Les effets de la sélection naturelle et de la dérive génétique sur le polymorphisme neutre

Sandrine Adiba

Ecole Normale Supérieure de Paris

PhD thesis of the Doctoral School European Interdisciplinary Frontier in Life Sciences

Natural selection and genetic drift effects on neutral polymorphism

Defended by

Sandrine ADIBA

On the 13th of December 2010

PhD comitee :

Dr. **Minus van Baalen**, advisor

Dr. **Frantz Depaulis**, advisor

Pr. **Amaury Lambert**

Dr. **Pierre Cosson**

Dr. **Christine Paillard**

Dr. **Delphine Sicard**

Invited:

Pr. **Erick Denamur**, Tutor

Dr. **Clément Nizak**, Tutor

Laboratoire d'Ecologie et Evolution
Unité Mixte de Recherche 7625 CNRS/ENS/UPMC
Bat.A 7e étage-CC237
7 Quai Saint Bernard 75252 PARIS Cedex 05 France

MERCI

Cette thèse doit beaucoup à Frantz Depaulis et Minus van Baalen. J'ai eu la chance de bénéficier d'une grande liberté tout au long de ce travail et de profiter de l'expérience et savoirs de mes deux directeurs par des discussions et séances de travail toujours stimulantes et enrichissantes.

Ce travail doit beaucoup à Clément Nizak et Erick Denamur, mes deux tuteurs. Merci Erick pour ton soutien, ta disponibilité et les discussions que l'on a pu avoir ensembles, sans ta participation, cette thèse n'aurait pas été la même. Je remercie Clément qui m'a accueillie dans son laboratoire à Grenoble et m'a fait découvrir les joies de la biologie moléculaire et une ville agréable.

Christine Paillard et Pierre Cosson que j'ai rencontrés lors de colloques ont accepté de faire partie de mon jury de thèse et méritent tous mes remerciements. Je remercie également Delphine Sicard pour avoir accepté de faire partie de mon jury de thèse.

Amaury Lambert fait de plus partie de mon jury de thèse, j'en suis à la fois très honorée et particulièrement heureuse car c'est une personne que j'apprécie à la fois pour ses qualités scientifiques et humaines.

Je tiens ensuite à remercier toute l'équipe EEM, Régis Ferrière pour les discussions et sans qui ce projet n'aurait pu avoir lieu, Silvia De monte pour les discussions stimulantes et enrichissantes. Les étudiants en thèses de cette équipe ont également été présents tout au long de ce parcours : Dominique Carval avec qui j'ai passé des moments de paillasse inoubliables, Sébastien Balesteros, Robin Aguilée, Anton Camacho.

Je remercie l'ensemble du laboratoire d'écologie et particulièrement Jean-Clobert pour m'avoir recrutée, Oliver Kaltz avec qui j'ai travaillé avant cette thèse et qui m'a beaucoup appris en analyses statistiques mais aussi en rigueur d'écriture.

Je tiens ensuite à remercier François Taddei, Ariel Lindner, Samuel Bottani pour m'avoir fait confiance et m'avoir permis de faire partie d'une école doctorale hors du commun où il fait bon passer du temps. Merci également à Laura Ciriani et Céline Guarrigues pour leur disponibilité, leur réactivité et leur bonne humeur, c'est un

plaisir d'être en contact avec elles. Merci aux étudiants que j'ai pu y rencontrer: Lulla, Tim, Livio, Amandine, David et bien d'autres.

Merci aux membres du laboratoire d'Ecologie: Stéphane Legendre et Jean-Marc Rossi pour leurs conseils en informatique, Murielle Richard pou sa disponibilité et ses connaissances en biologie moléculaire, Elisabeth Nguyen, Nathalie Guillory et Clarisse Coquemont pour la partie logistique et leur disponibilité.

Un grand merci à Christophe Benedet, qui a partagé la paillasse avec moi lors de ma dernière année de thèse, avec qui j'ai beaucoup ri, et sans qui cette dernière année n'aurait pas été la même. Merci pour ta sympathie, ta bonne humeur, ta disponibilité. Je suis sûre que ta thèse sera scientifiquement une réussite.

Je tiens maintenant à remercier les personnes qui on été particulièrement proches de moi pendant toutes ces années, qui m'ont supportée, soutenue, qui ont su parfois me faire oublier ma thèse.

Je remercie tout d'abord Michèle Huet de m'avoir beaucoup appris sur les manips en labo, pour ses conseils, sa bonne humeur, sa rigueur, pour tous les moments que l'on a pu partager ensembles, les rires et les discussions scientifiques. Merci d'être encore là, de me supporter même loin à n'importe quelle heure et de me secouer quand il le faut. Cette thèse n'aurait pas vu le jour sans toi.

Merci à Fatiha Checkroun sans qui le travail mené en laboratoire n'aurait pu être possible, elle m'a été d'un aide précieuse, m'a soutenue chaque jour avec sa bonne humeur et son dynamisme.

Je remercie Loïc Elard avec qui je partage mon bureau, et les pauses en plein air. Il a su être là dans des moments difficiles et je l'en remercie de tout cœur. Merci également pour tous tes conseils et discussions scientifiques partagés.

Je remercie également toute l'équipe de Badminton, et particulièrement Guillaume Delaval mon partenaire et ami, il a su être là dans des moments particuliers, m'a soutenue et divertie sur cette fin de thèse.

Merci à Valérie Cohen qui m'a accompagnée tout au long de ce projet, tu as su me faire sortir de ma bulle de travail quand il le fallait et a partagé tant de moments particuliers et importants de ma vie. C'est un peu grâce à toi que j'en suis là aujourd'hui.

Si je dois remercier quelqu'un ensuite, c'est Marie Baker avec qui j'ai partagé tant de discussions, terrasses de café, rires, pleurs et tellement d'autres émotions. Tu as su être là sur la fin de cette thèse, l'écriture de ce manuscrit n'aurait pas eu la même saveur. Merci pour ces instants de ciel bleu au milieu de tous ces nuages.

Enfin un grand merci à ma famille, mes parents pour leur éducation et la force qu'ils m'ont donnée pour affronter la vie; Marie-Christine et Stéphane mon frère qui sait être présent à des moments clés de ma vie, je te dédie cette thèse.

Et pour finir, une pensée pour mes chiens, pour leur compagnie ; ils m'ont permis de mettre le nez dehors chaque jour durant ce travail d'écriture. Leurs promenades ont souvent été un moment de recul et de réflexion nécessaires à l'avancée de ce travail.

For the reader…

This work is about ecological interactions and evolution.

Historical overviews were included in this manuscript and give some interesting aspects about what we called "biological and theoretical models".

When possible, I have therefore tried to make connections between empirical and theoretical studies. This complementary approach was in my mind all along this thesis.

The biological model used in this work was a new model in our laboratory, for this reason, some experimental set-ups were needed before focusing on our main issue about population genetics aspects. Interesting results came from these experiments and allowed us to study some bacterial traits such as virulence genes involved in ecological interactions.

I had also details some concepts about evolutionary ecology and theoretical models in ecology and evolution to provide (I hope) a better understanding of this work.

Contents

Chapter 1

I. Introduction

1.GENERAL INTRODUCTION

Ecology is the study of environmental systems and their components interact. It traditionally focuses on the abundance and distribution of individual species, including the causal factors thereof, and emphasizes understanding how species fit together and influence each other.

The inherited traits of a population change through successive generations. The study of evolution focuses on these changes, which result from interactions among various evolutionary processes which introduce or eliminate variation in a population.

A more recent development in the study of evolution is the integration of evolutionary processes in ecological theory.

Diversity is present at each level of a biological structure, starting with gene, and including phenotypic differences among individuals (even during an individual's lifetime), and in small groups of individuals, whole populations and entire species. Diversity is essential for the survival of all living organisms; it is the basis for the species evolvability and their ability to adapt to environmental variations. The main source of variation is mutation, which introduces genetic changes. Two main evolutionary processes cause variants to become more common or rare in a population. One is natural selection, which allows traits that favor survival and reproduction to become more common, while those that hinder survival and reproduction become more rare. Another causal factor of evolution is genetic drift, which leads to random changes in common traits in a population. Genetic drift is most important when traits do not strongly influence survival, particularly so in small populations in which chance plays a disproportionate role in frequency of traits passed on to the next generation. From a genetic viewpoint, evolution is a generation-to-generation change in the frequencies of alleles within a population.

Ecological and evolutionary processes interact in nature. Ecological interactions are essential to evolutionary dynamics, since they generate selective pressures on adaptive traits and can influence trait variation and inheritance. Thus, evolution and ecology govern the properties of ecosystems, with ecological dynamics determining aspects like abundance and spatial composition, and evolutionary

dynamics controlling features like traits and genetic composition. Adaptive traits determine the conditions under which ecological dynamics unfold, and thus affect abundances. The dynamics of abundance naturally underlies the dynamics of traits. This mutual dependence between ecology and evolution is captured by the so-called eco-evolutionary feedback loop [1, 2].

A central goal of evolutionary ecology is to understand the roles of different ecological processes in producing patterns of evolutionary diversification. Theoretical and empirical studies have shown that competition and antagonistic interactions such as prey/predator or host/parasite interaction can promote diversification or maintain biological diversity (see chapter I.3 p.37).

Concepts and Background

The potential of a population to adapt to an environment depends on its genetic variance [3]. In this context, determining the distribution of the genetic variance of a population and the factors contributing at the origin, as well as the maintenance and distribution of this genetic variance observed remain central and fundamental research issues [4]. Population genetic variation arises through random mutations. According to natural selection theory, environment influences the evolution of populations by selecting the best adapted individuals. Natural selection and therefore many factors like competition for resources, spatial structure, biotic factors (e.g., predation, parasitism) influence which individuals will survive and reproduce. The tendency of alleles to spread over time in a population is then due to survival and reproductive success.

Genetic drift is a random process occurring at each generation, involving random variations in the number of offspring [5]. According to the neutral theory of molecular evolution, some alleles arising by mutation fluctuate in frequency according to the – random – number of offspring of their carriers. In the long run, this process leads to the eventual loss or fixation of mutations in the population. The most persuasive argument in favor of the neutral theory is that the less functionally important one genome area is the more variable it is. It is thus important to determine the relative contributions of neutral and selective factors in influencing genetic variability.

1.1 Neutral polymorphism

Based largely on a brilliant series of papers by Kimura in the 1960s and 70s and by King & Jukes 1969 [6], the neutral model of evolution has become the standard by which positive selection must be separated from genetic observations. In the neutral theory, the vast majority of mutations are subdivided into two groups. The first group, after which the model is named, is selectively neutral (or effectively neutral) mutations which eventually become lost or fixed in a species by genetic drift. The later substitutions account for almost all the observable nucleotide changes between two species. The second group is selectively deleterious mutations, which may frequently arise continuously and are eliminated over time by natural selection. Because these mutations are eventually and generally eliminated from a species, they are rarely observed when comparing the genomes of two species. The vast majority is counter-selected or neutral. Only a small fraction of them are positively selected. Because they cause mutant phenotypes, these mutations are well known to functional geneticists, since they account for nearly all of the mutant strains and human diseases that are much studied throughout biology and medical research.

1.2 Demography and genetic drift

Stochastic variations in the environment lead to fluctuations in population size. The influence of such stochastic variation is proportionally greater for small populations than for large ones. The smaller the population the greater the probability that fluctuation will eventually lead to extinction is. They are subject to a higher chance of extinction and are more subjected to genetic drift, resulting in stochastic variation of their gene pool. Genetic variation is determined by the joint action of natural selection and genetic drift (chance). The relative importance of genetic drift in a small population is higher, since any allele, deleterious, beneficial or neutral is more likely to be lost or fixed.

1.3 Natural selection, biotic factors and spatial structure

1.3.1 Natural selection

Selection occurs when one genotype leaves on average a different number of progeny that of another. This may happen because of differences in survival, in mating or in fertility.

Positive selection is the process by which new advantageous genetic variants invade a population. Positive selection, also known as Darwinian selection, is the main mechanism that Darwin envisioned as being responsible for phenotypic evolution.

There are three general selection regimes (Figure I-1):

Directional selection occurs when selection favors one extreme phenotypic trait value over the other extreme. This typically results in a change in the mean value of the trait under selection.

Disruptive selection occurs when selection favors the extreme phenotypic trait values over the intermediate trait values. In this case, the variance increases as the population tends to be divided into two distinct groups.

Stabilizing selection occurs when selection favors the intermediate phenotypic trait value over the extreme values. Populations under this type of selection typically experience a decrease in the amount of additive genetic variation for the trait under selection.

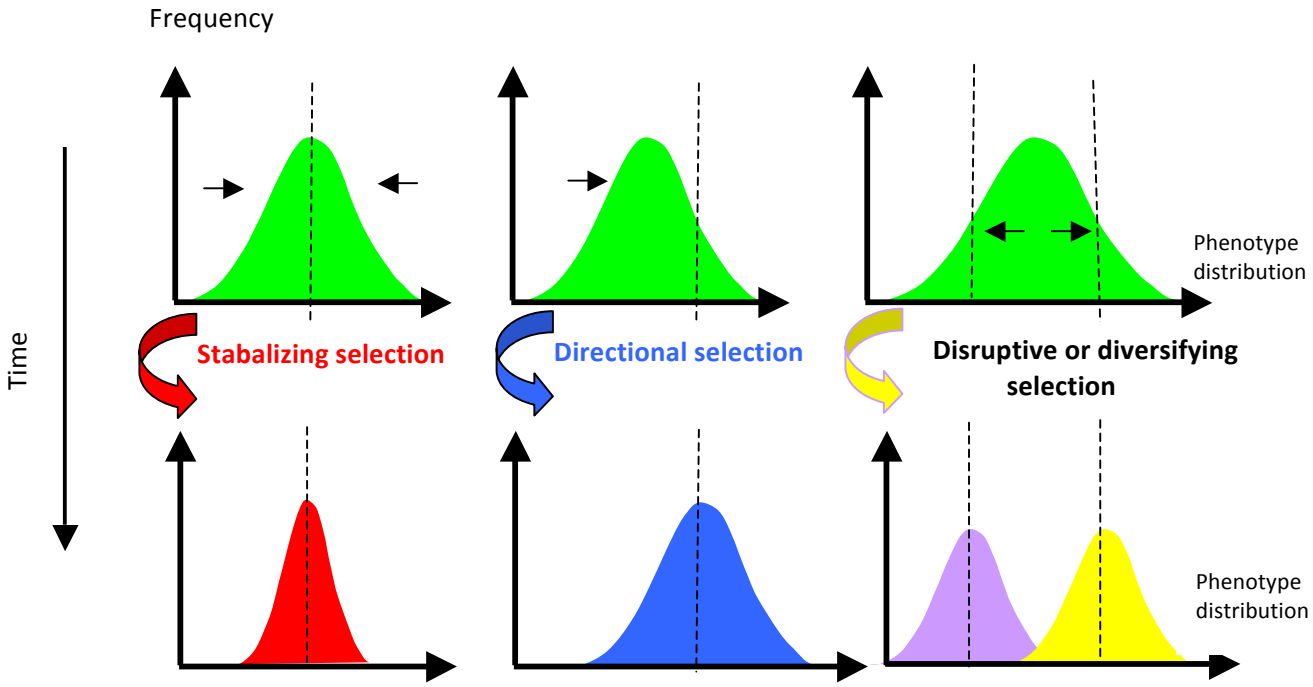


Figure I-1: The three basic selection regime
Dotted traits indicate the optimum expressions of the phenotype. Horizontal arrows show the directions of selective forces.

1.3.2 Biotic factors

Biological interactions result from the fact that organisms in an ecosystem interact with each other in the natural world, since no organism is an autonomous entity isolated from its surroundings, but it forms part of its environment. Ecological interactions such as competition and antagonistic interactions play a central role in diversification.

There is mounting evidence that natural selection and in particular *balancing selection*, has played a major role in many evolutionary diversifications and/or in maintaining polymorphism. By balancing selection, population geneticists mean any form of selection that promotes polymorphism, for example by overdominance (heterozygote superiority in fitness) or frequency-dependent selection (**FDS**), which favors different genotypes in different environments [7] (see Chapter I.3, p.37).

1.3.3 Spatial structure

Nature is definitely spatially structured. First, individuals interact preferentially with individuals in their neighborhood (populations are not homogeneous). Secondly, connections between individuals are often correlated with spatial proximity (geographical distance). Spatial structure is likely to have important evolutionary consequences in natural selection. Such effects on evolution and adaptation have been widely studied using theoretical models (see e.g. [8, 9]).

In asexual organisms, adaptation proceeds more slowly in a spatially structured environment than in one where competition is global [10]. An important phenomenon that limits adaptation in asexual is the Hill–Robertson effect [11], which is. clonal interference among advantageous haplotypes [12]. In asexual organisms, clones with different beneficial mutations compete with each other, only one becoming fixed in the population.

1.4 Selective sweep

Natural selection is the driving force for the phenotypic adaptation of organisms to a changing environment. Selective pressure applied to individuals leads to changes in the genetic content of the population. Organisms bearing a more adapted genotype would outcompete their peers. It results in the fixation of beneficial allele(s) in the population. Genes that are subject to selection are usually

found in the context of a chromosome. Adjacent genomic segments physically linked to the selected genes are therefore either dragged to fixation along with the beneficial allele or discarded with the less fit alleles during the process called genetic hitchhiking. The strength of hitchhiking decreases with the genetic distance from the selected locus and recombination can eventually separate the selected allele from adjacent loci (Figure I-2).

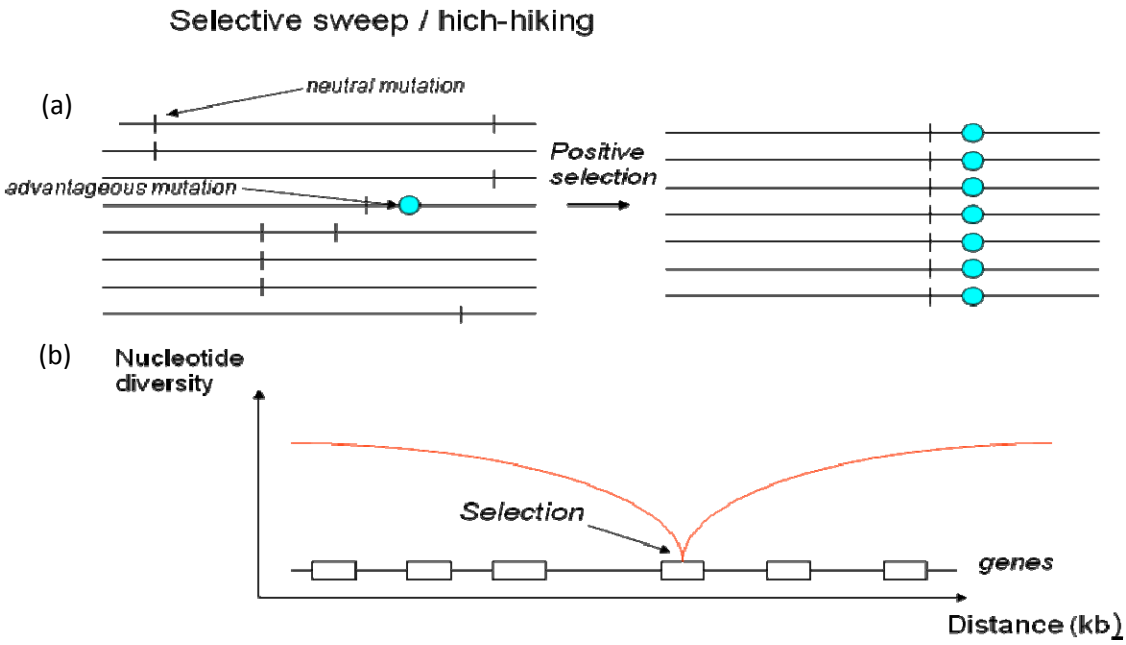


Figure I-2: Selective sweep/Hich-hiking
(a) The neutral mutation physically linked to the advantageous mutation will be fixed.
(b) The strength of hichhiking decreases with the distance from the selected locus. Hich-hiking leads to a local decrease of diversity. [13]

The processes of natural selection and genetic drift are complex. Our approach was to combine experimental evolution and modeling to disentangle the effect on neutral polymorphism of these evolutionary forces.

2.EXPERIMENTAL EVOLUTION AND MODELING

2.1 Experimental evolution

In evolutionary and experimental biology, the field of experimental evolution is concerned with testing hypotheses and theories of evolution by use of controlled experiments. Evolution may be observed in the laboratory as populations adapt to new environmental conditions and/or change by such stochastic processes as random genetic drift.

2.1.1 Early experimental evolution

One of the first to carry out a controlled evolution experiment was William Dallinger., who cultivated small unicellular organisms in a custom-built incubator (Figure I-3) over period of six years during the late 19th century (1880-1886) [14].

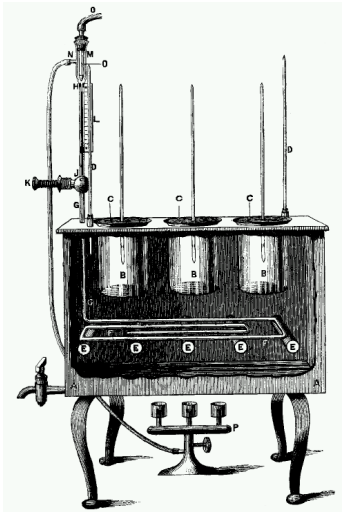


Figure I-3: Drawing of the incubator used by Dallinger in his controlled evolution experiment
The President's Address. *J. Royal Microscop. Soc.*, pp. 185-199, 1887. (Actual image is Fig. 31 on page 193).

Dallinger slowly increased the temperature of the incubator from an initial 15.5°C to 70°C. The early cultures had shown clear signs of distress at a temperature of 22.7°C, and were certainly not capable of surviving at 70°C. on the other hand, the organisms Dallinger had in his incubator at the end of the experiment were perfectly fine at 70°C. However, these organisms would no longer grow at the initial 15.5°C. Dallinger concluded that he had found evidence for Darwinian adaptation in his incubator, and that the organisms had adapted to live in a high-temperature environment. Unfortunately, Dallinger's incubator was accidentally destroyed in 1886, and Dallinger could not continue this line of research.

2.1.2 Modern experimental evolution

From the 1880s to 1980, experimental evolution was intermittently practiced by a variety of evolutionary biologists, including the highly influential Theodosius

Dobzhansky. Like other experimental research in evolutionary biology during this period, much of this work lacked extensive replication and was carried out only for relatively short periods of evolutionary time. But by 1980, a variety of evolutionary biologists realized that the key to successful experimentation lay in extensive parallel replication of evolving lineages as well as in using a greater number of generations of selection. One of the first of a new wave of experiments using this strategy was laboratory "evolutionary radiation" of *Drosophila melanogaster* populations that Michael R. Rose started in February, 1980. This system began with ten populations, five of which were cultured at older ages, and five at early ages. Since then more than 200 different populations have been created using laboratory radiation, with selection targeting multiple characteristics. Some of these highly differentiated populations have also been selected "backward" or "in reverse," by returning experimental populations to their ancestral culture regime. Hundreds of people have worked with these populations over the last three decades.

More recently, many groups carried out evolution experiments using diverse organisms including plants, vertebrates and in particular microorganisms. Many microbiologists realized that microbes offer powerful systems for evolutionary insights. (Box 1).

Box 1:| Advantages of microorganisms for evolution experiments

Microorganisms that have been used in experimental evolution include many bacteria and viruses, as well as unicellular algae and fungi. These organisms are well suited for such experiments for many practical reasons:

- They are easy to propagated.
- They reproduce quickly, which allows experiments to run for many generations.
- They allow large populations in small spaces, which facilitate experimental replication.
- They can be stored (i.e. frozen) in suspended animation and later revived which allow direct comparison of ancestral and evolved types.
- Many microbes reproduce asexually and the resulting clonality enhances the precision of experimental replication.
- Asexuality also maintains linkage between a genetic marker and the genomic background thus facilitating fitness measurements (Box 2 p.23).
- It is easy to manipulate environmental variables, such as resources, as well as the genetic composition of founding populations.
- There are abundant molecular and genomic data for many species, as well as biotechnological tools for their precise genetic analysis and manipulation.

[15]

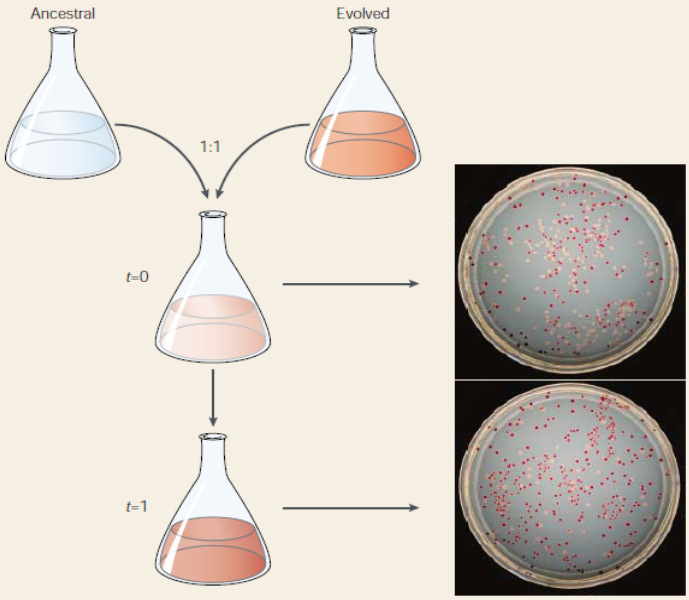
Importantly, adaptation can be quantified by measuring changes in fitness in the experimental environment, in which fitness reflects the propensity to leave descendants [16, 17].

Fitness can be measured by competition between an evolutionary line and its genetically marked ancestor (Box 2).

Box 2: | Measuring fitness

The fitness of an evolved type is generally expressed relative to its ancestor. Relative fitness is measured by allowing the ancestral and evolved types to compete with one another. Unless otherwise specified, the competition environment is the same as that used for the experimental evolution. The following description presents the protocol used in the long-term serial-transfer experiment with *Escherichia coli*[13,25,26], but similar procedures are used in experiments with many microorganisms.

The two competitors are grown separately in the competition environment to ensure that they are comparably acclimated to the test conditions. They are then mixed (usually at a 1:1 ratio) and diluted (100-fold in this case) in the competition environment. Initial densities at timepoint $t = 0$ are estimated by diluting and spreading the cells on an indicator agar that distinguishes the evolved and ancestral types by colony colour, which differs owing to an engineered marker that is selectively neutral. In this case, red and white colonies correspond to Ara^- and Ara^+ phenotypes, respectively. After one day ($t = 1$) (corresponding to the serial-transfer cycle in the evolution experiment), final densities are estimated by plating cells, as before, on the indicator agar. The growth rate of each competitor is calculated as the natural logarithm of the ratio of its final density to its initial density (adjusted for dilution during plating). Relative fitness is then defined simply as the ratio of the realized growth rates of the evolved and ancestral types.



[15]

Measuring Fitness:

Relative fitness W is calculated as the ratio of the growth rate of one strain relative that of another during their direct competition. Nevertheless, it is preferable to express performance in terms of selection rates, r (see below) (i) when one competitor is much less fit than the other [18] or (ii) when one or both competitors are declining in abundance, such as during competitions assays under starvation conditions or in the presence of an antibiotic (Box3).

Box 3: | Example of *W* and *r* calculation

Two strains *A* and *B* are growing up separately proliferating to a density of ~4.10^9 cells/ml. At time 0 of the competition experiment, the density of each strain is adjusted to 2.10^7 cells/ml. After an appropriate dilution (two successive 100 fold factors) and plating of 0.1 ml; a sample yields 192 colonies of *A* and 204 colonies of *B*.

At time 0:

A(0)= estimated density of *A* at time 0

A(0)= 192x10x100x100 = $1.92.10^7$ cells/ml

B(0)= estimated density of *B* at time 0

B(0)= 204x10x100x100

B(0)= $2.04.10^7$ cells/ml

The mixed population grows and competes over one day. At the end of which the total density has grown back to 4.10^9 cells/ml. After an appropriate dilution (three successive 100 fold factors) and plating of 0.1 ml, a sample yields 319 colonies of *A* and 107 colonies of *B*.

At time 1:

A(1)= estimated density of *A* at time 1 day

A(1)= 319x10x100x100x100

A(1)= $3.19.10^9$ cells/ml

B(1)= estimated density of *B* at time 1 day

B(1)= 107x10x100x100x100

B(1)= $1.07.10^9$ cells/ml

To estimate the relative fitness *W*, we can define m_A and m_B, as the realized Malthusian parameters respectively for strains *A* and *B*:

m_A= ln[A(1)/A(0)]/day

m_B=ln[B(1)/B(0)]/day

The relative (Darwinian) fitness *W* is the ratio of the Malthusian parameter:

$W= m_A/m_B$ = 1.29

During the competition, *A* increased about 29% faster than *B*.

The selection rate is equal to the difference between the two strains'Malthusian parameter during direct competition.

Hence the estimation of the selection rate, *r* is:

r= (ln[A(1)/A(0)]-ln[B(1)/B(0)])/day

r= 1.153 per day.

Over the course of one day of competition, the density of *A* increased by around 1.1 natural-logs more than the density of *B* did.

Many microbial evolution experiments were begun to explore the interesting dynamics that can emerge from complex interactions among several components of a larger system. There is a growing interest in evolution experiments that are carried out *in silico*.

From R. Lenski home page (http://myxo.css.msu.edu/ecoli/srvsrf.html)

2.2 Modeling biology

2.2.1 Modeling

Given the complexity of both ecological and evolutionary processes, it is necessary to make several simplifying assumptions in the development of theories. Modeling is an approach to assess complex natural phenomena using simplified assumptions in the development of theories.

2.2.2 Population genetics models

Mathematical population genetics addresses the question about gene frequency distributions in analytical form. Fisher, Haldane and Wright [3, 19, 20] together shaped the foundations of the field and are referred to as the 'great trinity' of population genetics. In particular they produced seminal papers in this field.

A primary focus of mathematical population genetics concerns fixation probabilities, which is the probability that a given allele in a population will ultimately invade the whole population. Empirically, fixation probabilities can be related to the rate at which a population adapts to a changing environment, the rate of loss genetic diversity or the rate of emergence of drug resistance. Analytical results may be derived under simple conditions to which more complex cases assessed through simulations should be compared.

There are three approaches to computing fixation probabilities.

When the population size is quite small, the state space of a population (exactly how many individuals have exactly which genotype) can be enumerated, using a markov chain approach to exactly determine the fixation probability (see Gale 1990) [21].

When the population size is large, methods based on discrete branching processes are often used. These methods build on the 'Haldane–Fisher' model [3, 19], and can be related to a Galton–Watson branching process. It provides an approximation to the true fixation probability, since it assumes that the wild type population is sufficiently large for the fate of each mutant allele to be independent of all the others [22-26].

Finally, when the population is large and the change in allele frequency per generation is small in each generation (i.e. weak selection), methods that incorporate a diffusion approximation may be used.

Haldane [19] showed that the probability of ultimate fixation, π, of an advantageous allele with selection coefficient s is $\pi = 2s$ when the allele is initially present as a single copy in a large and constant population size (and s is not too large). Other assumptions include discrete generations and a Poisson distribution of the number of offspring.

Kimura's approach was to use a diffusion approximation model. The main result is that the probability of ultimate fixation π of an allele with an initial frequency p is:

$$\pi \approx \frac{1 - e^{-4spNe}}{1 - e^{-4sNe}} \qquad (1)$$

Where N_e is the effective population size (the equivalent size of an 'ideal' population).

When $p=1/2N_e$, this probability of fixation tends to the previous result $2s$.

The same expression was obtained by Moran (1960) and Gillespie (1974) [27] using slightly different approaches.

Population of variable size:

- Growing and declining or cyclic population sizes

Ewens (1967) [28] used a discrete multiple branching process to study the survival probability of new mutants in a population that assumes cyclic sequence of population sizes that matches a typical experimental set up. Ewens derived corresponding ultimate probability of an initially unique mutant:

$$\pi \approx \frac{2s\tilde{N}}{\overline{N}} \qquad (2)$$

where $\tilde{N}$ is the harmonic mean and $\overline{N}$ is the arithmetic mean of the population sizes in the cycle.

Ewens found that the π value may be considerably less than $2s$, implying that a constant population size is favorable for the survival of new advantageous mutants whereas fluctuations slow adaptation.

Otto and Whitlock (1997) [22] later built on the work of Ewens and Kimura in populations modeled by exponential and logistic growth or decline.

Pollak (2000) [29] studied the fixation probability of beneficial mutations in a population changing cyclically in size and proved that the result found for Poisson-

distributed offsprings by Ewens (1967) [28] and Otto and Whitlock (1997) [22] still holds in situations where the Poisson offspring distribution is not necessary.

- Population Bottlenecks

A population bottleneck is a sudden, severe reduction in population size with subsequent recovery of greater population size. An important assumption to note is that after the population bottleneck, each individual—mutant or wild-type—is assumed to survive the bottleneck with the same probability. Thus the 'offspring' distribution of each individual at the bottleneck is the same, for either mutant or wild type, i.e., no selection occurs. In contrast, during growth any selective advantage for the mutant can be achieved.

Wahl & Gerrish (2001) [23] derived the probability that a beneficial mutation is lost due to population bottlenecks using both a branching process approach [3, 19] and a diffusion approximation [30-32]. When selection is weak, the two approaches provide the same approximation for the extinction probability X (X=1-π) of a beneficial mutation that occurs at time t between bottlenecks:

$$1 - X \approx 2srt\tau.e^{-rt} \qquad (3)$$

where s is the selective advantage of the mutant over the wild type strain, r is the Malthusian growth rate of the wild type population and τ is the time at which a bottleneck is applied. It was thus found that the fixation probability,π, drops rapidly as t increases, implying that mutations that occur late in the growth phase are unlikely to survive population bottlenecks, they have not enough time to increase in frequency and survive bottleneck.

Experimental techniques in microbial evolution have advanced to the point where the fate of a new mutant strain within a controlled population can be tracked over many generations.

2.3 Experimental evolution and modeling

Several experiments have provided some of the most relevant data for testing predictions about such branching points [33-35]. Most of them have been conducted using microbial cultures, such as *Pseudomonas* or *Escherichia coli*. Such model organisms are ideal for testing this theory because their short generation

times and well-defined genetic stocks make evolutionary experiments feasible. For example in single strain of *E. coli* propagating in glucose medium, evolutionary diversification eventually took place, resulting in the stable maintenance of two distinct physiological types [33, 36]. Similarly, in colonies of a single strain of *Pseudomonas* propagated on a complex liquid medium, evolutionary diversification eventually took place, resulting in three well-defined types that appear to coexist indefinitely (the “fuzzy spreader”, “wrinkly spreader” and “smooth” types; see [37]). These morphs appear to exploit spatial niches in the liquid medium.

Conclusion

In the absence of information concerning the selective regime, most simple models are classically used to study such processes: constant coefficient of selection across time, space, no frequency dependence, absence of interference between selective events, no interaction with demography. Natural conditions probably largely violate these assumptions, and the issue is then to assess the corresponding effect on genetic variability.

The occurrence of genetic drift is difficult to demonstrate *in natura* and to distinguish from selective effects. It is also difficult to test experimentally, because of the long time scales involved. Moreover, the results of evolutionary experiments are generally interpreted without taking into account stochastic processes and are often confounded with experimental error. Nevertheless, these stochastic phenomena can sometimes be the sole explanation for an observation. There is therefore relatively little data available on genetic drift (see however Buri 1956 [38]).

To study these evolutionary forces, the most direct approach, used here, relies on the analysis of temporal variations of allele frequencies [39]. However, one of the issues addressed in this thesis is to what extent such frequency variations are due to random phenomena, or to selective effects. Genetic drift leads to frequency fluctuations than can go either way thus allowing disentangling it from selection which typically is unidirectional. When an advantageous mutation appears in a genome, provided it escapes initial loss by drift, it tends to increase in frequency and eventually reaches fixation, thereby replacing pre-existing variation on the target locus. Furthermore, the genomic area in which it appeared is then swept of neutral mutations, which otherwise accumulate within non-functional DNA. This phenomenon, called selective sweep or hitchhiking, provides a powerful tool to detect rare selective events through their effect on the neighboring neutral loci [40]. During our experiments, the selective regime will be provided by a biotic factor of prey-predator type. In prey-predator interactions, environments often change quickly, and therefore impose a fluctuating selective pressure on its prey unlike the traditional assumptions of population genetics models. Similarly, the predator population should strongly affect the prey demography and vice versa.

The choice of the biological model is then primordial to perform evolution experiments. Such ecological factors will be included in the population genetics models. Since the influence of selection and genetic drift processes on diversity is complex, a relevant model is needed in order to provide quantitative predictions to be tested on the corresponding experimental data.

Literature cited

1. Metz, J., R. Nisbet, and S. Geritz, *How should we define [] fitness' for general ecological scenarios?* Trends in Ecology & Evolution, 1992. **7**(6): p. 198-202.
2. Dieckmann, U. and R. Ferrière, *Adaptive dynamics and evolving biodiversity.* Evolutionary conservation biology, 2004: p. 188–224.
3. Fisher, R., *The evolution of dominance in certain polymorphic species.* American Naturalist, 1930. **64**(694): p. 385-406.
4. Zhan, J., *et al.*, *Variation for neutral markers is correlated with variation for quantitative traits in the plant pathogenic fungus Mycosphaerella graminicola.* Molecular Ecology, 2005. **14**(9): p. 2683-2693.
5. Kimura, M., *The neutral theory of molecular evolution*. 1985: Cambridge Univ Pr.
6. King, J. and T. Jukes, *Non-darwinian evolution.* Science, 1969. **164**(881): p. 788-798.
7. Thoday, J., *Disruptive selection.* Proceedings of the Royal Society of London. Series B, Containing papers of a Biological character. Royal Society (Great Britain), 1972. **182**(67): p. 109.
8. Rousset, F., *Genetic structure and selection in subdivided populations*. 2004: Princeton Univ Pr.
9. Charlesworth, B., D. Charlesworth, and N. Barton, *The effect of genetic and geographics structure on neutral variation.* Annual Review of Ecology, Evolution, and Systematics, 2003. **34**(1): p. 99-125.
10. Gordo, I. and P. Campos, *Adaptive evolution in a spatially structured asexual population.* Genetica, 2006. **127**(1): p. 217-229.
11. Hill, W. and A. Robertson, *The effect of linkage on limits to artificial selection Genet.* Res, 1966. **8**: p. 269-294.
12. Gerrish, P. and R. Lenski, *The fate of competing beneficial mutations in an asexual population.* Genetica, 1998. **102**: p. 127-144.
13. Hartl, D. and A. Clark, *Principles of population genetics*. 1997: Sinauer associates Sunderland, MA.
14. Hall, A. and N. Colegrave, *How does resource supply affect evolutionary diversification?* Proceedings of the Royal Society B: Biological Sciences, 2007. **274**(1606): p. 73.
15. Elena, S. and R. Lenski, *Evolution experiments with microorganisms: the dynamics and genetic bases of adaptation.* Nature Reviews Genetics, 2003. **4**(6): p. 457-469.
16. Holland, J., et al., *Quantitation of relative fitness and great adaptability of clonal populations of RNA viruses.* Journal of virology, 1991. **65**(6): p. 2960.
17. Lenski, R., et al., *Long-term experimental evolution in Escherichia coli. I. Adaptation and divergence during 2,000 generations.* Am Nat, 1991. **138**: p. 1315.
18. Travisano, M., et al., *Experimental tests of the roles of adaptation, chance, and history in evolution.* 1995. **267**: p. 87-90.
19. Haldane, J. *A mathematical theory of natural and artificial selection. VI. Isolation*. in *Proceedings of the Cambridge Philosophical Society*. 1930.
20. WRIGHT, W., *S., 1931 Evolution in Mendelian populations.* Genetics. **16**: p. 97-159.
21. Gale, J., *Theoretical population genetics*. 1990: Allen & Unwin Australia.
22. Otto, S. and M. Whitlock, *The probability of fixation in populations of changing size.* Genetics, 1997. **146**(2): p. 723.
23. Wahl, L. and P. Gerrish, *The probability that beneficial mutations are lost in populations with periodic bottlenecks.* Evolution, 2001. **55**(12): p. 2606-2610.
24. Johnson, T. and P. Gerrish, *The fixation probability of a beneficial allele in a population dividing by binary fission.* Genetica, 2002. **115**(3): p. 283-287.

25. De Oliveira, V. and P. Campos, *Dynamics of fixation of advantageous mutations.* Physica A: Statistical and Theoretical Physics, 2004. **337**(3-4): p. 546-554.
26. Champagnat, N. and A. Lambert, *Evolution of discrete populations and the canonical diffusion of adaptive dynamics.* Annals of Applied Probability, 2007. **17**(5-6): p. 102.
27. Gillespie, J., *Natural selection for within-generation variance in offspring number II. Discrete haploid models.* Genetics, 1975. **81**(2): p. 403.
28. Ewens, W., *The probability of survival of a new mutant in a fluctuating environment.* Heredity, 1967. **22**: p. 438-443.
29. Pollak, E., *Fixation probabilities when the population size undergoes cyclic fluctuations.* Theoretical Population Biology, 2000. **57**(1): p. 51-58.
30. Wright, S., *The differential equation of the distribution of gene frequencies.* Proceedings of the National Academy of Sciences of the United States of America, 1945. **31**(12): p. 382.
31. Kimura, M., *Some problems of stochastic processes in genetics.* The Annals of Mathematical Statistics, 1957. **28**(4): p. 882-901.
32. Kimura, M., *On the probability of fixation of mutant genes in a population.* Genetics, 1962. **47**(6): p. 713.
33. Travisano, M. and P. Rainey, *Studies of adaptive radiation using model microbial systems.* Am Nat, 2000. **156**: p. S35-S44.
34. Rainey, P., et al., *The emergence and maintenance of diversity: insights from experimental bacterial populations.* Trends in Ecology & Evolution, 2000. **15**(6): p. 243-247.
35. Kassen, R., *The experimental evolution of specialists, generalists, and the maintenance of diversity.* Journal of Evolutionary Biology, 2002. **15**(2): p. 173-190.
36. Turner, P., V. Souza, and R. Lenski, *Tests of ecological mechanisms promoting the stable coexistence of two bacterial genotypes.* Ecology, 1996. **77**(7): p. 2119-2129.
37. Rainey, P. and M. Travisano, *Adaptive radiation in a heterogeneous environment.* Trans. R. Soc. B, 1977. **280**: p. 29-101.
38. Buri, P., *Gene frequency in small populations of mutant Drosophila.* Evolution, 1956. **10**(4): p. 367-402.
39. Nei, M. and F. Tajima, *Genetic drift and estimation of effective population size.* Genetics, 1981. **98**(3): p. 625.
40. Maynard Smith, J. and J. Haigh, *The hitchhiking effect of a favorable gene.* Genet Res, 1974. **23**(1): p. 23-35.

Objectives

The main objective of this thesis was to study the effect of natural selection and genetic drift on neutral polymorphism using an interdisciplinary approach incorporating experimentation with a bacterium, *Escherichia coli,* and an amoeba, *Dictyostelium discoideum,* and a theoretical model comparable to the experimental set up.

To achieve this objective, we developed and used various experimental and theoretical setups to address the following questions:

A first phase of studies sought to elucidate how much natural variability is there in ecological interactions between different *E.coli* strains and *D. discoideum*.

What are the bacterial traits responsible for this variation?

Then, we focused more on following aspects of population genetics:

- Studying the effects of spatial structure and of biotic factor on polymorphism. We measured temporal allele frequency variations, using a microbial system in four environmental conditions: with or without agitation and with or without biotic factor, taking demography into account. We carried out serial transfers, which allowed the populations to be maintained, and to increase the number of generations.
- Development of a theoretical model for studying the effects of demographic stochasticity on neutral allele fixation probabilities and the time to fixation in a first step, followed by study of combined demographic stochasticity and selection (work in progress).

3. Competition, antagonistic interactions and balancing selection

In our work, we were interested about selective forces acting during evolution and in ecological interactions on biological systems. Theoretical and empirical studies have shown that some ecological interactions such as competition and antagonistic interactions play on, promote, and maintain diversity.

Natural selection has played an important role in many evolutionary diversifications.

During the process of divergent selection, natural selection acts in contrasting directions between environments, which drives the fixation of different alleles that are advantageous in one environment but not in the other. Selection arising from environmental differences or ecological interactions acts in contracting directions on two populations and favors traits within populations.

Balancing selection, so called because many alleles are maintained in balanced, play also an important role in diversification. Nevertheless, the role given to balancing selection in maintaining polymorphism under ecological interactions is still unclear.

In this part, we highlight initial findings, namely experimental and theoretical studies related to the role given to balancing selection during ecological interactions.

Competition, antagonistic interactions and balancing selection

Sandrine Adiba[1,2*], Minus van Baalen[1,2] and Frantz Depaulis[2]

[1]Laboratoire d'Ecologie, CNRS UMR7625, Université Pierre et Marie Curie, Paris Universitas, 7 quai St-Bernard, 75252 Paris, France

[2]Laboratoire d'Ecologie, CNRS UMR7625, Ecole Normale Supérieure, 46 rue d'Ulm, 75230 Paris cedex 05, France

(In preparation)

Abstract

A central goal of evolutionary ecology is to understand the role of various ecological processes in producing patterns of evolutionary diversification. Balancing selection, so called because many alleles are maintained in balance, means either overdominance, frequency dependent selection or diversifying selection; it can promote polymorphism. To determine which of these types of selection is involved in ecological complex systems is a tough task that often requires diverse approaches. Theoretical and empirical studies have shown that competition and antagonistic interactions, such as prey/predator or host/parasite interactions can promote diversification or maintain biological diversity. Nevertheless, the role assigned to balancing selection by maintaining genetic polymorphism under ecological interactions remains unclear. Assessing its contribution requires a combination of approaches. This review highlights initial findings, and recent experimental and theoretical studies related to the role of balancing selection during ecological interactions such as competition and antagonistic interactions. Competitive interactions may increase or decrease diversity, which renders predictions about these patterns challenging. Several studies suggest that frequency dependent selection on one or the other partner involved in antagonistic interactions may maintain genetic variations. Related issues such as the nature of selection inducing MHC polymorphism, remain a matter of debate.

Abbreviations:

FD(S): Frequency Dependent (Selection)

NFDS: Negative Frequency Dependent Selection

ESS:Evolutionary Stable Strategy

MHC: Major Histocompatibility Complex

Introduction

In this year marking the bicentenary of Darwin and Wallace's initial presentation of the Theory of Natural Selection, a current issue for evolutionary biologists remains understanding how such an amazing level of evolutionary diversification can have taken place throughout life history of life [1]. Although Charles Darwin certainly recognized the origin of evolutionary diversification (i.e. species), he did not produce a fully compelling explanation for it [2].

Biological interactions result from the fact that organisms in an ecosystem interact with each other, in the natural world, no organism is an autonomous entity isolated from its surrounding, but is part of its environment. Some ecological interactions, such as competition and antagonistic interactions played a central role in diversification. Competition is defined as an interaction between individuals or populations that is mutually detrimental. Whereas in antagonistic interactions, one partner benefits at the expense of another, this definition covers all types of interactions including those in which one organism eats another, regardless of trophic level and closeness of association, including host/parasite interactions. There is mounting evidence that natural selection and in particular balancing selection, has played an important role in many evolutionary diversifications.

By balancing selection, population geneticists mean either overdominance (heterozygote superiority in fitness), frequency-dependent selection (**FDS**), or diversifying selection that favors different genotypes in different environments [3] (Figure I-4).

