constraints

Our initial experiments have shown that producers get a differential benefit when present at frequencies below 2%, thus, C<sub>4</sub>-HSL has to be supplied exogenously. This will not prevent us to study benefits of transfer, provided that C<sub>4</sub>-HSL is supplied in both with and without transfer conditions, and that benefits of transfer are significantly higher than benefits of this low concentration of C<sub>4</sub>-HSL.

Finally, the synthetic cooperation system was shown not to be robust at long time-scales: Julien Bénard-Capelle (post-doc collaborator from INSERM U1001) found that non-producers rapidly evolve to grow faster in the presence of Cm (unpublished data).

The success of non-producers may be due to mutations in genes responding to C<sub>4</sub>-HSL that allow for constitutive expression of Cm-resistance. Because of such genotype and phenotype changes, the system can be only used for experiments on ecological but not on evolutionary timescales.

### 1.5.3 Potential future system improvements

#### Fluorescence markers

Although the appearance of cells expressing both GFP and YFP is an issue in our experiments, it corresponds to a real biological phenomenon: plasmid superinfection [Smith, 2011]. So far we excluded plasmid superinfection from our studies. Analyzing it is possible, but would require a change in fluorescence markers, as GFP and YFP are too similar for their combinations to be accurately distinguished.

#### Changes in the public-good set-up

Another interesting modification of the system would be to change the response to cooperation. Currently, the response to public good presence is an increase in growth rate, resulting in transient and thus unstable benefits. Potential effects of transfer can be observed only during a limited window of time. Using effects on growth yield instead would be more robust measurements, not restrained to a precise time interval. Such effects would also correspond to the interesting situations in which cooperation gives individuals the access to new ecological niches, for example the ability to consume new nutrients. An option could be to replace the coding sequence for Cm-resistance by another gene like the  $\beta$ -galactosidase coding sequence that allows growth on lactose. I attempted to do this with the  $\beta$ -galactosidase, but with no success.

#### Modulating transfer and cooperation

It would be interesting to modulate the intensity of transfer or cooperation, instead of having only two states (high/absent). We can change effective transfer rates by manipulating temperature and cell densities, but we do not have strains with intermediate transfer or cooperation abilities. Moreover, no regulations of cooperation or transfer are implemented. Absence of regulation allows us to rigorously study the selection of transfer or cooperation abilities, but is quite different from real biological systems. We expect that cooperation and transfer are generally present only when they can be efficient (at high densities for instance), which probably strongly decreases net costs for these behaviours [Kümmerli and Brown, 2010, Haft et al., 2009].



## Chapter 2

# Genetic information transfer promotes cooperation

## 2.1 Introduction

Microorganisms cooperate via the production of secreted factors such as signalling, resource-scavenging, virulence or anti-competitor molecules. These molecules generate shared extracellular benefits, or "public goods", that affect crucial processes defining population growth and virulence of pathogenic bacteria. Public good molecules are costly to produce but benefit not just the organism secreting them but their neighbours as well [Crespi, 2001, West et al., 2007c]. Therefore, the maintenance of cooperative behaviours is threatened by the spread of cheaters, individuals that benefit from cooperation without contributing to it. Social evolution theory predicts that cooperation can be maintained when benefits of cooperation are directed preferentially to organisms carrying cooperation genes, as summarized by Hamilton's rule [Hamilton, 1964a, Frank, 1998]: a behaviour will be selected when  $r \times b > c$ , where  $b$  is the benefit to recipients,  $c$  is the cost to the actor, and  $r$  is the relatedness, or assortment, between actors and recipients of the behaviour.

Social evolution theory has been successfully applied to micro-organisms [Griffin et al., 2004, Gilbert et al., 2007, Diggle et al., 2007b], highlighting numerous factors that can promote cooperation. Particularly, positive assortment or relatedness among cooperators can arise through limited dispersal of relatives [Kümmerli et al., 2009a], or kin discrimination [Strassmann et al., 2011]. However, the meaning of relatedness is particularly ambiguous for bacteria. Bacterial genomes are extremely plastic, with high mutation rates and peculiar forms of sex [Levin and Bergstrom, 2000, Narra and Ochman, 2006]: genes transfer horizontally within and between bacterial lineages at high frequencies, mainly by association to mobile genetic elements such as plasmids or phages [Thomas and Nielsen, 2005, Frost et al., 2005]. Relatedness is therefore expected to be variable and locus-specific in bacteria [West et al., 2006], as opposed to more homogenous relatedness patterns across genomes determined by sexual reproduction in Eukaryotes.



Cooperative alleles could actually benefit from horizontal transfer in two ways. First, horizontal transfer of a cooperative allele may enforce cooperation by infection, converting previous cheaters to cooperators [Smith, 2001]. However, this enforcement is predicted to be unstable in the face of cheats [Nogueira et al., 2009, Mc Ginty et al., 2011], as a non-cooperative allele will displace the cooperative allele when it also transfers horizontally. Nevertheless, genes coding for cooperative traits are over-represented on mobile elements [Nogueira et al., 2009, Rankin et al., 2010], suggesting some link between social behaviours and horizontal transfer. To explain this trend, theoretical work suggests that horizontal gene transfer increases locus-specific relatedness by spreading mobile alleles within populations [Nogueira et al., 2009, Mc Ginty et al., 2011]. Genes associated with mobile genetic elements will thus have higher relatedness than other genes in the genome, which should favour cooperation at mobile loci. Both infection and kin selection could then influence cooperation in complementary ways [Mc Ginty et al., 2013]. However, so far there is no experimental evidence that either of those mechanisms is able to maintain cooperation. Here, we conduct experiments and simulations to test whether transfer can promote cooperation, either by pure infection or by increasing relatedness in structured populations.

## 2.2 Design of the study

To study cooperation and transfer, we developed a synthetic system with independent control of cooperation and conjugation in *Escherichia coli*. This system builds on a synthetic cooperation system where cooperators are producers of a secreted public good (the quorum sensing molecule C<sub>4</sub>-HSL) that provides faster growth in both producers and non-producers in the presence of the antibiotic chloramphenicol, by activating expression of a chloramphenicol resistance enzyme [Chuang et al., 2009]. Populations with varying proportions of public-good producing bacteria can provide assortment between cooperators, and the faster growth of populations enriched in cooperators leads to a case of Simpson’s paradox, allowing transient selection of cooperation at the metapopulation scale [Chuang et al., 2009].

We add plasmid transfer to this system to study the effect of transfer on cooperation when producer, non-producer or both alleles are able to transfer by conjugation to recipient cells (Figure 2.1). We use the conjugation machinery of the F plasmid [Frost et al., 1994]: a helper F plasmid ( $F_{HR}$ ) mobilises the transfer of F-*oriT* bearing producer

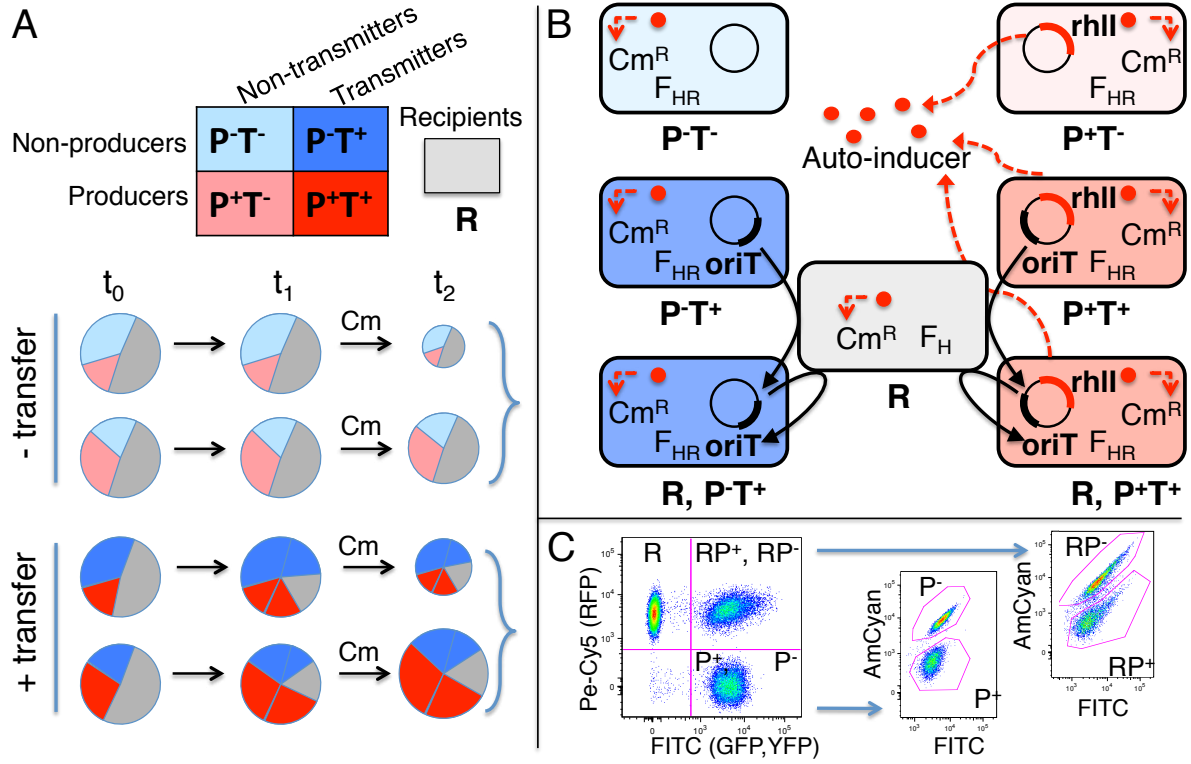


Figure 2.1: **Experimental system.** **A: Experiment design.** Public good producers ( $P^+$ , red) and non-producers ( $P^-$ , blue) are mixed at  $t_0$  with recipients ( $R$ ) in populations with varying ratios of  $P^+$  to  $P^-$ .  $T^-$  plasmids ( $T^-P^+$  and  $T^-P^-$ , pale colors) cannot be transmitted.  $T^+$  plasmids ( $T^+P^+$  and  $T^+P^-$ , bright colors) can be transmitted to recipients, yielding respectively new  $P^+$  and  $P^-$ . Populations are then diluted in chloramphenicol (Cm)-containing medium ( $t_1$ ), where subsequent growth (until  $t_2$ ) depends on  $P^+$  proportion. **B: Synthetic system for conjugation and cooperation.**  $P^+$  cells express the synthase RhlI that produces Rhl auto-inducer (red dots), which diffuses (red arrows) and activates expression of  $Cm^R$  in all cells.  $F_{HR}$  plasmid expresses F conjugation machinery and mobilizes plasmids bearing  $oriT$  ( $T^+$ ) to recipient cells (black arrows), leading to  $RP^+$  and  $RP^-$ , whereas plasmids without  $oriT$  ( $T^-$ ) are not mobilized. **C: Flow cytometry.** Recipients are marked with RFP, plasmids with YFP ( $P^+$ ) or GFP ( $P^-$ ). Initial plasmid bearers and recipients without or with plasmids are first distinguished with green and red filters, followed by separation of  $P^+$  and  $P^-$  using green and cyan filters.

( $P^+$ ) or non-producer ( $P^-$ ) plasmids to recipient cells (Figure 2.1-B). Recipients,  $P^+$  and  $P^-$  plasmids are marked with distinguishable fluorescent proteins to enable monitoring of strain and plasmid dynamics using flow cytometry (Figure 2.1-C and Methods). This synthetic approach allows us to study the effect of horizontal transfer without interference from other processes or plasmid-host co-adaptation [Tanouchi et al., 2012].

In parallel, we use mathematical modelling to analyze more deeply the effects of varying transfer, with realistic parameters derived from our experiments (see Methods). Briefly, we follow the two plasmids  $P^+$  and  $P^-$ , that can transfer by conjugation to recipient cells ( $R$ ) within populations. Public good production is costly and decreases the

growth rate of cells bearing  $P^+$  plasmid, but benefits the population in the presence of Cm, where the growth rate of cells is function of the proportion of  $P^+$ -bearing cells in the population.

## 2.3 Results

### 2.3.1 Enforcement of cooperation in a well-mixed population

We first model (Figure 2.2-A) and test experimentally (Figure 2.2-B) the effect of horizontal transfer in a single well-mixed population where initially the majority of cells (97.5%) are recipients. In the absence of transfer, producer plasmids are slightly outcompeted by non-producer plasmids, as producer cells pay the cost for public good production (3% decrease in  $P^+$  ratio, Figure 2.2 blue,  $p=0.03$ ): cooperation is not maintained. When  $P^+$  bear *F-oriT*, they are transferred to recipients. As expected [Smith, 2001], the transmissible producer plasmid outcompetes the static non-producer plasmid thanks to  $P^+$  invasion of recipients (Figure 2.2 green,  $p=0.004$ ), leading to enforcement of cooperation. However,  $P^-$  plasmids also invade recipients in the same way when bearing *F-oriT*, dis-

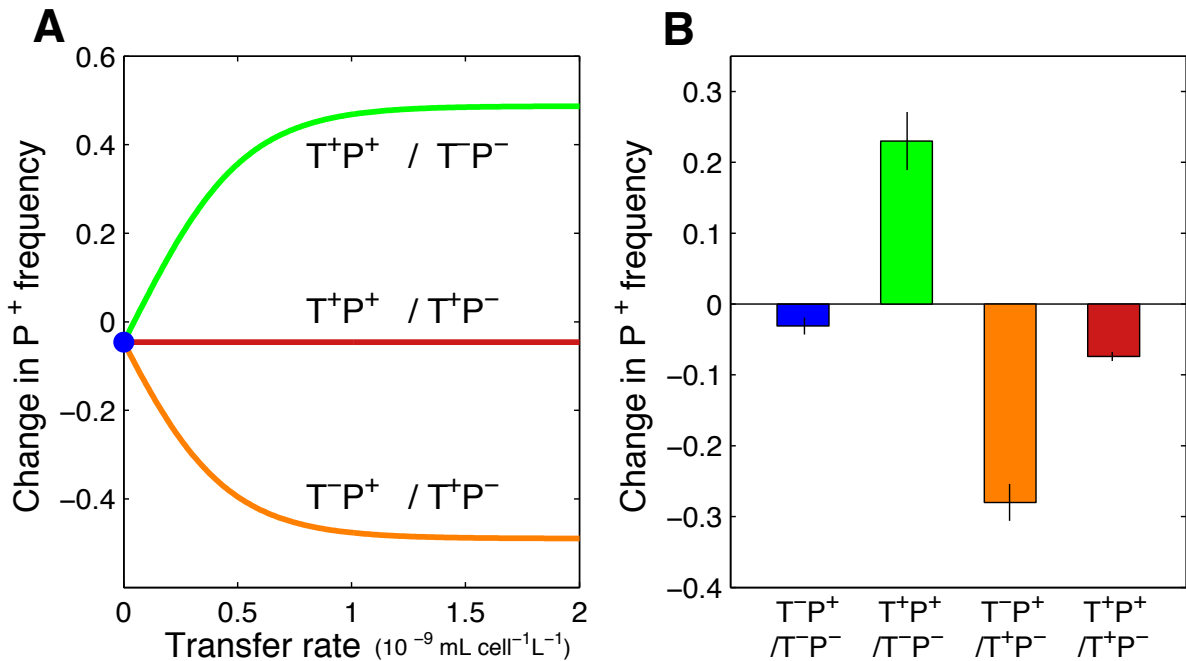


Figure 2.2: **Transfer favors producers as well as non-producers in a well-mixed population.** Initial frequencies are 1.25%  $P^+$ , 1.25%  $P^-$  and 97.5%  $R$ . The change in frequency of  $P^+$  among all plasmids  $P^+/(P^+ + P^-)$  is computed in the absence of Cm ( $t_0$  to  $t_1$ ). **A:** Simulated change in  $P^+$  frequency in the absence of transfer (blue point) and as a function of transfer rate, when  $P^+$  (green line),  $P^-$  (orange line) or both plasmids (red line) can be transferred. **B:** Experimental change in  $P^+$  frequency depending on  $P^+$  and  $P^-$  transmission genotypes (association to  $T^+$  or  $T^-$ ), shown as means  $\pm$  s.e.m. ( $N=9$ ).

favoring the selection of cooperation when  $P^+$  plasmids are not transmissible (Figure 2.2 orange,  $p=5.10^{-6}$ ). When both plasmids transfer at the same rate, the benefit of transfer cancels out and producers are outcompeted again (7% decrease in  $P^+$  ratio, Figure 2.2 red,  $p=3.10^{-6}$ ). Thus, as predicted by theory [Mc Ginty et al., 2011], transfer is not sufficient to maintain cooperation in a well-mixed population, if both  $P^+$  and  $P^-$  alleles are allowed to transfer.

### 2.3.2 Selection for cooperation in a structured population

We then move to a metapopulation consisting of distinct populations with varying proportions of producers, providing positive relatedness, or assortment. Sufficiently high relatedness can generate Simpson’s paradox: despite the local decrease in producer frequency in each population, the global frequency of producers increases thanks to the faster growth of populations enriched in producers [Chuang et al., 2009]. Our simple metapopulation consists of two populations with different ratios of  $P^+$  to  $P^-$  and a constant proportion of  $R$ . Populations grow separately until  $t_2$ , where the two populations are mixed (Figure 2.1-A). We focus on the case where both plasmids transfer.

Simulations predict that under this simple population structure, adding plasmid transfer to both plasmids (with the same transfer rate) will increase the global ratio of  $P^+$  to  $P^-$  plasmids across populations, allowing the selection of cooperation (Figure 2.3 inset). We then confirm this behaviour with our experimental system: producers are outcompeted in the absence of transfer, but they outcompete non-producers when both plasmids transfer, with a 10% increase in the frequency of  $P^+$  compared to  $P^-$  thanks to transfer ( $p=6.10^{-6}$ , Figure 2.3 left). Thus, in a metapopulation setup, the direction of selection is reversed thanks to transfer.

Increase of cooperation due to transfer could be the result of within- or among-population dynamics. To distinguish between the two, we analyze plasmid dynamics at the population level. The ratio of  $P^+$  to  $P^-$  within each population does not increase with transfer (Figure 2.3, right and inset), as observed previously in the case of a single well-mixed population, confirming that we do not observe here a direct infectious benefit of  $P^+$  plasmids. However, the part of the global  $P^+$  frequency explained by differential growth among populations, which can be calculated using the parameter  $d$  (see Methods) increases from 0.08 to 0.37 with transfer ( $p=4.10^{-5}$ ). Thus, the analysis of population dynamics shows that transfer promotes cooperation by increasing among-population vari-

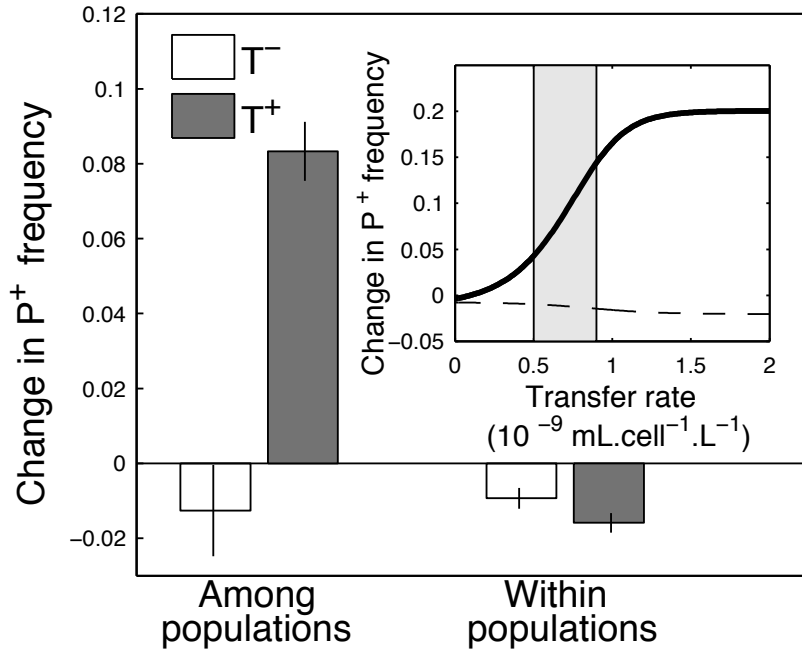


Figure 2.3: **Transfer selects for cooperation in structured populations.** The metapopulation consists of 2 populations, with initial  $P^+/P^-$  ratio of 1/4 and 4/1, and 97.5% R. The change in frequency of  $P^+$  among all plasmids  $P^+/(P^+ + P^-)$  is computed without (open columns) and with (filled columns) transfer of both  $P^+$  and  $P^-$  plasmids, and shown as means  $\pm$  s.e.m. ( $N=9$ ). The inset shows the simulated change in  $P^+$  frequency as a function of transfer rate, among (solid line) and within (dashed line) populations, the shaded area indicating estimated experimental transfer rates.

ation in growth, which amplifies the effect of Simpson's paradox.

### 2.3.3 Effect of transfer on population structure

This effect of transfer on among-population variation suggests a change in population structure, and can be further analyzed in terms of relatedness coefficients describing the association between alleles in the population [West et al., 2006] (see Methods for relatedness definition). We computed relatedness after transfer happens, but before cooperation becomes beneficial ( $t_1$ , see Figure 2.1-A). Relatedness at the producer plasmid locus increases strongly with transfer, both in simulations and in experiments ( $p=0.01$ , Figure 2.4-A), while relatedness at a non-mobile locus (the producer locus considering only founder  $P^+$  cells, excluding recipients) does not increase (it actually decreases, as founder  $P^+$  cells decrease in frequency when they transfer  $P^+$  plasmid:  $p=0.005$ ).

We then use simulations to further investigate the dynamics and effect of relatedness. Following plasmid invasion, relatedness increases strongly after mixing plasmid donors and recipients (Figure 2.4-B). This increase relies on the fact that plasmid transfer is a lo-

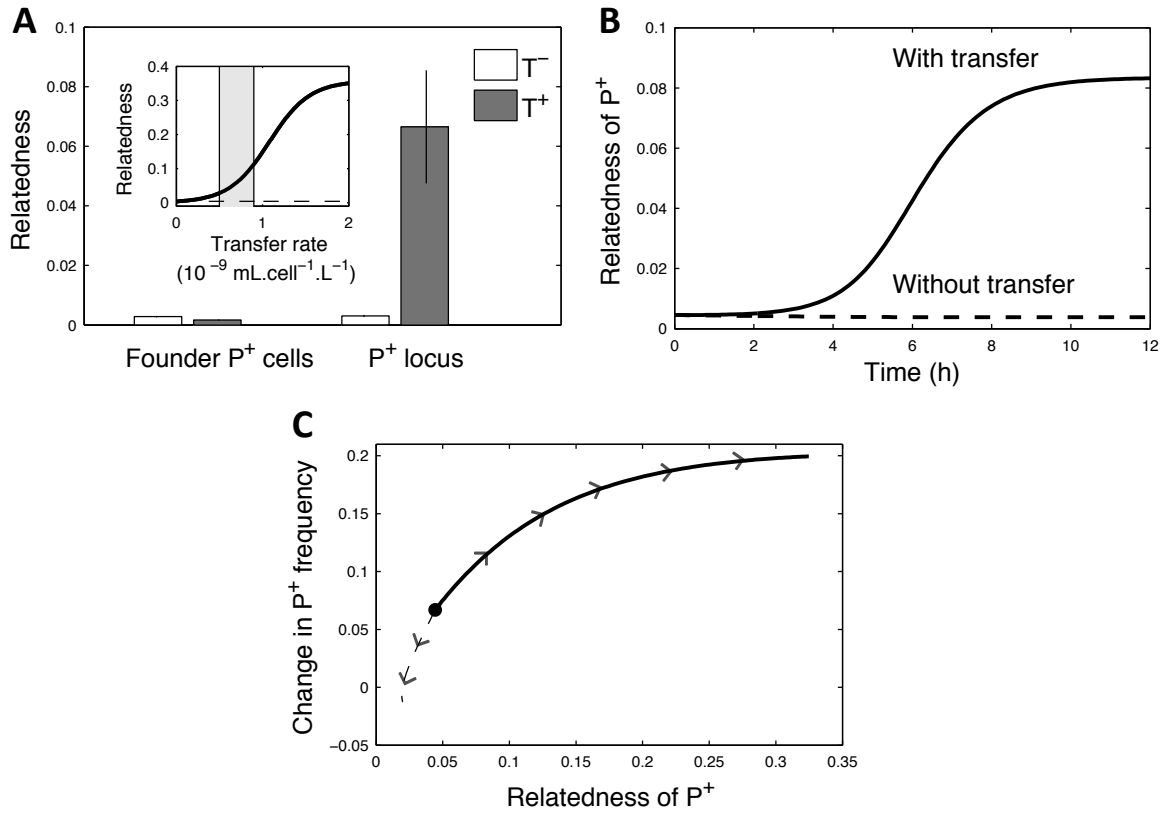


Figure 2.4: **Transfer selects for cooperation by increasing relatedness.** Parameters are the same as in Figure 2, except recipient's proportion which is 97.5% for Figure 2.4-A and 2.4-B, and 75% for Figure 2.4-C. **A:** Relatedness computed at  $t_1$ , without (open columns) and with (filled columns) transfer, shown as mean  $\pm$  s.e.m (N=9). The inset shows simulated relatedness function of transfer rate for mobile (solid line) and non-mobile (dashed line) loci, the shaded area indicating estimated experimental transfer rates. **B:** Simulated relatedness as a function of time, without transfer (dashed line) or with transfer rate  $8.10^{-10}$  at the cooperation locus  $P^+$  (solid line). **C:** Change in  $P^+$  frequency from  $t_1$  to  $t_2$  in simulations as a function of  $P^+$  relatedness, with transfer within populations (solid line) or randomized across populations (dashed line). Arrows indicate the direction of increasing transfer rates, the dot indicating absence of transfer.

cal event that happens strictly within populations, which amplifies pre-existing variations in producer proportions among populations and increases assortment at the metapopulation scale. Increasing the rate of transfer in simulations leads to a parallel increase in relatedness and in the proportion of producers (Figure 2.4-C, solid line).

To test if transfer indeed acts on cooperation through changing relatedness, we finally prevent the effect of transfer on relatedness in our simulations. As the increase in relatedness is due to within-population transfer, we simulated a global mixing of all transferred plasmids across populations at  $t_1$ , keeping other parameters constant. This treatment actually causes a decrease in relatedness with transfer, as the proportion of producers becomes more homogeneous across populations (Figure 2.4-C, dashed line). Subsequently, we observe a decrease in the  $P^+$  to  $P^-$  ratio, despite no change in costs or benefits of co-

operation, or in the total transfer events for both  $P^+$  and  $P^-$ . Thus, simulations confirm that the effect of transfer on cooperation is due to the local nature of transfer and the resulting increase in relatedness.

### 2.3.4 Effect of epidemic spread among recipients

As the effect of transfer relies on the presence of recipients, we expect that the fate of producers versus non-producers will be affected by variations in the proportion and properties of recipients. In simulations, lowering the proportion of plasmid-bearing cells in a population of recipients progressively decreases the strength of selection for producers in the absence of transfer (Figure 2.5, black arrows). As fewer producers are present in a population of recipients, they get less and less benefits from their own public good production and are counter-selected. Adding plasmid transfer (Figure 2.5, red arrows) restores the selection of producers, up to its maximal level when recipients are fully invaded by plasmids: transfer effectively opposes dilution of producer genes by allowing plasmids to invade recipients. Thus, transfer is more effective with a lower initial proportion of plasmids.

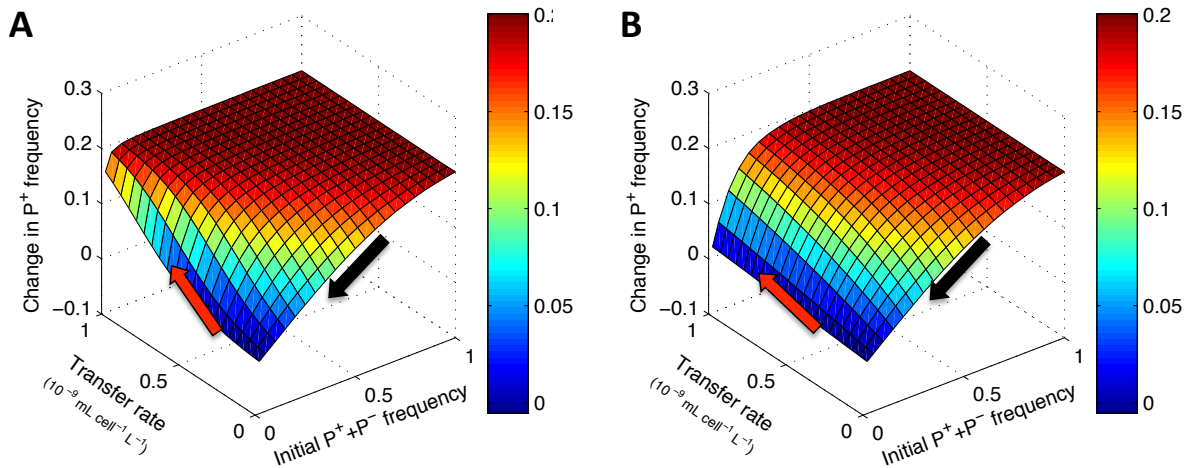


Figure 2.5: **Transfer has a more pronounced effect when recipients are abundant and amplify transfer.** The metapopulation consists of 2 populations, with initial  $P^+/P^-$  ratio of 1/4 and 4/1. Surfaces represent the change in frequency of  $P^+$  among all plasmids  $P^+/(P^+ + P^-)$  in the presence of Cm (from  $t_1$  to  $t_2$ ), as a function of transfer rate and initial frequency of plasmids ( $P^+ + P^-$ ). **A:** Change in  $P^+$  frequency when recipients do not transfer plasmids. **B:** Change in  $P^+$  frequency when recipients also transfer plasmids. Black arrows indicate decreasing plasmid proportion, red arrows indicate increasing transfer rates.

The effect of transfer will also depend on recipient's subsequent transfer ability. We compare here two extreme situations: recipients with the same transfer ability as donors,

thus behaving as secondary donors, and recipients with no transfer ability at all. We see a strong effect of recipients (Figure 2.5-A *vs* 2.5-B): transfer is more efficient when recipients behave as secondary donors of plasmids (which is the case in our experiments), especially when the initial proportion of plasmids is low. In the absence of secondary transfer, initial donors do not convert enough recipients into plasmid-bearers to have a strong impact, and effective transfer is low despite high transfer rates from initial plasmid-bearers. We can conclude that the impact of conjugation on cooperation depends not only on initial plasmid donors, but also on the behaviour of recipients. Transfer ability can be highly variable in recipients, and this heterogeneity has been shown to lead to a huge amplification of plasmid spread in bacterial communities [Dionisio et al., 2002].

### 2.3.5 Invasion of populations by rare production alleles.

In our experiments, transfer acts by amplifying the existing variation among populations but it does not explain the origin of such variation, which we previously created by fixing arbitrary ratios of producers to non-producers in initial populations. Variation in frequencies of producers could arise through different mechanisms, like founder effects or extinction and recolonisation in metapopulation dynamics. In the case of plasmid genes, such variation will arise when a few plasmid-bearing cells colonise a population of recipients. In this situation, we expect some synergy between infectious transfer and public good production: the relatedness of producers will be too low for them to be selected without transfer.

We simulated this process by applying strong dilutions to our simulated populations. Such dilutions lead to stochastic differences between populations [Chuang et al., 2009], with the number of founder cells in each population following a Poisson distribution. We vary both the dilution factor of founder cells and the transfer rate of plasmids, and show that producers are selected only when both the dilution and the transfer rate are high (Figure 2.6). If dilution is not strong enough, the stochastic variation in the number of founder cells is too low to provide sufficient difference in producer frequency across populations. In the absence of transfer, the proportion of producers is too low to allow for the selection of cooperation. Transfer simultaneously amplifies the variation created by dilution and overcomes the scarcity of cooperators created by the dilution. Notably, the producer as well as non-producer plasmids invade recipients thanks to horizontal transfer, but producers outcompete non-producers thanks to their ability to exploit increasing



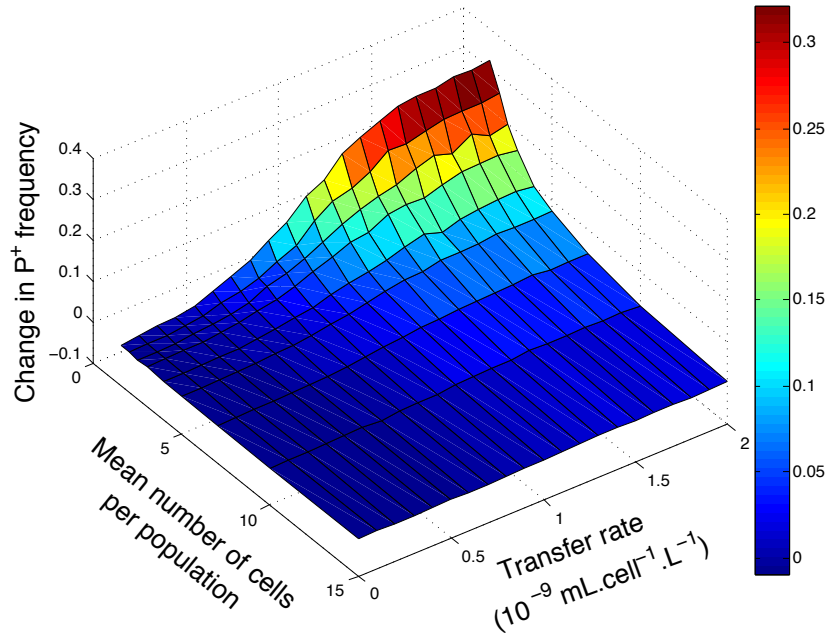


Figure 2.6: **Transfer allows rare producer genes to invade bacterial metapopulations.** The simulated metapopulation consists of 96 populations initiated from a strongly diluted mix of ( $1\%P^+$ ,  $1\%P^-$ ,  $98\%R$ ), giving rise to Poisson distribution of  $P^+$  and  $P^-$  across populations. The change of  $P^+$  frequency among all plasmids  $P^+/(P^+ + P^-)$  in the presence of Cm ( $t_1$  to  $t_2$ ) is shown as a function of transfer rate and mean initial number of plasmid-bearing cells per population, averaged over 20 simulations.

relatedness - generating the Simpson's paradox.

## 2.4 Discussion

Both our experiments and simulations demonstrate that horizontal transfer favours cooperators in structured populations. In non-structured populations, transfer directly benefits any transferred allele, but not particularly the producer one: cooperation can be very strongly enforced by infectious transfer but only on a short timescale. As soon as a cheating plasmid that can also transfer arises, the producer plasmids will lose their infectious benefit. In contrast, horizontal transfer specifically favors cooperative alleles in structured populations by increasing relatedness at the gene level, which ensures that benefits of cooperation feedback on producer genes. This mechanism requires two conditions: some initial variation in producer versus non-producer frequencies, and an initial abundance of plasmid-free cells.

We first arbitrarily set a defined population structure in order to dissect population dynamics, showing that the increase in relatedness at plasmid loci following transfer allows cooperative alleles to be selected at these sites. We then study a case where the initial population structure arises through strong dilution of plasmid-bearing cells lead-

ing to stochastic variations in frequencies. Simulations show that conjugation allows plasmid-encoded cooperative alleles to overcome rarity and invade bacterial populations, the initial variation being amplified thanks to infectious transfer. This scenario could be relevant in natural populations, with the migration of a few plasmid-bearing cells to a new environment that contains only plasmid-free cells. Populations of plasmid-free cells are probably encountered often, as a huge variability of plasmid distribution is observed across bacterial isolates [Boyd et al., 1996]. Benefits and costs of carrying a plasmid are also dependant on environmental conditions, which can change on a short spatial scale or be frequency-dependant [Ellis et al., 2007]. In naturally varying environments, we thus expect plasmids to be lost when disfavored, which will repeatedly ensure formation of plasmid-free cells.

In conclusion, horizontal transfer can increase the parameter space where cooperation is favoured, by infecting recipients and increasing assortment. In natural environments, assortment can be the result of dilution and limited dispersal, but it will be opposed by migration. Horizontal transfer amplifies assortment because it happens at a local scale: conjugation, the most prominent mechanism of horizontal transfer, is necessarily local as it requires cell contact. In that sense, horizontal transfer is similar to other clustering mechanisms like multicellularity [Pfeiffer and Bonhoeffer, 2003, Koschwanez et al., 2011], range expansion [Datta et al., 2013] or formation of groups [Powers et al., 2011], that can lead to increased relatedness. However, its striking particularity is that it affects only specific genes: in contrast with sexually reproducing organisms, the plasticity of bacterial genetic systems and the efficiency of horizontal gene transfer lead to a very different relatedness for different genes in the same genome: relatedness can vary extremely rapidly as horizontal transfer is decoupled from cell division (and can happen much faster). Our results highlight the fact that in the context of cooperation, relevant relatedness is measured at the locus controlling cooperation, and not at the whole genome scale [West et al., 2006]. They can also explain the observed overrepresentation of social genes on mobile genetic elements [Nogueira et al., 2009, Rankin et al., 2010], as these genes will especially benefit from increased relatedness. Additionally, mobile genetic elements are often associated with biofilm formation, another way of cell clustering, and plasmid conjugation itself has been shown to physically induce biofilm formation via conjugative pili [Ghigo, 2001] or other genes [Madsen et al., 2012]. Biofilms will thus be another way for mobile elements to increase relatedness. There, the increase in relatedness will not be gene-specific, but will still preferentially favour cooperation genes. Reciprocally, biofilms

are particularly favorable to horizontal gene transfer [Madsen et al., 2012], which could increase its gene-specific effect on relatedness.

Finally, our work suggests that horizontal transfer could provide additional adaptive benefits when associated with cooperative behaviours, and therefore cooperation may influence selection on horizontal transfer mechanisms, and contribute to the observed diversity in plasmid transfer rates among bacterial isolates [Dionisio et al., 2002]. In bacteria, cooperation would thus represent an indirect selection pressure for sex, defined as any process selected by the benefits of genetic exchange [Redfield, 2001], additionally to any existing advantages of transfer for mobile elements themselves. Reciprocally, targeting conjugation mechanisms [Lujan et al., 2007] could hinder cooperative behaviours involved in the virulence of pathogenic bacteria by decreasing the range of conditions favoring cooperation, in addition to preventing the dissemination of virulence and antibiotic resistance genes.

## 2.5 Materials and methods

### 2.5.1 Plasmids and strains.

The background strain is JC1191, an *E. coli* strain that can grow in low concentrations of chloramphenicol (Cm) in the presence of C<sub>4</sub>-HSL, thanks to the addition of *Pseudomonas* quorum sensing machinery [Chuang et al., 2009]: JC1191 contains the att::rhl-catLVA(Sp<sup>R</sup>) segment with both an Rhl auto-inducer-responsive promoter (P<sub>rhl\*</sub>) driving an unstable version of *cat* responsible for Cm resistance, and the *rhlR* gene under the weak constitutive promoter P<sub>lacI</sub><sup>q</sup>, integrated into the chromosome with the spectinomycin resistance gene Sp<sup>R</sup>.

$P^+$  plasmid carries an artificial operon of *YFP* and *rhlI* under control of the strong promoter  $P_R$ . RhlI enzyme produces the Rhl auto-inducer C<sub>4</sub>-HSL [Chuang et al., 2009].  $P^-$  plasmid carries GFP under control of  $P_R$ . Recipients (*R*) bear pSB3K3-RFP plasmid, that carries *mRFP1* under control of the strong promoter  $P_{lacI}$ .  $P^+$ ,  $P^-$  and pSB3K3-RFP plasmids are maintained with kanamycin resistance genes.

To provide transfer ability, all cells bear the helper plasmid  $F_{HR}$ , a mutant of pOX38::Tc [Anthony et al., 1994] with reduced mobilization efficiency by the F relaxase (1000-fold reduction compared to F), and a deletion of the *traS* gene primarily responsible for the entry exclusion of F plasmid [Achtman et al., 1977]: recipient cells bearing  $F_{HR}$  can receive  $T^+$  plasmids efficiently, and behave as secondary donors.  $T^+$  versions of  $P^+$  and  $P^-$

plasmids additionally carry the wild-type *oriT* sequence of F plasmid. Genotypes and relevant phenotypes of competitor plasmids are summarized in Figure 2.7.

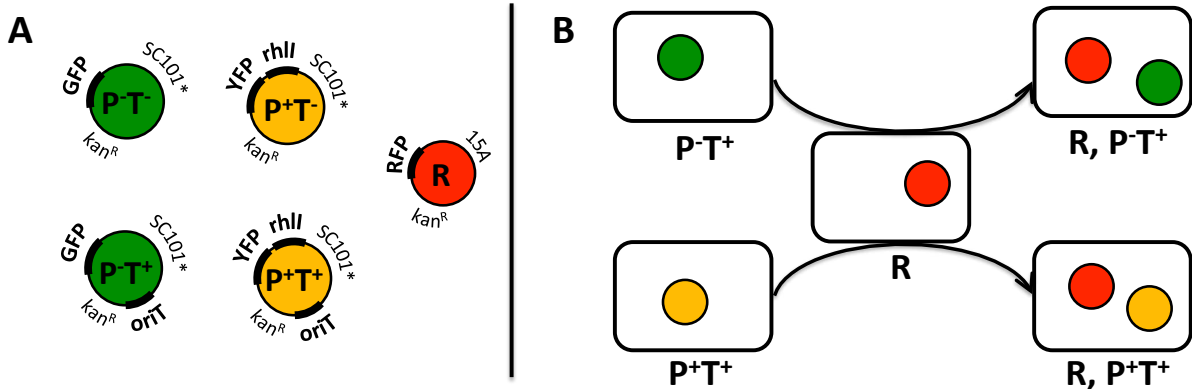


Figure 2.7: **Plasmids used in experiments and strain identification.** Plasmids are colored according to the fluorescence genes they bear. **A: Plasmids used in experiments.** Plasmids are colored according to the fluorescence genes they bear. pSC101-origin (pSC101\*) bearing plasmids are responsible for transfer and production phenotypes.  $P^-$  and  $P^+$  indicate public good production status:  $P^+$  plasmids express *rhII* synthase and *YFP*;  $P^-$  plasmids express only *GFP*.  $T^-$  and  $T^+$  indicate transfer status:  $T^+$  are transferable as they bear F *oriT*. Recipients bear a compatible plasmid (with p15A replication origin) expressing RFP. p15A-origin and pSC101-origin bearing plasmids all bear a kanamycin resistance gene (*kan<sup>R</sup>*). **B: Identification of strains with plasmid fluorescence genes.** Initial strains are marked with only one fluorescence plasmid (the case of  $T^+$  plasmids is represented here). With transfer, recipients bearing two plasmids arise, and are identified by the combination of RFP and GFP or RFP and YFP fluorescence.

### 2.5.2 Growth and experiment conditions.

Cells were grown in Luria-Bertani (BD Difco) medium with 25  $\mu\text{g}/\text{mL}$  spectinomycin (Sp, Sigma-Aldrich) and 50  $\mu\text{g}/\text{mL}$  kanamycin (Kn, Sigma-Aldrich), and with or without 6.25  $\mu\text{g}/\text{mL}$  chloramphenicol (Cm; Sigma-Aldrich) and 0.75  $\mu\text{M}$  C<sub>4</sub>-HSL (CAS# 67605-85-0, Cayman Chemical). Experiments were conducted under well-mixed conditions with 5mL medium in 50 mL tubes (Sarstedt).

For competition experiments, strains were mixed at various ratios (vol/vol) and first grown from a 10-fold dilution ( $t_0$ ; Figure 2.1-A) into medium lacking Cm at 35°C, up to an optical density of 3. This temperature allows transfer to happen, as F transfer is strongly reduced at 30°C. When the initial frequency of  $P^+$  and  $P^-$  plasmids was low, this step was repeated with maximum 2 successive dilutions into medium lacking Cm in order to increase plasmid transfer. Cultures were then diluted 10-fold until stationary phase at 30°C, which allows pre-induction of Cm-resistance by C<sub>4</sub>-HSL in non-producer as well as producer cells and enhances the fluorescence signal. Finally, for experiments

involving cooperation, cultures were diluted 100-fold into medium containing 6.25  $\mu\text{g/mL}$  Cm ( $t_1$ ; Figure 2.1-A) and grown for 12 to 16 hours ( $t_2$ ; Figure 2.1-A) at 30°C. 0.75  $\mu\text{M}$  C<sub>4</sub>-HSL was added to the medium from the 30°C-dilution step for competitions with less than 3% initial producers, as  $P^+$  were found to outcompete  $P^-$  at low initial frequencies, suggesting a differential benefit of very low public-good concentrations for producers. 0.75  $\mu\text{M}$  C<sub>4</sub>-HSL (mimicking the production of 2.5%  $P^+$ ) restored the apparent cost of  $P^+$ , maintaining the configuration of the system in an altruistic state.

While transfer can happen from  $T^+$  to any of the cells present, our experimental setup ensured transfer to happen mainly to  $R$  cells.  $T^+$  cells were placed in very low initial frequencies, making transfer events with other  $T$  cells as recipients rare. Double GFP+YFP positive cells, arising either from transfer between  $T^+$  cells or from transfer of both  $P^+$  and  $P^-$  plasmids to  $R$  cells, represented less than 2% of all cells analyzed and were excluded from the analysis. Cultures were analyzed for strain and plasmid proportions by flow cytometry.

It should be noted that markers are not stably maintained after transfer, as both  $T^+$  and pSB3K3- $P_{lacI}$ -RFP plasmids bear Kn-resistance markers, and one of the plasmids could be lost without the loss of resistance of the cell. However, this problem is minimized in our experimental set-up: both types of plasmids are compatible, as they have different replication origins, and cells are cultured for only a short time period after the transfer actually happens. In spite of potential for long-term marker loss, RFP fluorescence still accurately identifies cells on the timescale of our experiments. Indeed, we see no shift in the red fluorescence signal of transconjugants (that stays clearly distinct from the one of donor cells) compared to recipients during our experiments (data not shown).

