Predator control of diversity: case studies using microcosms.

Thesis submitted for the degree of Doctor of Philosophy

University College London

by

Alessandra Dupont

Department of Genetics, Evolution and Environment,
Center for Biodiversity Research
University College London
And

Department of Life Sciences
Natural History Museum
September 2016

Declaration

I, Alessandra Ö. C.- Dupont, confirm that the work presented in this thesis is my own. Where information had been derived from other sources, I confirm that this has been indicated in this thesis.

“The scientist does not study nature because it is useful; he studies it because he delights in it, and he delights in it because it is beautiful. If nature were not beautiful, it would not be worth knowing, and if nature would not be worth knowing, life would not be worth living.”

Henry Poincaré (1854-1912)

Abstract

Predation is a major mode of interaction in natural environments, and predators have an important impact on prey evolution, community composition and food web complexity. Bacterivorous protists are key components of aquatic and terrestrial environments, as well as major drivers of bacterial diversity and community composition. When grazing on prokaryotes, protists impact inter- and intraspecific interactions, biomass production and biogeochemical flows. Furthermore, the prey's ability to develop physiological and morphological defence mechanisms affects both bacterial diversity and predatory eukaryotes' grazing ability, survival and distribution. While theoretical work usually analyses simple to partially complex predator-prey systems, practical studies focus are often limited to one prey and one predator species.

Using a variety of ecologically and physiologically diverse bacterivorous protists, this study identifies the relationship between predator-prey interactions and its impact on species diversity. While bacterial communities are closely linked to their environment characteristics, protist diversity and distribution was shown to be dependent on both biotic and abiotic factors, and community composition differences driven by few major lineages. Similarly, bacterial communities submitted to varied protist predators were distinguished only by a handful of major lineages. In parallel, protist impact on bacterial diversity was highly modulated by prey community taxonomical composition and ecological strategies. Indeed, in this study, bacteria observed genotypic changes in line with short-term phenotypic plasticity resulting on the development of defence mechanisms against predators of distinct ecological niches.

Predator-prey interactions in the light of diversity are far more complex than what closed microcosm experiments can translate, but they encompass valuable information applicable to natural systems.

Acknowledgments

First of all, I would like to thank David & David, my supervisors, for giving me the opportunity to work on this project. I am grateful you believed in me at all times, even when I wasn't really sure of what I was doing anymore. I learned a great deal with both of you, and not only about protist ecology!

Little special note to Tom and his incredible R knowledge and will to help me out any given moment. I would also like to thank Ville for all his help and input. Without you, I would have been a chapter and two papers too short.

To the former "Bass lab", thank you for welcoming me into their headquarters; and Hannah for inducing me to the Museum's laboratories and their health & safety regulations... To Sigrid, Cedric, Aurélie, Finlay, Guy and Theresa for making me a part of their group and coffee breaks, Christmas meals and chocolate time. In the new "Bass lab", special thanks to Rachel for so much help in everything lab and admin related. I wouldn't haven managed without your help!

To my fellow PhD students – Ines, Francesca and Georgia in particular – for meeting on Friday for drinks and then again in the lab over the weekend.

To my friends in Paris, London and elsewhere for always asking if I'm a doctor already, and how are my 'beasties' doing (after finding out protists are a thing!).

To Alex, for being there with and for me when it all started.

Special thanks to Björn for giving the best working station one can get; for putting up with me, my mess, my procrastination and lack of will but never ever giving up on me.

Last but far from least, to my brother and his usual "Are you done yet or what?". Guess what, I finally did it. And most of all, to my parents for always believing and encouraging me. I did finish my 'studies' in 2016 after all…

Publications

Friman, V.-P., Dupont, A.Ö.C., Bass, D., Murrell, D.J., and Bell, T. (2015) Relative importance of evolutionary dynamics depends on the composition of microbial predator–prey community. *ISME J.* 1–11.