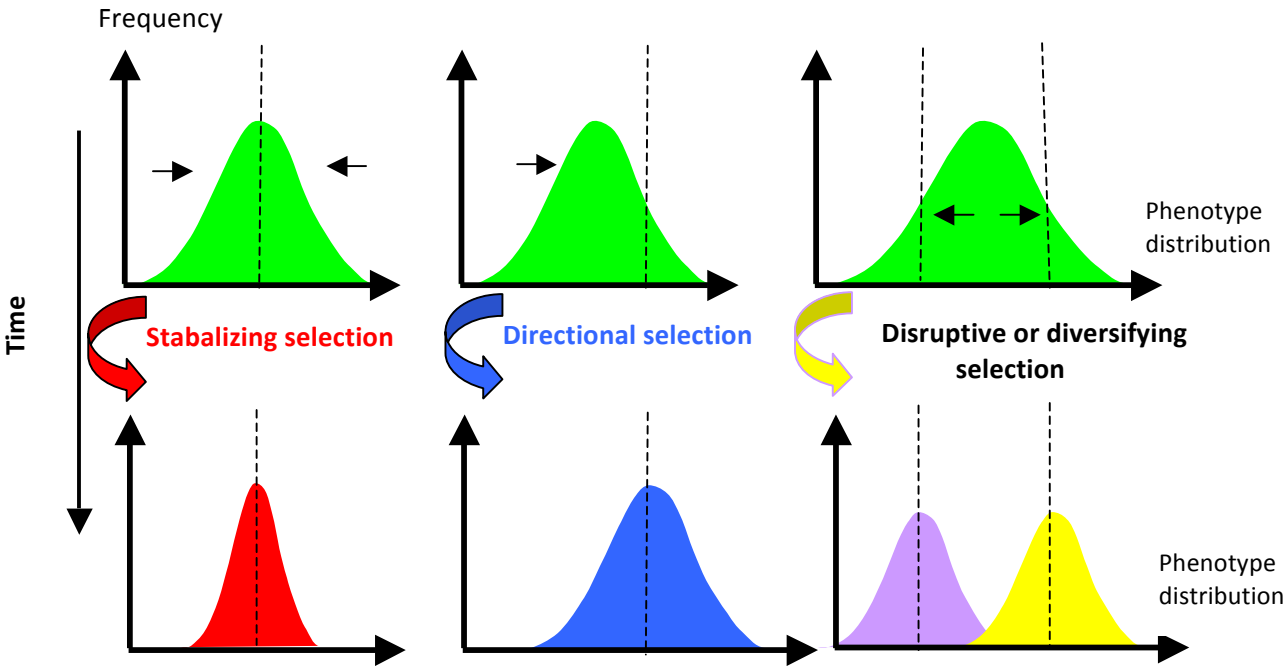


Figure I-4 : The three basic types of selection
Dotted traits indicate the optimum expressions of the phenotype. Horizontal arrows show the directions of selective forces.

The definition of FDS in its classical term applies to an evolutionary process where 'the relative fitness' of a type varies with the relative frequency to other types in the population [4]. But classical definitions are broad and can be related to:

- frequency of the genotypes in the population [5]
- fitness varying as a function of the frequencies of diploïd genotypes [4]
- fitness depending on the relative frequencies of other genotypes [6]
- fitness strategy according on its frequency relative to that of other strategies through fixed total population size
- fitness varying to favor rare types (Negative Frequency Dependence Selection: **NFDS)** [4].

Furthermore, a long standing debate exists concerning the role of balancing selection in maintaining genetic polymorphism, which was frequently used to explain genetic variation [7]. The neutral theory of molecular evolution provided a competing explanation for genetic polymorphism, but how common balancing selection really is remains unclear.

The twin realization that disruptive selection (also called diversifying selection) (Figure I-4) can persist over substantial periods of time and that many ecological

scenarios produce just this type of selection regime has triggered a massive effort among theoreticians to explore evolutionary consequences of such scenarios. The consequence that has attracted most attention is the phenomenon called evolutionary branching of a lineage [8,9,10], which allows the possibility of sympatric speciation [11]. So called 'adaptive dynamics' has arosen as a recent mathematical theory of the evolution of complex systems, the objective of which is to elucidate the long-term effects of the interactions between ecological and evolutionary processes. Evolutionary branching occurs when FDS splits a phenotypically monomorphic population into two distinct phenotypes. In this approach, ecological interactions are the evolutionary driving force in considering feedback between evolutionary change and the ecological conditions experienced by individuals. Disruptive selection, simultaneously favoring individuals at both extremes of the distribution, is not an externally imposed assumption: the system converges to a state, in which it experiences disruptive selection. Such events can occur under a wide range of ecological interactions. Simplifying assumptions are used to analyze these models, namely rare mutations, haploid and monomorphic resident population, and ecological equilibrium. This corresponds to separate ecological and evolutionary time scale, with the ecological dynamic occurring faster than the evolutionary dynamics. Invasion fitness [12] is determined by the phenotype of the resident population, reflecting the consequence of frequency dependant ecological interactions. The actual mechanisms by which a phenotypic cluster of individuals splits into two distinct descendant clusters are poorly understood. There are two traditional ways to understand this evolutionary process. The first one is the geographic isolation, after which the two populations follow separate evolutionary paths, such allopatric speciation, is quite well understood (see Rice W.R. and Hoster E.E 1993) [13] and not covered in this paper. The second traditional approach investigates the conditions under which speciation may occur in sympatry, which is when gene flow is possible between two species. The difficulties in the theory of sympatric speciation are twofold [14]. On the one hand, ecological conditions must induce disruptive selection. On the other hand, given such ecological conditions, the mating system must evolve to the reproductive isolation between the phenotypes that are favored by disruptive selection.

In this review, we address which type of ecological interactions give rise to balancing selection and then diversification. We consider competition and antagonistic interactions. Indeed, these ecological interactions played a central role in early ideas on the origin and maintenance of biological diversity. Furthermore, NFDS is often the result of these interactions (see summary table at the end of this paper). There is mounting evidence that natural selection has played an important role in many evolutionary diversifications and therefore it is of interest to know what ecological process causes natural selection to favor diversity. We summarize some historical progress on this topic, highlight some steps that recent evolutionary theory has made toward addressing these issues and discuss different ecological interactions about the role assigned to balancing selection generating or maintaining genetic polymorphism or diversity.

Throughout, we adopt Mayr's (1942) biological species concept, in which species are groups of actually or potentially interbreeding natural populations that are isolated reproductively from other such group.

Competition for resources

In the 1950s and 1960s, early ideas of evolutionary diversification were developed with the concept of ecological opportunity based on the notion that the environment is made up of different ecological niches (consumable resources and abiotic factors like temperature, humidity, etc.) [15]. It turns out that there are countless ways in which evolutionary diversification might fill up available niches. It was highlighted by two related ideas: the competitive exclusion principle, which states that no two species with the same niche can stably coexist [16,17], followed by the principle of limiting similarity [18,19,20] which states that the more similar two species are in their resource requirement, the more intensely they will compete (and then exclude each other though competition for resources). Furthermore, in a co-evolutionary context, each species affects the resource distribution in the environment and thus modifies the selective regime affecting its own evolution as well as that of other species, a process known as eco-evolutionary feed-back in particular in adaptive dynamics. Indeed, each species affects the evolutionary trajectory of competing species. This conceptual statement could be related to the theory of character displacement developed by Brown & Wilson (1956) [21],

Hutchinson (1959) [1], and MacArthur & Levins (1964) [17], prior to the concept of limiting similarity. It should result in phenotypic characters related to resource use in similar species being displaced from one another where their geographic ranges overlap.

Competition promoting diversifying selection has been observed in many instances under natural and experimental conditions. The corresponding review is the subject of the first part of our study. Several models have been developed, and more recently, adaptive dynamics theory has attempted to link ecology, evolution, population genetics, life history traits and game theory. Then, in theory, disruptive selection arising as a consequence of frequency dependent processes (e.g., competition for limiting resources) can result in adaptive diversification, that is, the splitting of an ancestral lineage into distinct descendent lineages as a consequence of natural selection acting on individuals within the population [22,23].

Empirical observations and experimental evolution

Circumstantial evidence for evolutionary branching in sympatry comes from the repeatable trophic polymorphisms observed in numerous species of postglacial fishes [24,25,26,27] and from studies concerning the freshwater three-spine stickleback fish (*Gasterosteus aculeatus*). Indeed, in a few lakes in the Strait of Georgia, southwestern British Columbia, two species of freshwater three-spine stickleback have evolved morphological adaptations for consuming two different resources. The first species, 'benthic' is morphologically and behaviourally specialized for benthivory, the second species, 'limnetic' is specialized for planktivory. Both comparative [28] and direct experimental evidence [29] indicate that these differences are the result of natural selection caused by frequency-dependant resource competition. The scenario of 'double-invasion' proposed by McPhail (1993) (Figure I-5) [30] is the result of two separate invasions, in which the second invader evolved into the modern limnetic, isolated reproductively and behaviorally [31].

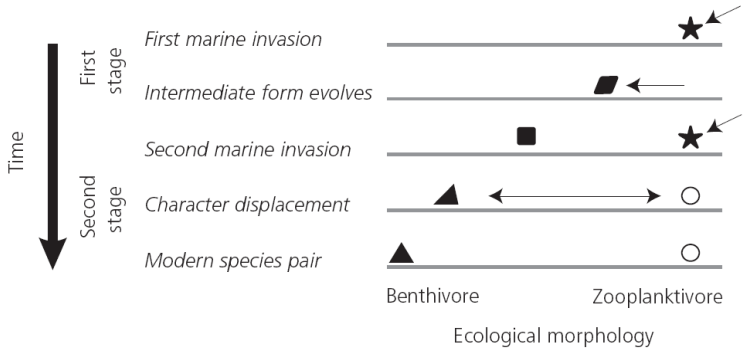


Figure I-5: Double invasion and stickleback speciation [30]

These divergent phenotypes observed are the evolutionary result of competition for resources and character displacement between the two groups of colonizing marine stickleback [28]. An alternative hypothesis is selection on premating isolation in sympatry.

Further empirical evidence for sympatric speciation via evolutionary branching arose from studies on organisms such as intertidal snails [32], Anolis Lizards [33] in which resource competition is likely to have been a major driving force toward speciation.

Other evidence for sympatric speciation comes from evolution experiments using microbial systems. They are useful for testing ideas of adaptive diversification because microbes have large population sizes, short generation times and are easy to culture in laboratory, as well as because evolved and ancestral genotypes can be frozen at -80°C for later comparisons [34]. Tyerman et al. [35] used 36 lines of *Escherichia coli B* evolved with daily serial transfers for 1,000 generations in glucose and glucose-acetate environment. In both environments, the ancestor had repeatedly diversified into two types, identified by the size of the colony: large (L) and small (S). This surprising finding of stable polymorphism appeared at first to contradict the competitive exclusion principle, but it was subsequently evidence there was resource partitioning of the limiting glucose resource by the numerically dominant genotype. Large (L) and Small (S) colonies correspond to two ecological types, a glucose specialist and an acetate specialist which were maintained by NFDS [36]. Competition experiments between S and L types carried out in various environments suggest that diversification in one trait (colony size) may be robust across environment: NFDS maintains the

balance between the different types specializing in different resources when the two partners are derived from the same evolutionary environment. But diversification in other traits (i.e., competitive ability) may not be robust: competitive exclusion occurring in one outcome and oscillatory dynamics in three others were observed when partners were derived from different evolutionary environments. Another evolution experiment, using twelve initially identical populations of *Escherichia coli*, founded in 1988 and evolved in a glucose-limited medium that also contains Citrate [37] showed that NFDS causes ecological diversity. A Citrate-using (*Cit*$^{+}$) variant evolved in one population by 31,500 generations resulting in some diversity. The *Cit*$^{+}$ phenotype did not achieve fixation during the population expansion, but instead, *Cit*$^{-}$ cells persisted as a minority. Different initial frequencies of *Cit*$^{+}$ and *Cit*$^{-}$ clones were mixed and stably coexisted over many generations. The *Cit*$^{-}$ cells gradually approached an equilibrium frequency of 1% of their initial frequency. This stable coexistence suggests that *Cit*$^{-}$ cells are superior to the *Cit*$^{+}$ cells in competition for glucose, allowing the former to persist as glucose specialist. Thus, the overall diversity increased with two members, one a resource specialist, and the other a generalist.

The work of Adams et al. [38], who performed long-term selection experiments with chemostat-propagated populations of *E.coli* grown on a single limiting carbon source, showed that polymorphism occurs during competitive interactions among genotypes. There was also evidence for resource partitioning of the limiting glucose resource by the numerically dominant genotype [39]. Diversity can also be maintained through Genotype X Genotype interactions that occur over small spatial scales and where strict competitive hierarchies are absent [40]. A number of studies using microbial organisms appear well adapted testing the role of NFDS in producing evolutionary branching points through disruptive selection [41,42,43]. Indeed, experimental evolution allows for the observation of a process otherwise inferred *a posteriori*. These results are challenging and have provided empirical data that are consistent with recent theoretical predictions (see, e.g. ref. [44]). Evolutionary diversification occurred as expected, and NFDS was strongly supported in those experiments.

Theoretical models were developed to study the mechanism by which evolutionary diversification into distinct phenotypes may occur in sympatry. The aim is to provide an evolutionary explanation for how selection favors the origin of different species and how they coevolve. The extension of earlier theory on ecological character displacement considered how diversifying selection acts on phenotypic variation on an intraspecific level [45]. Previous theory and more recent theory of adaptive dynamics show interesting relationship between results. The conditions under which evolutionary branching points occur are the same as those under which evolutionary divergence and coexistence occur in two-species or multiple-species models of interspecific competition, ecological character displacement [46,47], symmetric competition, and asymmetric competition [8].

Symmetric competition occurs when individuals with similar phenotypes compete more strongly with each other than individuals with divergent phenotypes do. Asymmetric competition occurs when the competition effect between similarity and competition level is stronger with any other kind of relation [8] (Figure I-6).

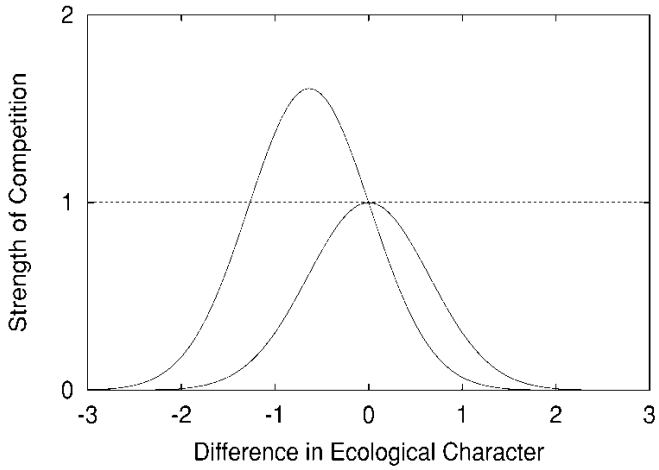


Figure I-6 : Strength of competition as a function of phenotype difference between competitors

The function describing symmetric competition is a symmetric function with a maximum at 0. The function describing asymmetric competition has a maximum for some negative difference in character values [8].

The model developed by Michael Doebeli and Ulf Dieckmann [8] based on classical Lotka-Volterra population dynamics with competition show that symmetric and asymmetric competition lead to evolutionary branching whenever

the frequency dependence is strong enough. Another simple model for asymmetric competition, also based on Lotka-Volterra population dynamics, was developed by Eva Kisdi [48]. The evolution of polymorphism by small mutational steps is possible under certain shape functions (concave or convex) of the intrinsic growth rate and the competition coefficient. Depending on a set of conditions, evolutionary branching, repeated branching, and extinction may be observed. But the author concluded that it is difficult to state that competitive asymmetry facilitates evolutionary branching in general. Finally, identifying the critical ecological factors that define whether asymmetric competition can lead to evolutionary branching should be the subject of further investigation. Similar results were obtained with a model for the evolution of seed size in plants, including evolutionary branching, repeated branching and extinction [49]. Strong competitive asymmetry and high resource level favor coexistence of plants with different seed size.

The asymmetric competition model describes by Law, Marrow and Dieckmann[50], links processes operating an individual scale to that of macroscopic evolution. For example, assessing phenotypic distribution across species, from microscopic encounter between individuals (depending on mechanisms of competition) and mesoscopic population dynamics (describing how the number of individuals in each population is affected by competitive encounters). This model shows that asymmetric competition between species can in principal show various evolutionary outcomes, including coexistence of species, because large individuals suffer fewer disadvantages when competition is asymmetric. But the smaller species can drive the larger one to extinction.

Considering sexual reproduction, frequency-dependent competition for unimodal distribution of resources can lead to sympatric speciation if mating is assortative [51]. But to make this genetic theory more general, it is rather advisable to consider models in which assortative mating is not based on the characters that determine ecological interactions but on selectively neutral markers. Such a more general model was developed by Dieckmann and Doebeli (1999) [22], with demographic stochasticity (resulting genetic drift) included in the description. They show that evolutionary branching is a robust phenomenon in sexual populations even when assortative mating is based on neutral marker traits.

The theory of adaptive dynamics is quite recent, and therefore there are still few explicit tests of its prediction. Nevertheless, coupling data analyses and the framework of adaptive dynamics can be a way to test evolutionary scenarios. This approach should provide a powerful way to test the double-invasion scenario of stickleback speciation [30]. The corresponding model tests the possible role of frequency-dependence in this simple scenario. The result shows two possible scenarios, with an intermediate stable attractor (when the population is subjected to stabilizing selection), or an intermediate branching point (where the population experiences disruptive selection). Experimental research [29] as well as assortative mating studies [52,53] support the theoretical prediction that size assortative mating could have led to a sympatric split into large and small forms that result from competition for resources.

Such competitive interactions have the potential to increase or decrease diversity, with the outcome being dynamic and challenging to predict. Other examples of co-evolution come from studies of predator-prey and host-parasite interactions. It is worth considering whether a pattern of NFDS is sufficient to rule out other possible ecological interactions as explanations for the diversification.

Prey-Predator interactions

In this paper, we use the classical definition of predation which is the consumption of one animal by another. However, predation may be broadly defined to include all transfers of energy from one organism to another, including herbivores, parasites, parasitoids and pathogens. One of the differences between competition and predation is that the predator cannot survive without its prey. In this section, we review studies related to the classical definition of predation, and studies of host/parasites interactions that can lead to some co-evolution process in the next. The term divergence is used to refer to increased differences between related species in various aspects of their phenotypes. The role of predation in biological evolution has received less attention than other ecological factors such as competition and host-parasite interactions. However, predation should represent a primary source of evolutionary changes. Its potentially key role in phenotypic divergence, speciation and diversification rate, even in the absence of interspecific competition for resources, has only recently attracted significant

attention and remains controversial [54]. Controversy arises from the fact that predation is difficult to investigate in the wild. Furthermore, phenotypes generally reflect the influence of multiple selective agents in addition to those of predators. The effects of a given predator on prey phenotypes may then be complex to differentiate. In natural populations, predators may also affect the selective regime of their preys by altering their interactions with other parameters, such as competition levels and other alternative predation pressures. Predators can drive phenotypic differentiation among prey populations by a number of different mechanisms [55]. Fitness trade-off in prey traits may depend on the specific predation regime (Figure I-7A), competition for 'enemy free space' (Figure I-7B), and predators altering the interactions of prey with other selective agents (Figure I-7C). Among all of these paths of diversification, divergent selection is the common issue and is responsible for phenotypic differences in each case (Figure I-7D). Predators can drive divergent selection on prey traits within predator regimes because predation is heterogeneously distributed in space and time and because many traits exhibit trade-offs over a range of predatory environments.

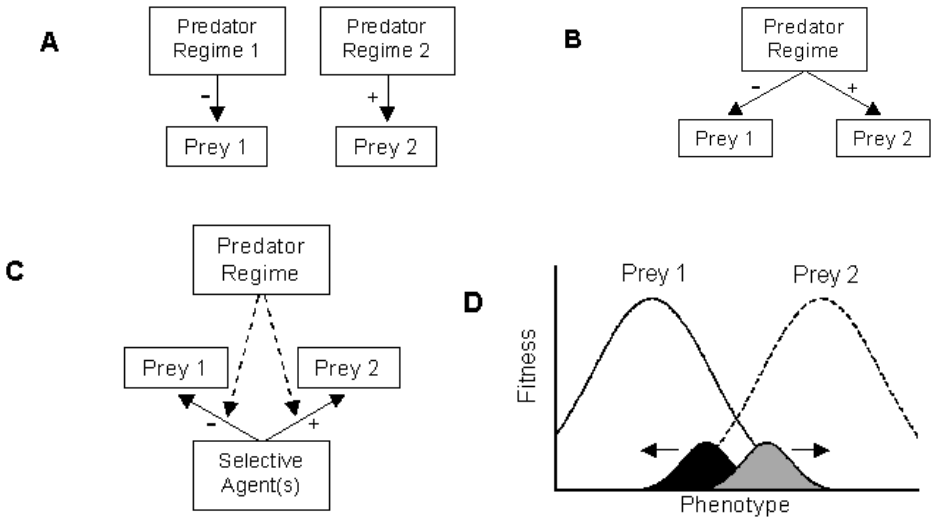


Figure I-7: Three common ways by which predators can drive phenotypic divergence among prey populations (or species).

Solid arrows depict selection on a prey trait.

The sign next to each arrow indicates the direction toward the optimal phenotype.

(A) Divergent selection arising from divergent predator regimes

(B) Competition for enemy free space within a given predator regime

(C) An interaction between predation and the prey's selective regime.

(D) Possible fitness functions arising from any of the above scenario

Shaded areas represent transient trait distributions for the two prey populations. The arrows illustrate the direction of selection [55].

The aim of the next sub-part of this paper is to describe a number of experiments that investigate prey-predator interactions, in particular with respect to the contribution of balancing selection. We then attempt to compare these results with those obtained using theoretical models.

Empirical investigations

To explore the contribution of predation to character divergence, Rundle and collaborators [56] tested whether predation alters the effect of competition on divergent selection. Empirical studies support the idea that predation reduces competition by increasing mortality (at least over a short term) [57], thereby weakening divergent selection. But given the diversity of ways in which predation may affect divergent selection between competitors (resource partitioning, anti-predator behaviour...), the level of divergent selection does not necessary increase with increasing competition. Their experiments show that divergent selection on the target population tends to be stronger in the predator-added treatment than with predation-free treatment. These results indicate that predation and other agents of mortality may facilitate trait displacement.

Meyer and Kassen [58] use *Pseudomonas fluorescens* microbial system to provide the evidence for the role of predation in promoting adaptive evolutionary changes through diversifying selection. *P. fluorescens* can propagate in spatially structured environments, diversifying and generating three niche specialist types: smooth morphotype (SM), which spreads in the liquid phase, wrinkly spreader morphotype (WS), which form a biofilm at the air-broth interface, and fuzzy spreader (FS) morphotype, which remains in the less aerobic bottom of vials [59] (Figure I-8).

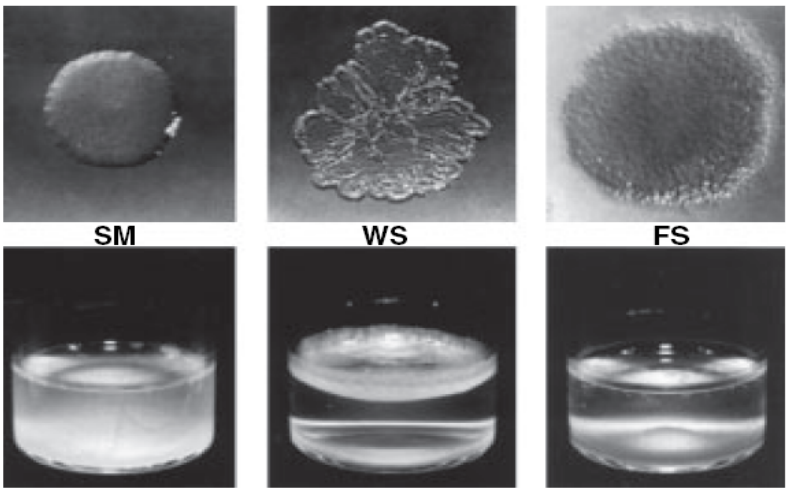


Figure I-8: adaptive radiation in *Pseudomonas*
Colony morph variants.
Figures reproduced from Rayney & Travisano [60].

The authors tested the effect of a predator (the protista *Tetrahymena thermophila*) on diversity shifts among bacterial genotypes. They estimated the frequency-dependent fitness functions of competing niche-specialist prey in the presence and in the absence of predation. They observed prey diversification with competition and predation independently generating diversifying selection. When faced with both competition and predation, bacterial ecomorphs showed NFDS, in which the fitness of a genotype decrease with its frequency in the population, thus maintaining several ecomorphs. In the absence of predator, FDS seemed to be promoted by competition for resources, while in the presence of predator FDS seemed to result from refuge from predators via a floating bacterial mat. But the most challenging of this study is that diversification of bacteria was delayed in the presence of predators. The reason for this seems to be that predation reduces the intensity of resource competition among bacterial ecomorphs, hence in the corresponding level of diversifying selection. All in all, the results suggest that predation may play a major role in adaptive radiation.

Another study [61] on *Timema* stick insects also tests the role of predation. Their findings provide another piece of experimental evidence that predation may drive phenotypic divergence through adaptive radiation. Results provided similar support for the role of competition and especially predation in promoting adaptive radiation. These findings contrast with previous studies that demonstrated that the

main effect of predation during adaptive radiation is to influence the levels of resource competition by reducing population density through mortality [56,59].

Experimental studies of predation and mechanisms of selection involved have however received limited attention and should greatly benefit from further investigations.

Theoretical studies

One of the first theory proposed to account for the effects of predators on diversification was a verbal and graphical approach [62]. The role of predators was then investigated in the field of coexistence and habitats [63]. Brown and Vincent [64] used evolutionary game theory to investigate the influence of predators on the diversity of prey species. Abrams [65] also employed an Evolutionary Stable Strategy approach but instead considered how the anti-predator traits of a local prey species evolved in response to the addition of a second prey species. They concluded that frequency-dependent predation could lead to the coexistence of multiple predator and prey species in evolutionary stable states. But they included FDS competition among the prey as an additional diversifying agent.

All these models suggest that sympatric prey may experience trait shifts as readily from interaction with predators as from interactions with competitors.

Doebelli and Dieckmann [8] developed an extension of the classical Lotka-Volterra predator-prey models involving an interaction parameter. This model describes predation efficiency as depending on two quantitative characters, one in the prey and the other one in the predator such that the more prey and predator traits are similar, the stronger the interactions.

They did not involve frequency-dependence in the competitive interactions among the prey. The FDS on the prey results from the presence of the predator.

When the frequency-dependence is strong enough, the singular point will be a branching point for the prey character while the predator does not experience evolutionary branching.

This model demonstrated that evolutionary branching in the prey trait is most likely when predation efficiency decrease sufficiently with increasing distance between prey and predator traits. A primary evolutionary branching in the prey may for instance promote a secondary branching in the predator, but this outcome depends on the details of the system and in particular on the divergence from the

(ancestral) branching state. If these two phenotypic branches further diverge, for example toward increased width of resource distribution and increasing predation efficiency, a secondary branching predator is induced, with each predator predating specific prey.

To introduce some population genetics in this model, the quantitative trait running ecological interactions are genetically coded by many equivalent, additive, diploid and diallelic unlinked loci. Evolutionary branching occurs in sexual multilocus versions of this model above some threshold of assortative mating [8].

To put populations in a more relevant biological context, models with explicit carrying capacity were developed. In such cases, individual and populational demographic parameters arise. Bowers et al. [66] demonstrated that evolutionary branching in a predator-prey system arose only if the carrying capacity of prey was an emergent property of the process (in response to crowding) rather than if an explicit carrying capacity was imposed on the prey. Branching (i.e., phenotypic divergence) can readily occur when there are trade-offs between predator avoidance and growth rate.

Many questions remain concerning the detailed nature of predation's role in generating diversity of phenotypes and species, how reproductive strategies might reflect anti-predator adaptations, and how predation might influence trait divergence and speciation. Such issues should greatly benefit from further theoretical, experimental and empirical investigations in particular employing a pluridisciplinary approach.

Another kind of antagonistic interaction concerns host-parasite interactions. Parasitism reveals a paramount effect on host survival and reproduction. Such interactions may lock host and parasite population into a runaway evolutionary process, but how these dynamical interactions rule polymorphism remains a matter of debate.

Host-Parasite interactions

It is clear that parasites affect host population sizes but the nature of interactions between hosts and parasites remains unclear. What is known is that

during host-parasite interactions, reciprocal natural selection plays a role on the evolution of host resistance and parasite infectivity. An apparent paradox results from the fast evolution of parasites, generally with short generation times and high mutation rates, compared to their host. The impact of parasites on the genetics of host populations is thought to have broad evolutionary implications. In particular, models [67,68,69] suggest that parasites are selected to overcome the resistance of common hosts. Consequently hosts with rare resistance genes have selective advantage. Selection may thus be frequency-dependent. This FDS on host and parasite may maintain genetic variation and the selective advantage of rare variants may promote the production of diverse offspring with rare resistance genes [67].

The other selective force which may maintain polymorphism via parasitism is overdominance. For instance, this may account for the outstanding polymorphism of the vertebrate immune system. The Major Histocompatibility Complex (MHC) loci in humans show high level of polymorphism. Some studies [70] invoke selection favoring MHC heterozygous host (overdominance) and others [71] involve selection for rare MHC alleles arising from host-pathogen co-evolution (FDS).

In this third and last part, we shall detail various related experimental studies, and then focus on the controversies concerning the types of selection involved in MHC polymorphism.

Experiments testing selection type

Carius *et al.* [72] tested for the presence of genetic variation in resistance and infectivity in populations of the freshwater crustacean *Daphnia magna* and its bacterial microparasite, *Pasteuria ramosa*. It was a complete cross-infection experiment. Reciprocal selective pressure acting in this association is likely to be strong because infected hosts are sterilized and parasites that cannot successfully invade the host produce no transmission stages at all. They observed genetic variation among bacterial isolates' infectivity to its host. Furthermore, clones possessed specific resistance to certain parasite isolates and bacteria isolates possessed host clone-specific infectivity. But selection favors parasites that are able to infect common host clones. These findings suggest that FDS can operate in this system and that the co-evolution may be oscillatory, as per the Red Queen Hypothesis.

Microorganisms also allow studying the role of parasites in sympatric (within population) and allopatric (between the populations) diversification. But co-evolution in most bacteria-phage systems appears to be limited to one or two cycles of reciprocal evolution of resistance and infectivity, for example, *E. coli* and *T-phage* [73] or *Pseudomonas aeruginosa* and *PP7* [74,75]. More persistent cycles of antagonistic coevolution were observed in several bacteria-phage systems [76]. The first studies testing coevolutionary theory to be conduct over coevolutionary time scale involve the interaction between *P. fluorescens SBW25* and the lytic phage *phi2*[77,78,79]. A virulent phage life-cycle results in bacterial death; invasion of bacterial cells, replication and lysis, hence impose a strong selection for bacterial resistance. In the absence of phage, sympatric diversity was greatly reduced. Phage reduced host density and thus resource harvesting, thereby reducing competition and differentiation. In contrast, allopatric diversity was greatly increased as a result of phage-imposed selection for resistance, which causes populations to follow divergent evolutionary trajectories. The explanation may involve a selection for resistance to phages. Long-term selection experiments have shown that antagonistic coevolution between *P .fluorescens* and phage *phi2* is persistent over an evolutionary time scale (50 transfers or roughly 400 bacterial generations) [59] leading to multiple rounds of reciprocal selection for resistance and infectivity. In this process, selection was mostly directional. The observed coevolutionary dynamics are broadly consistent with asymmetrical multilocus gene-for-gene host parasite interaction that allows for the evolution of generalists [76, 80]. Nevertheless, in heterogeneous environments, the absence of phage reduces diversity. This arises because phage causes a reduction in bacterial density, weakening NFDS [81]. In contrast, in homogeneous environments, where diversity is otherwise low, phage increases diversity as a result of fitness trade-offs between bacterial resistance and competitive ability [81].

However, the generality of the evolutionary patterns observed in this system may be somewhat limited to host-parasite systems that undergo a degree of directional selection. Such systems include some plant-pathogen interactions [82] that broadly comply with a multilocus gene-for-gene model of coevolutionary interaction, which allows for the evolution of generalist host and parasite types [80]. By contrast, in many host-parasite systems, infection relies upon highly specific

matching of host and parasite genotypes and selection is predominantly fluctuating (i.e., matching allele interactions) [83]. Furthermore, lytic phages are obligate killers and therefore do not form chronic associations with their hosts, as do most classic parasites. It is therefore challenging to define 'virulence' for such system.

Nevertheless, parasites may play a key ecological role promoting diversity, by initially mediating coexistence between competitors. Such a most interesting finding of Morgan *et al.* [84] demonstrated that parasites allow for the coexistence of two genotypes of bacterium *P. fluorescens*, one resistant and the other sensitive to the phage. In the absence of phage, the resistant genotype was rapidly driven to extinction in all populations. In the presence of the phages, the resistant genotype persisted.

Theoretical studies suggest that the effect of FDS on host and parasite may be to maintain genetic variations. Matching-alleles models have demonstrated the potential for FDS and selection for sex [85]. In contrast, gene-for-gene interactions do not seem to favor FDS and selection for sex because a single overall superior genotype may go to fixation [86].

Host-parasite interactions and MHC polymorphism

The major histocompatibility complex (MHC) is a large gene family or genomic region found in most vertebrates and plays an important role in the immune system. Genes encoding MHC molecules are among the most polymorphic genes known in vertebrates. Two selective pressures are thought to be involved: selection favoring MHC heterozygous hosts and selection for rare MHC alleles by host-pathogen coevolution (FDS) [87]. For instance Doherty and Zinkernagel [88] argued that in population exposed to an array of pathogens, heterozygous individuals are favored because they are able to present a broader array of antigens, therefore resist a broader array of pathogens than homozygotes. The mechanism of NFDS hypothesis was first developed by Clarck and Kirby [87]. This model proposes that parasite antigenicity is selected to avoid an MHC-based immune response of the most common parasite genotype.

The roles of these two mechanisms have been debated for decades [89,90] and we should provide some of their features.

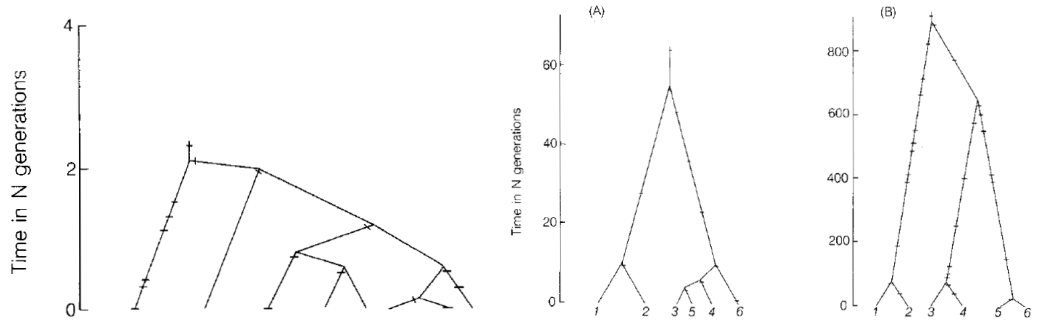
A commonly held view is that the polymorphism is due to selection favoring MHC-heterozygous hosts [89,91]. Different MHC molecules bind to different peptides

MHC heterozygous hosts may then present a greater variety of peptides and thus defend themselves against a wider variety of pathogens. Several methods have been used to test for selection acting on MHC polymorphism.

The Ewens [92] –Watterson [93] homozygosity test of neutrality is based on the fact that the effect of balancing selection can be detected by comparing the observed allele frequency distribution with the homozygosity expected under neutrality.

Over larger evolutionary time scale, Hill and Hastie [94] and then Hugues and Nei[95] proposed the *dN/dS* ratio test of positive selection (where *dN* and *dS* are respectively non-synonymous and synonymous substitution rates), but there is local variation of this ratio between sites. A ratio greater than one indicates diversifying selection.

A less commonly applied but potentially powerful method to assess the role of overdominant selection is to compare the structure and the shape of allelic genealogy with a coalescent model of the evolution of balancing genetic polymorphism [89] (Figure I-9). The main advantage of this approach is that it potentially allows for testing between two alternatives with respect to the overdominance model: Symmetric Balancing Selection, whereby all heterozygotes derive a similar advantage relative to homozygotes and Divergent Allelic Advantage by which heterozygotes carrying more divergent allelic sequences show greater selective advantage than individuals carrying relative similar alleles.



Allele genealogy for neutral alleles. Symbols (-) denotes the occurrence of mutations. *N*=100, *Ns*=0

Allelic genealogies for overdominant alleles. Symbols (-) denotes the occurrence of mutations. *N*=100, *Ns*=33 (A);*N*=200, *Ns*=100 (B).

Figure I-9: Allele genealogy for neutral versus overdominant alleles

Models testing the overdominance hypothesis [70] lead to the conclusion that this hypothesis is consistent with the main feature of MHC polymorphism and that there is currently no reason to reject it. The biological cause of selection at the MHC loci must reflect that T cells (bearing MHC proteins) have dual specificity and that they recognize viral (as well as other) antigens (nonself) and MHC molecules of the stimulating and target cells (self). Heterozygotes may benefit from enhanced immune surveillance relative to homozygotes, due to their ability to present an expanded spectrum of antigens by T cells. The view that overdominance rather than NFDS is responsible for the maintenance of MHC alleles comes largely from this functional insight.

This mechanism of host-pathogen co-evolution has been criticized because it would rise to a too dynamic polymorphism, in which MHC allele frequencies fluctuate over time [91].

Nevertheless, Borghans *et al.* (2004)[71] confirmed this dynamic behavior and showed that host-pathogen co-evolution can easily account for a realistic and high degree of MHC polymorphism, and that heterozygote advantage per se is insufficient to explain the observed level of polymorphism, even in very large populations. In coevolutionary simulations there were both heterozygote selection and FDS. Both FDS by host-pathogen coevolution and heterozygote advantage give rise to a dynamic equilibrium with fluctuating allele/haplotype frequencies.

The two models (overdominance and FDS) give rise to quite similar theoretical results and the critical experiments to distinguish between them have not yet been designed [95] .

Another difficulty raised by Satta *et al.* [96] is that the estimation of selection coefficient for human MHC showed low selection intensity, thereby placing severe restrictions on the possibility of measuring selection directly in vertebrate populations.

To conclude, host-parasite interactions can promote diversity or polymorphism either by FDS or by overdominance. These interactions are not well understood and mechanisms involved during the co-evolution of these two partners need more thorough investigation, partly because of the genetic and dynamic complexity of the processes involved.

Conclusion

In this review, we have attempted to demonstrate that in order to improve our understanding of the role of balanced selection with respect to ecological interactions, and in particular competition and antagonistic interactions, will require better models and additional experimental data.

Further empirical studies using micro-organisms are needed. The ones available have yielded very interesting results, and such real time hypothesis testing should be further developed.

Three additional potentially fruitful research directions are of primary importance for understanding what happens on an evolutionary timescale.

The first one focuses on environmental shifts, which have been shown to have a large effect on species diversity. In such case, interactions among populations or species vary because of the changing environmental conditions.

The second line of research focuses on interaction with multiple evolutionary branching. Experimental evolution using microbial systems is clearly a powerful way for developing such research.

Ecological interactions are very complex and the third line of research consists of a multidisciplinary approach coupling experiments and models related to the system, or using a variety of study systems.

Appendix : Summary Table

Type of Study : E (Experiments) N (observation in natura) M (Models)	Biological Model	Ecological interactions : C (Competition) P/p (Predator/prey) H/P (Host/parasite)	Proposed selection mechanism	References
N	***Gasterosteus aculeatus***	C	FD of resource competition and Character displacement	28,29,30,31
N	*Littorina saxatilis*	C	FD resource competition	32
N	Anolis Lizards	C	FD resource competition	33
E	*Escherichia coli*	C	FDS	35,36,37,38
M		C	FD	8,22,48,51
M		C	Stabilized selection or disruptive selection	30
E	***Gasterosteus aculeatus***	P/p	Character displacement	56
E	*Pseudomonas fluorescens/ Tetrahymena thermophila*	P/p	Diversifying selection NFDS	58
E	*Timema cristinae*	P/p	Diversifying selection	61
M		P/p	FDS	62,63,64,65
E	*Daphnia magna/ Pasteura ramosa*	H/P	FDS	72
E	*Pseudomonas fluorescens/ Phage Phi2*	H/P	Directional selection	59
E	*Linum mardinal/ Melampsora lini*	H/P	Directional selection	82
M		H/P	Directional selection	86
M		H/P	FDS	85
M	MHC	H/P	Overdominance	70
M	MHC	H/P	FDS	71

References

1. Hutchinson GE.1959. Homage to Santa Rosalia, or why are there so many kinds of animals? Am. Nat. 93: 145-159.

2.Darwin C. [1859]. 2006. The Origin of species by Means of Natural Selection, or the Preservation of Favored Races in the Struggle for life. Reprint ed. Dover Thrift Editions.

3. Thoday J.M. 1972. Disruptive selection. Proc. R. Soc. 182(B):109-143.

4. Heino M., Metz JAJ., Kaitala V. 1998. The Enigma of Frequency-Dependent Selection. IIASA Interim Report IR-97-061.Trends Ecol. Evol. 13: 367-370.

5. Ayala F.J. and Campbell C.A. 1974. Frequency-dependent selection, Annu.Rev. Ecol. Syst. 5: 115-138

6. Hartl D.L. and Clark A.G. 1989. Principles of Population Genetics, Sinauer Associates

7. Asthana S., Schmidt S., and Sunyaev S 2005. A limited role for balancing selection. Trends Gen.21: 30-32.

8. Doebeli M. Dieckmann U. 2000. Evolutionary Branching and Sympatric Speciation caused by Different Types of Ecological Interactions.IIASA Interim Report IR-00-040.

9. Kisdi E. and Geritz SAH. 1999. Adaptive dynamics in allele space: evolution of genetic polymorphism by small mutations in a heterogeneous environment. Evolution 53: 9931008.

10. Waxmann D., Gravilets S.2005.20 questions about adaptive dynamics. J. Evol. Biol. 18: 1139-1154

11. Gravilet S. 2004. Fitness landscapes and the origin of Species, Princeton University

12. Metz J.A.J., Nisbet R.M. and Geritz S.A.H. 1992. How should we define'fitness' for general ecological scenarios? Trends Ecol. Evol. 7: 198-202

13. Rice W.R. and Hostert E.E.1993.Laboratory experiments on speciation what have we learned in 40 years. Evolution 47: 1637-1653.

14. Kondrashov, A. S., and Mina M. V. 1986. Sympatric speciation: when is it possible? Biological Journal of the Linnean Society 27:201–223.

15. Day T. and Young K.A. 2004. Competitive and Facilitative evolutionary diversification. Bioscience 54: 101-109.

16. MacArthur RH. 1964. Environmental factors affecting bird species diversity.Am. Nat.98: 387-397.

17. MacArthur R and Levins R. 1964. Competition, habitat selection, and character displacement in a patchy environment. Proc. Natl. Acad. Sci. 51: 1207-1210.

18. MacArthur R, Levins R. 1967. The limiting similarity, convergence, and divergence of coexisting species. Am. Nat. 101: 377–385.

19. May RM. 1974. On the theory of niche overlap. Theoretical Population. Biology 5: 297-332.

20. May RM, MacArthur RH. 1972.Niche overlap as a function of environmental variability. Proc. Natl. Acad. Sci. 69:1109-1113.

21. Brown WL Jr, Wilson EO. 1956. Character displacement. Systematic. Zoology5: 49-64.

22. Dieckmann U, Doebeli M. 1999. On the origin of species by sympatric speciation. Nature 400: 354-357.

23. Friesen M. L., Saxer G., Travisano M. and Doebeli M. 2004. Experimental evidence for sympatric ecological diversification due to frequency dependent competition in *Escherichia coli*. Evolution 58: 245-260.

24. Robinson BW, Wilson DS. 1994. Character release and displacement in fishes: A neglected literature. Am. Nat. 144: 596-627

25. Schluter D.1996. Ecological speciation in postglacial fishes. Phil. Trans. R. Soc.B 351: 807-814.

26. Schliewen UK, Tautz D, Paabo S. 1994. Sympatric speciation suggested by monophyly of crater lake cichlids.Nature 368: 629-632

27. Schliewen U, Rassmann K,Markmann M,Markert J,Kocher T,Tautz D. 2001.Genetic and ecological divergence of a monophyletic cichlid species pair under fully sympatric conditions in Lake Ejagham. Cameroon. Mol. Ecol. 10: 1471-1488

28. Schluter D. and McPhail JD. 1992. Ecological character displacement and speciation in sticklebacks..Am. Nat. 140: 85-108

29. Schluter D. 1994 . Experimental evidence that competition promotes divergence in adaptive radiation. Science 266 :798-800

30. Rundle HD. and Schluter D. 2004. Natural Selection and Ecological Speciation in Sticklebacks. In Adaptive Speciation, eds. Dieckmann U, Doebeli M, Metz JAJ & Tautz D. : 192–209. Cambridge University Press.

31. Nagel L and Schluter D. 1998 . Body size, natural selection, and speciation in sticklebacks. Evolution 52:209-218.

32. Johannesson K., Rolan-Alvarez E., Ekendahl A. 1995. Incipient reproductive isolation between two sympatric morphs of the intertidal snail *Littorina saxatilis*. Evolution 49: 1180-1190

33. Losos B. J., Jackman T.R., Larson A., 1998. Contingency and Determinism in Replicated Adaptive Radiations of Island Lizards. Science 279: 2115-2118.

34. Elena S., Lenski R.E. 2003. Evolution experiments with microorganisms:the dynamics and genetic bases of adaptation. Nature Reviews-Genetics 4:457-469

35. Tyerman J., Naomi H., Saxer G, Travisano M., Doebeli M. 2005. Unparallel diversification in bacterial microcosms. Proc. R.Soc.2: 1393-1398

36. Levin B. 1988. Frequency-dependant selection in bacterial populations. Phil. Trans. R.Soc. B 319: 359-472.

37. Blount Z.D.,. Borland C.Z, and R.E. Lenski 2008. Historical contingency and the evolution of a key innovation in an experimental population of *Escherichia coli*. Pnas 105: 7899-7906.

38. Helling RB, Vargas CN, Adams J. 1987.Evolution of *Escherichia coli* during growth in a constant environment. Genetics 116:349-58

39. Rosenzweig RF, Sharp RR, Treves DS,Adams J. 1994. Microbial evolution in a simple unstructured environment: genetic differentiation in Escherichia coli. Genetics 137:903–17

40. de Visser JA. GM, Zeyl CW, Gerrish PJ, Blanchard JL, Lenski RE. 1999. Diminishing returns from mutation supply rate in asexual populations. Science 283:404–6

41. Rainey PB, Buckling A, Kassen R, Travisano M. 2000. The emergence and maintenance of diversity: Insights from experimental bacterial populations. Trends Ecol. Evol.15:243-247.

42. Travisano M, Rainey PB. 2000. Studies of adaptive radiation using model microbial systems. Am. Nat. 156: S35-S44.

43. Kassen R. 2002. The experimental evolution of specialists, generalists, and the maintenance of diversity. J. Evol. Biol. 15: 173-190.

44. Travisano M, Rainey PB. 2000. Studies of adaptive radiation using model microbial systems. Am. Nat. 156: S35-S44

45. Roughgarden J. 1972. Evolution of niche width. Am. Nat. 106:683-718

46. Slatkin M. 1980. Ecological character displacement. Ecology 66: 355–371

47. Day T. 2000.Competition and the effect of spatial resource heterogeneity on evolutionary diversification. Am. Nat. 155: 790-803

48. Kisdi E. 1999. Evolutionary Branching Under Asymmetric Competition. IIASA Interim Report IR-98-045.J. Theor. Biol. 197: 149-162.

49. Geritz SAH, van der Meijden E, Metz JAJ. 1999. Evolutionary Dynamics of Seed Size and Seedling Competitive Ability. IIASA Interim Report IR-97-071. Theoretical Population Biology 55: 324-343.

50. Law R, Marrow P, Dieckmann U. 1997. On Evolution under Asymmetric Competition. IIASA Working Paper WP-96-003. Evol. Ecol. 11: 485-501.

51. Doebeli, M. 1996. A quantitative genetic competition model for sympatric speciation. J. Evol. Biol. 9:893-909.

52. Nagel L. and Schluter, D. 1998. Body size, natural selection, and speciation in sticklebacks. Evolution 52: 209-218.

53. Rundle H.D. and Schulter D. 1998. Reinforcement of stickleback mate preferences: sympatry breeds contempt. Evolution 52:200-208.

54. Vamosi SM .2005. On the role of enemies in divergence and diversification of prey: a review and synthesis. Canadian Journal of Zoology 83: 894-910

55. Ashraf M. T. Elewa. 2007. Predation in Organisms:A Distinct Phenomenon, Springer Berlin Heidelberg 177-220

56. Rundle HD., Vamosi SM., Schluter D. 2003. Experimental test of predation's effect on divergent selection during character displacement in sticklebacks. Proc. Natl. Acad. Sci. 100: 14943-14948

57. Gurevitch J., Morrison J. A. and Hedges, L. V. 2000. The interaction between competition and predation: a meta-analysis of field experiments Am. Nat.155: 435-453.