### 2.5.3 Data analysis

#### Selection of cooperation.

The selection of cooperation was evaluated at the plasmid level as the change in the frequency of producer plasmids among all plasmids  $P^+/(P^+ + P^-)$ , excluding recipients that do not bear  $P^+$  or  $P^-$  plasmids. The global frequency at the metapopulation level was measured by pooling equal volumes of populations, effectively taking into account differential growth among populations. The frequency within populations was computed as the mean of ratios within populations, excluding the effect of differential growth. To quantify the difference of within and among-population ratios in a normalized way across experiments, we computed a coefficient  $d$  that represents how differential growth biases

the global ratio. Let  $Y_1$  and  $Y_2$  be ratios in the two populations, and  $Y_m$  be the global ratio,  $d$  is defined by :

$$Y_m = \frac{(1-d)Y_1 + (1+d)Y_2}{2} \quad (2.1)$$

thus

$$d = \frac{2Y_m - Y_1 - Y_2}{Y_2 - Y_1} \quad (2.2)$$

### **Relatedness.**

Relatedness is defined as the measure of how an individual's social environment covaries with the individual's genotype, which can be calculated as the linear regression coefficient connecting an individual's genotype with the genotype of its interactants [Damore and Gore, 2012]. Here, we focus on the relatedness of the cooperation allele ( $P^+$ ), considering that the social environment of an individual corresponds to the population it belongs to. Let  $p_i$  and  $n_i$  be respectively the proportion of producers and number of bacteria within population  $i$  ( $i \in (1, 2)$ ), and  $p_{tot}$  and  $n_{tot}$  be respectively the proportion of producers and number of bacteria in the metapopulation. Then, assuming populations are of the same size, which is the case at  $t_1$ , the relatedness of producers  $r_P$  can be calculated as follows:

$$r_x = \frac{\sum_i \frac{p_i}{n_i} \times \frac{p_i}{p_{tot}} - \frac{p_{tot}}{n_{tot}}}{1 - \frac{p_{tot}}{n_{tot}}} \quad (2.3)$$

### **Statistical analysis.**

Differences between conditions with and without transfer were tested with two-sample, two-sided t-tests. The normality of distributions was confirmed with Shapiro-Wilkinson tests, rejecting the null hypothesis of normality when  $p < 0.05$ . When normality was rejected (Figure 2.2-B when both plasmids transfer; Figure 2.4-A, relatedness at the  $P^+$  locus), the non-parametric Wilcoxon signed-rank test was used instead of the t-test.

#### **2.5.4 Model.**

We model the dynamics of producer ( $P^+$ ) and non-producer ( $P^-$ ) alleles, carried on horizontally transmitted, incompatible plasmids that can be transferred to recipients ( $R$ ) only. Producer cells pay a cost of cooperation  $c$ , non-producers do not.

Transfer happens at rates  $\gamma_p^+$  and  $\gamma_p^-$  respectively for  $P^+$  and  $P^-$ , following a mass-action law [Levin et al., 1979], and growth follows a logistic function saturating at carrying capacity  $K$ , mimicking growth to stationary phase. Transfer saturates in the same way

at carrying capacity  $K$ , as F transfer has been shown to strongly decrease when cells approach stationary phase [Frost and Manchak, 1998].

We explicitly follow the experimental set-up by modeling the two steps: first pre-incubation, where pre-induction of Cm resistance and transfer happen before cooperation is needed (from  $t_0$  to  $t_1$ ), then growth in the presence of Cm, where growth rate depends on the public-good concentration (from  $t_1$  to  $t_2$ ). We modelled the effect of the public good in a simplified way by assuming that it is proportional to the proportion of producing cells in the local population. During the pre-incubation step, the growth rate  $\psi$  is the constant rate  $\psi_1$ . In the presence of Cm, the growth rate  $\psi$  noted  $\psi_2$  depends on the proportion of producers as follows:  $\psi_2 = \psi_0(1 + b \times \frac{P^+}{n_{tot}})$ . General equations for growth and transfer are common to both steps. We modelled two cases, with or without amplification of transfer by recipients (that become secondary donors).

With amplification, initial and secondary plasmid donors are not distinguished:

$$\begin{aligned}\frac{dP^+}{dt} &= [\psi(1 - c) + \gamma_P^+ R] \times P^+ \times (1 - \frac{n_{tot}}{K}) \\ \frac{dP^-}{dt} &= [\psi + \gamma_P^- R] \times P^- \times (1 - \frac{n_{tot}}{K}) \\ \frac{dR}{dt} &= [\psi - (\gamma_P^+ P^+ + \gamma_P^- P^-)] \times R \times (1 - \frac{n_{tot}}{K}) \\ n_{tot} &= P^+ + P^- + R\end{aligned}$$

Without amplification of transfer in recipients, initial plasmid-bearers (transmitters, T) and recipients (R) are considered as separate genotypes:

$$\begin{aligned}\frac{dTP^+}{dt} &= \psi(1 - c) \times TP^+ \times (1 - \frac{n_{tot}}{K}) \\ \frac{dTP^-}{dt} &= \psi \times TP^- \times (1 - \frac{n_{tot}}{K}) \\ \frac{dR}{dt} &= [\psi - (\gamma_P^+ TP^+ + \gamma_P^- TP^-)] \times R \times (1 - \frac{n_{tot}}{K}) \\ \frac{dRP^+}{dt} &= [\psi(1 - c) \times RP^+ + \gamma_P^+ \times TP^+ \times R] \times (1 - \frac{n_{tot}}{K}) \\ \frac{dRP^-}{dt} &= [\psi \times RP^- + \gamma_P^- TP^- \times R] \times (1 - \frac{n_{tot}}{K}) \\ n_{tot} &= TP^+ + TP^- + R + RP^+ + RP^-\end{aligned}$$

The pre-incubation time from  $t_0$  to  $t_1$  was set to 12h after 100-fold initial dilution from stationary phase cultures at carrying capacity, and growth in presence of the antibiotic was allowed for 60h after a second 100-fold dilution. To analyze the effect of pure infectious

transfer (Figure 2.2), changes in  $P^+$  proportion were computed from  $t_0$  to  $t_1$ . When cooperation was involved (all other simulations), changes in  $P^+$  proportion were computed from  $t_1$  to  $t_2$ . As the public good acts on growth rate, the benefit of cooperation is only transient [Chuang et al., 2009] and  $t_2$  has to be chosen before all populations reach stationary phase. For each simulation,  $t_2$  is defined as the time point where the selection of  $P^+$  is maximal over all conditions tested.

Parameters values are shown below and were estimated from our experimental data:

Parameter	Symbol	Estimation
Carrying capacity	$K$	$4.10^9$ cells $\text{mL}^{-1}$
Growth rate in the absence of Cm	$\psi_1$	$0.96 \text{ h}^{-1}$
Growth rate coefficient in the presence of Cm	$\psi_0$	$0.12 \text{ h}^{-1}$
Benefit of cooperation on growth rate	$b$	4
Cost of public good production on growth rate	$c$	0.04

The rate of transfer  $\gamma$  ( $\gamma_P^+$  or  $\gamma_P^-$ ) was varied from 0 to  $2.10^{-9} \text{ mL.cell}^{-1}\text{L}^{-1}$ , which encompasses the range of transfer rates that can be measured, and knowing that derepressed plasmids transferring at a rate around  $10^{-9} \text{ mL.cell}^{-1}\text{L}^{-1}$  [Levin et al., 1979]. This rate was divided by 10 in the presence of Cm, to mimic experiments where growth in Cm happens at  $30^\circ\text{C}$  and transfer is reduced. We did not attempt to measure  $\gamma$  experimentally, as the transfer rate is not constant during the duration of the experiments (because of successive dilutions and shifts in temperature). However, we can estimate an effective transfer rate leading to the same plasmid invasion than observed at  $t_1$  in experiments. Effective transfer rates vary from  $5.10^{-8}$  to  $10^{-9} \text{ mL.cell}^{-1}\text{L}^{-1}$  in our experiments.

To study the effect of relatedness on the selection of cooperation (Figure 2.4-C), plasmid-bearing cells that arose from transfer ( $RP^+$  and  $RP^-$ ) were distributed in equal proportions in the two populations at  $t_1$ , keeping all other population parameters constant, and relatedness was computed after mixing. To study the effect of strong cell dilution (Figure 2.6), cells were distributed in 96 populations each of  $10 \mu\text{L}$ , following a Poisson distribution of parameter  $\lambda$  ( $\lambda \in [1 : 15]$  for  $P^+$  and  $P^-$ ) and  $98 \times \lambda$  (for  $R$ ), ensuring an initial frequency of 2% plasmids. Because of the strong initial dilution, the pre-incubation time was set at 24h and a second 10-fold dilution step was added before  $t_1$ . Results were averaged over 20 replicate simulations, as strong variance arises from Poisson distribution. Computer simulations were conducted using MATLAB.





## Chapter 3

# Indirect selection on chromosomal genes for horizontal transfer

### 3.1 Introduction

Bacteria bear numerous mobile elements that have the ability to undergo horizontal transfer between unrelated hosts [Frost et al., 2005]. Mobile elements are the main actors of genetic exchange among bacteria, because they carry not only the genes involved in their own maintenance and transmission, but also additional "accessory" genes. The collection of accessory genes is not random: some gene classes are overrepresented on mobile elements. Specifically, plasmids often encode optional traits, beneficial only in some environments [Eberhard, 1989], such as genes for antibiotic resistance or specific substrate metabolism. Mobile genes may also be involved in social interactions between cells [Rankin et al., 2010], such as the well-known example of  $\beta$ -lactamases, responsible for extracellular degradation of antibiotics [Yurtsev et al., 2013], or the secreted cholera toxin [Waldor and Mekalanos, 1986]. Bioinformatic studies have confirmed that genes coding for secreted proteins, ones likely to participate in social interactions, are overrepresented on mobile elements [Nogueira et al., 2009]. The observed enrichment in some classes of genes raises the question of gene mobility : why are some genes more mobile than others [Eberhard, 1989, Rankin et al., 2010] ?

Classically, horizontal transfer by conjugation has been considered as a parasitic feature promoting selfish plasmid spread [Turner et al., 1998]. Indeed, plasmid conjugation is costly to the host: it decreases host growth rate and fitness [Turner et al., 1998, Haft et al., 2009], and makes cells sensitive to male-specific phages [Jalasvuori et al., 2011]. At the gene level, mobility could allow accessory genes to persist in the environment in the absence of continued selection for these genes in a given local population [Bergstrom et al., 2000]. In a structured environment with sporadic selection, genes can be lost from a given population in the absence of selection but maintained in neighbouring populations from where they can re-infect the focal population. This hypothesis, that focuses on selection

acting at the level of plasmid genes, can explain the prevalence of optional genes on selfish plasmids - these genes would particularly benefit from mobility - but not the prevalence of social genes on plasmids. Also, in this scenario, the localization of specific genes on mobile elements does benefit the plasmid, but not directly the host.

On the other hand, the transfer of accessory genes could be beneficial for the host. Indeed, we must take into account that the mobility of a given gene can be influenced by other chromosomal - less mobile - genes. Both plasmidic and chromosomal factors are involved in the regulation of transfer in donor cells [Frost and Koraimann, 2010]. Diverse defense mechanisms like restriction-modification systems [Luria and Human, 1952] or CRISPRs [Marraffini and Sontheimer, 2008] also control the establishment of plasmids in recipient cells. Actual transfer rates have been shown to vary greatly, up to eight orders of magnitude for the same plasmid depending on host background [Gordon, 1992, Dionisio et al., 2002], and both donor and recipient ability have been shown to be extremely variable. The strong variation in host donor and recipient abilities remains to be explained, and suggests that selective pressures can act on host genes to modify transfer rates. Two potential benefits of transfer to the host have been highlighted:

Firstly, mobile elements have been described as constituting a "communal pool" of genes [Norman et al., 2009] that are available, when needed, to bacteria. The localization of accessory genes on plasmids and their horizontal transfer would generally benefit the whole population, which would promote horizontal transfer. Recipient ability for beneficial accessory genes can be directly selected here, as transfer to a recipient cell will directly affect its fitness. For instance, CRISPR immunity was shown to be lost in recipient cells in the presence of beneficial plasmids [Jiang et al., 2013]. However, benefits for donor hosts are less obvious: the donor cell does not directly benefit from transferring genes, and favoring competitors by providing them a beneficial allele does not seem to be profitable. Population-wide benefits are thus not sufficient to explain how host genes promoting costly transfer would be selected [Rankin et al., 2010]. Selection acting on donor ability has to be analyzed more precisely in a social context, taking into account both the costs and the benefits of plasmid transfer for a given host genotype.

Secondly, transfer can effectively manipulate recipient's behaviour to its neighbour's benefit, particularly in the case of social genes. One major social behaviour is cooperation, that includes public good production: the maintenance of cooperation is threatened by the spread of cheaters, individuals that benefit from cooperation without contributing to it [West et al., 2007a]. We have shown experimentally (Chapter 2) that the horizontal

transfer of public good genes promotes the maintenance of public good cooperation in two ways. Firstly, horizontal transfer can enforce cooperation in previously non-producing cells [Smith, 2001]. Secondly, in a structured population, horizontal transfer promotes cooperation even when non-producing plasmids also transfer (excluding a direct infectious benefit for producing plasmids). The effect on cooperation is due to the local properties of transfer, that increase relatedness for mobile alleles: new producer cells, arising from transfer, are more strongly associated to other producers, favoring cooperation [Nogueira et al., 2009]. Transfer promoting cooperation suggests there may also exist an indirect benefit of transfer when cooperative behaviours are beneficial, allowing for second-order selection of transfer through cooperation.

We thus focus on hosts, and analyze the selective pressures acting on chromosomal genes controlling transfer ability, in order to understand if and under which conditions the hosts can benefit from plasmid transfer. We first analyze theoretically selective pressures in a general case where transferred plasmids can carry both private goods and public goods genes, then distinguish between private and public goods for further analysis. Moreover, we analyze separately the selective pressures acting on donor and recipient abilities. We first use a general, direct fitness approach to generate predictions, which we test with further experiments and simulations by competing strains that differ strongly in either donor or recipient ability.

## **3.2 Selection of chromosomal control of transfer: direct fitness approach**

Here we model the consequences on host fitness of small changes in transfer rate, under the control of host genes that are not transferred themselves. For the sake of simplicity, we assume that the basal transfer rate is common to all plasmids, with no changes in plasmid genes controlling transfer.

### **3.2.1 Description of the model**

Our model uses a neighbour-modulated fitness approach, which partitions fitness into the effect of an individual's own genotype and the effect of social neighbors [Taylor and Frank, 1996]. We assume that the population is structured in distinct patches, and both the effects of social behaviours and horizontal transfer happen within patches. Indeed, conjugation requires cell-cell contact, so by definition it happens among neighbours. Furthermore,

we assume that there is only one plasmid type in the population. Each individual in the population is characterized by three different traits:

(1) The presence and effects of the plasmid in the cell. The plasmid may produce public and private goods, that translate into a private effect  $x_p$  on its host, and a public effect  $y_p$  on all individuals in the patch. In the case of public good production genes,  $x_p$  is negative (cost of public good production), and  $y_p$  is positive (benefit of public good presence). In the case of private good genes,  $x_p$  is positive and  $y_p$  is zero.

(2) The donor ability of the host,  $q$ . Donor ability confers a direct cost  $c_q$  to the host. The effect of donor ability at the group scale will depend on effective transfer in the patch and plasmid effects in recipient cells, and is thus not explicitly modelled.

(3) The recipient ability of the host,  $r$ . Recipient ability has a direct effect  $c_r$  to the host which may be negative (cost of receiving) or positive (cost of immunity in cells with low recipient ability). The effect of recipient ability at the group scale will, like the one of donor ability, depend on effective transfer and plasmid properties.

Notations concerning those three traits are summarized in Table 3.1. We indicate individual-level behaviours with lowercase letters, and patch-level behaviours (arising from the sum of individual's behaviours in the patch) with uppercase letters.

Property	individual behaviour	patch average	individual effect	group effect
good production	$p$	$P$	$x_p$	$y_p$
donor ability	$q$	$Q$	$c_q$	$\Delta P_t, \Delta p_t$
recipient ability	$r$	$R$	$c_r$	$\Delta P_t, \Delta p_t$

Table 3.1: **Notations of the direct fitness model.** 'Individual behaviour' indicates the intensity of trait expression in the focal cell, 'patch average' the mean of trait expression within a patch. All individual and group effects are effects on fitness  $W$ .  $\Delta P_t$  and  $\Delta p_t$  are proportions of cells that received transferred plasmids: they are not explicitly modelled, but derived from genotype's frequencies within patches (see text, Equations 3.2 and 3.3).

We assume that plasmid and host traits are distributed independently in the population, *ie* plasmid presence in a cell is independant from the cell's genotype concerning donor or recipient ability.

We model first the effects of the plasmid genes, that can include genes coding for both private and public goods, on the host fitness ( $W$ ). The private effect of plasmid genes on cells bearing the plasmid is proportional to their level of expression ( $x_p \times p$ ). At the patch-level, public good production modulates population size (elastic population regulation

with public goods), and the effect of plasmid public good production is proportional to the mean public good production in the patch ( $y_p \times P$ ). Thus,

$$W = 1 + y_p \times P + x_p \times p \quad (3.1)$$

Transfer modifies fitness directly, through the direct costs of transfer on the cell ( $c_q$  and  $c_r$ ), and indirectly, through the effects of the transferred plasmids: the indirect effects of transfer are the changes in private and public effects of the plasmid genes in recipient cells due to plasmid transfer. For simplicity, we assume that the plasmid effect is binary: cells can have either  $p = 1$  (plasmid-bearing cells) or  $p = 0$  (plasmid-free cells), and  $P$  is equivalent to the proportion of plasmid-bearing cells in the patch. Transfer happens within patches, and leads to the appearance of new plasmid-bearing cells from previously plasmid-free cells, at a rate proportional to cell donor and recipient abilities, and following a mass-action law: it is proportional to the density of both plasmid-bearing and plasmid-free cells [Levin et al., 1979]. Neglecting secondary transfer from recipients, we thus consider that for a genotype with  $q$  donor ability and  $r$  recipient ability, the number of transfer events is  $q \times r \times g \times P(1 - P)$ ,  $g$  being the basal plasmid transfer rate.

For simplicity, we also assume that the total donor ability in a patch  $Q$  is the weighted average of  $q$  for all genotypes present in the patch in proportion to their frequency, and the total recipient ability in a patch  $R$  is the weighted average of  $r$  for all genotypes present in the patch in proportion to their frequency. This is approximately the case when tested in simulations in the absence of secondary transfer. Secondary transfer from recipients leads to amplification of transfer, and the overall efficiency of transfer gets higher than predicted by the arithmetic averages. Secondary transfer also has an impact on the total amount of transfer across patches, as there is more transfer in patches enriched in cells with high transfer abilities. We neglect this phenomenon here, but will analyze it in simulations.

We use  $\Delta P_t$  to mark the proportion of cells that received the plasmid in a patch because of transfer, and  $\Delta p_t$  the proportion of cells sharing the focal genotype for host traits ( $q$  and  $r$ ) that received the plasmid in the patch. The mean proportion of cells that received the plasmid following transfer in a patch  $\Delta P_t$  is with our assumptions:

$$\Delta P_t = Q \times R \times g \times P(1 - P) \quad (3.2)$$

where  $P$  and  $(1 - P)$  are respectively the frequencies of plasmid-bearing and plasmid-free cells in the patch.

However, for cells sharing chromosomal alleles, the probability of receiving a plasmid  $\Delta p_t$  does depend on their own recipient ability  $r$  instead of the mean recipient ability in the patch  $R$ , thus:

$$\Delta p_t = Q \times r \times g \times P(1 - P) \quad (3.3)$$

Using  $P_0$  and  $P_t$  to mark the global proportion of plasmid-bearing cells in the population before and after transfer, and  $p_0$  and  $p_t$  to mark the proportion of plasmid-bearing cells in cells with the focal genotype before and after transfer, yields:  $P_t = P_0 + \Delta P$  and  $p_t = p_0 + \Delta p$ . Substituting Equation 3.2 and Equation 3.3 into Equation 3.1 (considering fitness and thus proportions of plasmids after transfer), and adding direct effects of transfer on fitness, yields:

$$W = 1 + y_p[P + Q \times R \times g \times P(1 - P)] + x_p[p + Q \times r \times g \times P(1 - P)] - c_q q - c_r r \quad (3.4)$$

Here we explicitly express fitness as a function of plasmid presence as well as added effects that are due to plasmid transfer.

### 3.2.2 Selection of chromosomal donor ability

#### Model analysis

We analyze the consequences on fitness of a small change in donor ability  $q$ , by calculating the derivative of  $W$  depending on  $q$  from 3.4. We obtain:

$$\frac{dW}{dq} = g \times P(1 - P) \times [y_p R + x_p r] \times \frac{dQ}{dq} - c_q \quad (3.5)$$

$g \times P(1 - P)$  term acts on the efficiency of transfer: transfer is proportional to plasmid basal transfer rate, plasmid-bearing and plasmid-free cell frequencies. The effect of transfer on fitness is a function of actual transfer events, and is stronger when  $P(1 - P)$  is high, *ie* when the plasmid is present in patches, but has not completely invaded them. We distinguish between global and individual recipient ability: global recipient ability within a patch affects the total transfer events in a patch, and thus the amount of public good produced by the plasmid ( $y_p R$ ); individual recipient ability affects the individuals plasmid-bearing status, and thus the private effects of the plasmid ( $x_p r$ ).

The  $\frac{dQ}{dq}$  term, the change in patch donor ability given a change in individual donor

ability, corresponds to the relatedness or assortment among cells concerning donor ability  $q$ .  $\frac{dQ}{dq}$  is positive when cells with higher donor ability within patches are associated with each other more than randomly. Since all effects of transfer events are weighted by this relatedness among donors, some structure in the population must exist in order for benefits of transfer to feed back on donor cells.

The last term  $c_q$  is the direct effect of transfer on the cell. When transfer is costly,  $c_q < 0$ . For an increase in transfer ability to be selected ( $\frac{dW}{dq} > 0$ ), benefits of transfer ( $y_p R + x_p r$  when beneficial) weighted by actual transfer  $g \times P(1 - P)$  and relatedness among donors  $\frac{dQ}{dq}$  have to outweigh the cost of transfer  $c_q$ .

### Interpretation

Except the direct cost of transfer  $c_q$ , all effects are indirect, as they act through the transfer to recipients, and play a role on donors only through the assortment between donor cells. Indirect effects can then be separated into private and public effects:

(1)  $y_p$  is the public effect of the plasmid. Its effect on the selection of transfer increases as a function of global transfer in the patch, which yields more public good production. From the donor's side, transfer to a related or non-related cell are equivalent.

(2)  $x_p$  is the private effect of the plasmid on the recipient. Its effect on the selection of donor ability increases as a function of the probability of transfer to a recipient of the same genotype. If the genotype's recipient ability  $r$  is 0, the private effect of the plasmid will not play any role in the selection of donor ability (transfer to kin is not possible). If its recipient ability  $r > 0$ , selection on donor ability will be modified by the direct effect of the plasmid on fitness of related recipients.

We can focus on two extreme cases of transferred plasmids.

(1) Private effects only. For a plasmid with only private effects on its host ( $y_p = 0$ ), the change in fitness depending on  $q$  is now (from Equation 3.5):

$$\frac{dW}{dq} = g \times P(1 - P) \times r \times x_p \times \frac{dQ}{dq} - c_q \quad (3.6)$$

Here, selection on transfer happens through plasmid effect  $x_p$  on the fitness of recipients. For related recipients (when  $\frac{dQ}{dq} > 0$ , so transfer is biased towards kin as cells sharing high donor ability are clustered together), increasing transfer can be positively selected for a plasmid with beneficial effects on fitness of the carrying cells, as it favors relatives of the donor cells, providing that these cells have some recipient ability for the plasmid



( $r > 0$ ).

(2) Public benefits only. For a plasmid with public effects, for instance one that codes for the production of a public good, we have  $y_p > 0$  (benefit of the public good) and  $x_p \leq 0$  (cost of public good production). Here, increasing transfer ability can be selected when the total amount of transfer is high, leading to a high level of public good production in the patch. But this will be hindered by the decrease in fitness caused by paying the production cost in related cells that received the plasmid.

Equation 3.5 can be seen as a form of second-order Hamilton's rule concerning donor ability: increasing donor ability can be selected when the benefits of transfer, weighted by relatedness between donors, outweigh the cost of transfer. This is complicated by the fact that transfer acts in two ways, directly on recipients's fitness, and indirectly on total public good production.

Moreover, actual transfer depends on population structure parameters, particularly the coexistence of both plasmid-bearing and plasmid-free cells within patches. A major issue here is the possibility for the existence of population structures that would lead to the selection of transfer. Indeed, the selection of transfer requires two apparently opposed conditions. Firstly, the population has to be structured in a way that results in a positive relatedness (preferential interactions) between the cells sharing donor ability ( $\frac{dQ}{dq} > 0$ ). Secondly, for transfer to happen, plasmid-bearers must encounter plasmid-free cells in order for  $P(1 - P)$  to be maximized, which will be hindered by a strong spatial population structuring where plasmid-bearing and plasmid-free cells will be separated. I will address this issue by conducting simulations described below (see section 3.5).

### 3.2.3 Selection of recipient ability

#### Model analysis

We analyze the consequences of a small change in recipient ability  $r$  on fitness, by calculating the derivative of  $W$  depending on  $r$  from Equation 3.4. We obtain:

$$\frac{dW}{dr} = g \times P(1 - P) \times Q \times [y_p \frac{dR}{dr} + x_p] - c_r \quad (3.7)$$

$c_r$  is the direct effect of recipient ability.  $c_r$  is positive if the immunity against plasmid entry is an active and costly mechanism.

$g \times P(1 - P) \times Q$  term modifies the efficiency of transfer: transfer is proportional to plasmid basal transfer rate, plasmid-bearing and plasmid-free cells frequencies, and global

donor ability in the patch. The effect of transfer on fitness is function of actual transfer events, and gets stronger when  $P(1 - P)$  increases.

Plasmid transfer acts in two ways: first through public good production, weighted by the relatedness between cells sharing the same recipient ability ( $y_p \frac{dR}{dr}$ ). Increased public good production following the appearance of new producers benefits all cells in the patch. Secondly, transfer acts through the private effect of the plasmid on fitness of recipients ( $x_p$ ). As the effect is direct, there is no weighting by relatedness among recipients here.

### Interpretation

Contrary to donor ability, under these conditions the recipient ability can be directly selected, when the transferred plasmid increases direct fitness of its host ( $x_p > 0$ ). For instance, this could be the case of antibiotic resistance plasmids in the presence of antibiotics: cells able to receive the plasmids will be directly selected for, without need for assortment among recipients.

In the case of a plasmid coding for public good production, however, we have  $x_p < 0$  and  $y_p > 0$ . Here, benefits of cooperation can select indirectly for recipient ability, if relatedness among cells sharing recipient ability ( $\frac{dR}{dr}$ ) is greater than zero: benefits from the public goods produced in neighbouring plasmid recipients feedback on recipient cells.

As for the selection of donor ability, the structure of the population seems to have opposite effects on relatedness and on the possibility for transfer to happen. This will be addressed in simulations.

### 3.2.4 Summary of the direct fitness model

We summarize here the selective pressures acting on the chromosomal control of transfer in our model, by separating the direct and indirect consequences of transfer on fitness, and separating further the effect of transferred plasmids on fitness (Table 3.2):

Trait	Direct effect	Effect through transfer $\times gP(1 - P)$		
		on cell fitness		via public good
		of the focal cell	of neighbours	
donor ability	$c_q$	-	$r \frac{dQ}{dq} x_p$	$R \frac{dQ}{dq} y_p$
recipient ability	$c_r$	$Q x_p$	-	$Q \frac{dR}{dr} y_p$

Table 3.2: **Selective pressures acting on chromosomal transfer control** for donor and recipient abilities. Effects acting through transfer are weighted by the factor  $gP(1 - P)$  that acts on transfer efficiency.

(1) Changing both donor and recipient abilities has a direct effect on fitness, linked to the mechanistic cost of modifying cell donor or recipient properties.

(2) Transferred plasmids have a direct effect on cell fitness in recipients. This translates into a direct effect of changing recipient ability for a given genotype. This also leads to an indirect effect of changes in donor ability through the effect on the donor's neighbours that receive the plasmid, when these neighbours are related to the focal donor cell.

(3) Finally, effects occurring through public good production are similar for donor and recipient abilities: both affect fitness generally when relatedness is non-zero.

In the model, we assumed very small changes in donor or recipient ability. However, mutations affecting these traits can have a large effect. Moreover, this model does not explore the population structure needed for relatedness to occur, or for transfer to be maximized, and neglects the potential associations between  $q$ ,  $r$  and  $p$  traits. In order to investigate this, we move to simulations and experimental tests, where we now compete strains that can bear a plasmid with genes involved in the production of either a private good or a public good, and also differ significantly in transfer ability. Simulations explicitly follow the experimental system, with realistic parameters derived from our experiments (see Methods), and results from experiments and simulations are thus presented together.

### 3.3 Selection for transfer in a simple metapopulation

Here we analyze the outcome of competition between two strains differing in transfer abilities in a simple metapopulation that consists of two well-mixed populations with arbitrarily chosen competitor frequencies. This set-up provides non-random assortment, or relatedness, at the metapopulation scale. We start by assuming a given initial proportion of plasmid-free cells, but will later vary proportions of plasmid-free cells.

#### 3.3.1 Selection for transfer of private good genes

In our experiments we use antibiotic resistance genes, which are very often present on plasmids, as an example of private good genes. The plasmids  $K$  carry the *aph* gene coding for the Aph enzyme, which inactivates the antibiotic kanamycin and provides resistance. Aph could theoretically also have a public effect, as it decreases antibiotic concentration in the growth medium, but at the antibiotic concentrations used, resistant cells provide no or very low protection to sensitive neighbours. The benefit of carrying a plasmid is

extreme in presence of antibiotics: plasmid-free cells die and only plasmid-bearing cells can survive and grow. In experiments described below, we focus on the selective pressures acting on donor ability.

### Design of the study

We focus on two strains that differ in the ability to donate the plasmids  $K$ :  $D^+$  strain transfers plasmids with donor ability  $q$ ,  $D^-$  strain does not (Figure 3.1-A). For both  $D^+$  and  $D^-$ , the same initial proportion of cells (called respectively  $D_p^+$  and  $D_p^-$ ) initially carries  $K$ . In experiments,  $D^+$  et  $D^-$  carry initially two versions of  $K$ , respectively  $K$ -YFP and  $K$ -GFP, that can be distinguished as they bear respectively YFP and GFP fluorescence genes. The metapopulation consists of two populations,  $s_1$  and  $s_2$ , differing in their  $D^+$  to  $D^-$  ratio:  $s_1$  is enriched in  $D^-$  cells, and  $s_2$  in  $D^+$  cells (Figure 3.1-B). Cells are mixed ( $t_0$ ), then transfer and growth take place within populations. After growth to saturation ( $t_1$ ), populations are pooled and diluted 100-fold in kanamycin-containing medium, where only plasmid-bearing cells are able to grow, until saturation ( $t_2$ ).

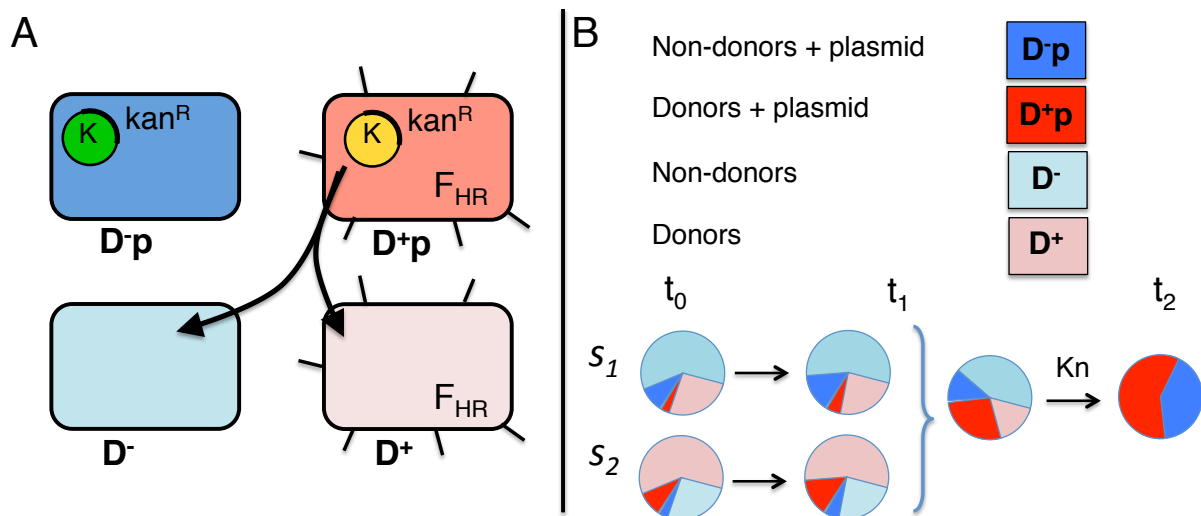


Figure 3.1: **Selection of donor ability for private good genes: experiment design.** **A: Experimental strains.** The two competing strains ( $D^+$  and  $D^-$ ) differ in donor ability for plasmids  $K$  that code for antibiotic resistance ( $kan^R$ ), and fluorescence (plasmid background color).  $D^+$  bears  $F_{HR}$  plasmid that confers high donor ability without interfering with recipient ability (indicated by pili),  $D^-$  does not bear  $F_{HR}$  and can receive plasmids but not transfer them.  $D_p^+$  transfers  $K$  plasmid to both  $D^-$  and  $D^+$  plasmid-free cells (black arrows). **B: Experimental setup.**  $D^+$  and  $D^-$  are competed: initially ( $t_0$ ), a small proportion of  $D^+$  and  $D^-$  cells ( $D_p^+$  and  $D_p^-$ ) bear  $K$  plasmids. The metapopulation consists of two populations,  $s_1$  and  $s_2$ , that differ in  $D^+$  vs  $D^-$  ratio. After growth and transfer from  $D_p^+$  ( $t_1$ ), populations are pooled and antibiotic (Kn) selection is applied: only  $K$ -bearing cells can grow to saturation ( $t_2$ ).

## Results

We follow the change in frequency of  $D^+$  strain from  $t_0$  (the beginning of the competition) to  $t_1$  (after transfer, before antibiotic selection) to  $t_2$  (after antibiotic selection), when  $K$  plasmids are initially present in 1% of both  $D^+$  and  $D^-$  cells. In simulations, when

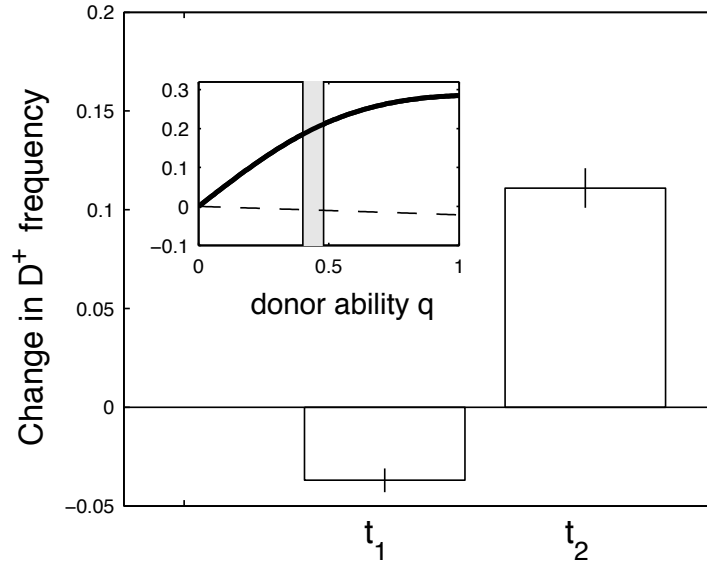
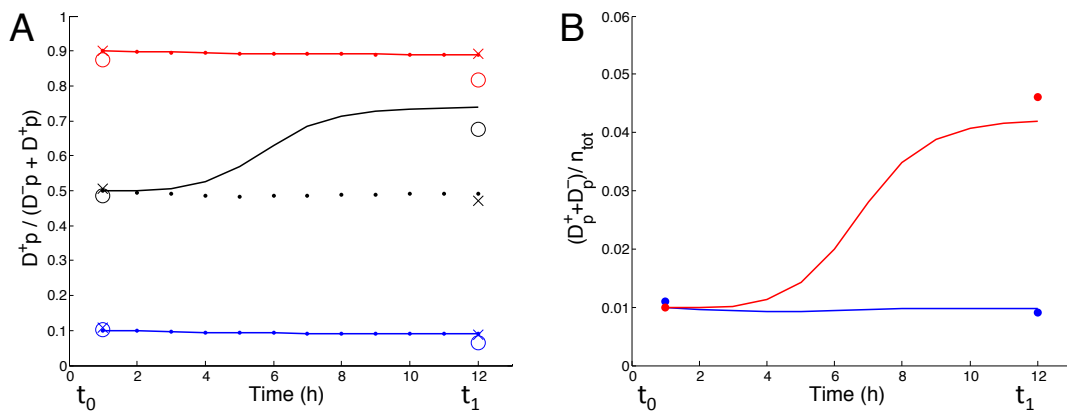


Figure 3.2: **Change in frequency of  $D^+$  strain transferring private good genes.** The metapopulation consists of 2 populations, with initial  $D^+/D^-$  ratios of 1/9 and 9/1, and 1%  $K$ -bearing cells for each strain. The change in frequency of  $D^+$  compared to  $D^-$ ,  $D^+/(D^+ + D^-)$  is computed from  $t_0$  to  $t_1$  (before antibiotic selection) or to  $t_2$  (after antibiotic selection) for the pool of  $s_1$  and  $s_2$  populations, and shown as means  $\pm$  s.e.m. ( $N=6$ ). The inset shows the simulated change in  $D^+$  frequency as a function of its donor ability  $q$ , from  $t_0$  to  $t_1$  (dashed line) and  $t_0$  to  $t_2$  (solid line), the shaded area indicating estimated experimental transfer rates.

measured at timepoint  $t_1$   $D^+$  frequency slightly decreases with an increase in donor ability (Figure 3.2 inset, dashed line) but at  $t_2$  it increases with an increase in donor ability (Figure 3.2 inset, solid line). Experimental results show the same pattern:  $D^+$  strain decreases in frequency at  $t_1$  compared to  $t_0$  (-4%,  $p=0.002$ ), but then increases at  $t_2$  (11% increase compared to  $t_0$ ,  $p=0.01$ ). The decrease from  $t_0$  to  $t_1$  is caused in simulations by the cost of transfer, assumed to be proportional to donor ability. In experiments,  $D^+$  bears a mutant of F plasmid,  $F_{HR}$  which may impose a strong cost on the carrying cell in comparison to  $D^-$ , which does not bear any version of F. In both cases, the decrease at  $t_1$  suggests that the subsequent increase in frequency at  $t_2$  is due to selection for antibiotic resistance, and not to a direct benefit of donor ability to  $D^+$  strain.

As antibiotics act by selecting for  $K$ -bearing cells, we next analyze the dynamics of  $K$  plasmids to understand the observed increase in frequency of  $D^+$ . Within each

population,  $K$  plasmids are present in  $D^+$  and  $D^-$  in proportion to the frequency of  $D^+$  and  $D^-$  themselves at  $t_0$  and at  $t_1$  (Figure 3.3-A), both in simulations and in experiments. This means that in well-mixed conditions,  $D_p^+$  cells may transfer the plasmid to both  $D^+$  and  $D^-$  cells, in proportion to their frequencies in the population (the slight decrease observed in  $D_p^+$  frequency follows the decrease in the frequency of  $D^+$  strain caused by the cost of transfer for  $D^+$  cells). However, we observe a bias in the transfer to recipients at the metapopulation level:  $D_p^+$  increases in comparison to  $D_p^-$  a few hours after  $t_0$  (when transfer happens) in simulations as well as in experiments (Figure 3.3-A, black line). Since the bias does not happen within populations, it must be due to among-population



**Figure 3.3: Biased transfer in simulations and experiments.** **A:** Frequency of  $D^+p$  vs  $D^-p$  function of time. The black line is the frequency in the metapopulation; blue and red lines are frequencies within populations  $s_1$  and  $s_2$  respectively. Dotted lines indicate frequencies of  $D^+$  vs  $D^-$  (including both plasmid-bearing and plasmid-free cells; they are superposed to  $D^+p$  vs  $D^-p$  frequencies within populations). Crosses indicate experimental ratios for strains, circles indicate experimental ratios for plasmids. **B:** Total proportion of plasmid-bearing cells in the populations  $s_1$  (blue) and  $s_2$  (red). Dots indicate experimental proportions. Simulations and experiments are the same as in Figure 3.2. Simulations focus on a strain  $D^+$  with donor ability  $q = 0.45$ .

dynamics. Indeed, total plasmid transfer is more prevalent in the  $s_2$  population enriched in  $D^+$  cells (Figure 3.3-B), biasing transfer towards  $D^+$  cells at the metapopulation level. Experiments and simulations thus confirm the major prediction of Equation 3.6: donor ability can be selected when effective transfer is biased towards cells sharing high donor ability, which arises through population structure.

### 3.3.2 Selection for transfer of public good genes

Here we examine the scenario in which the transferred plasmid carries public good genes. As an example of public good genes, we use the public good system characterized in

previous studies [Chuang et al., 2009, 2010]. The gene *rhII* encodes the rhII enzyme, which catalyzes the production of a secreted public good, the quorum-sensing  $C_4$ -HSL auto-inducer (see Figure 3.4-A).  $C_4$ -HSL provides faster growth in the presence of the antibiotic chloramphenicol (Cm) in all cells bearing the *cat* gene under control of the  $P_{rhI}$ \* promoter. Based on these properties we use a plasmid with the *rhII* gene as a public-good plasmid ( $P^+$ ), and a plasmid without *rhII* as a control plasmid ( $P^-$ ). Competitions aiming to test the effect of public good production are done in a growth medium containing Cm, where the public good  $C_4$ -HSL is beneficial.

### Design of the study

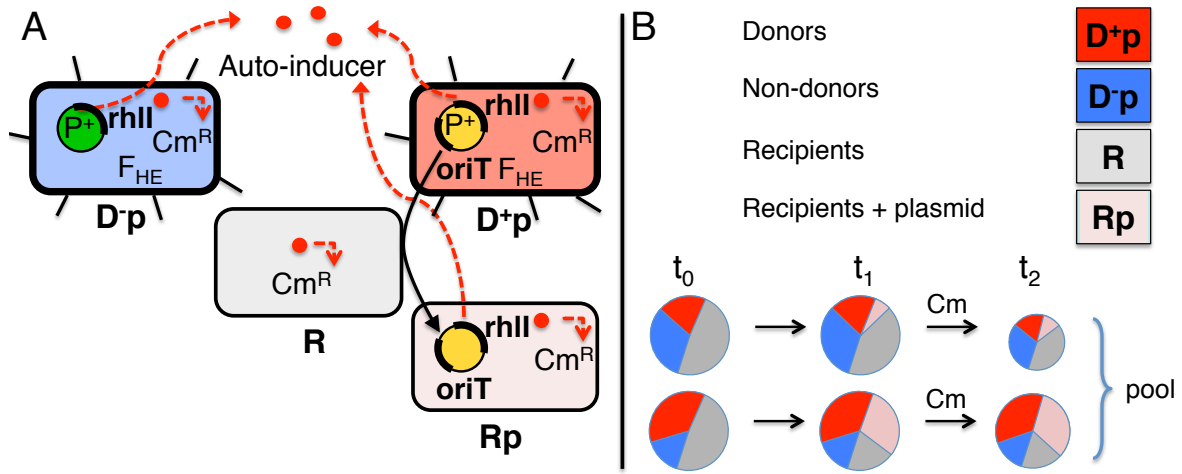


Figure 3.4: **Selection of donor ability for public good genes: experiment design.** **A: Experimental strains.** The two strains  $D^+$  and  $D^-$  carry  $F_{HE}$  plasmid (indicated by pili) that mobilizes *oriT*-bearing plasmids and provides entry exclusion (indicated by thick cell membranes).  $D_p^+$  carries a plasmid containing *oriT* ( $PT^+$  plasmid, see Methods), and can mobilize it to recipient cells  $R$ , yielding  $R_p$  cells.  $D_p^-$  carries a plasmid that cannot be mobilized ( $PT^-$  plasmid, see Methods). Plasmids can be either  $P^+$  or  $P^-$ : the  $P^+$  plasmid coding for public good production is represented. Plasmids also code for fluorescence, indicated by their background color. All cells can respond to the auto-inducer produced by  $P^+$ -bearing cells by expressing antibiotic resistance ( $Cm^R$ ). **B: Experimental setup.**  $D_p^+$  and  $D_p^-$  are competed. At  $t_0$ , the metapopulation consists in two populations differing in their  $D^+$  vs  $D^-$  ratio, with the same proportion of  $R$ . After growth and transfer ( $t_1$ ), cells are confronted to chloramphenicol (Cm), and subsequent growth is dependant on auto-inducer concentration. Populations are then pooled and measured at  $t_2$ .

Based on Equation 3.5, the selection of donor ability for public-good producing plasmids should be possible because of feedbacks from public good production. Equation 3.5 also predicts that transfer to kin does not bring additional benefits - on the contrary, related cells suffer from the additional cost of public good production. The effects on donor

cell fitness of this additional cost ( $c_{pr}$  in Equation 3.5) would be difficult to distinguish experimentally from the direct cost of transfer ( $c_q$  in Equation 3.5). For this reason, we chose to focus experimentally on the selection of donor ability to unrelated recipients. We analyze the case of transfer to kin later, in simulations.

We use distinct strains as primary plasmid-bearers ( $D_p^+$  and  $D_p^-$ , that differ in donor ability) and secondary plasmid-bearers (recipients,  $R$ ). We compare situations which differ by the presence of public good genes on plasmids. In the first condition,  $D_p^+$  and  $D_p^-$  both bear a public-good production plasmid ( $P^+$ ), while in the second they both bear an empty control plasmid ( $P^-$ ).  $D^+$  can transfer the plasmid to recipients ( $R$ , leading to  $R_p$ ),  $D^-$  and  $R$  cells cannot (Figure 3.4-A). The metapopulation consists in two populations differing in their  $D^+$  vs  $D^-$  ratio (Figure 3.4-B). Cells are mixed ( $t_0$ ), then transfer takes place within populations, similarly to the previous section. After transfer ( $t_1$ ), populations are diluted in the presence of Cm, where population growth is function of the concentration of C<sub>4</sub>-HSL produced by  $P^+$ -bearing cells. Finally, at  $t_2$  equal proportions of the two populations are mixed to evaluate frequencies at the metapopulation level.