Dupont, A.Ö.C., Griffiths, R.I., Bell, T., and Bass, D. (2016) Differences in soil micro-eukaryotic communities over soil pH gradients are strongly driven by parasites and saprotrophs. *Environ. Microbiol.*

Dupont, A.Ö.C., Bell, T. and Bass, D. Protist grazing effects on natural bacterial community taxonomic composition. *In prep.*

Dupont, A.Ö.C., Friman, V.-P., Bell, T., Murrell, D.J. and Bass, D. *Pseudomonas fluorescens* SBW25 transcriptomie analyses in response to protist predation. *In prep.*

Abbreviations

ANOSIM Analysis of Similarity

SIMPER Similarity Percentage analysis

NMDS Non-Metric Dimensional Scaling

PCA Principal Component Analysis

OTU Operational Taxonomic Unit

LB Luria-Bertani Broth

Microorganisms

19-3E *Cercomonas paraglobosa*

AP *Acanthamoeba polyphaga*

B13 *Cercomonas braziliensis*

BS *Bodo saltans*

Ca5HKv (=CS) *Cercomonas saepenatans*

CaSphII *Cercomonas pigra*

CP *Chilomonas paramecium*

Kv-Hf *Allapsa scotia*

NZ1-5c *Paracercomonas vonderheydenii*

SW2 *Paracercomonas minima*

TP *Tetrahymena pyriformis*

TV *Tetrahymena vorax*

TABLE OF CONTENTS

Table of figures

Table of tables

Chapter 1 INTRODUCTION

Humanity has been observing and documenting the natural world since the dawn of civilisation. Ernst Haeckel, a German scientist contemporaneous and fan of Darwin's work, was the first to use the term 'ecology' in order to describe 'the study of the interactions between organisms and their environment' (Begon *et al.*, 2007). However, in a world where species numbers estimates range between 1.5 million and 30 million different species of organisms (May 1990), and where every single one of them interacts with both their biotic and abiotic environment, ecology will always be the field of new and updated discoveries.

1.1 A microbial world

In the late 1590s, two Dutch spectacle makers – Zacharias Jansen and his father Hans started experimenting with lenses, and realised that when combining several of them in a single tube, the image at its end appeared much larger than what any magnifying glass could achieve. However, magnification wasn't that powerful – of about 9x only – and the image at the end of the tube was mostly blurry, and no Jansen 'microscope' survived. It is Anton van Leeuwenhoek, a Dutch scientist, who later developed the first real microscope by further working on lenses in order to make them better. With his powerful new tool – a magnifying tube with a 270x power of magnification – van Leeuwenhoek explored and studied the world around him. When he first observed and described life in a water droplet – teeming with what he termed "animalcules" – the Dutch scientist became a pioneer in the study of microorganisms, and later the 'Father of Microbiology'. It is Robert Hooke however, a British multidisciplinary scientist, who first used the term 'cell' to describe the microscopic structure of cork, e.g. plant cells.

In the late 17th and 18th centuries, many microscopists had accurately observed and described a series of 'animalcules' and categorised them into 'Animalia' or 'Plantae', the classification of living beings used at the time (Corliss, 1998). It would be Ernst Haeckel –again! – who would propose the term 'Protozoa' as

third kingdom of life, and evolutionary origin to the long accepted two previous kingdoms, although not in a single linear timeline anymore. Indeed, Haeckel proposed a 'tree' to illustrate kingdoms, groups and their relationships (fig 1.1).

Later scientific advances would allow scientists to distinguish bacteria – without a distinct nucleus – from protists, possessors of a distinct nucleus. For a longer time, the kingdom Protista contained a multitude of organisms that wouldn't fit any of the established Animalia, Plantae, Prokaryota or Fungi (defined by Witthaker in 1969). The advent of molecular techniques in the late 20th century however, distinguished Eukaryotes – amongst which Protists – and Prokaryotes (Woese, 1987).

Modern protist classification is based on a combination of morphological studies, and confirm that protists are not only morphologically diverse, but also phylogenetically divergent, belonging to a variety of eukaryotic supergroups (fig 2.2) (Cavalier-Smith and Chao, 2003; Adl *et al.*, 2005, 2012; Pawlowski, 2014).