58. Meyer, J. R. & Kassen, R. 2007. The effects of competition and predation on diversification in a model adaptive radiation. Nature 446: 432-435 .

59. Buckling, A. & Rainey, P. B. 2002. The role of parasites in allopatric and sympatric host diversification. Nature 420: 496-499.

60. Rainey, P.B. and Travisano, M. 1998. Adaptive radiation in a heterogeneous environment. Nature 394: 69-72.

61. Nosil P, Crespi BJ. 2006. Experimental evidence that predation promotes divergence in adaptive radiation. Proc. Natl Acad. Sci. 103: 9090-9095

62. Fryer, G. 1959. Some aspects of evolution in Lake Nyasa. Evolution 13: 440-451.

63. Holt, R.D. 1977. Predation, apparent competition, and the structure of prey communities. Theor. Popul. Biol. 12: 197-229.

64. Brown, J.S. and Vincent, T.L. 1992. Organization of predator–prey communities as an evolutionary game. Evolution, 46: 1269-1283.

65. Abrams, P.A. 2000. Character shifts of prey species that share predators. Am. Nat. 156(Suppl.): S45–S61.

66. Bowers RG, White A, Boots M, Geritz SAH, Kisdi E .2003. Evolutionary branching/speciation: contrasting results from systems with explicit or emergent carrying capacities. Evolutionary Ecology Research 5: 883-891

67. Hamilton, W. D. 1980. Sex versus non-sex versus parasite. Oikos 35:282-290.,

68. Bell, G., and J. Maynard Smith. 1987. Short-term selection for recombination among mutually antagonistic species. Nature 328:66-68.

69. Hamilton, W. D., R. Axelrod, and R. Tanese. 1990. Sexual reproduction as an adaptation to resist parasites. Proc. Natl. Acad. Sci.. USA 87:3566-3573

70. Takahata N., Satta Y, Klein J.1992. Polymorphism and Balancing selection at Major Histocompatibility Complex loci Genetics 130: 925-938

71. Borghans J.A. M., Beltman J. B., De Boer R.J. 2004. MHC polymorphism under host-pathogen coevolution Immunogenetics 55:732-739

72. Carius H.J, Little T.J, and Ebert D. 2001. Genetic variation in a host-parasite association: potential for co-evolution and frequency-dependent selection. Evolution 55: 1136-1145.

73. Bohannan B.J.M., Lenski R.E.2000. Linking genetic change to community evolution: insights from studies of bacteria and bacteriophage. Ecol. Lett. 3: 362-377

74. Brockhurst M.A., Buckling,A., Rainey, P.B. 2005. The effect of a bacteriophage on diversification of the opportunistic bacterial pathogen *Pseudomonas aeruginosa*. Proc. R. Soc. B 272: 1385-1391.

75. Brockhurs, M.A., Bucklin, A., Rainey P.B. 2006. Spatial heterogeneity and the stability of host parasite coexistence. J. Evol. Biol. 19: 374-379.

76. Buckling A., Rainey P.B. 2002. Antagonistic coevolution between a bacterium and a bacteriophage. Proc. R. Soc. B 269: 931-936

77. Brockhurst M.A., Morgan, A.D., Rainey P.B., Buckling A.2003. Population mixing accelerates coevolution. Ecol. Lett. 6: 975-979.

78. Buckling A., Rainey P.B.2002. Antagonistic coevolution between a bacterium and a bacteriophage. Proc. R. Soc. B 269: 931-936.

79. Buckling A., Wei Y., Massey R.C., Brockhurst M.A., Hochbert M.E.2006. Antagonistic coevolution with parasites increases the cost of host deleterious mutations. Proc. R. Soc. B 273: 45-49

80. Sasaki A. 2000. Host–parasite coevolution in a multilocus gene-for-gene system. Proc. R. Soc. B 267: 2183-2188

81. Brockhurst M.A., Rainey P.B., Buckling A. 2004. The effect of spatial heterogeneity and parasites on the evolution of host diversity. Proc. R. Soc. B 271: 107-111

82. Thrall P.H., Burdon, J.J. 2003. Evolution of virulence in a plant host-pathogen metapopulation. Science 299: 1735-1737

83. Agrawal, A.F., Lively C.M. 2002. Infection genetics: gene-for-gene versus matching alleles models and all points in between. Evol. Ecol. Res. 4: 79-90

84. Morgan A.D., MacClean C. R., Buckling A. 2009.Effects of antagonistic coevolution on parasite-mediated host coexistence, J. Evol.Biol 22:287-292

85. Otto, P. O., and Y. Michalakis. 1998. The evolution of recombination in changing environments. Trends Ecol. Evol. 13:145-151

86. Clay, K., and P. Kover. 1996. The Red Queen hypothesis and plant/pathogen interactions. Annu. Rev. Phytopathol. 34:29-50

87. Clarke, B.C. and Kirby, D.R.S. 1966. Maintenance of histocompatibility polymorphism. Nature 211: 999-1000

88. Doherty, P.C. & Zinkernagel, R.M. 1975. Enhanced immunological surveillance in mice heterozugous at the H-2 gene complex. Nature 256: 50-52

89. Takahata N, Nei M. 1990.Allelic genealogy under overdominant and frequency-dependent selection and polymorphism of major histocompatibility complex loci. Genetics 124: 967-978.

90. Hughes AL, Lewontin RC, Ginzburg LR, Tuljapurkar SD. 1978.Heterosis as an explanation for large amounts of genic polymorphism. Genetics 88:149-170

91. Hughes, A.L.and Nei M. 1988. Pattern of nucleotide substitution at major histocompatibility complex class I loci reveals overdominant selection. Nature 335: 167-170.

92. Ewens, W.J. 1972. The sampling theory of selectively neutral alleles. Theo. Pop. Bio. 3: 87-112.

93. Watterson, G.A. 1978. The homozygosity test of neutrality. Genetics 88: 405-417.

94. Hill, R.E. & Hastie, N.D. 1987. Accelerated evolution in the reactive centre regions of serine protease inhibitors. Nature 326: 96-99

95. Hughes AL, Yeager M .1998. Natural selection at major histocompatibilitycomplex loci of vertebrates. Annual Review of Genetics, 32: 415-435

96. Satta Y., O'Huigin C. Takahata N, Klein J. 1994. Intensity of natural selection at the major histocompatibility complex loci. Proc. Nat. Acad. Sci. USA. 91: 7184-7188

Chapter 2

II. State of the art on biological models

Two organisms were well suited to carry out our evolution experiments.

The first one was the Gram negative bacteria *E. coli*, which has a short generation time, large population size, can be stored and for which it is very easy to manipulate environmental variables.

The second one was the social amoeba *D. discoideum* which is a good bacterial predator and which could be used as biotic factor on *E. coli*. *D. discoideum* has also short generation time, large population size, and can be stored and used under different environmental conditions.

I will provide a historical view of these two organisms as well as some of their characteristics and ecology.

II STATE OF THE ART ON BIOLOGICAL MODELS

1. *Escherichia coli*

1.1 History: 135 years of *Escherichia coli*

1.1.1 *Bacterium coli commune*

Theodor Escherich (1857-1911) was the first person to describe *Bacterium coli commune*. After one-and a-half years of research, in 1886 he presented his results on "The intestinal bacteria of the infant and their relationship to the physiology of digestion". He mainly noticed two types of bacteria in infant faeces. One was found particularly frequent in the lower sections of the gut, and for this reason Theodor Escherich described them as typical "colonic bacteria" and named the species "*Bacterium coli commune*", or the common colonic bacterium (Figure II-1)

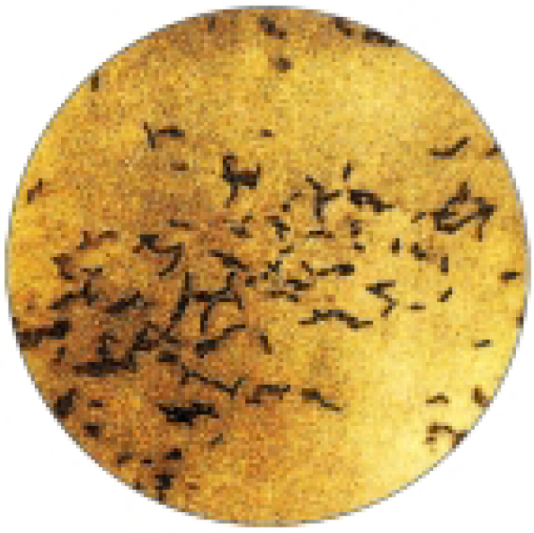


Figure II-1: The first illustration of *Bacterium coli commune*
Permanent preparation by Theodor Escherich in 1886: stained with gentian violet in aqueous aniline. In: "The intestinal bacteria of the infant and their relationship to the physiology of digestion". Ferdinand-Enke-Verlag. Stuttgart.

1.1.2 *E. coli* K-12: The favorite “experimental model” of the molecular geneticist

In 1922, it was a pure chance that a harmless *E. coli* variant was isolated from the stools of a patient suffering from diphtheria. It found a place in the bacteriological strain collection of the Medical School of Stanford University in California under the laboratory designation “K-12”. Like conventional animal and plant breeders, microbiologists started looking through their “domestic species” *E. coli* and looked for new strains. With the aid of X-rays and UV rays, they introduced various mutations into the K-12 “wild strain”. Very soon, many laboratories world-wide were working with the original strain and the derived variants.

1.1.3 Intra and extra-intestinal pathogenic *E. coli* variants

Theodor Escherich was initially convinced that *Bacterium coli commune* was a “harmless commensal”. He was to a large extent right, because most of the approximately 50,000 serotypes should be counted among commensal gut organisms. However, Theodor Escherich reported on “cystitis in children provoked by *Bacterium coli commune*” as early in 1894. This early hypothesis that *E. coli* bacteria, which could be pathogenic in the urinary tract under certain circumstances, persist without symptoms in the intestine and for various reasons find their way into the organs disposing of urine, where they might cause inflammation, has now been confirmed by modern biochemical and molecular biological methods [1, 2]. There is now a huge extensive collection of publications on the host-specific pathogenic potential of various *E. coli* variants.

The genetic structure of *E. coli* population is mainly clonal, with four main phylogenetic groups denoted A, B1, B2 and D [3, 4]. Each phylogenetic group displays various life history traits and differs in its phenotypic and genotypic characteristics [4, 5]. Compared to commensal strains, extra-intestinal pathogenic *E. coli* (ExPEC) have additional virulence factors, such as adhesins (e.g., P and I fimbriae), iron-acquisition systems (e.g., aerobactin), host defense-avoidance systems (e.g., capsule, lipopolysaccharide), and toxins (e.g., hemolysin, cytotoxic necrotizing factor1) [6, 7]. ExPEC virulence factors, encoded by genes often clustered on so-called pathogenicity island (PAI), are mainly found in strains of the B2 and D phylogenetic groups but

scarcely in the A and B1 phylogenetic groups, to which the majority of commensal and extra-intestinal pathogenic *E. coli* strains belong [8-10].

1.2 The organism

E. coli is a member of the family of Enterobacteriaceae, related to the genus *Shigella* and closely related to the genera *Citrobacter, Salmonella* and *Klebsellia*. It is a Gram-negative bacteria rod of around 1.1-1.5 μm x 2.0-6.0 μm in size. It grows under aerobic and anaerobic conditions (facultative anaerobic), because it possesses two different redox systems (menaquinone and ubiquinone) which enable it to derive energy from catabolic metabolism under both aerobic and anaerobic conditions. Under optimal growing condition, the rate of cell division of the *E. coli* bacteria is very rapid; the number of bacteria cells can double every 20 minutes.

The first genetic map of the *E. coli* K-12 genome was produced as early as 1964 [11]. One of the first strains of bacteria whose genome was completely sequenced was a strain of *E. coli* K-12 (MG 1655) [12]. The complete genome sequence of the *E. coli* K-12 strain published in 1997 revealed about 4,300 genes, with a size of 4.7×10^6 base pairs, the genome of *E. coli* K-12 is smaller than the human genome by a factor of 1,000.

1.3 Ecology

Most strains of *E. coli* are commensal and use their host as a suitable habitat without any noteworthy effect (benefit or detriment) for the host, but pathogenic strains exist that cause harm or even death [13]. Even so-called extra-intestinal pathogenic *E. coli* strains found in mammalian intestinal tracts do not normally cause disease, but they do when they incidentally invade sterile niches such as urinary tract, blood [14, 15] or cerebrospinal fluid. Besides this primary habitat, *E. coli* can use the external environment, including water, sediment, and soil [16-18] where some specific strains may grow [19-21] (Figure II-2).

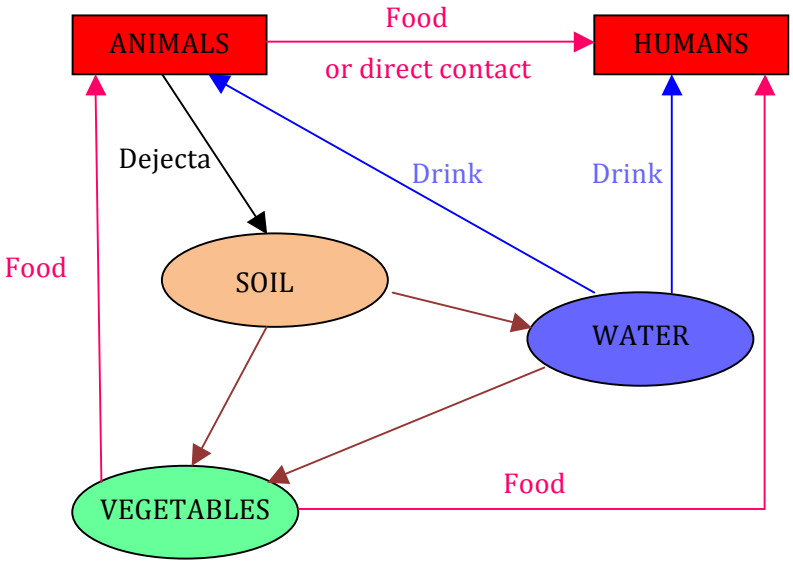


Figure II-2: Potential *E. coli* life cycle
The main reservoir is the mammalian intestinal tract (squares)
The second habitat is the external environment: water, sediments, soil (circles)

1.4 Biological model

E. coli is probably the most well-understood bacterium, and it is an extremely important model organism in many fields of research, particularly in molecular biology, genetics, and biochemistry. It is easy to cultivate under laboratory conditions, and research strains are very safe to work with. As with many bacteria, *E. coli* grows quickly, allowing for many generations to be studied over a short time period (*E. coli* cells can double in number after only twenty minutes).

Furthermore, a very large number of *E. coli* bacteria can be grown in a small volume—many millions in a drop of broth, for example.. Microbes offer powerful systems for experimental evolution (Chapter I, Box 1 p.22).

2. *Dictyostelium discoideum*

2.1 History

The mycologist Oskar Brefeld (1869) first described the slime mould, *Dictyostelium*. He observed *Dictyostelium mucoroides* while examining the fungal flora in horse dung, and then grew purer cultures in rabbit dung. He realized that the amoeba aggregated to give rise to fructifications. He named the species *Dictyostelium* – *Dicty* means net-like and *stelium* means tower- because the aggregation territories

looked like nets and fruiting bodies like towers. He added the qualifier *mucoroides* because the new organism resembled the fungus *Mucor*.

In 1899, G.A. Nadson reported that *D. mucoroides* grew with a known species of bacteria – *Bacillus fluorescens liquefaciens*- he assumed that the two organisms were symbionts rather than predator and prey [22]. It was in 1902 that Potts developed media more sophisticated than simple dung and realized that the growth of the amoeba depended on the presence of bacteria [23]. Potts was the first to maintain the amoeba in a persistent vegetative stage with fresh bacteria food sources. How the cells of the aggregate assembled was an issue that had been considered as early as 1902 by Olive and Potts, who suggested the possibility of chemotaxis, but provided no evidence to support the hypothesis [24, 23].

In 1903, Vuillemin possibly under the influence of Metchinikoff, realized that amoeba engulfed the bacteria, that the digestion was intracellular, and that *D. mucoiroides* and its bacterial food source were not symbionts [25].

The social amoebae were assumed to be of no interest for experimental embryology, until embryology moved out of the shadow of the evolutionary theories of Haeckel. One of the earliest workers to stress the developmental aspects of *Dictyostelium* development was Arndt (1937). Raper, encouraged by Harper, started to study the *Acrasiales* as a young student [26]. Raper started by isolating a previously unknown species *D. discoideum*, from the hardwood forest of little Butt's Gap, in the periphery of Ashville, North Carolina [27]. Raper detected on *D. discoideum* and other species from soil samples, whereas previously organisms had been recovered from dung. Raper showed that any species of bacteria, spread as a lawn, would support the growth of *D. discoideum.* He settled on a two member system, employing *E. coli* or *Aerobacter* (now *Klebsellia*) *aerogenes*, established the media for growth. In the early 1940s, Raper and Thom published their studies about the developmental capacities of *Dictyostelium* [28].

Bonner (1947) [29] solved the problem of chemotaxis and aggregation. He found a substance named acrasin, a generic name for any chemotactic molecule produced by a member of the order *Acrasiales*. In *D. discoideum* and a number of other species, the acrasin is cAMP [30]. Shaffer proposed that the chemotactic agent was emitted in pulses that diffused locally and stimulated nearby cells to release more acrasin in a runaway process. The relay theory predicted a pace-maker consisting of a group of

cells in the center of the aggregated. All that we know of the pace-maker cells is that they are the first cells among equals, generated by chance, and are not genetically distinct [31].

There were two other major contributors to the establishment of *Dictyostelium* as an experimental organism. If the experiments of Raper, Bonner, and their students and colleagues were largely biological, Maurice Sussman realized that he could take advantage of the natural synchrony of development to study biochemistry of this development. The period from the late 1960s to the mid 1980s saw an exponential increase in the number of research articles involving *D. discoideum*.

2.2 Organism

D. discoideum is a simple haploid eukaryote with a cellular organization typical of higher eukaryotes. It has a plasma membrane and no cell wall. The 35 Mb genome was sequenced [32] and found to contain many genes that are homologous to those of higher eukaryotes but missing in *Saccharomyces cerevisiae*.

D. discoideum reproduces through mitotic division of single cells that feed by phagocytosis on bacterial lawn, or on simple axenic liquid medium. Amoebae grow on bacterial lawns or in liquid cultures of defined media with doubling times of four and twelve hours, respectively. The term ‘social amoebae’ derives from the observed behavior of cells when the food supply is exhausted or removed. Upon onset of starvation, undifferentiated single cells stop division and enter a multicellular cycle [33] (Figure II-3).

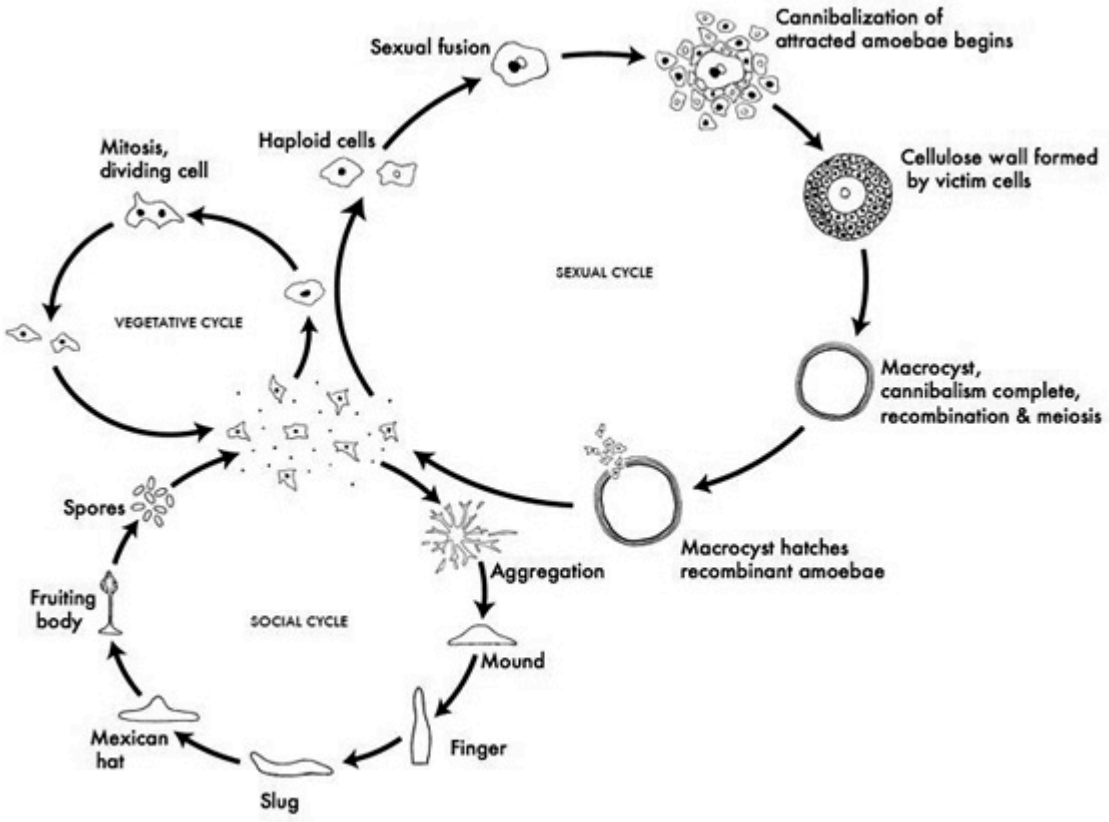


Figure II-3: *D. discoideum* life cycle

Vegetative cycle: haploid amoeba, dividing mitotically. When food is scarce, either the sexual cycle or the social cycle begins.

In the social cycle, amoeba aggregate toward cAMP by the thousands, forming a slug, then a fruiting body, in which about 20% of the cells die, to form a stalk for lifting remaining cells up to a better place for sporulation and dispersal.

In the sexual life cycle, amoeba aggregate to cAMP and sex pheromone, and two cells of opposite mating types fuse and then begin to consume the other attracted cells. When cannibalism is complete, the giant diploid cell forms a robust macrocyst which eventually undergoes recombination and meiosis.

Brown D. and Strassmann J.E.

The initiation of development occurs when on the order of 10^2 to 10^5 cells aggregate to a single point by chemotaxis according to a gradient of cAMP signal, which is relayed to cells that are more distant from the aggregation center [34].

Following the formation of multi-cellular aggregates, cells differentiate into two basic cell types, pre-stalk and pre-spore, that ultimately become stalk and spore cells, respectively. Cell differentiation results in the formation of a mature fruiting body in which a mass of up to 80,000 spores rests on top of a multicellular stalk composed of up to 20,000 dead, vacuolated stalk cells (Figure II-4).

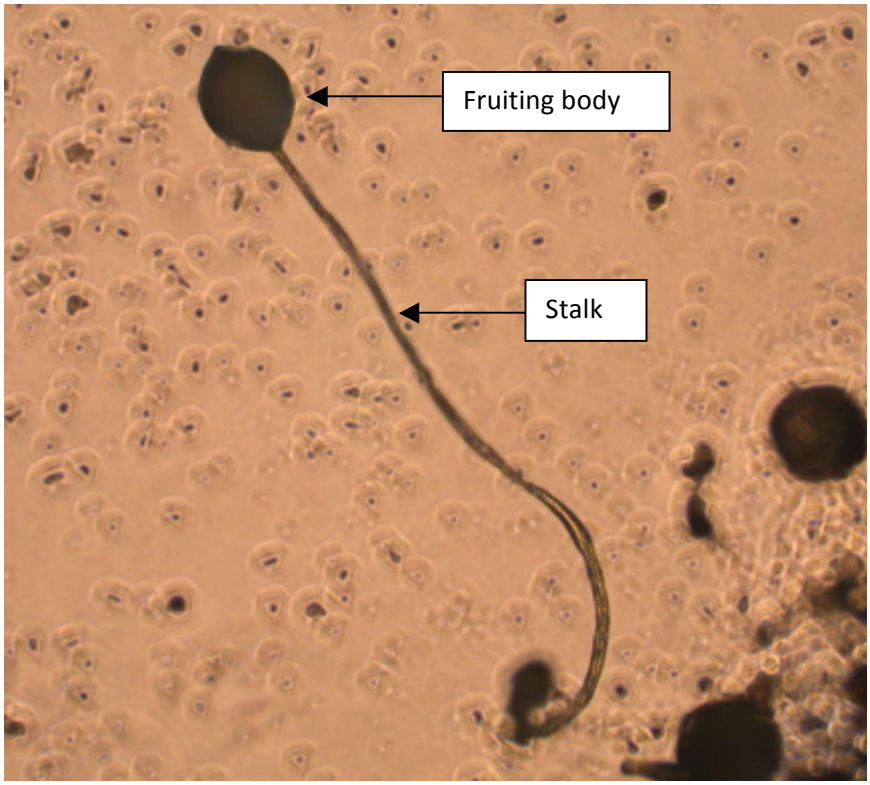


Figure II-4: *D. discoideum* fruiting body and stalk

This organism is particularly suited for studies of cytokinesis, motility, phagocytosis, chemotaxis, signal transduction, the evolution of sociality/metazoans, and aspects of development. Many of these processes, which play important roles in health and diseases, are either absent or are less accessible in other model organisms. Free-living amoebae feed mainly on bacteria, fungi, and algae through phagocytosis. The digestion occurs within the phagolysosomes. Some micro-organisms (such as Legionella) have evolved resistance to protists, since these micro-organisms are not internalized or are able to survive, grow, and exit free-living amoebae after internalization.

2.3 Ecology

Amoebae that occur in the human intestinal tract (*Entamoeba* species) [35-37] typically also occur in external environment as well, in habitats such as fresh water, moist soil, at various interfaces (water-soil, water-plant, water-air, Figure II-5) [38].

Figure II-5: *D. discoideum* natural habitat
D. discoideum occurs in natural habitats such as moist soil.

They are distributed worldwide (Figure II-6), but the set of species in a particular location depends on abiotic factors and food availability. Amoebae are particularly effective bacterial predators.

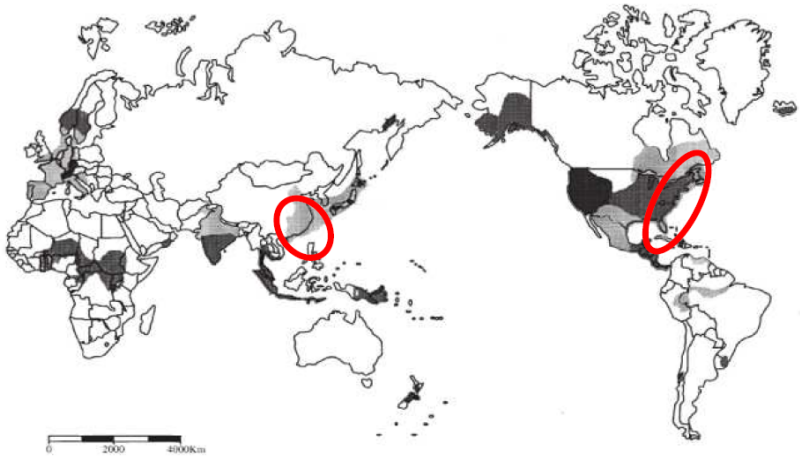


Figure II-6: Amoeba geographical distribution
Locations of cellular slime mould sampling around the world. Red circle localises *D. discoideum* [39].

The group of free-living amoeba is the only group of organisms that can cause a decrease in a bacterial population from 10^8 down to 10^5 per gram of soil [40-42].

D. discoideum is a cellular slime mould which natural habitat are soil and leaf litter in which it predates bacteria by phagocytosis.

2.4 Biological model

The haploid social amoeba *D. discoideum* has been used as a model to study host-factors involved in cellular aspects of host-pathogen interactions [43, 44]. The

common mechanisms of infection needed to infect both mammals and *D. discoideum* by a wide variety of pathogens reinforces the value of the *E. coli/D. discoideum* system as an appropriate model to study host-pathogen relations and their evolution [43, 45]. In addition, elaborate genetic tools including genomic sequence [32] and plasmids allowing constitutive or inducible expression of genes are now available.

Furthermore, *D. discoideum* was described as showing the same pathway as macrophage phagocytosis (Figure II-7).

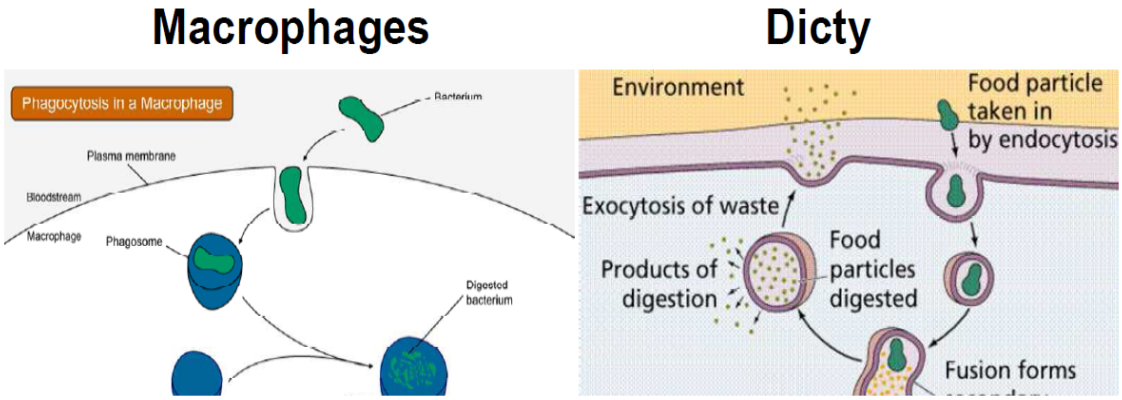


Figure II-7:Comparison of phagocytosis pathway between macrophage and *D. discoideum*

D. discoideum had been described as showing the same pathway as macrophage phagocytosis.

The bacteria is internalised by endocytosis on a phagosome and then digested after the phagosome and lysosome fusion. Products of digestion are then exocytised.

The similarity in the intracellular infections of macrophages and protozoa by *Legionella pneumophila* is quite remarkable. Within the cells of both of these evolutionary distant hosts, the bacterium multiplies within a phagosome that does not fuse with lysosomes [46]. The strategy of resistance to macrophage microbial effectors varies from species and might have been acquired following exposure to environmental predators such as free-living amoebae.

Selective predation has given rise to diverse routes of bacterial defense, including adaptive mechanisms in bacterial biofilms, and has promoted major transitions in bacterial evolution such as pathogenesis. Bacterial anti-predator adaptations are diverse and can act either before ingestion or afterwards inside the food vacuole [47] (Figure II-8).

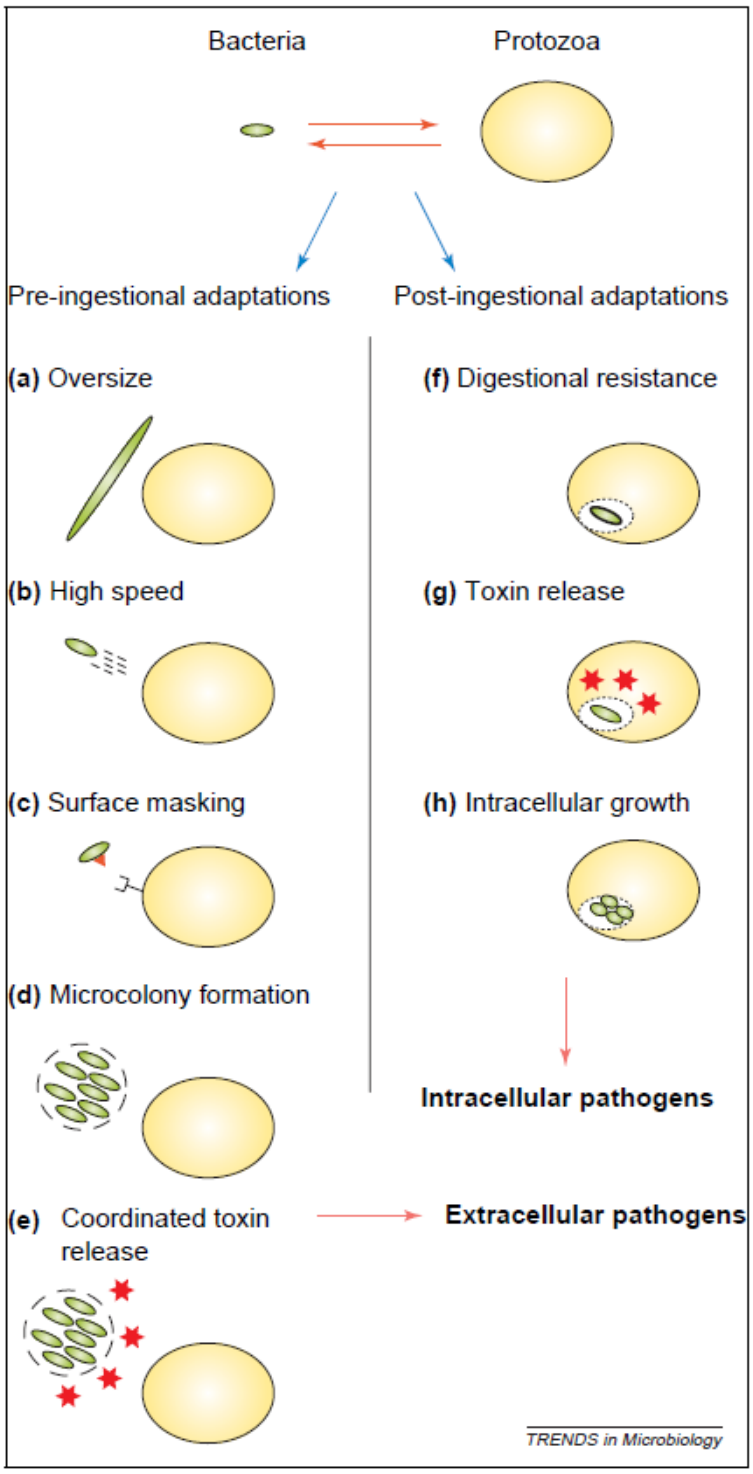


Figure II-8: Schematic overview of the potential routes of bacterial adaptation against predation by protozoa and the hypothetical transition from grazing resistance to pathogenesis

Blue arrows indicate the two divergent strategies of pre-ingestional (a-e) and post-ingestional (f-h) adaptations emerging from bacteria-protozoa interactions.
Red arrows show the potential origin of pathogenesis. [47]

Literature cited

1. Blum, G., J. Hacker, and R. Marre, *Properties ofEscherichia coli strains of serotype O6.* Infection, 1995. **23**(4): p. 234-236.
2. Plos, K., et al., *Intestinal carriage of P fimbriated Escherichia coli and the susceptibility to urinary tract infection in young children.* The Journal of infectious diseases, 1995. **171**(3): p. 625-631.
3. Desjardins, P., et al., *Sex in Escherichia coli does not disrupt the clonal structure of the population: evidence from random amplified polymorphic DNA and restriction-fragment-length polymorphism.* Journal of molecular evolution, 1995. **41**(4): p. 440-448.
4. Herzer, P., et al., *Phylogenetic distribution of branched RNA-linked multicopy single-stranded DNA among natural isolates of Escherichia coli.* Journal of bacteriology, 1990. **172**(11): p. 6175-6181.
5. Escobar-Paramo, P., et al., *Identification of forces shaping the commensal Escherichia coli genetic structure by comparing animal and human isolates.* Environmental Microbiology, 2006. **8**(11): p. 1975.
6. Johnson, J., *Virulence factors in Escherichia coli urinary tract infection.* Clinical Microbiology Reviews, 1991. **4**(1): p. 80-128.
7. Johnson, J. and T. Russo, *Extraintestinal pathogenic Escherichia coli:" The other bad E coli".* Journal of Laboratory & Clinical Medicine, 2002. **139**(3): p. 155.
8. Picard, B., et al., *The link between phylogeny and virulence in Escherichia coli extraintestinal infection.* Infection and immunity, 1999. **67**(2): p. 546.
9. Escobar-Paramo, P., et al., *A specific genetic background is required for acquisition and expression of virulence factors in Escherichia coli*, in *Molecular biology and evolution*. 2004, SMBE. p. 1085-1094.
10. Tenaillon, O., et al., *The population genetics of commensal Escherichia coli.* Nature Reviews Microbiology, 2010. **8**(3): p. 207-217.
11. Taylor, A. and M. Thoman, *The genetic map of Escherichia coli K-12.* Genetics, 1964. **50**(4): p. 659.
12. Blattner, F., et al., *The complete genome sequence of Escherichia coli K-12.* Science, 1997. **277**(5331): p. 1453.
13. Steinert, M., U. Hentschel, and J. Hacker, *Symbiosis and pathogenesis: evolution of the microbe-host interaction.* Naturwissenschaften, 2000. **87**(1): p. 1-11.
14. Smith, H., *Observations on the flora of the alimentary tract of animals and factors affecting its composition.* The Journal of Pathology and Bacteriology, 1965. **89**(1).
15. Berg, R., *The indigenous gastrointestinal microflora.* Trends in Microbiology, 1996. **4**(11): p. 430-435.
16. Van Donsel, D., E. Geldreich, and N. Clarke, *Seasonal variations in survival of indicator bacteria in soil and their contribution to storm-water pollution.* Applied and Environmental Microbiology, 1967. **15**(6): p. 1362-1370.
17. VanDonsel, D. and E. Geldreich, *Relationships of salmonellae to fecal coliforms in bottom sediments.* WATER RESEARCH, VOL 5, NO 11, P 1079-1087, NOVEMBER 1971. 3 FIG, 3 TAB, 21 REF., 1971.
18. Faust, M., *Coliform bacteria from diffuse sources as a factor in estuarine pollution.* Water Res., 1976. **10**(7): p. 619-624.
19. Barker, J., T. Humphrey, and M. Brown, *Survival of Escherichia coli O157 in a soil protozoan: implications for disease.* FEMS microbiology letters, 1999. **173**(2): p. 291.
20. Power, M., et al., *Phenotypic and genotypic characterization of encapsulated Escherichia coli isolated from blooms in two Australian lakes.* Environmental Microbiology, 2005. **7**(5): p. 631.
21. Solo-Gabriele, H., et al., *Sources of Escherichia coli in a coastal subtropical environment.* Applied and Environmental Microbiology, 2000. **66**(1): p. 230-237.
22. Nadson, G., *Des cultures du Dictyostelium mucoroides Bref. et des cultures pures des amibes en general.* Scripta Bot. Horti Univ. Imp. Petropolitanae, 1889. **16**: p. 153-190.
23. Potts, G., *Zur physiologie des Dictyostelium mucoroides.* 1902.
24. Olive, E., *Monograph of the Acrasieae*. 1902: Printed for the [Boston] Society [of Natural History].
25. Vuillemin, P., *Une Acrasiee bacteriophage.* Comp. Roud. Acad. Sci. Paris, 1903. **137**: p. 387-389.
26. Raper, K. and A. Rahn, *The dictyostelids*. 1984: Princeton University Press Princeton, New Jersey.
27. Raper, K., *Dictyostelium discoideum, a new species of slime mold from decaying forest leaves.* J. Agr. Res, 1935. **50**: p. 135-147.
28. Raper, K., *Pseudoplasmodium formation and organization in Dictyostelium discoideum.* J. Elisha Mitchell Sci. Soc, 1940. **56**: p. 241-282.

29. Bonner, J. and L. Savage, *Evidence for the formation of cell aggregates by chemotaxis in the development of the slime mold Dictyostelium discoideum.* Journal of Experimental Zoology, 1947. **106**(1): p. 1-26.
30. Konijn, T., et al., *Cyclic AMP: a naturally occurring acrasin in the cellular slime molds.* American Naturalist, 1968. **102**(925): p. 225-233.
31. Bonner, J., *Development in the cellular slime molds: the role of cell division, cell size and cell number.* Developing Cell Systems and their Control, 1960: p. 3-20.
32. Eichinger, L., et al., *The genome of the social amoeba Dictyostelium discoideum.* Nature, 2005. **435**(7038): p. 43-57.
33. Kessin, R., *Dictyostelium: evolution, cell biology, and the development of multicellularity*. 2001: Cambridge Univ Pr.
34. Van Haastert, P. and P. Devreotes, *Chemotaxis: signalling the way forward.* Nature Reviews Molecular Cell Biology, 2004. **5**(8): p. 626-634.
35. Fotedar, R., et al., *PCR detection of Entamoeba histolytica, Entamoeba dispar, and Entamoeba moshkovskii in stool samples from Sydney, Australia.* Journal of Clinical Microbiology, 2007. **45**(3): p. 1035.
36. Hamzah, Z., et al., *Differential detection of Entamoeba histolytica, Entamoeba dispar, and Entamoeba moshkovskii by a single-round PCR assay.* Journal of Clinical Microbiology, 2006. **44**(9): p. 3196.
37. Troll, H., H. Marti, and N. Weiss, *Simple differential detection of Entamoeba histolytica and Entamoeba dispar in fresh stool specimens by sodium acetate-acetic acid-formalin concentration and PCR.* Journal of Clinical Microbiology, 1997. **35**(7): p. 1701.
38. Rivera, F., et al., *Pathogenic and free-living protozoa cultured from the nasopharyngeal and oral regions of dental patients: II.* Environmental research, 1986. **39**(2): p. 364.
39. Swanson, A., E. Vadell, and J. Cavender, *Global distribution of forest soil dictyostelids.* Journal of Biogeography, 1999. **26**(1): p. 133-148.
40. Habte, M. and M. Alexander, *Further evidence for the regulation of bacterial populations in soil by protozoa.* Archives of Microbiology, 1977. **113**(3): p. 181-183.
41. Bryant, R., et al., *Interactions of bacterial and amoebal populations in soil microcosms with fluctuating moisture content.* Applied and Environmental Microbiology, 1982. **43**(4): p. 747-752.
42. Clarholm, M., *Protozoan grazing of bacteria in soil—impact and importance.* Microbial Ecology, 1981. **7**(4): p. 343-350.
43. Burns, D., et al., *Intracellular growth of Legionella pneumophila in Dictyostelium discoideum, a system for genetic analysis of host-pathogen interactions.* Infection and immunity, 2000. **68**(5): p. 2939-2947.
44. Steinert, M. and K. Heuner, *Dictyostelium as host model for pathogenesis.* Cellular microbiology, 2005. **7**(3): p. 307-314.
45. Cardelli, J., *Phagocytosis and macropinocytosis in Dictyostelium: phosphoinositide-based processes, biochemically distinct.* Traffic, 2001. **2**(5): p. 311.
46. Gao L.Y, Harb O.S and Kwaik Y.A, *Utilization of similar mechanisms by Legionella pneumophila to paratize two evolutionary distant host cells, mammalian macrophages and protozoa.* Infection and Microbiology, 1997. **65**(11): p. 4738-4746.
47. Matz C. and Kjelleberg S., *Off the hook-How bacteria survive protozoan grazing*. Trends in microbiology. **13**(7):p. 302-307.

Chapter 3

III.Experimental set up

It was the first time that different *E. coli* strains were faced to the amoeba *D. discoideum*. For this reason, we had to carry out preliminary experiments to set up our main experiment of coevolution.

Our main coevolution experiment was design to study the temporal variation of two neutral alleles over several generations under various environmental conditions: demography with and without spatial structure and with or without biotic factor.

First, we used a set of strains to study the interactions between *E. coli* strains and *D. discoideum*, confronting a set of *E. coli* strains to *D. discoideum* both in liquid culture media and on agar plates (solid media). We then estimated the growth of *D. discoideum* and of seven *E. coli* strains in the liquid culture media over a period of seven days. Then, we plated on agar plates the entire set of *E. coli* in the presence of *D. discoideum*. Six days later, we observed the appearance of lysis plaques indicating the grazing trait or the absence of lysis plaques indicating grazing resistance trait, for each *E. coli* strain tested.

Then we tested various culture media. Since *E. coli* and *D. discoideum* do not grow on the same standard culture media, this parameter had to be adjusted to allow for sustainable coexistence of the two protagonists.

Third, the bottleneck regime (dilution rate and rhythm) needed adjustment on order to increase the number of generations, while sustaining our culture throughout the entire experiment.

Finally, we studied the evolution of the invasion of various *E. coli* strains within amoeba over three months.

1. STRAINS AND CULTURE CONDITIONS

1.1 Bacteria strains

We used a collection of 37 *E. coli* strains (Chapter IV, Table IV.1, p.111) as previously described [1]. Thirty-one of these strains were isolated from extraintestinal infection and from feces of healthy humans. They represent ten commensal, 21 ExPEC and three laboratory adapted strains.

Commensal strains belong mainly to the phylogenetic group A and B1 whereas ExPEC are mainly from the B2 phylogenetic group. One pathogenic strain belonged to the D phylogenetic group [1].

An *E. coli* HPI mutant and its complemented isogenic strain were tested for *D. discoideum* grazing: the IAI51 *irp1*- mutant (with the *irp1* inactivated gene) and the IAI51 pCP1 mutant (with the re-complementation of the *irp1* gene on plasmid pCP1) (for more details see chapter IV.1, p.104). The HPI-encoded yersiniobactin system of the strain was inactivated by insertion of a kanamycin resistance gene cassette into the *irp1* gene coding for the yersiniobactin protein HMWPI. These strains were tested in a mouse model of extra-intestinal virulence [2].

The laboratory adapted strain B/r (Pasteur Institute) constitutes a standard food for *D. discoideum*. We also used *E. coli* K-12 MG1655 and B REL606, REL 607 (Richard Lenski, Michigan State University) strains which are also phagocytized by *D. discoideum* and harmless. *E. coli* REL 606 and REL 607 are respectively the Ara^- and Ara^+ ancestors. They are isogenic, with the exception of the spontaneous mutation to Ara^+ in REL 607. We used TA agar medium to disentangle Ara^- and Ara^+ strains. Colonies of Ara^- strains appear red on TA agar, while those of Ara^+ strains appear white.

1.2 Bacteria culture conditions

An overnight bacterial culture of 10 ml of HL5 liquid medium with 130 rpm/min agitation was run in a 50 ml Falcon tube, incubated at 37°C. The day of the experiment, the culture was centrifuged (20,000 rpm/min, 20 min), washed once, and re-suspended in MCPB.

1.3 Amoeba strain and culture conditions

The amoeba *D. discoideum* AX3, an axenic strain was grown in 10 ml of HL5 liquid culture medium. Cells from mid-logarithmic cultures were centrifuged (2,000 rpm/min, 7 min) and washed once with MCPB.

1.4 Culture media

We adjusted the culture media for each experiment. We used HL5 the standard rich media for *D. discoideum*,, another standard *D. discoideum* media SM/25 and a *D. discoideum* buffer: MCPB (APPENDIX 1: CULTURE MEDIA, p.197).

1.5 Freezer protocol

A volume of 900 µl of the bacterial culture was mixed with 300 µl of glycerol (60%) and frozen at -80°C. Two replicates were run.

A volume of 500 µl of the co-culture (bacteria and amoebae) was mixed with 500µl DMSO (20%) frozen at -80°C. Three replicates were carried out (APPENDIX 2: FREEZE, p.198).

2. ASSAY PROTOCOL

2.1 Population sizes estimates

2.1.1 Bacteria

In order to carry out the evolution experiment using several replicates, we had to optimize the plating assays. We tested 100 µl or 10 µl of the bacterial culture for plating.

Various percentages of *E. coli* REL 606 were mixed with *E. coli* REL 607: 1%, 10%, 30%, 50%, 70%, 90%, 99%. A volume of 100 µl and 10 µl were plated to test if plating of the two volumes lead to significantly different results. We estimated various proportions of each *E. coli* type in the culture (APPENDIX 3: PROTOCOL: Plating assay, p.199).

There was no significant difference between plating 100 µl or 10 µl (Figure III-1). Thus, for the co-evolution experiments, we plated 10 µl.

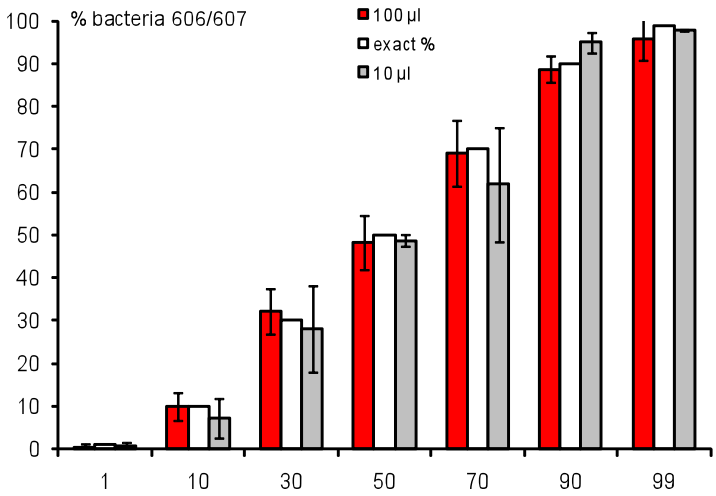


Figure III-1: Test of 100 µl versus 10 µl bacterial culture plated
Different percentages of *E. coli* REL 606 were mixed with *E. coli* REL 607: 1%, 10%, 30%, 50%, 70%, 90%, 99%.