## Results

We tested experimentally the case in which 75% of cells are initially  $R$  cells. Faster growth is observed in the presence of Cm for populations enriched in  $D^+$  cells when the transferred

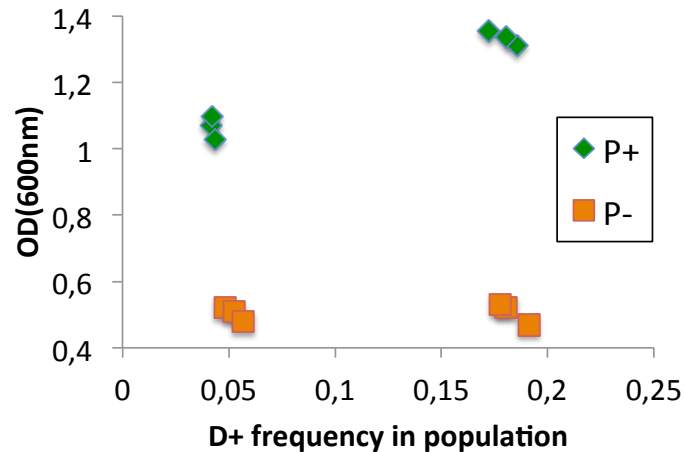


Figure 3.5: **Effect of transfer by  $D^+$  cells on population growth.** The optical density at 600nm ( $OD_{600}$ ) was measured after 14h of growth in the presence of Cm ( $t_2$ ), in two populations differing in  $D^+$  strain proportion, and carrying either a public-good production plasmid ( $P^+$ ) or a control plasmid ( $P^-$ ), that  $D^+$  can transfer.  $OD_{600}$  is shown as a function of  $D^+$  proportion at  $t_1$ . Results shown are three replicate experiments done the same day.



plasmid carries public-good genes ( $P^+$  plasmid) but not when it lacks them ( $P^-$  plasmid) (Figure 3.5). The difference suggests that selection can act among populations when public good genes are transferred: transfer would yield benefits at the group level.

We then analyze the dynamics of the frequency of  $D^+$ , in the experiments as well as in simulations. Simulations (Figure 3.6 inset) predict that  $D^+$  cells are outcompeted when a non-producing plasmid is transferred, and outcompete  $D^-$  cells when production of public goods from the plasmid genes increases. We observe the same behaviour with

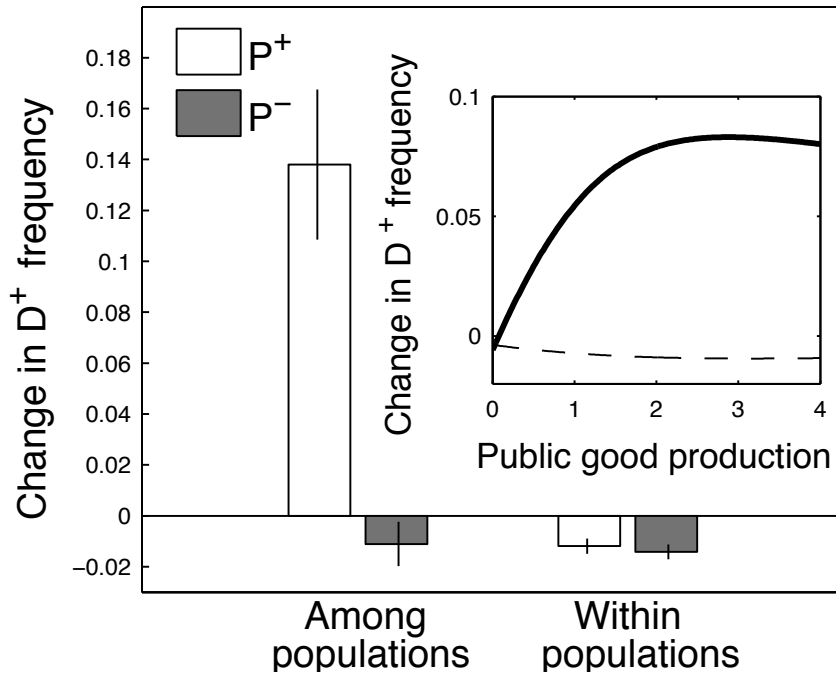


Figure 3.6: **Selection of chromosomal donor ability for public good genes.** The metapopulation consists of 2 populations, with initial  $D^+/D^-$  ratio of 1/4 and 4/1, and 80% R. The change in frequency of  $D^+$  from  $t_1$  to  $t_2$ ,  $D^+/(D^+ + D^-)$ , is computed for strains carrying a plasmid with (open columns) or without (filled columns) public good production, and shown as mean  $\pm$  s.e.m. ( $N=9$ ). The inset shows the simulated change in  $D^+$  frequency as a function of public good production rate by plasmid genes, among (solid line) and within (dashed line) populations.

our experimental system: at the metapopulation scale (among populations, Figure 3.6 left), plasmid donor proportion tends to decrease when they transfer  $P^-$  plasmid (1% decrease among populations in  $D^+$  frequency, not significant,  $p=0.23$ ), but outcompete non-donor cells when they transfer  $P^+$  plasmid (14% increase in the frequency of  $D^+$ ,  $p=0.002$ ). Donor ability is thus selected here because of public good gene transfer.

To confirm that this effect is the result of metapopulation dynamics, we analyze strain dynamics at the population level. The frequency of  $D^+$  compared to  $D^-$  within each population decreases in experiments as well as simulations (Figure 3.6, right and inset),

with a 1% experimental decrease for both  $P^+$  and  $P^-$  plasmid transfer ( $p=0.006$  and  $p=0.001$  respectively): within populations, transfer is costly. The decrease confirms that the selection of donor ability is not a direct benefit of plasmid transfer for the donor cells within populations. However, the difference between among-population and within population changes is strongly significant ( $p=0.001$ ). Thus, the analysis of population dynamics shows that host genes promoting plasmid transfer are selected when the transferred plasmid carries public good production genes. Selection is due to faster growth of populations enriched in plasmid donors, which is strong enough to compensate for the direct cost of transfer for donor cells.

This experiment further shows that some structure in the metapopulation is necessary to observe selection of transfer, as donor ability is not selected within well-mixed populations. The results are in accordance with the role of relatedness among donors in Equation 3.5.

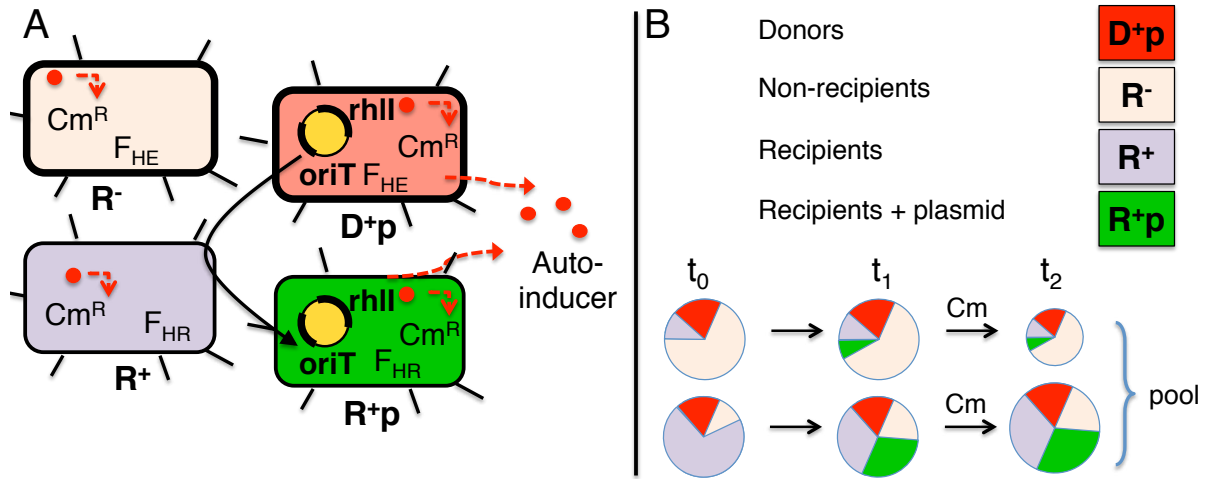
### 3.3.3 Selection of recipient ability for public good genes

Direct selection on recipient ability for private traits has been demonstrated with CRISPR systems [Jiang et al., 2013]. We focus here on recipient ability for public good genes, and use the same plasmids that in section 3.3.2. We assume that variation in recipient ability confers no direct fitness cost to the host, but can have a long-term cost associated to plasmid burden. The public good is costly to produce and provides a population-wide benefit.

#### Design of the study

Two strains  $R^+$  and  $R^-$  are competed here:  $R^+$  cells can receive efficiently plasmids,  $R^-$  cells receive them at a much lower rate. As for the selection of donor ability, we experimentally use a distinct strain as the initial plasmid donor,  $D_p^+$ , now in order to exclude differences in donor ability as reasons for selection. Indeed, we take advantage of the entry exclusion properties of F plasmid to control for recipient ability: the  $F_{HE}$  plasmid expresses functional entry exclusion, so receives plasmids very poorly. The  $F_{HR}$  plasmid is mutated in entry exclusion genes, so receives plasmids at a high rate (see Figure 1.10). As we use different F plasmids in  $R^+$  and  $R^-$  cells to control for recipient ability (Figure 3.7-A), donor ability could also differ if strains were used as both recipients and donors. In simulations, we assume that  $R^-$  cells do not receive plasmids at all.

The metapopulation consists in two populations differing in their  $R^+$  vs  $R^-$  ratio (Fig-



**Figure 3.7: Selection for recipient ability: experimental design. A: Experimental strains.** The two strains  $R^+$  and  $R^-$  do not initially carry plasmids.  $D_p^+$  strain transfers  $P^+$  or  $P^-$  plasmid ( $P^+$  plasmid that codes for public good production is represented here) to  $R^+$  with high efficiency, and to  $R^-$  with very low efficiency, as  $F_{HE}$  plasmid present in  $R^-$  cells expresses entry exclusion genes (represented by a thick black cell membrane). All cells are able to respond to the auto-inducer by expressing antibiotic resistance ( $Cm^R$ ). **B: Experimental setup.**  $R^+$  and  $R^-$  strains are competed. The metapopulation consists initially in two populations differing in  $R^+$  vs  $R^-$  ratio, with the same proportion of  $D_p^+$  cells ( $t_0$ ). After growth and transfer ( $t_1$ ), cells are confronted to chloramphenicol (Cm), and subsequent growth is dependant on auto-inducer concentration. Populations are then pooled and measured at  $t_2$ .

ure 3.7-B). Cells are mixed ( $t_0$ ), then growth and transfer take place within populations. After growth to saturation ( $t_1$ ), populations are diluted in the presence of Cm and further grown until  $t_2$ , similarly to the setup in section 3.3.2.

## Results

We tested experimentally the case in which 20% of cells are initially  $D_p^+$  cells. Differences in recipient ability effectively lead to a strong difference in the proportion of cells that receive the plasmid: in the presence of 20% donor cells, at  $t_1$ , almost no ( $< 2\%$  of the total population)  $R^-$  cells receive the plasmids, whereas 36% of  $R^+$  cells receive it (data not shown). Faster growth is observed in the presence of Cm for populations enriched in  $R^+$  cells when the transferred plasmid carries public-good genes ( $P^+$  plasmid) but not when it lacks them ( $P^-$  plasmid) (Figure 3.8). The difference suggests that selection can act among populations when public good genes are transferred: transfer would yield benefits at the group level.

We next analyze the dynamics of the frequency of  $R^+$  compared to  $R^-$  cells in experiments and simulations when  $D_p^+$  cells are initially 20%. Simulations (Figure 3.9 inset)

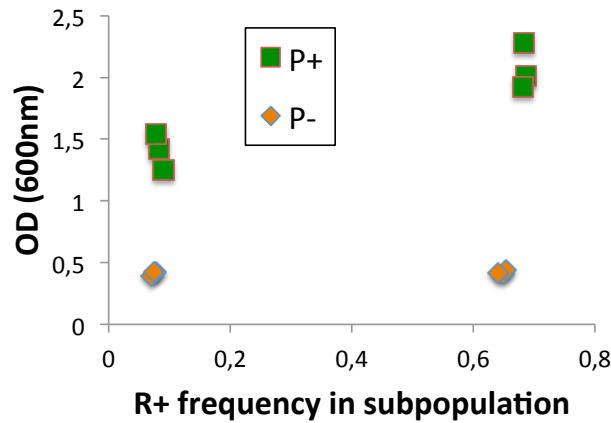


Figure 3.8: **Effect of transfer to  $R^+$  cells on population growth.** The optical density at 600nm ( $OD_{600}$ ) was measured after 14h of growth in the presence of Cm ( $t_2$ ), in two populations differing in the initial  $R^+$  strain proportion, with a donor strain transferring either a public-good production plasmid ( $P^+$ ) or a control plasmid ( $P^-$ ).  $OD_{600}$  is shown as a function of  $R^+$  proportion at  $t_1$ . Results shown are three replicate experiments done the same day.

predict that  $R^+$  cells outcompete  $R^-$  cells when the production of public goods from the plasmid increases. In the absence of public good production,  $R^+$  and  $R^-$  cells do equally well (as we assumed no direct cost for recipient ability, and no cost of the plasmid except the cost of public good production). In experiments, plasmid recipients strongly outcompete non-recipient cells at the metapopulation scale (among populations, Figure 3.7 left) when the  $P^+$  plasmid is transferred (16% increase in the frequency of  $R^+$  from  $t_1$  to  $t_2$ ,  $p=2.9 \cdot 10^{-4}$ ). Here, recipients also outcompete non-recipients when a non-producing ( $P^-$ ) plasmid is transferred (3% increase in the frequency of  $R^+$  from  $t_1$  to  $t_2$ ,  $p=0.029$ ). Still, the difference between  $P^+$  and  $P^-$  transfer is highly significant ( $p=2 \cdot 10^{-4}$ ).

The frequency of  $R^+$  decreases with public good production in simulations within each population: recipient ability for a public good producing plasmid is costly (as we assumed a cost proportional to public good production). This confirms that the selection of recipient ability is not a direct benefit of transfer to the recipients within populations, but is due to among-population selection. In our experiments, there is no decline but a slight increase in  $R^+$  frequency within populations: the cost of public good production added on recipients that received the plasmid seems be negligible on this timescale. Moreover,  $R^+$  and  $R^-$  cells differ in both the  $F_H$  plasmid and fluorescence markers they express: the net cost on fitness may actually be higher for  $R^-$  cells. Still, we can conclude that the main part of the increase in  $R^+$  frequency happens because of among population variations when the transferred plasmid bears public-good production genes: among-

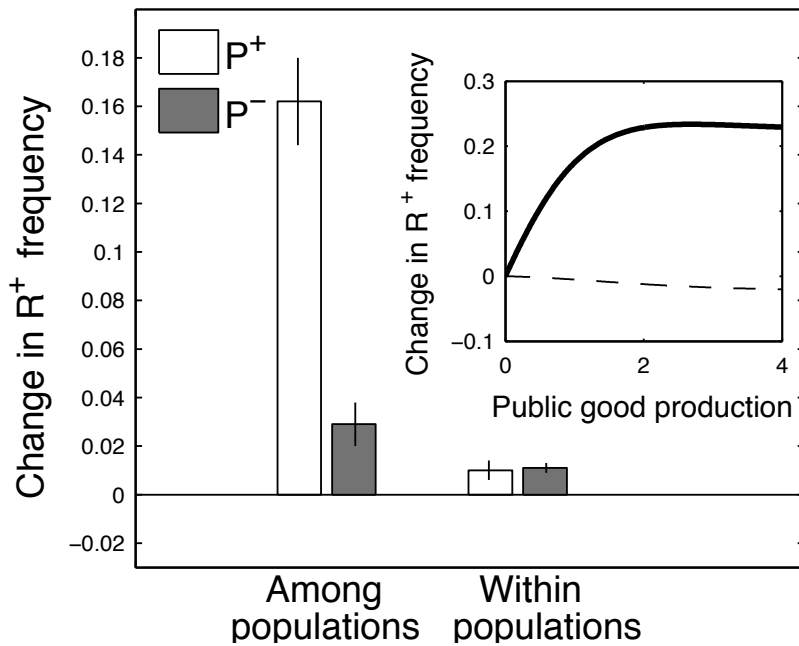


Figure 3.9: **Selection of chromosomal recipient ability for public good genes.** The metapopulation consists of 2 populations, with initial  $R^+/R^-$  ratio of 7/1 and 1/7, and 20%  $D^+p$  cells. The change in frequency of  $R^+$  compared to  $R^-$ ,  $R^+/(R^+ + R^-)$  from  $t_1$  to  $t_2$  is computed for strains carrying a plasmid with (open columns) or without (filled columns) public good production, and shown as means  $\pm$  s.e.m. ( $N=6$ ). The inset shows the simulated change in  $R^+$  frequency as a function of public good production rate of the plasmid transferred by donor cells, among (solid line) and within (dashed line) populations.

and within populations changes are significantly different when  $P^+$  plasmid is transferred ( $p=3.10^{-4}$ ), not when  $P^-$  plasmid is transferred ( $p=0.12$ ). Host genes promoting the cell recipient ability can thus be selected for when the transferred plasmid carries public-good production genes, because of faster growth of populations enriched in plasmid recipients.

Here again, some structure in the metapopulation is necessary to observe selection of transfer: within well-mixed populations, recipient ability is not selected. This is in accordance with the role of relatedness among recipients in Equation 3.7.

### 3.4 Influence of the availability of plasmid-free cells

Both experiments and simulations show that selection of transfer ability is possible. We next analyze in simulations the influence on the selection of transfer of the availability of plasmid-free cells, which is a key parameter governing the effect of transfer efficiency.

### 3.4.1 Transfer of private good genes

We measure the change in  $D^+$  frequency at the metapopulation scale while varying both  $D^+$  donor ability  $q$  and the proportion of plasmid-free cells in  $D^+$  and  $D^-$  strains (Figure 3.10).  $D^+$  frequency increases when both donor ability is high and plasmid-free cells are

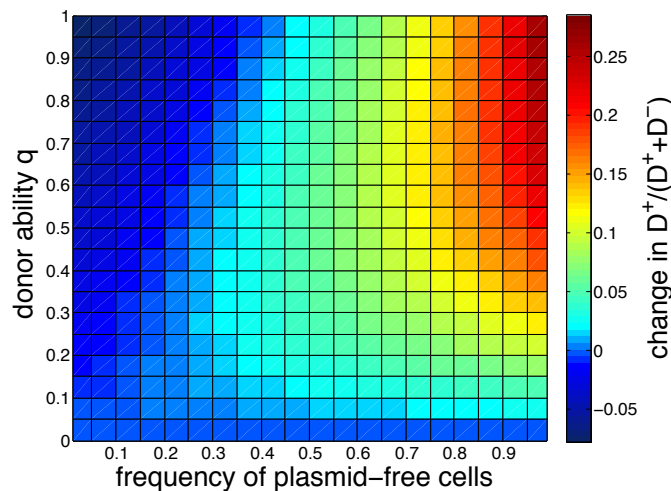


Figure 3.10: **Selection of donor ability for private good genes: effect of the abundance of plasmid-free cells.** Population structure is the same as in Figure 3.2, but now with varying initial frequency of plasmid-free cells. The color scale represents the change in frequency of  $D^+$ ,  $D^+/(D^+ + D^-)$ , computed between  $t_0$  and  $t_2$  (after antibiotic selection) as a function of  $D^+$  donor ability  $q$  and plasmid-free cell frequency at  $t_0$ .

abundant (Figure 3.10).

When less than 10% plasmid-free cells are available,  $D^+$  strain actually decreases in frequency: among-population selection does not compensate for the cost of donor ability. Effective changes in numbers of cells bearing  $K$  are low, and transferred plasmids represent only a small proportion of cells that survived antibiotic selection (for instance, maximally 10% if initially 10% cells are plasmid-free). Moreover, the selection of  $D^+$  is not maximal for an equal frequency of plasmid-bearing and plasmid-free cells, as was suggested from Equation 3.6, but for higher frequencies of plasmid-free cells. The difference between theoretical predictions and simulations results can be explained in two ways. First,  $D^+$  cells that received the plasmid may transfer it again, amplifying transfer until  $t_1$ , which was not modelled in Equation 3.6. Secondary transfer shifts maximal transfer towards lower initial frequencies of plasmid-bearing cells, resulting in maximal change in numbers of cells bearing  $K$  plasmids not around 50% but 60% of plasmid-free cells for low donor ability, and up to 85% of plasmid-free cells for higher donor ability (Figure 3.11-A). Secondly, varying the initial proportion of plasmid-bearing cells modifies not only the

efficiency of transfer, but also the initial proportion of cells that can survive antibiotic selection even in the absence of transfer. Indeed, when only a few plasmids are present initially, the majority of plasmid-bearing cells after transfer actually arises from transfer. However, when more plasmids are present initially, the proportion of plasmid-bearing cells arising from transfer decreases (the proportion of plasmids that arise because of transfer increases with the frequency of plasmid-free cells, Figure 3.11-B). This affects the global importance of transfer in explaining ratios of  $D_p^+$  vs  $D_p^-$  at  $t_1$ , which are then selected in presence of the antibiotic. Donor ability is thus under stronger selection when few cells initially bear the gene upon which selection acts than when such cells are abundant.

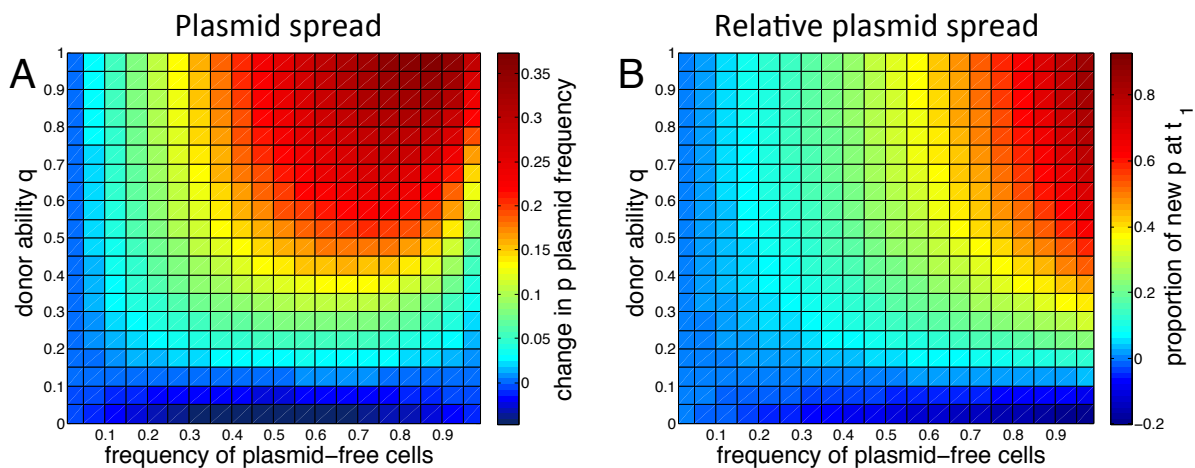


Figure 3.11: **Effect of plasmid initial proportion on the efficiency of transfer.** Results reported here are from the same simulations as in Figure 3.10, with the same axis. In **A**, the colorscale represents the change in total plasmid abundance from  $t_0$  to  $t_1$ ,  $p_{t1} - p_{t0}$  (plasmid spread, that approximates plasmid transfer). In **B**, the colorscale represents the proportion of plasmids present at  $t_1$  that were not present at  $t_0$   $(p_{t1} - p_{t0})/p_{t0}$  (relative plasmid spread).

### 3.4.2 Transfer of public good genes

We now turn to the transfer of public good genes, and first analyze what happens when recipients are a strain distinct from  $D^+$  and  $D^-$  strains, similar to the experimental setup and previous simulations (Figure 3.12-A). Transfer is counter-selected for low proportions of recipients: donor ability is costly and in the absence of efficient transfer, it brings no sufficient benefits to plasmid donors. For higher proportions of recipients, donor ability is selected. Again, this is entirely due to among-population selection, as  $D^+$  frequency always decreases within populations (data not shown). The maximal selection of  $D^+$  depends on both donor ability and initial  $R$  proportion: intermediate transfer rates are

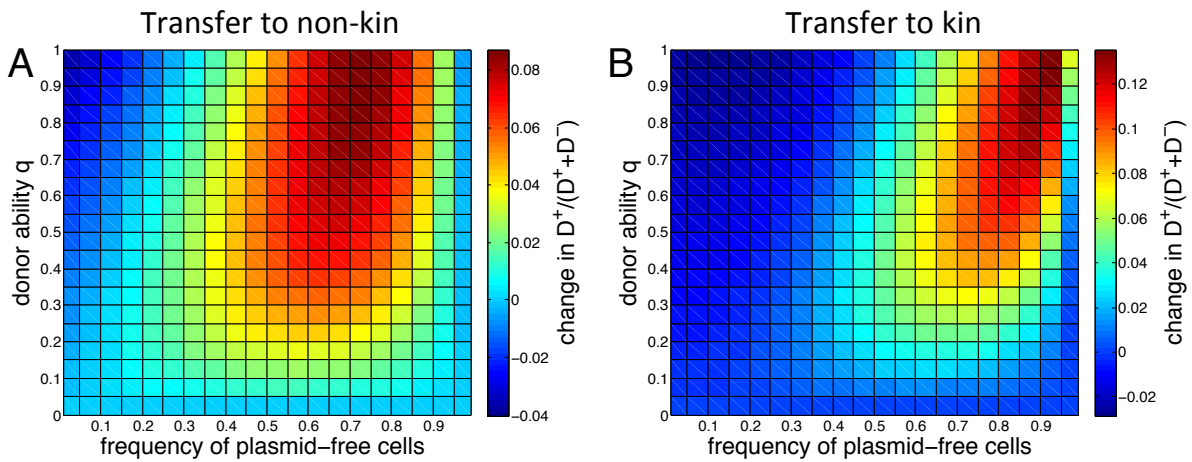


Figure 3.12: **Selection of donor ability for public good genes: effect of the abundance of plasmid-free cells.** Population structure is the same as in Figure 3.6, but now with varying proportion of plasmid-free cells and public good production  $\alpha = 1$ . The colorscale represents the change in frequency of  $D^+$  compared to  $D^-$ ,  $D^+/(D^+ + D^-)$  from  $t_1$  to  $t_2$  (when public good production leads to benefits) for strains carrying a plasmid with public good production of 1, as a function of  $D^+$  donor ability and the initial frequency of plasmid-free cells. In **A**,  $R$  are modelled as a distinct strain (transfer to non-kin). In **B**,  $D^-$  and  $D^+$  are able to receive plasmids (transfer to kin), and  $R$  proportion indicates the proportion of both  $D^+$  and  $D^-$  cells that are initially plasmid-free.

maximally selected for intermediate proportions of recipients, and higher transfer rates for a higher proportion of recipients. However, for a very high proportion of recipients ( $> 90\%$ ), donor ability is counter-selected again.

This result is qualitatively in agreement with Equation 3.5: selection of donor ability is higher with high effective transfer. However, maximal selection of donor ability is biased towards proportions of recipients higher than 50%. Here, this is not due to a bias in transfer itself: effective total transfer is maximal around 50% of recipients (Figure 3.13-A) (there is no amplification of transfer, as  $R$  do not act as secondary donors). However, what is also needed to select for donor ability is that populations enriched in  $D^+$  cells actually get significantly more feedback from transfer than other populations: the part of public good production that arises from transfer (and not from initial plasmid-bearers, present in equal proportions in both populations) has to be maximized. This part is higher when initial recipients are present in higher proportion (Figure 3.13-C), as initial plasmid-bearers make then a lower proportion of total producers. Feedback to donors (the relative benefit of transferring the public good) are thus maximized at recipient proportions that are higher than 50%.

Additionally, we analyze in simulations what happens when transfer takes place be-



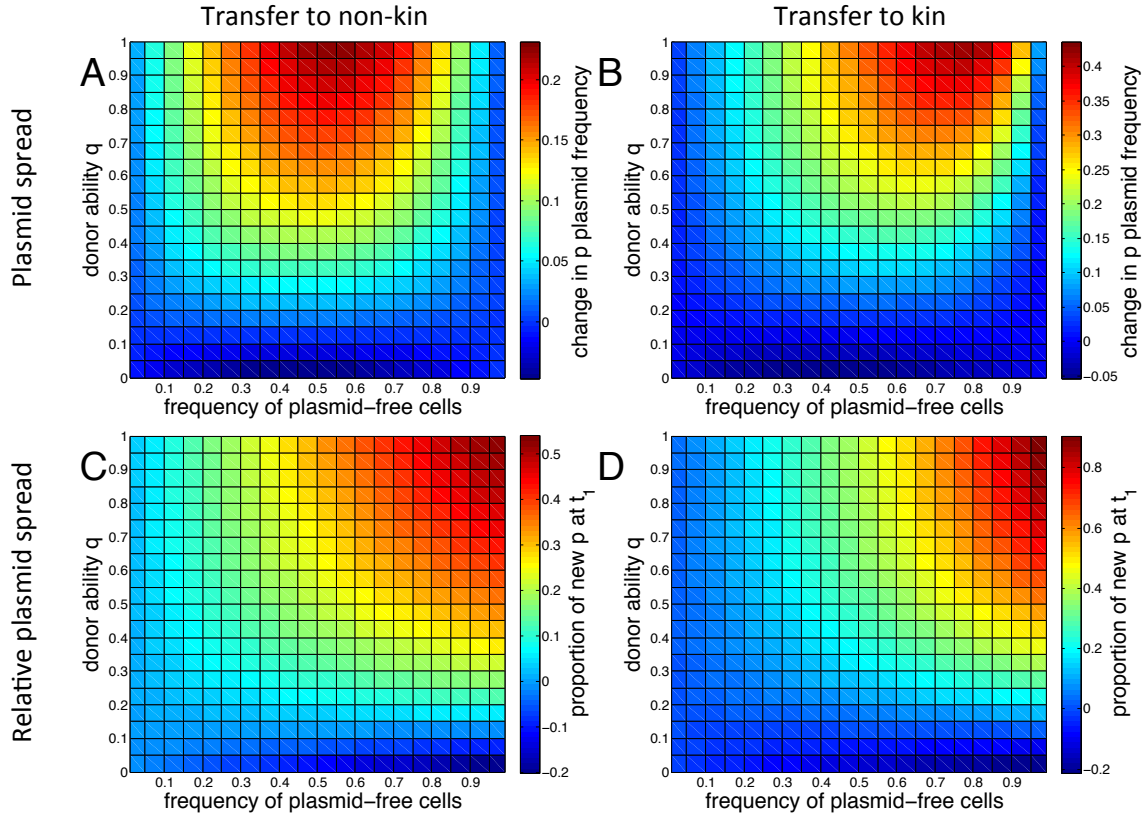


Figure 3.13: **Effect of recipient abundance on total transfer efficiency.** Results of the simulations shown in Figure 3.12, shown as a function of  $D^+$  donor ability  $q$  and the initial frequency of plasmid-free cells. In **A** and **B**, the colorscale represents the change in total plasmid abundance from  $t_0$  to  $t_1$   $p_{t1} - p_{t0}$  (plasmid spread, that approximates plasmid transfer). In **C** and **D**, the colorscale shows the proportion of plasmids present at  $t_1$  that were not present at  $t_0$ ,  $(p_{t1} - p_{t0})/p_{t0}$  (relative plasmid spread). In **A** and **C**,  $R$  is modelled as a distinct strain (transfer to non-kin). In **B** and **D**,  $D^-$  and  $D^+$  can receive plasmids (transfer to kin).

tween  $D$  strains, allowing transfer to kin (without a distinct recipient strain). Two strains,  $D^+$  and  $D^-$ , are competed, but both of them now have a varying initial proportion of plasmid-bearing cells.  $D_p^+$  can transfer the plasmid to both  $D^+$  and  $D^-$  plasmid-free cells. Adopting the same population structure as in Figure 3.12-A, we observe a slightly different pattern of selection (Figure 3.12-B *vs* 3.12-A): the amplitude of  $D^+$  frequency change is higher, and maximal selection happens for a higher initial proportion of plasmid-free cells. The difference is due to an increase in the actual transfer efficiency for populations enriched in  $D^+$  cells: plasmids can be transferred again from newly created  $D_p^+$  cells, yielding more final plasmid-bearers and public-good production (Figure 3.13-B *vs* 3.13-A), as we saw concerning the transfer of private good plasmids. Secondary transfer also increases the proportion of producers explained by transfer (Figure 3.13-D *vs* 3.13-C). Still, we observe the same general pattern, and even stronger selection for donor ability

when donors ( $D^+$  cells) transfer to kin, the opposite from the predictions of Equation 3.5. Indeed, amplification due to secondary transfer was not modelled, but it compensates here for the additional costs of public good production in cells related to the initial donor cells.

Selection of donor ability thus depends on population structure and transfer dynamics, with stronger selection when transfer is more effective but also when the feedback from public goods to donor cells is maximized.

### 3.4.3 Recipient ability for public good genes

Again, we first analyze what happens when plasmid donors are a strain distinct from  $R^+$  and  $R^-$  strains, similar to the experimental setup and previous simulations (Figure 3.14-A). Transfer is counter-selected for low proportions of recipients: recipient ability is costly (indirectly, because of public good production by cells that received the plasmid), and a low proportion of recipients does not strongly affect the total number of producers and the subsequent among-population selection. For higher proportions of recipients, recipient ability is selected, which is entirely due to among-population selection, as  $R^+$  frequency always decreases within populations (data not shown). The maximal increase in

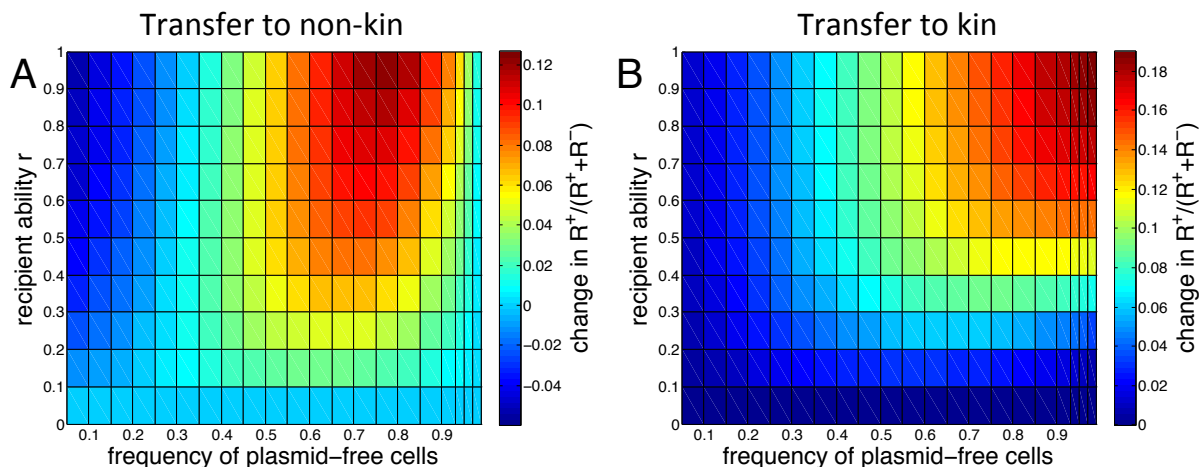


Figure 3.14: **Effect of recipient proportion on the selection of recipient ability.** Population structure is the same as in Figure 3.9, but now with varying initial proportion of plasmid-free cells ( $R^+ + R^-$ ), and public good production  $\alpha = 1$ . The colorscale shows the change in frequency of  $R^+$  compared to  $R^-$ ,  $R^+/(R^+ + R^-)$  from  $t_1$  to  $t_2$  (when public good production leads to benefits) in the presence of cells carrying a plasmid with public good production of 1, as a function of  $R^+$  recipient ability and the proportion of plasmid-free cells at  $t_0$ . In **A**,  $R$  are modelled as a distinct strain (transfer to non-kin). In **B**,  $R^-$  and  $R^+$  are able to donate plasmids (transfer to kin).

$R^+$  frequency happens for intermediate proportions of plasmid-free cells. When transfer

happens among  $R$  cells themselves (with no distinct  $D_p^+$  strain) (Figure 3.14-B), transfer is again counter-selected for low proportions of recipients. However, the maximal increase in  $R^+$  frequency happens here for almost 100% initially plasmid-free cells. The observed patterns can be explained in a similar way as in the case of the selection of donor ability for public good genes (section 3.4.2). In the case of transfer from a distinct  $D_p^+$  cell, effective transfer is maximized at intermediate plasmid-free cells frequencies (data not shown). However, the final proportion of producers that actually arise from transfer is higher for higher initial proportions of plasmid-free cells, which enhances the benefit in populations enriched in  $R^+$  cells. When transfer happens among  $R$  cells, secondary transfer occurs from  $R^+$  cells that receive the plasmid, amplifying the efficiency of transfer when plasmid-free cells are initially abundant, and promoting stronger selection of  $R^+$  cells.

In conclusion, selection for transfer abilities depends on population structure and transfer dynamics, with selection being stronger when (1) transfer is more effective and (2) feedbacks from public goods or (3) the proportion of cells that benefit privately from transfer are maximized. We next focus on the possibility of realistic population structures where both relatedness and transfer are sufficient to select for transfer abilities.

## 3.5 Population structure requirements

To mimic a more natural population structure, we model strong dilution of cells: with enough strong dilution, variations in the frequencies of genotypes at the metapopulation level arise simply from stochastic fluctuations in founder frequencies in populations (following Poisson distributions). We model a mix of all cell types present in previous simulations, now strongly diluted in multiple populations. We vary both the dilution factor applied to founding populations and transfer abilities.

We additionally assume here that donor ability is not costly, in order to avoid a strong decrease in donor cell frequency during the high number of generations needed to recover sufficient cell density from strong dilution. This assumption corresponds well with situations in natural systems, where transfer is often regulated, with no expression of transfer genes at low cell densities (where transfer is not efficient).

### 3.5.1 Transfer of private good genes

We follow the change in frequency of  $D^+$  strain from  $t_0$  to  $t_2$  in a metapopulation of 200 populations, initiated from a strongly diluted mix of equal proportions of  $D^+$  and  $D^-$

cells, each with initially 20%  $K$ -bearing cells. We vary the strength of dilution and  $D^+$  donor ability (Figure 3.15). Within populations,  $D^+$  frequency does not decline strongly

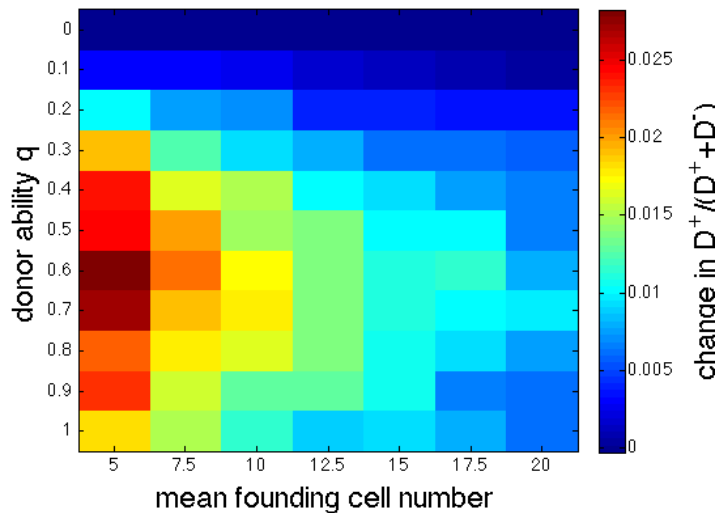


Figure 3.15: **Selection of donor ability for private good genes after strong initial dilution.** The simulated metapopulation consists of 200 populations initiated from a strongly diluted mix of equal proportions of  $D^+$  and  $D^-$  cells, each with initially 20%  $K$ -bearing cells, giving rise to a Poisson distribution with parameter  $\lambda$  for each  $D_p^+$  and  $D_p^-$  cells. The colorscale represents the change in  $D^+$  frequency from  $t_0$  to  $t_2$  (after antibiotic selection), shown as a function of  $D^+$  donor ability and mean founding cell number ( $5 \times \lambda$ , that depends on the strength of dilution), averaged over 50 simulations.

(as there is no cost of transfer in these simulations), but does not increase significantly either (data not shown), so all  $D^+$  selection is due to among-population selection.  $D^+$  strains are selected for with strong initial dilution and intermediate to high donor ability. This is due to the combination of two factors:

First, dilution up to low founding cell numbers provides sufficient variation in  $D^+$  frequencies among populations for among-population selection to act. The variation can be quantified by measuring the relatedness for  $D^+$  strain ( $r_D$ , measured as the regression coefficient between individual donor ability and donor ability of neighbouring cells, see Methods):  $r_D$  at  $t_0$  increases with increasing dilution (Figure 3.16-A). Sufficient dilution is thus needed to select for  $D^+$  cells.

Then, plasmid transfer should be sufficient to provide benefits to populations enriched in  $D^+$  cells. Plasmid invasion from  $t_0$  to  $t_1$  (Figure 3.16-B) depends mainly on donor ability  $q$ . However, it also depends on the strength of dilution: it declines with decreasing initial mean number of cells. The decrease in plasmid invasion is the effect of strong spatial structuring in populations caused by dilution: only a few cells are present initially

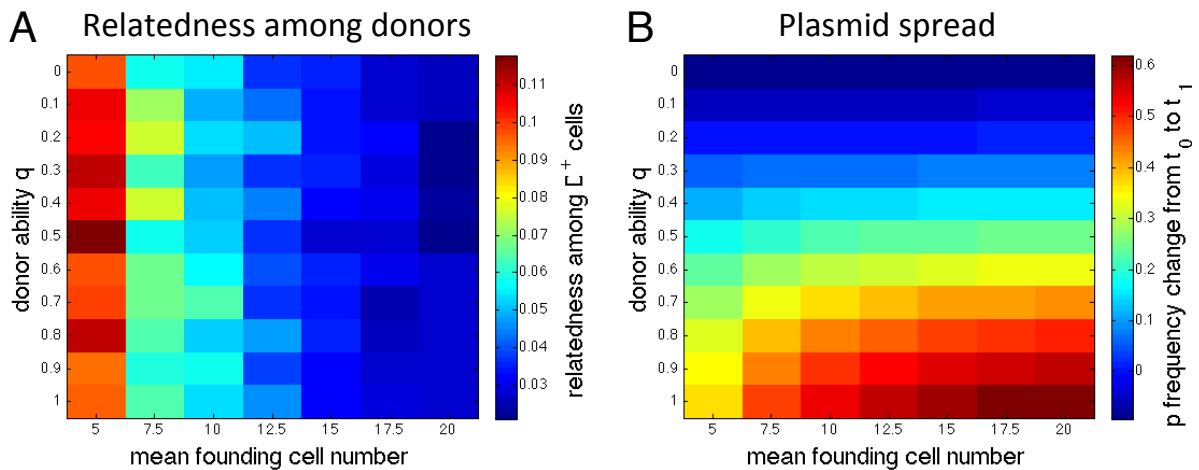


Figure 3.16: **Change in population characteristics after strong initial dilution.** Results reported here are from the same simulations as results shown in Figure 3.15. The colorscales represent relatedness among  $D^+$  cells  $r_D$  at  $t_0$  (**A**) and the total change in proportion of plasmid-bearing cells from  $t_0$  to  $t_1$  (plasmid spread, **B**), shown as functions of  $D^+$  donor ability  $q$  and mean founding cell number.

in each population, and often not cells of all cell types, which decreases the number of possible encounters between plasmid-bearing and plasmid-free cells in an increasing percentage of the populations. Still, even with the segregation of cell types arising from dilution, transfer is quite abundant, with up to a 35% increase in plasmid frequency at the metapopulation scale.

Both the transfer and the relatedness among donors are thus high enough to promote the selection of donor ability with the combination of low founding cell number (high dilution) and high donor ability.

### 3.5.2 Transfer of public good genes

We now focus on the transfer of public good genes and follow the change in frequency of  $D^+$  strain from  $t_1$  to  $t_2$  in a metapopulation of 200 populations initiated from a strongly diluted mix of equal proportions of  $D^+$  and  $D^-$  cells, each with initially 20%  $P^+$ -bearing cells. We vary the strength of dilution and  $D^+$  donor ability (Figure 3.17).

Again,  $D^+$  cells are selected when both the dilution and  $D^+$  donor ability are high (Figure 3.17). When dilution is weak (high  $\lambda$ ), it does not provide sufficient variation in  $D^+$  frequencies among populations (relatedness among donor cells is low: the presence of a given  $D^+$  cell is not strongly associated to the presence of other  $D^+$  cells in the neighbourhood, that would benefit from increased public good production). When dilution is strong enough, relatedness among donors becomes higher, and variations in frequencies

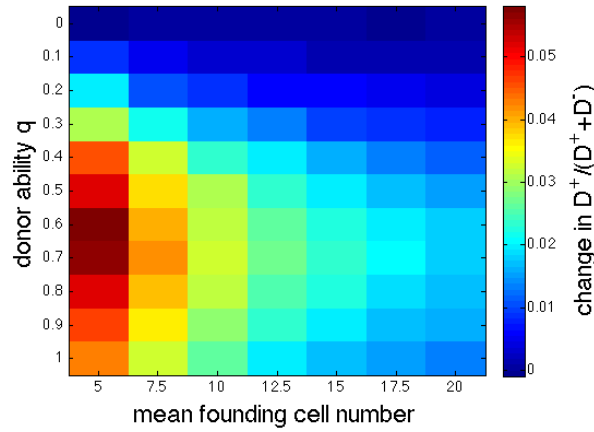


Figure 3.17: **Selection for donor ability for public good genes after strong initial dilution.** The simulated metapopulation consists of 200 populations initiated from a strongly diluted mix of equal proportions of  $D^+$  and  $D^-$  cells, each with initially 20%  $P^+$  plasmid-bearing cells, giving rise to Poisson distribution of parameter  $\lambda$  for each for each  $D_p^+$  and  $D_p^-$  cells.  $P^+$  plasmid has public good production  $\alpha = 1$ . The colorscale shows the change in  $D^+$  frequency from  $t_1$  to  $t_2$  (when public goods lead to benefits), as a function of  $D^+$  donor ability and the mean founding cell number ( $5 \times \lambda$ , that depends on the strength of dilution), and averaged over 50 simulations.

are strong enough so that selection can take place among populations. Plasmid spread, however, also depends on population structure: it declines with higher dilution. Similarly to the case of the transfer of private good genes, transfer is still quite abundant, enough for donor ability to be selected through feedbacks of public good production: donor ability can be selected for without choosing an arbitrary repartition of plasmids and strains within populations.