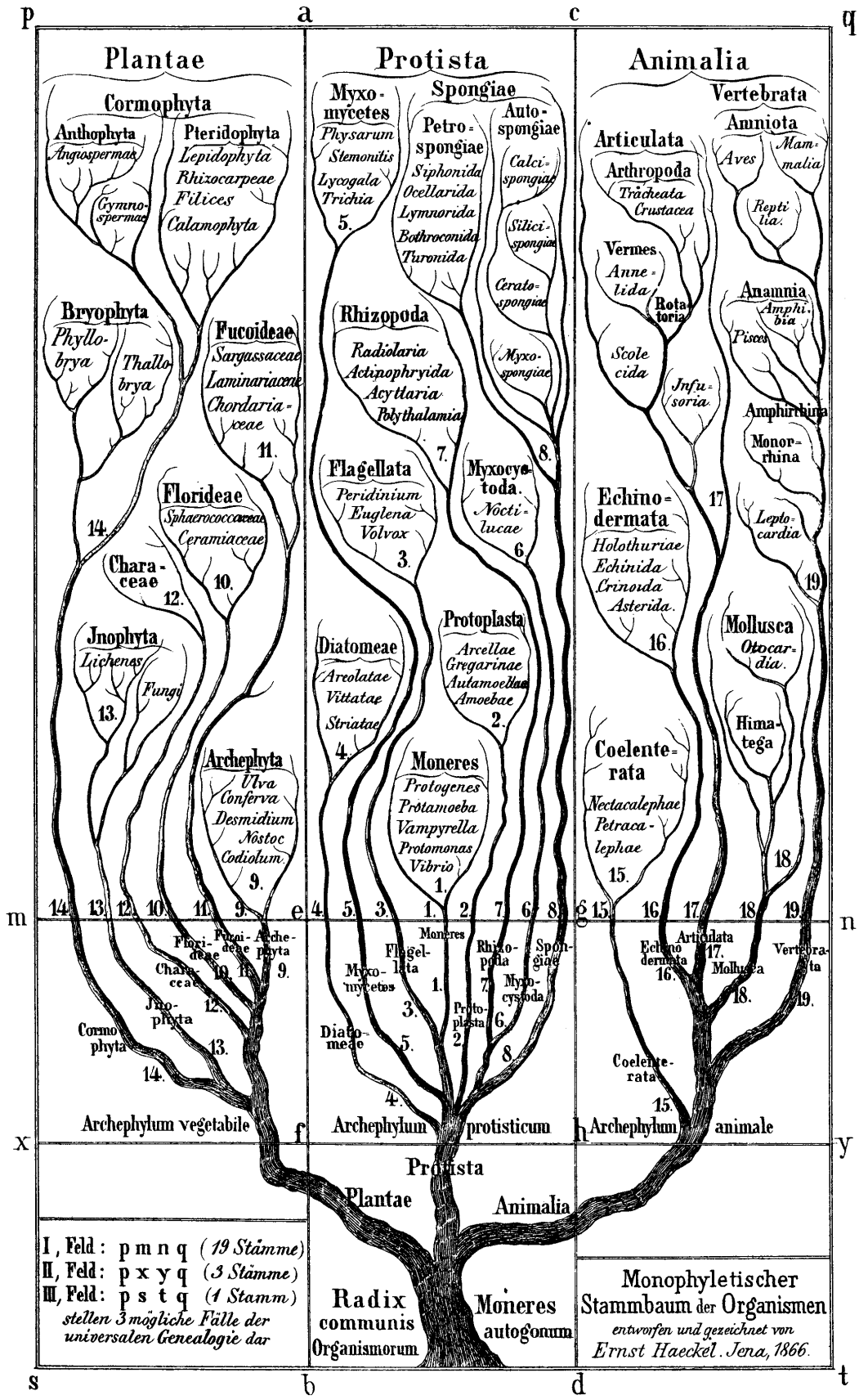


Figure 1.1 Reproduction of Haeckel's tree of life, highlighting the three principal 'branches' of organisms (Protista, Plantae and Animalia kingdoms).