The number of bacteria (unit=CFU colony forming unit) was estimated by plating 10 µl at different dilution rates on TA medium (Figure III-2).

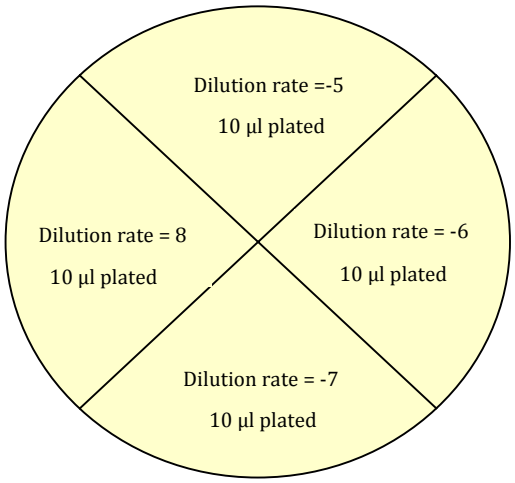


Figure III-2: Bacterial cell number estimation
The plate was divided in four parts, 10 µl culture media plated on each part with different dilution rates.

2.1.2 Amoeba *D. discoideum*

The number of *D. discoideum* was estimated by hemocytometer.

Cell density (culture) = number of *D. discoideum* x 90,000/number square count (Figure III-3).

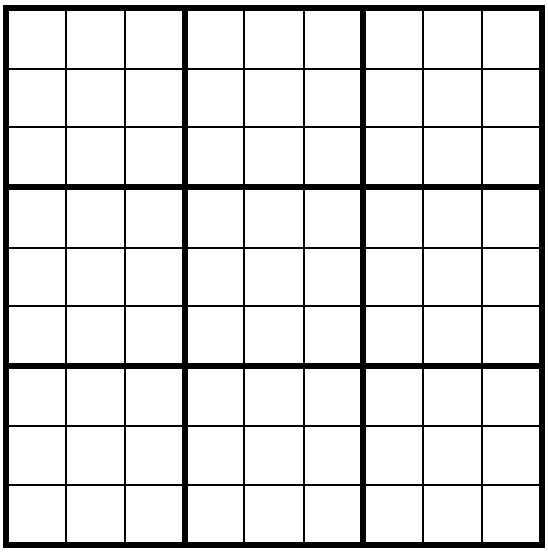


Figure III-3: *D. discoideum* cell number estimation
81-squares grid

2.2 Allele frequency estimation

We used *E. coli* REL 606 and REL 607, the Ara^- and Ara^+ ancestors, respectively. The number of each bacteria strain (unit=CFU colony forming unit) was estimated by plating 10 µl at different dilution rates on TA medium (Figure III-2). We used TA agar medium for distinguishing Ara^- and Ara^+ strains. Colonies of Ara^- strains appear red on TA agar, while those of Ara^+ strains appear white. Different dilutions were prepared in order to obtain the most precise allele frequency estimates.

3. CULTURE CONTROLS

When conducting evolution experiments using different bacteria strains, we controlled for putative contamination among the various cultures each week before freezing them.

This control was performed by adjusting the PCR protocol of Clermont *et al.* [3]. DNA was directly provided by 50 µl of bacterial culture. Each reaction was carried out using a 20 µl mixture containing: 13.35 µl H_20, 0.9 µl dNTP, 0.15 µl *Taq* polymerase, 2 µl of 10x buffer and 20 pmol of each primer.

The primer pairs used were ChuA.1 (GACGAACCAACGGTCAGGAT) and ChuA.2 (TGCCGCCAGTACCAAAGACA), YjaA.1 (TGAAGTGTCAGGAGACGCTG) and YjaA.2 (ATGGAGAATGCGTTCCTCAAC), TspE4C2.1 (GAGTAATGTCGGGGCATTCA) and TspE4C2.2 (CGCGCCAACAAAGTATTACG), which generate 279, 211 and 152 bp fragments, respectively. The PCR was conduct using a thermal cycler under the following conditions: DNA denaturation 15 min at 95°C; 30 cycles of 30 s at 95°C, 30 s at 55°C, 30 s at 72°C and a final extension step of 7 min at 72°C. A volume of 2 µl charged buffer (DNA loading 6X, fermentas) was added to a volume of 10 µl of PCR products and loaded on agarose (1g agarose 1.2%, 85 ml TBE 1X, 5 µl BET) at 110 Volts, 80 mA for the migration. This triplex PCR profile is specific for each *E. coli* phylogenetic group (Figure III-4).

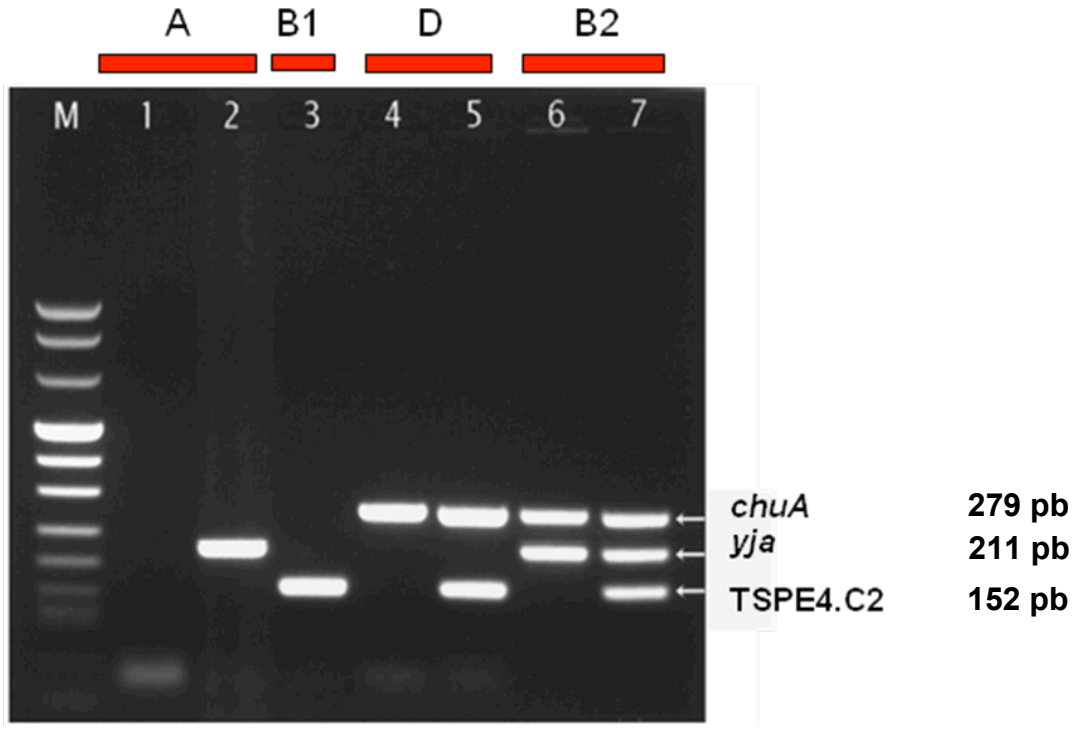


Figure III-4: Triplex PCR profiles for each *E. coli* phylogenetic group
Each combination of *chuA* and *yjaA* gene and DNA fragment TSPE4.C2 amplification allowed phylogenetic group determination of strains. Lanes 1 and 2, group A; lane 3, group B1; lanes 4 and 5, group D; lanes 6 and 7, group B2. Lane M contained markers.

4. *E. COLI* AND *D. DICTYOSTELIUM* INTERACTIONS IN LIQUID MEDIA

4.1 Strains and culture conditions

4.1.1 Bacteria

We used two laboratory adapted *E. coli* strains, two human commensal strains, and one human pathogenic strain (Table III-1).The laboratory strains, *E. coli* B/r (Pasteur Institute) and *E. coli* K-12 MG1655 constitute a standard food for *D. discoideum*, they are phagocytized by *D. discoideum* and are harmless to *D. discoideum*. The *E. coli* strain IAI1 is a human commensal strain, belonging to the B1 phylogenetic group without virulence genes and harmless to the mouse model.

The *E. coli* strain ED1a is a human commensal strain that belong to the B2 phylogenetic group and carries virulence genes, but is harmless to the mouse model.

The *E. coli* strain RS218 is a human pathogenic strain belonging to the B2 phylogenetic group that carries some virulence genes and is virulent to the mouse model.

Table III-1: Characteristics of *E. coli* strains used

E. coli Strain	Effect on human[a]	Effect on mouse model[b]	Presence of virulence genes[c]	Phylogenetic group[d]
B/r	L	NA	-	A
K12-MG1655	L	NK	-	A
IAI1	C	NK	-	B1
ED1a	C	NK	+	B2
RS218	P	K	+	B2

a: L: Laboratory adapted strain
C: human commensal strain
P: human pathogenic strain

b: NA: non tested
NK: Non mouse killer [1]
K: mouse killer [1]

c: (-): absence of virulence genes
(+): presence of virulence genes

d: phylogenetic groups determined as in [1]

An overnight bacterial culture of 10 ml of HL5 liquid medium with 130 rpm/min agitation was prepared in a 50 ml Falcon tube, incubated at 37°C. The day of the experiment, the culture was centrifuged (20,000 rpm/min, 20 min), washed once and re-suspended in MCPB.

4.1.2 Amoeba

The amoeba *D. discoideum* AX3, an axenic strain, was grown in 10 ml of HL5 culture liquid medium. Cells from mid-logarithmic cultures were centrifuged (2,000 rpm/min, 7 min) and washed once with MCPB.

4.2 Assays

4.2.1 Growth in co-culture

Each *E. coli* strain was mixed with *D. discoideum* at 5.10^4 cells/ml concentration at the beginning of the experiment. Each day, the density of each population was estimated. Controls were conducted with *E. coli* alone and *D. discoideum* alone. Three replicates were carried out for each sample.

4.2.2 Test of resource use

To test whether various effects of bacteria strains on amoeba could be explained by different level of resource used, we first prepared a bacteria/amoeba co-culture and an amoeba control culture. After four days of these co-cultures, we filtered them (thus removing the cells) and added fresh amoebae. The density of amoeba was monitored daily. Three replicates were run for each sample.

4.3 Statistical analyses

For the growth experiment, a χ^2 test was applied to test for a significant difference at each time between the control (*D. discoideum* growth alone) and the set of bacterial strains co-cultured with amoeba.

For the test of resource used, a χ^2 test was carried out to test for a significant difference between the set of bacteria/ amoeba co-cultures.

4.4 Results

4.4.1 Growth in co-culture

We observed a substantial natural variability among bacteria strains for their effect on amoeba (Figure III-5). Nevertheless, the human pathogenic strain (RS218) was not the one that showed the most negative effect on amoeba growth compared to the two commensal strains (IAI1 and ED1a). Amoeba did not substantially affect bacterial growth (Figure III-6).

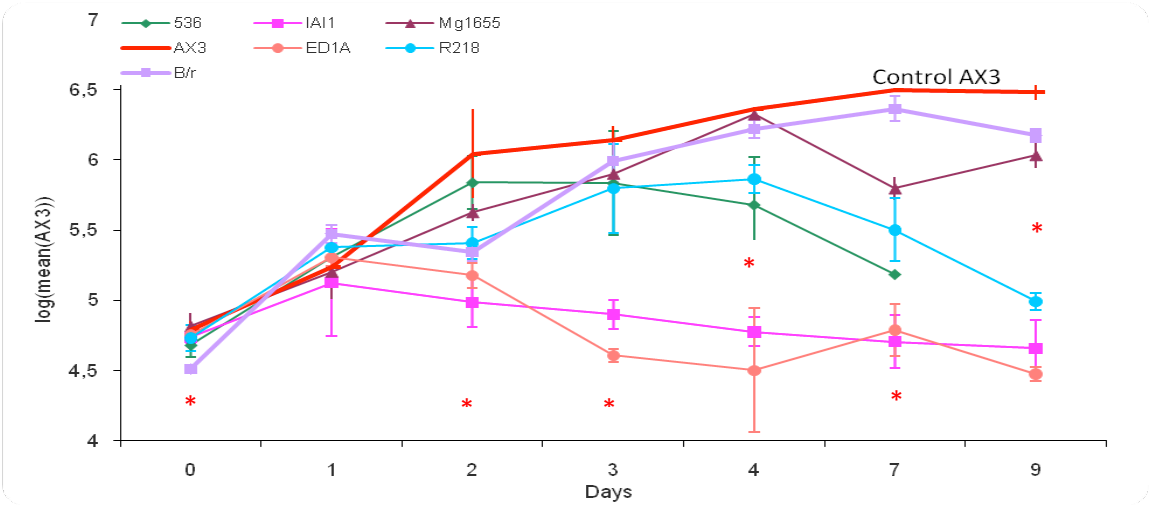


Figure III-5: *D. discoideum* growth
D. discoideum density was assessed over seven days with different *E. coli* strains. (*) indicates significant differences with the control (*D. discoideum* without *E. coli*).

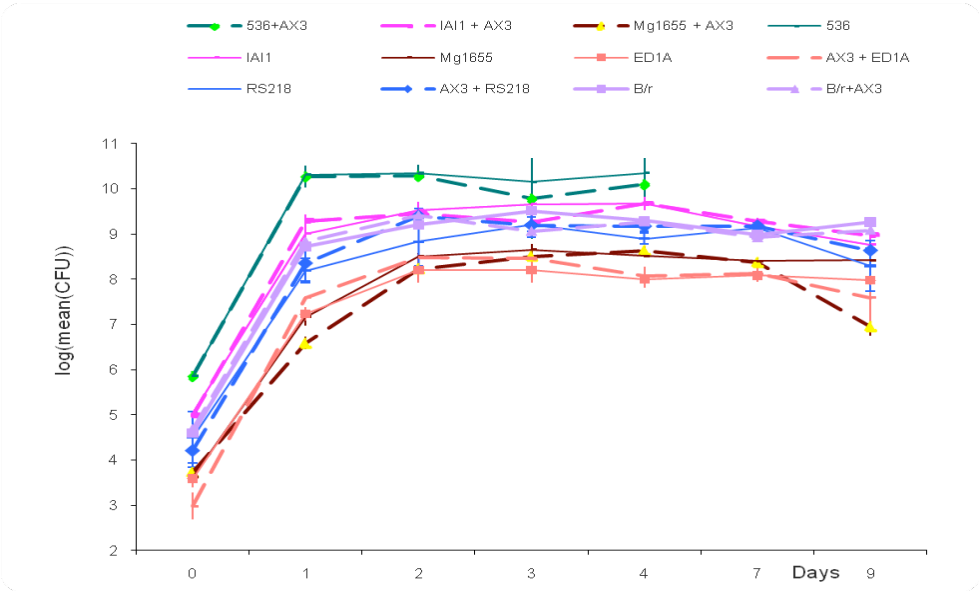


Figure III-6: *E. coli* growth
E. coli density was assessed over seven days with (--) and without (-) *D. discoideum*.

Both species (bacteria and amoeba) survived and could be maintained when mixed over one week, but a significant amoeba growth decrease was observed.

The co-culture of each bacterial strain showed no differential growth that could explain their differential effects on amoeba (Figure III-6).

4.4.2 Test of resource use

The media exhausted by the various *E. coli* strains showed no difference on amoeba growth suggesting that their resource use did not differ (Figure III-7).

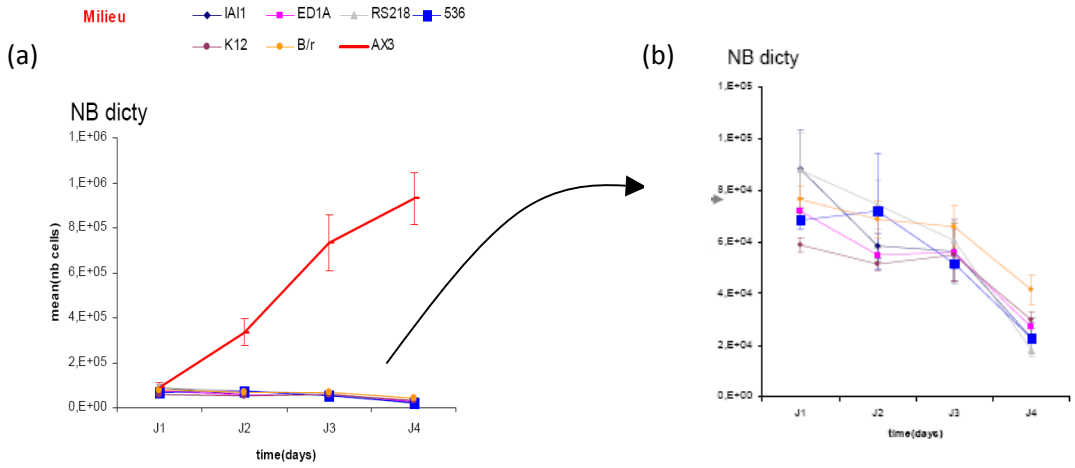


Figure III-7: Test of resource use

After four days of a bacteria/amoeba co-culture the culture was filtered (thereby removing cells) and fresh amoebae were added. The density of amoeba was monitored daily.

(a) *D. discoideum* growth

(b) zoom of the *D. discoideum* growth with the bacteria/amoeba co-culture media filtered

5. *E. COLI* AND *D.DISCOIDEUM* INTERACTION ON AGAR PLATES (solid medium)

The material and methods and results of these experiments are details in the chapter IV p.104.

In this paper, we assessed how resistance to grazing correlates with some bacterial traits such as the presence of virulence genes by subjecting a well-characterised collection of 31 *Escherichia coli* strains (human commensal or extra-intestinal pathogenic) to grazing by the social haploid amoeba *Dictyostelium discoideum*.

Whatever the relative population size (bacteria/amoeba) for a non-pathogenic bacteria strain, *D. discoideum* was able to phagocytise, digest and grow. In contrast, a

pathogenic bacterium strain killed *D. discoideum* above a certain bacteria/amoeba population size. A plating assay was then carried out using the *E. coli* collection confronted with *D. discoideum* grazing. *E. coli* strains carrying virulence genes such as *iroN, irp2, fyuA* involved in iron uptake, which belong to the B2 phylogenetic group, and which are virulent in a mouse model of septicaemia were resistant to grazing from *D. discoideum*. Experimental proof of the key role of the *irp* gene in the grazing resistance was evidenced with a mutant strain lacking this gene. These virulence determinants may well have been originally selected and (or) further maintained for their role in the natural habitat, including resistance to digestion by free-living protozoa, rather than for virulence *per se*.

6. CULTURE MEDIA AND BOTTLENECK REGIME

6.1 Strains and culture conditions

To carry out co-evolution experiments with *E. coli* and *D. discoideum*, we tested various culture media that allowed these two organisms to survive.

We used the laboratory adapted strains *E. coli* REL 606 and the amoeba *D. discoideum*.

For the bottleneck regime experiment, we used only amoeba.

Cultures were prepared as in Chapter III.1.2, p.85.

6.2 Assays

6.2.1 Test of culture media

The density of each culture was adjusted to 5.10^4 cells/ml at the beginning of the experiment. We tested two culture media: MCPB (buffer) and SM/25. The density of each population was estimated each day. Three replicates were run for each sample.

6.2.2 Bottleneck regime

To maintain the cultures over a long period and maximize the number of generations, we had to dilute each sample at the same dilution rate and rhythm.

Each sample was diluted at the same rate and rhythm. Two dilution rates were tested: 10^{-1} and 10^{-2} with a daily rhythm (APPENDIX 4: PROTOCOL: Bottleneck regime, p.199). Three replicates of each sample were run.

6.3 Results

6.3.1 Test of culture media

To carry out co-evolution experiments with *E. coli* and *D. discoideum*, we tested different culture media that allowed these two organisms to survive. With the MCPB media, cells (amoeba and bacteria) were hardly able to grow in one day in comparison with the SM/25 media, in which amoebae were able to grow compared to bacteria (Figure III-8).

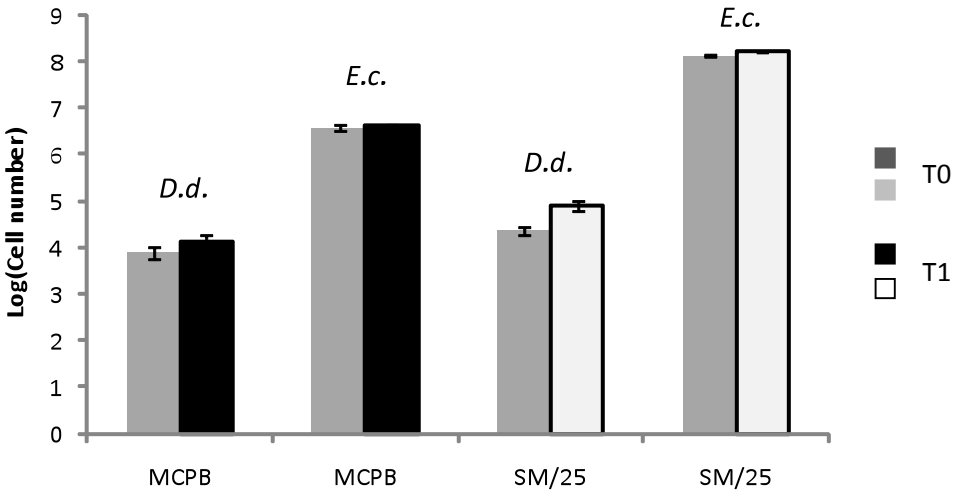


Figure III-8: Growth of *D. discoideum* and *E. coli* on different culture media
The density was assessed over two days in two different media: SM/25 and the buffer MCPB.
D.d.: D. discoideum
E.c. : E. coli

6.3.2 Bottleneck regime

The amoeba growth rates showed that the dilution rates should not be greater than 1/100th and the frequency should not exceed one dilution over two days (Figure III-9 and Figure III-10).

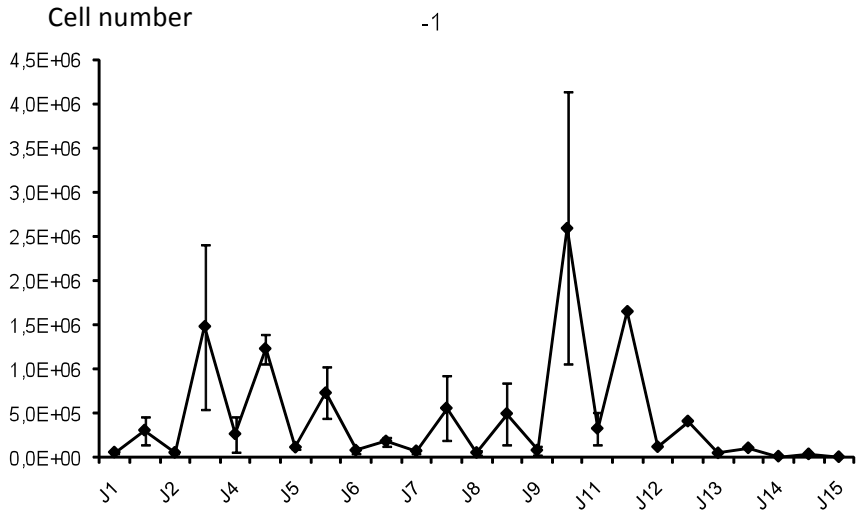


Figure III-9: *D. discoideum* growth with a 1/10 dilution

The *D. discoideum* density was measured daily with 1/10 dilution per day

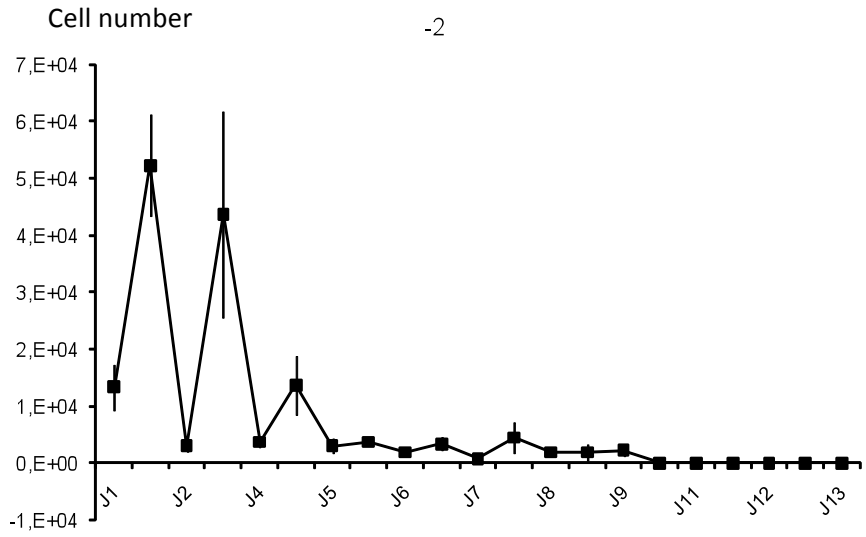


Figure III-10: *D. discoideum* growth with 1/100 dilution

The *D. discoideum* density was measured daily with 1/100 dilution per day.

7. THE EVOLUTION OF BACTERIA RESISTANCE TO AMOEBA PHAGOCYTOSIS

7.1 Strains and culture conditions

Five *E. coli* strains were used for this experiment. One human pathogenic strain and its isogenic strain without virulence traits, two commensal strains with different phylogenetic groups and one laboratory strain without virulence trait (Table III-2).

Table III-2: Characteristics of *E. coli* strains used

Strain	Origin	Group	*Sfa/foc*	*iroN*	*aec*	*papC*	*papG*	*hly*	*cnfI*	*hra*	*fyuA*	*Irp2*	PAI I	PAI II	PAI III	HPI
536	U	B2	+	+	-	+	III	+	-	+	+	+	-	+	+	+
536ΔHPI	U	B2	-	-	ND	-	-	-	ND	ND	ND	ND	-	-	-	-
REL606	L	A	-	-	-	-	-	-	-	-	-	-				
IAI1	C	B1	-	-	-	-	-	-	-	-	-	-	-	-	-	-
IAI4	C	A	-	-	-	-	-	-	-	-	-	-	-	-	-	-

Origin: U (urinary tract), L (laboratory), C(commensal); phylogenetic Group and virulence genes (+) presence, (-) absence.

Amoebae and bacteria cultures were prepared as in Chapter III.1.2, p.85.

7.2 Assays

7.2.1 Evolution experiments

Five *E. coli* strains were cultured with or without (control) *D. discoideum* (three replicates per strain), in 10 ml SM/25 culture media without agitation, at 24°C. Culture media was replaced every three days, with 5 ml transferred to a new bottle containing 5 ml fresh media or 10 ml removed and fresh media added in the same bottle.

7.2.2 Control

A triplex PCR was carried out weakly before freezing to control each culture for possible contamination (see Chapter III.3 p.87).

7.2.3 Assessment of the evolution of bacteria resistance to amoeba phagocytosis

For each culture, an antibiotic (Gentamycin 5 mg/ml) was added during 45 minutes and subsequently removed by centrifugation (2,000 rpm/min, 15 min). This protocol allowed us to keep only bacteria inside amoeba and remove free bacteria in the media [4]. Amoeba density was estimated by the use of a hemocytometer. Amoebae were then lysed with SDS 20% for two minutes, then removed by centrifugation (10,200 rpm/min, 15 min). Intracellular bacteria were estimated by plating 10µl of this culture with a set of dilution rates.

7.2.4 Statistical analyses

An ANOVA was applied using JUMP (SAS) software. We tested for each effect: strains, phylogenetic group, replicate, number of bacteria generation.

7.3 Results

7.3.1 Evolved *E. coli* with evolved *D. discoideum*

Statistical analyses showed that two factors had significant effects: the number of bacterial generations and the bacterial phylogenetic group. When *D. discoideum* had coevolved with *E. coli*, the number of intracellular bacteria increased with the number of bacterial generations ($F_{7,7}$=23,1058, p<0,0001) (Figure III-11).

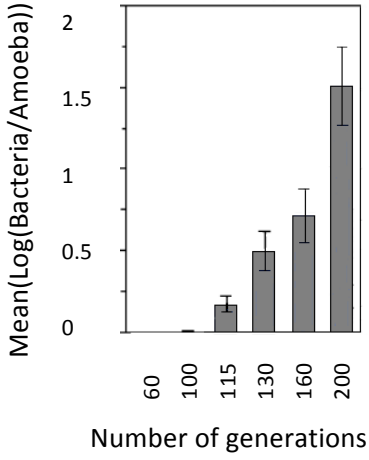


Figure III-11: Evolution of intracellular bacteria evolved *E. coli* confronted with evolved *D. discoideum*
The number of intracellular bacteria increased with the number of bacterial generations.

Statistical analyses showed that depending on the bacteria phylogenetic group, the number of intracellular bacteria were different ($F_{2,2}$=3.61; p=0.0302) (Figure III-12). The strain belonging to the phylogenetic group (B2) showed the greatest number of

intracellular bacteria. *E. coli* strains belonging to this group are generally the most virulent strains [5].

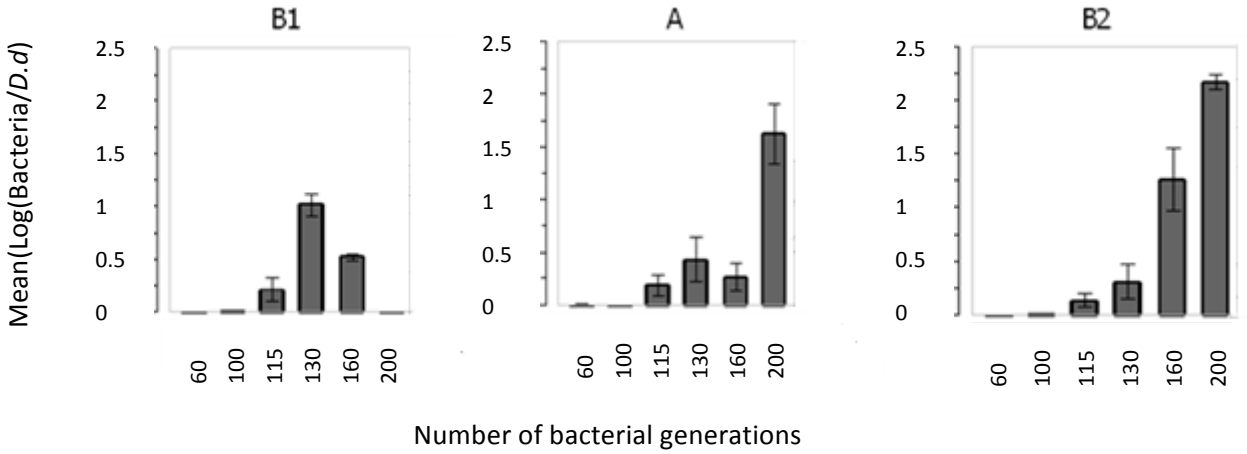


Figure III-12: Evolution of intracellular bacteria: evolved *E. coli* confronted with evolved *D. discoideum*, by phylogenetic group

For the phylogenetic group B1, intracellular bacteria increased at the beginning, before decreased. For the phylogenetic groups A and B2, intracellular bacteria increased with the number of bacteria generations.

7.3.2 Evolved *E. coli* confronted with ancestral *D. discoideum*

We confronted evolved *E. coli* to ancestral and evolved *D. discoideum*, estimating the number of intracellular bacteria at three different bacteria generation numbers. Depending on the *D. discoideum* status (evolved or ancestral), the number of intracellular bacteria was significantly different (Figure III-13).

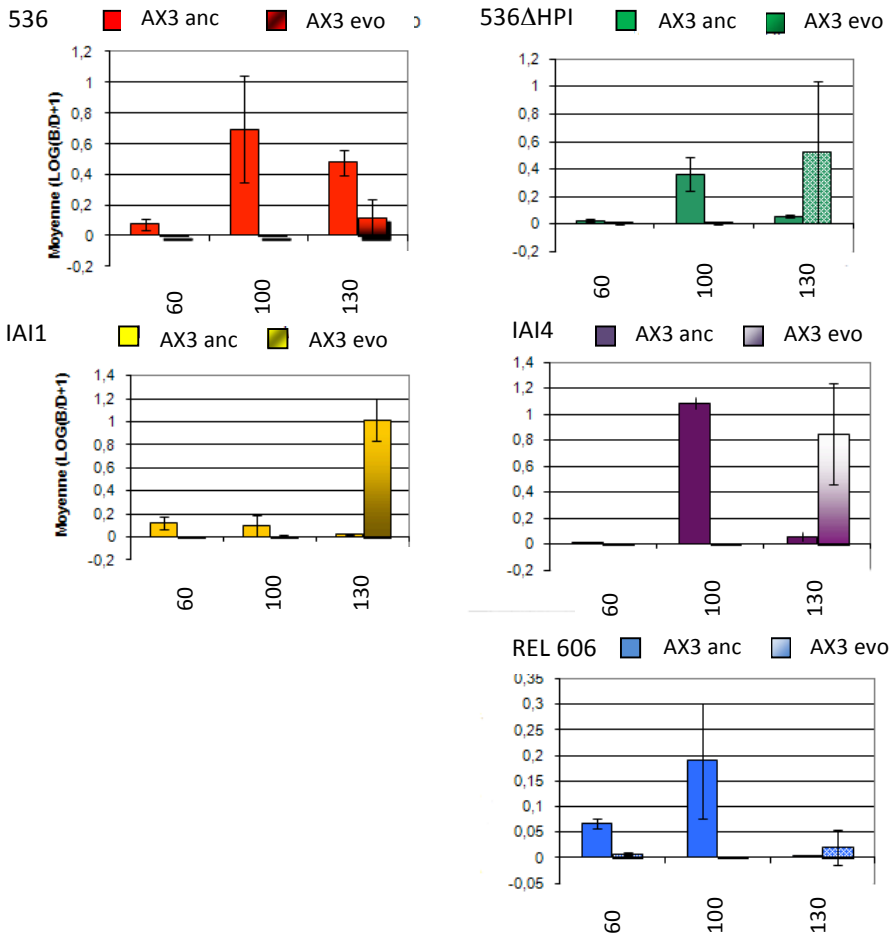


Figure III-13: Comparison of the evolution of intracellular bacteria level between evolved *D. discoideum* and ancestral *D. discoideum* by *E. coli* strain

For each bacteria strain, we measured intracellular bacteria when confronted with evolved or ancestral *D. discoideum* at different bacteria generation numbers of the coevolution experiment.

For the pathogenic strain *E. coli* 536, the number of intracellular bacteria was greater with the ancestral *D. discoideum* than to the evolved strain for each time studied.

For the *E. coli* Δ536, IAI1, IAI4 and REL606 strains, we also observed a greater number of intracellular bacteria with the ancestral *D. discoideum* compared to the evolved one, but only for the first two times. The third time showed a greater number of intracellular bacteria with evolved *D. discoideum* compared to the ancestral one, suggesting the evolution of amoeba resistance to bacterial virulence and its ability to phagocytize bacteria.

This co-evolution experiment showed that the number of intracellular bacteria evolved. After several bacteria generations, the ability of amoeba to phagocytize

bacteria was greater. We will now have to compare ancestral and evolved bacteria confronted with ancestral amoeba in order to confirm the evolution of the bacterial resistance to digestion as well as to carry out the same experiments with other bacteria strains.

Literature cited

1. Diard, M., et al., *Caenorhabditis elegans as a simple model to study phenotypic and genetic virulence determinants of extraintestinal pathogenic Escherichia coli.* Microbes and infection, 2007. **9**(2): p. 214-223.
2. Schubert, S., et al., *Yersinia high-pathogenicity island contributes to virulence in Escherichia coli causing extraintestinal infections.* Infection and immunity, 2002. **70**(9): p. 5335.
3. Clermont, O., S. Bonacorsi, and E. Bingen, *Rapid and simple determination of the Escherichia coli phylogenetic group.* Applied and Environmental Microbiology, 2000. **66**(10): p. 4555.
4. Alsam, S., et al., *Escherichia coli interactions with Acanthamoeba: a symbiosis with environmental and clinical implications.* Journal of medical microbiology, 2006. **55**(6): p. 689.
5. Picard, B., et al., *The link between phylogeny and virulence in Escherichia coli extraintestinal infection.* Infection and immunity, 1999. **67**(2): p. 546.

Chapter 4

IV. Results

From grazing resistance to pathogenesis: the coincidental evolution of virulence factors

1. FROM GRAZING RESISTANCE TO PATHOGENESIS: the coincidental evolution of virulence factors

The interactions between *E. coli* and *D. discoideum* were complex and we used liquid culture media and agar plates to study them.

Free-living protozoa affect bacterial populations in multiple ways. They may act as predators but may also provide a protective environment and even act as vectors. Moreover, mechanisms that improve resistance to digestion by predators such as free-living amoeba may well express themselves as virulence factors in other organisms. These virulence factors understand adherence, digestion resistance, intracellular toxin production and intracellular replication.

Various theories for the evolution and the maintenance of virulence factors have been proposed. We tested here one of these theories here: the coincidental evolution of virulence factors.

From grazing resistance to pathogenesis: the coincidental evolution of virulence factors

Sandrine Adiba[1,4*], Clément Nizak[2], Minus van Baalen[1,4],
Erick Denamur[3**] and Frantz Depaulis[4**]

[1]Laboratoire d'Ecologie, CNRS UMR7625, Université Pierre et Marie Curie, Paris Universitas, 7 quai St-Bernard, 75252 Paris, France

[2]Laboratoire de Spectrométrie de Physique, CNRS UMR5588,
140 Avenue de la Physique, 38402 Saint Martin d'Hères, Grenoble, France

[3]INSERM U722 and Université Paris Diderot
16 rue Henri Huchard, 75018 Paris, France

[4]Laboratoire d'Ecologie, CNRS UMR7625, Ecole Normale Supérieure
46 rue d'Ulm, 75230 Paris cedex 05, France

*To whom correspondence should be addressed. E-mail:sadiba@snv.jussieu.fr.
**The two last authors contributed equally to this work.

Published in PLoS One – August 2010

Abstract

To many pathogenic bacteria, human hosts are an evolutionary dead end. This raises the question of why evolutionary forces have shaped their virulence traits. Why are these bacteria so virulent? The coincidental evolution hypothesis suggests that such virulence factors result from adaptation to other ecological niches. In particular, virulence traits in bacteria might result from selective pressure exerted by protozoan predator. Thus, grazing resistance may be an evolutionarily exaptation for bacterial pathogenicity.

This hypothesis was tested by subjecting a well characterized collection of 31 *Escherichia coli* strains (human commensal or extra-intestinal pathogenic) to grazing by the social haploid amoeba *Dictyostelium discoideum*. We then assessed how resistance to grazing correlates with some bacterial traits, such as the presence of virulence genes.

Whatever the relative population size (bacteria/amoeba) for a non-pathogenic bacteria strain, *D. discoideum* was able to phagocytise, digest and grow. In contrast, a pathogenic bacterium strain killed *D. discoideum* above a certain bacteria/amoeba population size. A plating assay was then carried out using the *E. coli* collection faced to the grazing of *D. discoideum*. *E. coli* strains carrying virulence genes such as *iroN, irp2*, *fyuA* involved in iron uptake, belonging to the B2 phylogenetic group, and being virulent in a mouse model of septicaemia, were resistant to grazing by *D. discoideum*. Experimental proof of the key role of the *irp* gene in the grazing resistance was evidenced with a mutant strain lacking this gene. Such determinant of virulence may well have been originally selected and (or) further maintained for their role in natural habitat: resistance to digestion by free-living protozoa, rather than for virulence *per se*.

Key words: coincidental evolution, ecology, virulence, grazing protozoa, *Escherichia coli*, *Dictyostelium discoideum*, amoeba

Introduction

The evolution and the maintenance of virulence are among the most important topics addressed in recent years by evolutionary biologists. Different theories have been proposed and become the focus of intense debate. The “conventional wisdom” [1] also known as the “avirulence hypothesis” suggests that parasites would always evolve to become avirulent. Virulence is then considered to be a maladaptation of new or recent associations between parasites and hosts. This hypothesis was challenged during the 1980’s. A number of theories [2,3,4,5,6,7] proposed that virulence would be maintained by natural selection and should depend on the mechanism of transmission. This “trade-off” hypothesis between virulence and transmission is still investigated [8]. Two alternative models, namely short-sighted within host-selection and the coincidental evolution, are two other ways by which natural selection can favor virulence, irrespective of any relationship between virulence and transmission. Our work will focus on the last hypothesis: the coincidental evolution which hypothesizes that virulence is a coincidental by-product of the adaptation to other ecological niches.

It is well known that free living protozoa may affect bacterial populations in multiple ways. They may act as predators but also provide a protective environment and even act as vectors [9]. Their long co-evolutionary history suggests that a series of adaptations ensuring bacterial survival should have emerged. Moreover, mechanisms that improve resistance to digestion by predators such as free-living amoebae may well express themselves as virulence factors in other organisms [10]. Thus bacterial traits such as adherence, digestion resistance, intracellular toxin production and intracellular replication [11,12] or outer-membrane structures [13] might facilitate a transition from free-living organism to intracellular parasite. The ability of bacteria to resist and exploit a normally hostile niche to ensure survival was discovered in the bacteria *Legionella pneumophila,* which is able to multiply within certain species of free-living amoeba [14]. The ability of *L. pneumophila* to parasitize macrophages and cause human disease may thus well be a consequence of its resistance to amoebae. For many diseases such as Legionnaire disease, Lyme disease, pneumonia disease caused by Hantavirus, human play no (or at best, a negligible one) role in the transmission, and are then an evolutionary dead-end [15]. Genes that cause these effects must therefore be favored by their effects elsewhere. Such cases in

which virulence is obviously of no selective value (neutral) [16], led to the hypothesis of the coincidental evolution of virulence factors. Adaptation to commensal and/or saprophyte habitats may thus coincidentally promote some virulence factors, rather than selection for virulence *per se* [17]. However, the coincidental evolution hypothesis has not been tested.

Escherichia coli is a good biological model to investigate such coincidental evolution. This Gram-negative and facultatively anaerobic bacterium occurs both as commensal and pathogenic strains. Commensal *E. coli* (most strains) use their host as a suitable habitat without any noteworthy benefit or detriment for the host, whereas pathogenic strains may cause harm or even death [18]. Even so-called extra-intestinal pathogenic *E. coli* (ExPEC) strains found in mammal intestinal tracts do not normally cause disease, but they do, when they incidentally invade sterile niches such as urinary tract, blood [19,20] or cerebrospinal fluid. Besides this primary habitat, *E. coli*'s secondary habitat involves the external environment, such as water, sediment, and soil [21,22,23] where some specific strains may grow [24,25,26].

E. coli population genetic structure is mainly clonal, with four main phylogenetic groups denoted A, B1, B2 and D [27,28]. Each phylogenetic group shows various life history traits and differs in its phenotypic and genotypic characteristics [27,29]. Compared to commensal strains, ExPEC have additional virulence factors such as adhesins (e.g., P and I fimbriae), iron-acquisition systems (e.g., aerobactin), host defence-avoidance systems (e.g., capsule, lipopolysaccharide), and toxins (e.g., hemolysin, cytotoxic necrotizing factor1) [30,31]. ExPEC virulence factors, encoded by genes often clustered on so-called pathogenicity island (PAI), are mainly found in strains of the phylogenetic group B2 and D, but rarely in the phylogenetic groups A and B1, to which the majority of commensal and extraintestinal pathogenic *E. coli* strains belong [32,33,34]. The maintenance and the evolution of these genes associated with virulence within the phylogenetic group B2 suggest that virulence could be a coincidental by-product, since these genes are implicated in complex commensal niche colonization [15,35], in which bacteria are subjected to grazing by protista, such as free-living amoeba.

In a similar fashion, amoebae that occur in the human intestinal tract (*Entamoeba* species) [36,37,38] typically also occur in external environment too, in habitats such as fresh water, moist soil, at various interfaces (water-soil, water-plant, water-air)

[39]. They are distributed worldwide, but the set of species in a particular location depends on abiotic factors and food availability. Amoebae are particularly effective bacterial predators. Free-living amoeba is the only group of organisms that can lead to a decrease of a bacterial population from 10^8 down to 10^5 per gram of soil [40,41,42]. The haploid social amoeba *Dictyostelium discoideum* has been used as a model to study host-factors involved in cellular aspects of host-pathogen interactions [43,44]. The common mechanisms of infection required to infect both mammals and *D. discoideum* by a wide variety of pathogens reinforces the use of the *E. coli/D. discoideum* system as a valid model to study host-pathogen relations and their evolution [43,45]. In addition, elaborate genetic tools including genomic sequence [46] and plasmids allowing constitutive or inducible expression of genes are now available.

In the context of the variety of ecological and evolutionary factors involved in the interconnected life histories of bacteria and amoeba, we investigated here evolutionary forces selecting and maintaining virulence factors. To test the coincidental evolution hypothesis, a first goal was to test whether ecological interactions between *D. discoideum* and *E. coli* have an effect. The first experiment details the relative population size effect (bacteria/amoeba) on amoeba grazing resistance. Secondly, the ability for pathogenic bacteria to kill or inhibit amoeba growth was assessed. Our main experiment compares the grazing resistance of a collection of 31 human commensal and pathogenic *E. coli* strains to predation by the social amoeba *D. discoideum* and assesses the correlation of resistance with known virulence genes. The results we obtained suggested that the so-called High pathogenicity island (HPI) of *E. coli* plays an important role. This role was further assessed using a mutant bacterial strain with inactivated genes and its complemented isogenic strain.

Material and methods

Strains and culture conditions

Bacterial strains

The collection of *E. coli* strains (Table IV-1) has been previously described [72,73]. These strains were isolated from extraintestinal infections and from feces of healthy humans. They represent 10 commensal, 21 ExPEC and three laboratory adapted strains.

Commensal strains belong mainly to the phylogenetic group A and B1 whereas ExPEC are mainly from the B2 phylogenetic group. One pathogenic strain belongs to the D phylogenetic group.

An *E. coli* HPI mutant and its complemented isogenic strain were tested for *D. discoideum* grazing: the IAI51 *irp1*- mutant (with the *irp1* inactivated gene) and the IAI51 pCP1 mutant (with the recomplementation of the *irp1* gene on plasmid pCP1). The HPI-encoded yersiniobactin system of the strain was inactivated by insertion of a kanamycin resistance gene cassette into the *irp1* gene coding for the yersiniobactin protein HMWPI. These strains were tested in a mouse model of extra-intestinal virulence [48].

The laboratory adapted strain B/r (Pasteur Institute) is a standard food for *D. discoideum*. We also used *E. coli* K-12 MG1655 and B REL606 (Richard Lenski, Michigan State University) strains which are also phagocytized by *D. discoideum* and are harmless.