### 3.5.3 Recipient ability for public good genes

In the final set of simulations, we model a mix of  $R^+$ ,  $R^-$ ,  $R^+p$  and  $R^-p$  cells and follow the change in frequency of  $R^+$  strain from  $t_1$  to  $t_2$  in a metapopulation of 200 populations initiated from a strongly diluted mix of equal proportions of  $R^+$  and  $R^-$ , each with initially 20%  $P^+$ -bearing cells.  $P^+$  plasmid has public good production  $\alpha = 1$ . We vary the strength of dilution and  $R^+$  recipient ability (Figure 3.18).

Similarly to what we saw for the selection of donor ability,  $R^+$  cells are selected when both the dilution and  $R^+$  recipient ability are high (Figure 3.18). The reasons for such pattern are also similar: when dilution is too weak, variation in  $R^+$  frequencies among populations is low. Moreover, plasmid spread depends on both recipient ability and population structure, and declines with higher dilution. Still, even with the segregation between plasmid-bearing and plasmid-free cells arising from dilution, transfer is enough

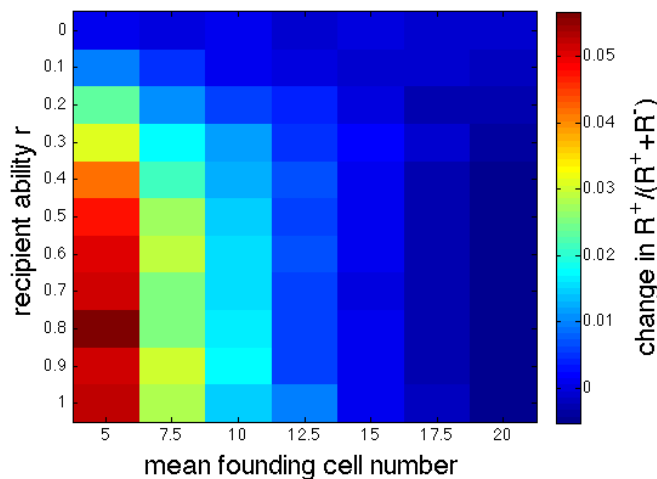


Figure 3.18: **Selection for recipient ability after strong initial dilution.** The simulated metapopulation consists of 200 populations initiated from a strongly diluted mix of equal proportions of  $R^+$  and  $R^-$  cells, each with initially 20%  $P^+$  plasmid-bearing cells, giving rise to Poisson distribution for each type of cell, of parameter  $\lambda$  for plasmid-bearing cells  $D_p^+$ .  $P^+$  plasmid has public good production  $\alpha = 1$ . The colorscale shows the change in  $R^+$  frequency from  $t_1$  to  $t_2$  (when public goods lead to benefits) as a function of  $R^+$  recipient ability and the mean founding cell number ( $5 \times \lambda$ , that depends on the strength of dilution), and averaged over 50 simulations.

for recipient ability to be selected through feedbacks of public good production. As was the case for donor ability, recipient ability can thus be selected for without constraining the repartition of plasmids and strains within populations.

## 3.6 Discussion

Using experiments as well as simulations that follow them, coupled to a general, direct fitness approach, here we show that transfer ability can be increased by selection on non-mobile host genes. The results from the different approaches are generally in agreement, and show that selection of transfer ability is generally favored - whatever the type of good encoded on the plasmid - at two conditions: (1) there must be variability in the composition of populations for the considered trait, that allows significant relatedness among cells, and (2) transfer must be efficient.

### 3.6.1 Selection on transfer ability concerning private traits

Direct selection can act on recipient ability if the incoming mobile elements are beneficial to host fitness. However, there is no similar direct selection on donor ability, as donor ability does not directly modify the phenotype of the donor cell relatively to the transferred genes: why bacteria would transfer beneficial genes to their apparent competitors is thus an intriguing question. Previous analyses have focused on explaining the presence of

host beneficial genes on plasmids through benefits to plasmids themselves, and shown that carrying intermittently selected genes can benefit plasmids [Bergstrom et al., 2000]. We demonstrate here that plasmid mobility can also benefit the donor host, and genes promoting plasmid mobility can furthermore be selected through this indirect benefit. This finding can be related to general considerations where mobile elements are described as a "communal pool" of genes [Norman et al., 2009], or "agents of open source evolution" [Frost et al., 2005], leading to benefits at the population level [Heuer et al., 2008]. However, population-wide benefits would not be sufficient to select for costly chromosomal alleles promoting transfer: transfer can be considered as an altruistic trait as it is costly to the host [Rankin et al., 2010]. We show here that donor ability for private traits is actually selected for when it does not happen towards any competitor cell, but is biased towards related cells (that share the allele for high donor ability): donor ability can be selected through kin selection. As is often the case in the context of kin selection [Hamilton, 1964a, West et al., 2007b], selection requires a population structure that ensures that benefits of the behaviour feedback preferentially on cells sharing the same allele. We derive a Hamilton's rule-like condition for the selection of donor ability: benefits of the transferred gene, weighted by the efficiency of transfer among related cells and the relatedness between donor cells in the population, have to outweigh the cost of transfer (Equation 3.6). In other words, chromosomal, non-mobile alleles benefit from sharing other beneficial genes with kin, as this will increase their inclusive fitness.

These results can be compared to conclusions about the evolution of teaching in animals (where teaching can be defined as a transfer of information that is costly for the sender and beneficial for the receiver [Riboli-Sasco et al., 2008]). Indeed, it has been shown that teaching of adaptive information can be selected when teachers and pupils are related [Fogarty et al., 2011]. Interestingly, the same study also concluded that teaching was favored only if the information could not be acquired easily without teaching. In the case of genetic information transfer, a key difference from teaching is that genes can be transmitted both vertically (from parent to offspring) and horizontally (from donor to receiver). Here we demonstrate the selection of transfer in a case where initially only a few cells bear the gene of interest and horizontal transfer allows more efficient spread to related cells than vertical transfer.

A strong assumption in the scenario we implement is that transfer has to happen before selection on transferred genes. It seems unrealistic that hosts randomly transfer genes that may but may not be useful in the future. However, transfer is in reality regulated, and



could occur in conditions that are correlated with a specific selective pressure. A striking case that could fit with this idea is the one of tetracycline-resistance mobile elements, whose transfer is induced by sub-inhibitory concentrations of tetracycline [Rashtchian et al., 1982]: transfer happens in conditions where mobile elements are likely to increase the fitness of hosts in the near future (the diffusion of antibiotics can create antibiotic gradients, with antibiotics first present only in sub-inhibitory concentrations).

### 3.6.2 Selection on transfer ability concerning public good genes

Transfer ability concerning public good genes is selected through a quite different mechanism: it does not require that transfer happens towards related cells. However, relatedness must still be generally high in the population: cells sharing alleles that favor transfer have to be associated more than randomly, in order for benefits of additional public goods produced following transfer (to related or unrelated cells) to feedback preferentially on them. This is comparable to the conditions required for interspecies mutualism: cooperation towards another species necessitates within-species relatedness [Frank, 1994, Foster and Wenseleers, 2006], so that the response from individuals of the other species (equivalent to public good production here) benefits preferentially cells related to the initial cooperating cells. Once again, the second-order selection for transfer can be expressed in a form analogous to Hamilton’s rule: benefits of public goods produced by neighbours, weighted by the efficiency of transfer and relatedness, have to outweigh the cost of transfer. Here selection acts in a similar way on recipient and donor abilities, as they both lead to an increase in the concentration of public goods in the environment.

We here analyzed separately selective pressures acting on donor and recipient abilities, but they will generally interact. If transfer happens within a strain (no separation between plasmid donors and recipients), it simply increases the proportion of cells in the population that act as cooperators. If recipient ability is very low, high donor ability will bring less benefits as effective transfer will be low, and conversely: high recipient ability will be less beneficial if donor ability is very low. If transfer happens between distinct strain or species, transfer could actually be a form of manipulation of the recipient strain by the donor strain. We have not modelled multi-species populations explicitly here: we did not assume any differences between donors and recipients other than plasmid presence and transfer, like differences in plasmid cost or gene expression which would be likely in the case of manipulation. In natural systems, it is possible that only one of the interacting strains or species benefits from transfer. The most striking example of this is T-DNA

transfer from the Ti plasmid of *Agrobacterium tumefaciens* to host plant cells. T-DNA integrated in the plant genome induces tumor formation and the production of specific metabolites, opines, that are used by *Agrobacterium* cells only [Platt et al., 2012]. Here, there is selection of T-DNA donor ability, but obviously not of recipient ability by the parasitized plant. However, in other cases both donor and recipient strains or species could possibly benefit from transfer. This possibility will probably be highly dependent on the details of the exact interaction mechanisms and the coevolution of donor and recipient strains.

### 3.6.3 Plasmid abundance requirements

A recurrent conclusion of this study is that plasmid-free cells need to be abundant for transfer ability to be selected. In the long-term, this implies that plasmids must be repeatedly lost from some proportion of the population and not spread to fixation. This is more generally required for the selection of plasmid transfer itself: horizontal transfer benefits plasmids mainly when many recipient, plasmid-free cells are available. Still, repeated gene loss seems more paradoxical here, as we focus on genes that bring a benefit to the host: why do these genes not stay associated with the cells they benefit? However, the genes such as antibiotic resistance we considered here are either only transiently beneficial or beneficial only at a group scale with individual costs: they could be repeatedly lost when they are not selected for. In the case of antibiotic resistance genes, loss could happen for instance when bacteria migrate from antibiotic-treated hosts or environments to non-treated ones. In the case of public-goods, non-producing bacteria that do not pay the cost of production could prevent the fixation of plasmids carrying public-good genes.

Another factor that will influence the selection of transfer is its cost to the cell. Recipient ability is probably not costly, as defense mechanisms are typically active so cells can receive plasmids by default. On the other hand, donor ability is typically costly, as it involves accrued production of pili and sensitivity to male-specific phages [Haft et al., 2009, Jalasvuori et al., 2011]. To overcome potential costs, transfer is generally strongly regulated as a function of environmental conditions [Frost and Koraimann, 2010], and could be induced specifically when it benefits indirectly the cell.

The variability of transfer rates observed across isolates [Dionisio et al., 2002] could then be explained by the relative importance of all these factors in different environmental conditions. A well-known example is the one of antibiotic resistance plasmids (which can lead to either private or public traits), for which selective pressures are likely variable:

they will be strongly beneficial to the host in the presence of antibiotics, but costly in the absence of antibiotics. Transfer itself will be more costly when the density of male-specific phages is high.

### 3.6.4 Population abundance and environmental requirements

The selection of transfer further requires a quite specific population structure, with sufficient positive assortment among cells but no strong assortment among plasmids. Using simulations, we analyze the possibility of occurrence of such populations. We show that with strong dilution, the strength of selection for transfer is indeed reduced by the opposing effects of dilution on relatedness and on efficient transfer, but selection of transfer is still possible in some range of parameters. Moreover, we made the simplifying assumption that plasmids are initially present in equal proportions in every strain. This will often not be true, as plasmid association to host genotypes will be influenced by plasmid-bacteria coevolution history, which will in turn modify population structure. For instance, when transferring a plasmid leads to increased fitness to both the donor cell and the transferred plasmid, the association between plasmids and cells with high donor ability could be favored on longer timescales, modifying population structure and increasing the possibility of plasmid transfer.

Relatedness among cells sharing high transfer ability has to be sufficient for transfer to be selected. We also showed in Chapter 2 that plasmid transfer actually increases relatedness at the transferred loci. This increase is due to the local properties of transfer, and can be extremely high in epidemic dynamics, where recipients become secondary donors and the whole population of cells is invaded. Some similar dynamics is observed here with chromosomal loci involved in transfer: even if they are not transferred themselves (so there is no strict invasive dynamics), local amplification of transfer is favored when populations are already structured, as cells with high transfer ability are clustered together (provided that some of these cells initially do not bear plasmids). Finally, donor ability could itself promote an increase in the spatial structuring of populations through physical effects. Indeed, conjugation has been shown to promote biofilm formation, partly because of the aggregation properties of conjugative pili [Ghigo, 2001]. Biofilms are very structured environments, so their formation could favor physically the selection of transfer by ensuring high relatedness.

### 3.6.5 Conclusion

In conclusion, the selective pressures we describe here could contribute to the very large variability of transfer rates observed among bacterial isolates [Gordon, 1992, Dionisio et al., 2002]. They will favor the transfer of genes that benefit the host, either directly or through the production of public goods. The same dynamics could in addition contribute to the explanation of the pattern of genes found on mobile elements [Rankin et al., 2010], as both plasmids and chromosomes could benefit from the location of those genes on plasmids.

In our analysis, we did not consider selective pressures acting on plasmid genes themselves: plasmids gain direct fitness benefits from horizontal transfer, which are also influenced by multiple parameters. However, we show that chromosomal genes do not only suffer the cost of parasitic plasmid transfer, but can also benefit from it. The coevolution of plasmid and chromosomal genes controlling transfer remains an open question, as their interests may or may not converge, depending on ecological and evolutionary scenarios.

## 3.7 Materials and methods

### 3.7.1 Experiments

#### Plasmids and strains

The background strain is JC1191, an *E. coli* strain that can grow in low concentrations of chloramphenicol (Cm) in the presence of C<sub>4</sub>-HSL, thanks to the addition of *Pseudomonas* quorum sensing machinery [Chuang et al., 2009]: JC1191 contains the att::rhl-catLVA(Sp<sup>R</sup>) segment with both an Rhl auto-inducer-responsive promoter (P<sub>rhl\*</sub>) driving an unstable version of *cat* responsible for Cm resistance, and the *rhlR* gene under the weak constitutive promoter P<sup>q</sup><sub>lacI</sub>, integrated into the chromosome with the spectinomycin resistance gene Sp<sup>R</sup>.

Plasmids used in experiments are shown in Figure 3.19. They all carry the *aph* gene providing kanamycin resistance (kan<sup>R</sup>). *K*-YFP and *K*-GFP plasmids both bear F-*oriT*, and express respectively *YFP* and *GFP* under control of the strong promoter *P<sub>R</sub>*. *P*<sup>-</sup>*T*<sup>-</sup> plasmid carries *GFP* under control of *P<sub>R</sub>*, *P*<sup>+</sup>*T*<sup>-</sup> plasmid additionally carries the *rhlI* gene under control of the same promoter: RhlI enzyme produces the Rhl auto-inducer C<sub>4</sub>-HSL [Chuang et al., 2009]. *P*<sup>-</sup>*T*<sup>+</sup> plasmid carries *YFP* under control of *P<sub>R</sub>* and F-*oriT*, *P*<sup>+</sup>*T*<sup>+</sup> plasmid additionally carries the *rhlI* gene under control of the same promoter *P<sub>R</sub>* (*P*<sup>-</sup>*T*<sup>+</sup>

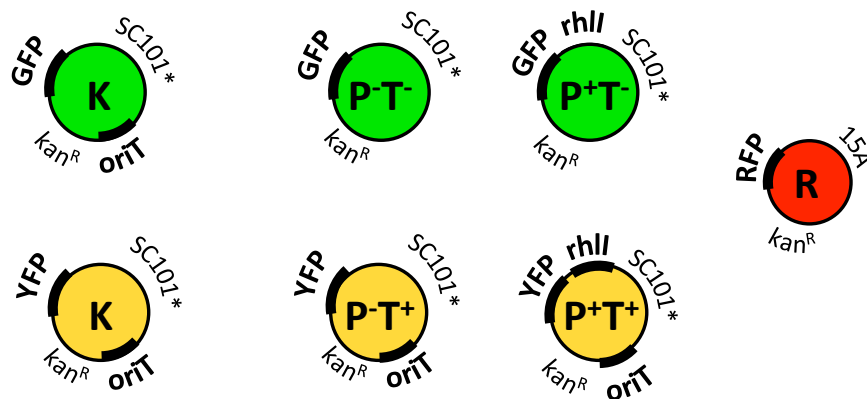


Figure 3.19: **Plasmids used in experiments.** p15A-origin and pSC101-origin bearing plasmids all bear kanamycin resistance. Plasmids are colored according to the fluorescence genes they bear. *K* plasmids bear F *oriT* in addition to kanamycin resistance (used as the private good gene). pSC101-origin bearing plasmids are involved in transfer and private and public good production phenotypes.  $P^-$  and  $P^+$  indicate public good production status:  $P^+$  plasmids bear *rhII* synthase,  $P^-$  plasmids do not.  $T^-$  and  $T^+$  indicate transfer status:  $T^+$  are transferable as they bear F *oriT*,  $T^-$  plasmids are not. Recipients bear a plasmid with p15A replication origin expressing RFP, compatible with pSC101-origin bearing plasmids.

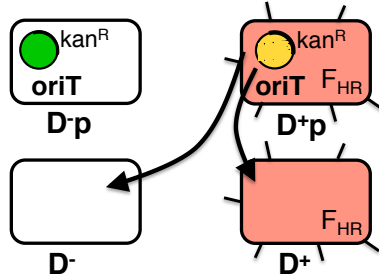
and *K*-YFP are actually the same plasmid).

For experiments concerning donor ability for public good genes (section 3.3.2), recipients *R* bear pSB3K3-RFP plasmid, a middle-copy plasmid that carries *mRFP1* under control of the strong promoter  $P_{lacI}$ . For other experiments (sections 3.3.1 and 3.3.3), one of the competitor strains (respectively  $D^+$  and  $R^+$ ) is marked with the *td-Cherry* gene under control of a  $pRNA1$  promoter, inserted in the *attTn7* site of the chromosome.

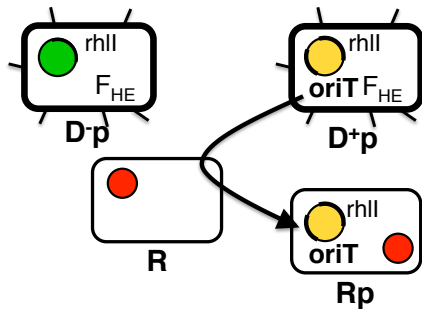
To modulate transfer ability, cells bear either  $F_{HR}$  or  $F_{HE}$  helper plasmid, or no F plasmid. Both  $F_{HR}$  and  $F_{HE}$  are mutants of the F plasmid pOX38::Tc [Anthony et al., 1994] with two point mutations in *oriT* sequence that lead to reduced self-mobilization efficiency by the F relaxase (1000-fold reduction compared to F).  $F_{HE}$  and  $F_{HR}$  mobilize other plasmids bearing *oriT* efficiently.  $F_{HE}$  has the wild-type *traS* gene encoding F plasmid entry exclusion (thus  $F_{HE}$ , for helper with entry exclusion), leading to very low recipient ability of cells bearing it.  $F_{HR}$  additionally bears a deletion of the *traS* gene: recipient cells bearing  $F_{HR}$  are able to receive plasmids efficiently (100-fold more efficiently than cells bearing  $F_{HE}$ , see Figure 1.10) and behave as secondary donors of *oriT*-bearing plasmids (thus  $F_{HR}$ , for helper receiver plasmid).

Transfer and fluorescent markers for identification by flow cytometry are summarized in Figure 3.20 for each competition experiment setup.

## A: donor ability for private good genes



## B: donor ability for public good genes



## C: recipient ability for public good genes

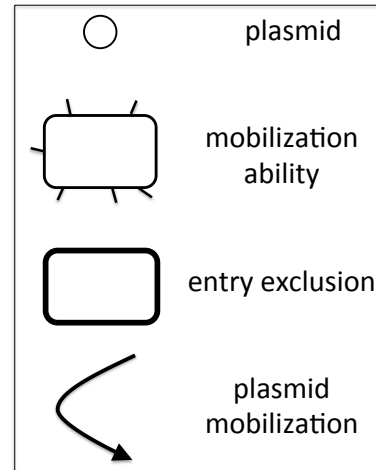
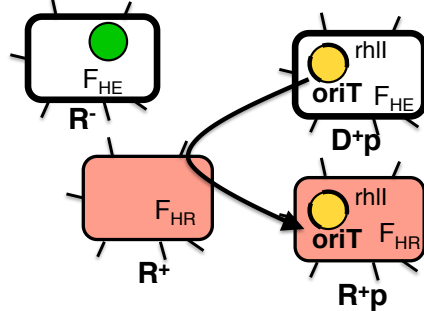


Figure 3.20: **Summary of the strains used in competition experiments: control of transfer and fluorescence markers.** Colors indicate fluorescence markers associated with each strain (cell background color) or plasmid (plasmid background color). Pili indicate the ability to mobilize plasmids bearing *oriT*; a thick black line indicates functional entry exclusion inhibiting recipient ability. Arrows indicate plasmid mobilization. Only relevant plasmid genes are indicated (see Figure 3.19 for plasmid maps).

## Growth and experiment conditions.

Cells were grown in Luria-Bertani (BD Difco) medium with 25  $\mu\text{g}/\text{mL}$  spectinomycin (Sp, Sigma-Aldrich) and with or without 50  $\mu\text{g}/\text{mL}$  kanamycin (Kn, Sigma-Aldrich) and 6.25  $\mu\text{g}/\text{mL}$  chloramphenicol (Cm; Sigma-Aldrich). Experiments were conducted under well-mixed conditions with 5mL medium in 50 mL tubes (Sarstedt).

For competition experiments, strains were mixed at various ratios (vol/vol) and first grown from a 10-fold dilution from stationary phase cultures ( $t_0$ ) at 35°C, up to an optical density of 3. This temperature allows transfer to happen, as F transfer is strongly reduced at 30°C. This step was repeated once with 5-fold dilution in order to increase plasmid transfer. Cultures were then diluted 10-fold until stationary phase at 30°C ( $t_1$ ), which allows pre-induction of Cm-resistance by C<sub>4</sub>-HSL in non-producer as well as producer cells and enhances the fluorescence signal. Finally, cultures were diluted 100-fold at  $t_1$  and grown for 12 to 16 hours at 30°C until  $t_2$ .

For competition experiments not involving public good production (section 3.3.1), the

growth medium contained Sp from  $t_0$  to  $t_1$ , and Sp + Kn from  $t_1$  to  $t_2$ . Cultures with various ratios were mixed at  $t_1$ . For competition experiments between donors involving public good production (section 3.3.2), the growth medium contained Sp + Kn from  $t_0$  to  $t_1$ , and Sp + Kn + Cm from  $t_1$  to  $t_2$ . For competition experiments between recipients involving public good production (section 3.3.3), the growth medium contained Sp from  $t_0$  to  $t_1$ , and Sp + Cm from  $t_1$  to  $t_2$ .

Cultures were analyzed for strain and plasmid proportions by flow cytometry. Optical density at 600nm was measured on a BioPhotometer plus (Eppendorf).

### 3.7.2 Data analysis

#### Estimations of selection

Selection acting on a given trait is globally estimated by measuring changes in strain frequencies over time. Frequencies at the metapopulation level are measured by pooling equal volumes of populations, effectively taking into account differential growth among populations. Frequencies within populations were computed as the mean of each individual frequency within populations, in order to exclude the effect of differential growth among populations.

#### Computing relatedness

Relatedness is defined as the measure of how an individual's social environment covaries with the individual's genotype, which can be calculated as the linear regression coefficient connecting an individual's genotype with the genotype of its interactants [Damore and Gore, 2012]. Here, we consider that the social environment of an individual corresponds to the population it belongs to.

With  $x_i$  and  $n_i$  being respectively the proportion of  $x$  allele and number of bacteria within population  $i$ , and  $x_{tot}$  and  $n_{tot}$  being respectively the proportion of  $x$  allele and number of bacteria in the whole metapopulation, the relatedness of  $x$  allele can be calculated as followed (assuming populations are of the same size, which is true at  $t_1$ ):

$$r_x = \frac{\sum_i \frac{x_i}{n_i} \cdot \frac{x_i}{x_{tot}} - \frac{x_{tot}}{n_{tot}}}{1 - \frac{x_{tot}}{n_{tot}}} \quad (3.8)$$

#### Statistical analysis

The significance of differences between conditions was tested with two-sample, two-sided t-tests. The normality of distributions was confirmed with Shapiro-Wilkinson tests, re-

jecting the null hypothesis when  $p < 0.05$ .

### 3.7.3 Simulations

#### General model assumptions

Our models were constructed to mimic experimental conditions of strain growth and plasmid transfer. We use notation similar to the one used in the direct fitness model (section 3.2).

Growth follows a logistic function saturating at carrying capacity  $K$ , mimicking growth to stationary phase. Plasmid transfer happens at a basal rate  $\gamma$ , following a mass-action law [Levin et al., 1979], and saturates in the same way as growth rate at carrying capacity  $K$ , as F transfer has been shown to strongly decrease when cells approach stationary phase [Frost and Manchak, 1998]. Strains are characterized by donor ability  $q$  or recipient ability  $r$  that modulate effective transfer  $\gamma_e$  ( $\gamma_e = q \times \gamma$  when varying donor ability,  $\gamma_e = r \times \gamma$  when varying recipient ability). We assume a cost of transfer for the donor cell  $c_q$  so that the cost on growth is proportional to donor ability, but no cost of receiving plasmids ( $c_r=0$ ). All costs act on growth rate.

The rate of transfer was divided by 10 from  $t_1$  to  $t_2$ , to mimic experiments where growth happens at 30°C and transfer is reduced. We did not attempt to measure  $\gamma_e$  experimentally, as the transfer rate is not constant during the duration of the experiments (because of successive dilutions and shifts in temperature). However, we can estimate an effective transfer rate leading to the same plasmid invasion than observed at  $t_1$  in experiments. Effective transfer rates are around from  $5 \cdot 10^{-8}$  mL.cell<sup>-1</sup>L<sup>-1</sup> in our experiments.

We explicitly follow the experimental set-up by modeling the two steps: first pre-incubation, where pre-induction of Cm resistance (in the cases of public goods) and transfer happen (from  $t_0$  to  $t_1$ ), then growth in conditions where the plasmid genes affect growth (from  $t_1$  to  $t_2$ ): in the presence of Cm (for public good genes) or of Kn (for private good genes). The pre-incubation time (from  $t_0$  to  $t_1$ ) was set to 12h after 100-fold initial dilution from carrying capacity, and growth from  $t_1$  was allowed for 60h after a second 100-fold dilution.

To study the effect of strong cell dilution (section 3.5), cells were distributed in 200 populations each of 10  $\mu$ L, following a Poisson distribution of varying parameter  $\lambda$  ( $\lambda \in [1 : 4]$  for  $D_p^+$  cells). Because of the strong initial dilution, the pre-incubation time was set at 24h and a second 10-fold dilution step was added before  $t_1$ . Results were averaged over 50 replicate simulations, as strong variance arises from Poisson distribution. The cost of



transfer  $c_q$  was set to 0 to prevent excessive fitness loss for donor cells.

### Plasmid effects on growth

We note the basal growth rate  $\psi_p$  for plasmid-bearing cells and  $\psi_\emptyset$  for plasmid-free cells. In the case of private goods, there is no growth for plasmid-free cells ( $\psi_\emptyset = 0$ ) after  $t_1$ . For cells bearing plasmids, the growth rate is the same as the growth rate before  $t_1$ .  $t_2$  is the end of growth. We do not vary the production rate for private goods.

The public effect of public goods on growth is modelled as proportional to the proportion of public-good producing cells at a given time in the local population ( $\frac{ptot}{ntot}$ ), and to the good production rate  $\alpha$  ( $\psi = \psi_0(1 + b_p\alpha \times \frac{ptot}{ntot})$ ), and is the same for plasmid-bearing and plasmid-free cells. The public good production rate in experiments is fixed arbitrarily at  $\alpha = 1$ , as costs and benefits parameters were fitted experimentally for this value.

Parameters values shown were estimated from our experimental data.

Parameter	Symbol	Estimation
Carrying capacity	$K$	$4.10^9$ cells $\text{mL}^{-1}$
Basal rate of plasmid transfer	$\gamma$	$10^{-9}$ $\text{mL cell}^{-1}\text{L}^{-1}$
Basal growth rate from $t_0$ to $t_1$	$\psi_1$	$0.96 \text{ h}^{-1}$
Basal growth rate from $t_1$ to $t_2$	$\psi_2$	$0.12 \text{ h}^{-1}$
Cost of good production on growth rate	$c_p$	0.04
Benefit of public good production on growth rate	$b_p$	4
Cost of donor ability on growth rate	$c_q$	0.05

Plasmids effects on growth are summarized below:

Notation	$t_0$ to $t_1$	$t_1$ to $t_2$ (private goods)	$t_1$ to $t_2$ (public goods)
$\psi_\emptyset$	$\psi_1$	0	$\psi_2(1 + b_p\alpha \times \frac{ptot}{ntot})$
$\psi_p$	$\psi_1(1 - c_p\alpha)$	$\psi_1(1 - c_p\alpha)$	$\psi_2(1 + b_p\alpha \times \frac{ptot}{ntot})$

When the public good acts on growth rate, the benefit of cooperation is only transient [Chuang et al., 2009] and  $t_2$  has to be chosen before all populations reach stationary phase. For each simulation,  $t_2$  was defined as the time point where the selection of  $P^+$  was maximal over all conditions tested.

**Selection of donor ability**

In simulations described in sections 3.3.1 and 3.3.2, we follow the frequency of  $D^+$  strain compared to  $D^-$  strain. In the first scenario,  $D^+p$  cells transfer the plasmid  $p$  only to a distinct  $R$  strain, that is not capable of secondary transfer:

$$\begin{aligned}
\frac{dD_p^+}{dt} &= \psi_p(1 - c_q q) \times D_p^+ \times (1 - \frac{n_{tot}}{K}) \\
\frac{dD_p^-}{dt} &= \psi_p \times D_p^- \times (1 - \frac{n_{tot}}{K}) \\
\frac{dR}{dt} &= [\psi_\emptyset - \gamma_e D_p^+] \times R \times (1 - \frac{n_{tot}}{K}) \\
\frac{dR_p}{dt} &= [\psi_p \times R_p + \gamma_e D_p^+ \times R] \times (1 - \frac{n_{tot}}{K}) \\
n_{tot} &= D_p^+ + D_p^- + R + R_p \\
p_{tot} &= D_p^+ + R_p
\end{aligned}$$

In the second scenario,  $D^+p$  cells transfers the plasmid to both  $D^+$  et  $D^-$  plasmid-free cells:

$$\begin{aligned}
\frac{dD_p^+}{dt} &= [\psi_p(1 - c_q q) + \gamma_e D^+] \times D_p^+ \times (1 - \frac{n_{tot}}{K}) \\
\frac{dD_p^-}{dt} &= [\psi_p \times D_p^- + \gamma_e D_p^+ \times D^-] \times (1 - \frac{n_{tot}}{K}) \\
\frac{dD^+}{dt} &= [\psi_\emptyset(1 - c_q q) - \gamma_e D_p^+] \times D^+ \times (1 - \frac{n_{tot}}{K}) \\
\frac{dD^-}{dt} &= [\psi_\emptyset - \gamma_e D_p^+] \times D^- \times (1 - \frac{n_{tot}}{K}) \\
n_{tot} &= D_p^+ + D_p^- + D^+ + D^- \\
p_{tot} &= D_p^+ + D_p^-
\end{aligned}$$

**Selection of recipient ability**

In simulations described in section 3.3.3, we follow the frequency of  $R^+$  strain compared to  $R^-$  strain. We assume that  $c_r$  is zero, and  $c_q$  also, as we are not concerned with the selection of donor ability here. In the first scenario, a distinct  $D_p^+$  strain transfers the

plasmid to  $R^+$  cells, that are not capable of secondary transfer:

$$\begin{aligned}
 \frac{dD_p^+}{dt} &= \psi_p \times D_p^+ \times \left(1 - \frac{n_{tot}}{K}\right) \\
 \frac{dR^-}{dt} &= \psi_\emptyset \times R^- \times \left(1 - \frac{n_{tot}}{K}\right) \\
 \frac{dR^+}{dt} &= [\psi_\emptyset - \gamma_e D_p^+] \times R^+ \times \left(1 - \frac{n_{tot}}{K}\right) \\
 \frac{dR_p^+}{dt} &= [\psi_p \times R_p^+ + \gamma_e D_p^+ \times R^+] \times \left(1 - \frac{n_{tot}}{K}\right) \\
 n_{tot} &= D_p^+ + R^- + R^+ + R_p^+ \\
 p_{tot} &= D_p^+ + R_p^+
 \end{aligned}$$

In the second scenario, both  $R_p^+$  and  $R_p^-$  cells transfer the plasmid to  $R^+$  cells:

$$\begin{aligned}
 \frac{dR_p^+}{dt} &= [\psi_p \times R_p^+ + \gamma_e (R_p^+ + R_p^-) \times R^+] \times \left(1 - \frac{n_{tot}}{K}\right) \\
 \frac{dR_p^-}{dt} &= \psi_p \times R_p^- \times \left(1 - \frac{n_{tot}}{K}\right) \\
 \frac{dR^+}{dt} &= [\psi_\emptyset - \gamma_e (R_p^+ + R_p^-)] \times R^+ \times \left(1 - \frac{n_{tot}}{K}\right) \\
 \frac{dR^-}{dt} &= \psi_\emptyset \times R^- \times \left(1 - \frac{n_{tot}}{K}\right) \\
 n_{tot} &= R_p^+ + R_p^- + R^+ + R^- \\
 p_{tot} &= R_p^+ + R_p^-
 \end{aligned}$$

Computer simulations were conducted using MATLAB.

## Chapter 4

# Co-selection of transfer and public good genes

### 4.1 Introduction

In the previous chapters we have identified some of the interactions between public good production and horizontal gene transfer. Namely, in structured populations, horizontal transfer of public good genes promotes cooperation by increasing relatedness at transferred loci (Chapter 2). Reversely, host genes promoting transfer can be selected for by the feedbacks of public good production from newly converted recipients, also when some structuring in the population is present (Chapter 3). However, when studying the effect of transfer we did not consider the competition between genotypes with and without transfer. Similarly, in the case of cooperation promoting transfer, we did not consider the competition between public-good producing and non-producing genotypes.

Here we remedy these simplifications by addressing the remaining question of how the selective pressures acting on both public good production and horizontal transfer interact. Will horizontal transfer be selected as soon as public good production is beneficial? Can the competition between public-good producers and non-producers affect the selection of transfer and vice-versa? Moreover, both chromosomal and plasmid genes are able to influence plasmid transfer [Frost and Koraimann, 2010]. Studies of the evolution of transfer have generally focused on mobile elements themselves [Turner et al., 1998, Turner, 2004, Smith, 2011]: plasmids, like other mobile elements or parasites, gain a selfish, direct advantage from transfer, which allows them to spread in host populations. This benefit is modulated by the availability of recipients and by the cost of horizontal transfer to the host that leads to a trade-off with vertical transfer [Turner et al., 1998, Haft et al., 2009]. On the contrary, so far we have not considered selective pressures acting on plasmids, but focused on chromosomes, which gain no direct advantage from transfer.

Thus, now we ask how the indirect selection of transfer through public good benefits interacts with the direct selection of transfer on the plasmid side, and with the selection of public good production itself: when are transfer and cooperation co-selected? In other words, in which conditions does the selection of one trait require the presence of the other

trait? We use a simulation approach and, as before, consider simple metapopulations with subpopulations differing in the frequency of competing genotypes. We first analyze the interaction between public good production and transfer from the plasmid side (with plasmid genes controlling transfer, section 4.2), then the interaction between public good production and host genes controlling transfer (section 4.3).

## 4.2 Co-selection of plasmid transfer and public good production genes

### 4.2.1 Design of the study

We analyze here the interplay between plasmid transfer and public good production: rates of transfer and production could be combined in numerous ways on competitor plasmids. We limit the analysis to the case of two competing plasmids, and vary the rates of transfer and production for one or two of the competitors. We focus mainly on the case where both public good production and transfer genes vary on the same plasmid: we compete a plasmid  $P_\alpha T_\gamma$  with varying rate of transfer  $\gamma$  and rate of public good production  $\alpha$  with a control plasmid  $P_0 T_0$  that has neither transfer nor public good production ability, and serves as a reference (Figure 4.1-A). We also briefly analyze the case where both plasmids produce public goods at the same rate  $\alpha$  but only one plasmid transfers at rate  $\gamma$  (Figure 4.1-B), and the case where both plasmids transfer at the same rate  $\gamma$  but only one plasmid produces public goods at rate  $\alpha$  (Figure 4.1-C), in order to assess the generality of our conclusions. We assume costs of production and transfer proportional to  $\alpha$  and  $\gamma$  respectively, and acting on the host's growth rate. We do not consider here within-cell plasmid competition, and assume that the presence of a plasmid within a cell prevents superinfection by another plasmid.

We model a simple metapopulation that consists in two populations differing in the ratio of  $P_\alpha T_\gamma$  plasmid to the reference plasmid, with a constant proportion of plasmid-free cells ( $R$ ) (Figure 4.1-D):  $s_1$  population is impoverished in  $P_\alpha T_\gamma$  cells, while  $s_2$  population is enriched in  $P_\alpha T_\gamma$  cells. We model the effects of public production in the same way as in the previous chapters, following closely the experimental protocol that was used in order to allow comparisons with previous results. Competitors are first mixed ( $t_0$ ) and growth and transfer happen from  $t_0$  to  $t_1$ , with no effect from public good production except its cost. Populations are then diluted in a selective medium where growth until

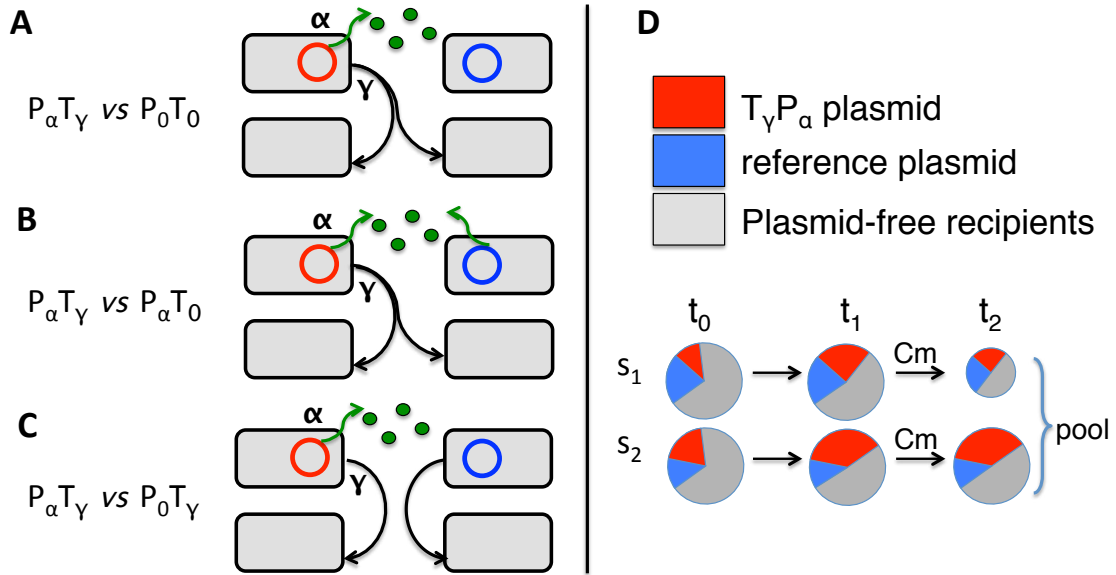


Figure 4.1: **Co-selection of plasmid transfer and public good production: simulation design.** **Competing strains and plasmids (A, B, C):** Two plasmids (red and blue) are competed in the presence of plasmid-free recipient cells (gray) that can acquire plasmids. Green dots represent the public good, black arrows indicate plasmid transfer. A plasmid  $P_\alpha T_\gamma$  (red) with rate of transfer  $\gamma$  and rate of public good production  $\alpha$  is competed against a reference plasmid (blue). The reference is different across panels: a null plasmid  $P_0 T_0$  that neither transfers horizontally nor produces public goods (A), a plasmid  $P_\alpha T_0$  that produces public goods at the same rate  $\alpha$  but does not transfer (B) or a plasmid  $P_0 T_\gamma$  that transfers at the same rate  $\gamma$  but does not produce public goods (C). **Experiment design (D):** Cells bearing  $P_\alpha T_\gamma$  plasmid (red) and cells bearing the reference plasmid (blue) are mixed in a metapopulation consisting of two populations that differ in their initial ratios of  $P_\alpha T_\gamma$  to the reference plasmid (at  $t_0$ ), and present the same promotion of plasmid-free cells:  $s_1$  is impoverished of  $P_\alpha T_\gamma$ -bearing cells,  $s_2$  is enriched in  $P_\alpha T_\gamma$ -bearing cells. After growth and transfer (function of  $\gamma$ ), populations are diluted in the selective medium (Cm) at  $t_1$ , where subsequent population growth (until  $t_2$ ) depends on the proportion of  $P_\alpha$  cells and  $\alpha$ .

$t_2$  is a function of the proportion of public-good producing cells, and of their production rate  $\alpha$ . As the benefits of public good production are transient, we arbitrarily define  $t_2$  as the timepoint where  $P_\alpha T_\gamma$  proportion at the metapopulation level is maximal across all conditions tested.

#### 4.2.2 Plasmid selection: transfer and cooperation interact

We focus first on the case where both transfer and public good genes are present only on  $P_\alpha T_\gamma$  plasmid (Figure 4.1-A). We measure the change in the frequency of  $P_\alpha T_\gamma$  compared to  $P_0 T_0$  from  $t_0$  to  $t_2$  as a function of  $\gamma$  and  $\alpha$  in a case where initially 25% of cells carry plasmids (Figure 4.2). In the absence of public good production ( $\alpha = 0$ ), increasing  $\gamma$  increases strongly the frequency of  $P_\alpha T_\gamma$  compared to  $P_0 T_0$ , as the population gets

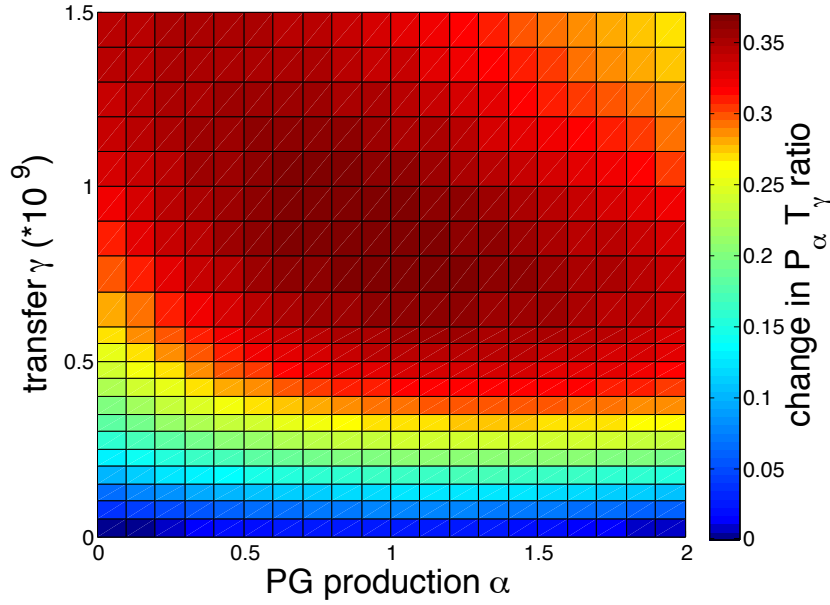


Figure 4.2: **Selection of  $P_\alpha T_\gamma$  plasmid with varying transfer and public good production rates.** The metapopulation consists of 2 populations, with initial  $P_\alpha T_\gamma$  vs  $P_0 T_0$  ratios of 1/4 and 4/1, and 25% total plasmid-bearing cells. The colorscale represents the change in the frequency of  $P_\alpha T_\gamma$  compared to  $P_0 T_0$  from  $t_0$  to  $t_2$ , shown as a function of public good production rate  $\alpha$  and transfer rate  $\gamma$  by  $P_\alpha T_\gamma$ .

progressively invaded by mobile  $P_\alpha T_\gamma$  plasmids. The effect of transfer rate  $\gamma$  on invasion saturates around  $10^{-9}$  mL cell $^{-1}$ h $^{-1}$ . In the absence of transfer ( $\gamma = 0$ ), there is a very slight increase in  $P_\alpha T_\gamma$  frequency for intermediate production rates ( $\alpha$  around 1).

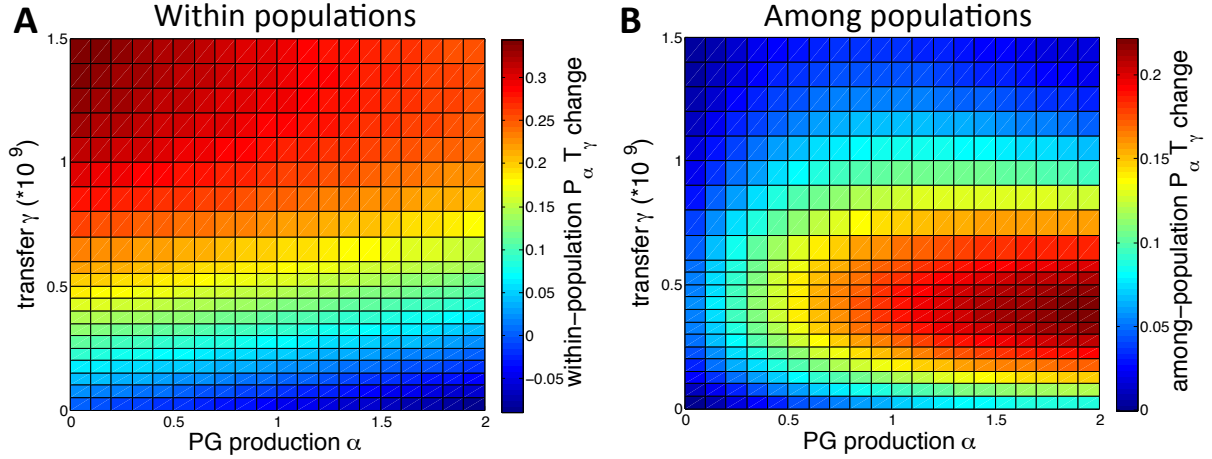
When  $P_\alpha T_\gamma$  plasmid both transfers and produces public goods ( $\gamma > 0$  and  $\alpha > 0$ ), we observe an interaction between transfer and public good production: the selection of  $P_\alpha T_\gamma$  plasmid is maximized for intermediates rates of both transfer and public good production. The interaction is not only additive, but transfer and public good production act in synergy: the increase in  $P_\alpha T_\gamma$  frequency with given  $\gamma$  and  $\alpha$  is greater than the sum of independant effects of  $\gamma$  and  $\alpha$  (data not shown). This synergistic effect happens only for intermediate rates: for higher  $\gamma$  and  $\alpha$ , the selection of  $P_\alpha T_\gamma$  declines. Strikingly, for production rates  $\alpha > 0$ , the transfer rate maximizing plasmid fitness is not the maximal transfer rate tested, contrary to what happens in the absence of public good production. Reversely, high production rates are also counterselected for a given transfer rate  $\gamma$ .

The results presented here are dependant of the timepoint chosen for  $t_2$ , as populations progressively reach saturation even in the absence of public good production. With different timepoints, the optimal range of parameters that maximize fitness is modified. Still, the two main conclusions from the analysis are conserved (data not shown): public good production enhances plasmid fitness for a given transfer rate, and decreases optimal

transfer rate. Two general conclusions can be drawn from this analysis: first, carrying public good production genes can benefit a conjugative plasmid. Secondly, adding public good production can decrease the optimal transfer rate for a given plasmid.

### 4.2.3 Transfer and cooperation interact via population structure

We next analyze the reasons for the observed interaction: we can distinguish within- and among-population dynamics. Within populations, increasing public good production



**Figure 4.3: Effect of transfer and public good production on changes within- and among populations.** The metapopulation parameters are the same as in Figure 4.2. The colorscales show the mean change in the frequency of  $P_\alpha T_\gamma$  compared to  $P_0 T_0$  within populations (**A**) and the part of the change in  $P_\alpha T_\gamma$  frequency due to among-populations differences (total change minus change within populations, **B**), as a function of public good production rate  $\alpha$  and transfer rate  $\gamma$  by  $P_\alpha T_\gamma$ .

lowers the frequency of  $P_\alpha T_\gamma$  plasmid (Figure 4.3-A), in the absence but also in the presence of transfer. Maximal invasion happens for the highest transfer rate and no public good production. Thus, the observed interaction at the metapopulation scale does not arise through within-population dynamics: public-good production only brings a cost to host cells, decreasing plasmid vertical transmission. Among populations however, we again observe a strong interaction between transfer and public good production concerning the selection of  $P_\alpha T_\gamma$  plasmid (Figure 4.3-B), that explains the observed interaction at the metapopulation level: among-population selection peaks for high public good production and intermediate transfer rates.

Why is the among-population increase in  $P_\alpha T_\gamma$  maximal for intermediate transfer rates? The effect can be explained by the actual growth of populations (Figure 4.4-A): the difference in total growth between the two populations is maximal for high public good production (as the public good enhances growth rate for all cells in the population) and



intermediate transfer rates. The effect of transfer rate can be understood by analyzing plasmid spread at  $t_1$  with varying transfer rates (Figure 4.4-B): in both populations,  $P_\alpha T_\gamma$  frequency at  $t_1$  increases with transfer rate  $\gamma$  (shown here for  $\alpha = 1$ ). For the  $s_2$  population enriched in  $P_\alpha T_\gamma$  plasmid (red line), plasmid spread saturates when it reaches 90% of the population. For the  $s_1$  population where  $P_\alpha T_\gamma$  is initially unfrequent (blue line), plasmid invasion still increases for higher transfer rates: the difference in producer cells abundance among populations decreases with increasing transfer rate. Differences among populations are thus maximized for intermediate transfer rates. The saturation

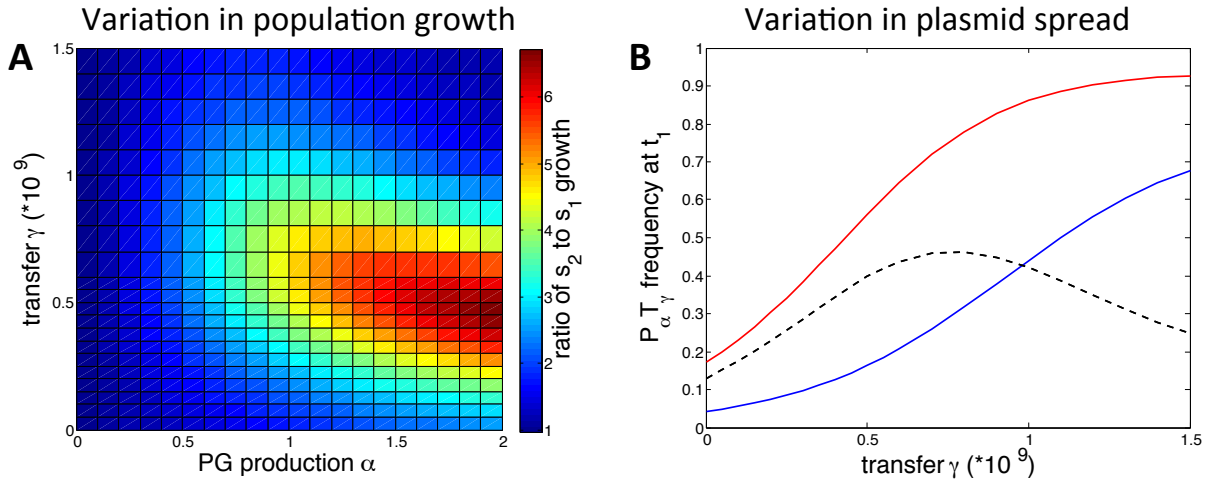


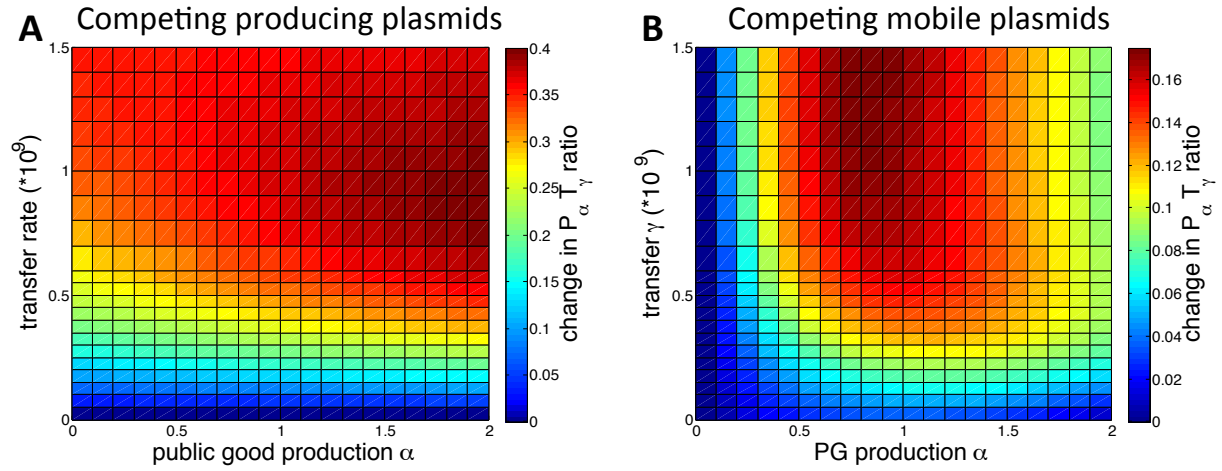
Figure 4.4: **Reasons for the effects of transfer rate on fitness.** The metapopulation parameters are the same as in Figure 4.2. **A:** The colorscale shows the ratio of  $s_2$  to  $s_1$  populations growth at  $t_2$ , as a function of public good production rate  $\alpha$  and transfer rate  $\gamma$ . **B:**  $P_\alpha T_\gamma$  spread within populations  $s_1$  (blue) and  $s_2$  (red), and its difference between  $s_2$  and  $s_1$  (dashed black line) are shown at  $t_1$  (after transfer but before public goods benefit growth) as a function of  $P_\alpha T_\gamma$  transfer rate  $\gamma$ .

of transfer is dependant on the assumption that  $P_\alpha T_\gamma$  plasmid is initially present in all populations: extremely high transfer rates lead to complete invasion of all populations, irrespectively of the plasmid initial proportion. With a metapopulation where  $P_\alpha T_\gamma$  and  $P_0 T_0$  plasmids do not coexist at all in populations, the saturation disappears, and  $P_\alpha T_\gamma$  plasmid gets the highest fitness with highest transfer rate (data not shown).

#### 4.2.4 Co-selection of transfer and public good production with other competition setups

We next analyze the competitions where either public good production or transfer rate are the same in both plasmids. In one case,  $P_\alpha T_\gamma$  plasmid competes with  $P_\alpha T_0$  plasmid that produces public goods at the same rate but does not transfer (Figure 4.1-B). In the

other case,  $P_\alpha T_\gamma$  plasmid competes with  $P_0 T_\gamma$  plasmid that transfers at the same rate but does not produce public goods (Figure 4.1-C).



**Figure 4.5: Fitness with different combinations of transfer and public good production genes on plasmids.** The metapopulation structure is the same as in Figure 4.2, with a different reference plasmid as competitor. In **A**,  $P_\alpha T_\gamma$  plasmid competes against  $T_0 P_\alpha$  plasmid. In **B**,  $P_\alpha T_\gamma$  plasmid competes against  $T_\gamma P_0$  plasmid. The colorscale shows the change in the frequency of  $P_\alpha T_\gamma$  compared to the reference plasmid from  $t_0$  to  $t_2$ , as a function of public good production rate  $\alpha$  and transfer rate  $\gamma$ .

When both plasmids produce the public good (Figure 4.5-A), transfer provides a direct benefit to the mobile plasmid ( $P_\alpha T_\gamma$ ), which increases with public good production: public good production and transfer still act synergistically on plasmid fitness. As previously, this synergy brings a decrease in the optimal transfer rate (compared to its optimum in the absence of public good production), due to the fact that among-population selection is maximized for intermediate transfer rates which yield the highest variation in producer proportion among populations (data not shown). The main difference from Figure 4.2 is that here high production rates are not counter-selected. Indeed, plasmids do not suffer from the cost of higher production here, as the competing plasmid in this scenario also produces the public good. For a given production rate, a mutant plasmid with higher or lower transfer rate can be selected depending on its effect on the strength of among-population selection.

When both plasmids transfer (Figure 4.5-B), we again observe an interaction of transfer and public good production:  $P_\alpha T_\gamma$  plasmid is more strongly favored with higher transfer rates. The competition is actually exactly the same as the one described in Chapter 2 (but focusing on plasmids and not on public good production itself). Increases in public good production rate are more strongly selected when plasmids have high transfer rates.

The synergy between plasmid transfer and plasmid public good genes is present for various competition setups: intermediate transfer rates generally promote fitness in synergy with public good production. Still, we have not considered actual competition over transfer between plasmids: indeed, either  $P_\alpha T_\gamma$  was the only mobile plasmid or both plasmids transferred at the same rate. Can a decrease in optimal transfer still be observed when the competing plasmid transfers at a high rate? We consider an additional case, where  $P_\alpha T_\gamma$  plasmid now competes against a non-producing but highly mobile plasmid,  $P_0 T_{high}$  (Figure 4.6).

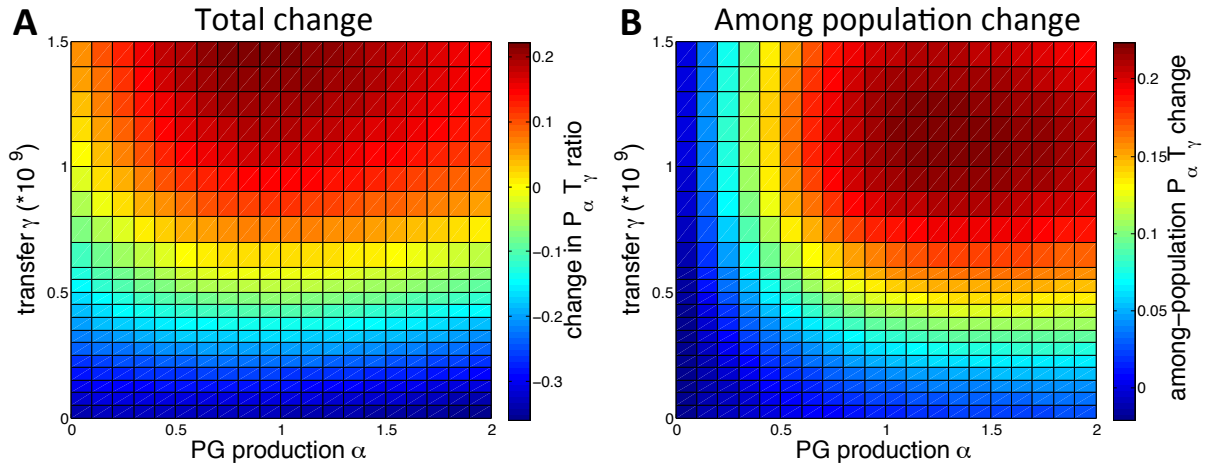


Figure 4.6: **Selection of a public-good producing plasmid *vs* a non-producing, highly mobile plasmid.** The metapopulation structure is the same as in Figure 4.2, with  $P_\alpha T_\gamma$  competing against  $P_0 T_{high}$  plasmid that has  $\gamma = 10^{-9} \text{mL} \cdot \text{cell}^{-1} \text{h}^{-1}$ . The colorscales show the total change in the frequency of  $P_\alpha T_\gamma$  compared to  $P_0 T_{high}$  from  $t_0$  to  $t_2$  (A) and the part of the change due to among-population dynamics (B), as a function of  $P_\alpha T_\gamma$  public good production rate  $\alpha$  and transfer rate  $\gamma$ .

At the metapopulation scale (Figure 4.6-A),  $P_\alpha T_\gamma$  is generally outcompeted in the absence of public-good production when its transfer rate is lower than  $\gamma = 10^{-9} \text{mL} \cdot \text{cell}^{-1} \text{h}^{-1}$  (the rate of transfer of the competing plasmid  $P_0 T_{high}$ ) and outcompetes  $P_0 T_{high}$  with higher transfer rate. Public-good production modifies the issue of the competition: for a given transfer rate, it increases  $P_\alpha T_\gamma$  frequency. The increase is again due to among-population dynamics (Figure 4.6-B): selection among populations is enhanced with high  $P_\alpha T_\gamma$  transfer rates, that amplify among-population variations in  $P_\alpha T_\gamma$  frequency. However, the optimal transfer rate is strongly modified compared to what we observed in Figure 4.2: a decrease in transfer with public good production is not favored when the competing plasmid transfers at a high rate. The decrease in among-population selection observed in Figure 4.3-B does not occur (Figure 4.6-B). Indeed,  $P_\alpha T_\gamma$  transfer does not

lead to the invasion of all populations at saturation, as the competing plasmid also transfers and prevents total invasion by  $P_\alpha T_\gamma$  plasmid. The presence of a competing plasmid leads to higher among-population selection when both plasmids transfer.

Overall, public good production still benefits the mobile  $P_\alpha T_\gamma$  plasmid in the presence of a competing highly mobile plasmid, but  $P_\alpha T_\gamma$  optimal transfer rate is modified because the range of parameters where among-population selection is the strongest is altered.

### 4.2.5 Influence of plasmid abundance on the interaction

We next analyze the influence of the initial abundance of plasmid-bearing cells, as it modulates plasmid infectious spread. We focus on a case where the conjugative plasmid  $P_\alpha T_\gamma$  has a production rate  $\alpha = 1$ , and competes against  $P_0 T_0$  plasmid (Figure 4.7).

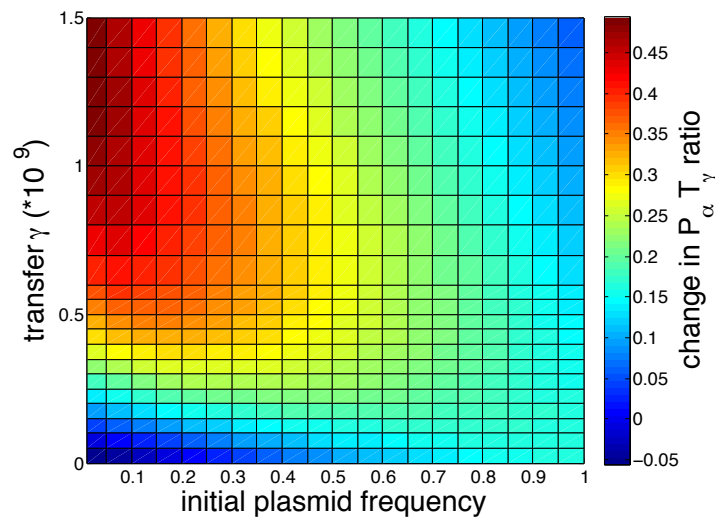


Figure 4.7: **Effect of initial plasmid abundance on the interaction between transfer and public good production.** The metapopulation consists of 2 populations, with initial  $P_\alpha T_\gamma$  vs  $P_0 T_0$  ratios of 1/4 and 4/1, and a varying initial ratio of plasmid-bearing to plasmid-free cells.  $P_\alpha T_\gamma$  has a public-good production rate  $\alpha = 1$ . The colorscale shows the change in frequency of  $P_\alpha T_\gamma$  compared to  $P_0 T_0$  from  $t_0$  to  $t_2$ , as a function of  $P_\alpha T_\gamma$  transfer rate and initial plasmid proportion.

In the absence of transfer ( $\gamma = 0$ ), the selection of  $P_\alpha T_\gamma$  increases with increasing plasmid proportion: when there are more plasmids, relatedness among public good producers is higher and populations enriched in public good producing plasmids get more feedbacks from public good production. The effect of transfer then depends on the initial plasmid proportion: in the absence of plasmid-free cells (initial plasmid proportion = 1), there is no effective transfer, so  $P_\alpha T_\gamma$  decreases in frequency with increased transfer ability, as  $P_\alpha T_\gamma$  cells still bear the cost of transfer. When many plasmid-free cells are present (initial plasmid proportion close to 0), transfer strongly increases  $P_\alpha T_\gamma$ , as there is opportunity

for infectious spread. For intermediate proportions of plasmid-free cells, an intermediate transfer rate is optimal.

Again, the effects of within- and among-population dynamics can be compared in order to understand the global dynamics (Figure 4.8). Within populations, the change in  $P_\alpha T_\gamma$  frequency is maximal for high transfer rates and low initial plasmid proportion (Figure 4.8-A), because these conditions allow a complete invasion of populations by  $P_\alpha T_\gamma$  infectious spread. As less plasmid-free cells are available, the opportunity for fitness gains by infectious spread within populations declines and becomes compensated by the cost of transfer: high transfer ability does not bring high benefits anymore within populations.

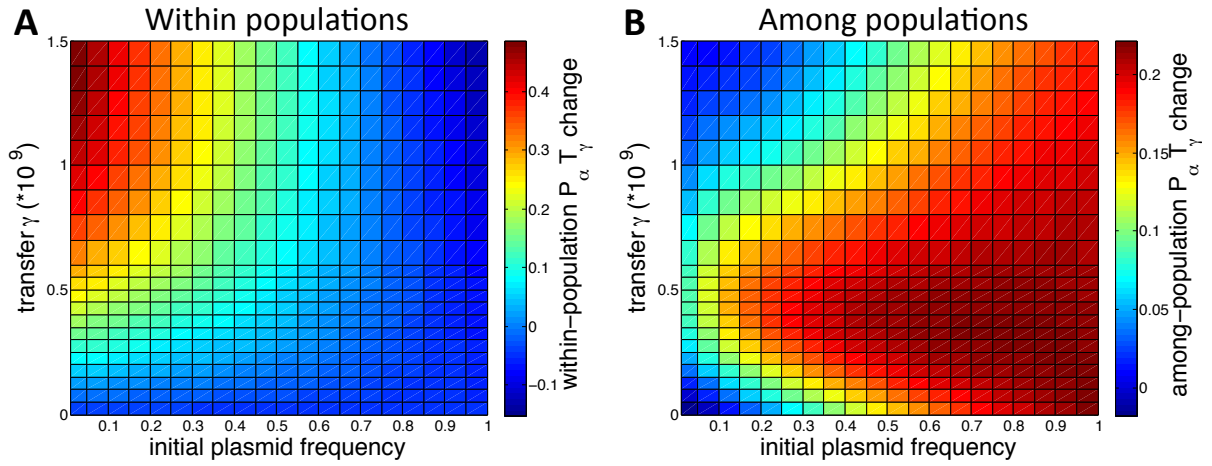


Figure 4.8: **Effect on recipient abundance: within- and among-population dynamics.** Population structure and parameters are the same as in Figure 4.7. The colorscales show the mean change in the frequency of  $P_\alpha T_\gamma$  compared to  $P_0 T_0$  within populations (A) and the change in  $P_\alpha T_\gamma$  frequency compared to  $P_0 T_0$  due to among-populations differences (B) from  $t_0$  to  $t_2$ , as a function of transfer rate  $\gamma$  and initial plasmid proportion.

The among-population differences in growth linked to public-good production, however, are highest when plasmids are more abundant, which maximizes feedbacks among public-good producing cells. In the absence of transfer, the among-population component of  $P_\alpha T_\gamma$  change in frequency is thus stronger with higher initial plasmid frequency (Figure 4.8-B). The effect of transfer is maximized for intermediate proportions of plasmid-free cells: in the absence of plasmid-free cells (initial plasmid proportion close to 1), among-population selection is already strong in the absence of transfer, and does not increase a lot with transfer, that is not efficient. When many plasmid-free cells are available, transfer can actually increase among-population selection, as it is low without transfer and there is more opportunity for efficient transfer. For low initial plasmid frequencies (close to

zero), transfer first increases among-population selection. However, the among population change in frequency then decreases for higher values of transfer rate. The dynamics here is similar to the one in Figure 4.3-B: the difference in transfer efficiency in populations differing in initial plasmid proportion actually decreases as populations get completely invaded by  $P_\alpha T_\gamma$  plasmid (see Figure 4.4-B). The within-population component of plasmid success increases and the among-population component decreases.

The influence of public good production on plasmid fitness is thus globally dependant on the opportunities for transfer. When plasmids are initially rare, the effect of infectious transfer is predominant, and strongly increases plasmid frequency. When plasmids are more abundant, the effect of transfer on among-population dynamics becomes a significant part of plasmid selection. Synergy between plasmid transfer and production rates happens mainly when benefits of public goods are strong enough to influence significantly plasmid dynamics.

### 4.3 Co-selection of chromosomal transfer and plasmid cooperation genes

We now analyze the interaction between the selection of host genes controlling transfer in *trans* (located on the chromosome, unlike in section 4.2 when they were on the plasmid) and public good production: we compete simultaneously two strains differing in donor ability  $q$ , and two plasmids differing in their public good production rate  $p$ . We assume that both plasmids can be transferred at the same rate by a given strain (the chromosome has no ability to discriminate between plasmids). We analyze the selective pressures acting simultaneously on public good production (competition between plasmids) and transfer ability (competition between strains).

#### 4.3.1 Design of the study

We compete the two strains  $D_q$  and  $D_0$  and the two plasmids  $P_\alpha$  and  $P_0$ :  $D_q$  transfers both  $P_\alpha$  and  $P_0$  plasmids to both  $D_q$  and  $D_0$  plasmid-free cells (Figure 4.9-A). We assume a cost for both public-good production and donor ability, proportional to respectively  $\alpha$  and  $\gamma$  and acting on the host's growth rate. To detect selection acting at the group level, we need populations that differ in the initial frequency of competitors, here both  $D$  strains and  $P$  plasmids. To manipulate strains and plasmids independantly, we model a metapopulation that consists in 4 populations (Figure 4.9-B): the total frequency of

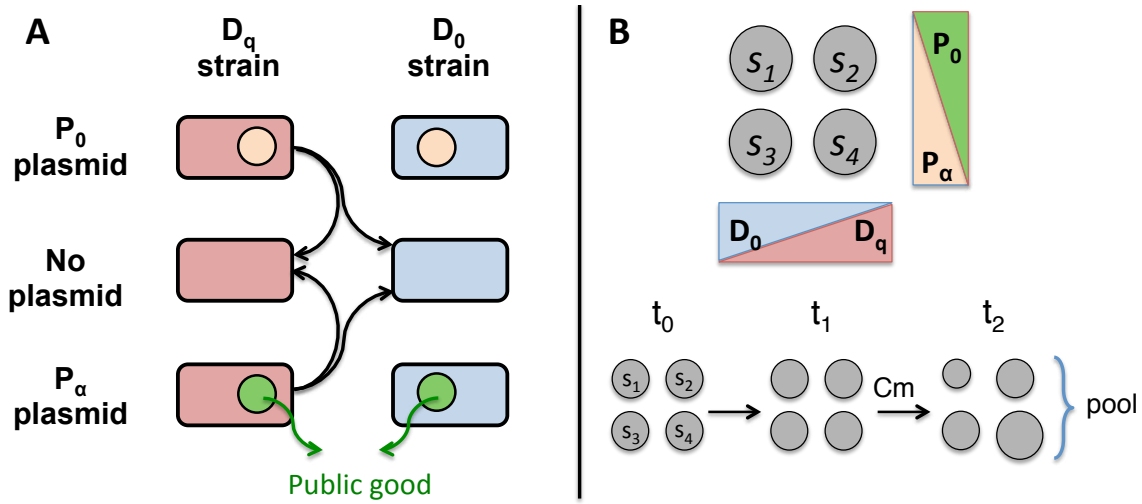


Figure 4.9: **Co-selection between donor ability and public good production: simulations design. Competing strains and plasmids (A):** Two strains are competed:  $D_q$  with donor ability  $q$  (red) and  $D_0$  without donor ability (blue). Some cells of each strain bear the non-producing plasmid  $P_\alpha$  (yellow) while others carry the public-good producing plasmid  $P_0$  (green). Plasmid-bearing  $D_q$  cells transfer plasmids to plasmid-free cells of either  $D_q$  or  $D_\alpha$  strain (black arrows). **Experiment design (B):** The metapopulation consists in 4 populations:  $s_1$  and  $s_2$  are enriched in  $P_\alpha$ -bearing cells,  $s_3$  and  $s_4$  in  $P_0$ -bearing cells.  $s_1$  and  $s_3$  are enriched in  $D_q$  strain,  $s_2$  and  $s_4$  are enriched in  $D_0$  strain. Each population has the same total proportion of plasmids  $p$ . After founding of populations ( $t_0$ ), cells grow and transfer happens. At  $t_1$ , populations are diluted in the presence of the antibiotic (Cm). Subsequent growth depends on the proportion of  $P_\alpha$ -bearing cells and  $P_\alpha$  production rate  $\alpha$ , until  $t_2$  where populations are pooled.

plasmids ( $p = P_\alpha + P_0$ ) is the same for all populations.  $P_\alpha$  vs  $P_0$  ratio is low for the two populations  $s_1$  and  $s_2$ , high for the two others,  $s_3$  and  $s_4$ .  $D_q$  to  $D_0$  ratio is low for  $s_1$  and  $s_3$ , high for  $s_2$  and  $s_4$ . Variations in plasmids and strains are independent, and allow us to follow changes in frequencies within and among populations for each trait.

The experiment design is the same as in the previous section: competitors are first mixed ( $t_0$ ) and growth and transfer happen from  $t_0$  to  $t_1$ . Populations are then diluted in medium containing the antibiotic chloramphenicol (Cm), where growth until  $t_2$  is a function of the proportion of public-good producing cells and their production rate  $\alpha$ .

### 4.3.2 Co-selection of chromosomal transfer genes and cooperation

We begin with a case where previous analyses have shown a benefit of transfer (Chapter 2) or its selection (Chapter 3): an abundance of plasmid-free cells (we assume here  $p = 0.1$ ) and non-random population structure for both plasmids and strains. We vary the public-good production rate  $\alpha$  for  $P_\alpha$ -bearing cells and the donor ability  $q$  for  $D_q$  cells, and measure the change in frequency of competitors from  $t_0$  to  $t_2$  (Figure 4.10).



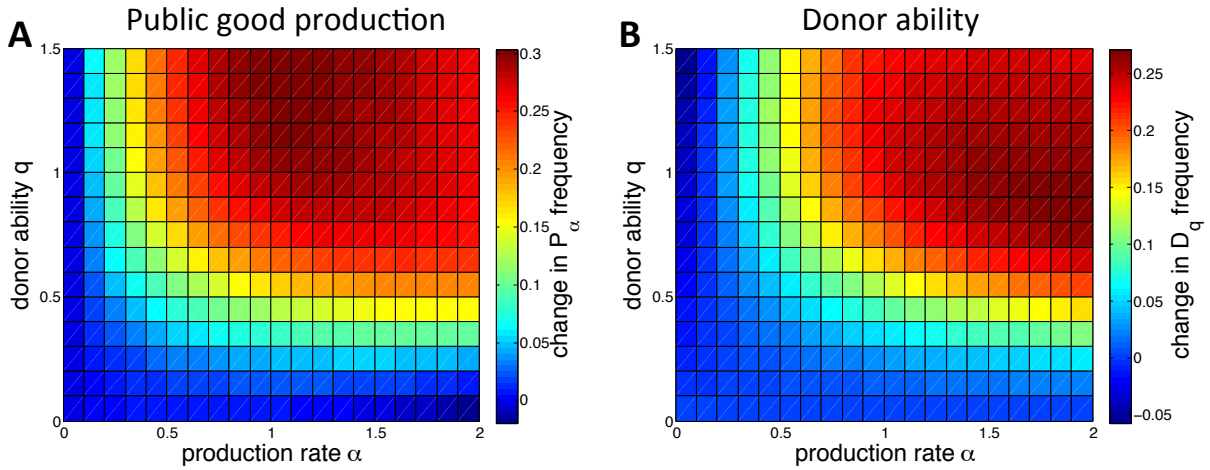


Figure 4.10: **Co-selection of donor ability and public good production.** The metapopulation consists in 4 populations (see Figure 4.9), each with initially 10% plasmid-bearing cells ( $p = 0.1$ ). The ratio of  $P_\alpha$  to  $P_0$  plasmids is 1/9 in  $s_1$  and  $s_2$ , 9/1 in  $s_3$  and  $s_4$ ; the ratio of  $D_q$  to  $D_0$  strains is 1/9 in  $s_1$  and  $s_3$ , 9/1 in  $s_2$  and  $s_4$ . The colorscales show the change in  $P_\alpha$  to  $P_0$  frequency (A) and in  $D_q$  to  $D_0$  frequency (B) from  $t_0$  to  $t_2$ , as a function of  $D_q$  donor ability  $q$  and  $P_\alpha$  public-good production  $\alpha$ .

In the absence of donor ability ( $q = 0$ ),  $P_\alpha$  plasmids decline in frequency compared to  $P_0$  plasmids (Figure 4.10-A). When  $D_q$  donor ability increases, higher values of  $P_\alpha$  production rate are selected. The maximal selection of  $P_\alpha$  happens for high  $D_q$  donor ability and intermediate  $P_\alpha$  production rate. All the observed increase is due to among-population dynamics:  $P_\alpha$  decline in frequency within populations for all values tested (data not shown). Why does among-population dynamics increase  $P_\alpha$  frequency? At the metapopulation scale, the relatedness at  $P_\alpha$  locus increases with transfer at  $t_1$  (Figure 4.11-A): as plasmids spread within populations, the assortment between  $P_\alpha$  plasmids increases, equivalently to what we observed in Chapter 2. The effect of plasmid transfer is the same as in Chapter 2, even if now only one part of the total population ( $D_q$  cells) are transferring the plasmid.

In the absence of public good production ( $\alpha = 0$ ),  $D_q$  cells decline in frequency compared to  $D_0$  cells (Figure 4.10-B). As  $D_q$  cells transfer a plasmid that does not bring public good benefits, donor ability is not selected. When public good production increases, donor ability is selected. The maximal selection of  $D_q$  happens for maximum  $P_\alpha$  production rate and intermediate  $D_q$  transfer rate. Again, all the increase is due to among-population dynamics:  $D_q$  decline in frequency within populations for all values tested (data not shown).  $D_q$  is selected at the metapopulation level, as populations enriched in  $D_q$  cells grow faster than populations with low frequencies of  $D_q$  cells: they have



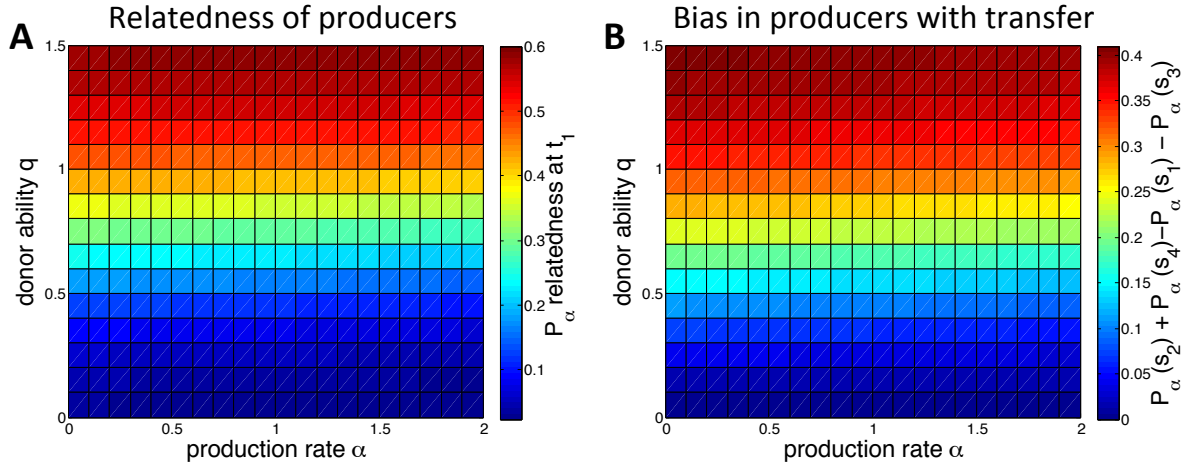


Figure 4.11: **Parameters controlling the co-selection of donor ability and public good production.** The metapopulation is the same as shown in Figure 4.10. Colorscales show the relatedness at the  $P_\alpha$  locus (A) and the difference in  $P_\alpha$  frequency in  $D_q$ -enriched populations compared to  $D_0$ -enriched populations ( $P_\alpha(s_2) + P_\alpha(s_4) - P_\alpha(s_1) - P_\alpha(s_3)$ , B) at  $t_1$ , as a function of  $D_q$  donor ability  $q$  and  $P_\alpha$  public-good production  $\alpha$ .

a higher proportion of  $P_\alpha$ -bearing cells (Figure 4.11-B). The selection of donor ability is analogous to what we observed in Chapter 3, even if  $D_q$  cells now transfer both  $P_\alpha$  and  $P_0$  plasmids.

Chromosomal genes promoting transfer and plasmid genes for public good production can thus be co-selected for some range of donor ability and public good production: here, public good production is selected only in the presence of transfer, and vice-versa. Interestingly, the optimum range of parameters (where selection is maximal) is not the same for donor ability and public good production:  $P_\alpha$  cells are maximally favored for maximal transfer rates (that ensures maximal relatedness), but not for maximal public good production, as  $P_\alpha$ -bearing cells pay the cost of increasing production. Reversely,  $D_q$  cells are maximally favored for maximal production rates (that ensure strong feedbacks from transfer), but not for maximal donor ability, as  $D_q$  cells pay the cost of transfer.

### 4.3.3 Conditions for co-selection

We focused here on a case where we suspected that co-selection of donor ability and public good production would be favored. We next analyze the effect of some key population parameters on the co-selection of donor ability and public good production.

#### Plasmid abundance

We first analyze the effect of the initial proportion of plasmid-bearing cells  $p$ . We assume that  $P_\alpha$  public-good production  $\alpha = 1$ , and follow the change in frequency of  $P_\alpha$  and  $D_q$

from  $t_0$  to  $t_2$ , while varying plasmid proportion  $p$  and  $D_q$  donor ability  $q$  (Figure 4.12).

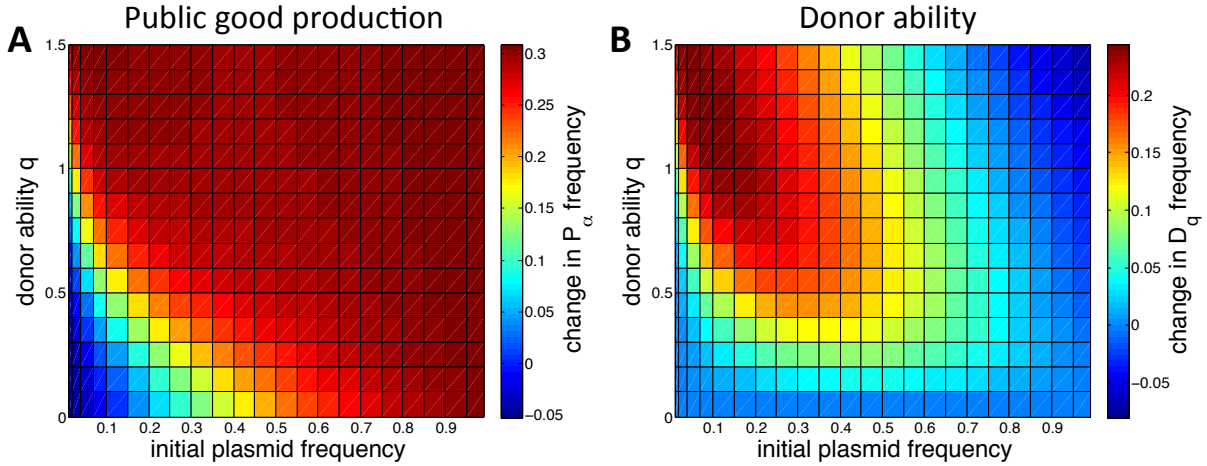


Figure 4.12: **Effect of plasmid abundance on the selection of donor ability and public good production.** The metapopulation consists in 4 populations (see Figure 4.9), each with initial proportion of  $p$  plasmid-bearing cells. The initial ratio of  $P_\alpha$  to  $P_0$  plasmids is  $1/9$  in  $s_1$  and  $s_2$ ,  $9/1$  in  $s_3$  and  $s_4$ ; the initial ratio of  $D_q$  to  $D_0$  strains is  $1/9$  in  $s_1$  and  $s_3$ ,  $9/1$  in  $s_2$  and  $s_4$ .  $P_\alpha$  plasmid has the production rate  $\alpha = 1$ . Colorscales show the changes in  $P_\alpha$  to  $P_0$  ratio (**A**) and  $D_q$  frequency (**B**) from  $t_0$  to  $t_2$ , as a function of  $D_q$  donor ability  $q$  and initial plasmid frequency  $p$ .

When plasmids are initially abundant,  $P_\alpha$  plasmids are favored independently from  $D_q$  donor ability (Figure 4.12-A): relatedness among  $P_\alpha$  is already high in the absence of transfer, and a high donor ability does not lead to many transfer events, as plasmid-free cells are rare. When plasmids are initially rare,  $P_\alpha$  plasmids are favored only with high donor ability  $q$ : transfer is needed to select for cooperation, as relatedness among  $P_\alpha$  plasmids is too low without transfer. On the contrary,  $D_q$  cells are selected only when plasmids are initially rare (Figure 4.12-B): when plasmids are initially abundant,  $D_q$  decline in frequency when they have high donor ability. A higher donor ability is selected with lower initial proportion of plasmids. Indeed, donor ability brings benefits on population growth only when transfer is effective and provides strong benefits in terms of public good production:  $D_q$  cells are selected in the same range of parameters for which  $P_\alpha$  cells cannot be selected in the absence of transfer, *ie* a low initial plasmid frequency.

Public good production and donor ability are thus co-selected when plasmid-free cells are abundant: in these conditions, cooperation is counter-selected in the absence of transfer, and transfer benefits growth, allowing for second-order selection of transfer. When plasmid-free cells are scarce, public-good production is selected, but donor ability is not.

### Population structure requirements

We assumed until now a fixed population structure with given variation among populations for both  $D$  and  $P$  genotypes. We now modulate population structure by changing the initial asymmetry among populations concerning either  $P$  or  $D$  genotypes: we vary the difference in  $P_\alpha$  vs  $P_0$  or  $D_q$  vs  $D_0$  proportions among populations from 0 to 1. (When asymmetry is 0, there is no difference among populations; when asymmetry is 1, each population contains only one allele for the locus considered. So far we considered asymmetries of 0.8 for both  $P$  and  $D$  genotypes.) We follow the change in  $P_\alpha$  and  $D_q$  frequency at the metapopulation scale from  $t_0$  to  $t_2$  (Figure 4.13).

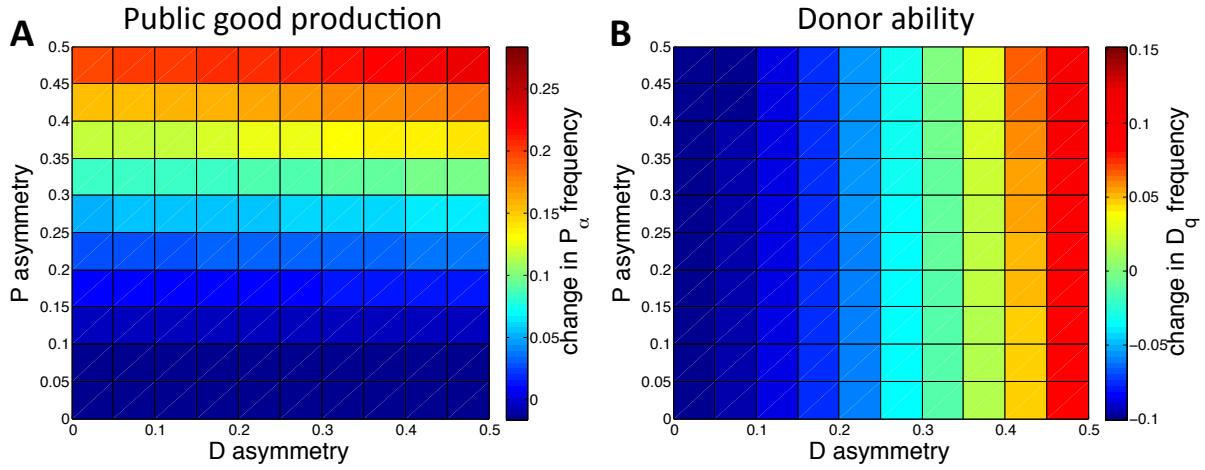


Figure 4.13: **Effects of population structure on the selection of donor ability and public good production.**  $P_\alpha$  plasmid has the production rate  $\alpha = 1$ ,  $D_q$  strain has the donor ability  $q = 1$ . The metapopulation consists in 4 populations (see Figure 4.9), each with initial proportion of 20% plasmid-bearing cells, and varying distribution of  $P$  and  $D$  genotypes: D asymmetry measures the initial difference in  $D_q$  frequency between  $D_q$ -rich and  $D_q$ -poor populations; P asymmetry measures the initial difference in  $P_\alpha$  frequency between  $P_\alpha$ -rich and  $P_\alpha$ -poor populations. Colorscales show the changes in  $P_\alpha$  to  $P_0$  frequency (A) and  $D_q$  to  $D_0$  frequency (B) from  $t_0$  to  $t_2$ , as a function of  $P$  and  $D$  genotypes asymmetry.

$P_\alpha$  plasmids are selected over  $P_0$  plasmids (Figure 4.13-A) only when there is some initial asymmetry in  $P_\alpha$  vs  $P_0$ . Reversely,  $D_q$  strain is selected over  $D_0$  strain (Figure 4.13-B) only when there is some initial asymmetry in  $D_q$  vs  $D_0$ . Both donor ability and public good production genes thus need some variation in their frequencies among populations, or relatedness at their own locus, in order to be selected.

When  $P_\alpha$  and  $P_0$  do not present initial asymmetry, the ratio of  $P_\alpha$  to  $P_0$  plasmids within populations does not change, even with transfer, as  $D_q$  transfers equally both plasmids. There is thus no possibility for among-population selection of public-good

producers. The total amount of  $P_\alpha$ -bearing cells, however, still gets higher in populations enriched in  $D_q$  cells when there is asymmetry concerning  $D$  genotype. The asymmetry in  $P_\alpha$ -bearing cells further provides increased feedbacks from public-good benefits to  $D_q$  cells, and allows among-population selection of donor ability.

When there is no asymmetry in  $D_q$  vs  $D_0$  frequencies,  $P_\alpha$  cells are still selected if some asymmetry in  $P_\alpha$  vs  $P_0$  frequencies preexists: this is exactly what we described in Chapter 2. Transfer by  $D_q$  cells promotes selection of  $P_\alpha$  by increasing relatedness at the  $P_\alpha$  locus. However, since in this case there is no initial relatedness concerning donor ability, the  $D_q$  cells do not get more feedbacks than  $D_0$  cells from public-good production and their frequency does not increase.

Finally, when both  $D$  and  $P$  genotypes have non-random population distribution, we observe that asymmetry promotes the selection of both  $D_q$  and  $P_\alpha$ . Strong asymmetry in  $D$  distribution increases the selection of  $P_\alpha$  plasmids (Figure 4.13-A); strong asymmetry in  $P$  distribution also increases the selection of  $D_q$  strain (Figure 4.13-B). Indeed,  $D_q$  cells amplify the pre-existing variation in  $P_\alpha$  proportion among populations, thanks to within-population transfer, and themselves get feedbacks from  $P_\alpha$  proportion.

Non-random assortment among  $D_q$  cells and among  $P_\alpha$  plasmids is thus needed to select respectively for donor ability and public-good production. Still, we assumed here for simplification that the distributions of  $D$  and  $P$  genotypes across populations are independent (so we can independently manipulate the population distribution of each of them). Due to this independence, the assortment for one genotype does not lead to assortment for the other one. However, the interaction between donor ability and public good production can modify population distribution itself, which we address below.

#### 4.3.4 Association between plasmids and host transfer genes

To follow changes in the association between  $P$  and  $D$  genotypes, we compute linkage disequilibrium between the  $P_\alpha$  and  $D_q$  alleles.

##### Linkage disequilibrium between transfer and plasmids

Linkage disequilibrium (LD) is a measure used in population genetics that quantifies the deviation from random association between alleles at different loci. Measures analogous to LD have been applied to quantify the linkage between host and symbiont [Sanchez et al., 2000] or gene-culture disequilibrium [Feldman and Cavalli-Sforza, 1984]. We apply

it here to the case of  $P_\alpha$  and  $D_q$  alleles:

$$LD_{D_q P_\alpha} = f(D_q P_\alpha) - f(D_q) \times f(P_\alpha)$$

where  $f(D_q P_\alpha)$  is the frequency of  $D_q P_\alpha$  allele combination ( $P_\alpha$  plasmid in  $D_q$  cells) in the metapopulation,  $f(D_q)$  the total frequency of  $D_q$  strain and  $f(P_\alpha)$  the total frequency of  $P_\alpha$  plasmid. Generally, positive linkage indicates that  $D_q$  and  $P_\alpha$  are found together associated with each other more than what would be expected from their individual frequencies.