Table IV-1: Phenotypic and genotypic characteristics of the E. coli strains studied

	Strain ID	Phylogenetic groups [a]	Virulence gene [b]											Phenotype [c]					Model tested [d]	
			K1	afa/draBC	fyuA	iucD	papC	papG	hly	cnf1	hra	kpsMII	traT	RSer	RLysoz	RBil	Mot	GenTim	Mouse	Amoeba
Laboratory adapted strain	K-12 MG1655	A	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	NK	G
	B/r		ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	G
	B REL 606	A	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	G
Commensal strain	Ben13f	B2	-	+	+	-	+	III	-	-	+	+	+	+	-	+	-	+	K	GR
	ColF6c	B2	-	-	-	-	+	II	+	+	+	+	+	+	+	+	-	+	K	GR
	ECOR58	B1	-	+	+	-	-	-	-	-	-	-	-	+	+	+	+	+	K	GR
	ED1a	B2	-	-	-	+	-	-	-	-	-	+	+	-	-	-	-	-	NK	GR
	IAI1	B1	-	-	-	-	-	-	-	-	-	-	-	+	-	+	-	+	NK	GR
	IAI12	A	-	-	-	-	-	-	-	-	-	-	-	+	-	+	-	+	NK	G
	IAI13	A	-	-	-	-	-	-	-	-	+	-	-	-	-	-	-	-	NK	G
	IAI15	A	-	-	-	-	-	-	-	-	-	+	+	-	-	-	+	-	NK	G
	IAI2	B1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	NK	G
	IAI4	A	-	-	-	-	-	-	-	-	-	-	-	-	-	+	-	+	NK	G
Pathogenic strain	536	B2	-	+	+	-	+	III	+	-	+	+	+	+	+	+	+	+	K	GR
	CFT073	B2	-	+	+	+	+	II	+	-	-	+	+	+	+	-	+	-	K	GR
	EC7372	B2	-	-	-	+	+	II	+	-	-	+	+	+	-	+	+	+	K	GR
	ECOR64	B2	-	+	+	-	-	-	-	-	-	+	+	+	+	-	-	-	K	GR
	F11	B2	-	+	+	-	+	III	+	+	+	+	+	+	+	+	+	+	K	GR
	F63	B2	+	+	+	-	+	III	+	+	+	+	+	+	+	+	+	-	K	GR
	IAI19	A	-	-	-	-	-	-	-	-	-	-	-	+	+	-	+	-	NK	G
	IAI21	B1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	NK	G
	IAI39	D	+	-	-	+	+	II	-	-	-	+	+	-	-	-	-	-	K	G
	IAI44	A	-	-	-	-	-	-	-	-	-	+	+	-	-	-	+	-	NK	GR
	IAI48	B2	-	-	-	+	-	-	-	-	-	-	-	+	-	+	-	-	NK	GR
	IAI49	B2	-	-	+	-	-	-	-	-	-	+	+	+	+	-	+	+	K	GR
	IAI51	B2	+	+	+	-	-	-	-	-	-	+	+	-	+	-	-	+	K	GR
	IAI52	B2	+	+	+	-	-	-	-	-	-	+	+	-	+	-	-	+	K	G
	IAI60	B2	+	-	+	+	+	II	-	-	-	+	+	-	-	+	+	+	NK	GR
	IAI64	B2	-	+	+	-	+	III	+	+	+	+	+	-	-	-	+	+	K	G
	IAI72	B2	-	+	+	+	+	II	+	+	+	+	+	-	+	+	-	-	K	GR
	IAI73	B2	-	+	+	-	+	III	+	+	+	+	+	+	+	+	+	-	K	GR
	IAI74	B2	+	+	+	-	+	III	+	+	+	+	+	-	+	-	+	+	K	GR
	J96	B2	-	+	+	-	+	I+III	+	+	+	+	+	-	+	+	+	-	NK	G
	R3218	B2	+	+	+	-	+	III	+	+	+	+	+	+	+	+	+	+	K	GR

(a) Phylogenetic groups determined as in [74]

(b) Virulence genes determined as in [73], (+) presence, (-) absence. All these data are from [72] except the amoeba model data that have been generated in this work

(c) Phenotypes tested as in [72], abbreviations used are: RSer: serum resistance, RBil: bile resistance, RLys: lysozym/lactoferrin resistance, Mot: motility, GenTim: generation time. Binarization of data in [72]

(d) Biological model tested: mouse [73, 72], K: mouse killer (9 or 10 mice killed over the 10 inoculated), NK: mouse non-killer (no mouse lethality or 1-5 mice over the 10 inoculated), *D.discoideum*, G: grazing, GR: grazing resistance.

Bacterial culture conditions

For all the experiments and strains, an overnight culture in 10 ml of HL5 medium (5 g/l proteose peptone, 5 g/l thiotone E peptone, 10 g/l glucose, 5 g/l yeast extract, 0.35 g/l Na_2HPO_4, 7 H_2O, 0.35 g/l KH_2PO_4, pH 6.5) with 130 rpm/min agitation was prepared in a 50 ml Falcon tube, incubated at 37°C. On the day of the experiment, the culture was centrifuged (20,000 rpm/min; 20 min), washed once, and re-suspended in MCPB (1.42 g/l Na_2HPO_4, 1.36 g/l KH_2PO_4, 0.19 g/l $MgCl_2$, 0.03 g/l $CaCl_2$, pH 6.5). Cells were then plated on 5.5 cm diameter plates of HL5 agar.

Amoeba strain and culture conditions

The amoeba *D. discoideum* AX3, an axenic strain was used in all the experiments. Amoebae were grown in 10 ml of liquid HL5. Cells from mid-logarithmic cultures were then centrifuged (2,000 rpm/min; 7 min) and washed once with MCPB.

For all experiments, a volume of 300 µl was plated. It corresponds to the minimum volume covering all the Petri dish (5.5 cm diameter).

Assays

Relative population size effects

Bacterial strains used in these experiments were the pathogenic strain 536 and the non pathogenic strain B REL606.

In the first part of this experiment, the bacterial population size effect was tested with 10^4 to 10^8 bacterial cells faced to 100 amoebas. In the reverse experiment, the amoeba population size effect was tested with 10^2 to 10^6 amoebae cells faced to 10^8 bacteria. Relative population size for both experiments were 10^2, 10^3, 10^4, 10^5, 10^6 bacteria cells/amoebae.

A volume of 300 µl bacterial culture was plated on HL5 agar Petri dishes and allowed to dry for 20 min under a sterile air flow. The same volume of amoeba culture was added to these plates and allowed to dry for 20 min under a sterile air flow. Plates were covered with parafilm and incubated for six days at 24°C. Three replicates were performed for each experiment. Plates were screened at day three

and day six to assess the occurrence of bacterial lysis plaques formed by *D. discoideum*.

Death versus growth inhibition

In the case where no bacterial lysis was observed, it is not possible to distinguish between killing or growth inhibition of amoebae. Another experiment was therefore carried out under the same conditions as above, using 10^8 bacterial cells (REL606 or 536 strains) and 10^2 amoebae. Plates were examined two and six days after plating. Three replicates were performed. To disentangle death and growth inhibition, the layer was transferred from plates to a liquid MCPB medium and amoeba and bacterial cells were screened with a Bac-live Kit that colors viable cells in green fluorescent and membrane damages in red [26]. This experiment was also done with the strains B/r and IAI51.

Grazing resistance

All strains in the collection were used in this experiment. Bacteria and amoebae were plated according to the same protocol that was used in the relative population size experiment. This experiment was performed with two different bacterial population size: low (10^4 cells) and high (10^8 cells). Each bacterial population size was associated with three amoeba population size: 10, 10^2, 10^3 cells. Three replicates were run for the entire collection strains and each population size combination. Five replicates were run for the IAI51 mutants.

Plates were checked for the occurrence of bacterial lysis plaques formed by *D. discoideum,* three and six days after plating for the low and high bacterial population size, respectively. This experiment was repeated two times with the strains IAI51 and mutants, IAI52, B/r, RS218, 536 and REL606.

Statistical analysis

For the relative population size experiment, a χ^2 test was applied to test for the effects of various factors on amoeba growth: strain effect (536 *vs* REL606), relative population size effects, and the interaction between strain and relative population size.

For the grazing experiment, in order to investigate association among possible grazing resistance and various phenotypic and genotypic bacterial traits, a factorial

analysis of correspondence (FAC) was carried out using R software and a binary table, with 31 lines corresponding to the collection of strains and 28 columns corresponding to the origin of the strain (pathogen or commensal), the four *E. coli* phylogenetic groups, 13 virulence factors, serum resistance, lysozyme/lactoferrin resistance, bile resistance, motility, growth rate, mouse killer (mouse lethality in a septicaemia model), mouse non-killer [72], grazing resistance, and grazing. The "grazing resistance" phenotype corresponds to the plates for which we did not observe a lysis plaque, for the three replicates and for all amoebae densities. The "grazing" phenotype corresponds to the plates where we observed lysis plaques for at least one replicate for each amoeba population size, or for all three replicates with one or two of the three amoeba densities. For each column, we coded each strain with a binary code: present=1, absent=0. The influence of the various factors on grazing resistance was tested with a κ test of correspondence.

For the IAI51 mutants, the presence of the *irp1* gene was tested with a χ^2 test.

Results

Relative population size effect on grazing resistance

To test whether the relative population size of *E. coli* and amoeba cells affect the grazing of bacteria by amoeba, a plating assay was carried out. *D. discoideum* cells were plated on HL5 agar plates with either the non-pathogenic *E. coli* strain REL606 or the pathogenic *E. coli* strain 536. We tested different bacteria/amoebae cells ratios by measuring grazing capacity at different bacteria (*vs*. amoebae) population sizes for a given amoeba (*vs.* bacteria) population size (summing up to 60 plates). Over the course of a few days, the bacteria formed lawns with amoeba embedded in them on the plates. The bacteria phagocytized by *D. discoideum* were assessed through the occurrence of bacterial lysis plaques.

Three days after plating, with 10^2 amoeba cells (Figure IV-1A), plaques were observed for three REL 606 bacterial population sizes: 10^4,10^5,10^6 cells and one population size of strain 536: 10^4 cells (Figure IV-1B). In the symetrical experiment (10^8 bacteria cells), we observed plaques only with the strain REL606 for two amoeba population sizes, 10^6 and 10^5 (Figure IV-1A), and not for the strain 536 (Figure IV-1B).

Six days after plating, plaques were noticed with the strain REL606 for both types of experiments and for all population sizes (bacteria *vs.* amoebae) (Figure IV-1C). But for the pathogenic strain 536, plaques appeared for three relative population sizes: 10^2, 10^3 and 10^4 bacteria/amoeba (Figure IV-1D). In the symmetrical experiment, using 10^4 amoebae, plaques were observed for a single replicate (Figure IV-1D).

These results show that with the pathogenic strain 536, up to a relative population size of 10^4, bacterial lysis plaques arise, but not with greater relative population sizes.

D. discoideum feeds on *E. coli* REL606 whatever the relative population size (Figure IV-1C) and form plaques that can be noticed after six days. *D. discoideum* does not form plaques on the lawns of *E. coli* 536 for a relative population size greater than 10^4 bacteria/amoeba (strain effect: χ^2= 4.945, p=0.0262; population size effect: χ^2=4.04, p=0.0444; strain*population size effect: χ^2=4.04, p=0.0444). Therefore, in contrast with *E. coli* REL606, *E. coli* 536 resists digestion by amoebae. *E. coli* 536 may kill amoeba, or at least substantially inhibit their growth.

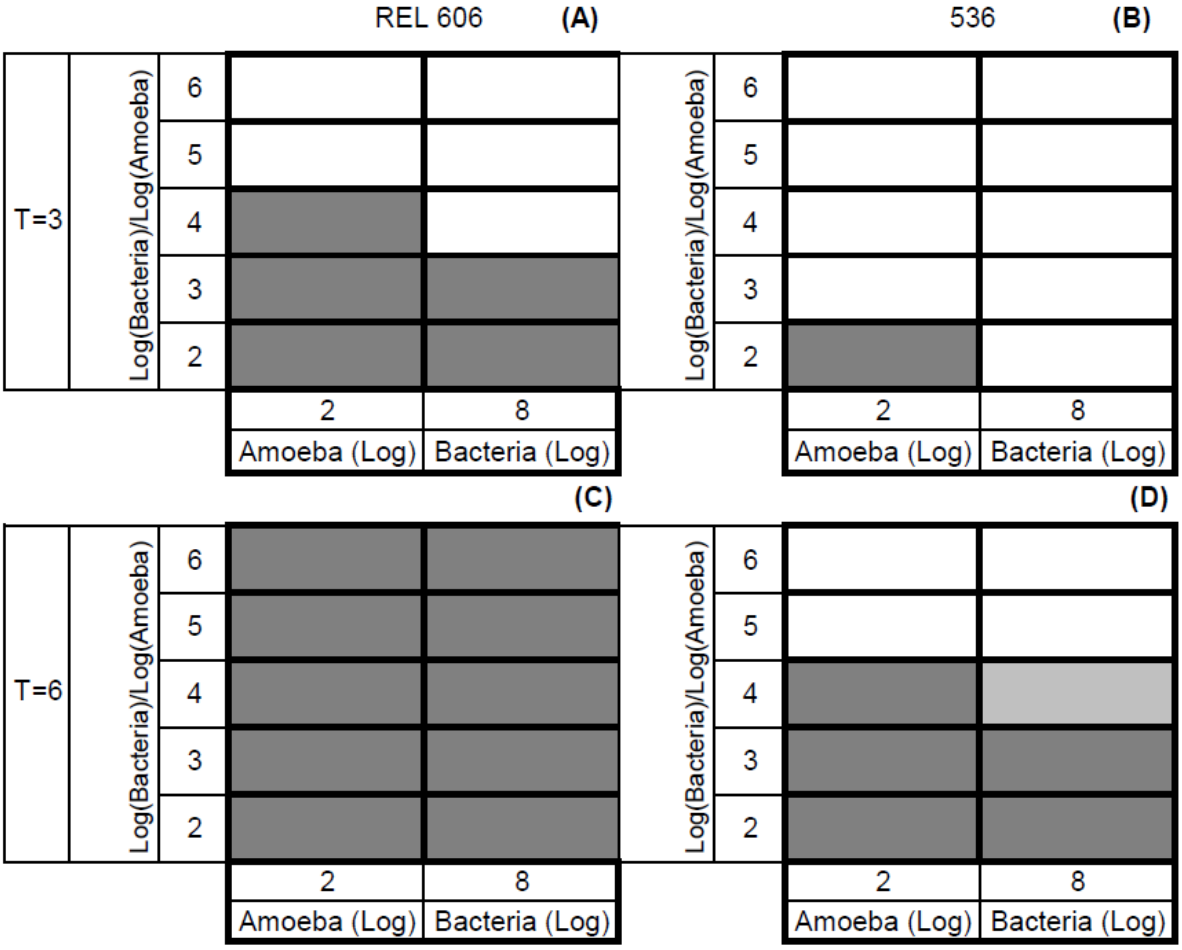


Figure IV-1: Relative population size effect on *E. coli*/*D. discoideum* interactions

Bacterial lysis plaques (strains REL 606 and 536) observed three days (T=3) or six days (T=6) after plating.

Dark gray color represents three replicates with bacterial lysis plaques.

Gray color represents one replicate with bacterial lysis plaques.

White color represents zero replicate with bacterial lysis plaques.

(A) Three days after plating, strain REL 606

(B) Tree days after plating, strain 536

(C) Six days after plating, strain REL 606

(D) Six days after plating, strain 536

Death versus growth inhibition

To test the ability of the pathogenic strain *E. coli* 536 to kill *D. discoideum* or inhibit its growth, bacteria (the pathogenic strain 536 and the non-pathogenic REL606) were plated with amoebae at 10^6 relative population size. Two and six days after plating, the top layers were transferred in a MCPB liquid medium. This

liquid culture was then screened with BacLight kit, in which healthy living cells fluoresce in green, whereas damaged cells fluoresce in red [26].

Two days after plating, both bacteria (536 and REL606 strains) and amoebae fluoresce in green (Figure IV-2 B, first and second rows), indicating living cells. Six days after plating, for the experiment with strain REL606, we observed some bacteria with green and red fluorescent color outside and inside amoeba (white arrow), red fluorescence indicating membrane cell damages. Amoeba in green fluorescent stain (Figure IV-2B, third row) indicated live cells. Thus, *D. discoideum* remained alive and had digested some *E. coli* 606 bacteria cells. Similar results were observed with the B/r strain (data not shown). The culture with *E. coli* REL606 thus revealed a high density of amoebae which implies that amoeba can achieve high growth rates in this diet. In contrast, six days after plating, the picture with the pathogenic strain *E. coli* 536 revealed three or more living *E. coli* cells in the *D. discoideum* one (Figure IV-2B, fourth row) and other bacteria outside the amoeba in red fluorescence. Staining with BacLight kit revealed green fluorescent, living bacteria (white arrow) inside amoeba. Red fluorescent staining indicated damaged membranes of the *D. discoideum* that should lead to cell death. Similar results were obtained using pathogenic strain IAI51 (data not shown). This fluorescent straining showed living *E. coli* 536 bacteria cells and *D. discoideum* dead cells.

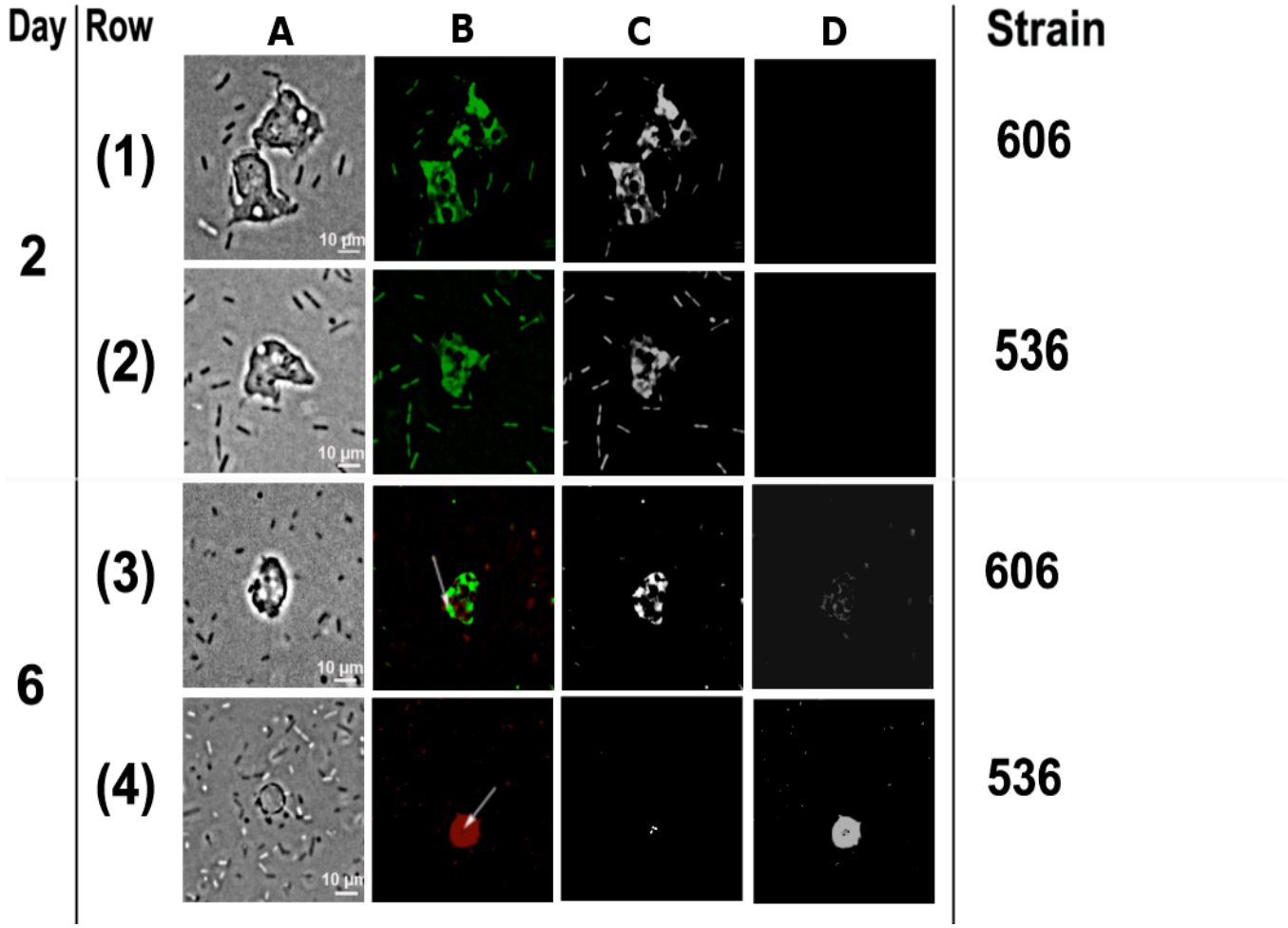


Figure IV-2: Visualization of the *E. coli*/*D. discoideum* viability

(A) Phase contrast microscopy (20x),

(B) Combined Red and Green: Green fluorescent stain indicate living cells, red fluorescence indicate cells with damaged membranes.

(C) Green channel alone

(D) Red channel alone

(1): Amoeba with the strain *E. coli* B REL 606 two days after plating, all cells are green reflecting living cells.

(2): Amoeba with the strain *E. coli* 536 two days after plating, all cells are green reflecting living cells.

(3): Amoeba with the strain *E. coli* B REL 606 six days after plating. Red fluorescent bacteria are cells with damaged membranes (white arrow) and green fluorescent cells are alive (amoeba and bacteria).

(4): Amoeba with the strain *E. coli* 536 six days after plating, green fluorescent bacteria are living cells (white arrow), red fluorescent amoeba and bacteria are cells with damaged membranes.

Grazing resistance

In order to compare the ability of pathogenic *E. coli* strains to resist grazing protozoa with that of commensal strains, and to identify bacterial virulence traits that could have been favoured and maintained by selection for other functions than virulence *per se*, a plating assay was carried out for each strain in the bacterial collection (including pathogenic and commensal ones) at a low (10^4) and a high (10^8) initial bacterial population size with three amoeba population sizes (10, 10^2, 10^3). Three replicates were run for each combination (N=558 plates). A few days after plating, the bacteria formed lawns on these plates in which the amoebae were embedded. At low bacterial population size (10^4), in agreement with the previous experiments, *D. discoideum* formed plaques on lawns with all *E. coli* strains. At high bacterial population size (10^8), *D. discoideum* fed on some *E. coli* strains, but did not form any bacterial lysis plaques for some other strains (Figure IV-3). The results were highly reproducible with identical results in all replicates for the IAI51, IAI52, B REL606, RS218 and B/r strains.

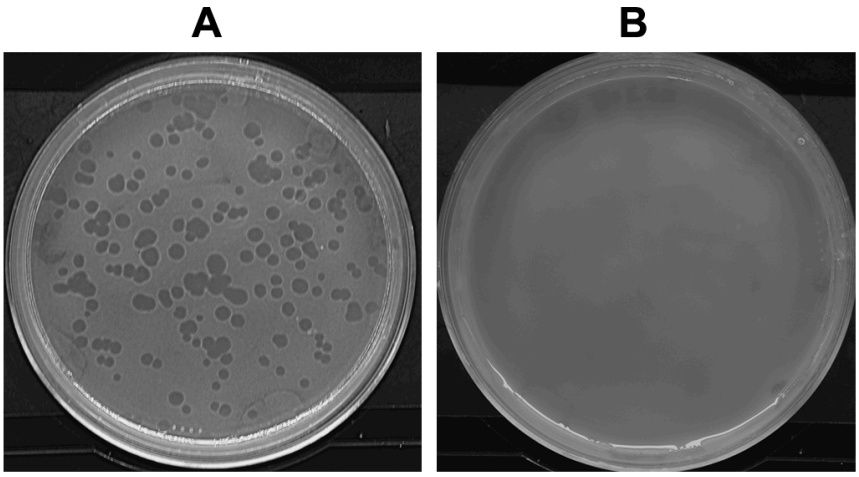


Figure IV-3: Bacterial lysis plaque occurrence

(A) A non-virulent *E. coli* strain (B REL606; 10^8 cells) was plated with *D. discoideum* (10^2 cells). Bacterial lysis plaques were observed, which is characteristic of the grazing phenotype.

(B) A virulent *E. coli* strain (536; 10^8 cells) was plated with *D. discoideum* (10^2 cells). No bacterial lysis plaques were observed, which is characteristic of the grazing resistance phenotype.

Association between grazing resistance and bacterial virulence traits

To identify associations between grazing resistance to protozoa and various bacterial traits, a FAC (Factorial Analysis of Correspondence) was applied using all

data available for the strain collection (Table IV-1). The FAC (Figure IV-4) is based on χ^2 distances, using a covariance matrix. It detects existing positive and negative associations, and measures the contribution of each factor to the total inertia. The projection of variables onto the F1/F2 plane accounted for 47.47% of the total variance and distinguished two main groups of variables by positive and negative values for the first F1 factor. The first group (F1 negative values and F2 positive and negative values) encompasses the following traits: grazing, commensal origin, the absence of mouse lethality and the phylogenetic groups A, B1 and D. The second main group (F1 positive values and F2 positive and negative values) encompasses grazing resistance trait, mouse lethality, pathogenic origin of the strains, phylogenetic group B2, virulence factors, the resistances to serum, bile and lysozyme/lactoferrin, the motility and fast growth rate.

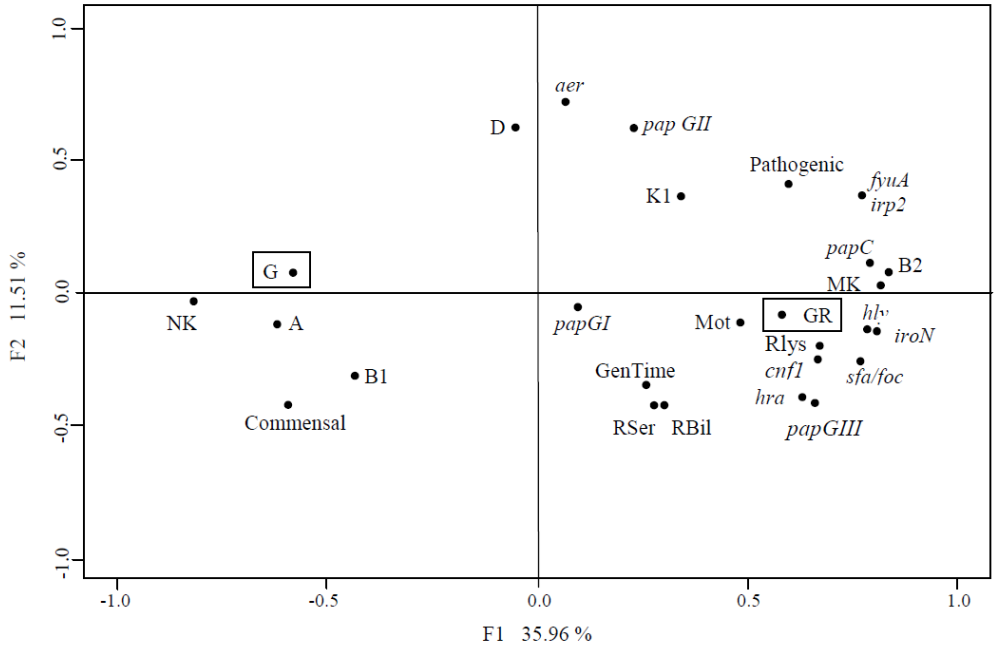


Figure IV-4: Factorial analysis of correspondences

Projection onto the F1/F2 plane of the 28 bacterial traits characterized for the 31 *E. coli* collection strains. A, B1, B2 and D: phylogenetic groups; Com: commensal, Path: pathogenic; *papGI, II, III: papG* and corresponding three existing alleles; K1: K1 antigen (capsule), *sfa/foc, iroN, hly, cnf1, hra, fyuA, irp2, aer*: virulence genes; RSer: Serum resistance; RBil: Bile resistance; RLys: lysozym/lactoferrin resistance; Mot: motility; GenTim: generation time; K: mouse killer, NK: mouse non-killer; G: grazing, GR: grazing resistance (boxed). The projection of the variables onto the F1/F2 plane accounts for 47.47% of the total variance. The first group (F1 negative values and F2 positive and negative values) encompasses following traits: the grazing, the commensal origin, the absence of mouse lethality and the phylogenetic groups A, B1 and D. The second main group (F1 positive values and F2 positive and negative values) encompasses the grazing resistance trait, the mouse lethality, the pathogenic origin of the strains, the phylogenetic group B2, virulence genes, the resistance to serum, bile and lysozyme/lactoferrin, the motility and fast growth rate.

The FAC thus shows that virulence is strongly associated with grazing resistance.

To identify possible correlations between grazing resistance to protozoa and the presence of virulence genes and other traits, we performed a κ test of correspondences. This analysis indicated that grazing resistance was correlated to the mouse killer phenotype (κ=0.459; *p*=0.00998), to the phylogenetic group B2 (κ=0.497; *p*=0.00558), to the resistance to serum (κ=0.479, *p*=0.00574), to the resistance to lysozyme (κ=0.349; *p*=0.0443), to the resistance to the bile (κ=0.349; *p*=0.0443), and to the presence of the genes *iroN* (κ=0.403; *p*=0.0222), *fyuA* (κ=0.412; *p*=0.0203), *irp2* (κ=0.412; *p*=0.0203). Both *fyuA* and *irp2* genes belong to the high pathogenicity island (HPI). This 35- to 45-kb stretch of DNA carries a siderophore-mediated iron uptake system called 'the yersiniobactin locus' [47]. In summary, κ analyses revealed that pathogenicity of various *E. coli* strains on mouse, the B2 phylogenetic group, and the presence of virulence genes are significantly correlated with grazing resistance to protozoa.

Experimental demonstration of the role of the HPI in the grazing resistance

The IAI51 strain showed a high-virulence phenotype on mouse [48] and *D. dictyostelium.* This strain was thus selected to test experimentally the impact of HPI on *D. dictyostelium* that was already suggested from the correlation analysis. Two mutant IAI51 strains were tested for the amoeba grazing: IAI51 *irp1-* (bearing a deletion covering the *irp1* gene) and the IAI51 pCP1 (IAI51*irp1-* re-complemented with the *irp1* gene on a plasmid). These mutant strains were already tested in a mouse model of the extra-intestinal virulence and have indicated that HPI plays a role in mouse lethality [48].

Five replicates of the three bacterial strains (IAI51, IAI51*irp1-*, IAI51pCP1) were plated at a high bacterial population size (10^8 cell) and three amoeba population sizes (10, 10^2, 10^3 cells) (N=45 plates). The presence of the bacterial lysis plaques was assessed six days after plating (Table IV-2). This experiment was repeated twice. The presence of the *irp1* gene was highly significantly associated with the grazing resistance to amoeba since deletion of the *irp1* gene made the

strain sensitive to the amoeba (grazing) whereas complementation of this mutant by the cPC1 plasmid restored grazing resistance (three strains model: IAI51/IAI51 *irp1*-/IAI51 pCP1, ddl=2, χ^2 =30.11, p<0.0001; two strains model: IAI51 *irp1*-, IAI51 pCP1, ddl=1, χ^2= 20.49, p<0.001; two strains model: IAI51, IAI51 *irp1*-, ddl=1, χ^2 = 42.71, p<0.001).

Table IV-2: Demonstration of the role of the HPI in the grazing resistance

Strain				
	IAI51 pCP1	+	+	++
	IAI51 irp-	+++	++++	++++
	IAI51	-	+	-
	606	+++++	+++++	+++++
		1	2	3
		log(number amoeba)		

Bacterial lysis plaques (strains REL606, IAI51, IAI51 *irp*-, IAI51 *pCP1*) observed six days after plating. A density of 10^8 bacteria was plated with 10, 10^2, 10^3 amoebae. (+) represents the number of replicates with lysis plaques, (-) means no replicates showing plaques.

Discussion

This study investigated the coincidental evolution hypothesis of virulence factors. In particular, we tested the hypothesis that some bacterial traits favor the survival of *E. coli* when confronted with grazing protozoa. A close relationship was observed between the mouse killer phenotype, which represents intrinsic extra-intestinal virulence of the bacterial strain [32], and the grazing resistance trait. *D. discoideum* was able to survive and phagocytize *E. coli* strains that (i) did not harbor virulence genes involved in iron capture (*iroN, fyuA, irp*), (ii) are not resistant to serum, bile, lysozyme/lactoferine, or that do not belong to the B2 phylogenetic group. *D. discoideum* cells are specialized phagocytes that essentially feed on a range of bacteria, their natural prey. After adhesion, engulfment of the bacteria in a phagosome is the first step prior to successive maturation steps that involve sequential delivery and retrieval of components to phagosomes, where lysosomal hydrolases are delivered in order to digest bacteria. The association between grazing resistance and various resistance phenotypes: resistance to human serum, lysozyme/lactoferine, bile, reflects the importance of the capacity of ExPEC cells to protect their membrane against lysosomal hydrolases from amoeba

phagosomes. Other bacterial strains, such as *Mycobacterium avium* bacilli [49], *E. coli O157* vero-cytotoxigenic strain [26], are known to resist digestion by amoeba (for a review, see [50]). However, mechanisms that determine resistance to grazing amoeba are complex and not fully understood. For instance, we found no correlation between the presence of the K1 capsule and the resistance of bacteria to amoeba grazing. Previous studies showed that this capsule enables *E. coli* to survive in human endothelial cells by preventing lysosomal fusion [51], as well as in *Acanthamoeba* [52]. Nevertheless, even if this capsule enhances bacterial survival, it does not seem to interfere with amoeba survival.

This study shows that various ExPEC strains can resist to digestion by amoeba. The percentage of grazing resistant *E. coli* strains that carry virulence factors is greater (76%) than the percentage of grazing resistant *E. coli* strains that do not (16%). This is consistent with the hypothesis that predation by protozoa can favor bacterial traits that increase survival, but emerge as virulence determinants in human infections. Genes responsible for such effects can thus be thought of as being maintained by coincidental selection. An earlier study [9] had already established that some intracellular pathogens normally persist in the environment in close association with protozoa and especially with amoeba. King *et al.* (1988) [10] suggested that grazing resistance to protozoa was an evolutionary precursor of bacterial pathogenicity. This was supported by the finding of predation-mediated variation at the *rfb* virulence locus of *Salmonella enterica* which illustrates the potential impact of protozoan grazing on the origin of bacterial pathogenicity [53]. Recently, Steinberg and Levin [54] showed that the Shiga toxin genes carried by some intestinal pathogenic *E. coli* strains increase bacterial survival in the presence of *Tetrahymena pyriformis,* indicating that such prey-predator interactions might have driven the evolution of this virulence factor. Similarly, our experiments showed that *E. coli* genes involved in the iron uptake such as *irp* and *fyuA* –which are located in the HPI- and *iroN* were involved in the grazing resistance to *D. discoideum*. Iron is a vital resource for bacteria, necessary for many functions, from respiration to DNA replication. While it is one of the most abundant elements on earth, its rarity in the biosphere makes it often a limiting resource for microbial growth. Furthermore, in the presence of oxygen, iron is oxidised to the ferric state, forming ferric hydroxide, which is insoluble in aqueous solution. Many bacteria can exploit *D. dictyostelium* phagocytosis and grow

intracellularly [26,55]. As a counter-strategy, many hosts have evolved iron withholding strategies to disable micro-organisms [56]. In *D. dictyostelium*, the transport of iron across the phagolysosomal membrane contributes to pH homeostasis, while at the same time depleting the lysosomal milieu from the iron required for bacterial growth. In response to this competition for iron, counter-mechanisms have been developed by microbes to overcome the scarcity of this metal. These mechanisms involve two principal methods: (i) either the synthesis of high iron affinity compounds called siderophores (iron-scavenging agents), or (ii) by a method of direct capture at the bacterial cell membrane [57]. Siderophores are ferric-specific microbial iron chelator compounds whose biosynthesis is regulated by the availability of iron in surrounding environment [58]. Bacteria respond to the low-iron environment by increasing expression of their iron acquisition system as well as of other various virulence factors [59]. Our IAI51 *irp1*- mutant showed less virulence on *D. dictyostelium* compared to the mutant IAI51 pCP1. This result was consistent with the closed relationship between iron and virulence. For the host, the competition for iron necessitates careful orchestration of the iron regulation in order to prevent bacteria invasion and to assure optimal iron homeostasis. Thus, some of the virulence factors involved in human disease may have an ecological function within natural communities or even have their origin specifically in successful anti-predator adaptation. Other papers [15,16] showed a possible link between bacterial evolution in the natural environment and their pathogenicity toward animals. They suggested that the pathogenicity genes were selected for alternative ecological functions.

Our grazing resistance experiment showed a strong correlation between lethality in mice and grazing resistance. This is the first demonstration that *D. discoideum* can be used as host model for specifically detecting ExPEC strain virulence. More generally, previous studies [60,61] reported a high correlation between virulence, as estimated from its resistance to *D. discoideum* grazing, and death in mammalian infection model. Our results suggest that for ExPEC *E. coli* strains, similar mechanisms determine virulence in mammalian systems and amoeba. This relation is corroborated by correlation between resistance to lysozyme or serum and grazing resistance in our experiments. From a bacterial point of view, the similarities between *D. discoideum* and mammalian cells include

membrane trafficking, endocytosis, phagocytosis and exocytosis pathways [45,55]. Moreover, *D. discoideum* shares its chemotactic capacity with leucocytes. Thus *D. discoideum* grazing looks like killing by macrophages [62,63]. The amoeba host model (*D. discoideum*) has already been used with a variety of bacterial species: including *L. pneumophila* [55], *Pseudomonas aeruginosa* [61], *Vibrio cholerae* [64], and *Aeromonas* spp (*A.salmonicida* and *A.hydrophila*) [65]. The relevance of this model is not only based on the observation that many pathogens show a low species specificity, but also that a large range of hosts will encounter the same bacteria in their natural environment. The *D. discoideum–E. coli* model is a very simple and powerful system. Using a simple plating assay, *D. discoideum* forms a phagocytosis plaque on a lawn of non-pathogenic bacteria but does not on a lawn of pathogenic ones. The virulence of the bacteria used may thus be extrapolated from the ability *of D. discoideum* to survive and grow in their presence [61,66]

The fact that free living amoebae graze on bacteria has important implications from ecological and evolutionary viewpoints [67]. Ecological interactions are often grouped according to the net effect that each species acts on its partner (positive, negative or null). Such definitions of species interactions are then classified either as competitive, antagonistic, or mutualistic. But the interaction between bacteria and amoebae reveals many facets. The interactions between *E. coli* and *D. discoideum* are complex, and the balance between conflict and interest may be delicate. *E. coli* can find food inside *D. discoideum* but at the expense of a substantial risk to be digested. Our death *vs.* growth inhibition experiment showed that once inside *D. Dictyostelium*, a struggle for life begins. Bacteria may exploit amoeba phagocytosis to find key nutrients, but depending on the strain characteristics, the bacteria either resists digestion by amoeba and kills *D. discoideum* (Figure IV-2B fourth row) or is digested by *D. discoideum* (Figure IV-2B third row). As soon as bacteria and amoeba show sufficient common interest, they may both sustain themselves (Figure IV-2B first and second rows), but if selfish interests prevail over common interest, the harmonious relationship stops and the struggle for life ensues until the death of one of the partners or the other. If bacteria are able to resist predation, invading amoebae present a number of potential advantages: the amoebae protect them against adverse environmental conditions and may even act as vector for bacteria toward new resources. So even

if their defence against predation is not perfect, bacteria may be selected to allow themselves be phagocytized notwithstanding the risk. Such partly antagonistic, partly mutualistic interactions have been called 'dangerous liaisons': partners team up but nevertheless relentlessly pursue their private interests [68]. Amoeba could not only be protective niches under stressful conditions or vehicles but also trojan horses –*e.g;* amoebal vesicales containing bacilli could be inhaled by humans, who the bacteria may then infect [14,69].

Other ecological factors, such as abundance, seem to be related to grazing resistance. Our study demonstrated that the relative population size of bacteria/amoeba could play a key role in grazing resistance (Figure IV-1). When the relative population size of pathogenic bacteria compared to amoeba is sufficiently high, no plaque lysis was observed (10^7 bacteria for 100 amoeba cells or 10^8 bacteria for 1,000 amoeba cells). However, if relative population size of bacteria to amoeba is lower than a certain threshold, we observed plaques and no grazing resistance. Many species of bacteria including *E. coli* K-12 MG1655 [70] and O157:H7 [71] were seen to regulate expression of certain genes in response to change in population size. In the primary (the gut) and secondary (in natura) *E. coli* habitats, bacteria and amoeba populations coexist. In natura, they are dispersed in soil patches. Clarholm [42] observed that amoebae were as abundant as 10^5 individuals per gram of soil. Their numbers may increase up to 20-fold in four days, thus amounting to 2.10^6 individuals per gram of soil. In nature, amoebae have been documented to cause a decline of prey bacteria from 10^8 to 10^5 cells per gram of soil. Thus, in soil, the relative amoeba/bacteria population size is likely to be high enough to permit the amoeba to have a great impact on bacterial population. Nevertheless, the establishment of a community depends on the history of the site, the order in which the species arrived, what kind of interactions develop among species and the quality of the patch. All these factors determine whether such a community would achieve a permanent state of equilibrium [42]. The primary *E. coli* habitat, gut, forms a complex ecosystem with various commensal amoeba species that coexist with more than 500 bacterial species, a total of 10^{10} to 10^{11} cells per gram of the large intestinal content. Commensal bacterial strains may coexist with amoeba in the gut and our experiments showed that whatever the relative amoeba/commensal bacteria population sizes amoeba may survive and grow.

Concluding remarks

In this paper, we argued that ecological factors are associated with the evolution of bacterial traits responsible for virulence. Our observations support the coincidental hypothesis for the evolution of virulence: the capacity to resist grazing by protozoa manifests itself in a human host as increasing virulence. Amoeba and bacteria coexist in a closed relationship in their different habitats (nature *vs.* gut).

The use of *D. discoideum* to detect ExPEC strain virulence has several advantages. The simplicity of the experimental protocol and its reproducibility surpasses those of mammalian systems. Furthermore, numerous genetic tools permit the use of various well-known *D. discoideum* mutants for studying the mechanisms underlying pathogenicity (endocytosis, phagocytosis), and sensitivity to pathogens, or to identify genes involved in virulence. Despite some limits (temperature culture of 24°C, simple cell), *D. discoideum* may prove to be a fruitful model system for preliminary studies prior conducting experiments with mammalian host models.

References

1. May R, Anderson R (1983) Epidemiology and genetics in the coevolution of parasites and hosts. Proceedings of the Royal Society of London Series B, Biological Sciences 219: 281-313.

2. Levin S, Pimentel D (1981) Selection of intermediate rates of increase in parasite-host systems. American Naturalist 117: 308.

3. Anderson R, May R (1982) Population biology of infectious diseases: Springer-Verlag, New York.

4. Ewald P (1983) Host-parasite relations, vectors, and the evolution of disease severity. Annual Review of Ecology and Systematics 14: 465-485.

5. Read A, Harvey P (1994) The evolution of virulence. Trends in Microbiology 2: 73-76.

6. Ebert D, Herre E (1996) The evolution of parasitic diseases. Parasitology Today 12: 96-101.

7. Frank S (1996) Models of parasite virulence. The Quarterly review of biology 71: 37.

8. Alizon S, van Baalen M (2005) Emergence of a convex trade off between transmission and virulence. Am Nat 165: E155-E167.

9. Brown M, Barker J (1999) Unexplored reservoirs of pathogenic bacteria: protozoa and biofilms. Trends in Microbiology 7: 46-50.

10. King C, Shotts E, Wooley R, Porter K (1988) Survival of coliforms and bacterial pathogens within protozoa during chlorination. Applied and environmental microbiology 54: 3023-3033.

11. Casadevall A (2008) Evolution of intracellular pathogens.

12. Matz C, Kjelleberg S (2005) Off the hook–how bacteria survive protozoan grazing. Trends in Microbiology 13: 302-307.

13. Weekers P, Bodelier P, Wijen J, Vogels G (1993) Effects of grazing by the free-living soil amoebae Acanthamoeba castellanii, Acanthamoeba polyphaga, and Hartmannella vermiformis on various bacteria. Applied and Environmental Microbiology: Am Soc Microbiol. pp. 2317-2319.

14. Rowbotham T (1980) Preliminary report on the pathogenicity of Legionella pneumophila for freshwater and soil amoebae. British Medical Journal 33: 1179-1183.

15. Levin B (1996) The evolution and maintenance of virulence in microparasites. Emerg Infect Dis 2: 93-102.

16. Levin B, Svanborg Eden C (1990) Selection and evolution of virulence in bacteria: an ecumenical excursion and modest suggestion. Parasitology(London Print) 100: 103-115.

17. Gould S, Lewontin R (1979) The spandrels of San Marco and the Panglossian paradigm: a critique of the adaptationist programme. Proceedings of the Royal Society of London Series B, Biological Sciences: 581-598.

18. Steinert M, Hentschel U, Hacker J (2000) Symbiosis and pathogenesis: evolution of the microbe-host interaction. Naturwissenschaften 87: 1-11.

19. Berg R (1996) The indigenous gastrointestinal microflora. Trends in Microbiology 4: 430-435.

20. Smith H (1965) Observations on the flora of the alimentary tract of animals and factors affecting its composition. The Journal of Pathology and Bacteriology 89.

21. Van Donsel D, Geldreich E, Clarke N (1967) Seasonal variations in survival of indicator bacteria in soil and their contribution to storm-water pollution. Applied and Environmental Microbiology 15: 1362-1370.

22. VanDonsel D, Geldreich E (1971) Relationships of salmonellae to fecal coliforms in bottom sediments. WATER RESEARCH, VOL 5, NO 11, P 1079-1087, NOVEMBER 1971 3 FIG, 3 TAB, 21 REF.

23. Faust M (1976) Coliform bacteria from diffuse sources as a factor in estuarine pollution. Water Res 10: 619-624.

24. Power M, Littlefield-Wyer J, Gordon D, Veal D, Slade M (2005) Phenotypic and genotypic characterization of encapsulated Escherichia coli isolated from blooms in two Australian lakes. Environmental Microbiology 7: 631.

25. Solo-Gabriele H, Wolfert M, Desmarais T, Palmer C (2000) Sources of Escherichia coli in a coastal subtropical environment. Applied and Environmental Microbiology 66: 230-237.

26. Barker J, Humphrey T, Brown M (1999) Survival of Escherichia coli O157 in a soil protozoan: implications for disease. FEMS microbiology letters 173: 291.

27. Herzer P, Inouye S, Inouye M, Whittam T (1990) Phylogenetic distribution of branched RNA-linked multicopy single-stranded DNA among natural isolates of Escherichia coli. Journal of bacteriology 172: 6175-6181.

28. Desjardins P, Picard B, Kaltenböck B, Elion J, Denamurl E (1995) Sex in Escherichia coli does not disrupt the clonal structure of the population: evidence from random amplified polymorphic DNA and restriction-fragment-length polymorphism. Journal of molecular evolution 41: 440-448.

29. Escobar-Paramo P, Menac'h L, Le Gall T, Amorin C, Gouriou S, et al. (2006) Identification of forces shaping the commensal Escherichia coli genetic structure by comparing animal and human isolates. Environmental Microbiology 8: 1975.

30. Johnson J, Russo T (2002) Extraintestinal pathogenic Escherichia coli:" The other bad E coli". Journal of Laboratory & Clinical Medicine 139: 155.

31. Johnson J (1991) Virulence factors in Escherichia coli urinary tract infection. Clinical Microbiology Reviews 4: 80-128.

32. Picard B, Garcia J, Gouriou S, Duriez P, Brahimi N, et al. (1999) The link between phylogeny and virulence in Escherichia coli extraintestinal infection. Infection and immunity 67: 546.

33. Escobar-Paramo P, Clermont O, Blanc-Potard A, Bui H, Le Bouguenec C, et al. (2004) A specific genetic background is required for acquisition and expression of virulence factors in Escherichia coli. Molecular biology and evolution: SMBE. pp. 1085-1094.

34. Tenaillon O, Skurnik D, Picard B, Denamur E (2010) The population genetics of commensal Escherichia coli. Nature Reviews Microbiology 8: 207-217.

35. Le Gall T, Clermont O, Gouriou S, Picard B, Nassif X, et al. (2007) Extraintestinal virulence is a coincidental by-product of commensalism in B2 phylogenetic group Escherichia coli strains. Molecular Biology and Evolution 24: 2373.

36. Fotedar R, Stark D, Beebe N, Marriott D, Ellis J, et al. (2007) PCR detection of Entamoeba histolytica, Entamoeba dispar, and Entamoeba moshkovskii in stool samples from Sydney, Australia. Journal of Clinical Microbiology 45: 1035.

37. Hamzah Z, Petmitr S, Mungthin M, Leelayoova S, Chavalitshewinkoon-Petmitr P (2006) Differential detection of Entamoeba histolytica, Entamoeba dispar, and Entamoeba moshkovskii by a single-round PCR assay. Journal of Clinical Microbiology 44: 3196.

38. Troll H, Marti H, Weiss N (1997) Simple differential detection of Entamoeba histolytica and Entamoeba dispar in fresh stool specimens by sodium acetate-acetic acid-formalin concentration and PCR. Journal of Clinical Microbiology 35: 1701.

39. Rivera F, Rosas I, Castillo M, Chávez M, Gómez R, et al. (1986) Pathogenic and free-living protozoa cultured from the nasopharyngeal and oral regions of dental patients: II. Environmental research 39: 364.

40. Habte M, Alexander M (1977) Further evidence for the regulation of bacterial populations in soil by protozoa. Archives of Microbiology 113: 181-183.

41. Bryant R, Woods L, Coleman D, Fairbanks B, McClellan J, et al. (1982) Interactions of bacterial and amoebal populations in soil microcosms with fluctuating moisture content. Applied and Environmental Microbiology 43: 747-752.