At  $t_0$ , we have  $LD_{D_q P_\alpha} = 0$  by design, as we start with an independant distribution of  $P$  and  $D$  genotypes across populations: there is no association between  $D_q$  and  $P_\alpha$ . We then compute  $LD_{D_q P_\alpha}$  after growth and transfer ( $t_1$ , Figure 4.14-A), and after further growth in presence of the antibiotic ( $t_2$ , Figure 4.14-B).

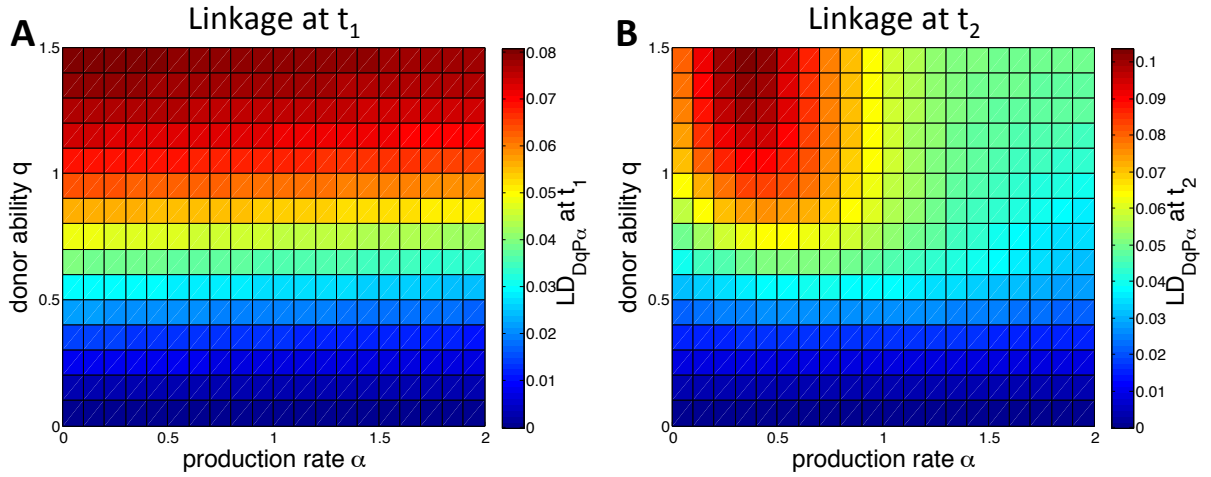


Figure 4.14: **Linkage disequilibrium between  $D_q$  and  $P_\alpha$  alleles.**  $LD_{D_q P_\alpha}$  is computed for the metapopulation described in Figure 4.10, and shown as a colorscale at  $t_1$  (A) and  $t_2$  (B) as a function of  $D_q$  donor ability  $q$  and  $P_\alpha$  public-good production  $\alpha$ .

At  $t_1$ ,  $LD_{D_q P_\alpha}$  is strongly dependent on  $D_q$  donor ability  $q$  (Figure 4.14-A). In the absence of transfer ( $q = 0$ ), there is no association between  $D_q$  and  $P_\alpha$  alleles. With increasing donor ability, positive linkage arises between  $D_q$  and  $P_\alpha$  alleles. Increasing  $P_\alpha$  production rate very slightly decreases linkage for a given donor ability. At  $t_2$  however,  $P_\alpha$  public good production has a strong impact on  $LD_{D_q P_\alpha}$  (Figure 4.14-B).  $LD_{D_q P_\alpha}$  is still low for low donor ability, and close to its values at  $t_1$  in the absence of public-good production ( $\alpha = 0$ ). For high donor ability, intermediate rates of  $P_\alpha$  public good production strongly increase  $LD_{D_q P_\alpha}$  compared to both its  $t_1$  value and its value in the absence of public good production. Further increase in production rates causes a decrease in  $LD_{D_q P_\alpha}$ .

At  $t_1$ , the emergence of linkage is not related to cooperation: linkage actually also appears between  $P_0$  and  $D_q$  alleles, following the same pattern ( $LD_{D_q P_0} > 0$  for  $q > 0$ , data not shown). We can thus infer that the linkage at  $t_1$  is primarily due to plasmid transfer. Indeed, the metapopulation structure ensures that more transfer takes place within  $s_2$  and  $s_4$  populations enriched in  $D_q$  cells: transfer happens mainly towards  $D_q$  cells at the metapopulation scale, and both  $P_\alpha$  and  $P_0$  become associated to  $D_q$  cells.

At  $t_2$  the situation is different:  $P_\alpha$  public good production also plays a role in linkage, which suggests that public goods modify linkage. As the public good acts through enhancing the growth rate in populations, we suspect that its effect on linkage may be due to among-population selection. To test this idea we computed  $LD_{D_q P_\alpha}$  while artificially removing the effect of differential growth (assuming equal size for all populations). The effect of public good production on  $LD_{D_q P_\alpha}$  disappeared (data not shown): the increase in  $LD_{D_q P_\alpha}$  at  $t_2$  is due to differential growth. Growth is mainly promoted in the  $s_4$  population (see Figure 4.9) initially enriched in both  $D_q$  cells and  $P_\alpha$  plasmids, where transfer by  $D_q$  cells increased the frequency of public-good producing cells. Since  $s_4$  (where  $D_q$  cells and  $P_\alpha$  are initially associated) is the most successful and  $s_1$  the less successful population,  $LD_{D_q P_\alpha}$  will increase at the metapopulation level. At  $t_2$ , the association between  $D_q$  strain and  $P_\alpha$  plasmid is thus selected through among-population selection of populations enriched in producer cells.

### Parameters governing linkage disequilibrium

Linkage disequilibrium in previous simulations relies on initial differences in population frequencies. How does population structure influence the emergence of linkage? We now modulate population structure in the same way as in the section 4.3.3, by independently varying the initial asymmetry in  $P$  and  $D$  genotypes.

At  $t_1$ , positive linkage between  $D_q$  strain and  $P_\alpha$  plasmids appears as soon as there is some initial asymmetry in  $D$  genotype among populations (Figure 4.15-A). The asymmetry in  $P$  genotype does not affect linkage. The linkage due to transfer thus needs some initial variation in genotype frequencies for donor ability  $q$  across populations. This observation can be related to what we observed in Chapter 3: a bias of transfer towards kin (other  $D_q$  cells) relies on some assortment between  $D_q$  cells at the metapopulation scale. Public good production does not play a role in the emergence of linkage at  $t_1$ , so population structure concerning  $P$  genotypes has no effect here.

At  $t_2$ ,  $D$  asymmetry still promotes positive linkage between  $D_q$  strain and  $P_\alpha$  in the

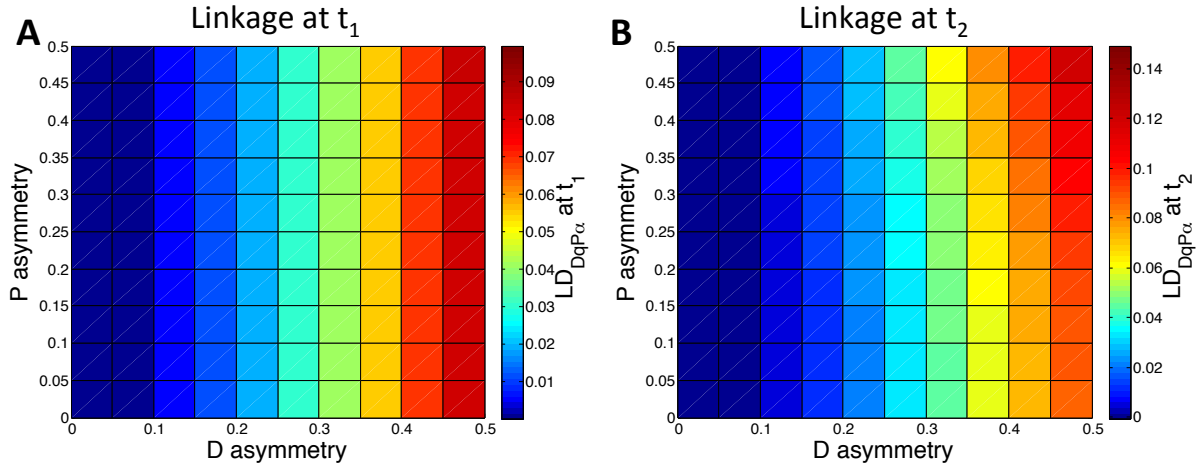


Figure 4.15: **Linkage disequilibrium and population structure.** The simulation is the same as in Figure 4.13, with  $\alpha = 1$  and  $q = 1$ .  $LD_{DqP_\alpha}$  is computed at the metapopulation scale, and shown as a colorscale at  $t_1$  (A) and  $t_2$  (B) as a function of  $P$  and  $D$  genotypes asymmetry among populations.

same way as at  $t_1$  (Figure 4.15-B). However, the asymmetry in  $P$  genotypes now also increases  $LD_{DqP_\alpha}$ , mainly when asymmetry for  $D$  is high: asymmetries in  $P$  and in  $D$  genotypes have a synergistic effect on linkage. In the absence of asymmetry for  $P$ , among-population selection cannot act on public good production: linkage due to  $P_\alpha$  public-good production thus arises through among-population selection, and is enhanced with increasing asymmetry in  $P$  (see Figure 4.13).

Our analysis shows that without initially enforcing any association between  $D$  and  $P$  genotypes, linkage between  $D_q$  strain and  $P_\alpha$  plasmids arises as soon as there is some asymmetry in  $D_q$  and  $P_\alpha$  distributions among populations. Part of it is simply due to the local nature of transfer and also happens for  $P_0$  plasmid (this requires only asymmetry for  $D$  genotype); but additional linkage specifically caused by  $P_\alpha$  arises through among-population selection because populations enriched in both  $D_q$  and  $P_\alpha$  are more successful.

## 4.4 Discussion

### 4.4.1 Conditions for the co-selection of transfer and cooperation

Our simulations have shown that transfer ability and public good production interact in terms of selective pressures acting on both plasmid and chromosomal genes:

On the plasmid side, the optimal transfer rate is modulated by the presence of a given public good production rate. By its indirect effects on among-population selection, transfer ability can be more strongly selected when public good production genes are present on mobile plasmids. However when it maximizes the variability in producer

frequencies among populations, a lower transfer rate can be selected. The evolution of plasmid transfer rate will thus be influenced by the public good genes that the plasmid may carry. The relative importance of selective pressures linked to public good production will depend on the strength of the direct benefit given by infectious spread (highest when plasmid-free cells are abundant) and on population structure (assortment is needed for indirect selection of public good production).

On the chromosomal side, selection for donor ability occurs in interaction with public good production, when transfer is needed in order for public good production to be selected; *ie* when plasmid-bearing cells are initially present in low frequencies. When plasmid-bearing cells are already abundant in the absence of transfer, public good production is selected alone, with no advantage brought by plasmid transfer ability. When plasmid-bearing cells are scarce, donor ability and public good production are co-selected: each trait necessitates the other trait in order to increase in frequency.

As all simulations share the same metapopulation setup and parameters for transfer and public good production, the cases of plasmid and chromosomal control of transfer can be roughly compared: co-selection between transfer and public good production generally occurs for effective transfer rates around  $10^{-9}\text{mL cell h}^{-1}$  and public production around 1, with low frequencies of initial plasmid-bearing cells (5 to 25%). This range of parameters corresponds to cases where transfer promotes the most the selection of public good production by increasing production and assortment among producers. From the plasmid side, this can have the effect of decreasing the optimal transfer rate compared to a case without public goods (where the optimal transfer rate is greater than  $10^{-9}\text{mL cell h}^{-1}$ ), but this depends on the competitor plasmid's transfer. From the host side, adding public good genes on plasmids promotes donor ability (as the optimum donor ability is zero for plasmids not bearing public good genes).

In both cases, an abundance of plasmid-free cells favors the selection of transfer: partly because of selfish spread from the plasmid side and purely because of among-population selection for public-good production from the chromosomal side. In conditions where transfer is likely to be efficient, both plasmid and chromosomal genes promoting plasmid transfer and the presence of public-good producing genes on plasmids should be favored. We analyzed here a specific example of public good, that acts on cell growth rate, and for which we have experimentally measured the actual parameters. However, the general conclusion of these simulations should hold in other cases: transfer and public good production could be co-selected as soon as transfer enhances the population benefits



provided by public goods. An interesting point we observed in some simulations modelling a public good effect on growth yield (and not growth rate) is that co-selection of transfer and public good production can occur also within populations: public good production leads to higher cell densities, most favorable to plasmid spread.

#### 4.4.2 Population structure effects

We modelled here a metapopulation divided in a small number of populations. Our analysis confirmed that some structure in the distribution of genotypes among populations is needed to provide indirect selection for both public good production and transfer ability, which was shown in Chapters 2 and 3. First, in order for public good production to be selected, producer plasmids have to present some initial structuring among populations, that is further amplified by transfer when plasmid-free cells are present. Secondly, transfer ability can be selected by indirect benefits provided by transferred genes, if there is initial structure in populations concerning transfer ability itself. However, our novel finding is that variation among populations for both transfer and production alleles leads to synergistic effects on fitness, and on population structure itself. Populations where both transfer and public good production are abundant get maximum benefits, which amplifies among-population selection for production and transfer. Moreover, this synergy favors populations where transfer and production alleles are associated within cells, leading to linkage between these alleles at the metapopulation scale.

The linkage between horizontal transfer and transferred alleles may seem paradoxical, since horizontal transfer has been shown to oppose associations between alleles [Brandvain et al., 2011]. However, linkage arises here because the association is selected at the metapopulation scale. The appearance of linkage can be compared to models of mutualism, that showed that genetic correlations among mutualists with horizontal transmission can arise because the association is favored [Frank, 1994]. Moreover, in structured populations, transfer is actually biased towards kin, which also favors the association between transferring strains and plasmids.

The appearance of linkage in these simple simulations, where we assumed initially independence between transfer and production alleles, suggests that strains with donor ability could become associated to public-good producing plasmids at larger timescales. This association could favor mutualistic coevolution between the host and the plasmids, even in the presence of horizontal transfer.

### 4.4.3 Other selective pressures acting on plasmids

The suggested coevolution between transferring strains and plasmids also hints at the issue of plasmid specialization towards a specific host background. We have not considered here plasmids that could be more adapted to one of several hosts: we assumed that host backgrounds differ only in plasmid transfer ability, not in plasmid effects on fitness. A differential effect of plasmids would be outside the scope of this study, as it would involve issues related to plasmid host range and specialization. For this reason, we also excluded here the analysis of transfer of private good genes (described in Chapter 3): transfer ability is selected on the chromosomal side only when transfer happens towards kin, which would require us to consider plasmids with limited host range. Instead, we focused on the case of transfer of public good genes, where transfer is beneficial through public goods produced by every new producer cell, no matter if it is kin or non-kin.

We also deliberately excluded here other selective pressures that can influence the evolution of plasmid transfer rates. First, the evolution of transfer rate is thought to depend on the tradeoff between horizontal and vertical transmission: the cost of horizontal transmission to the host decreases vertical transmission [Turner et al., 1998, Haft et al., 2009]. However, this effect is quite low in the conditions we model, as we assume high densities of plasmid-free hosts, that bring a strong benefit to horizontal transmission. Secondly, we do not consider within-host plasmid competition. Within-host competition is generally thought to be limited in nature, as conjugative plasmids present entry exclusion mechanisms that block the entry of related plasmids [Garcillán-Barcia and de la Cruz, 2008], and bottlenecks through cell division and conjugation lead to reduced diversity in plasmids within cells [Paulsson, 2002]. Still, superinfection between plasmids can evolve under conditions such as intense competition when no plasmid-free cells are available [Smith, 2011]. As superinfection is strongly deleterious to the host [Smith, 2012], it could prevent the evolution of mutualism between host and plasmid. The interplay of the benefits of transfer we describe in our work and selective pressures tending to decrease horizontal transmission or make it deleterious to the host could be studied in future work.

More generally, we did not consider here the coevolution between plasmids and hosts concerning plasmid transfer, but considered only the cases in which transfer is controlled either just by chromosomal or just by plasmid genes. However, since our analysis shows that horizontal transfer can benefit both plasmid and chromosomal genes separately when plasmids bear public good genes, we expect that the interests of plasmid and chromosomal

genes controlling transfer may often converge in the case of public good genes-bearing plasmids.

## 4.5 Methods

### 4.5.1 Description of the model

#### General model assumptions

We model growth of strains and plasmid transfer mimicking experimental conditions described in Chapters 2 and 3 for the case of plasmids carrying public good genes. Both growth and transfer follow a logistic function, stopping at carrying capacity  $K$ . Plasmids have a cost  $c_p$  on cells carrying them and a public benefit  $b_p$ . The plasmid affects cell growth proportionally to public good production rate  $\alpha$  ( $c_p \times \alpha$ ). We assume a cost of transfer for the donor cell  $c_q$  so that the cost affects cell growth and is proportional to donor ability ( $c_q \times q$ ).

We model two steps: first pre-incubation, where transfer happens (from  $t_0$  to  $t_1$ ), then growth in a selective medium where the public good benefits growth (from  $t_1$  to  $t_2$ ) and the transfer rate is divided by 10. The global effect of public good production is modelled as being proportional to the frequency of public-good producing cells at a given time in the local population. We thus assume that  $c_p < 0$  for the whole duration of the experiment (cost of public good production),  $b_p = 0$  from  $t_0$  to  $t_1$  (no benefit of the public-good in the absence of Cm) and  $b_p > 0$  from  $t_1$  to  $t_2$ . The pre-incubation time (from  $t_0$  to  $t_1$ ) was set to 12h after 100-fold initial dilution from carrying capacity, and growth from  $t_1$  was allowed for 60h after a second 100-fold dilution. The benefit of cooperation is only transient [Chuang et al., 2009] and  $t_2$  has to be chosen before all populations reach stationary phase. For each simulation,  $t_2$  was defined as the time point where the selection of public-good producers was maximal over all conditions tested.

Parameters values shown below were estimated from previous experimental data; the basal growth rate  $\psi$  is  $\psi_1$  from  $t_0$  to  $t_1$  and  $\psi_2$  from  $t_1$  to  $t_2$ .

Parameter	Symbol	Estimation
Carrying capacity	$K$	$4.10^9$ cells mL <sup>-1</sup>
Basal rate of plasmid transfer	$\gamma$	$10^{-9}$ mL cell <sup>-1</sup> L <sup>-1</sup>
Basal growth rate - from $t_0$ to $t_1$	$\psi_1$	$0.96$ h <sup>-1</sup>
- from $t_1$ to $t_2$	$\psi_2$	$0.12$ h <sup>-1</sup>
Cost of good production on growth rate	$c_p$	$0.04$
Benefit of public good production on growth rate	$b_p$	$4$
Cost of donor ability on growth rate	$c_q$	$0.05$

### Co-selection of donor ability and plasmid production

Here we compete both the two strains  $D_q$  and  $D_0$ , and the two plasmids  $P_\alpha$  and  $P_0$ .

$D_q$  cells transfers plasmids to both  $D_q$  and  $D_0$  cells: transfer happens at a basal rate  $\gamma = 10^{-9}$  mL cell<sup>-1</sup>h<sup>-1</sup>, following a mass-action law.  $D_q$  strain donor ability  $q$  modulates effective transfer  $\gamma_e$  ( $\gamma_e = q \times \gamma$ ).

$$\begin{aligned}
\frac{dD_q P_\alpha}{dt} &= [\psi(1 - c_p \alpha)(1 - c_q q)(1 + b_p \alpha \frac{p_{tot}}{n_{tot}}) + \gamma_e D_q] \times D_q P_\alpha \times (1 - \frac{n_{tot}}{K}) \\
\frac{dD_q P_0}{dt} &= [\psi(1 - c_q q)(1 + b_p \alpha \frac{p_{tot}}{n_{tot}}) + \gamma_e D_q] \times D_q P_0 \times (1 - \frac{n_{tot}}{K}) \\
\frac{dD_0 P_\alpha}{dt} &= [\psi(1 - c_p \alpha)(1 + b_p \alpha \frac{p_{tot}}{n_{tot}}) \times D_0 P_\alpha + \gamma_e D_q P_\alpha \times D_0] \times (1 - \frac{n_{tot}}{K}) \\
\frac{dD_0 P_0}{dt} &= [\psi(1 + b_p \alpha \frac{p_{tot}}{n_{tot}}) \times D_0 P_0 + \gamma_e D_q P_0 \times D_0] \times (1 - \frac{n_{tot}}{K}) \\
\frac{dD_q}{dt} &= [\psi(1 - c_q q)(1 + b_p \alpha \frac{p_{tot}}{n_{tot}}) - \gamma_e (D_q P_\alpha + D_q P_0)] \times D_q \times (1 - \frac{n_{tot}}{K}) \\
\frac{dD_0}{dt} &= [\psi(1 + b_p \alpha \frac{p_{tot}}{n_{tot}}) - \gamma_e (D_q P_\alpha + D_q P_0)] \times D_0 \times (1 - \frac{n_{tot}}{K}) \\
n_{tot} &= D_q P_\alpha + D_0 P_\alpha + D_q P_0 + D_0 P_0 + D_q + D_0 \\
p_{tot} &= D_q P_\alpha + D_0 P_\alpha
\end{aligned}$$

### Selection of production and transfer in plasmids

Here, two plasmids  $T_\gamma P_\alpha$  and  $T_{ref} P_{ref}$  are competed, in the presence of plasmid-free recipients  $R$ .  $T_\gamma P_\alpha$  has transfer rate  $\gamma_e = q_e \times \gamma$  and public good production  $\alpha$ .  $T_{ref} P_{ref}$  can be either  $T_0 P_0$ ,  $T_0 P_\alpha$  or  $T_\gamma P_0$ : it has transfer rate  $\gamma_r = q_{ref} \times \gamma$  and public good production  $\alpha_r$ . Transfer happens only to plasmid-free cells. Equations for growth and transfer are as follows (cells bearing  $T_\gamma P_\alpha$  are abbreviated as  $T_\gamma P_\alpha$  and cells bearing  $T_{ref} P_{ref}$

as  $T_{ref}P_{ref}$ , as there can be only one plasmid per cell):

$$\begin{aligned}\frac{dT_{\gamma}P_{\alpha}}{dt} &= [\psi(1 - c_p\alpha)(1 - c_qq_e)(1 + b_{tot}) + \gamma_e R] \times T_{\gamma}P_{\alpha} \times (1 - \frac{n_{tot}}{K}) \\ \frac{dT_{ref}P_{ref}}{dt} &= [\psi(1 - c_p\alpha_r)(1 - c_qq_{ref})(1 + b_{tot}) + \gamma_r R] \times T_{ref}P_{ref} \times (1 - \frac{n_{tot}}{K}) \\ \frac{dR}{dt} &= [\psi(1 + b_{tot}) - \gamma_e T_{\gamma}P_{\alpha} - \gamma_r T_{ref}P_{ref}] \times R \times (1 - \frac{n_{tot}}{K}) \\ n_{tot} &= T_{\gamma}P_{\alpha} + T_{ref}P_{ref} + R \\ b_{tot} &= b_p(\alpha T_{\gamma}P_{\alpha} + \alpha_{ref} T_{ref}P_{ref})\end{aligned}$$

## 4.5.2 Data analysis

### Estimations of selection

Selection acting on a given trait is globally estimated by measuring changes in allele frequencies over time. Global frequencies at the metapopulation level are measured by pooling equal volumes of populations, effectively taking into account differential growth among populations. Mean within populations frequencies are computed as the mean of each individual frequency within populations, excluding the effect of differential growth among populations.

### Computing relatedness

Relatedness is defined as the measure of how an individual's social environment covaries with the individual's genotype, which can be calculated as the linear regression coefficient connecting an individual's genotype with the genotype of its interactants [Damore and Gore, 2012]. Here, we focus on the relatedness of  $P_{\alpha}$ , considering that the social environment of an individual corresponds to the population it belongs to. With  $p_i$  and  $n_i$  being respectively the proportion of producers and number of bacteria within population  $i$ , and  $p_{tot}$  and  $n_{tot}$  being respectively the proportion of producers and number of bacteria in the whole metapopulation, the relatedness of producers  $r_P$  at the level of the metapopulation can be calculated as followed (assuming populations are of the same size, which is the case at  $t_1$ ):

$$r_P = \frac{\sum_i \frac{p_i}{n_i} \cdot \frac{p_i}{p_{tot}} - \frac{p_{tot}}{n_{tot}}}{1 - \frac{p_{tot}}{n_{tot}}}$$

## General discussion



## Chapter 5

# General discussion

## 5.1 Conclusions from models and experiments

### 5.1.1 An interplay between transfer and public good production

My work highlights different aspects of the complex interplay between the selective pressures acting on plasmid transfer and public good production. First, horizontal transfer generally promotes the maintenance of public good production, even when both production and non-production alleles can be transferred (Chapter 2). Secondly, the transfer of public good production alleles can promote the selection of chromosomal genes that enhance transfer (Chapter 3). Finally, selective pressures acting on both transfer and public good production can act in synergy, leading to a change in the optimal transfer rate for plasmids when they carry public good genes, and to co-selection of chromosomal genes enhancing transfer and of plasmids carrying public good genes (Chapter 4).

Altruistic behaviours (ones with only an indirect fitness benefit) are well-known to require positive relatedness or assortment in order to be selected in bacterial populations [Griffin et al., 2004, Diggle et al., 2007b, Ackermann et al., 2008, Chuang et al., 2009]. In this work, we first experimentally demonstrate that in a metapopulation context horizontal transfer enhances relatedness at transferred loci because it amplifies allele frequencies at a local scale, and thus promotes the selection of public good production alleles. This effect is maximized when plasmid epidemic spread takes place, because such spread lowers the amount of initial variation needed in order to provide sufficient assortment. We next show that the same social framework can be applied to transfer: transfer ability can be understood as a social, altruistic behaviour influenced by chromosomal genes, and promoted in structured populations. Moreover, when variation in genotype frequencies is present for both public good production and transfer alleles, both act in synergy on among-population selection, as the most successful populations are those where transfer and public good production are associated. As a result, conjugative plasmids with public-good production alleles can then be favored and alleles favoring transfer residing on the chromosome are co-selected with public good production alleles.



### 5.1.2 Second-order selection for transfer ability

Our analysis has revealed that transfer ability can be promoted by different kinds of second-order selection [Arber, 2000, Tenaillon et al., 2001]. For example, selection on donor ability acts through the effect of transferred genes on recipient cells, and not through a direct effect of transfer on the cell performing transfer. In this case, we exclude direct benefits of transfer, obvious for plasmids themselves and potentially also present for the host, such as the induction of biofilm formation by conjugative pili (see section III.1.2). Because selection acts through neighbouring cells, it can be seen as both second-order selection and social selection, and thus needs to be understood in a social context.

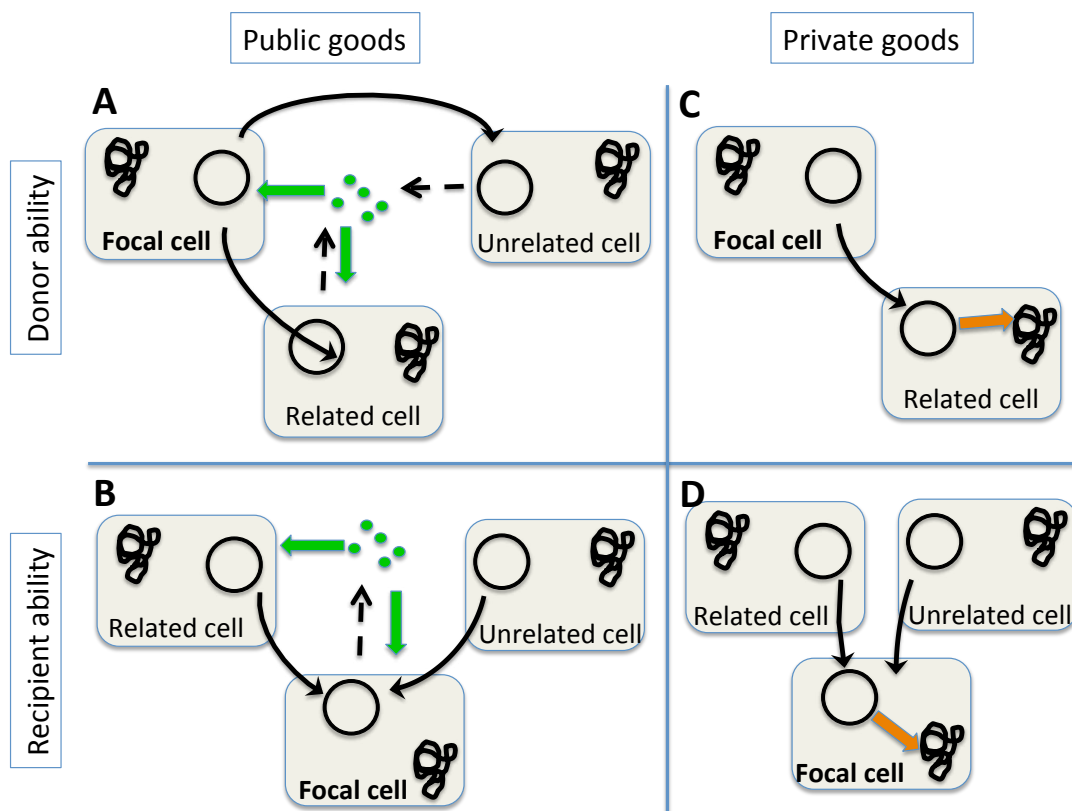


Figure 5.1: **Indirect selective pressures favoring chromosomal transfer ability.** Transfer is indicated as black arrows from the donor to the recipient cell, and can occur between related or unrelated cells. In the case of plasmids with public good production genes (**A** and **B**), transfer brings about public good production (dashed black arrows). Public goods (green dots) then benefit the focal cell but also its related cells (green arrows). In the case of plasmids with private good genes (**C** and **D**), transfer to kin leads to a direct increase in fitness for related cells (orange arrows). Effects with consequences on donor ability (**A** and **C**) and recipient ability (**B** and **D**) are shown.

The main and initial focus of this work has been on public-good production genes. In the case of public-good gene transfer, the transfer acts indirectly through a change in the environment shared by both donor and recipient cells, as well as their relatives (Figure

5.1-A and B). Increasing public good production can be selected when populations are structured, and related cells generally interact together. However, an unexpected result from the direct fitness analysis (Chapter 3) was that selection can also promote the transfer of private traits, that increase the direct fitness of recipients (Figure 5.1-C and D). This is still second-order selection for donor ability because the fitness of the donor cell does not directly increase through transfer. However, selection does not happen through changes in the common environment of donors and recipients, but through increases in the direct fitness of related cells. Recipient ability can be directly selected, when acquired genes directly benefit fitness.

### **5.1.3 Co-selection of altruistic alleles**

A final insight of this work (Chapter 4) is that not only do transfer and public good production increase fitness in synergy, but this synergy also promotes the linkage between both traits. In conditions where the association benefits fitness (when plasmid transfer promotes the selection of public good production), the association of transfer and public good alleles is selected indirectly via the success of populations where both traits coexist.

Plasmids that bear both public good production and transfer genes can outcompete others (such as null plasmids, plasmids with only public good production or plasmids with only transfer ability), because and when transfer amplifies the benefits from public good production. Co-selection can easily happen between plasmid genes controlling transfer and public good production genes, as the two genes are present on the same genetic unit, and are transferred together. On the other hand, chromosomal alleles promoting public-good production could be separated from public-good alleles precisely because of horizontal transfer. However, we show here that linkage can arise between plasmid public good production alleles and chromosomal alleles promoting transfer, because of the benefits of the association. More generally, altruistic alleles coding for transfer and public good production are co-selected on both plasmids and chromosomes.

## **5.2 Benefits and drawbacks of the approaches used**

### **5.2.1 A simplified synthetic and model system**

Our experimental results are all based on a synthetic system integrating public good production and plasmid transfer. Simulations themselves are tightly coupled to the experimental work, use the same design, incorporate experimentally measured parameters, and

generally mimic the experimental system. The synthetic approach leads to both benefits and drawbacks that can be generally associated with synthetic systems. In addition to this, I used in Chapter 3 a general fitness approach in order to get insight about the selective pressures acting on chromosomal genes and generate predictions.

### **Benefits of the synthetic approach**

The main benefit of the synthetic system is that it allows us to test precise hypotheses: that horizontal transfer can promote alleles coding for public good production and reversely, transferring those alleles can promote the selection of transfer. We manipulate independently transfer and public good production, and are thus able to prove that observed fitness measurements are linked to the transfer of specific alleles. For instance, selection of donor ability can be attributed in our experiments to the transfer of a public good gene in a structured population, and not to direct benefits of expressing conjugative pili. Results obtained with the synthetic system have allowed us to confirm the main predictions generated by our models, showing that we have not neglected to model some crucial aspect of real biological systems.

Moreover, we can reduce interactions between transfer, public goods and other cell traits that would be difficult to avoid otherwise. As mentioned in the introduction, many such unwanted interactions could take place in natural systems. Natural cooperative or transfer abilities are strongly regulated, in ways that could interfere with each others, or mask an actual interaction in laboratory conditions. Moreover, they could interact in complex ways precisely because of previous coevolution, including the co-selection described in our models. Such history could hide initial selective pressures that would have led to co-selection itself. These interactions would then be highly interesting to study as such in natural systems, but represent a further step in the analysis.

Finally, fluorescent markers have enabled us to obtain high precision fitness and transfer measurements, without resorting to plating that usually leads to huge variation between replicates. The combination of strain and plasmid markers also allows us to separate the effects of transfer and selection on fitness.

### **Generality of experimental findings**

The major limitation of a synthetic system is linked to its very goal: it can lead to a proof of concept, but cannot prove that the described phenomena are prevalent, or even take place, in actual organisms and real ecosystems. The key insight compared to modelling studies is that these phenomena can take place within real organisms and

environments, and with at least realistic parameters concerning transfer, benefits and costs, and population structure. Synthetic systems can generally be seen as intermediates between models and natural biological systems [Momeni et al., 2011, Tanouchi et al., 2012].

The full range of parameters values in nature remains unknown. As we were looking for a "proof of principle", we chose experimental conditions for which we expected to observe selection. For instance, one conclusion of experiments is that described effects can happen with "realistic" transfer rates. The F plasmid, that mobilizes plasmids at such efficient rates in our experiments, is actually a derepressed mutant selected in the laboratory [Cheah and Skurray, 1986], and was not found as a natural isolate. Still, natural isolates with such high plasmid transfer rates exist [Dionisio et al., 2002], so our synthetic system could correspond quite well to some natural ecosystems and help explain the existence of such strains. Some of the features of this system are actually likely to be less favorable to both transfer and public good production than the natural ones. For example, the costs of transfer and secretion are probably stronger than in nature, due to the absence of natural regulations decreasing the costs of public good production [Kümmerli and Brown, 2010]. Finally, our experiments can only detect changes in frequencies that are greater than a few percent, but natural selection could act on much weaker effects.

The public good response system used in this system is also quite peculiar, as it acts on growth rate and not on growth yield: in the absence of public goods, cells still reach maximal density after some time. This mechanism leads to constraints in terms of experimental setup, and to unstable benefits of cooperation. When we modelled the growth yield instead of growth rate effect, we saw that the maintenance of cooperation was more robust (data not shown). This is also supported by the work in another synthetic system, based on swarming in *Pseudomonas aeruginosa* [de Vargas Roditi et al., 2013]. Swarming was shown to expand the population carrying capacity, leading to more robust cooperation than in the case of transient benefits. Many cooperative behaviours, providing the ability to grow on new substrates, should lead to increased carrying capacity by opening a new niche for bacteria (see section I.2.3). Overall, although our setup may not cover all potential natural scenarios, it provides a convincing example, and one under less favorable conditions.

## Experimental limitations of the synthetic approach

Another drawback of synthetic approaches is that constructed systems are rarely stable, as their parts have not coevolved in interaction. In our case, the system required numerous adjustments in order to behave as expected. For instance, we found that different parts of the system were not all functional in the same range of temperatures. This is hardly surprising, as fluorescence proteins arise from either jellyfish or corals living in oceans [Ormö et al., 1996, Matz et al., 1999], and F plasmid is an *Escherichia coli* plasmid [Tatum and Lederberg, 1947], adapted to the gut environment of mammals. We also observed "sick" phenotypes (described in Chapter 1) we suspect are linked to perturbations of the initiation of transfer, that inadvertently happened when we mutated F plasmid in order to get mobilization of small plasmids, but not self-transfer.

Some of the results presented with simulations have not been reproduced experimentally, especially in Chapter 4. In some cases, it is a matter of time and interest: the effect of plasmid initial frequency for instance could be tested but would be labor intensive without providing groundbreaking results. In other cases however, the experimental system is not robust enough to test the modelling predictions. For instance, I did try to test the conclusions of Chapter 4 concerning plasmid gene combinations that maximize fitness. However, we observed a strong non-additivity on costs and benefits of public good production and transfer. Recipients were also poorly induced in the presence of chloramphenicol, which modifies the fitness of both strains and plasmids. The analysis of frequency changes allowed us to understand which factors were acting against the simplifying assumptions of models. With a synthetic system, this insight is however not very interesting: the poor induction of recipients, that we suspect arises from the cost of fluorescent markers on a middle-copy plasmid, is likely not related to any general constraint acting on transfer or public goods in real systems.

Still, these drawbacks did not prevent us from testing our main hypotheses. The experimental design moreover gives us internal controls to check if strains are behaving as expected. For instance, we can verify that benefits are due to among-population dynamics, and are not a direct benefit from some unsuspected modification of cells, confirming the validity of the synthetic approach.

## Generality of the modelling approaches

I have used in this work two kinds of models: simulations of cell growth and transfer mimicking the experimental setup (Chapters 2, 3 and 4) and a more general approach to

model fitness (Chapter 3).

Simulations have been designed in order to mimic as much as possible the experiments, and use experimentally measured parameters. Their conclusions are validated by their convergence with experiments when both were done in parallel. Their major contribution was to allow more flexibility and power in the number and kinds of hypotheses we wanted to explore. I could expand the parameter range (mainly concerning transfer and public good production, that are constitutive in experiments), and investigate population structures not tested in experiments.

I also used in Chapter 3 a general fitness approach, modelling more generally the indirect changes in host fitness brought by gene transfer in a structured population. The equations and predictions derived with this approach are based on numerous simplifying assumptions, like the linearity of public good fitness effects, the independence between the different traits modelled, and small effect mutations. Still, the key role of this approach was to provide general insights about the selective pressures acting on chromosomal genes concerning plasmid transfer, clarifying the distinction between different kinds of indirect effects. These insights have then been further confirmed or modified with simulations and experiments.

### **5.2.2 Extreme population structures**

Except for the direct fitness approach that does not explicitly describe the structure of populations, I mostly analyzed two kinds of population structure: (1) a simple metapopulation with two populations, and (2) populations subject to strong dilution. The first approach is sufficient to separate within- and among-population dynamics, and is amenable to experimental tests. The second has the advantage of not enforcing arbitrarily the population structure, particularly when a quite specific structure seems to be required for selection to occur.

Both setups are similar to the one presented in [Chuang et al., 2009], and capture the process of bacteria growing in well-mixed populations, that are completely separated from each other. Encounters between individuals in different populations occur when all populations are pooled, mimicking global competition. This setup ensures that both public-good production and plasmid transfer happen at a local scale, but competition happens at a global scale. If cooperation and competition occur at the same scale, benefits of cooperation would be compensated by increased competition between relatives [West et al., 2002]. Competition happens at a larger scale than cooperation when populations

can increase in size because of the benefits of cooperation (elastic populations). Mixing takes place at a time when cells benefiting from cooperation or transfer outnumber others: this should therefore be likely to occur by chance when cooperation stably increases growth yield of natural populations.

Horizontal transfer by conjugation is assumed to take place within populations, which will hold also in natural environments: conjugation happens between neighbouring cells. Its effect on public good cooperation will again depend on the scale of competition [Giraud and Shykoff, 2011], but no more than any public good cooperation in general. The scale of competition will generally depend on the life cycle, and is not specific to the effect of transfer [Rankin et al., 2011]. Finally, we have not considered any migration between populations. Still, the effect of transfer on relatedness was shown to hold even in the presence of migration in the population genetics model developed in [Nogueira et al., 2009]: some amount of migration decreases relatedness, but for a wide range of parameters, transfer still increases relatedness. Migration thus should not affect strongly our conclusions, if not present at rates so high that the effect of transfer would be completely suppressed.

The experimental and simulation setup we use to model population structure may thus be specific or even extreme but they allow us to understand the effect and selection of transfer, and are supported by previous work.

## 5.3 Possible selection of horizontal transfer in nature

The major conclusion of modelling work is that interactions between transfer and public good production (both the effect of transfer on cooperation and the subsequent selection of transfer) are most likely to take place when (1) horizontal transfer is efficient at short spatial scales and (2) there is an abundance of plasmid-free cells. Here I consider both conditions in the context of natural systems. In this section I also discuss a case not addressed in my work: that transfer could be selected in the context of antagonistic, and not only cooperative, interactions.

### 5.3.1 Infectious transmission at short spatial scales

The effect of horizontal transfer on the maintenance of cooperative behaviours fundamentally relies on infectious transmission at short spatial scales: infectious transmission spreads alleles at a local scale, and the assortment in turn increases at a higher scale. Maintenance of cooperative behaviours through horizontal transfer will thus be maxi-

mized when infectious transmission is strong and efficient, leading to rapid invasion of local patches by transferred alleles.

Here we focused on the case of bacterial conjugation, that naturally happens at a local scale, between neighbouring cells. Moreover conjugative plasmids often have the ability to undergo epidemic spread with secondary transfer from recipients, leading to efficient colonization within patches. Other mobile elements that transfer horizontally by conjugation, such as integrative conjugative elements, should behave in the same way as plasmids in this respect. Bacteriophages also move horizontally in a very efficient way. Their transmission is not so strongly constrained spatially as plasmid transmission, since phages can exist extracellularly. Still, they are likely to re-infect cells close to the ones that they originated from. The spread of phages by lysis makes them into virulent parasites, likely to modify the selective pressures acting on hosts for horizontal transfer. But some bacteriophages, such as the filamentous phages involved in *Vibrio cholerae* virulence, spread by secretion from the cell without lysis [Waldor and Mekalanos, 1986], and are thus likely to confer similar costs to the host as the plasmids.

The spread of mobile genetic elements can be seen as a very specific bacterial way of producing assortment. This mechanism relies only on active limited dispersal (without kin discrimination). Moreover, dispersal is gene-specific, so the population-level assortment increases specifically for the genes located on mobile elements. The gene-specific increase in assortment seems to be specific to bacteria: mobile elements also occur in Eukaryotes, but they spread through sexual reproduction [Burt and Trivers, 2006]. Sex also allows them to invade host populations [Goddard et al., 2001], but without the same epidemic dynamics that leads to faster spread than with vertical transmission.

A parallel with cultural evolution (in humans) can be made here: cultural transmission of a given behaviour (including altruistic behaviours) was indeed shown to increase assortment, as it increases variation among groups when it happens within groups [Boyd and Richerson, 1982]. Initial models of the effect of horizontal gene transfer on relatedness were actually based on the cultural transmission formalism [Nogueira et al., 2009]. Horizontal gene transfer is analogous here to unbiased cultural transmission, happening with a rate independent from the gene content and fitness effects of transferred elements. Both processes share a speed of transmission that is higher than the speed of respectively vertical and genetic transmission. Both processes also have a high specificity: only a subset of genes or behaviours are transmitted horizontally.

The selection of chromosomal genes increasing transfer ability additionally requires



assortment at the chromosomal scale, in order for the benefits of gene transmission to feedback on the individuals transmitting genes. The assortment for chromosomal alleles will not be increased through horizontal transfer, but can arise through classical mechanisms, like limited dispersal. When transferred genes code for private goods, selection happens for transfer between kin, a process again favored by limited cell dispersal. Other factors could also favor effective transfer to kin: related cells could share high recipient ability if donor and recipient ability have co-evolved. For instance, they could express the same restriction-modification systems that target foreign DNA, which would then be less likely to be transmitted between unrelated cells. The transfer of a given gene could also simply not bring fitness benefits to a phylogenetically distant competitor: for instance, an antibiotic resistance gene was not expressed after interspecies transfer [Guiney et al., 1984]. In this way, competitor species would not benefit from transfer, even in the absence of discrimination in the range of transfer.

Overall, different mechanisms in nature could enable infectious transmission of the type we modelled and analyzed experimentally in bacteria, ensuring both the generality and relevance of our results.

### **5.3.2 Occurrence of plasmid-free cells**

The availability of plasmid-free cells is key to the possibility of transfer itself, and so will control both the selection for transfer and occasions where transfer could promote cooperation. The occurrence of plasmid-free cells in nature is not well known. However, despite the lack of quantitative data, the very existence of plasmid transfer mechanisms demonstrates that plasmid-free cells occur frequently enough so that transfer can be selected. Conjugative plasmids thus likely often experience conditions where plasmid spread takes place: it is exactly in these conditions that they could also promote the maintenance of cooperative behaviours. The co-selection of transfer and public-good genes on mobile elements should thus generally be favored, as infectious transfer would anyway favor both.

Explaining sustained selection for chromosomal transfer ability genes is more difficult: it implies that gene movement benefits the host, but simple long-term gene maintenance does not. If maintenance were sufficient, vertical transmission would suffice and there would be no selection for transfer. Instead we expect that transfer will be selected via recurrent loss of transferred genes. However the question remains: under which conditions would horizontal transfer coupled to recurrent loss be more favorable than vertical

transmission? One major possibility is that the genes of interest are uniformly costly but rarely beneficial. If transfer is not too costly, the cost of gene maintenance could outweigh the occasional cost of transfer. Interestingly, this possibility corresponds to genes that are well-known to be located on plasmids (see section III.2.1): genes beneficial in transient or local conditions, such as antibiotic resistance genes. The presence of these genes on plasmids has previously been interpreted in two ways. First, localization on mobile elements would benefit accessory genes themselves, allowing them to persist in the absence of sustained selection [Bergstrom et al., 2000]. Secondly, the presence of a "communal gene pool" would allow prokaryotes to have access to a diversity of adaptive traits, without the burden of maintaining them permanently [Norman et al., 2009]. Here we propose a similar explanation, but without resorting to "for the good of the species"-like arguments. Selection could act at the scale of one genotype, and promote mechanisms that lead to repeated transfer and loss of accessory genes in structured populations where benefits of transfer feedback on the transmitter genotypes.

Overall, repeated loss and horizontal transfer could explain the variability of plasmids and gene content across isolates or environments, and reversely transfer would be promoted by the repeated occurrence of high amounts of plasmid-free cells.

### **5.3.3 Selection of interspecies horizontal transfer**

Interspecies plasmid transfer is one particular case we did not address in this work. We have always assumed that the host backgrounds differ only in plasmid transfer ability, but not in plasmid effects on fitness or public-good production or benefits.

Transfer between different species or ecotypes could be selected in ways different from those we described so far, depending on species properties and on their ecological interactions. (I use here species in a very broad sense, meaning any bacterial strains or ecotypes that differ more strongly than only in the trait of interest we concentrated on for now, donor or recipient ability. This could involve other traits involved in plasmid-host co-adaptation.)

#### **Interspecies transfer and bacterial competition**

The level of competition between two strains or species will depend on the degree of overlap between their ecological niches. If they share the same niche, competition will be strong: each species will benefit from a decrease in fitness from the other. Indeed, when modelling interspecies transfer we observed that cooperation was favored in the absence

of niche overlap: one species could benefit from transferring a public good plasmid to the other, and the second simultaneously from receiving it, because of increased global public good production. When donor and recipient species were sharing the same resources, the benefit gained by one species affected negatively the fitness of the other (data not shown). Cooperation between species mediated by plasmid transfer should thus be maximized when species occupy different ecological niches. This would for instance be the case for bacteria consuming different nutrients that are affected in the same way by an antibiotic, and thus both benefit from antibiotic degradation.

If two species are competing for resources, plasmids could on the contrary be used as biological weapons by the donor species: this would be a case of apparent competition [Holt and Lawton, 1994, Dionisio, 2007]. This dynamics has been demonstrated in the case of other mobile elements, bacteriophages [Bossi et al., 2003, Brown et al., 2006, Duerkop et al., 2012, Gama et al., 2013], but not yet for plasmids. In the case of plasmids, the interaction may not be as clearly competitive (lytic phages have a drastic effect on fitness by rapidly killing cells), but could involve some intermediate scenario between interference competition and active manipulation of recipients in order to produce a public good.

### **Interspecies transfer and manipulation**

Plasmid transfer could be a way for donor cells to manipulate recipient's behaviour. In this case, recipients are themselves the resource that can be "exploited" by plasmids. The most striking example of manipulation is *Agrobacterium tumefaciens* : bacteria transfer the T-DNA (not a plasmid, but a part of one) to plant cells, where the T-DNA is integrated into the chromosome and manipulates its gene expression in order to produce metabolites used by bacteria [Otten et al., 1992]. In this extreme case, the recipient cells are not competitors, but belong to the host of a parasitic bacterium. Other scenarios could be more similar to competition, with recipients that are simultaneously competitors and exploited in order to produce public goods. Such situation could evolve towards or from a cooperative scenario. For instance, selection for transfer could first act within species, but the mutations leading to increased public good production in unrelated recipients could then be favored in a plasmid coevolving with a given donor strain.

In the two cases of interspecies or intraspecies transfer, a further challenge is to understand why and how plasmids and host bacteria could evolve towards mutualism, despite the very existence of horizontal plasmid transfer favoring competition.

## 5.4 Coevolution between plasmids and chromosomes

In this work, I described the existence of mutualistic interactions between plasmids and their hosts, that involve plasmid horizontal transfer. However, horizontal transfer is usually seen as a factor that acts against the establishment of mutualism, as it decouples symbiont and host interests [Herre et al., 1999] (see section III.1.2). In this section, I discuss how plasmids could evolve towards mutualism while being horizontally mobile elements.

### 5.4.1 An alignment of interests robust to horizontal transfer

We observe mutualism between plasmids and their hosts despite horizontal transfer. As chromosomal genes can benefit from transfer, an alignment of interests occurs between chromosomal and transferred genes, because of the modifications caused by transferred genes in recipients: both plasmids and chromosomal genes benefit from transfer. This leads to linkage between chromosomal genes promoting transfer and transferred genes. At a more global scale, horizontal transmission thus actually leads to a pattern similar to vertical transmission: beneficial plasmids stay associated with their hosts. Paradoxically, the benefits of symbiosis are here brought by horizontal transmission itself, and not by the mere presence of plasmids within their hosts.

In the case of horizontal transfer among distant species, one reason for alignment of interests could be different: plasmids adapted mainly to the donor host would not benefit from adaptation to the recipient host in the absence of efficient transmission from this second host (I discuss here the case described above in section 5.3.3, where plasmid transfer would be a competitive trait of the donor species; not the case of "true" broad-host-range plasmids, that do not seem to specialize towards one specific host). From the plasmid side, transfer can then be seen purely as an extreme specialization to the primary host, as the secondary host would be an evolutionary dead-end for the transferred plasmid. From the chromosomal side, transfer would be adaptive only to the donor species, and not to the recipient. This can be compared to the case of quorum sensing, where quorum sensing between species will probably often not be true signalling, but coercion of one species by another [Diggle et al., 2007a].

In both cases of intraspecies and interspecies transfer, both plasmids and chromosomes benefit from plasmid localization of transferred genes. Public good genes localization on plasmids can give rise to higher selection for public good production; private good genes

localization on plasmids can benefit the host as well if transfer occurs within kin. Benefits of transfer thus feedback on both plasmid and chromosomal partners of the symbiosis, despite the occurrence of horizontal transfer.

### 5.4.2 Coadaptation of plasmids and chromosomes

We can infer from the described alignment of interests that coadaptation between plasmids and chromosomes will be favored. Some known examples of regulations could then be interpreted as coadaptation. Particularly, the complex regulation of plasmid transfer involving both plasmid and chromosomal genes [Frost and Koraimann, 2010] (see section II.3.1), may be a coregulated trait, and not just a product of antagonistic coevolution (with the evolution of both chromosomal defenses against plasmid transfer, and plasmid exploitation of chromosomal signalling pathways as cues for transfer).

This coregulation of transfer seems to be extreme in the case of *Enterococcus* pheromone plasmids, where transfer is induced by chromosomally encoded pheromones expressed by plasmid-free cells [Dunny, 2007] (see section II.3.2). Other mechanisms controlling plasmid transfer in this system are unknown, but could exist and induce transfer in conditions where plasmid genes benefit the host. Moreover, sex pheromone plasmids are known only in *Enterococcus* species: such coregulation seems to imply a strong specialization of plasmids towards a specific host, and would then concern mainly narrow-host-range plasmids.

Finally, coadaptation can be considered in relation to the enforcement hypothesis for the horizontal transfer effect on cooperation [Smith, 2001]. Indeed, it was argued and confirmed in our experiments that this scenario is not stable *a priori* against "cheater plasmids" that transfer at the same rate as public-good plasmids. However, the possible coadaptation and linkage between public-good plasmids and transferring strains suggests that public-good plasmids would be favored in the long-term, therefore public-good alleles would be actually more transferred than cheater alleles.

### 5.4.3 The effect of plasmid competition

We generally excluded the possibility of direct competition between plasmids in our analyses, particularly the case of within-cell plasmid competition: we assumed in simulations and ensured experimentally that only one plasmid type is present in a given cell. In nature, repeated bottlenecks during cell division and conjugation [Paulsson, 2002], and entry exclusion between related conjugative plasmids [Garcillán-Barcia and de la Cruz, 2008] will limit within-cell competition. Still, competition between plasmids sharing replication

mechanisms will occur, leading to incompatibility. Within-host competition is thought to favor increased virulence of pathogens [Frank, 1996]. In the case of plasmids, multiple infection can for instance select for post-segregational killing [Cooper and Heinemann, 2005], where cells that lose one plasmid at division are killed. Superinfection can also evolve, leading to strong costs to the host [Smith, 2011].

We briefly studied in Chapter 4 what happens when a plasmid competes against a plasmid with a different transfer rate. Transfer of the competitor plasmid modifies the optimal transfer rate, as the first plasmid also needs to transfer at a high rate to compensate for its competitor invasion. Moreover, transfer also affects among-populations patterns of variation, influencing the selection of public-good production. Competition between plasmids should be modelled more explicitly in order to study its consequences in a more detailed way. We can only hypothesize that plasmid competition would generally hinder the evolution of mutualism, as it is deleterious to the host. But mutualism could also be favored by the mechanisms we describe in our work: for instance, the linkage that arises between plasmids and strains favoring transfer could favor mutualist plasmids in competition with deleterious ones.

#### **5.4.4 Plasmids: symbionts integrated to the bacterial genome**

A more general issue related to this study is the exact status of plasmids themselves. Plasmids can be considered as symbionts, evolving on a parasitism-mutualism continuum with a complex role of horizontal transmission in this evolution [Harrison and Brockhurst, 2012]. But they simultaneously are extremely strongly connected to their hosts: they are part of the bacterial genome, share the cytoplasm with chromosomes, and at a longer timescale exchange genes with and between bacterial chromosomes.

Plasmids are also considered as a tool for bacterial sex. The occurrence of chromosomal recombination, linked to plasmid integration into the chromosome, speeds bacterial adaptation [Cooper, 2007]. In the scenarios we describe in this study, genetic exchange is not merely a side-effect of plasmid random integration into the chromosome, but occurs with every event of plasmid transfer, for genes with specific properties that are permanently present on plasmids. Some definitions of sex insist on the fact that it should concern processes "selected by the benefits of genetic exchange", and not every mechanism including genetic exchange as a side-effect [Redfield, 2001]. Following this definition, the selection of horizontal transfer we demonstrate in our work would precisely be a case of selection for sex concerning specific genes: selective pressures on genes modulating transfer are a

function of which degree of mobility for transferred genes most enhances host fitness.

Plasmids seem to be fundamentally at the limit between a symbiont and an integral part of the genome, and potentially evolve to globally benefit the genome while (and even thanks to) maintaining horizontal transmission. Because of the presence of horizontal gene transfer and its influence on cooperation, we also reach the limit of the kin selection vocabulary. Indeed, kin selection has been initially defined as the process by which traits are favoured "because of their effects on the survival and fertility of relatives" [Maynard Smith, 1964]. It was then shown that what really matters for kin selection is relatedness in the sense of genetic correlations among loci [Frank, 1998]. In microorganisms, genetic correlations can precisely arise because of horizontal gene transfer and the concept of kin is blurred: kin has to be defined at the gene level, because of the existence of mobile genetic elements that can be kin even if the rest of the genome is not. Selection in general, but also kin selection, both act simultaneously on different genetic units with diverging or converging interests.

## 5.5 Perspectives

### 5.5.1 Other approaches to the transfer-cooperation interplay

#### **Insight on the prevalence of the transfer-cooperation interplay in nature**

To extend the experimental and modelling proof of concept, a very interesting approach would be to analyze real biological systems where public good production is known to be coupled to horizontal transfer. Efficient transfer is known to occur for a few systems where transferred genes are proven or likely to be public goods, like the cholera toxin [Waldor and Mekalanos, 1986], the endotoxin in *Bacillus thuringiensis* [González et al., 1982], the T-DNA in *Agrobacterium tumefaciens* [Platt et al., 2012] or plasmid  $\beta$ -lactamases [Barlow and Hall, 2002]. In some of these systems, the ecology has been studied and it would be possible to simultaneously investigate both the real benefits and costs of public good production, and the transfer rates and their consequences. For instance, the endotoxin production in *Bacillus thuringiensis* has been shown to be a cooperative behaviour and studied in natural conditions in the field: plasmid presence can be identified, and relatedness assessed in natural settings [Raymond et al., 2012], and could be an promising model system [Raymond and Bonsall, 2013]. Other properties of natural systems, like the multiple mechanisms regulating both transfer and public good production, could also be taken into account. For instance, in *Agrobacterium tumefaciens*, both plasmid conjuga-

tive transfer and T-DNA transfer to the plant (here the cooperative behaviour) are under control of quorum-sensing systems [Platt et al., 2011, Cho et al., 2009].

Genomic and bioinformatics approaches can also bring information about the diversity and prevalence of plasmids bearing public good production genes in nature. The main result that motivated our research was the bioinformatic indication that secreted proteins are often located on plasmids [Nogueira et al., 2009]. Further genomic studies could give insight into the actual distribution of plasmids within host species: what proportion of cells in a given species bears plasmids? How does plasmid prevalence vary over time? These poorly known properties should have a significant impact on the importance of the interplay between plasmid transfer and public good production.

### **Other aspects of the transfer-cooperation interplay**

Another interesting approach would be to explicitly model real structured populations that are abundant in natural environments, like biofilms [Kolter and Greenberg, 2006]. Biofilms can present high rates of horizontal transfer [Molin and Tolker-Nielsen, 2003], but effective transfer will be modulated by many parameters of biofilm growth, like the possibility of contact between donors and recipients [Krol et al., 2011] (see section III.1.2). Moreover, plasmids themselves induce biofilm formation [Ghigo, 2001], and likely increase the spatial structuring of populations. Conjugation could thus physically enhance relatedness between neighbouring cells, in addition to the effect of gene transfer. Modelling these effects was outside the scope of this study, but they likely play a complex but important role in the interplay between plasmids and social behaviours.

In the same way, the coevolution between hosts and plasmids could be studied in a more detailed and realistic way. First, I have neglected the competition between plasmids, that is likely to move the plasmid-host interaction towards parasitism. Secondly, I have not analyzed simultaneously the selective pressures acting on hosts and on plasmids concerning transfer rates (as I have not modelled a dual control of transfer by plasmids and hosts, but each case separately). Finally, analyzing the effect of the costs and benefits of plasmid carried genes, as well as both transfer and plasmid maintenance properties, could help us to understand when a host would benefit more from plasmid continued maintenance or from repeated loss and transfer.



## 5.5.2 Applications of plasmid transfer?

Finally, if the selection of transfer and its role in the maintenance of cooperation are actually widespread in nature, it could be possible to interfere with it in conditions where transfer favors the virulence of pathogenic bacteria.

Better understanding of the effects of transfer and their mechanisms could be used to targeted cooperative pathogens. In the cases where virulence is a cooperative behaviour [Brown, 1999], it has been already suggested that targeting cooperation would decrease the evolution of resistance. Indeed, as cooperation is a global property of infections and beneficial only at the group level, it has been shown that evolution of resistance would be slower for cooperative behaviours compared to classical antibiotics that target individual cells [André and Godelle, 2005]. Further, introducing less virulent strains into natural pathogen populations could be a strategy to treat bacterial diseases, as cheats would have a competitive advantage and invade the population [Brown et al., 2009].

Similarly, when horizontal transfer is involved in the maintenance of public-good cooperation, targeting transfer could prevent cooperative virulence. Molecules targeting plasmid conjugation are known: for instance, inhibitors that target the relaxase of F plasmid have been shown to prevent conjugation [Lujan et al., 2007]. Since these inhibitors were also shown to selectively kill F-bearing cells, they could maybe target bacterial strains characterized by high transfer ability. Some of these molecules were already used in clinical applications for other applications, so they could be repurposed.

As for other cooperative behaviours, mutations in chromosomal genes promoting transfer could be counter-selected in the short-term. However, this is unlikely to be the case for plasmid resistance mutations, as plasmids directly benefit from transfer. Still, male-specific bacteriophages have been shown to select for the loss of antibiotic-resistance plasmids, and for the loss of conjugation in plasmids, even in the presence of antibiotic selective pressure [Ojala et al., 2013]. It thus seems possible to design treatments that would efficiently select against conjugation. Finally, the consequences of targeting transfer will likely depend on the dynamics of infection and the possibility of treatment before horizontal transfer happens. Potential pitfalls include horizontal transfer not happening during every infection, or occurring before an infection can be detected and targeted. Some of these applications are very long-term goals, but they can motivate the basic research of the type presented here.

# Annexes



## Appendix A

# Strains and plasmids

### A.0.3 The background strain

The background strain JC1191 [Chuang et al., 2009] has the genotype *att::rhl-catLVA(Sp<sup>R</sup>) ΔsdiA::FRT rrnB3 ΔlacZ4787 hsdR514 Δ(araBAD)567 Δ(rhaBAD)568 rph-1*. The *att::rhl-catLVA(Sp<sup>R</sup>)* segment contains both an Rhl auto-inducer-responsive promoter ( $P_{rhl*}$ ) driving an unstable version of *cat*, and the *rhlR* gene under the weak constitutive promoter  $P_{lacI}^q$ . This cassette is integrated into the chromosome with the spectinomycin resistance gene *Sp<sup>R</sup>*.

### A.0.4 Low-copy plasmids: production, transfer and fluorescence

#### Description

Low-copy plasmids are modified versions from pZS\*2R plasmids [Chuang et al., 2009] with pSC101\* replication origin, maintained with kanamycin resistance. They are constructed based on a combination of production/non-production, transfer/non-transfer, and fluorescence alleles (Figure A.1).

Production and fluorescence genes are in an artificial operon, separated by a Shine-Dalgarno translation signal, under control of the strong promoter  $P_R$ . The production allele expresses the enzyme RhlI, producing the Rhl auto-inducer C<sub>4</sub>-HSL [Chuang et al., 2009]. The non-production allele is a linker sequence only. Fluorescence alleles are GFP and YFP. The allele conferring transfer ability is the wild-type sequence of *F-oriT*.

#### Plasmid construction

Plasmids are derived from pZS\*2R-GFP,*rhlI* and pZS\*2R plasmids [Chuang et al., 2009]. YFP coding sequence [Elowitz, 2002] was amplified by PCR and cloned in pZS\*2R-GFP,*rhlI* (production plasmid from [Chuang et al., 2009]) between XhoI and HindIII sites, replacing GFP. RhlI sequence was removed from pZS\*2R-GFP,*rhlI* by digestion with HindIII and XbaI, and religation with the linker sequence AGCTTAATTAGCT-GAGTCTAG.

The wild-type sequence of *F-oriT* between coordinates 66002 to 66494 from GenBank

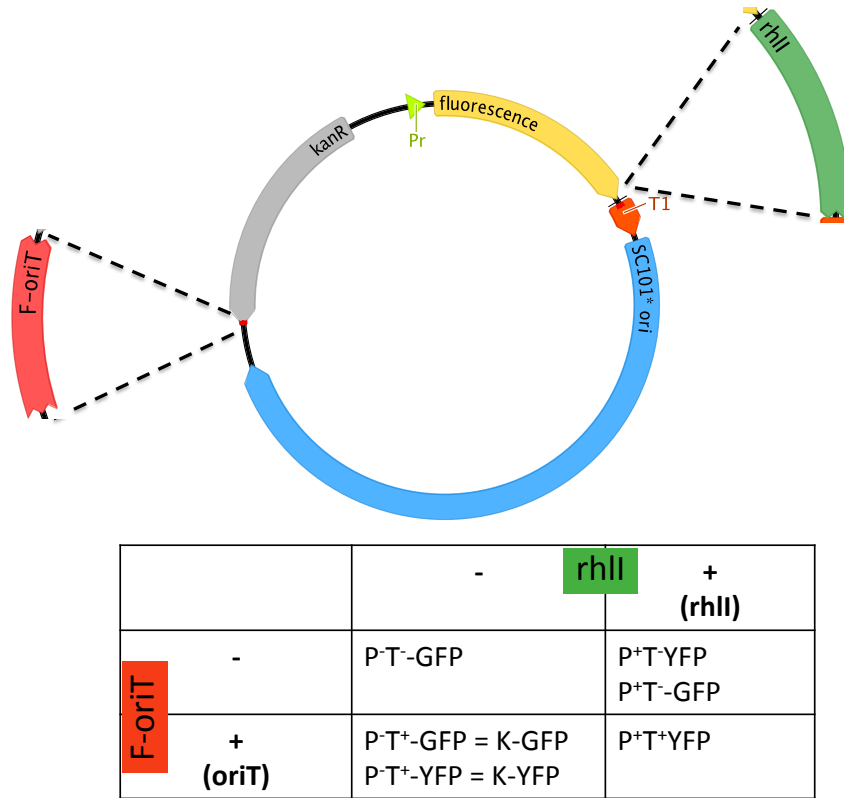


Figure A.1: **Map of pZS plasmids used in experiments.** All pZS plasmids bear the *aph* gene coding for kanamycin resistance (*kanR*, gray) and pSC101\* origin of replication (SC101\*ori, blue). T<sup>+</sup> plasmids bear *F-oriT* (red), T<sup>-</sup> plasmids do not. They bear in addition a fluorescence gene (GFP or YFP, yellow), and the *rhII* gene in the same operon in the case of P<sup>+</sup> plasmids. P<sup>-</sup> plasmids do not bear *rhII*. The table indicates the plasmid names, as used in the main text and figures, and their fluorescence.

NC\_002483 was amplified by PCR (using pOX38::Tc as template) and inserted in the common *SacI* site of P<sup>+</sup> and P<sup>-</sup> plasmids, in direct orientation.

## A.0.5 Strain fluorescence markers

### Description

Strains are identified with non-mobilizable red fluorescence markers. For experiments in Chapter 2 and section 3.3.2 of Chapter 3, recipients are identified by the *mRFP1* gene located on pSB3K3-RFP. pSB3K3-RFP is a medium-copy plasmid with p15A replication origin, carrying *mRFP1* gene under control of the strong promoter P<sub>*lacI*</sub>, and maintained with kanamycin resistance. For experiments in sections 3.3.1 and 3.3.3 of Chapter 3, one of the competitor background strains is JC1191<sub>red</sub> identified by the *tdCherry* gene under control of a p<sub>RNA1</sub> promoter, inserted in the attTn7 site of the chromosome. The other competitor has no red fluorescence.

## Construction

pSB3K3-RFP plasmid was obtained from the Registry of Standard Biological Parts ([www.partsregistry.org](http://www.partsregistry.org)). JC1191<sub>red</sub> was constructed by integration of the pRNA1-tdCherry gene construction on pNDL-32: pNDL-32 is an integration vector obtained from Nathan Lord (Paulsson laboratory), which carries a Tn7 integration construct for a very bright colE1 RNA1 promoter-tdCherry variant (<http://openwetware.org/wiki/Paulsson:Strains>). It is a temperature-sensitive, ampicillin-resistant vector, where pRNA1 promoter-tdCherry sequence is included between Tn7 insertion sites.

pNDL32 plasmid was transformed into JC1191 strain (containing  $F_{HR}$  plasmid) with selection on 100 $\mu$ g/mL ampicillin, then streaked twice at 30°C on LB-agar. Next, two colonies were streaked overnight on LB-agar at 42°C. Finally, plasmid loss was confirmed by checking that clones were ampicillin-resistant.

### A.0.6 F helper plasmids

#### Description

$F_{HE}$  is a mutant of pOX38::Tc [Anthony et al., 1994], where *oriT* contains 2 substitutions (A141T et C144G) that reduce binding and mobilization efficiency by the F relaxase compared to the wild-type F sequence (substitutions are numbered according to [Stern and Schildbach, 2001]). The other mutants that were tested are F-oriT::frt, F-oriT-R100 (C145A and C147T substitutions) and F-oriT-7 (A141T substitution).  $F_{HR}$  is derived from  $F_{HE}$  and bears a further deletion of the *traS* gene.

#### Construction

*oriT* point mutations were made by allelic exchange of pOX38::Tc [Anthony et al., 1994] with a modified *oriT* including point mutations made by PCR. Mutated *oriTs* were first cloned into the SacI site of pDS132 plasmid [Philippe et al., 2004], then integrated in pOX38::Tc by allelic exchange with the wild-type sequence [Philippe et al., 2004]. F-oriT::frt was obtained from the Registry of Standard Biological Parts ([www.partsregistry.org](http://www.partsregistry.org)).

$F_{HR}$  was obtained by deleting *traS* gene from  $F_{HE}$  using  $\lambda$ /red homologous recombination [Datsenko and Wanner, 2000]: the sequence between F coordinates 88274 and 88606 was replaced by the cat cassette from pKD3, which was then removed with pCP20 [Datsenko and Wanner, 2000].

Helper F plasmids were transferred by conjugation to all strains in which they were needed, as the plasmids retained a low conjugation ability.

## Appendix B

# Identification of strains, plasmids and transfer

### B.0.7 Measures of plasmid transfer rates

Instantaneous transfer rates were measured by measuring the transfer of antibiotic resistance markers (kanamycin resistance for mobilizable pZS plasmids, tetracycline resistance for F plasmids) to a standard recipient strain which is a spontaneous mutant resistant to nalidixic acid and rifampicin (MG Nal<sup>R</sup>Rif<sup>R</sup>).

Donors and recipients were separately grown to OD<sub>600</sub> = 0.6, then washed and mixed at 1/10 dilution after resuspension in LB. Conjugation was allowed to happen for 30 min at the temperature indicated, under well-mixed conditions. The mix was then diluted to appropriate concentrations and plated on LB-agar containing nalidixic acid and rifampicin (to select for recipients), spectinomycin (to select for donors) and nalidixic acid, rifampicin and either kanamycin or tetracycline to select for transconjugants. The concentrations used were 40µg/mL (nalidixic acid), 100µg/mL (rifampicin), 25µg/mL (spectinomycin), 100µg/mL (kanamycin) and 12.5µg/mL (tetracycline).

The conjugation rate  $\gamma$  was estimated as follows :  $\gamma = \frac{T}{D \times R}$  cell mL<sup>-1</sup>h<sup>-1</sup> where  $T$ ,  $D$  and  $R$  denote final densities of transconjugant, donor and recipient cells, respectively. Neglecting secondary transfer that would happen within 30 min, this measure of transfer follows the transfer rates defined and used in our simulations (where transfer follows a mass-action law depending on both donor and recipient cell concentrations), where in a short period of time,  $T = \gamma \times D \times R$ .

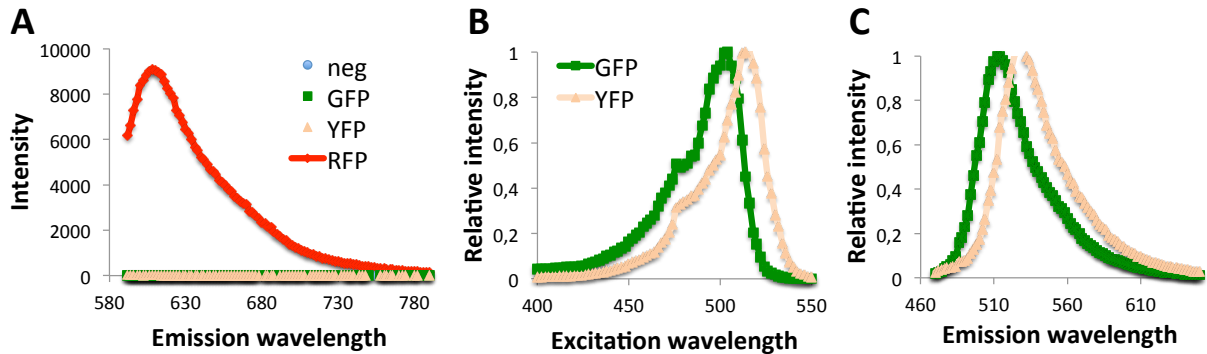
### B.0.8 Flow cytometry analysis of cell fluorescence

Cultures were analyzed for strain and plasmid proportions by flow cytometry. Flow cytometry technology measures optical parameters for thousands of individual particles per second, as particles are hydrodynamically focused and pass in front of a series of lasers and associated detectors. Each particle absorbs and scatters light, and may emit fluorescence if it contains fluorescent molecules excited by the wavelength of light emitted by lasers. Light is then splitted and filtered to detectors.



For flow cytometry analysis of plasmid and strain proportions, cultures were fixed in 1% PFA (Thermo Scientific) for 10 minutes, then resuspended in PBS (Life Technologies) and stored at 4°C. Data acquisition was performed on the Cochin Cytometry and Immunobiology Facility. For each sample, 50,000 or 100,000 cells (100,000 when competitors were present in frequencies  $< 1\%$ ) were analyzed using a BD LSR Fortessa cell analyzer (BD Biosciences) with 405 nm, 488 nm and 561 nm excitation lasers. Data analyses were performed using FlowJo (TreeStar).

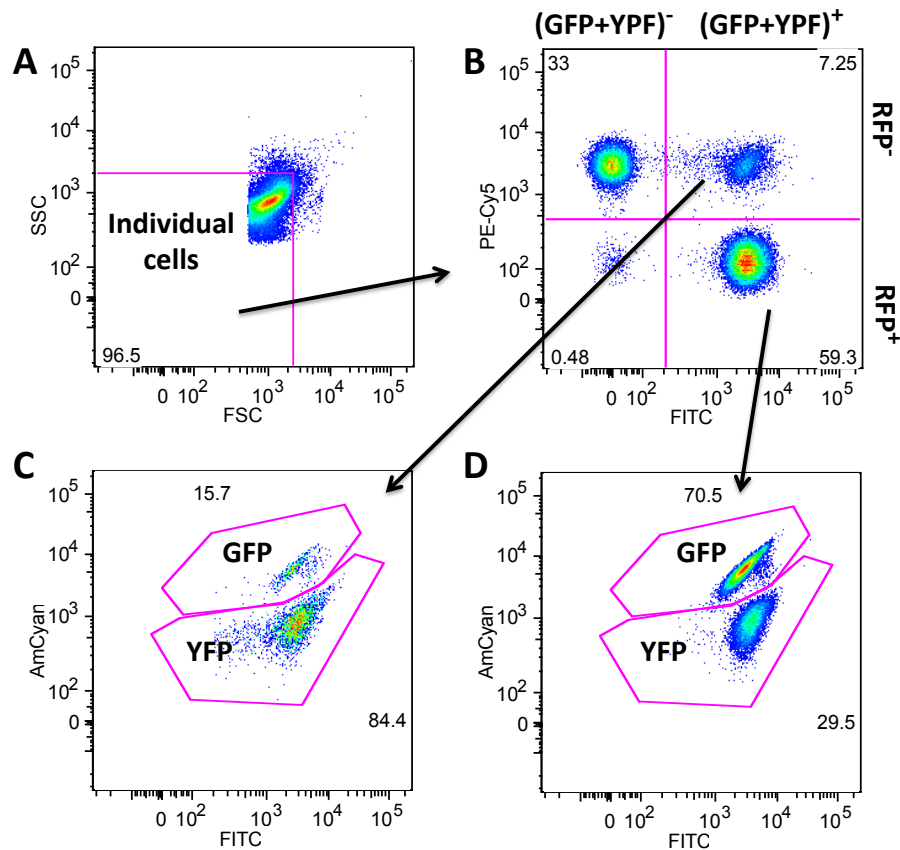
In our experiments, cells generally express a combination of GFP, YFP and RFP proteins or do not fluoresce at all. The flow cytometry analysis is based on the differences in fluorescence excitation and emission of fluorescent proteins (Figure B.1). RFP fluoresces at high wavelengths, with minimal background fluorescence from the other fluorescent proteins (Figure B.1-A); RFP-bearing cells are detected easily with excitation by the 561 nm laser and detection by the Pe-Cy5 filter around 670 nm. GFP and YFP have spectra that are very similar. To distinguish between them, we exploit the small differences in their spectra (Figure B.1, B and C): GFP excitation and emission peaks have slightly lower wavelengths than the YFP ones. Excitation with the 488 nm laser leads to a relatively stronger signal around 530 nm (FITC filter) for YFP; excitation with the 405 nm laser is weak for both proteins, but stronger for GFP relatively to YFP.



**Figure B.1: Fluorescence spectra of fluorescent proteins.** Fluorescence spectra were measured with an Infinite M200 reader (Tecan), on overnight cultures of cells carrying no fluorescence gene (neg), P<sup>+</sup>GFP plasmid (GFP), P<sup>-</sup>YFP plasmid (YFP) or tdCherry inserted into the chromosome (RFP). Fluorescence is shown after LB background subtraction. In **A**, fluorescence excitation at 561nm is shown as a function of the emission wavelength. In **B**, relative fluorescence for GFP and YFP is shown with 580 nm emission, as a function of the excitation wavelength. In **C**, relative fluorescence for GFP and YFP is shown with 440 nm excitation, as a function of the emission wavelength.

An example of flow cytometry analysis is given in Figure B.2. Gating is first performed on forward and side scattering (FSC and SSC, Figure B.2-A), in order to exclude the ma-

jority of doublets (agregates of cells that are counted as an unique event). Four quadrants



**Figure B.2: Example of flow cytometry analysis.** In each graph, each point represents an individual event detected by the cytometer, superposed points being indicated by the color scale (blue for low density to red for maximum density). The analysis is based on successive gating steps (pink sectors), where events are selected based on their light scattering or fluorescence properties. In **A**, gating is first performed based on FSC and SSC: FSC (forward scatter) measures diffracted light at small angles and roughly indicates cell size; SSC (side scatter) measures diffracted light at higher angles, and roughly indicates cell internal complexity. Gating excludes events with high FSC and SSC, that contain more than one cell. Events are next gated in **B** as a function of their fluorescence detected from FITC and Pe-Cy5 filters: the FITC filter measures fluorescence emitted around 530 nm after excitation by the 488 nm laser, which arises from GFP and YFP fluorescence. The Pe-Cy5 filter measures fluorescence emitted around 670 nm after excitation by the 561nm laser, and corresponds to RFP fluorescence. In **C** and **D**, events of each FITC-positive quadrant are finally separated based on the fluorescence emitted in the AmCyan filter around 525 nm, after excitation from the 405 nm laser. This indicates if FITC-detected fluorescence arises from GFP or YFP.

are then separated based on 530/30nm (FITC) and 670/30nm (PE-Cy5) filters (Figure B.2-B). PE-Cy5-positive quadrants contain cells expressing RFP. FITC-positive quadrants contain cells expressing either GFP or YFP, that both emit fluorescence around 530nm. In each of the FITC-positive quadrants, cells are finally separated based on 530/30nm and 525/50 filters, distinguishing GFP and YFP (Figure B.2-C and D).



# Bibliography



# Bibliography

- C. Absalon, K. Van Dellen, and P. Watnick. A communal bacterial adhesin anchors biofilm and bystander cells to surfaces. *PLoS Pathogens*, 7(8):e1002210, 2011.
- M. Achtman, N. Kennedy, and R. Skurray. Cell-cell interactions in conjugating *Escherichia coli*: role of traT protein in surface exclusion. *Proceedings of the National Academy of Sciences*, 74(11):5104–5108, 1977.
- M. Ackermann, B. Stecher, N. Freed, P. Songhet, W.-D. Hardt, and M. Doebeli. Self-destructive cooperation mediated by phenotypic noise. *Nature*, 454(7207):987–990, 2008.
- M. Adamczyk and G. Jagura-Burdzy. Spread and survival of promiscuous IncP-1 plasmids. *Acta Biochimica Polonica*, 50(2):425–454, 2003.
- S. Alizon, A. Hurford, N. Mideo, and M. Van Baalen. Virulence evolution and the trade-off hypothesis: history, current state of affairs and the future. *Journal of Evolutionary Biology*, 22(2):245–259, 2009.
- C. E. Alvarez-Martinez and P. J. Christie. Biological diversity of prokaryotic type IV secretion systems. *Microbiology and Molecular Biology Reviews*, 73(4):775–808, 2009.
- R. Anderson and R. M. May. Coevolution of hosts and parasites. *Parasitology*, 85(2):411–426, 1982.
- J.-B. André and B. Godelle. Multicellular organization in bacteria as a target for drug therapy. *Ecology Letters*, 8(8):800–810, 2005.
- K. G. Anthony, C. Sherburne, R. Sherburne, and L. S. Frost. The role of the pilus in recipient cell recognition during bacterial conjugation mediated by F-like plasmids. *Molecular microbiology*, 13(6):939–953, 1994.
- K. G. Anthony, W. A. Klimke, J. Manchak, and L. S. Frost. Comparison of proteins involved in pilus synthesis and mating pair stabilization from the related plasmids F and R100-1: insights into the mechanism of conjugation. *Journal of bacteriology*, 181(17):5149–5159, 1999.
- W. Arber. Genetic variation: molecular mechanisms and impact on microbial evolution. *FEMS microbiology reviews*, 24(1):1–7, 2000.
- D. Arutyunov and L. S. Frost. F conjugation: Back to the beginning. *Plasmid*, 70(1):18–32, 2013.
- G. F. Audette, J. Manchak, P. Beatty, W. A. Klimke, and L. S. Frost. Entry exclusion in F-like plasmids requires intact TraG in the donor that recognizes its cognate TraS in the recipient. *Microbiology*, 153(2):442–451, 2007.
- O. Avery, C. MacLeod, and M. McCarty. Studies on the chemical nature of the substance inducing transformation of pneumococcal types. *The Journal of experimental medicine*, 79(2):137–158, 1944.

- A. Babic, A. B. Lindner, M. Vulic, E. J. Stewart, and M. Radman. Direct visualization of horizontal gene transfer. *Science*, 319(5869):1533–1536, 2008.
- M. I. Bahl, L. H. Hansen, T. R. Licht, and S. J. Sorensen. Conjugative transfer facilitates stable maintenance of IncP-1 plasmid pKJK5 in *Escherichia coli* cells colonizing the gastrointestinal tract of the germfree rat. *Applied and Environmental Microbiology*, 73(1):341–343, 2007.
- M. I. Bahl, L. H. Hansen, and S. J. Sorensen. Persistence mechanisms of conjugative plasmids. In *Horizontal Gene Transfer: Genomes in Flux*. Maria B. Gogarten, 2009.
- J. E. Bailey, N. A. Da Silva, S. W. Peretti, J. H. Seo, and F. Srienc. Studies of host-plasmid interactions in recombinant microorganisms. *Annals of the New York Academy of Sciences*, 469(1):194–211, 1986.
- D. Baltrus. Exploring the costs of horizontal gene transfer. *Trends in Ecology and Evolution*, 28(8):489–495, 2013.
- D. Baltrus, K. Guillemin, and P. Phillips. Natural transformation increases the rate of adaptation in the human pathogen *Helicobacter pylori*. *Evolution*, 62(1):39–49, 2007.
- M. Barlow and B. Hall. Phylogenetic analysis shows that the OXA  $\beta$ -lactamase genes have been on plasmids for millions of years. *Journal of Molecular Evolution*, 55(3):314–321, 2002.
- R. Barrangou, C. Fremaux, H. Deveau, M. Richards, P. Boyaval, S. Moineau, D. A. Romero, and P. Horvath. CRISPR provides acquired resistance against viruses in prokaryotes. *Science*, 315(5819):1709–1712, 2007.
- L. Becks and A. Agrawal. The evolution of sex is favoured during adaptation to new environments. *PLoS Biology*, 10(5):e1001317, 2012.
- B. Belliure, A. Janssen, P. Maris, D. Peters, and M. Sabelis. Herbivore arthropods benefit from vectoring plant viruses: Thrips benefit from vectoring TSWV. *Ecology Letters*, 8(1):70–79, 2004.
- W. Bentley, N. Mirjalili, D. Andersen, R. Davis, D. Kompala, and W. Bentley. Plasmid-encoded protein: The principal factor in the “metabolic burden” associated with recombinant bacteria. *Biotechnology and Bioengineering*, 102(5):1283–1297, 2009.
- C. T. Bergstrom, M. Lipsitch, and B. R. Levin. Natural selection, infectious transfer and the existence conditions for bacterial plasmids. *Genetics*, 155(4):1505–1519, 2000.
- L. Beutin and M. Achtman. Two *Escherichia coli* chromosomal cistrons, *sfrA* and *sfrB*, which are needed for expression of F factor *tra* functions. *Journal of bacteriology*, 139(3):730–737, 1979.
- H. Blanca-Ordóñez, J. J. Oliva-García, D. Pérez-Mendoza, M. J. Soto, J. Olivares, J. Sanjuan, and J. Nogales. pSymA-Dependent mobilization of the *Sinorhizobium meliloti* pSymB megaplasmid. *Journal of Bacteriology*, 192(23):6309–6312, 2010.
- L. Bossi, J. A. Fuentes, G. Mora, and N. Figueroa-Bossi. Prophage contribution to bacterial population dynamics. *Journal of Bacteriology*, 185(21):6467–6471, 2003.
- J. E. Bouma and R. E. Lenski. Evolution of a bacteria/plasmid association. *Nature*, 335(6188):351–352, 1988.

- E. F. Boyd, C. W. Hill, S. M. Rich, and D. L. Hartl. Mosaic structure of plasmids from natural populations of *Escherichia coli*. *Genetics*, 143(3):1091–1100, 1996.
- R. Boyd and P. Richerson. Cultural transmission and the evolution of cooperative behavior. *Human Ecology*, 10(3):325–351, 1982.
- D. Bradley, D. E. Taylor, and D. Cohen. Specification of surface mating systems among conjugative drug resistance plasmids in *Escherichia coli* K-12. *Journal of Bacteriology*, 143(3):1466–1470, 1980.
- Y. Brandvain, C. Goodnight, and M. J. Wade. Horizontal transmission rapidly erodes disequilibria between organelle and symbiont genomes. *Genetics*, 189(1):397–404, 2011.
- S. P. Brown. Cooperation and conflict in host–manipulating parasites. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, 266(1431):1899–1904, 1999.
- S. P. Brown, M. Hochberg, and B. Grenfell. Does multiple infection select for raised virulence? *Trends in microbiology*, 10(9):401–405, 2002.
- S. P. Brown, L. Le Chat, M. De Paepe, and F. Taddei. Ecology of microbial invasions: Amplification allows virus carriers to invade more rapidly when rare. *Current Biology*, 16(20):2048–2052, 2006.
- S. P. Brown, S. A. West, S. P. Diggle, and A. S. Griffin. Social evolution in microorganisms and a Trojan horse approach to medical intervention strategies. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 364(1533):3157–3168, 2009.
- A. Buckling and M. A. Brockhurst. Kin selection and the evolution of virulence. *Heredity*, 100(5):484–488, 2008.
- J. J. Bull, I. J. Molineux, and W. Rice. Selection of benevolence in a host-parasite system. *Evolution*, 45(4):875–882, 1991.
- P. Bundock, A. den Dulk-Ras, A. Beijersbergen, and P. J. Hooykaas. Trans-kingdom T-DNA transfer from *Agrobacterium tumefaciens* to *Saccharomyces cerevisiae*. *The EMBO Journal*, 14(13):3206, 1995.
- A. Burt and R. Trivers. *Genes in conflict. The biology of selfish genetic elements*. Belknap Press, Harvard, 2006.
- C. Canchaya, G. Fournous, S. Chibani-Chennoufi, M.-L. Dillmann, and H. Brüssow. Phage as agents of lateral gene transfer. *Current Opinion in Microbiology*, 6(4):417–424, 2003.
- S. Casjens, N. Palmer, R. Van Vugt, W. Mun Huang, B. Stevenson, P. Rosa, R. Lathigra, G. Sutton, J. Peterson, and R. Dodson. A bacterial genome in flux: the twelve linear and nine circular extrachromosomal DNAs in an infectious isolate of the Lyme disease spirochete *Borrelia burgdorferi*. *Molecular microbiology*, 35(3):490–516, 2000.
- L. Chao and B. R. Levin. Structured habitats and the evolution of anticompetitor toxins in bacteria. *Proceedings of the National Academy of Sciences*, 78(10):6324–6328, 1981.
- K. Cheah and R. Skurray. The F plasmid carries an IS3 insertion within *fnO*. *Journal of general microbiology*, 132(12):3269–3275, 1986.



- H. Cho, U. M. Pinto, and S. C. Winans. Transsexuality in the rhizosphere: Quorum sensing reversibly converts *Agrobacterium tumefaciens* from phenotypically female to male. *Journal of Bacteriology*, 191(10):3375–3383, 2009.
- J. Chuang, O. Rivoire, and S. Leibler. Cooperation and Hamilton’s rule in a simple synthetic microbial system. *Molecular Systems Biology*, 6, 2010.
- J. S. Chuang, O. Rivoire, and S. Leibler. Simpson’s paradox in a synthetic microbial system. *Science*, 323(5911):272–275, 2009.
- S. H. Chung, C. Rosa, E. D. Scully, M. Peiffer, J. F. Tooker, K. Hoover, D. S. Luthe, and G. W. Felton. Herbivore exploits orally secreted bacteria to suppress plant defenses. *Proceedings of the National Academy of Sciences*, 110(39):15728–15733, 2013.
- T. F. Cooper. Recombination speeds adaptation by reducing competition between beneficial mutations in populations of *Escherichia coli*. *PLoS biology*, 5(9):e225, 2007.
- T. F. Cooper and J. A. Heinemann. Selection for plasmid post-segregational killing depends on multiple infection: evidence for the selection of more virulent parasites through parasite-level competition. *Proceedings of the Royal Society B: Biological Sciences*, 272(1561):403–410, 2005.
- B. J. Crespi. The evolution of social behavior in microorganisms. *Trends in Ecology and Evolution*, 16(4):178–183, 2001.
- R. Curtiss, L. G. Caro, D. P. Allison, and D. R. Stallions. Early stages of conjugation in *Escherichia coli*. *Journal of bacteriology*, 100(2):1091–1104, 1969.
- C. Dahlberg and L. Chao. Amelioration of the cost of conjugative plasmid carriage in *Escherichia coli* K12. *Genetics*, 165(4):1641–1649, 2003.
- J. Damore and J. Gore. Understanding microbial cooperation. *Journal of Theoretical Biology*, 299:31–41, 2012.
- C. Darwin. *On the Origin of Species by Means of Natural Selection, or, the Preservation of Favoured Races in the Struggle for Life*. John Murray, London, UK, 1859.
- K. A. Datsenko and B. L. Wanner. One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. *Proceedings of the National Academy of Sciences*, 97(12):6640–6645, 2000.
- M. S. Datta, K. S. Korolev, I. Cvijovic, C. Dudley, and J. Gore. Range expansion promotes cooperation in an experimental microbial metapopulation. *Proceedings of the National Academy of Sciences*, 110(18):7354–7359, 2013.
- R. Dawkins. *The Selfish Gene*. Oxford University Press, New York, 1976.
- L. De Gelder, J. J. Williams, J. M. Ponciano, M. Sota, and E. M. Top. Adaptive plasmid evolution results in host-range expansion of a broad-host-range plasmid. *Genetics*, 178(4):2179–2190, 2008.
- L. de Vargas Roditi, K. Boyle, and J. Xavier. Multilevel selection analysis of a microbial social trait. *Molecular Systems Biology*, 9, 2013.