42. Clarholm M (1981) Protozoan grazing of bacteria in soil—impact and importance. Microbial Ecology 7: 343-350.

43. Burns D, Solomon J, Rupper A, Cardelli J, Isberg R (2000) Intracellular growth of Legionella pneumophila in Dictyostelium discoideum, a system for genetic analysis of host-pathogen interactions. Infection and immunity 68: 2939-2947.

44. Steinert M, Heuner K (2005) Dictyostelium as host model for pathogenesis. Cellular microbiology 7: 307-314.

45. Cardelli J (2001) Phagocytosis and macropinocytosis in Dictyostelium: phosphoinositide-based processes, biochemically distinct. Traffic 2: 311.

46. Eichinger L, Pachebat J, Glöckner G, Rajandream M, Sucgang R, et al. (2005) The genome of the social amoeba Dictyostelium discoideum. Nature 435: 43-57.

47. Carniel E (2001) The Yersinia high-pathogenicity island: an iron-uptake island. Microbes and Infection 3: 561-569.

48. Schubert S, Picard B, Gouriou S, Heesemann J, Denamur E (2002) Yersinia high-pathogenicity island contributes to virulence in Escherichia coli causing extraintestinal infections. Infection and immunity 70: 5335.

49. Steinert M, Birkness K, White E, Fields B, Quinn F (1998) Mycobacterium avium bacilli grow saprozoically in coculture with Acanthamoeba polyphaga and survive within cyst walls. Applied and Environmental Microbiology 64: 2256-2261.

50. Greub G, Raoult D (2004) Microorganisms resistant to free-living amoebae. Clinical Microbiology Reviews 17: 413-433.

51. Kim K, Elliott S, Di Cello F, Stins M, Kim K (2003) The K 1 capsule modulates trafficking of E. coli-containing vacuoles and enhances intracellular bacterial survival in human brain microvascular endothelial cells. Cellular Microbiology 5: 245-252.

52. Jung S, Matin A, Kim K, Khan N (2007) The capsule plays an important role in Escherichia coli K1 interactions with Acanthamoeba. International journal for parasitology 37: 417-423.

53. Wildschutte H, Wolfe D, Tamewitz A, Lawrence J (2004) Protozoan predation, diversifying selection, and the evolution of antigenic diversity in Salmonella. Proceedings of the National Academy of Sciences 101: 10644-10649.

54. Steinberg M (2007) Grazing protozoa and the evolution of the Escherichia coli O157: H7 Shiga toxin-encoding prophage. Proceedings of the Royal Society B 274: 1921.

55. Solomon J, Isberg R (2000) Growth of Legionella pneumophila in Dictyostelium discoideum: a novel system for genetic analysis of host–pathogen interactions. Trends in Microbiology 8: 478-480.

56. Latunde-Dada G (2009) Iron metabolism: microbes, mouse, and man. BioEssays 31: 1309-1317.

57. Doherty C (2007) Host-pathogen interactions: the role of iron. Journal of Nutrition 137: 1341.

58. Orsi N (2004) The antimicrobial activity of lactoferrin: current status and perspectives. Biometals 17: 189-196.

59. Litwin C, Calderwood S (1993) Role of iron in regulation of virulence genes. Clinical microbiology reviews 6: 137.

60. Jander G, Rahme L, Ausubel F (2000) Positive correlation between virulence of Pseudomonas aeruginosa mutants in mice and insects. Journal of Bacteriology 182: 3843.

61. Cosson P, Zulianello L, Join-Lambert O, Faurisson F, Gebbie L, et al. (2002) Pseudomonas aeruginosa virulence analyzed in a Dictyostelium discoideum host system. Journal of bacteriology 184: 3027-3033.

62. Noegel A, Schleicher M (2000) The actin cytoskeleton of Dictyostelium: a story told by mutants. Journal of Cell Science. pp. 759-766.

63. Rupper A, Cardelli J (2001) Regulation of phagocytosis and endo-phagosomal trafficking pathways in Dictyostelium discoideum. BBA-General Subjects 1525: 205-216.

64. Pukatzki S, Ma A, Sturtevant D, Krastins B, Sarracino D, et al. (2006) Identification of a conserved bacterial protein secretion system in Vibrio cholerae using the Dictyostelium host model system. Proceedings of the National Academy of Sciences 103: 1528-1533.

65. Froquet R, Cherix N, Burr S, Frey J, Vilches S, et al. (2007) Alternative host model to evaluate Aeromonas virulence. Applied and Environmental Microbiology 73: 5657.

66. Pukatzki S, Kessin R, Mekalanos J (2002) The human pathogen Pseudomonas aeruginosa utilizes conserved virulence pathways to infect the social amoeba Dictyostelium discoideum. Proceedings of the National Academy of Sciences 99: 3159.

67. Marciano-Cabral F, Cabral G (2003) Acanthamoeba spp. as agents of disease in humans. Clinical Microbiology Reviews 16: 273-307.

68. van Baalen M, Jansen V (2001) Dangerous liaisons: the ecology of private interest and common good. Oikos 95: 211-224.

69. Barker J, Brown M (1994) Trojan horses of the microbial world: protozoa and the survival of bacterial pathogens in the environment. Microbiology 140: 1253-1259.

70. Sperandio V, Torres A, Kaper J (2002) Quorum sensing Escherichia coli regulators B and C(QseBC): a novel two-component regulatory system involved in the regulation of flagella and motility by quorum sensing in E. coli. Molecular microbiology 43: 809-821.

71. Sperandio V, Torres A, Giron J, Kaper J (2001) Quorum sensing is a global regulatory mechanism in enterohemorrhagic Escherichia coli O157: H7. Journal of Bacteriology 183: 5187.

72. Diard M, Baeriswyl S, Clermont O, Gouriou S, Picard B, et al. (2007) Caenorhabditis elegans as a simple model to study phenotypic and genetic virulence determinants of extraintestinal pathogenic Escherichia coli. Microbes and infection 9: 214-223.

73. Johnson J, Clermont O, Menard M, Kuskowski M, Picard B, et al. (2006) Experimental mouse lethality of Escherichia coli isolates, in relation to accessory traits, phylogenetic group, and ecological source. The Journal of infectious diseases 194: 1141-1150.

74. Clermont O, Bonacorsi S, Bingen E (2000) Rapid and simple determination of the Escherichia coli phylogenetic group. Applied and Environmental Microbiology 66: 4555.

Aknowledgments

We thank O. Tenaillon for discussions, B. Picard for statistical analysis discussions, S. Schubert for providing the mutant strains and the reviewer for helpful corrections.

Effects of Ábiotic and bioÔÉÃ factors on temporal allele frequency variations

2. Effects of abiotic and biotic factors on temporal allele frequency variations

We measured variations in the temporal frequency of two neutral alleles in order to address the interaction of evolutionary forces such as natural selection and genetic drift. The problem we addressed was the extent to which such frequency shifts are due to random phenomena or selective effects.

In our experiments, we followed allele frequency variations and the population sizes in four different environments: homogeneous (with agitation: no spatial structure), heterogeneous (without agitation: spatial structure), with and without biotic factor (with the bacteria predator *D. discoideum*). We crossed 2x2 treatments. Medium removal proceeded by serial transfer of the population every three days.

We applied these different environmental conditions to take into account the demography and to test for (i) the spatial structure and (ii) the biotic factor effects on neutral polymorphism.

We estimated changes in allele frequencies roughly every 17 bacterial generations.

The selection coefficient and the effective population sizes were estimated from temporal allele frequency variations.

Effects of abiotic and biotic factors on temporal allele frequency variations

Sandrine Adiba[1,2*], Minus van Baalen[1,2] and Frantz Depaulis[2]

[1]Laboratoire d'Ecologie, CNRS UMR7625, Université Pierre et Marie Curie, Paris Universitas, 7 quai St-Bernard, 75252 Paris, France

[2]Laboratoire d'Ecologie, CNRS UMR7625, Ecole Normale Supérieure
46 rue d'Ulm, 75230 Paris cedex 05, France

In preparation.

Abstract

Biotic and abiotic factors are responsible for the most changes of temporal allele frequency. Which alleles survive or perish is determined partly by natural selection and partly by pure chance (generic drift). Understanding evolutionary processes responsible for such variations in allele frequency is a fundamental problem in population genetics.

The maintenance of polymorphism in a population determined the rate at which a population is able to evolve and adapt to new environmental conditions.

Using the biological system bacteria (*Escherichia coli*) and amoeba (*Dictyostelium discoideum*), we follow temporal allele frequency variations under four environmental conditions: homogeneous environment, with or without biotic factor and heterogeneous environment, with or without biotic factor. We estimated allele frequency variations, population dynamics and effective population size during around 300 bacterial generations. The neutral polymorphism was maintained for one replicate out of five in the heterogeneous environment with biotic factor. Our results did not differ from genetic drift predictions.

Key words:

Biotic factor, abiotic factor, spatial structure, neutral polymorphism, experimental evolution, allele frequency, bacteria, amoeba.

Introduction

Investigation of temporal allele frequency variation is the most direct approach in population genetics, since it provides a thorough understanding of the evolutionary process. Accounting for the abundance of genetic variation relative to natural selection remains a central problem in evolutionary biology. Genetic polymorphisms constantly arise through mutation, and although most are promptly eliminated. Which particular alleles survive and which perish is determined partly by natural selection, partly by chance. The random process implies that chance alone can change allele frequency from generation to generation. This random process, known as genetic drift, does not change the allele frequencies in any systematic direction. In our work, we followed temporal allele frequency variations in four different environments: homogeneous, heterogeneous, both with and without biotic factor. We monitored the evolution of population sizes and allele frequencies. The selection coefficient and effective population sizes were estimated from the temporal allele frequency variations.

Our serial transfer regime used in our batch culture is analogous to a range of periodic environmental changes such as seasonal variations in which resources are abundant early in the season but become scarce as the population approaches its saturation density. Such an environment may, in theory promote the stable coexistence of two types on a single resource with no cross finding, provided one genotype has a growth rate advantage if a resource is abundant and the other has an opposing advantage when the resource is scarce (for more details see chapter I.3 p.37). In a recent study, Tuner *et al.* (1996) [1] examined the relative importance of both demographic trade-offs and cross-finding as explanation for coexistence. His results are consistent with theoretical work, showing that the conditions for the maintenance of polymorphisms under temporal variation are unlikely to be encountered even in simple bacterial populations.

Rainey and Travisano (1998) [2] explicitly examined the effect of spatial structure on the emergence of diversity, by allowing replicate populations of plant-colonizing bacterium *Pseudomonas fluorescens* to evolve in two kinds of microcosm that differed solely in the degree of spatial structure they provided. A spatially homogeneous environment was obtained by incubating broth cultures in an orbital shaker, and a spatially heterogeneous environment was obtained by incubating

cultures without shaking. After seven days of cultures, populations propagated in the heterogeneous environment had diversified, while populations propagated in the homogeneous environment remained genetically uniform. In order to assess the significance of spatial structure over long term maintenance of the polymorphism, populations initially propagated in the spatially heterogeneous environment were switched to the homogeneous one. Within one week after switching diversity began to fall, whereas diversity of control, unshaken (heterogeneous) populations remained high. Spatial structure appears essential both for the emergence and maintenance of these polymorphic populations. In our experiments, we used the same conditions to test for the maintenance of diversity (of neutral marker) in homogenous and spatially structured environments.

The arms race of adaptation and counter adaptation in predator-prey system is an evolutionary dynamic with broad implications, including the maintenance of diversity. Such antagonistic coevolution is suspected to be widespread in nature. We investigated here the contribution of natural selection and genetic drift to the maintenance of neutral polymorphism, using bacteria populations stained with a neutral marker and the haploid social amoeba *Dictyostelium discoideum* as biotic factor. The latter organism is a particular effective bacterial predator.

The prey diversity appeared dependent upon the presence of both the biotic factor and spatial structure. Ecological conditions can shape the evolutionary trajectory of this predator-prey system.

Material and methods

Strains and culture conditions

Bacteria

We used the laboratory adapted strains: *E. coli* REL 606 and *E. coli* REL 607, which are harmless and phagocytized by *D. discoideum* [3]. *E. coli* REL 606 and *E. coli* REL 607 are respectively Ara^+ and Ara^- ancestors. They are isogenic, with the exception of the spontaneous mutation to Ara^+ in REL 607. We used TA agar medium to set apart Ara^+ and Ara^- strains. Colonies of *Ara-* strains appear red on TA agar, while those of Ara^+ appear white.

An overnight culture in 10 ml of HL5 liquid medium with 130 rpm/min agitation was prepared in a 50 ml Falcon tube, incubated at 37°C. On the first day

of the experiment, the culture was centrifuged (20,000 rpm/min, 20 min), washed once and re-suspended in MCPB.

A sample of each replicate was frozen weekly (Note that these are not clonal samples).

A volume of 900 µl of the bacterial culture was mixed with 300 µl of Glycerol (60 %) and frozen at -80°C. Two tubes of each replicate were assayed.

The number of bacteria (unit=CFU colony forming unit) was estimated by plating 10 µl within a range of dilution rates on TA medium.

Amoeba

The amoeba *D. discoideum* AX3, an axenic strain was grown in 10 ml of HL5 culture liquid medium 25 cm^2 bottles, incubated at 24°C. Cells from mid-logarithmic cultures were centrifuged (2,000 rpm/min, 7 min) and washed once with MCPB.

A volume of 500 µl of the co-culture (bacteria and amoebae) was mixed with 500 µl DMSO (20%) frozen at -80°C. Three replicates were prepared.

The number of *D. discoideum* was estimated using a hemocytometer.

Cell density (culture) = number of cells x 90,000/number square count.

Assays

A mixture of 50% *E. coli* REL 606 and 50% *E. coli* REL 607 was prepared in 10 ml SM/25 on 50 ml Falcon tube and on bottle. Five replicates over the ten tubes (*vs.* bottles) were prepared with no biotic factor *D. discoideum* and in the five others we added amoeba. The ten Falcon tubes were incubated at 24°C and shaken at 100 rpm, to remove spatial structure. The ten bottles were incubated at 24°C without agitation to test the effect of spatial structure.

Every three other days, a bottleneck was performed with a dilution of 100 µl into 10 ml SM/25 fresh medium (APPENDIX 5: PROTOCOL: Coevolution experiments, p.200).

Statistical analyses

Population dynamics

Differences between the various environments, replicates and generation times were tested using ANOVA with JUMP software (full model and interactions),

separately for the time of each environment. The cell number was transformed in a logarithmic scale for statistical analyses.

The bacterial population size was estimated weekly before a bottleneck over 126 days (= 294 bacterial generations). For each replicate, a sample (300 µl for *E. coli* culture and 500 µl for *E. coli/ D. discoideum* co-culture) was frozen and plated (10 µl) to estimate the various proportions of each *E. coli* strain and bacterial density.

Amoeba population size was estimated before each bottleneck (every three days).

Estimates of the number of generations

Prior to every 1/100th dilution, the bacterial density was around 10^9 cells/ml. After dilution, the number of bacteria was thus around 10^7 cells/ml. Roughly seven generations are needed to increase the cell densities from 10^7 to 10^9. The dilution was applied every three days, but for two days the bacterial density may be constant (carrying capacity).

Thus across the 126 days of the experiments, the bacterial number of generations was:

$\frac{126 \times 7}{3}$ =294 bacterial generations

Allele frequency variation

Effective population size

In a simple Wright-Fisher model, the per generation drift variance of allele frequency is $p(1-p)/2N_e$ where p is the population frequency in the previous generation and N_e the effective population size. The effect of initial allele frequency can be compensated for by using some variation of Wright's standard variance F and this approach has formed the basis for several efforts to relate N_e to observe changes in allele frequencies.

If p_0 is the population frequency in generation 0 and p_t is the population frequency in generation t, the expected parametric F takes the form:

$$E((p_0 - p_t)^2) = p_0(1 - p_0)(1 - [1 - 1/2N_e]^t) \quad (1)$$

$$\text{such that } E(F) \approx 1 - [1 - 1/2N_e]^t \quad (2)$$

and *Ne* may be estimated from *t/(2F)* if *t* is not too large.

However, since one has generally access to sample and not population allele frequencies, *F* should be estimated by $\hat{F}$, which also takes into account sampling effects and induced by the method used to estimate $p_0(1-p_0)$. During initial use of the temporal method to estimate effective population size, Krimbas and Tsakas (1971) [4] computed the mean $\hat{F}$ and subtracted the quantity $[1/(2S_0)+1/(2S_t)]$ with S_t the sampling sizes, to account for sampling error. Krimbas and Tsakas [4] used the following expression to calculate $\hat{F}$ for a single locus:

$$\hat{F}_a = \frac{1}{K}\sum_{i=1}^{K} \frac{(x_i - y_i)^2}{x_i(1-x_i)} \qquad (3)$$

with *K* being the number of segregating alleles. However, $\hat{F}_a$ is infinitely large when an allele is found at time *t* but not at time 0.

Nei and Tajima (1981) [5] proposed an alternative estimator which overcomes this limitation:

$$\hat{F}_c = \frac{1}{K}\sum_{i=1}^{K} \frac{(x_i - y_i)^2}{(x_i + y_i)/2 - x_i y_i} \qquad (4)$$

Another method was suggested by Pollak (1983) [6]:

$$\hat{F}_k = \frac{1}{K-1}\sum_{i=1}^{K} \frac{(x_i - y_i)^2}{(x_i + y_i)/2} \qquad (5)$$

The resulting formula to estimate $\hat{N}_e$ is then:

$$\hat{N}_e = \frac{t}{2[\hat{F}_c - 1/(2S_0) - 1/(2S_t)]} \qquad (6)$$

And similarly for $\hat{F}_k$.

The differences between the various environments for all replicates were tested using an ANOVA (JUMP software), with a model involving environment, replicate, and environment*replicate effects.

We compared this estimation with the census population size before each bottleneck and taking the dilution rate into account.

Selection versus genetic drift effects

The selection coefficient was estimated for all replicates and all environments (spatial structure, biotic factor, and the combination spatial structure*biotic factor) using the following equation for the frequency over time linear regression:

$$\ln(\frac{p(t)}{q(t)}) = \ln(\frac{p(0)}{q(0)}) + st \quad (7)$$

with $p(t)$ *Ara*$^{+}$ the allele frequency at time *t*, $q(t)$ the *Ara*$^{-}$ allele frequency at time *t, s* the selective coefficient [7].

To set the significance of the slope on such time series, correlated data, a R script was developed.

For each replicate, time was re-sampled 1,000 times (bootstrap). For each bootstrap replicate, a similar linear regression was carried and the associative selective coefficient estimated, thereby empirically providing empirically the corresponding null distribution. The selection coefficient estimated from the raw data was then compared to this distribution in order to assess its *p* value. This analysis allowed us to test the significance of selection with respect to genetic drift.

Time to fixation or extinction

For a selectively neutral allele, assuming an initial frequency *p*, Kimura and Ohta (1969) [8] showed that the mean time $\bar{t}_1(p)$, to fixation of an allele is

$$\bar{t}_1(p) = -4N\left[\frac{(1-p)}{p}\right]\ln(1-p) \quad (8)$$

Similarly, they showed that the mean time to loss $\bar{t}_0(p)$ is

$$\bar{t}_0(p) = -4N\left[\frac{p}{(1-p)}\right]\ln(p) \quad (9)$$

Differences among the various environments for all replicates were tested using ANOVA (JUMP software) with a model involving environment, replicate, environment*replicate effects.

Results

Bacterial population sizes

Bacterial population sizes were estimated by plating 10 µl of the culture at different dilution rates (10^{-3} to 10^{-8}) once a week before the bottlenecks.

During the first 20 day (around the 40th generation), bacterial carrying capacity was increasing quickly (Figure IV-5 a,b,c) except for one ecological condition: biotic pressure and spatial structured population (Figure IV-5d). Under these conditions, the carrying capacity decreased at the beginning and then increased but less quickly than under other ecological conditions. In the meantime, allele frequency variations were directional. The same carrying capacity as under other conditions was reached after the 40th day (around 96th generation) and we subsequently observed bidirectional fluctuations of allele frequency (Figure IV-11, p.149).

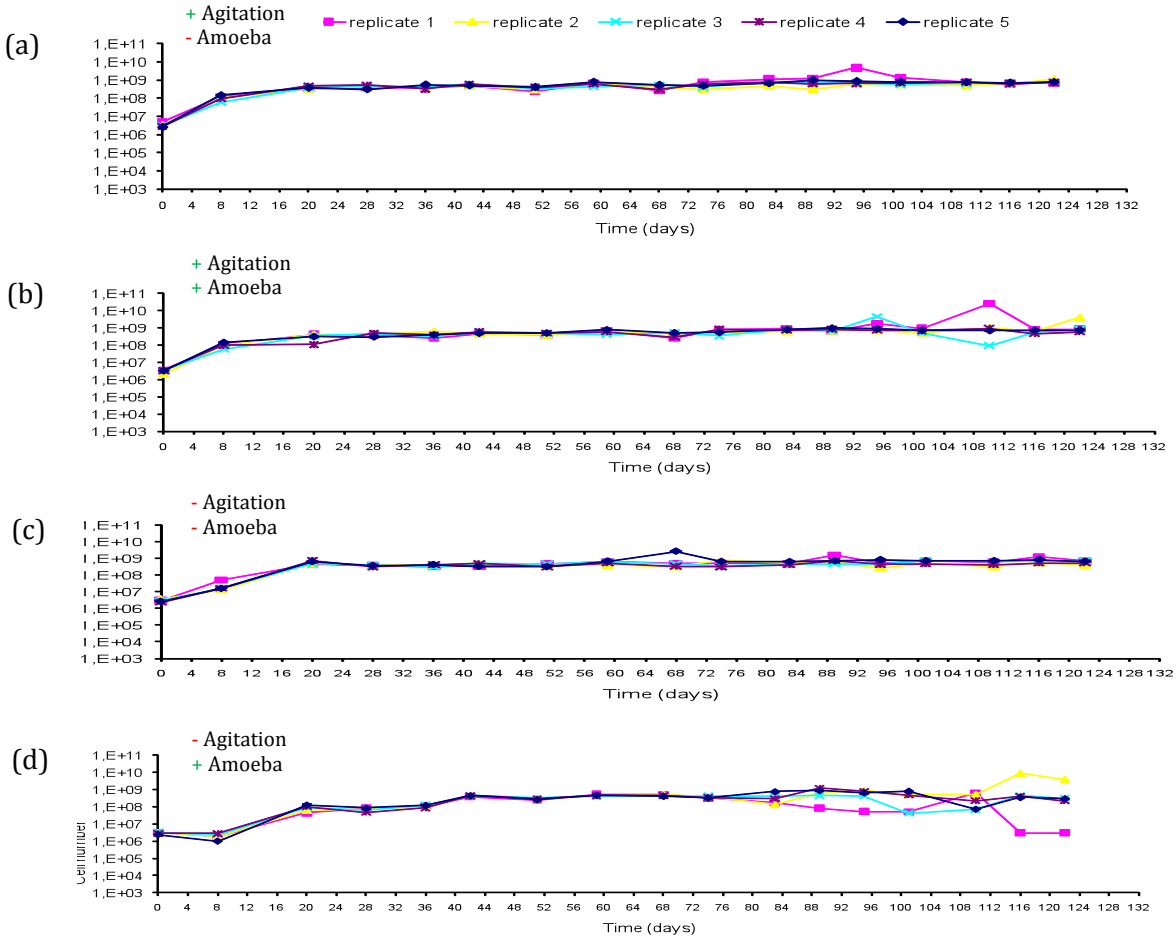


Figure IV-5: *E. coli* population dynamic

Each color represents one replicate

(a) With agitation (+) and without amoeba (-)

(b) With agitation (+) and with amoeba (+)

(c) Without agitation (-) and without amoeba (-)

(d) Without agitation (-) and with amoeba (+)

We compared population sizes in all environments. The statistical analyses showed a significant difference among the various environments and across time (Table IV-3).

Table IV-3: Bacterial population sizes differences in various environments

Source	Nombre de coefficients	Degré(s) de liberté	Somme des carrés	Rapport F	Prob. > F
environment	3	3	56,59370	12,6828	<,0001*
time(Gen)	1	1	312,18389	209,8837	<,0001*

The bacteria carrying capacity is lower in the spatial structured population with the biotic pressure *D. discoideum* and increased continuously with the time (Figure IV-6).

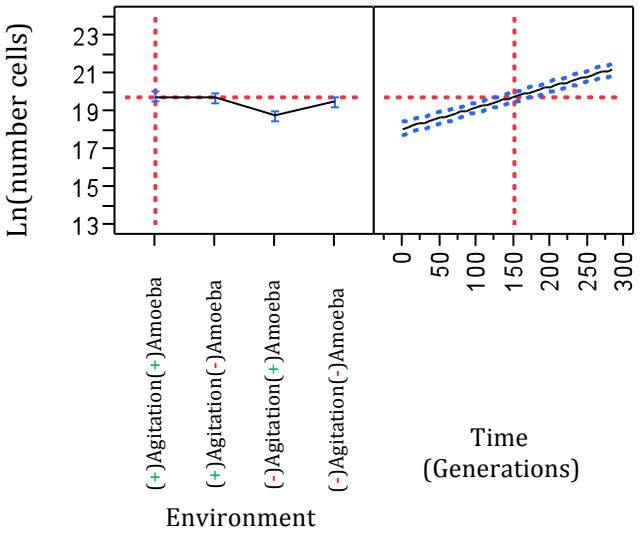


Figure IV-6: Bacterial carrying capacity in the various environments and accross time

Finally, we tested the time effects separately for each environment with ANOVA. The carrying capacity significantly increased with time in all environments (Table IV-4, Figure IV-7).

Table IV-4: Time effect in bacterial population sizes on each environment

Environment	F	*p*	
(+) agitation (-)amoeba	64.47	<0.0001	*
(+)agitation (+)amoeba	64.49	<0.0001	*
(-) agitation (-)amoeba	56.64	<0.0001	*
(-)agitation (+)amoeba	39.99	<0.0001	*

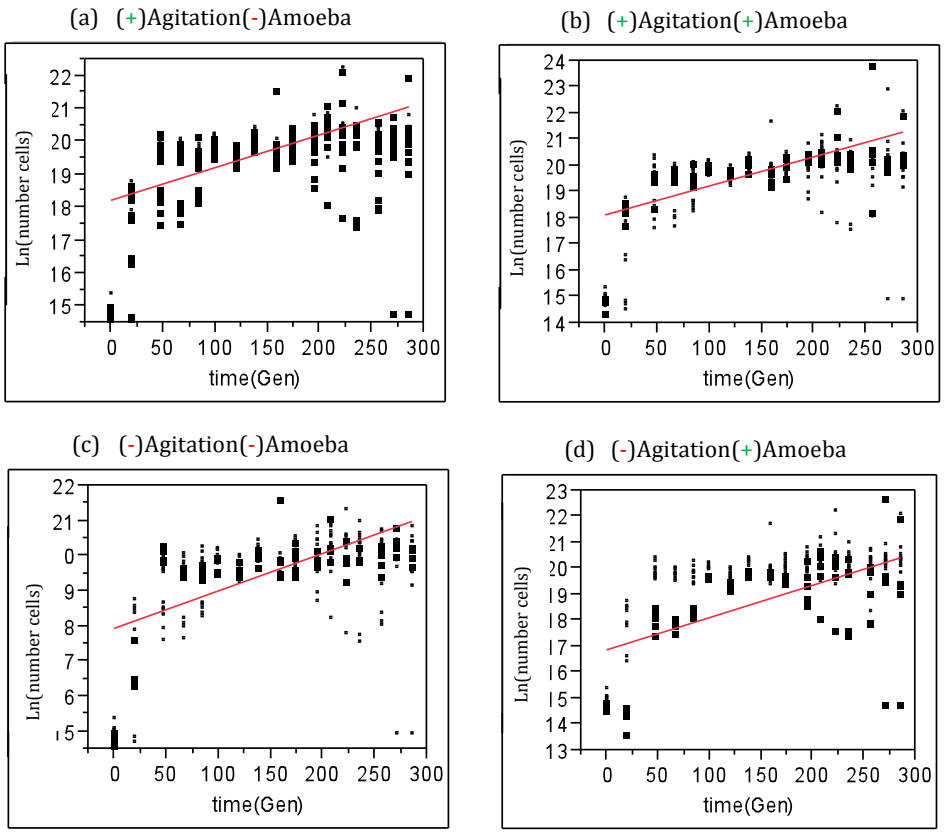


Figure IV-7: Linear regression of bacterial population sizes in each environment

(a) without spatial structure and without biotic factor.
(b) without spatial structure and with biotic factor.
(c) with spatial structure and without biotic factor.
(d) with spatial structure and with biotic factor.

To test separately shaking and amoeba effects, we applied ANOVA analyses pairwise across environments (Table IV-5).

Table IV-5: Statistical test between two environments

Environment 1	Environment 2	Effect tested	*F*	*p*	
(+) Agitation (-) Amoeba	(+) Agitation (+) Amoeba	Biotic factor		n.s	
(-) Agitation (-) Amoeba	(-) Agitation (+) Amoeba	Biotic factor	13.815	0.0003	*
(+) Agitation (+) Amoeba	(-) Agitation (+) Amoeba	Spatial Structure	24.3517	<0.0001	*
(+) Agitation (-) Amoeba	(-) Agitation (-) Amoeba	Spatial structure		n.s	
(+) Agitation (-) Amoeba	(-) Agitation (+) Amoeba	Spatial Structure *Biotic factor	22.112	<0.0001	*
(-) Agitation (-) Amoeba	(+) Agitation (+) Amoeba	Spatial Structure *Biotic factor		n.s	

With agitation, the population dynamic was not different with or without amoeba. With biotic factor, there was no difference with or without agitation.

We observed significant difference in bacterial population dynamics for the interaction spatial structure*biotic factor except between the environment without agitation-without amoeba *versus* agitation-amoeba.

Amoebae population dynamic

The population densities of amoebae were measured before each serial transfer (every third day) for *D. discoideum*. (Figure IV-8, Figure IV-9).

The growth rates oscillated for all replicates much more in the homogeneous than in heterogeneous environment (Figure IV-10) and become more stable after the 84th day (around 200th generation) and 68th day (around 160th generation) respectively.

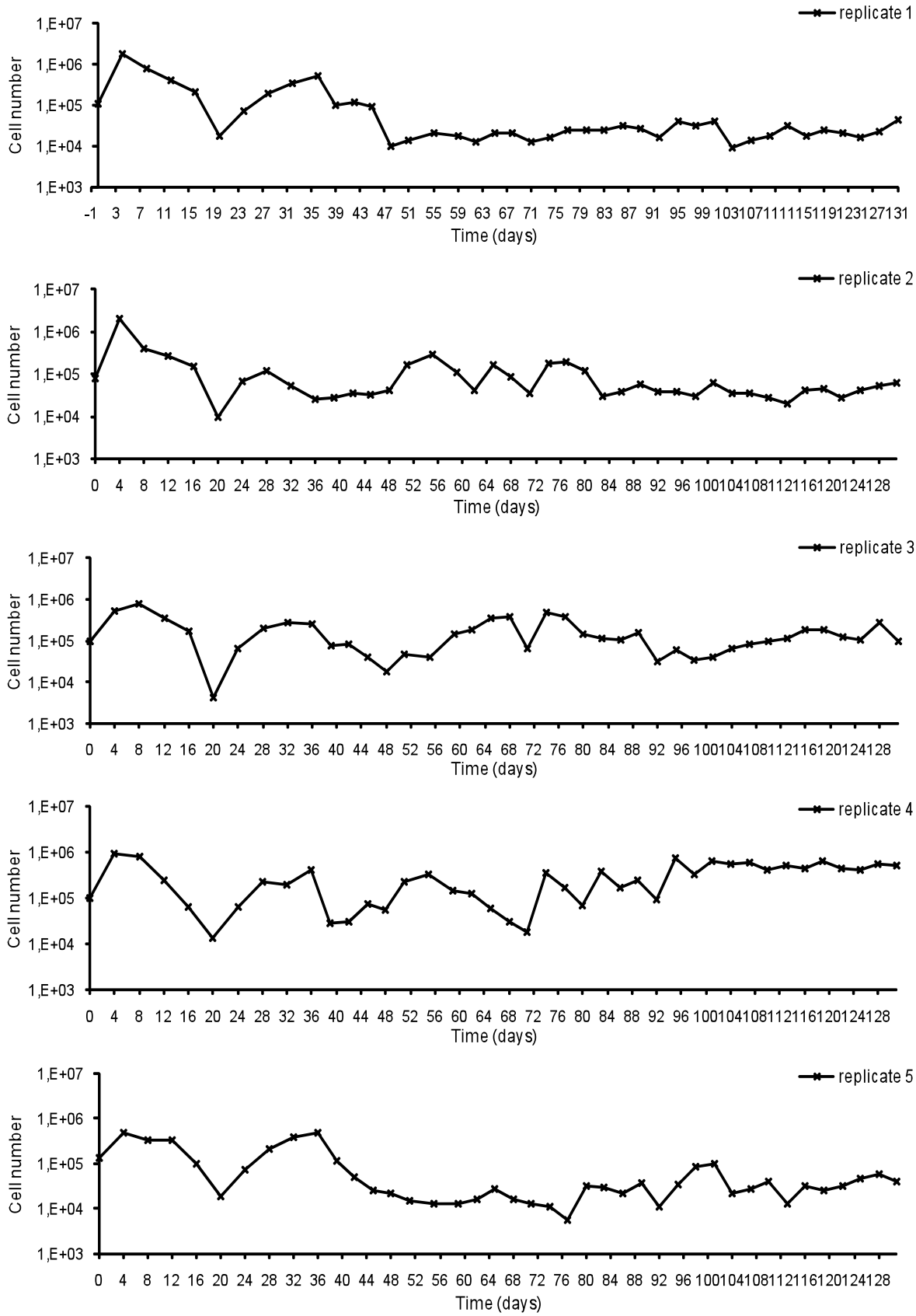


Figure IV-8: *D. discoideum* population dynamic without spatial structure (replicate one to five)

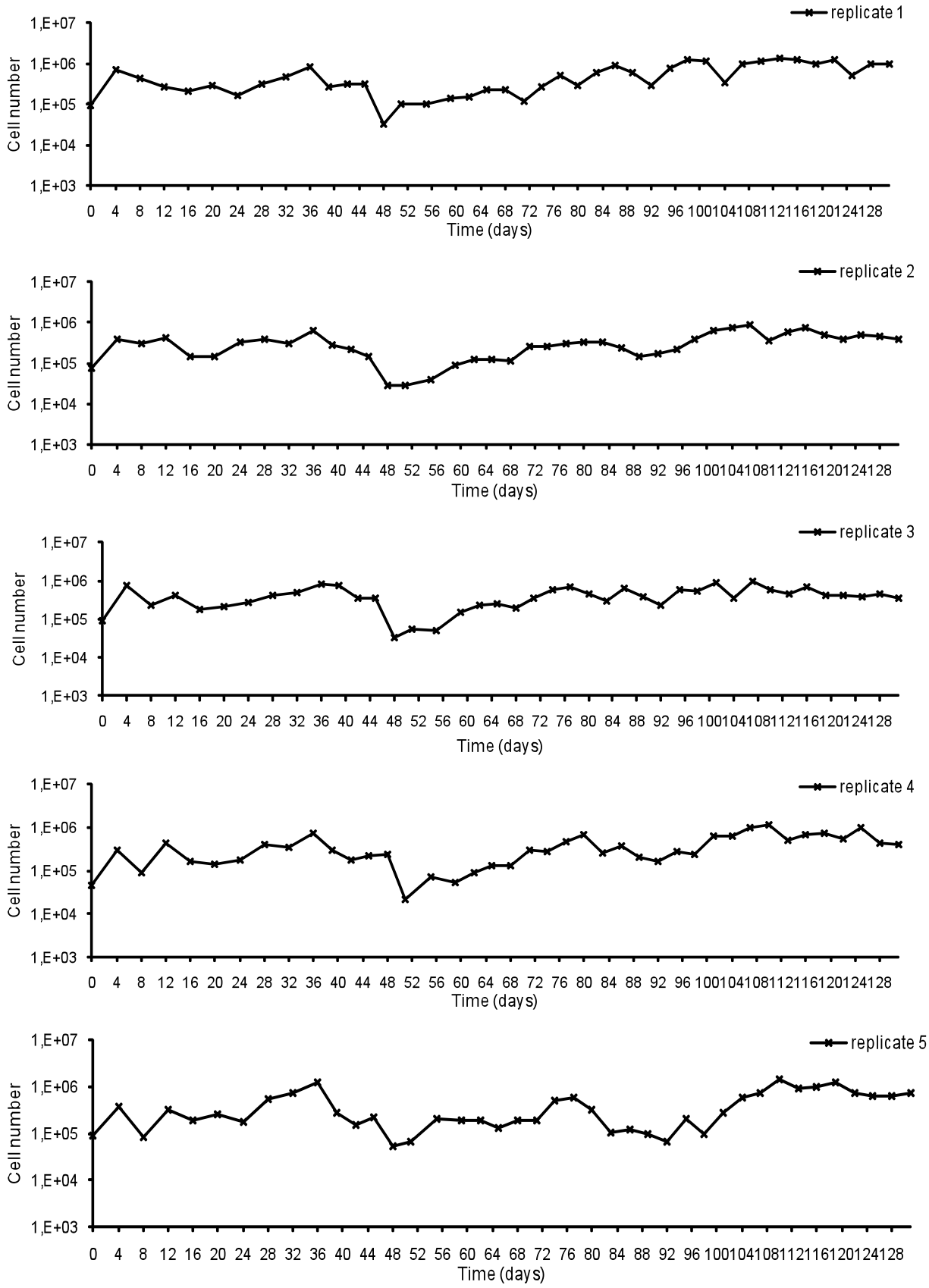


Figure IV-9: *D. discoideum* population dynamic in spatially structured environment (Replicate one to five).

The average population size was greater in the spatially structured population than on the spatially non-structured population after 126 days (300 generations). Without spatial structure environment, the mean population size decreased with the time (Figure IV-10, Table IV-6).

Mean by environment :

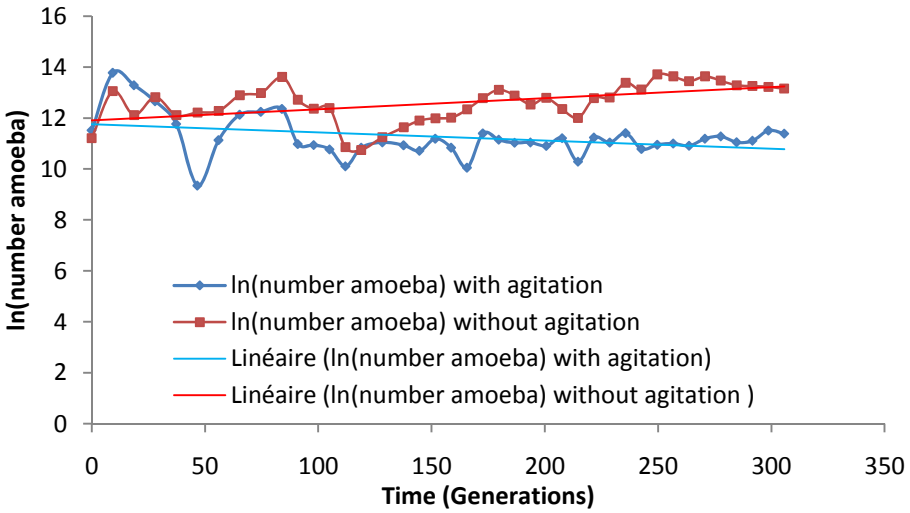


Figure IV-10: Mean ln(*D. discoideum*) per environment
Mean ln(*D. discoideum*) of five replicates per environment and linear regression

Table IV-6: Statistical interactions for the two environments: with and without spatial structure

Source	Nombre de coefficients	Degré(s) de liberté	Somme des carrés	Rapport F	Prob. > F
Environment	1	1	194,00977	188,9262	<,0001*
Time(Generations)	1	1	1,00339	0,9771	0,3235
Replicate	1	1	0,55448	0,5399	0,4629
Environment*Time	1	1	44,90078	43,7243	<,0001*
Environment*Replicate	1	1	6,07208	5,9130	0,0155*
Environment*Time*Replicate	1	1	5,17753	5,0419	0,0253*

Allele frequency variations

Demography:

To maintain our culture during the time of our experiments and to increase the number of generations, we carried out serial transfers with 1/100 dilution of 10 ml mixed bacterial cultures (50%;*E. coli* 606/50% *E. coli* 607 at time 0) every third day. Cultures were kept at 24°C with or without agitation at 100 rpm. To take the

demography effect into account, we monitored temporal allele frequency variations.

Bacterial allele frequencies were measured for 126 days corresponding to roughly 300 bacterial generations. We observed bidirectional fluctuations at the beginning of the experiment and all replicates were fixed at the end of this time (Figure IV-11a).

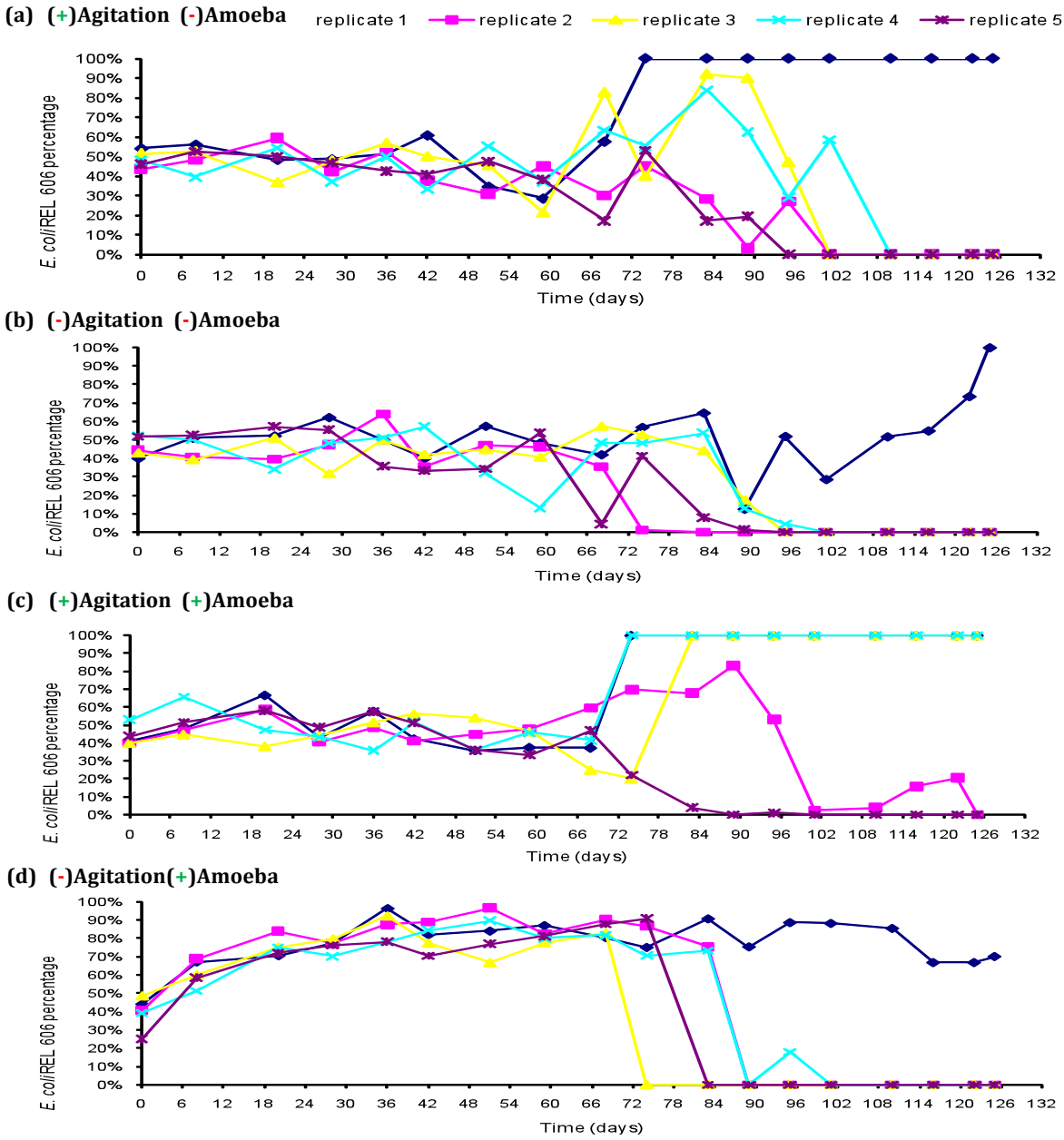


Figure IV-11: *E. coli* REL606 temporal allele frequency variations

(a) Without spatial structure and without biotic factor.
(b) With spatial structure and without biotic factor.
(c) Without spatial structure and with biotic factor.
(d) With spatial structure and with biotic factor.

Spatial structure

To test for the spatial structure effect on temporal allele frequency variations, cultures were not shaken.

We observed bidirectional fluctuations and fixations in all replicates by the end of the experiment (Figure IV-11b).

Biotic pressure

The same experiments were carried out with biotic selective pressure, with the addition of the amoeba *D. discoideum* that may predate bacteria. In a homogeneous environment (with shaking), we observed bidirectional fluctuations and fixations in all replicates (Figure IV-11c) after roughly 300 bacterial generations. In a heterogeneous environment (spatially structured), we observed repeated and rapid increase in allele frequency at the beginning of the experiment, then bidirectional fluctuations and fixation in four of the five replicates (Figure IV-11d).

Statistical analyses

Selection versus drift effects

Selection coefficients were estimated from allele frequency variations using a linear regression (Eq (7)).

We found significant selection coefficients only with the environment in the presence of agitation and amoeba and for replicate number two ($\hat{s}$=0.00773, p=0.047), four ($\hat{s}$ =0.0068; p=0.028) and five ($\hat{s}$ =0.0143; p=0).

For the other environments, selection coefficients were not significant thus the allele frequency variations may be considered as neutral. Genetic drift acted more strongly than natural selection for these environments or the selection was non-linear (homogeneous) over the generation time.

Time to fixation or extinction

We compared time to fixation among all environments for all replicates (Eqs. (8,9), Table IV-7). No significant differences were observed in time to fixation or extinction (Table IV-8).

Table IV-7: Time to fixation observed and under the neutral model and constant population size

Environment	R	$\hat{F}_c$	$\hat{F}_k$	Time($\hat{F}_c$)	Time($\hat{F}_k$)	Time (obs)	N_e($\hat{F}_c$)	N_e($\hat{F}_k$)	N_e(obs)
(+) Agitation (-) Amoeba	1	0.54	0.21	237	645	173	164	446	3.74E+05
	2	0.36	0.35	517	538	236	350	365	1.45E+05
	3	0.47	0.40	361	431	236	267	318	1.83E+05
	4	0.33	0.30	596	668	257	421	472	1.66E+05
	5	0.42	0.40	372	388	222	281	293	1.32E+05
(+) Agitation (+) Amoeba	1	0.66	0.30	169	381	173	136	307	2.19E+05
	2	0.28	0.22	936	1241	292	611	810	1.20E+05
	3	0.61	0.29	201	436	194	165	357	2.13E+05
	4	0.55	0.22	261	713	194	182	498	2.70E+05
	5	0.45	0.44	404	418	236	274	283	1.32E+05
(-) Agitation (-) Amoeba	1	0.21	0.13	929	1614	292	760	1321	1.90E+05
	2	0.50	0.50	296	301	194	202	205	1.06E+05
	3	0.36	0.35	498	513	222	334	344	1.14E+05
	4	0.48	0.45	354	375	236	262	277	1.21E+05
	5	0.57	0.55	293	305	236	216	225	1.20E+05
(-) Agitation (+) Amoeba	1	0.52	0.25	NF	NF	NF	294	647	3.92E+04
	2	0.74	0.53	221	311	208	145	205	6.77E+04
	3	0.69	0.60	185	214	173	130	151	6.42E+04
	4	0.59	0.45	322	437	236	208	282	7.10E+04
	5	0.68	0.52	256	341	194	148	197	4.10E+04

(+) Agitation= without spatial structure, (-) Agitation= with spatial structure, (+) Amoeba= with biotic factor, (-) Amoeba= without biotic factor. R=Replicate, NF: not fixed, Time($\hat{F}_c$), Time($\hat{F}_k$), N_e($\hat{F}_c$), N_e($\hat{F}_k$) estimated with Eqs (8,9 and 6) respectively, *Ne*(obs) was the harmonic mean of the cell number before the dilution times the dilution rate.