- G. del Solar, R. Giraldo, M. Ruiz-Echevarria, M. Espinosa, and R. Diaz-Orejas. Replication and control of circular bacterial plasmids. *Microbiology and Molecular Biology Reviews*, 62(2):434–464, 1998.
- J. C. Diaz Ricci and M. E. Hernández. Plasmid effects on *Escherichia coli* metabolism. *Critical reviews in biotechnology*, 20(2):79–108, 2000.
- S. P. Diggle, A. Gardner, S. A. West, and A. S. Griffin. Evolutionary theory of bacterial quorum sensing: when is a signal not a signal? *Philosophical Transactions of the Royal Society B: Biological Sciences*, 362(1483):1241–1249, 2007a.
- S. P. Diggle, A. S. Griffin, G. S. Campbell, and S. A. West. Cooperation and conflict in quorum-sensing bacterial populations. *Nature*, 450(7168):411–414, 2007b.
- F. Dionisio. Plasmids survive despite their cost and male-specific phages due to heterogeneity of bacterial populations. *Evolutionary Ecology Research*, 7(8):1089–1107, 2005.
- F. Dionisio. Selfish and spiteful behaviour through parasites and pathogens. *Evolutionary Ecology Research*, 9(7):1199, 2007.
- F. Dionisio, I. Matic, M. Radman, O. R. Rodrigues, and F. Taddei. Plasmids spread very fast in heterogeneous bacterial communities. *Genetics*, 162(4):1525–1532, 2002.
- F. Dionisio, I. Conceição, A. Marques, L. Fernandes, and I. Gordo. The evolution of a conjugative plasmid and its ability to increase bacterial fitness. *Biology Letters*, 1(2):250–252, 2005.
- M. Doyle, M. Fookes, A. Ivens, M. W. Mangan, J. Wain, and C. J. Dorman. An H-NS-like stealth protein aids horizontal DNA transmission in bacteria. *Science*, 315(5809):251–252, 2007.
- G. Dubey and S. Ben-Yehuda. Intercellular nanotubes mediate bacterial communication. *Cell*, 144(4):590–600, 2011.
- B. A. Duerkop, C. V. Clements, D. Rollins, J. L. M. Rodrigues, and L. V. Hooper. A composite bacteriophage alters colonization by an intestinal commensal bacterium. *Proceedings of the National Academy of Sciences*, 109(43):17621–17626, 2012.
- L. A. Dugatkin, M. Perlin, J. S. Lucas, and R. Atlas. Group-beneficial traits, frequency-dependent selection and genotypic diversity: an antibiotic resistance paradigm. *Proceedings of the Royal Society B: Biological Sciences*, 272(1558):79–83, 2005.
- Z. Dumas and R. Kümmerli. Cost of cooperation rules selection for cheats in bacterial metapopulations. *Journal of Evolutionary Biology*, 25(3):473–484, 2012.
- G. Dunny. Enterococcal sex pheromones: Signaling, social behavior, and evolution. *Annual Review of Genetics*, 47(1):130919210319006, 2012.
- G. M. Dunny. The peptide pheromone-inducible conjugation system of *Enterococcus faecalis* plasmid pCF10: cell-cell signalling, gene transfer, complexity and evolution. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 362(1483):1185–1193, 2007.
- W. G. Eberhard. Why do bacterial plasmids carry some genes and not others? *Plasmid*, 21(3):167–174, 1989.

- W. G. Eberhard. Evolution in bacterial plasmids and levels of selection. *Quarterly Review of Biology*, 65(1):3–22, 1990.
- D. Ebert. The epidemiology and evolution of symbionts with mixed-mode transmission. *Annual Review of Ecology, Evolution, and Systematics*, 44(1):623–643, 2013.
- R. Ellis, A. Lilley, S. Lacey, D. Murrell, and H. Godfray. Frequency-dependent advantages of plasmid carriage by *Pseudomonas* in homogeneous and spatially structured environments. *The ISME Journal*, 1(1):92–95, 2007.
- M. B. Elowitz. Stochastic gene expression in a single cell. *Science*, 297(5584):1183–1186, 2002.
- J. D. Elsas and M. J. Bailey. The ecology of transfer of mobile genetic elements. *FEMS microbiology ecology*, 42(2):187–197, 2002.
- P. W. Ewald. Transmission modes and evolution of the parasitism-mutualism continuum. *Annals of the New York Academy of Sciences*, 503(1):295–306, 1987.
- M. W. Feldman and L. L. Cavalli-Sforza. Cultural and biological evolutionary processes: gene-culture disequilibrium. *Proceedings of the National Academy of Sciences*, 81(5):1604–1607, 1984.
- L. Fogarty, P. Strimling, and K. N. Laland. The evolution of teaching. *Evolution*, 65(10):2760–2770, 2011.
- N. Forns, R. C. Banos, C. Balsalobre, A. Juarez, and C. Madrid. Temperature-dependent conjugative transfer of R27: Role of chromosome- and plasmid-encoded Hha and H-NS proteins. *Journal of Bacteriology*, 187(12):3950–3959, 2005.
- K. R. Foster. A defense of sociobiology. *Cold Spring Harbor Symposia on Quantitative Biology*, 74(0):403–418, 2010.
- K. R. Foster and T. Wenseleers. A general model for the evolution of mutualisms. *Journal of Evolutionary Biology*, 19:1283–1293, 2006.
- S. Frank. Genetics of mutualism: the evolution of altruism between species. *Journal of Theoretical Biology*, 170(4):393–400, 1994.
- S. A. Frank. Models of parasite virulence. *Quarterly Review of Biology*, 71(1):37–38, 1996.
- S. A. Frank. *Foundations of social evolution*. Princeton University Press, Princeton, NJ, 1998.
- S. A. Frank. Natural selection. III. Selection versus transmission and the levels of selection. *Journal of Evolutionary Biology*, 25(2):227–243, 2012.
- L. S. Frost and G. Koraimann. Regulation of bacterial conjugation: balancing opportunity with adversity. *Future Microbiology*, 5(7):1057–1071, 2010.
- L. S. Frost and J. Manchak. F phenocopies: characterization of expression of the F transfer region in stationary phase. *Microbiology*, 144(9):2579–2587, 1998.
- L. S. Frost, K. Ippen-Ihler, and R. A. Skurray. Analysis of the sequence and gene products of the transfer region of the F sex factor. *Microbiological reviews*, 58(2):162–210, 1994.

- L. S. Frost, R. Leplae, A. O. Summers, and A. Toussaint. Mobile genetic elements: the agents of open source evolution. *Nature Reviews Microbiology*, 3(9):722–732, 2005.
- J. Gama, A. Reis, I. Domingues, H. Mendes-Soares, A. Matos, and F. Dionisio. Temperate bacterial viruses as double-edged swords in bacterial warfare. *PLoS ONE*, 8(3):e59043, 2013.
- M. Garcillán-Barcia and F. de la Cruz. Why is entry exclusion an essential feature of conjugative plasmids? *Plasmid*, 60(1):1–18, 2008.
- A. Gardner. The Price equation. *Current Biology*, 18(5):198–202, 2008.
- A. Gardner and S. West. Greenbeards. *Evolution*, 64(1):25–38, 2010.
- A. Gardner and S. A. West. Spite and the scale of competition. *Journal of Evolutionary Biology*, 17(6):1195–1203, 2004.
- J. M. Ghigo. Natural conjugative plasmids induce bacterial biofilm development. *Nature*, 412(6845):442–445, 2001.
- O. M. Gilbert, K. R. Foster, N. J. Mehdiabadi, J. E. Strassmann, and D. C. Queller. High relatedness maintains multicellular cooperation in a social amoeba by controlling cheater mutants. *Proceedings of the National Academy of Sciences*, 104(21):8913–8917, 2007.
- T. Giraud and J. A. Shykoff. Bacterial cooperation controlled by mobile elements: kin selection versus infectivity. *Heredity*, 107(3):277–278, 2011.
- M. R. Goddard, D. Greig, and A. Burt. Outcrossed sex allows a selfish gene to invade yeast populations. *Proceedings of the Royal Society B: Biological Sciences*, 268(1485):2537–2542, 2001.
- J. González, J. Brown, and B. Carlton. Transfer of *Bacillus thuringiensis* plasmids coding for S-endotoxin among strains of *B. thuringiensis* and *B. cereus*. *Proceedings of the National Academy of Sciences*, 79:6951–6955, 1982.
- D. M. Gordon. Rate of plasmid transfer among *Escherichia coli* strains isolated from natural populations. *Journal of general Microbiology*, 138(1):17–21, 1992.
- A. S. Griffin, S. A. West, and A. Buckling. Cooperation and competition in pathogenic bacteria. *Nature*, 430:1024–1027, 2004.
- F. Griffith. The significance of pneumococcal types. *Journal of Hygiene*, 27:113–159, 1928.
- D. G. Guiney. Promiscuous transfer of drug resistance in gram-negative bacteria. *Journal of Infectious Diseases*, 149(3):320–329, 1984.
- D. G. Guiney, P. Hasegawa, and C. E. Davis. Plasmid transfer from *Escherichia coli* to *Bacteroides fragilis*: differential expression of antibiotic resistance phenotypes. *Proceedings of the National Academy of Sciences*, 81(22):7203–7206, 1984.
- R. J. F. Haft, J. E. Mittler, and B. Traxler. Competition favours reduced cost of plasmids to host bacteria. *The ISME journal*, 3(7):761–769, 2009.

- S. Halary, J. W. Leigh, B. Cheaib, P. Lopez, and E. Baptiste. Network analyses structure genetic diversity in independent genetic worlds. *Proceedings of the National Academy of Sciences*, 107(1):127–132, 2009.
- J. Haldane. *The causes of evolution*. Longmans, Green and Co., London, UK, 1932.
- O. Hallatschek, P. Hersen, S. Ramanathan, and D. Nelson. Genetic drift at expanding frontiers promotes gene segregation. *Proceedings of the National Academy of Sciences*, 104(50):19926–19930, 2007.
- H. Hamilton, N. Domínguez, K. Schwartz, K. Hackett, and J. Dillard. *Neisseria gonorrhoeae* secretes chromosomal DNA via a novel type IV secretion system. *Molecular Microbiology*, 55(6):1704–1721, 2005.
- W. D. Hamilton. The genetical evolution of social behaviour. II. *Journal of Theoretical Biology*, 7:17–52, 1964a.
- W. D. Hamilton. The genetical evolution of social behaviour. I. *Journal of Theoretical Biology*, 7:1–16, 1964b.
- W. D. Hamilton. Selfish and spiteful behaviour in an evolutionary model. *Nature*, 228:1218–1220, 1970.
- G. Hardin. The tragedy of the commons. *Science*, 162:1243–1248, 1968.
- M. J. Harley and J. F. Schildbach. Swapping single-stranded DNA sequence specificities of relaxases from conjugative plasmids F and R100. *Proceedings of the National Academy of Sciences*, 100(20):11243–11248, 2003.
- E. Harrison and M. Brockhurst. Plasmid-mediated horizontal gene transfer is a coevolutionary process. *Trends in Microbiology*, 20(6):262–267, 2012.
- E. Harrison, V. Koufopanou, A. Burt, and R. C. MacLean. The cost of copy number in a selfish genetic element: the 2- $\mu$ m plasmid of *Saccharomyces cerevisiae*. *Journal of Evolutionary Biology*, 25(11):2348–2356, 2012.
- F. Harrison, L. Browning, M. Vos, and A. Buckling. Cooperation and virulence in acute *Pseudomonas aeruginosa* infections. *BMC biology*, 4(1):21, 2006.
- M. Hausner and S. Wuertz. High rates of conjugation in bacterial biofilms as determined by quantitative in situ analysis. *Applied and Environmental Microbiology*, 65(8):3710–3713, 1999.
- W. Hayes. Observations on a transmissible agent determining sexual differentiation in *Bacterium coli*. *Journal of General Microbiology*, 8(1):72–88, 1953.
- E. A. Herre, N. Knowlton, U. G. Mueller, and S. A. Rehner. The evolution of mutualisms: exploring the paths between conflict and cooperation. *Trends in Ecology and Evolution*, 14(2):49–53, 1999.
- H. Heuer, Z. Abdo, and K. Smalla. Patchy distribution of flexible genetic elements in bacterial populations mediates robustness to environmental uncertainty. *FEMS Microbiology Ecology*, 65(3):361–371, 2008.
- B. T. Ho, M. Basler, and J. J. Mekalanos. Type 6 secretion system-mediated immunity to type 4 secretion system-mediated gene transfer. *Science*, 342(6155):250–253, 2013.

- R. Holt and J. Lawton. The ecological consequences of shared natural enemies. *Annual Review of Ecology and Systematics*, 25:495–520, 1994.
- M. Jalasvuori, V.-P. Friman, A. Nieminen, J. K. H. Bamford, and A. Buckling. Bacteriophage selection against a plasmid-encoded sex apparatus leads to the loss of antibiotic-resistance plasmids. *Biology Letters*, 7(6):902–905, 2011.
- K. Jensen, T. Little, A. Skorpington, and D. Ebert. Empirical support for optimal virulence in a castrating parasite. *PLoS biology*, 4(7):e197, 2006.
- W. Jiang, I. Maniv, F. Arain, Y. Wang, B. R. Levin, and L. Marraffini. Dealing with the evolutionary downside of CRISPR immunity: Bacteria and beneficial plasmids. *PLoS Genetics*, 9(9):e1003844, 2013.
- O. Johnsborg, V. Eldholm, and L. Håvarstein. Natural genetic transformation: prevalence, mechanisms and function. *Research in Microbiology*, 158(10):767–778, 2007.
- D. Johnson, R. Everett, and N. Willetts. Cloning of F DNA fragments carrying the origin of transfer *oriT* and the fertility inhibition gene *finP*. *Journal of molecular biology*, 153:187–202, 1981.
- B. Kerr, C. Neuhauser, B. J. M. Bohannan, and A. Dean. Local migration promotes competitive restraint in a host–pathogen ‘tragedy of the commons’. *Nature*, 442(7098):75–78, 2006.
- E. Kiers, R. Rousseau, S. A. West, and R. F. Denison. Host sanctions and the legume–rhizobium mutualism. *Nature*, 425(6953):78–81, 2003.
- R. Kolter and E. P. Greenberg. Microbial sciences: the superficial life of microbes. *Nature*, 441(7091):300–302, 2006.
- J. H. Koschwanez, K. R. Foster, and A. W. Murray. Sucrose utilization in budding yeast as a model for the origin of undifferentiated multicellularity. *PLoS Biology*, 9(8):e1001122, 2011.
- J. E. Krol, H. D. Nguyen, L. M. Rogers, H. Beyenal, S. M. Krone, and E. M. Top. Increased transfer of a multidrug resistance plasmid in escherichia coli biofilms at the air-liquid interface. *Applied and Environmental Microbiology*, 77(15):5079–5088, 2011.
- R. Kümmerli and S. P. Brown. Molecular and regulatory properties of a public good shape the evolution of cooperation. *Proceedings of the National Academy of Sciences*, 107(44):18921–18926, 2010.
- R. Kümmerli, A. S. Griffin, S. A. West, A. Buckling, and F. Harrison. Viscous medium promotes cooperation in the pathogenic bacterium *Pseudomonas aeruginosa*. *Proceedings of the Royal Society B: Biological Sciences*, 276(1672):3531–3538, 2009a.
- R. Kümmerli, N. Jiricny, L. S. Clarke, S. A. West, and A. S. Griffin. Phenotypic plasticity of a cooperative behaviour in bacteria. *Journal of Evolutionary Biology*, 22(3):589–598, 2009b.
- A. Lang, O. Zhaxybayeva, and J. Beatty. Gene transfer agents: phage-like elements of genetic exchange. *Nature Reviews Microbiology*, 10:472–482, 2012.
- B. Lau, P. Malkus, and J. Paulsson. New quantitative methods for measuring plasmid loss rates reveal unexpected stability. *Plasmid*, 70(3):353–361, 2013.

- J. Lederberg. Cell genetics and hereditary symbiosis. *Physiol. Rev*, 32(4):403–30, 1952.
- J. Lederberg, L. L. Cavalli, and E. M. Lederberg. Sex compatibility in *Escherichia coli*. *Genetics*, 37(6):720, 1952.
- L. Lehmann, L. Keller, S. West, and D. Roze. Group selection and kin selection: two concepts but one process. *Proceedings of the National Academy of Sciences*, 104(16):6736–6739, 2007.
- E. G. Leigh Jr. The evolution of mutualism. *Journal of Evolutionary Biology*, 23(12):2507–2528, 2010a.
- E. G. Leigh Jr. The group selection controversy. *Journal of Evolutionary Biology*, 23(1):6–19, 2010b.
- R. E. Lenski and J. E. Bouma. Effects of segregation and selection on instability of plasmid pACYC184 in *Escherichia coli* B. *Journal of bacteriology*, 169(11):5314–5316, 1987.
- B. R. Levin. The evolution of sex in bacteria. In *The evolution of sex: an examination of current ideas*, pages 194–211. Sinauer Associates, Sunderland, MA, 1988.
- B. R. Levin. Nasty viruses, costly plasmids, population dynamics, and the conditions for establishing and maintaining CRISPR-Mediated adaptive immunity in bacteria. *PLoS Genetics*, 6(10):e1001171, 2010.
- B. R. Levin and C. T. Bergstrom. Bacteria are different: observations, interpretations, speculations, and opinions about the mechanisms of adaptive evolution in prokaryotes. *Proceedings of the National Academy of Sciences*, 97(13):6981–6985, 2000.
- B. R. Levin, F. M. Stewart, and V. A. Rice. The kinetics of conjugative plasmid transmission: fit of a simple mass action model. *Plasmid*, 2(2):247–260, 1979.
- B. R. Levin, J. Antonovics, and H. Sharma. Frequency-dependent selection in bacterial populations. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 319:459–472, 1988.
- R. Lewontin. The units of selection. *Annual Review of Ecology and Systematics*, 1:1–18, 1970.
- M. Lipsitch, M. A. Nowak, D. Ebert, and R. M. May. The population dynamics of vertically and horizontally transmitted parasites. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, 260:321–327, 1995.
- M. Lipsitch, S. Siller, and M. A. Nowak. The evolution of virulence in pathogens with vertical and horizontal transmission. *Evolution*, 50(5):1729–1741, 1996.
- S. Lucchini, G. Rowley, M. Goldberg, D. Hurd, M. Harrison, and J. C. D. Hinton. H-NS mediates the silencing of laterally acquired genes in bacteria. *PLoS Pathogens*, 2(8):e81, 2006.
- S. A. Lujan, L. M. Guogas, H. Ragonese, S. W. Matson, and M. R. Redinbo. Disrupting antibiotic resistance propagation by inhibiting the conjugative DNA relaxase. *Proceedings of the National Academy of Sciences*, 104(30):12282–12287, 2007.

- P. D. Lundquist and B. R. Levin. Transitory derepression and the maintenance of conjugative plasmids. *Genetics*, 113(3):483–497, 1986.
- H. Luo, K. Wan, and H. H. Wang. High-frequency conjugation system facilitates biofilm formation and pAMB1 transmission by *Lactococcus lactis*. *Applied and environmental microbiology*, 71(6):2970–2978, 2005.
- S. E. Luria and M. Human. A nonhereditary, host-induced variation of bacterial viruses. *Journal of Bacteriology*, 64(4):557–569, 1952.
- R. Lutz and H. Bujard. Independent and tight regulation of transcriptional units in *Escherichia coli* via the LacR/O, the TetR/O and AraC/I1-I2 regulatory elements. *Nucleic acids research*, 25(6):1203–1210, 1997.
- J. Madsen, M. Burmølle, L. Hansen, and S. Sørensen. The interconnection between biofilm formation and horizontal gene transfer. *FEMS Immunology & Medical Microbiology*, 65(2):183–195, 2012.
- J. Manchak and L. S. Frost. Mutational analysis of F-pilin reveals domains for pilus assembly, phage infection and DNA transfer. *Molecular microbiology*, 43(1):195–205, 2002.
- C. Manoil and J. Rosenbusch. Conjugation-deficient mutants of *Escherichia coli* distinguish classes of functions of the outer membrane OmpA protein. *Molecular and General Genetics MGG*, 187(1):148–156, 1982.
- J. Marmur, R. Rownd, S. Falkow, L. S. Baron, C. Schildkraut, and P. Doty. The nature of intergeneric episomal infection. *Proceedings of the National Academy of Sciences*, 47(7):972, 1961.
- L. Marraffini and E. Sontheimer. CRISPR interference: RNA-directed adaptive immunity in bacteria and archaea. *Nature Reviews Genetics*, 11(3):181–190, 2010.
- L. A. Marraffini and E. J. Sontheimer. CRISPR interference limits horizontal gene transfer in *Staphylococci* by targeting DNA. *Science*, 322(5909):1843–1845, 2008.
- M. Matz, A. Fradkov, Y. Labas, A. Savitsky, A. Zaraisky, M. Markelov, and S. Lukyanov. Fluorescent proteins from nonbioluminescent Anthozoa species. *Nature biotechnology*, 17(10):969–973, 1999.
- T. May, K. Tsuruta, and S. Okabe. Exposure of conjugative plasmid carrying *Escherichia coli* biofilms to male-specific bacteriophages. *The ISME Journal*, 5(4):771–775, 2010.
- J. Maynard Smith. Group selection and kin selection. *Nature*, 201:1145–1147, 1964.
- J. Maynard Smith. Group selection. *Quarterly Review of Biology*, 51(2):277–283, 1976.
- S. E. Mc Ginty, D. J. Rankin, and S. P. Brown. Horizontal gene transfer and the evolution of bacterial cooperation. *Evolution*, 65(1):21–32, 2011.
- S. E. Mc Ginty, L. Lehmann, S. P. Brown, and D. J. Rankin. The interplay between relatedness and horizontal gene transfer drives the evolution of plasmid-carried public goods. *Proceedings of the Royal Society B: Biological Sciences*, 280(1761):20130400–20130400, 2013.



- L. M. Merlo, J. Pepper, B. J. Reid, and C. C. Maley. Cancer as an evolutionary and ecological process. *Nature Reviews Cancer*, 6(12):924–935, 2006.
- S. L. Messenger, I. J. Molineux, and J. J. Bull. Virulence evolution in a virus obeys a trade off. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, 266(1417):397–404, 1999.
- R. I. Modi, C. M. Wilke, R. F. Rosenzweig, and J. Adams. Plasmid macro-evolution: selection of deletions during adaptation in a nutrient-limited environment. *Genetica*, 84(3):195–202, 1991.
- S. Molin and T. Tolker-Nielsen. Gene transfer occurs with enhanced efficiency in biofilms and induces enhanced stabilisation of the biofilm structure. *Current Opinion in Biotechnology*, 14(3):255–261, 2003.
- B. Momeni, C.-C. Chen, K. Hillesland, A. Waite, and W. Shou. Using artificial systems to explore the ecology and evolution of symbioses. *Cellular and Molecular Life Sciences*, 68(8):1353–1368, 2011.
- C. D. Nadell, K. R. Foster, and J. B. Xavier. Emergence of spatial structure in cell groups and the evolution of cooperation. *PLoS Computational Biology*, 6(3):e1000716, 2010.
- H. Narra and H. Ochman. Of what use is sex to bacteria? *Current Biology*, 16(17):R705–R710, 2006.
- T. Netherwood, R. Bowden, P. Harrison, A. G. O’Donnell, D. S. Parker, and H. J. Gilbert. Gene transfer in the gastrointestinal tract. *Applied and Environmental Microbiology*, 65(11):5139–5141, 1999.
- T. Nogueira, D. Rankin, M. Touchon, F. Taddei, S. P. Brown, and E. Rocha. Horizontal gene transfer of the secretome drives the evolution of bacterial cooperation and virulence. *Current Biology*, 19(20):1683–1691, 2009.
- A. Norman, L. Hansen, and S. J. Sorensen. Conjugative plasmids: vessels of the communal gene pool. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 364:2275–2289, 2009.
- C. P. Novotny and K. Lavin. Some effects of temperature on the growth of F pili. *Journal of bacteriology*, 107(3):671–682, 1971.
- V. Ojala, J. Laitalainen, and M. Jalasvuori. Fight evolution with evolution: plasmid-dependent phages with a wide host range prevent the spread of antibiotic resistance. *Evolutionary Applications*, 6(6):925–932, 2013.
- S. Okasha. Levels of selection. *Current Biology*, 20(7):R306–R307, 2010.
- L. Orgel and F. Crick. Selfish DNA: the ultimate parasite. *Nature*, 284:604–607, 1980.
- M. Ormö, A. Cubitt, K. Kallio, L. Gross, R. Tsien, and S. Remington. Crystal structure of the *Aequorea victoria* green fluorescent protein. *Science*, 273:1392–1395, 1996.
- A. Osborn and D. Böltner. When phage, plasmids, and transposons collide: genomic islands, and conjugative- and mobilizable-transposons as a mosaic continuum. *Plasmid*, 48:202–212, 2002.

- L. Otten, J. Canaday, J. Gérard, P. Fournier, P. Crouzet, and F. Paulus. Evolution of *Agrobacteria* and their Ti plasmids - a review. *Molecular plant-microbe interactions*, 5(4):279–287, 1992.
- S. Otto and T. Lenormand. Resolving the paradox of sex and recombination. *Nature Reviews Genetics*, 3(4):252–261, 2002.
- K. L. Palmer and M. S. Gilmore. Multidrug-resistant Enterococci lack CRISPR-cas. *mBio*, 1(4):e00227–10–e00227–19, 2010.
- J. Paulsson. Multileveled selection on plasmid replication. *Genetics*, 161(4):1373–1384, 2002.
- D. Pérez-Mendoza and F. de la Cruz. *Escherichia coli* genes affecting recipient ability in plasmid conjugation: Are there any? *BMC Genomics*, 10(1):71, 2009.
- T. Pfeiffer and S. Bonhoeffer. An evolutionary scenario for the transition to undifferentiated multicellularity. *Proceedings of the National Academy of Sciences*, 100(3):1095–1098, 2003.
- N. Philippe, J.-P. Alcaraz, E. Coursange, J. Geiselmann, and D. Schneider. Improvement of pCVD442, a suicide plasmid for gene allele exchange in bacteria. *Plasmid*, 51(3):246–255, 2004.
- T. Platt and J. Bever. Kin competition and the evolution of cooperation. *Trends in Ecology and Evolution*, 24(7):370–377, 2009.
- T. Platt, C. Fuqua, and J. Bever. Resource and competitive dynamics shape the benefits of public goods cooperation in a plant pathogen. *Evolution*, 66(6):1953–1965, 2012.
- T. G. Platt, J. D. Bever, and C. Fuqua. A cooperative virulence plasmid imposes a high fitness cost under conditions that induce pathogenesis. *Proceedings of the Royal Society B: Biological Sciences*, 279(1734):1691–1699, 2011.
- S. Powers, A. Penn, and R. Watson. The concurrent evolution of cooperation and the population structures that support it. *Evolution*, 65(6):1527–1543, 2011.
- G. Price. Extension of covariance selection mathematics. *Annals of human genetics*, 35:485–490, 1972.
- P. Price, M. Westoby, B. Rice, P. Atsatt, R. Fritz, J. Thompson, and K. Mobley. Parasite mediation in ecological interactions. *Annual Review of Ecology and Systematics*, 17:487–505, 1986.
- D. C. Queller. Quantitative genetics, inclusive fitness, and group selection. *The American Naturalist*, 139(3):540–558, 1992.
- D. C. Queller. Single-gene greenbeard effects in the social amoeba *Dictyostelium discoideum*. *Science*, 299(5603):105–106, 2003.
- D. C. Queller and J. E. Strassmann. Kin selection and social insects. *BioScience*, 48(3):165–175, 1998.
- D. Rankin, K. Bargum, and H. Kokko. The tragedy of the commons in evolutionary biology. *Trends in Ecology and Evolution*, 22(12):643–651, 2007.

- D. J. Rankin, E. P. C. Rocha, and S. P. Brown. What traits are carried on mobile genetic elements, and why? *Heredity*, 106(1):1–10, 2010.
- D. J. Rankin, S. E. Ginty, T. Nogueira, M. Touchon, F. Taddei, E. P. C. Rocha, and S. P. Brown. Bacterial cooperation controlled by mobile elements: kin selection and infectivity are part of the same process. *Heredity*, 107(3):279–281, 2011.
- A. Rashtchian, G. Dubes, and S. Booth. Tetracycline-inducible transfer of tetracycline resistance in *Bacteroides fragilis* in the absence of detectable plasmid DNA. *Journal of bacteriology*, 150(1):141–147, 1982.
- B. Raymond and M. Bonsall. Cooperation and the evolutionary ecology of bacterial virulence: The *Bacillus cereus* group as a novel study system. *BioEssays*, 35(8):706–716, 2013.
- B. Raymond, S. A. West, A. S. Griffin, and M. B. Bonsall. The dynamics of cooperative bacterial virulence in the field. *Science*, 337(6090):85–88, 2012.
- R. Redfield. Do bacteria have sex? *Nature Reviews Genetics*, 2(8):634–639, 2001.
- A. M. Reiner. *Escherichia coli* females defective in conjugation and in adsorption of a single-stranded deoxyribonucleic acid phage. *Journal of bacteriology*, 119(1):183–191, 1974.
- A. Reisner, B. M. Höller, S. Molin, and E. L. Zechner. Synergistic effects in mixed *Escherichia coli* biofilms: conjugative plasmid transfer drives biofilm expansion. *Journal of bacteriology*, 188(10):3582–3588, 2006.
- A. Reisner, H. Wolinski, and E. Zechner. In situ monitoring of IncF plasmid transfer on semi-solid agar surfaces reveals a limited invasion of plasmids in recipient colonies. *Plasmid*, 67(2):155–161, 2012.
- L. Riboli-Sasco, S. P. Brown, and F. Taddei. Why teach? The evolutionary origins and ecological consequences of costly information transfer. *Sociobiology of Communication: an interdisciplinary perspective*, pages 265–274, 2008.
- M. A. Riley and D. M. Gordon. The ecological role of bacteriocins in bacterial competition. *Trends in Microbiology*, 7(3):129–133, 1999.
- M. A. Riley and J. E. Wertz. Bacteriocin diversity: ecological and evolutionary perspectives. *Biochimie*, 84(5):357–364, 2002.
- M. Rodríguez, M. Holcik, and V. N. Iyer. Lethality and survival of *Klebsiella oxytoca* evoked by conjugative IncN group plasmids. *Journal of bacteriology*, 177(22):6352–6361, 1995.
- D. Rubinsztein, W. Amos, J. Leggo, S. Goodburn, R. Ramesar, J. Old, R. Bontrop, R. McMahon, D. Barton, and M. Ferguson-Smith. Mutational bias provides a model for the evolution of Huntington’s disease and predicts a general increase in disease prevalence. *Nature Genetics*, 7:525–530, 1994.
- J. L. Sachs, U. G. Mueller, T. Wilcox, and J. J. Bull. The evolution of cooperation. *Quarterly Review of Biology*, 79(2):135–160, 2004.
- J. L. Sachs, M. O. Ehinger, and E. L. Simms. Origins of cheating and loss of symbiosis in wild *Bradyrhizobium*. *Journal of Evolutionary Biology*, 23(5):1075–1089, 2010.

- J. L. Sachs, R. G. Skophammer, and J. U. Regus. Evolutionary transitions in bacterial symbiosis. *Proceedings of the National Academy of Sciences*, 108(Supplement 2): 10800–10807, 2011.
- M. Sanchez, J. Arnold, and M. Asmussen. Symbiont survival and host-symbiont disequilibria under differential vertical transmission. *Genetics*, 154:1347–1365, 2000.
- K. M. Sandoz, S. M. Mitzimberg, and M. Schuster. Social cheating in *Pseudomonas aeruginosa* quorum sensing. *Proceedings of the National Academy of Sciences*, 104(40): 15876–15881, 2007.
- M. Schuster, C. P. Lostroh, T. Ogi, and E. P. Greenberg. Identification, timing, and signal specificity of pseudomonas aeruginosa quorum-controlled genes: a transcriptome analysis. *Journal of Bacteriology*, 185(7):2066–2079, 2003.
- A. Serna, E. Espinosa, E. M. Camacho, and J. Casadesus. Regulation of bacterial conjugation in microaerobiosis by host-encoded functions ArcAB and SdhABCD. *Genetics*, 184(4):947–958, 2010.
- L. Simonsen. The existence conditions for bacterial plasmids: Theory and reality. *Microbial Ecology*, 22(2):187–205, 1991.
- L. Simonsen, D. M. Gordon, F. M. Stewart, and B. R. Levin. Estimating the rate of plasmid transfer: an end-point method. *Journal of general microbiology*, 136(11):2319–2325, 1990.
- R. A. Skurray and P. Reeves. Characterization of lethal zygotis associated with conjugation in *Escherichia coli* K-12. *Journal of Bacteriology*, 113(1):58–70, 1973.
- F. Slater, M. Bailey, A. Tett, and S. Turner. Progress towards understanding the fate of plasmids in bacterial communities. *FEMS Microbiology Ecology*, 66(1):3–13, 2008a.
- F. R. Slater, K. D. Bruce, R. J. Ellis, A. K. Lilley, and S. L. Turner. Heterogeneous selection in a spatially structured environment affects fitness tradeoffs of plasmid carriage in Pseudomonads. *Applied and Environmental Microbiology*, 74(10):3189–3197, 2008b.
- C. Smillie, M. P. Garcillan-Barcia, M. V. Francia, E. P. C. Rocha, and F. de la Cruz. Mobility of plasmids. *Microbiology and Molecular Biology Reviews*, 74(3):434–452, 2010.
- J. Smith. The social evolution of bacterial pathogenesis. *Proceedings of the Royal Society B: Biological Sciences*, 268(1462):61–69, 2001.
- J. Smith. A Gene’s-Eye view of symbiont transmission. *The American Naturalist*, 170(4): 542–550, 2007.
- J. Smith. Superinfection drives virulence evolution in experimental populations of bacteria and plasmids. *Evolution*, 65(3):831–841, 2011.
- J. Smith. Tragedy of the commons among antibiotic resistance plasmids. *Evolution*, 66(4):1269–1274, 2012.
- S. Smukalla, M. Caldara, N. Pochet, A. Beauvais, S. Guadagnini, C. Yan, M. Vincés, A. Jansen, M.-C. Prevost, J.-P. Latgé, G. R. Fink, K. R. Foster, and K. J. Verstrepen. FLO1 is a variable green beard gene that drives biofilm-like cooperation in budding yeast. *Cell*, 135(4):726–737, 2008.

- M. Sota, H. Yano, J. M. Hughes, G. W. Daughdrill, Z. Abdo, L. J. Forney, and E. M. Top. Shifts in the host range of a promiscuous plasmid through parallel evolution of its replication initiation protein. *The ISME journal*, 4(12):1568–1580, 2010.
- V. Souza, M. Rocha, A. Valera, and L. Eguiarte. Genetic structure of natural populations of *Escherichia coli* in wild hosts on different continents. *Applied and Environmental Microbiology*, 65(8):3373–3385, 1999.
- B. Sowa, D. Moore, and K. Ippen-Ihler. Physiology of F-pilin synthesis and utilization. *Journal of Bacteriology*, 153(2):962–968, 1983.
- H. Steinmoen, E. Knutsen, and L. Haavarstein. Induction of natural competence in *Streptococcus pneumoniae* triggers lysis and DNA release from a subfraction of the cell population. *Proceedings of the National Academy of Sciences*, 99(11):7681–7686, 2002.
- J. Stern and J. Schildbach. DNA recognition by F factor TraI36: highly sequence-specific binding of single-stranded DNA. *Biochemistry*, 40(38):11586–11595, 2001.
- F. M. Stewart and B. R. Levin. The population biology of bacterial plasmids: a priori conditions for the existence of conjugationally transmitted factors. *Genetics*, 87(2):209–228, 1977.
- M. Strand and G. Burke. Polydnviruses as symbionts and gene delivery systems. *PLoS Pathogens*, 8(7):e1002757, 2012.
- J. E. Strassmann, Y. Zhu, and D. C. Queller. Altruism and social cheating in the social amoeba *Dictyostelium discoideum*. *Nature*, 408(6815):965–967, 2000.
- J. E. Strassmann, O. M. Gilbert, and D. C. Queller. Kin discrimination and cooperation in microbes. *Annual Review of Microbiology*, 65(1):349–367, 2011.
- H. Strohmaier, R. Noiges, S. Kotschan, G. Sawers, G. Högenauer, E. L. Zechner, and G. Koraimann. Signal transduction and bacterial conjugation: characterization of the role of ArcA in regulating conjugative transfer of the resistance plasmid R1. *Journal of molecular biology*, 277(2):309–316, 1998.
- D. Sun, Y. Zhang, Y. Mei, H. Jiang, Z. Xie, H. Liu, X. Chen, and P. Shen. *Escherichia coli* is naturally transformable in a novel transformation system. *FEMS Microbiology Letters*, 265(2):249–255, 2006.
- E. Szathmary and J. M. Smith. The major evolutionary transitions. *Nature*, 374:227–232, 1995.
- Y. Tanouchi, R. Smith, and L. You. Engineering microbial systems to explore ecological and evolutionary dynamics. *Current Opinion in Biotechnology*, 23(5):791–797, 2012.
- E. L. Tatum and J. Lederberg. Gene recombination in the bacterium *Escherichia coli*. *Journal of bacteriology*, 53(6):673, 1947.
- D. Taylor, C. Zeyl, and E. Cooke. Conflicting levels of selection in the accumulation of mitochondrial defects in *Saccharomyces cerevisiae*. *Proceedings of the National Academy of Sciences*, 99(6):3690–3694, 2002.
- P. D. Taylor and S. A. Frank. How to make a kin selection model. *Journal of Theoretical Biology*, 180(1):27–37, 1996.

- O. Tenaillon, F. Taddei, M. Radman, and I. Matic. Second-order selection in bacterial evolution: selection acting on mutation and recombination rates in the course of adaptation. *Research in Microbiology*, 152:11–16, 2001.
- C. Thomas. Paradigms of plasmid organization. *Molecular microbiology*, 37(3):485–491, 2000.
- C. Thomas and K. Nielsen. Mechanisms of, and barriers to, horizontal gene transfer between bacteria. *Nature Reviews Microbiology*, 3(9):711–721, 2005.
- C. M. Thomas. Evolution and population genetics of plasmids. In *Plasmid biology*, pages 509–528. Funnell and Phillips, ASM press edition, 2004.
- M. Tock and D. T. Dryden. The biology of restriction and anti-restriction. *Current Opinion in Microbiology*, 8(4):466–472, 2005.
- A. Toussaint and C. Merlin. Mobile elements as a combination of functional modules. *Plasmid*, 47(1):26–35, 2002.
- R. Trivers. The evolution of reciprocal altruism. *Quarterly Review of Biology*, 46(1):35–37, 1971.
- P. Turner. Phenotypic plasticity in bacterial plasmids. *Genetics*, 167:9–20, 2004.
- P. Turner, V. Cooper, and R. E. Lenski. Tradeoff between horizontal and vertical modes of transmission in bacterial plasmids. *Evolution*, 52(2):315–329, 1998.
- N. van der Hoeven. Evolution of bacterial surface exclusion against incompatible plasmids. *Journal of theoretical biology*, 117(3):431–452, 1985.
- G. J. Velicer, L. Kroos, and R. E. Lenski. Developmental cheating in the social bacterium *Myxococcus xanthus*. *Nature*, 404:598–601, 2000.
- G. Vilas-Bôas, L. Vilas-Bôas, D. Lereclus, and O. Arantes. *Bacillus thuringiensis* conjugation under environmental conditions. *FEMS Microbiology Ecology*, 25(4):369–374, 1998.
- M. Waldor and J. J. Mekalanos. Lysogenic conversion by a filamentous phage encoding cholera toxin. *Science*, 272:1910–1914, 1986.
- T. Watanabe. Infective heredity of multiple drug resistance in bacteria. *Bacteriological Reviews*, 27(1):87, 1963.
- S. A. West, I. Pen, and A. S. Griffin. Cooperation and competition between relatives. *Science*, 296(5565):72–75, 2002.
- S. A. West, A. S. Griffin, A. Gardner, and S. P. Diggle. Social evolution theory for microorganisms. *Nature Reviews Microbiology*, 4(8):597–607, 2006.
- S. A. West, S. P. Diggle, A. Buckling, A. Gardner, and A. S. Griffin. The social lives of microbes. *Annual Review of Ecology, Evolution, and Systematics*, 38(1):53–77, 2007a.
- S. A. West, A. S. Griffin, and A. Gardner. Evolutionary explanations for cooperation. *Current Biology*, 17(16):R661–R672, 2007b.

- S. A. West, A. S. Griffin, and A. Gardner. Social semantics: altruism, cooperation, mutualism, strong reciprocity and group selection. *Journal of Evolutionary Biology*, 20(2):415–432, 2007c.
- E. Westra, R. Staals, G. Gort, S. Høgh, S. Neumann, F. de la Cruz, P. Fineran, and S. Brouns. CRISPR-Cas systems preferentially target the leading regions of MOBF conjugative plasmids. *RNA Biology*, 10(5), 2013.
- W. R. Will and L. S. Frost. Characterization of the opposing roles of H-NS and TraJ in transcriptional regulation of the F-plasmid tra operon. *Journal of bacteriology*, 188(2):507–514, 2006.
- D. S. Wilson. A theory of group selection. *Proceedings of the National Academy of Sciences*, 72(1):143–146, 1975.
- R. A. F. Wozniak and M. Waldor. Integrative and conjugative elements: mosaic mobile genetic elements enabling dynamic lateral gene flow. *Nature Reviews Microbiology*, 8(8):552–563, 2010.
- J. B. Xavier, W. Kim, and K. R. Foster. A molecular mechanism that stabilizes cooperative secretions in *Pseudomonas aeruginosa*. *Molecular Microbiology*, 79(1):166–179, 2011.
- X. Yang, Q. Ma, and T. K. Wood. The R1 conjugative plasmid increases escherichia coli biofilm formation through an envelope stress response. *Applied and Environmental Microbiology*, 74(9):2690–2699, 2008.
- S. Yaron, G. L. Kolling, L. Simon, and K. R. Matthews. Vesicle-mediated transfer of virulence genes from *Escherichia coli* O157: H7 to other enteric bacteria. *Applied and environmental microbiology*, 66(10):4414–4420, 2000.
- E. Yurtsev, H. Chao, M. Datta, T. Artemova, and J. Gore. Bacterial cheating drives the population dynamics of cooperative antibiotic resistance plasmids. *Molecular Systems Biology*, 9, 2013.
- D. Zahrl, M. Wagner, K. Bischof, and G. Koraimann. Expression and assembly of a functional type IV secretion system elicit extracytoplasmic and cytoplasmic stress responses in *Escherichia coli*. *Journal of Bacteriology*, 188(18):6611–6621, 2006.
- M. Zatyka and C. Thomas. Control of genes for conjugative transfer of plasmids and other mobile elements. *FEMS Microbiology Reviews*, 21:291–319, 1998.
- N. Zinder and J. Lederberg. Genetic exchange in *Salmonella*. *Journal of bacteriology*, 64(5):679, 1952.





## Abstract

Bacteria are social organisms which participate in multiple cooperative and group behaviours. They moreover have peculiar genetic systems, as they often bear mobile genetic elements like plasmids, molecular symbionts that are the cause of widespread horizontal gene transfer and play a large role in bacterial evolution. Both cooperation and horizontal transfer have consequences for human health: cooperative behaviours are very often involved in the virulence of pathogens, and horizontal gene transfer leads to the spread of antibiotic resistance. The evolution of plasmid transfer has mainly been analyzed in terms of infectious benefits for selfish mobile elements, however chromosomal genes can also modulate horizontal transfer. A huge diversity in transfer rates is observed among bacterial isolates, suggesting a complex co-evolution between plasmids and hosts. Moreover, plasmids are enriched in genes involved in social behaviours, and so could play a key role in bacterial cooperative behaviours.

We study here the coevolution of gene mobility and sociality in bacteria. To investigate the selective pressures acting on plasmid transfer and public good production, we use both mathematical modelling and a synthetic system that we constructed where we can independently control public good cooperation and plasmid conjugation in *Escherichia coli*. We first show experimentally that horizontal transfer allows the specific maintenance of public good alleles in a structured population by increasing relatedness at the gene-level. We further demonstrate experimentally and theoretically that this in turn allows for second-order selection of transfer ability: when cooperation is needed, alleles promoting donor and recipient abilities for public good traits can be selected both on the plasmid and on the chromosome in structured populations. Moreover, donor ability for private good traits can also be selected on the chromosome, provided that transfer happens towards kin. The interactions between transfer and cooperation can finally lead to an association between transfer and public good production alleles, explaining the high frequency of genes related to cooperation that are located on plasmids. Globally, these results provide insight into the mechanisms maintaining cooperation in bacteria, and may suggest ways to target cooperative virulence.

**Keywords:** horizontal gene transfer; plasmids; bacterial cooperation; mobile genetic elements

## Résumé

Les bactéries sont des organismes extrêmement sociaux, qui présentent de multiples comportements de coopération. De plus, les génomes bactériens sont caractérisés par la présence de nombreux éléments génétiques mobiles, tels que les plasmides. Ces éléments mobiles sont la cause de transferts génétiques horizontaux, et jouent un rôle important dans l'évolution bactérienne. La coopération et le transfert horizontal ont tous deux des conséquences importantes sur la santé humaine: des comportements coopératifs sont souvent à l'origine de propriétés de virulence chez les bactéries pathogènes, et le transfert horizontal entraîne la dissémination de gènes de résistance aux antibiotiques. L'évolution du transfert horizontal a jusqu'ici été analysée principalement en termes de bénéfices infectieux apportés à des éléments génétiques égoïstes. Cependant, le taux de transfert des plasmides est extrêmement variable et partiellement contrôlé par les gènes des bactéries hôtes, suggérant une co-évolution complexe entre hôtes et plasmides. De plus, les plasmides sont particulièrement riches en gènes liés à des comportements coopératifs, et semblent donc jouer un rôle-clé dans les phénomènes de socialité bactérienne.

Ce travail porte sur la coévolution entre mobilité génétique et socialité chez les bactéries. Nous analysons ici les pressions de sélection agissant sur le transfert de plasmides et la production de biens publics, à l'aide de modèles mathématiques et d'un système synthétique que nous avons construit chez *Escherichia coli*, dans lequel nous pouvons contrôler indépendamment la coopération et la conjugaison. Dans un premier temps, nous montrons expérimentalement que le transfert horizontal favorise le maintien de la coopération dans une population structurée, en augmentant la sélection de parentèle agissant au niveau des gènes transférés. Dans un second temps, nous montrons expérimentalement et théoriquement que l'échange génétique lui-même peut être sélectionné: les bactéries transférant des plasmides codant pour des biens publics sont favorisées dans une population structurée. Le transfert de gènes codant pour des biens privés peut également être sélectionné, à condition que ce transfert s'effectue entre bactéries apparentées. Finalement, ces interactions entre transfert horizontal et coopération peuvent mener à une association entre allèles de coopération et de transfert, expliquant la fréquence élevée de gènes sociaux situés sur des plasmides. Ces résultats permettent de mieux comprendre le maintien de comportements coopératifs chez les bactéries, et suggèrent des moyens de cibler certains cas de virulence bactérienne.

**Mots-clés:** transfert horizontal; plasmides; coopération bactérienne; éléments génétiques mobiles