Table IV-8: Selection versus genetic drift from time to fixation or extinction

Source	Nombre de coefficients	Degré(s) de liberté	Somme des carrés	Rapport F	Prob. > F
Environment	3	3	960,4000	0,1769	0,9100
Replicate	1	1	207,0250	0,1144	0,7410
Replicate*Environment	3	3	4599,8750	0,8474	0,4942

Effective population size

For all replicates and environments, we compared the effective sizes estimated with the $\hat{F}_C$ or $\hat{F}_k$ estimator (Eq. (6), Table IV-8). The two estimators did not reveal any substantial difference. Nevertheless, when we compared these effective population sizes with the effective population size estimated using the harmonic mean of the census population size, the differences were roughly about 1,000 fold.

Discussion

The aim of our study was to address to what extent allele frequency trajectories where affected by selection and drift according to the presence of predator and spatial structure.

In our experiments, we followed bacteria allele frequency variations and the population sizes in four different environments: homogeneous (with agitation, no spatial structure), heterogeneous (without agitation, spatially structured), homogeneous with biotic factor (with agitation and with the bacteria predator *D. discoideum*), heterogeneous and biotic factor (without agitation and with the bacteria predator *D. discoideum*). Populations were maintained by serial transfer every third day: 100 µl of culture was transferred into 10 ml fresh media. Dilution by a factor of 1/100 yielded roughly seven generations per transfer. Before the transfer, bacterial density was on the order of $\sim 10^7$ to 10^9 cells/ml. Bacterial densities were obtained by plating cultures. Samples from each population were periodically (weekly ~17 bacterial generations) stored at -80°C. Changes in allele frequencies during the experiment were assessed weekly by plating 10 µl of various culture dilutions (10^{-4} to 10^{-8}). Five parallel cultures were started with a mix of 50% *E. coli Ara*$^+$ and 50% *E. coli Ara*$^-$. The cultures were propagated for around 300 generations. Analyses of the bacterial culture provided allele frequency trajectories and population dynamics. Hence, for each of the five replicates and four environmental conditions, we surveyed the allelic distribution every other 17 bacterial generations. Changes in allele frequency could in principle have three different causes: (i) new mutation, (ii) genetic drift, (iii) selection. The first two forces are not directed and affect all alleles to the same extent. However, selection increases only the frequency of the neutral allele associated with the advantageous mutation. Hence, it is possible to trace adaptive events by significance of the increase in neutral allele frequency.

Bacterial and *D. discoideum* population dynamics

Bacteria: dynamic of adaptation

The increased carrying capacity with time demonstrated bacterial adaptation to a new environment. Such dynamic initially rapid and decelerating over time

indicates that populations, after being placed in a new environment, are evolving from a region of low fitness towards an adaptive peak or plateau [9].

Beneficial substitutions tend to show larger effects early in an experiment, when the population is far from being at an adaptive peak, rather than later as it approaches a local peak. Two replicate populations might reach different final fitness levels if, through the random effects of mutation and drift, they move toward different adaptive peaks [9].

In asexual populations such as common bacteria, beneficial mutations that occur in different clones may be affected by 'clonal interference' that is caused by competition among beneficial mutations. This affects the delay of their spread and partly account for the observed plateau.

The bacterial carrying capacity was lower in the heterogeneous environment and with biotic factor (Figure IV-5). The impact of *D. discoideum* was higher in this spatial structure environment than in the homogeneous environment. This may be explained by possibily better *D. discoideum* predation in an environment without shaking.

D. discoideum population dynamic

We observed some oscillations in the *D. discoideum* population sizes (Figure IV-8 and Figure IV-9), which was measured every third day. These fluctuations could have been due to *D. discoideum* coevolution with bacteria.

Bacteria could evolve during the course of this coevolution, and some of them could have become resistant to digestion by *D. discoideum*. An arm race [10] could have begun between resistance of the bacteria to the digestion by *D. discoideum* and increased ability of *D. discoideum* to eat bacteria (see chapter III.7 p.96). Nevertheless, these oscillations stopped after a while, and we observed a sort of stability (Figure IV-8 and Figure IV-9). Such observations (oscillations at the beginning followed by stability) may be due to evolution in the resistance to digestion of bacteria and/or to evolution of the ability of *D. discoideum* to eat bacteria, up to a point at which we observed equilibrium between the two actors. This seems to involve negative-frequency dependant selection as in the red-queen theory (advantage to the rare phenotype) [11]. However, in such cases, oscillations in allele frequency should be observed and neutral polymorphism should be maintained (as in R1 spatial structure and Biotic Factor, Figure IV-11d and Figure

IV-9) and we did not observed the maintenance of polymorphism in all environments.

In the spatially structured environment, we observed greater mean population sizes (Figure IV-10), which should imply greater impact on bacteria and on the variation of allele frequency. It is in this environment that neutral polymorphism was best maintained (Figure IV-11d).

Natural selection versus genetic drift

As defined by Wright [12] random genetic drift includes bidirectional fluctuations in gene frequencies, which are of indeterminate direction. Such variation is caused by accidents in the generation sampling of genes from a population of finite effective size. However, random drift may be important in conjunction with systematic pressures on the gene frequencies, particularly with natural selection. What is most needed is the kind of experimental evidence that would allow analysis of the combined effects of random drift and natural selection.

In the latter environmental condition, the estimation of the selection coefficient with the linear regression of the equation (7) showed a significant value for three out of the five replicates (R2, R4 and-R5). For the other environments, we found no significant selective coefficient. This means that for all these environments and replicates, genetic drift may only be sufficient to explain the allele frequency variations.

The effective population sizes were estimated with $\hat{F}_C$ and $\hat{F}_k$ values. We found no significant difference between the four environments and demonstrated no evidence for natural selection using the effective population sizes. But these estimated effective population sizes were lower than the census population size (roughly 10^7 cells/ml) and the estimation with the harmonic mean (Table IV-7). This may be explained by the assumptions used to calculate the estimators $\hat{F}_C$ and $\hat{F}_k$ which are low allele frequency variations and constant population sizes. Furthermore, in our experiments, generation times were not constants.

We then compared the times to fixation with the expected values calculated base on the assumption of a neutral model. If selection was acting, we should have observed a time that was shorter than the expected value. We found such a substantial difference for one replicate (R1, Figure IV-11) in the homogeneous

environment with biotic factor. All other comparisons showed no evidence for natural selection and we cannot reject the assumption of genetic drift.

In nature, populations frequently experience changes in size under the influence of a number of factors including varying physical conditions, resource availability, habitat, and predator density. In growing populations, selection is more effective, increasing the probability that beneficial alleles are fixed and that deleterious alleles are lost. Conversely, in shrinking populations, selection is less effective: beneficial mutations are less likely to go to fixation whereas deleterious alleles are more likely to go to fixation. If the population is growing, this will reduce the amount of sampling error or genetic drift that occurs. In declining populations, the amount of genetic drift in the population increases. Such changes in population size increase genetic drift and slow the selective spread of beneficial alleles, but also slow the selective elimination of deleterious alleles. Bottlenecks affect the evolutionary dynamic, reducing the probability that a beneficial mutation will reach fixation. Wahl and Gerrish [13] found that the selective advantage of mutations that eventually survive bottlenecks is about twice as large as the mean selective advantage of all beneficial mutations that occur. They determined that the dilution ratio 0.1-0.2 allows the largest number of beneficial mutation to survive. In our experiments, we used a dilution ratio of 0.01. This serial transfer regime can in theory promote stable coexistence of two genotypes on a single resource with no cross-feeding, provided one genotype has a growth rate advantage when the resource is abundant (at the beginning of the transfer) and the other has an opposing advantage when the resource is scarce [1, 14, 15]. In a recent study of a polymorphism steadily maintained under a serial transfer regime, Turner *et al.* [1] examined the relative importance of both demographic trade-offs and cross-feeding as explanation for coexistence. Trade-offs were observed between growth rates at high and low concentrations of glucose, however, the magnitude of the trade-off alone was insufficient to explain the coexistence of genotypes (subsequent experiments revealed the polymorphism to be maintained by cross-feeding). This outcome and our results are consistent with theoretical work showing that the conditions required for the maintenance of polymorphisms under temporal variations are stringent and unlikely to be encountered even in simple bacterial populations [16, 17].

Maintenance of polymorphism

Our experiments demonstrated that in both a homogeneous (non-spatial structure) or heterogeneous (spatial structure) environment, neutral polymorphism was not ultimately maintained. We observed the maintenance of the neutral polymorphisms for the entire course of the experiment for only one replicate (R1, Figure IV-11) in the heterogeneous environment - with biotic factor but not in the homogeneous environment – with biotic factor. This result is consistent with other studies [2] showing that spatial structure and predation may promote polymorphism. Nevertheless, in this last environment, allele frequency variations were all in the same direction for the same maker. This brought up the question of the neutrality of the marker we used under our experimental conditions; further investigations are needed to confirm this neutrality.

Competition

Competition is for resources or space. In our experimental setup, we switched between rich environment and a scarce one, which means that resources fluctuated with time but increase evolutionary rates. In the heterogeneous environment (with spatial structure) competition for space might have an impact on the maintenance of neutral polymorphism. The competitive exclusion principle [18] predicts that evolution in an environment containing a single limiting resource should proceed by a series of clonal replacements. However, in apparent contradiction to this competitive-exclusion principle, stable polymorphism may evolve in asexual populations. Selection based solely on the exploitation of available resources can explain patterns of diversification and the outcomes of this process are strongly affected by the availability of resources in the environment. Competition for resources may generate divergent selection, favoring the evolution of genotypes adapted to different niches. Dykhuizen and Davis (1980) [19] used two genotypes of *E. coli*, one of which was deleted for lactose use, to study competition in mixtures of lactose and maltose. Over the course of the experiment, a novel mutation occurred in the *Lac*$^{+}$ generalist, conferring the ability to grow at low level of lactose. This mutant was able to coexist with the ancestral strain, apparently because of functional interference between mechanisms responsible for the uptake of the two substrates. Maclean *et al.* (2005) [20] found that metabolic diversity evolved within experimental *Pseudomonas* line and was

maintained by negative-frequency-dependant selection, but they observed no selective sweep. Thus, adaptation to mixtures of substrates may or may not be associated with the periodic fixation of a single genotype. Hall & Colegrave (2007) [21] found that diversity was greatest in environments with intermediate resource supply compared to scarce or rich resource environment.

In an environment that provides ecological opportunity and in which selection has favored the evolution of niche specialists, the maintenance of coexisting genotypes is ensured by density-dependent process [16]. In a constant resource environment, selection will operate in a negative-frequency-dependant manner, favoring genotypes when they are rare (because corresponding resources will be relatively more abundant) but not when they are common (because resources will be scarce and competition intense [22, 23]). Elena and Lenski [24] found that negative-frequency-dependent selection operated in the six long term *E. coli* populations. Negative-frequency-dependent selection is a powerful force maintaining variation in these experimental bacterial populations. It is also possible that it plays a significant role in maintaining the genetic diversity observed in most bacterial species [25, 26].

Our data suggest that even if there is competition for resources or space, ecological opportunity is necessary but not sufficient for the evolution of stable polymorphism. Without specialization, stable coexistence of genotypes is generally not possible.

Spatial structure and biotic pressure

Other factors tested for the maintenance of the polymorphism were spatial structure and biotic factor. The combination of spatial structure and biotic factor was found to be a key factor in the emergence or maintenance of polymorphism. Our spatially homogeneous environment was obtained by incubating broth cultures in an orbital shaker, whereas our spatially heterogeneous environment was obtained by incubating cultures without shaking. We found that neutral polymorphism was not maintained in homogeneous or heterogeneous environments. Spatial structure is an important component of natural environments; most species do not occupy spatially distinct niches and instead exploit a subset from among wide range of simultaneously available resources. Rainey and Travisano [27] found that populations propagated in a heterogeneous

environment had diversified, whereas populations propagated in a homogeneous environment remained genetically uniform. The various genotypes that emerged in a heterogeneous environment were adapted to different spatial niches and were identified by their different colony morphologies. When spatial structure was eliminated by constantly shaking the tubes, diversity was quickly lost, since there was no longer a range of distinct niches. In our experiments, we followed only the neutral marker *Ara*$^{+}$, but the other bacterial traits may have changed and diversified.

Our experiments showed that neutral polymorphism was maintained only for one replicate in the heterogeneous environment and biotic pressure. Predation is one of the factors believed to have a major impact on competitive interactions and to favor diversity. But depending on the nature of predation, predators may increase, decrease or have no effect on diversity. Our experiments showed for all replicates that predation did not maintain the neutral polymorphism. Predation is a common hypothesis for high biological diversity. However, the argument that predators maintain diversity by keeping populations below levels at which they compete is now known to be highly simplistic [28, 29]. Predators may add density dependence to their prey populations [30]. They may promote diversity when they diversified in specialists. Predators are often invoked as biological disturbance agents inflicting mortality patchily in space and time. Predators may have a role analogous to a spatiotemporally variable environment, which may lead to coexistence by the spatial storage effect [31].

Using our experimental design, we were able to demonstrate that biotic factor only or spatial structure alone were not sufficient to maintain neutral polymorphism, whereas when we combined the two elements, we observed for one replicate the maintenance of neutral polymorphism. This result must be confirmed with further replicates.

Ecological opportunity is necessary but insufficient for the evolution of stable polymorphism. Without specialization, stable coexistence of genotypes sometimes seems impossible.

Literature cited:

1. Turner, P., V. Souza, and R. Lenski, *Tests of ecological mechanisms promoting the stable coexistence of two bacterial genotypes.* Ecology, 1996. **77**(7): p. 2119-2129.
2. Rainey, P.B. and M. Travisano, *Adaptive radiation in a heterogeneous environment.* 1998. **394**(6688): p. 69-72.
3. Adiba, S., et al., *From Grazing Resistance to Pathogenesis: The Coincidental Evolution of Virulence Factors.* PloS one, 2010. **5**(8): p. e11882.
4. Krimbas, C. and S. Tsakas, *The genetics of Dacus oleae. V. Changes of esterase polymorphism in a natural population following insecticide control-selection or drift?* Evolution, 1971. **25**(3): p. 454-460.
5. Nei, M. and F. Tajima, *Genetic drift and estimation of effective population size.* Genetics, 1981. **98**(3): p. 625.
6. Pollak, E., *A new method for estimating the effective population size from allele frequency changes.* Genetics, 1983. **104**(3): p. 531.
7. Hartl, D. and A. Clark, *Principles of population genetics*. 1997: Sinauer associates Sunderland, MA.
8. Kimura, M. and T. Ohta, *The average number of generations until extinction of an individual mutant gene in a finite population.* Genetics, 1969. **63**(3): p. 701.
9. Elena, S. and R. Lenski, *Evolution experiments with microorganisms: the dynamics and genetic bases of adaptation.* Nature Reviews Genetics, 2003. **4**(6): p. 457-469.
10. Dawkins, R. and J. Krebs, *Arms races between and within species.* Proceedings of the Royal Society of London. Series B, Biological Sciences, 1979. **205**(1161): p. 489-511.
11. Van Valen, L., *A new evolutionary law.* Evolutionary theory, 1973. **1**(1): p. 1-30.
12. Wright, S., *Population structure in evolution.* Proceedings of the American Philosophical Society, 1949. **93**(6): p. 471-478.
13. Wahl, L. and P. Gerrish, *The probability that beneficial mutations are lost in populations with periodic bottlenecks.* Evolution, 2001. **55**(12): p. 2606-2610.
14. Stewart, F. and B. Levin, *Partitioning of resources and the outcome of interspecific competition: a model and some general considerations.* American Naturalist, 1973. **107**(954): p. 171-198.
15. Tilman, D., *Resource competition and community structure*. 1982: Princeton Univ Pr.
16. Dempster, E. *Maintenance of genetic heterogeneity*. in *Cold Spring Harbor Symposia on Quantitative Biology*. 1955: Cold Spring Harbor Laboratory Press.
17. Felsenstein, J., *The theoretical population genetics of variable selection and migration.* Annual Review of Genetics, 1976. **10**(1): p. 253-280.
18. Hardin, G., *The competitive exclusion principle.* Science, 1960. **131**(3409): p. 1292-1297.
19. Dykhuizen, D. and M. Davies, *An experimental model: bacterial specialists and generalists competing in chemostats.* Ecology, 1980. **61**(5): p. 1213-1227.
20. MacLean, R., A. Dickson, and G. Bell, *Resource competition and adaptive radiation in a microbial microcosm.* Ecology Letters, 2005. **8**(1): p. 38-46.
21. Hall, A. and N. Colegrave, *How does resource supply affect evolutionary diversification?* Proceedings of the Royal Society B: Biological Sciences, 2007. **274**(1606): p. 73.
22. Haldane, J., *S. 1932. The causes of evolution*, Harper, New York.
23. Ayala, F. and C. Campbell, *Frequency-dependent selection.* Annual Review of Ecology and Systematics, 1974. **5**(1): p. 115-138.
24. Elena, S. and R. Lenski, *Long-term experimental evolution in Escherichia coli. VII. Mechanisms maintaining genetic variability within populations.* Evolution, 1997. **51**(4): p. 1058-1067.
25. Haubold, B., et al., *Detecting linkage disequilibrium in bacterial populations.* Genetics, 1998. **150**(4): p. 1341.

26. Levin, B., J. Antonovics, and H. Sharma, *Frequency-Dependent Selection in Bacterial Populations [and Discussion].* Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences, 1988: p. 459-472.
27. Rainey, P. and M. Travisano, *Adaptive radiation in a heterogeneous environment.* Trans. R. Soc. B, 1977. **280**: p. 29-101.
28. Abrams, P., *The competitive exclusion principle: other views and a reply.* Trends in Ecology & Evolution, 1986. **1**(5): p. 131-132.
29. Chesson, P. and N. Huntly, *The roles of harsh and fluctuating conditions in the dynamics of ecological communities.* The American Naturalist, 1997. **150**(5): p. 519-553.
30. Murdoch, W. and J. Bence, *General predators and unstable prey populations.* Predation: direct and indirect impacts on aquatic communities, 1987: p. 17–30.
31. Pacala, S. and D. Tilman, *Limiting similarity in mechanistic and spatial models of plant competition in heterogeneous environments.* American Naturalist, 1994. **143**(2): p. 222-257.

Demographic stochasticity effect on neutral allele fixation

3. DEMOGRAPHIC STOCHASTICITY EFFECT ON NEUTRAL ALLELE FIXATION

To understand how stochastic (genetic drift) and deterministic (natural selection) factors interact to shape allele frequency variations in a population, we shifted to focus our work on stochastic events. We will include natural selection in future work.

In natural environments, populations frequently change in size, fluctuating under the influence of factors including a biotic factor (temperature, physical conditions), resource availability, competition and predation. These variations affect the maintenance of mutations and then the rate of adaptation of a population. In a Moran model[1], the probability of fixation of a new neutral allele is $1/N$ for a haploid population of constant size N. In a population of finite size, the change in frequency of a mutant allele depends only on random sampling from generation to generation. The random sampling results in chance changes in allele frequency (random genetic drift).

This study examines changes in the probability of fixation and time to fixation under several scenarios including stochasticity with respect to demography and time.

A first scenario we considered was one in which populations experience fluctuations in size with events (birth or death) that occur either at constant or stochastic intervals.

Another scenario is a population that experiences cycles of pure births, then a decline in population size.

The last scenario we considered is an alternation of population increases and declines with stochastic birth and death

For all these scenarios, we studied the effects of stochasticity with respect to demography and in time on neutral allele fixation probability and time to fixation.

DEMOGRAPHIC STOCHASTICITY EFFECT ON NEUTRAL ALLELE FIXATION

Sandrine Adiba[1,2*], Minus van Baalen[1,2] and Frantz Depaulis[2]

[1]Laboratoire d'Ecologie, CNRS UMR7625, Université Pierre et Marie Curie, Paris Universitas, 7 quai St-Bernard, 75252 Paris, France

[2]Laboratoire d'Ecologie, CNRS UMR7625, Ecole Normale Supérieure
46 rue d'Ulm, 75230 Paris cedex 05, France

Work in progress…

Abstract

Most of the models developed in population genetics consider either a constant population size or simple growth and decline scenarios. However, natural environments change continuously, thereby affecting population size and diversity. Such fluctuations in population size may be purely stochastic and should affect whether a new mutant can invade a resident population.

We studied an individual-based population genetics model, for a haploid population, with overlapping generations over a continuous interval of time. We examined how such stochasticity in demography can alter mutant fixation probability and time to fixation. We developed three models: (i) a population stochastically fluctuating around a fixed value, (ii) repeated alternation of growth by births and bottleneck by deaths, (iii) the repeated alternation of increase and decline of population size both with stochasticity.

We used simulations to show that demographic stochasticity alters the fixation probability and the time to fixation compared to a population of constant size.

Key words: Allele fixation probability, time to fixation, stochasticity in demography, bottleneck

Introduction

Theoretical population genetics is concerned with the understanding of how stochastic and deterministic factors interact to shape the allele frequency in a population. In this paper, we focus on stochastic factors, which result from demographic changes and which may be due to fluctuating environmental conditions. The estimation of the fixation probability of a new allele and its time to fixation have been greatly developments since the outset of mathematical population genetics [2-4]. Assessing the relative characteristic of fixation probability is of primary importance, since it influences the rate at which a population could adapt to a changing environment and the level of diversity, and is therefore essential to most models of adaptation. Such models require a number of assumptions which are often not relevant in natural populations. For simplicity, previous studies assumed fixed population sizes: birth events are associated with death events. However, in natural environment, births and deaths are independent and the population size varies stochastically. Several studies [5-7] involving changes in population size, presumed deterministic growth between bottlenecks. These models treated only extinction due to the bottleneck but beneficial mutations may also be lost due to stochasticity during growth or decline in population size [7, 8]. More importantly, neutral models assumed discrete generation times, implicitly assuming fixed generation times with Poisson distributed numbers of offspring [9]. Parson [10] developed a model that allows for stochastic population sizes with density dependence that is therefore more realistic than the standard Wright-Fisher formulation [4, 11, 12]. Some studies [10, 13, 14]} have demonstrated that when the population size varies stochastically, the probability of fixation differs from that of fixed or deterministic population size models as in the Wright-Fisher model [4, 15] in discrete time or the Moran model [16] in continuous time. Nevertheless, they showed that for neutral types with the same birth and death rates, classical expressions hold with the total population number replaced by the carrying capacity. They demonstrated that under strict neutrality, classical results remain valid even when the assumption of fixed population size is relaxed.

Numbers of models [13, 14, 17, 18] have highlighted the extreme sensitivity of fixation probabilities to all these different assumptions.

Some of the adaptations that interest us most in this paper are adaptations to (i) abiotic factors, such as environmental changes, colonization (founder event), and antibiotic resistance and (ii) biotic factors such as prey/predator or host/parasite interactions. All these factors may involve changes in the biological environment that are severe enough to cause a big reduction in population size. Bottleneck may be due to predators eating preys resulting in a sharp decrease in population size; or to virulent parasites infecting some hosts, producing a high lethality in the population. This reduction in size is often purely random, such as in evolution experiments in which serial and regular transfers are necessary to maintain populations over several generations, to add fresh medium, and to maximize evolution by increasing the number of generations. Such bottlenecks are generally modeled by a binomial sampling, and no stochastic event is taken into account. In *natura*, a decrease in population size could be a purely stochastic death. To compare fixation probability and time to fixation in a stochastic and fluctuating demography, we developed three models. Allele fixation probability and time to fixation were compared to our model of reference: the Moran model with constant population size.

The first model is a stochastic birth and death model, implying no trend for size variation (fluctuating). This model included no growth or decline in population size. The second model is a pure birth process followed by a bottleneck with a pure death process. The bottleneck mimics natural periodic variations (e.g. seasonal). The third model includes a growth in population size by birth and death process followed by a decline of population size with a stochastic birth and death process.

We separated the model to include (a) the time (in generation) for each event was fixed in order to compare model with fixed generation time versus stochastic time, (b) stochastic time (following an exponential law). We compared our fixation probability that we estimated using simulations to that of Moran model simulations and approximations derived from branching processes.

Our model is a generalization of the Moran model [16], while relaxing the assumption of fixed population size by making birth and death independent, thereby allowing the population size to vary stochastically.

The models

We used population genetics haploid models with two types, mutants and residents, representing individuals carrying different neutral alleles respectively. These individual based models were in continuous time with overlapping generations.

We introduced a single mutant (p_0=1/N) in the resident population. This mean that for a population size of 100, p_0=0.01 and for a population size of 1,000, p_0=0.001 therefore p_0 decreases with the increased population size.

The Moran model was chosen as referenced. This model is a stochastic process that describes a finite population of constant size. It proceeds by a series of events (equivalent to 1/N generations), each of which characterized by the death of a random individual and replaced by a newborn of allelic type randomly chosen from those of extant putative parents. Classical results and approximations are known for the fixation probability and time to fixation from the branching process (Haldane), and will be used as reference results [4, 12, 19, 20].

For all models developed, we separated the fixation probability and the time to fixation according to whether the population survives or its size went to one. In the last case, the fixation or extinction had occurred and the population had a chance to go to extinction. Furthermore, we introduced a carrying capacity to not have an explosion of the population size because of stochasticity.

Models 1 involve stochastic birth and death processes. The model 1a (stochasticity in demography) proceeds by a serie of events, each (equivalent to 1/N generations) characterized by the birth or the death of a random individual with an allelic type randomly chosen from those of extent putative parents. Thus, population size fluctuates, but its mean value is constant. Model 1b (stochasticity with respect to demography and time follows the same mechanism than model 1a, except that the time of event is not constant and follows an exponential law (events follows a Poisson process).

Models 2 proceeds by a serie of events that are either equivalent to 1/N generations (model 2a) or following a Poisson process (Model 2b). The population undergoes cyclic changes in population size, with a pure birth process followed every T_B units of time by a pure death process. Allelic types are randomly chosen among extend putative parents. Since the bottleneck frequency (1/T_B) is constant

and the birth rate not equal to one, the bottleneck occurs each time for a given population size N_{max}, different at each bottleneck. Thus the magnitude of the bottleneck was not constant.

Model 3 (stochasticity in growth and decline in population size) consists of a serie of events each of which is equivalent to 1/N generation (Model 3a) or that follows a Poisson process (Model 3b). The growth phase is characterized by the birth (with the probability $b/(b+d)$) or death (with the probability $d/(b+d)$) with $b>d$ of a random individual of an allelic type randomly chosen among extent putative parents. The death phase follows the same process with $d>b$. The frequency of the bottleneck ($1/T_B$) is constant, thus the population size at the bottleneck was not always the same.

Fixation probability, time to fixation and effective population size

The fixation probability of a single mutant in the referenced Moran model (as well as in the classical Wright-Fisher model) is $1/N$. We compared this result to our simulations in order to study the effects of stochasticity on demography and time.

For a selectively neutral allele, the probability of ultimate fixation is equal to its initial allele frequency. Using the analysis of the Kolmogorov backward equation [20], under a haploid Moran model, the expected time for a neutral allele with initial frequency p to go to fixation (given that it is eventually fixed) or to loss (given that it is eventually lost) are as follow:

Mean fixation time (in generations)

$$\bar{t}_1(p) = -N\left[\frac{(1-p)}{p}\right]\ln(1-p) \qquad (1)$$

Mean extinction time (in generations)

$$\bar{t}_0(p) = -N\left[\frac{p}{(1-p)}\right]\ln(p) \qquad (2)$$

When $p=1/N$, that is, when a new neutral mutation has just occurred and there is only one copy in the population, the time to fixation (given that it is eventually fixed) is approximately N generations. On the other hand, the time it takes for a new neutral mutation to eventually be lost (given that it is eventually lost) is $1-1/N$ generations.

We compared these results of our simulation with these approximations.

Roughly speaking, the effective population size of a population is the number of individuals in a theoretically ideal population (following a neutral constant size model, generally the Wright-Fisher model) that displays the same magnitude of random generic drift as the focus population.

When confronted with demographic variations, the effective size N_e may be approximated by the harmonic mean of the census sizes over time.

$1/Ne=(1/t)x(1/N_0+1/N_1+....+1/N_{t-1})$ (3) [1]

We used this formula in our simulations to estimate the demographic effective population size in each model.

Results

Population dynamics

In the model 1, the total population size varies stochastically around a typical value (Figure IV-12a). In the model 2, the growth is governed by births only, and bottlenecks occur at regular frequency with a death process (Figure IV-12b). For model 2, growth was simulated with a birth rate higher than the death rate, which allows the population to fluctuate around an average growth trend. The bottleneck was similarly simulated with a death rate greater than the birth rate, leading the population to fluctuate and decline (Figure IV-12c).

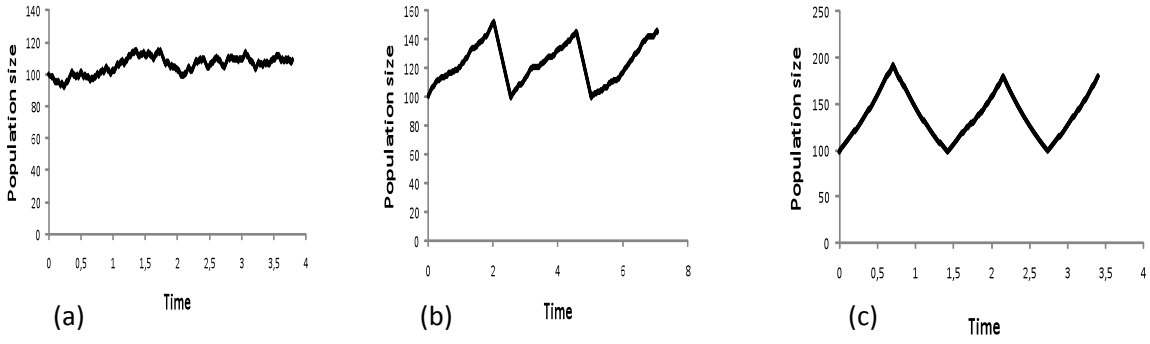


Figure IV-12: Population dynamics

(a) Model 1: the population size fluctuates around a mean value

(b) Model 2: the population size increases with a regular birth process and declines with a death process

(c) Model 3: the population size increases with a birth-death process and decrease with the same process

Effective population size

Due to stochasticity in the population sizes, the model 1 yielded a lower effective population size compared to the Moran model (Figure IV-13) except for small values. In the model 2 and 3, we found a lower effective population size compared to the neutral model, due to bottlenecks and the stochasticity of events.

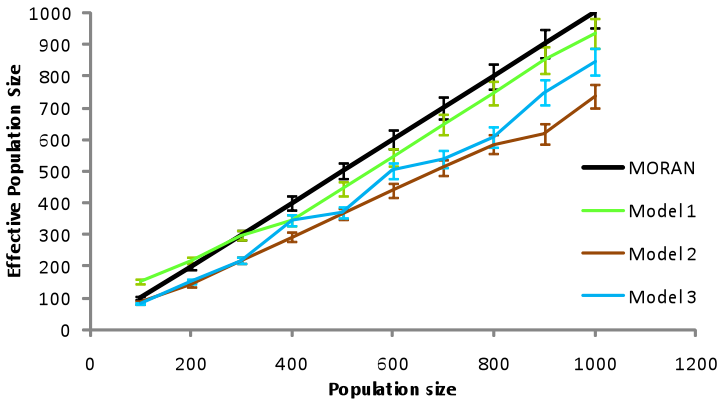


Figure IV-13: Effective population size
The effective population size was estimated with the harmonic mean for the three models.

Fixation probabilities

The fixations probabilities were assessed by simulations (500 simulations). We compared our model with the referenced Moran model with constant population size.

The model 1a (stochasticity in demography) showed higher fixation probability compared to the Moran model (Figure IV-14a). Demographic stochasticity allowed the population size to fluctuate around a mean value. Nevertheless, due to this of the population size stochasticity, the population can go to extinction (model 1 extinction). In this case, the mutant allele displayed a greater chance to become fixed.

For model 2 (growth and bottleneck by death), the fixation probability was lower than in a population of constant population size (Moran model) (Figure IV-14 a and b) excepted for the population sizes of 100 and 300, for which the fixation probabilities were greater.

The model 3a (stochasticity in population growth and decline), showed greater fixation probabilities compared to the Moran model when the population sizes

were small (<400) and fixation probabilities were lower when the population size was larger (Figure IV-14a).

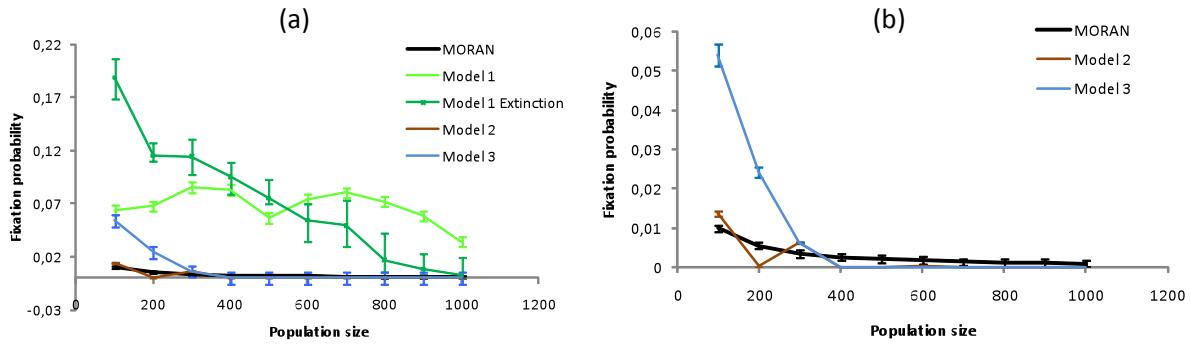


Figure IV-14: Fixation probabilities

(a) Fixation probability for the three models developed and the Moran model

(b) Comparison of the fixation probability between the Moran model and model 2 and 3.

Time to fixation

For the first model (stochasticity in demography), the time to fixation deviates from the simple neutral prediction (Moran model and approximation) (Figure IV-15a). Model 1 demonstrated shorter time to fixation compared to the Moran model, and this time was shorter when the population goes to the extinction due to the stochasticity (Figure IV-15 a and b). Model 1 (Figure IV-15b) yielded shorter fixation times when the events were stochastic (model 1b) and the population size was small.

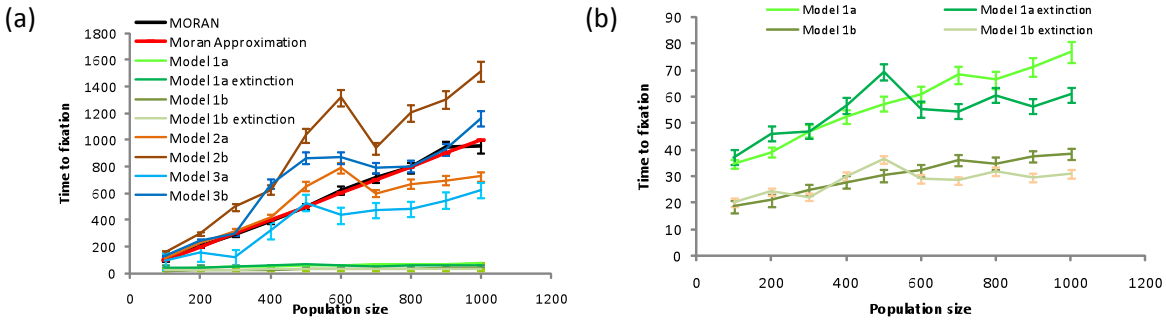


Figure IV-15: Time to fixation

(a) Time to fixation for all models and Moran approximation

(b) Comparison of the time to fixation for the model 1

When we had no stochasticity in the time of events and pure birth followed by pure death process (model 2a), above a population size 500, the time to fixation did not differ from the Moran model (Figure IV-15a). For the intermediate values

(500 and 600 individuals) and greater values (up to 1,000 individuals), the time to fixation differed from the reference model and was greater and shorter, respectively. With stochasticity in time (Model 2b), time to fixation was greater.

When we added stochasticity in birth and death processes (model 3a), time to fixation did not differ from the reference model for small population sizes (lower than 500 individuals) except for 300 individuals, and was lower for large population sizes. Adding stochasticity to the time of events (model 3b), time to fixation was greater for intermediate values of population sizes (between 400 and 700 individuals) than in the Moran model of constant population size (Figure IV-15a).

Discussion

Our study focused on the effect of demographic stochasticity on neutral fixation process. We developed three models that allowed for stochastic population sizes, with or without bottlenecks. We studied fixation probabilities and time to fixation for the neutral case, with overlapping generations and in continuous time.

This first model tested for the demographic stochasticity. This model simulated a population with independent birth and death processes, the total population size fluctuating stochastically, which caused changes in allele number. The second model simulated a population in growth with pure birth followed by decline with a pure death process. The third model was that of a population in growth with stochasticity (the birth rate exceeding the death rate) and in cyclic decline with deaths (the death rate exceeding the birth rate).

For all these individual based models in continuous time and with overlapping generations, the fixation probability and time to fixation were estimated by simulations (500 runs) and compared with those of the Moran model with constant population sizes. Due to the stochasticity of events and/or bottlenecks, effective population sizes were lower compared to the Moran model (Figure IV-13) except for small values in the model 1. The stochasticity in this model allows population sizes to be greater than with constant population size. In small populations, the stochasticity may have a great impact on population sizes.

Model 1 of demographic stochasticity showed that when the population size varies stochastically around a mean value, the fixation probability and the time to

fixation differ qualitatively from that of a fixed population size model (Moran model) [16]. Due to stochasticity, a new mutant arising in a resident population had a better chance to fixe in a resident population compared to a population of constant size. The time to fixation was also reduced. When the new mutant went to fixation, it was able tp invade the resident population in a shorter time, due only to chance. Parson and Quince [10, 13] recently developed a model in which birth and death were independent and the population size varied stochastically around a mean value governed by density dependence. They showed that when the population size varies stochastically, the probability of fixation is different from that of a constant population size. The deviation was ascribed to two separated phenomena. The first was an advantage in growing populations for the type that had the greatest birth rate. The second was a disadvantage for this type, due to larger fluctuations in population size. This disadvantage of the type with higher birth rate was an example of "r and K selection" [21, 22]. However, they found that the expected absorption and fixation times for an arbitrary number of neutral types, all of which have the same birth and death rates, classical results with the total population number replaced by the carrying capacity hold. This demonstrates that under strictly neutrality, the classical results remain valid even when the assumption of fixed population size is relaxed. This result is due in large part to the fact that the total population will grow rapidly toward the carrying capacity and spend a long period fluctuating around the carrying capacity thereafter. This model was defined as a quasi-neutral model, both types have the same equilibrium population densities (birth rates are proportional to death rates). In this case, the proportion of the type with higher birth rate increased. In our model, birth and death rates are independent and equals, we found greater fixation probability and a shorter time to fixation. Offsprings variance may affect this probability. Lambert (2006) [14] showed that variance in reproduction may affect fixation probabilities. He displayed a mean variance trade-off (reduced growth rate compared to reduced reproduction variance) therefore different alleles may have different trait values but still be neutral to each other. In this case, a small departure from that trade-off was advantageous for those alleles that initially had small variances in reproduction and growth rate rather than high ones. This proves that classical models tend to underestimate fixation probabilities in growing resident population and overestimate then in declining resident population. Another study [23]

showed that strong stochasticity in offspring variance means smaller fixation probability for beneficial alleles and higher fixation probability for deleterious alleles, either reduced growth rate of their increase variance in reproduction. We intend to study our model 1 in more details, in order to confirm our results regarding fixation probabilities and analyze whether this is due to offspring variance or whether in our model the population had a tendency to grow after a long time.

In model 2 and 3, growth and decline stages occured in addition of demographic stochasticity (model 3). Stochasticity in the increase and decrease population size made the probability of fixation greater than with a fixed population size for small population sizes. But above 200 individuals (model 2) or 300 (model 3), the probability of fixation was lower than with a constant population size (Moran model and Figure IV-14). Thus for small population sizes, a new mutant arising in a resident population had a better chance to invade than in a large population size.

Wahl and Gerish [7] studied the probability that a beneficial mutation was ultimately eliminated by periodic bottleneck. In a population with growth between bottlenecks, the probability of fixation drops with *s* (the selective advantage) and mutations having small selective advantage are unlikely to survive. They showed that the number of surviving mutations was reduced by the bottleneck effect, which is in agreement with our results. Mutations may also be lost during the growth phase due to drift [24] or competition [25]. In our model, we found that demographic stochasticity did not change the probability of fixation when bottleneck occurred (model 2 and 3) except for small population size (<300), in which case a new mutant had greater chance to survive when stochasticity in growth and death occurred.

Compared to the model 1 (a versus b), the stochasticity of events in model 2 and 3 increased the time to fixation (model 2a versus 2b and 3a versus 3b, Figure IV-15). First, the stochasticity in birth and death (model 2 versus model 3) decrease time to fixation; second, stochasticity in event occurrence increased the time to fixation (model a versus b, Figure IV-15). Thus, combining stochasticity in birth and death processes but maintaining the series of events constant (no stochasticity in time), the time to fixation was shorter.

Our models showed that demographic stochasticity may influence the probability of fixation and the time to fixation compared to a fixed population size model. The results of our model 1 must be confirmed by other analyses.

Literature cited

1. Hartl, D. and A. Clark, *Principles of population genetics*. 1997: Sinauer associates Sunderland, MA.
2. Fisher, R. *A, 1922 On the dominance ratio*. in *Proc. R. Soc. Edinb*.
3. Haldane, J. *S., 1927 A mathematical theory of natural and artificial selection, Part V: Selection and mutation*. in *Proc. Camb. Philos. Soc*.
4. WRIGHT, W., *S., 1931 Evolution in Mendelian populations.* Genetics. **16**: p. 97-159.
5. Kojima, K. and T. Kelleher, *Survival of mutant genes.* American Naturalist, 1962. **96**(891): p. 329-346.
6. Otto, S. and M. Whitlock, *The probability of fixation in populations of changing size.* Genetics, 1997. **146**(2): p. 723.
7. Wahl, L. and P. Gerrish, *The probability that beneficial mutations are lost in populations with periodic bottlenecks.* Evolution, 2001. **55**(12): p. 2606-2610.
8. Lambert, A., *The branching process with logistic growth.* Annals of Applied Probability, 2005: p. 1506-1535.
9. Sibert, A., F. Austerlitz, and E. Heyer, *Wright-Fisher revisited: the case of fertility correlation.* Theoretical Population Biology, 2002. **62**(2): p. 181-197.
10. Parsons, T. and C. Quince, *Fixation in haploid populations exhibiting density dependence I: the non-neutral case.* Theoretical population biology, 2007. **72**(1): p. 121-135.
11. Fisher, R., *The general theory of natural selection.* Denver Publication, New York, 1958.
12. Ewens, W., *Mathematical population genetics: theoretical introduction*. 2004: Springer Verlag.
13. Parsons, T. and C. Quince, *Fixation in haploid populations exhibiting density dependence II: The quasi-neutral case.* Theoretical population biology, 2007. **72**(4): p. 468-479.
14. Lambert, A., *Probability of fixation under weak selection: a branching process unifying approach.* Theoretical population biology, 2006. **69**(4): p. 419-441.
15. Fisher, R., *The genetical theory of natural selection, 272 pp*. 1930, Clar endon Press, Oxford.
16. Moran, P. *Random processes in genetics*. in *Mathematical Proceedings of the Cambridge Philosophical Society*. 1958: Cambridge Univ Press.
17. Hubbarde, J., G. Wild, and L. Wahl, *Fixation Probabilities When Generation Times Are Variable: The Burst Death Model.* Genetics, 2007. **176**(3): p. 1703.
18. Alexander, H. and L. Wahl, *Fixation probabilities depend on life history: fecundity, generation time and survival in a burst-death model.* Evolution, 2008. **62**(7): p. 1600-1609.
19. Kimura, M., *On the probability of fixation of mutant genes in a population.* Genetics, 1962. **47**(6): p. 713.
20. Kimura, M. and T. Ohta, *The average number of generations until extinction of an individual mutant gene in a finite population.* Genetics, 1969. **63**(3): p. 701.
21. MacArthur, R. and E. Wilson, *The theory of island biogeography, 203 pp.* New Jersey: Princeton, 1967.
22. Pianka, E., *r and K Selection or b and d Selection?* American naturalist, 1972. **106**(951): p. 581-588.
23. Gillespie, J., *Natural selection for variances in offspring numbers: a new evolutionary principle.* American Naturalist, 1977. **111**(981): p. 1010-1014.
24. Kimura, M. and T. Ohta, *Probability of fixation of a mutant gene in a finite population when selective advantage decreases with time.* Genetics, 1970. **65**(3): p. 525.
25. Gerrish, P. and R. Lenski, *The fate of competing beneficial mutations in an asexual population.* Genetica, 1998. **102**: p. 127-144.

Chapter 5

V. Discussion

A central goal of this thesis was to understand the role of natural selection and genetic drift processes in maintaining neutral polymorphism. Theoretical and empirical studies have shown that antagonistic interactions such as prey/predator or host/parasite interactions, can maintain biological diversity. Our review highlighted experimental and theoretical studies related to the role assigned to balancing selection during such ecological interactions in maintaining polymorphism.

The biological system consisting of the bacteria *Escherichia coli* and the social amoeba *Dictyostelium discoideum*, were chosen for various reasons. First, *D. discoideum* is a major predator of several species of bacteria. Therefore, according to the bacterial strain (even from the same species), bacteria could be digested and harmless for *D. discoideum* as *E. coli* K12 MG1655, B/r or REL606 and REL607. Other bacterial strain, such as *E. coli* O157, could resist digestion by amoeba and multiply within them [1]. Then, depending on the bacterial strain, we may observe prey/predator or host/parasite interactions.

For this reason, in our first experiments, we tested for the natural variability of interactions between various *E. coli* strains and *D. discoideum*. Since our amoeba could be cultured in liquid media or on agar plates, we tested for variability under both conditions, analyzing the traits responsible for this variation. We then focus our work on population genetics aspects, by studying various environmental conditions that influence neutral polymorphism.

In the first part of this chapter, we will discuss interactions between bacteria and amoebae. Several bacterial strains [2] are known to resist digestion by amoeba. Environmental pressures may be responsible for bacterial adaptation and survival strategies. Thus, before carrying out our main coevolution experiment, we studied the evolution of bacterial resistance to phagocytosis. This topic is covered in the second part of our discussion.

The collection of *E. coli* strains we had was well characterized in virulence genes. Results regarding interactions between a set of *E. coli* strains and *D. discoideum* revealed a correlation between carrying of virulence genes and the effect on *D. discoideum*. Our experimental system enabled us to test for the coincidental hypothesis of virulence factors. In the third part of this chapter, we will more generally discuss empirical and theoretical studies concerned with the maintenance and evolution of virulence.

Finally, we will discuss the maintenance of neutral polymorphism, including our complementary experimentation and modeling approaches.

1. Interactions between bacteria and amoebae

The various interactions between bacteria and amoebae

The ability of amoebae to feed on Gram-negative bacteria, as well as to harbor potential pathogens such as *Legionella pneumophila* [3, 4], *E. coli O157* [1], and *Pseudomonas aeruginosa* [5] suggest that both amoeba and bacteria are involved in complex interactions that may play important roles in the environment and for human health. We used the amoeba *D. discoideum* and studied its interactions with various *E. coli* strains (laboratory adapted, human commensal and pathogenic strains) in two kinds of environments (liquid and solid).

We found no correlation between virulence genes and their effect on *D. discoideum* growth in liquid media, with agitation. On agar plates (solid media), we found a correlation between virulence genes and the ability of *D. discoideum* to feed on bacteria. Bacteria bearing virulence genes involved in iron uptake metabolism are resistant to the amoeba phagocytosis. Since iron is a vital resource for bacteria that is necessary for many functions (see Chapter IV.1 p.104), this suggests that bacteria may exploit *D. discoideum* and phagocytosis in order to obtain key nutrients.

It is well known that amoebae acts as predator and host for several species of bacteria, and its ability to host bacterial pathogens, such as *Legionella pneumophila* [3, 4], *E. coli* O157 [1], *Pseudomonas aeruginosa* [5], *Vibrio cholerae* [6] has gained particular attention. However, the precise mechanisms associated with bacteria-amoebae interactions are still incompletely understood. Further complexity is added by the fact that amoeba feed on bacteria. Thus, why bacteria interact with amoeba and how pathogenic bacteria maintain survival inside amoeba while non-pathogenic are killed remains unclear. Using clinical and laboratory isolates of *E. coli*, we demonstrated that the outcome of *E. coli* interactions with *D. discoideum* varies, depending (i) according to environmental conditions, (ii) in the presence of virulence genes of *E. coli* strains.

The interactions between amoebae and bacteria are not only complex but also diverse. It was shown that amoebae may be predators and hosts. Four decades ago

Jeon [7] demonstrated that the interaction between an amoeba (*Amoeba proteus* strain D) and the pathogenic so-called X-bacteria could relatively rapidly evolve towards an obligate dependency [8]. When X bacteria are phagocytised by *A. proteus*, some survive in phagolysosomes and multiply. After the order of 200 host generations, the X bacteria and the amoebae establish a stable symbiotic relationship that results in mutual dependence for survival – since evolved amoeba (called xD) die if their bacterial symbionts are removed [9, 8]. Obligate mutual dependence evolved within a short period of evolutionary time. The work of Jeon and his collaborators has paved the way for recreating the process of symbiont integration in vitro. However, one major difficulty with Jeon's system is that the bacteria cannot be cultivated in vitro. Furthermore, the origin of X bacteria is not known, and genome-wide information on both the host and the symbiont are limited.

The "dangerous liaisons" between bacteria and amoeba

Using our experimental system, we were able to show that the interactions between bacteria and amoebae had many facets. The balance between conflict and interest may be delicate. Bacteria can find resources inside amoeba but take the risk of being digested. Bacteria may exploit amoeba phagocytosis when the environment is scarce and resist digestion by amoeba. As soon as bacteria and amoeba have sufficient common interest they reveal sustainable interactions. These interactions may in fact be a mixture of antagonistic and mutualistic aspects. Amoeba feed on bacteria, but can also protect them when the environment is hard. Jeon [7] demonstrated that the transition from parasite to symbiont could be relatively rapid. The X-bacteria may confer benefits to the amoeba host. The difficulty in understanding this interaction is in determining how such common interests arise. As long as bacteria and amoeba demonstrate sufficient common interest, they can live together and cooperate, but if the conditions change, private interests prevail and a struggle begins between the two partners [10].

For host/parasite interactions, the common interest is quite clear: survival of the host. The parasites may modulate their virulence, whereas the host may modify its energetic and resource investment for use in fighting against the parasite. Common interest is less intuitive for prey/predator interaction. The fact that

amoebae are able to engulf a great number of bacteria may prevent them against more pathogenic and virulent bacteria. In return, bacteria may not resist to phagocytosis in exchange for protection, nutrients or a vector.

D. discoideum an alternative host model system

In Chapter IV.1, p.104, we reported an experiment confronting a collection of 31 *E. coli* strains to grazing by *D. discoideum* amoeba. We found a correlation between the virulence genes carrying by *E. coli* strains and the amoeba grazing ability.

By definition, the virulence of a given bacterial strain is measured by confrontating it with a host. Mice are often used to assess the virulence of bacterial pathogens, based on the fact that murine immune defenses are similar to those of the human body. However, these experiments are difficult to carry out, expensive, and ethically problematic, since they inflict significant suffering on infected animals. Alternative hosts, such as the nematode *Caenorhabitis elegans* and the insect *Drosophila melanogaster* or even unicellular *Acanthamoeba castellanii* or *Dictyostelium discoideum*, have proven useful to study bacterial virulence [11, 12]. Due to the universality of virulence factors implicated in the infectious process and based on the observation that many pathogens display low species specificity, alternative hosts models are considered as relevant. Furthermore, these biological models are naturally confronted with these pathogens in their natural environments and our paper (Chapter IV.1 p.104, [13]) demonstrates that many virulence factors can evolve to fight these natural predators [14]. The unicellular amoeba *D. discoideum* may be used to carry out a very simple assessment of the bacterial virulence of clinical isolates. When plated *D. discoideum* with non-pathogenic bacteria, we observed plaque lysis, indicating that *D. discoideum* fed on this bacterial strain, whereas with a pathogenic bacteria, no plaque lysis was observed. In this case, we showed that *D. discoideum* died while bacteria may be alive. Thus, bacterial virulence can be assessed from the ability of *D. discoideum* to eat bacteria as previously shown for *Klebsellia pneumonae* [15] or *Pseudomonas aeruginosa* [5, 16]. These previous studies also showed a correlation between virulence in the *D. discoideum* host model and in a mouse infection model.

D. discoideum may be used as a model eukaryotic cell that in aspects to its cell biology and interaction with microbes mimics a mammalian macrophage. Several

environmental pathogenic bacteria including *Legionella pneumophila* [4], *Mycobacterium marinum* and *Pseudomonas aeruginosa* [17] resist *D. discoideum* predation by producing factors that either kill amoeba or allow successful intracellular survival and multiplication. In these cases, the same virulence mechanisms that operate against mammalian cells have also been implicated in resistance to *D. discoideum* predation. For example, it has been shown that *Pseudomonas aeruginosa* can kill *D. discoideum* by using type III secretion system (T3SS) to deliver the cytotoxic ExoU effector protein into target cells [5] and that this T3SS is also essential for mammalian virulence [18].

All these studies indicate that in various hosts, bacterial pathogenicity relies on a largely overlapping set of bacterial virulence genes [19]. *D. discoideum* may be used to assess the virulence of many bacterial species including Klebsellia [15], Aeromonas [20] or Pseudomonas [5, 16]. Klebsellia virulence against *D. discoideum* relies on its ability to resist intracellular killing [15]. Pseudomonas virulence is due to secreted toxins that can kill *D. discoideum* cells [5, 16].

A consequence of the finding that amoeba may host bacterial pathogens under harsh environmental conditions may help to explain how bacterial pathogens are transmitted to susceptible hosts, thus offering a potential route of entry into human body. The fact that *D. discoideum* resembles human macrophages in many ways, particularly in its phagocytic activity and cell surface receptor and that they exhibit parallel mechanisms in their interactions with various *E. coli* strains (clinical isolates) may provide an alternative model for studying *E. coli* pathogenesis and to understand its immune evasion mechanisms. Amoeba may act as a bacterial predator, or as a reservoir or 'Trojan Horse' [2, 14] for bacteria with environmental and clinical implications.

Advantages of non-mammalian hosts include their ease of cultivation and short generation time, as well as the availability to elaborate genetic, biochemical and cell biological tools. Genetic tools available for *D. discoideum* include plasmids which allow the constitutive or inducible expression of genes, GFP (green fluorescent protein) fusion constructs, targeted deletions of multiple genes, and restriction enzyme mediated integration mutagenesis. These features render *D. discoideum* a powerful protozoan model for studying host factors involved in cellular aspects of pathogen–host interactions [4, 17]. A possible disadvantage of *D. discoideum* is that amoebae do not survive temperatures greater than 27°C, which

is not compatible with an induction temperature of 37°C required for virulence gene expression of many pathogens.

2. The evolution of bacterial resistance to phagocytosis

Environmental pressures

In natural environment, bacteria always live in relationships with other organisms and are particularly threatened by bacteriovorous predators, such as free-living protozoa or nematodes. The harsh environment in which bacteria live is characterized by a constant competition for nutrients and the menace of protozoan and metazoan predators. Over the course of this long co-evolutionary history and under these evolutionary pressures, some bacteria have developed sophisticated defense mechanisms, including the secretion of toxins, or the capacity to avoid lysosomal killing, as well as to replicate intracellularly within protozoa. Protozoa are primordial phagocytes and share many features with mammalian phagocytes, particularly macrophages. In the course of these interactions with protozoa, bacteria may adapt and then become resistant to bactericidal mammalian macrophages and thereby cause human disease. With the selective pressure the protozoa exert, they not only select for virulence traits that allow survival and intracellular replication within macrophages, but also can serve as protective reservoir when the environment is not adequate, facilitating the transmission of infectious microbes to human [2, 21-23].

In our experiments on the evolution of bacterial resistance to phagocytosis (Chapter III, p.96) , we mixed a set of bacterial strains with *D. discoideum* and let them coevolved for three months. We then measured the number of intracellular bacteria (i) at different times and (ii) confronted them to evolved versus ancestral amoeba. We found that when *E. coli* and *D. discoideum* coevolved, the number of intracellular bacteria increased with time. This increase was observed especially for the B2 bacterial phylogenetic group. It was shown that bacteria belonging to the B2 phylogenetic group were more virulent than bacteria belonging to other phylogenetic groups [24]. When we compared evolved versus ancestral *D. discoideum* for each bacteria strains used, the evolved pathogenic strain was more invasive with the ancestral *D. discoideum* than with the evolved one. This means that with time, the pathogenic strain became more resistant to the amoeba

digestion. For the other evolved bacteria strains (commensal and mutant), the pattern was discontinuous over generation time, and at the last time step we observed that bacteria were more invasive with evolved *D. discoideum* than with ancestral ones. The carriage of virulence traits permits bacteria to become more invasive to ancestral *D. discoideum*. We now have to compare ancestral and evolved bacterial strains with ancestral amoebae in order to confirm the evolution of bacterial resistance to phagocytosis and to test more bacterial strains and study various traits responsible for this evolution.

Survival strategies

Our work showed that bacteria were more able to survive amoeba grazing when they carried virulence traits. Thus, predation by protozoa might select for specific antipredator adaptations in bacteria. These adaptations could act before ingestion (extracellular) or afterward, inside food vacuoles (intracellular).

Extracellular adaptations include the acquisition of oversized morphology, increased bacterial motility [25], cell-cell contact discrimination [26], biofilm formation, and toxin release [27] (see Chapter II, Figure II.8, p.78). Inside food vacuole, the bacteria have to resist digestion, particularly to face enzymatic degradation. Intracellular adaptation could be render more drastic by the production of chemical products (toxins) [27].

Bacteria that are internalized in amoebae encompass three main groups: those which multiply and cause lysis of amoebae cells such as *Legionella* and *Listeria*; those which multiply without causing cell lysis (*Vibrio cholera*), and those which survive without multiplication (certain coliforms and mycobacteria) [14].

Implications for human health

Among Gram-negative bacteria, *E. coli* is the leading cause of several diseases. We are only beginning to understand the mechanisms associated with *E.coli* survival and multiplication in the blood-stream. It has been shown that the ability to growth or survive intracellularly in amoeba allowed some species to enable them to survive and replicate within animal macrophages and to escape host immunity. *Legionella pneumophila* survival in both natural and man-made environments (hospital water systems and moist sanitary areas) is related to its

ability to survive in the amoebal hosts. There is evidence that the ability of some bacterial species to cause human disease may be a consequence of evolutionary selection for intracellular survival and replication within protozoa [14, 28]. Amoebae act as a host for several bacteria and their ability to host bacteria pathogens under harsh environmental conditions may help transmission to susceptible hosts and offer a potential route of entry into the human body.

Amoeba may then act as a reservoir, vector or 'Trojan Horse' for bacteria. This finding has environmental, evolutionary and clinical implications. Thom [29] showed that the association between *Vibrio cholera* and amoeba trophozoites might favor the microorganisms' survival in an aquatic environment.

It has been shown that after a coculture with protozoa, bacteria (Coliforms: *Escherichia coli, Citrobacter freudeii, Enterobacter agglomerans, Enterobacter cloacae, Klebsiella pneumonia, Klebsiella oxytoca* and bacterial pathogens *Salmonella typhimirium, Yersinia enterocolitica, Shigella somnei, Legionella gormanii and Campylobacter jejuni*) were more resistant to free chlorine [30]. This raises questions of chlorine disinfection of coliform and pathogenic bacteria within water distribution systems and indicates important clinical implications. Protozoa commonly found in natural environment (lakes, rivers) as well as in drinking-water reservoirs may act as reservoirs for bacteria. The resistance of amoebal cysts to extremes of temperature [31, 32] and to the effects of biocides [33] may contribute to difficulties in eradicating some bacterial species, such as Legionella from contaminated water systems using conventional disinfectant procedures. It was thus concluded that amoebic intracellular environment was responsible for the bacterial protection against chlorine disinfection.

The fact that amoebae may act as an environmental reservoirs for bacteria is well established. There is growing evidence that these organisms play a major role in the survival and virulence trait of some pathogenic species and through this requires further investigations.

3. The maintenance and the evolution of virulence

A number of hypotheses have been proposed to explain the maintenance and the evolution of virulence for obligate pathogens. However, for facultative pathogens such as *E.coli*, the evolution of virulence is poorly understood. Many observations are consistent with the conventional wisdom [34] concerning host-parasite coevolution. In this theory, parasite-host coevolution is in the direction of commensalim or mutualism; parasite should not harm the host needed for their survival and transmission from one host to another. The virulence is thus explained here as demonstrating new or recent associations among parasites and hosts.

Some observations were consistent with this conventional wisdom such as for emerging diseases (Legionaires' diseases, lyme diseases), but for some virulent pathogens like Shigella and *Neisseria gonorrhoeae*, humans appear to be unique or dominant host and for some microparasitic diseases like malaria and tuberculosis, there is evidence for a long history with humans. However, for conventional wisdom, apparently a 'long' time may still not be enough for these microparasites to evolve in the direction of commensalism.

In the early 1980s, taking into account epidemiology and the ecology of the microparasite and in particular the relationship between virulence and transmission (coupling these two parameters positively) in contrast with conventional wisdom, it was proposed that natural selection could favor and maintain some level of virulence. Assuming such a trade-off, Paul Ewald [35] proposed scenarios for the evolution of virulence of some microparasites, including those responsible for cholera, influenza, dysentery and AIDS [35, 36]. Nevertheless, some pathogenic microbes do not require such a positive association between virulence and transmission. For example, some *E. coli* strains that are commensal of the gut of warm-blooded animals and humans, can cause severe diseases (septicemia, meningitis) although hosts represent evolutionary dead-end for these microbes since they rapidly lead to host death with poor inter-host transmission. For these facultative pathogens, the evolution of virulence remains unclear.

Our work (Chapter IV, p.104) tested the coincidental evolution hypothesis. According to this hypothesis, factors responsible for bacterial virulence may be

favored and maintained by natural selection for other ecological factors such as survival in natural environment rather than to provide the parasite an advantage within a host or its transmission to other hosts. By subjecting a collection of 31 *E.coli* strains well characterized in virulence to amoeba grazing, we found that bacteria strains carrying virulence factors were resistant to the amoeba grazing. These virulence factors may have been favored by selection in external environmental conditions. Further investigations are needed to test whether these virulence factors provide an advantage to the bacteria expressing them.

Our experimental work proposed basic protocols for experimental studies of bacterial epidemiology, and it would be relevant to test theories related to the evolution of the maintenance of virulence in human pathogens.

4. The maintenance of neutral polymorphism: experimental and theoretical approaches

During this thesis, evolution experiments were carried out to study the maintenance of polymorphism under four environmental conditions: homogeneous versus heterogeneous environment and with or without biotic factor. We followed temporal allele frequency variations over roughly 300 generations. Cultures were maintained by serial transfer introducing fresh media and increasing the number of generations. We found the maintenance of polymorphism for one replicate out of five in the heterogeneous environment with biotic factor. Our experiments suggest that genetic drift played a major role compared to natural selection (Chapter IV, p.134) .

The effect of demographic stochasticity and our serial transfer regime may have an impact on the maintenance of polymorphism. Gerrish and Wahl [37] showed that the bottleneck regime have an impact on allele fixation probability. Furthermore, previous studies [38-40] have demonstrated the influence of stochastic variation in population size on allele fixation probability and on time to fixation. In our experiments, like in natural environment, population sizes vary, fluctuating stochastically, since the environment is not fixed but variable.

Thus, to study first this effect of demographic stochasticity and bottleneck regime, we developed three models (Chapter IV, p.162). Selection will be covered

in future work. The first model addressed the effect of demographic stochasticity by allowing the population size to fluctuate around a mean value.

The second model focus on the effect of bottlenecks, and the third model combined demographic stochasticity and the bottleneck regime. Using simulations, we estimated fixation probability and time to fixation of a single neutral mutant arising in a resident population. We found that the fixation probability was greater in model 1 with a fluctuating in population size compared to a model of constant population size, and that the time to fixation was shorter. The bottleneck (model 2) decreases the probability of fixation, which was lower than in a population of constant size.

When we combined demographic stochasticity with a bottleneck (model 3), the fixation probability was greater compared to a population of constant size for small population sizes but decreased rapidly with the population size. For small population sizes, stochasticity acts in favor of the mutant.

The effective population sizes were lower than in a model of constant size. Stochasticity in demography (model 2 versus 3) increased the effective population size. These results have to be confirmed in future studies.

Combining our experimental results with our theoretical has helped to interpret stochastic forces acting in our experimental system. Our model demonstrated that only demographic stochasticity and a serial transfer regime could be responsible of our lost of polymorphism. Bottlenecks had a big effect on temporal allele frequency variations, thus on polymorphism. Our models must be further developed in order to address questions of how the two evolutionary forces, natural selection and genetic drift interact. We will compare this model to our experimental data.

Literature cited

1. Barker, J., T. Humphrey, and M. Brown, *Survival of Escherichia coli O157 in a soil protozoan: implications for disease.* FEMS microbiology letters, 1999. **173**(2): p. 291.
2. Greub, G. and D. Raoult, *Microorganisms resistant to free-living amoebae.* Clinical Microbiology Reviews, 2004. **17**(2): p. 413-433.
3. Hägele, S., et al., *Dictyostelium discoideum: a new host model system for intracellular pathogens of the genus Legionella.* Cellular Microbiology, 2000. **2**(2): p. 165-171.
4. Solomon, J. and R. Isberg, *Growth of Legionella pneumophila in Dictyostelium discoideum: a novel system for genetic analysis of host–pathogen interactions.* Trends in Microbiology, 2000. **8**(10): p. 478-480.
5. Pukatzki, S., R. Kessin, and J. Mekalanos, *The human pathogen Pseudomonas aeruginosa utilizes conserved virulence pathways to infect the social amoeba Dictyostelium discoideum.* Proceedings of the National Academy of Sciences, 2002. **99**(5): p. 3159.
6. Pukatzki, S., et al., *Identification of a conserved bacterial protein secretion system in Vibrio cholerae using the Dictyostelium host model system.* Proceedings of the National Academy of Sciences, 2006. **103**(5): p. 1528-1533.
7. Jeon, K. and M. Jeon, *Endosymbiosis in amoebae: recently established endosymbionts have become required cytoplasmic components.* Journal of Cellular Physiology, 1976. **89**(2): p. 337-344.
8. Jeon, K. and I. Lorch, *Unusual intra-cellular bacterial infection in large, free-living amoebae* 1.* Experimental Cell Research, 1967. **48**(1): p. 236-240.
9. Jeon, K., *Bacterial endosymbiosis in amoebae.* Trends in Cell Biology, 1995. **5**(3): p. 137-140.
10. van Baalen, M. and V. Jansen, *Dangerous liaisons: the ecology of private interest and common good.* Oikos, 2001. **95**(2): p. 211-224.
11. Goy, G., et al., *The Neff strain of Acanthamoeba castellanii, a tool for testing the virulence of Mycobacterium kansasii.* Research in microbiology, 2007. **158**(4): p. 393-397.
12. Kurz, C. and J. Ewbank, *Infection in a dish: high-throughput analyses of bacterial pathogenesis.* Current opinion in microbiology, 2007. **10**(1): p. 10-16.
13. Adiba, S., et al., *From Grazing Resistance to Pathogenesis: The Coincidental Evolution of Virulence Factors.* PloS one, 2010. **5**(8): p. e11882.
14. Barker, J. and M. Brown, *Trojan horses of the microbial world: protozoa and the survival of bacterial pathogens in the environment.* Microbiology, 1994. **140**(6): p. 1253-1259.
15. Benghezal, M., et al., *Specific host genes required for the killing of Klebsiella bacteria by phagocytes.* Cellular Microbiology, 2006. **8**(1): p. 139-148.
16. Cosson, P., et al., *Pseudomonas aeruginosa virulence analyzed in a Dictyostelium discoideum host system.* Journal of bacteriology, 2002. **184**(11): p. 3027-3033.
17. Steinert, M. and K. Heuner, *Dictyostelium as host model for pathogenesis.* Cellular microbiology, 2005. **7**(3): p. 307-314.
18. Vance, R., A. Rietsch, and J. Mekalanos, *Role of the type III secreted exoenzymes S, T, and Y in systemic spread of Pseudomonas aeruginosa PAO1 in vivo.* Infection and immunity, 2005. **73**(3): p. 1706.
19. Hilbi, H., et al., *Environmental predators as models for bacterial pathogenesis.* Environmental Microbiology, 2007. **9**(3): p. 563-575.
20. Froquet, R., et al., *Alternative host model to evaluate Aeromonas virulence.* Applied and Environmental Microbiology, 2007. **73**(17): p. 5657.
21. Rowbotham, T., *Preliminary report on the pathogenicity of Legionella pneumophila for freshwater and soil amoebae.* British Medical Journal, 1980. **33**(12): p. 1179-1183.
22. Brown, M. and J. Barker, *Unexplored reservoirs of pathogenic bacteria: protozoa and biofilms.* Trends in Microbiology, 1999. **7**(1): p. 46-50.

23. Molmeret, M., et al., *Amoebae as training grounds for intracellular bacterial pathogens*, in *Applied and environmental microbiology*. 2005, Am Soc Microbiol. p. 20-28.
24. Picard, B., et al., *The link between phylogeny and virulence in Escherichia coli extraintestinal infection.* Infection and immunity, 1999. **67**(2): p. 546.
25. Matz, C. and K. Jurgens, *High motility reduces grazing mortality of planktonic bacteria.* Applied and environmental microbiology, 2005. **71**(2): p. 921.
26. Wildschutte, H., et al., *Protozoan predation, diversifying selection, and the evolution of antigenic diversity in Salmonella.* Proceedings of the National Academy of Sciences, 2004. **101**(29): p. 10644-10649.
27. Matz, C., et al., *Impact of violacein-producing bacteria on survival and feeding of bacterivorous nanoflagellates.* Applied and environmental microbiology, 2004. **70**(3): p. 1593.
28. Fields, B., *The molecular ecology of legionellae.* Trends in Microbiology, 1996. **4**(7): p. 286-290.
29. Thom, S., D. Warhurst, and B. Drasar, *Association of Vibrio cholerae with fresh water amoebae.* Journal of medical microbiology, 1992. **36**(5): p. 303.
30. King, C., et al., *Survival of coliforms and bacterial pathogens within protozoa during chlorination.* Applied and environmental microbiology, 1988. **54**(12): p. 3023-3033.
31. Chang, S., *Resistance of pathogenic Naegleria to some common physical and chemical agents.* Applied and environmental microbiology, 1978. **35**(2): p. 368.
32. Biddick, C., L. Rogers, and T. Brown, *Viability of pathogenic and nonpathogenic free-living amoebae in long-term storage at a range of temperatures.* Applied and environmental microbiology, 1984. **48**(4): p. 859.
33. De Jonckheere, J. and H. Van de Voorde, *Differences in destruction of cysts of pathogenic and nonpathogenic Naegleria and Acanthamoeba by chlorine.* Applied and environmental microbiology, 1976. **31**(2): p. 294.
34. May, R. and R. Anderson, *Epidemiology and genetics in the coevolution of parasites and hosts.* Proceedings of the Royal Society of London. Series B, Biological Sciences, 1983. **219**(1216): p. 281-313.
35. Ewald, P., *Host-parasite relations, vectors, and the evolution of disease severity.* Annual Review of Ecology and Systematics, 1983. **14**(1): p. 465-485.
36. Ewald, P., *Evolution of infectious disease*. 1994: Oxford University Press, USA.
37. Wahl, L. and P. Gerrish, *The probability that beneficial mutations are lost in populations with periodic bottlenecks.* Evolution, 2001. **55**(12): p. 2606-2610.
38. Parsons, T. and C. Quince, *Fixation in haploid populations exhibiting density dependence I: the non-neutral case.* Theoretical population biology, 2007. **72**(1): p. 121-135.
39. Parsons, T. and C. Quince, *Fixation in haploid populations exhibiting density dependence II: The quasi-neutral case.* Theoretical population biology, 2007. **72**(4): p. 468-479.
40. Lambert, A., *Probability of fixation under weak selection: a branching process unifying approach.* Theoretical population biology, 2006. **69**(4): p. 419-441.

Chapter 6

VI. Conclusion and perspectives

This thesis describes studies of various ecological and evolutionary aspects of our biological system. This system was a new biological model devised in our laboratory, which required that we set up of experiments aimed at first understanding various ecological interactions between various *E. coli* strains and *D. discoideum*. This system is a powerful biological model for investigating the natural variability we found for the interactions between bacteria and amoeba. It opens the door to complementary and innovating experiments in ecology, evolution as well as for medical research.

We found that *E. coli* and *D. discoideum* can be used as a prey/predator model system, and to test for alternative theories regarding the maintenance of virulence. We demonstrated that bacterial strains carrying virulence genes are more resistant to amoeba grazing than bacteria without these traits. Jeon [1] reported a rapid symbiosis between X-bacteria and *D. discoideum*, demonstrating that the interaction between *D. discoideum* and bacteria could be not only complex, but also diverse and very interesting for further evolution experiments.

In our coevolution experiments (temporal allele frequency variation, chapter IV, p.134), we calculated absolute fitness to estimate the selective coefficient under the various environmental conditions we investigated. Using our frozen cultures, we still have to (i) estimate the selection coefficient at various steps of the experiment using competition experiments between ancestral and evolved *E. coli* strains (relative fitness) under our four different environmental conditions; (ii) estimate the evolution of bacterial resistance to digestion by amoeba. Our experiments on the evolution of resistance to phagocytosis in various bacterial strains demonstrated that the evolution of the resistance may well be rapid.

It was possible to differentiate bacterial strains we used with a presumably neutral marker, allowing us to culture them without antibiotics, which may be a selective environment. We followed temporal allele frequency variation in the bacteria prey. It would also be interesting to follow the temporal allele frequency in the predator. With the active participation of Clement Nizak, we constructed during this thesis fluorescent markers for the predator *D. discoideum*, and are now able to follow the system on the both sides, which promises exciting and

innovating experiments using this biological system. In parallel, we have developed methods of picture analysis that allow us to very quickly and precisely estimate population sizes and the frequencies of each marker.

We plan to address the repeatability of evolutionary trajectories and outcomes, since our populations have evolved by natural selection and by genetic drift acting on variation generated solely by spontaneous mutation that occurred during the experiments. Thus, our experiments will allow us to examine the effects of contingency. The role of *historical contingency* (Stephan Jay Gould) in evolution has been much debated but rarely tested. On the other hand, Conway Morris [2] countered that natural selection constrains organism to a relatively few highly adaptive options, such that "*the evolutionary routes are many but the destinations are limited*". We will be able to conduct such experiments with our biological model.

Finally in this work, we used *D. discoideum* as a unicellular organism; however, while they are starving, cells cooperate to form multicellular organisms. Some cells died (20%) and constitute a stalk, leading to consider them 'altruistic', and others form spores (80%), which germinate and can disseminated. Using the fluorescent markers developed for *D. discoideum* and the methods devised to estimate various proportions of *D. discoideum* stained, we plan to conduct experiments on the evolution of cooperation.

Literature cited

1. Jeon, K. and M. Jeon, *Endosymbiosis in amoebae: recently established endosymbionts have become required cytoplasmic components.* Journal of Cellular Physiology, 1976. **89**(2): p. 337-344.
2. Conway Morris, S., *Life's solution.* Inevitable humans in a lonely universe, 2003.

VII. Appendix

APPENDIX 1: CULTURE MEDIA

LB:

25g LB

1L water

Autoclaving

LBA:

40g LBA

1L water

Autoclaving

HL5:

5g Proteose peptone

5g Thiotione E Peptone

5g Yeast Extract

10g Glucose

0,35g Na_2HPO_4, $7H_2O$

0,35g KH_2PO_4

1L water

pH=6,4-6,6 (+HCl)

Autoclaving

HL5A:

HL5

15g Agar

Autoclaving

SM25:

0,4g BactoPeptone

0,04g Yeast Extract

0,4g Glucose

1,9g KH_2PO_4

0,1g $MgSO_4$

1L water

pH=6,4-6,6 (+HCl)

Autoclaving

TA:

10g Tryptone 1g Yeast Extract 5g NaCl 16g Agar 800 ml water Autoclaving	10 g Arabinose 200 ml water Autoclaving

Mix and add 1ml of 5% stock of Tetrazolium filtred (2,3,4 triphenyl tetrazolium chloride).

MCPB:

1,42 g Na_2HPO_4

1,36g KH_2PO_4

0,19g $MgCl_2$

0,03g $CaCl_2$

1l water

pH=6,5

Filtered

APPENDIX 2: FREEZE

Bacteria:

300µl Glycerol 60%

900µl Bacterial culture

-80°C

Amoeba:

500µl DMSO (Dimethyl Sulfoxide) 20%

500µl *D. discoideum* culture

-80°C

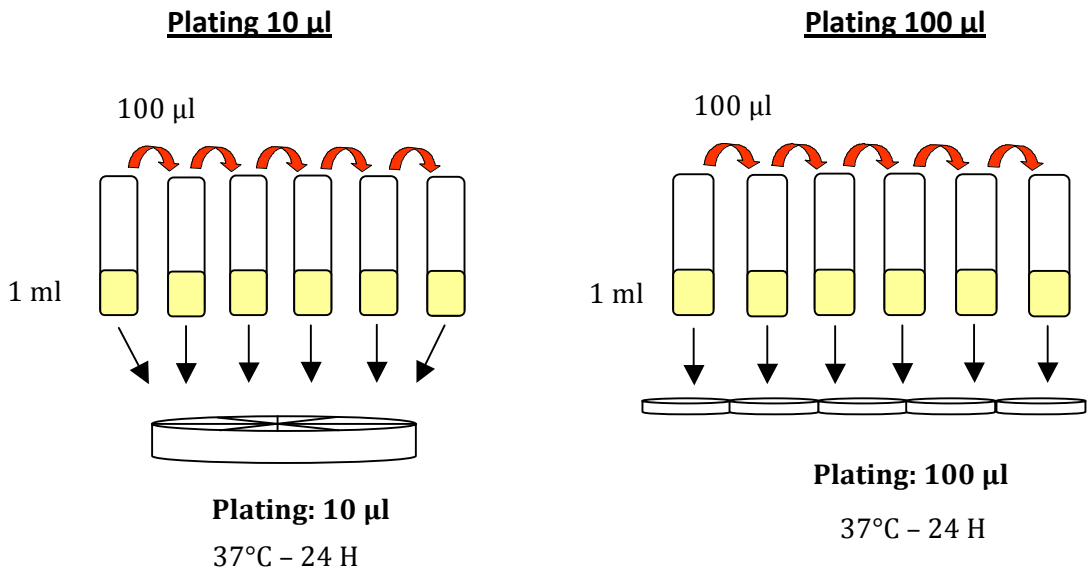
Thawing *D. discoideum* cells:

Transfer the cells to a bottle containing 10ml of HL5 media

Allow the cells to sit (30 min) to recover

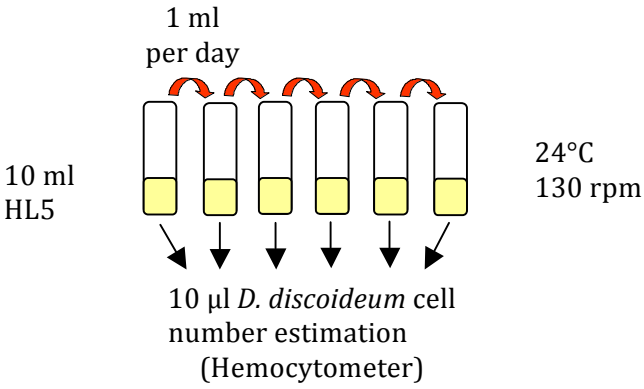
Change the media to remove DMSO.

APPENDIX 3: PROTOCOL: Plating assay

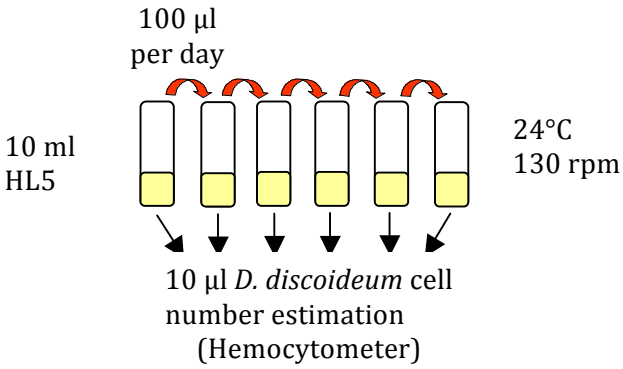


APPENDIX 4: PROTOCOL: Bottleneck regime

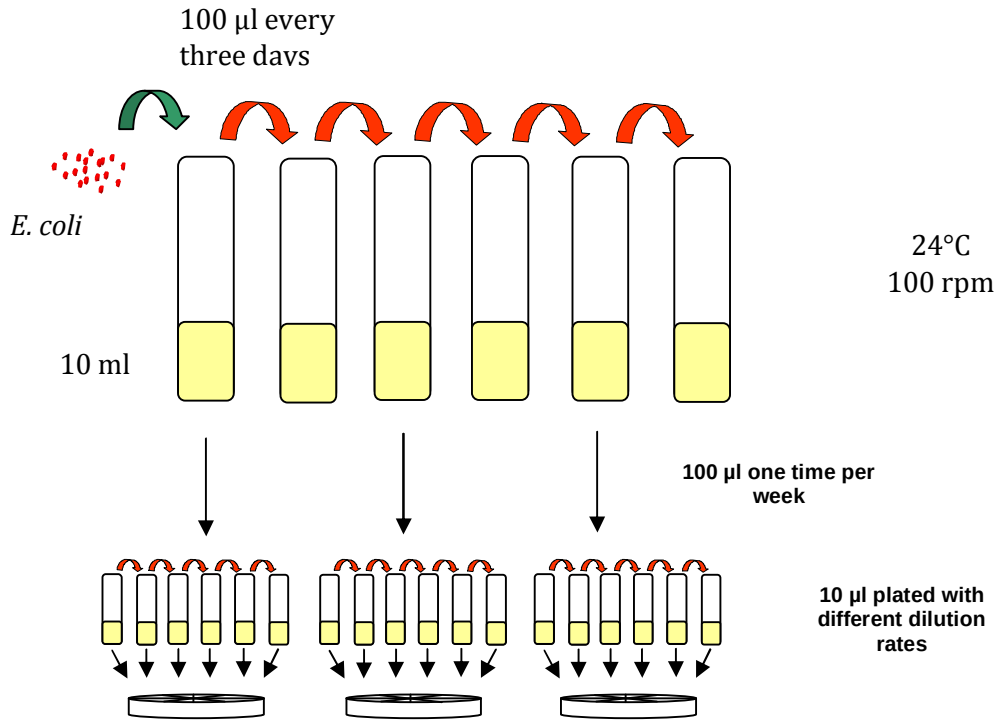
Dilution rate: 1/10



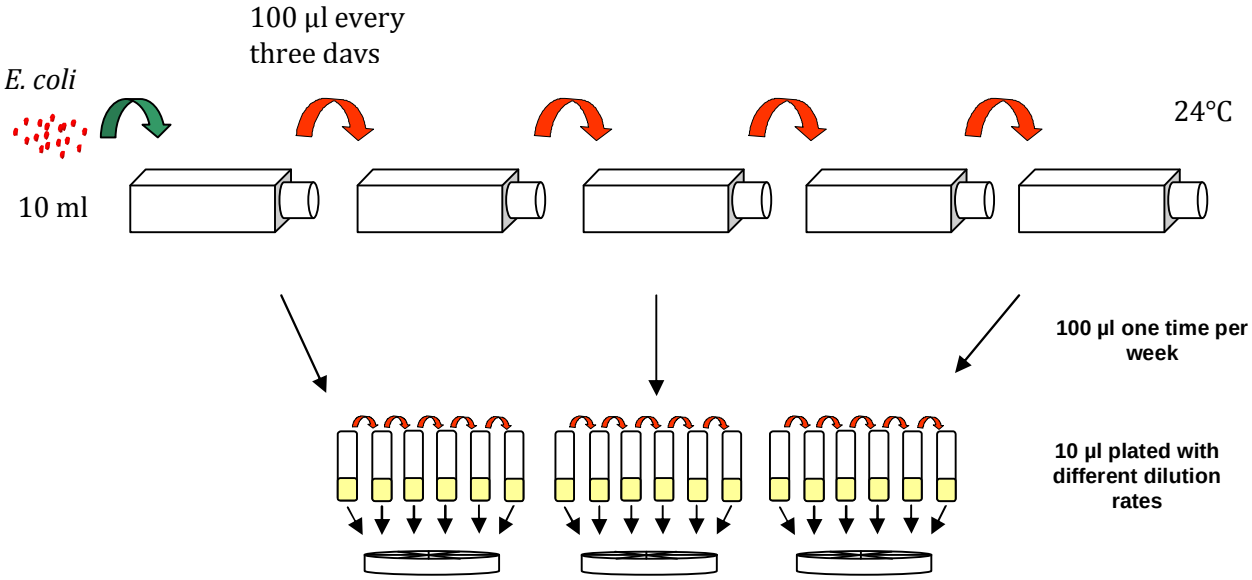
Dilution rate: 1/100



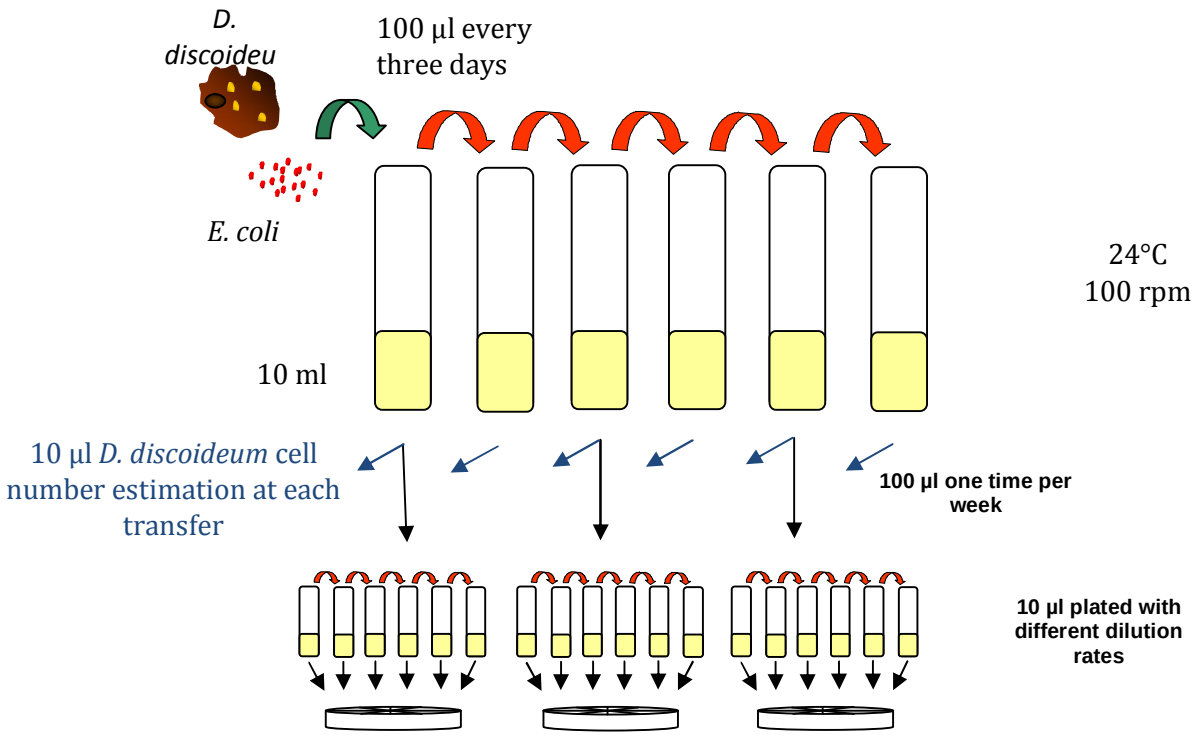
Demography : *E.coli* without spatial structure and without biotic factor



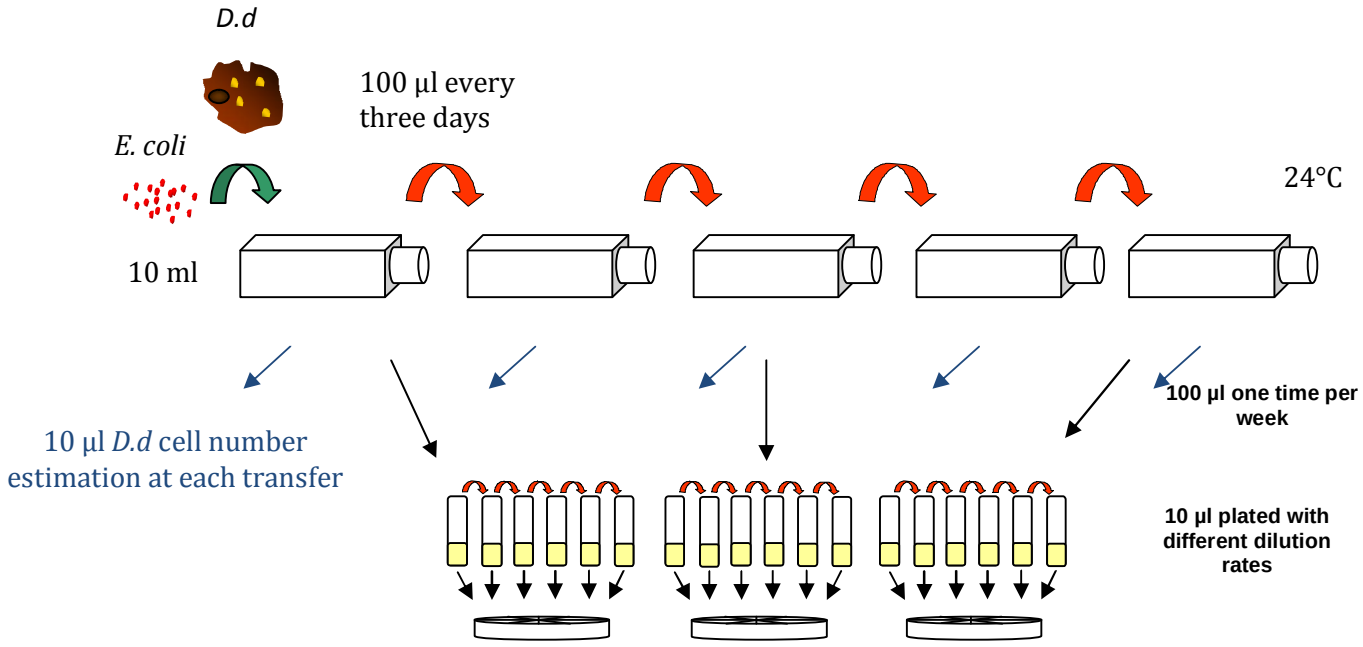
Demography with spatial structure: *E. coli* with spatial structure and without biotic factor



Demography with biotic factor: *E. coli* without spatial structure and with biotic factor (*D. discoideum*)



Demography with spatial structure and biotic factor: *E. coli* with spatial structure and with biotic factor (*D. discoideum*)



VIII. GLOSSARY

Absolute Fitness: of a genotype is defined as the ratio between the number of individuals with that genotype after selection to those before selection. It is calculated for a single generation and may be calculated from absolute numbers or from frequencies.

Acrasin: chemical messenger

Balancing selection: refers to a number of selective processes by which multiple alleles (different versions of a gene) are actively maintained in the gene pool of a population *at frequencies above that of gene mutation*.

Branching process: is a Markov process that models a population in which each individual in generation n produces some random number of individuals in generation $n + 1$, according to a fixed probability distribution that does not vary from individual to individual.

Commensal: is a class of relationship between two organisms in which one organism benefits but the other is unaffected.

Directional selection: occurs when natural selection favors a single phenotype and therefore allele frequency continuously shifts in one direction.

Disruptive or diversifying selection: describes changes in population genetics in which extreme values for a trait are favored over intermediate values. In this case the variance of the trait increases and the population is divided into two distinct groups. This evolutionary process is believed to be the driving force behind sympatric speciation.

Fitness: is an individual's ability to propagate its genes.

Frequency dependent selection: is the term given to an evolutionary process where the fitness of a phenotype is dependent on its frequency relative to other phenotypes in a given population.

Genetic drift: is change in the frequency of a gene in a population due to random sampling.

Hill Robertson effect or clonal interference: describes an evolutionary advantage to genetic recombination.

Hitchhicking or Selective sweep: is the process by which an evolutionarily neutral or deleterious allele or mutation may spread through the gene pool by virtue of being linked to a gene that is positively selected.

Isogenic: refers to a group of genetically identical individuals.

Negative frequency dependent selection: the fitness of a phenotype increases as it becomes rarer. Negative frequency-dependent selection is an example of balancing selection.

Polymorphism: having multiple alleles of a gene within a population.

Relative fitness: is quantified as the average number of surviving progeny of a particular genotype compared with average number of surviving progeny of competing genotypes after a single generation, *i.e.*, one genotype is normalized at $w = 1$ and the fitness of other genotypes is measured with respect to that genotype.

Speciation: evolutionary process by which new biological species arise.

Stabilizing selection: type of natural selection in which genetic diversity decreases as the population stabilizes on a particular trait value.

List of figures

List of tables

List of Boxes

Abstract

The diversity of living organisms is essential for their capacity to evolve and adapt to environmental changes. Therefore, determining the factors responsible for the origin of diversity and for the maintenance of the genetic variance observed remains central and fundamental research objective. The aim of this thesis was to understand the evolutionary factors maintaining neutral polymorphism. Since the influence of evolutionary processes such as natural selection and genetic drift are complex, we developed complementary experimental and theoretical approaches in order to disentangle their contributions.

Using a biological model consisting of the bacterium *Escherichia coli* and the social amoeba *Dictyostelium discoideum* enable us to study the natural variability of interactions between the two species. In the second part of this work, we studied the bacterial traits involved in this natural variability. We showed that bacteria carrying virulence genes were resistant to grazing by amoeba, a result which was in agreement with the coincidental evolution hypothesis of virulence factors.

We then focus on population genetics aspects of our biological system. In coevolution experiments, we followed temporal allele frequency variations over 300 bacterial generations under four sets of environmental conditions: with or without biotic factor and with or without spatial structure. Our results did not differ from genetic drift predictions. The aim of theoretical model we developed was to address the demographic stochasticity effects on neutral allele fixation probability and time to fixation. We found that fixation probability and the time to fixation were affected by the demographic stochasticity compared with a model using a population of constant size (Moran model).

Key words: natural selection, genetic drift, polymorphism, coevolution, virulence, bacteria, amoeba, fixation.

Résumé

La diversité des organismes est essentielle pour leur capacité à évoluer et s'adapter aux variations environnementales. De ce fait, déterminer les facteurs responsables de l'origine de cette diversité ainsi que de la maintenance de la variabilité génétique observée reste un objectif fondamental en recherche. L'objectif de cette thèse était de comprendre les facteurs évolutifs maintenant le polymorphisme neutre. L'influence des processus évolutifs tels que la sélection naturelle et la dérive génétique étant complexes, nous avons combiné des approches complémentaires expérimentale et théorique.

Le système expérimental utilisé, la bactérie *Escherichia coli* et l'amibe sociale *Dictyostelium discoideum* nous a permis d'étudier dans un premier temps la variabilité naturelle des interactions existant entre les deux espèces. Dans une seconde partie, nous avons étudié les traits bactériens impliqués dans cette variabilité. Nous avons montré que les bactéries portant des facteurs de virulence sont plus résistantes à la digestion des amibes, ce qui est en accord avec l'hypothèse de coïncidence évolutive des facteurs de virulence.

Le deuxième volet de cette thèse concerne les aspects de génétique des populations de ce système. La troisième partie de notre expérimentation était de suivre les variations temporelles des fréquences alléliques de populations bactériennes comportant un marqueur neutre, durant 300 générations et sous quatre conditions environnementales : avec ou sans structuration spatiale et avec ou sans facteur biotique. Nous avons observé que les variations des fréquences alléliques observées étaient compatibles avec la dérive génétique.

L'objectif du modèle théorique a été dans un premier temps d'étudier les effets de la stochasticité démographique sur les probabilités de fixation d'un nouvel allèle neutre arrivant dans une population résidente ainsi que sur le temps de fixation. La probabilité de fixation ainsi que le temps de fixation sont modifiés par les effets stochastiques lorsque l'on compare nos modèles à taille de population fluctuantes à un modèle à taille de population constante tel que le modèle de Moran.

Mots clés : sélection naturelle, dérive génétique, polymorphisme, coévolution, virulence, bactérie, amibe, fixation.