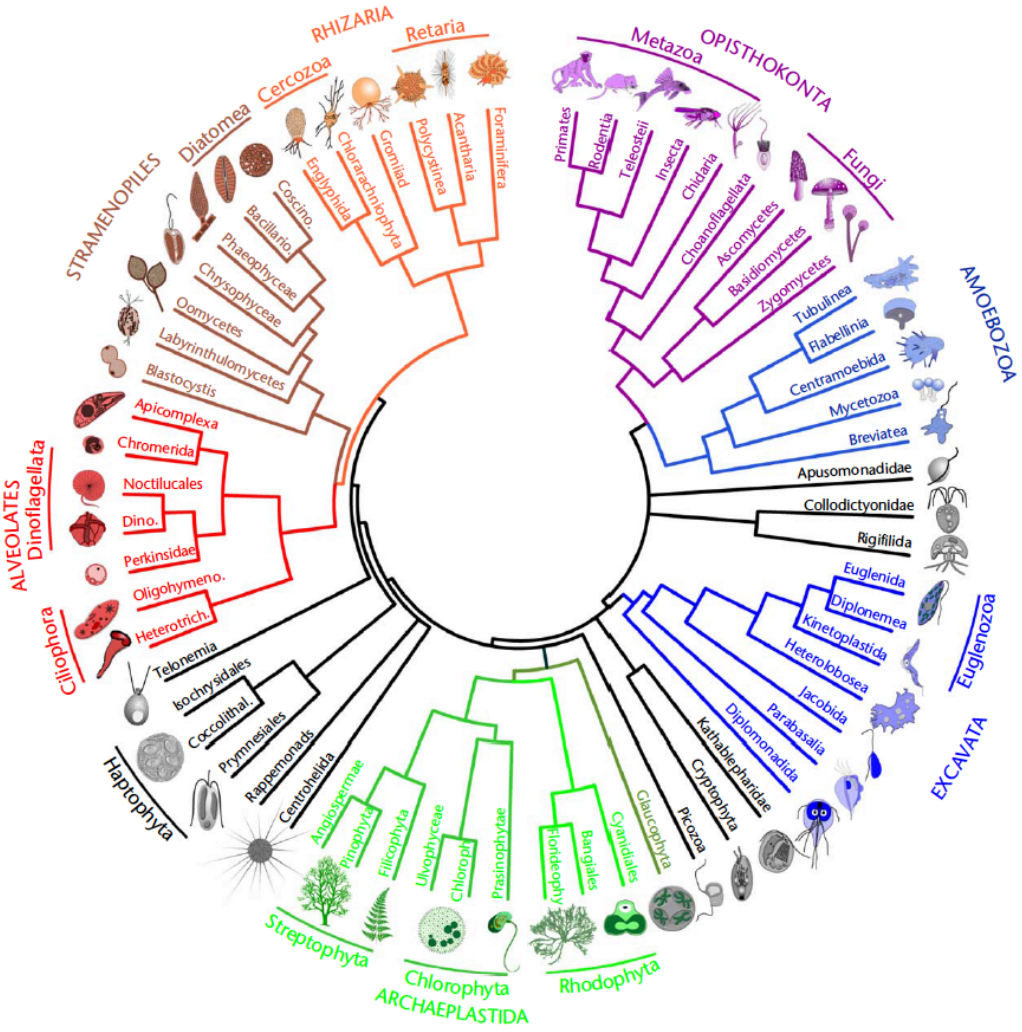


Figure 1.2 Phylogenomic tree of eukaryotes (Pawlowski, 2014)

1.2 Protists

Coined by Haeckel, the term 'Protista' (from the Greek '*protistos*') means "the very first". Protists are unicellular microbial eukaryotes (Adl *et al.* 2005) and key components of microbial communities in both aquatic and terrestrial environments. If some parasitic forms responsible for well-known human diseases, such as toxoplasmosis for example, are rather well known, protists remain understudied. Indeed, their life stories tend to be complex, often including multiple and distinct morphologies.

1.2.1 Where is everybody?

Ecological biogeography – as opposed to historical biogeography, which focuses on dispersal events and evolutionary relationships between organisms occurring over geological time-scales – aims to understand the factors affecting species distribution. While macroecology can easily assess the distribution of two distinctive species, on the microscopic scale it can often be difficult to differentiate morphologically similar organisms. Knowing which protist species – or even taxa – occur where easily becomes a harder task.

Two long debated views regarding protist distribution, still up to date, propose that species display either ubiquity or moderate endemism. Baas Becking (1934) stated that, in the case of microorganisms 'Everything is everywhere, but the environment selects'. However, the second part of the sentence – that underlies the importance of ecological restrictions on protist distribution – got lost in time, and Becking's statement misinterpretation persisted for long, raising the debate on microorganisms dispersion and endemicity (Foissner, 1999, 2006; Fenchel and Finlay, 2004; Finlay, 2004). If protist dispersion would be only limited by ecological and not historical factors (historical biogeography), background rates of resting and dispersal units such as spores and cysts would be expected to be found in almost all environments.

Finlay and colleagues (Fenchel and Finlay, 2004) argue that microbial organisms are so abundant that their distribution ought to be global, if not by dispersal (Finlay, 2004), then for pure statistical reasons (Finlay, 2002). On the other hand, Bates *et al.* (2013) observed distribution patterns in protist species according to biome types, thus inferring a certain degree of endemism. This agrees with Foissner's "moderate endemicity model" (Foissner, 1999, 2006) that compares protist dispersal to higher organisms' dispersal. Whether these are a limited species (i.e. morphotypes, such as amoeba) world widely distributed or endemic evolutionarily diversified species, true diversity of unicellular eukaryotes is still an open question. Meanwhile, new molecular

taxonomy techniques are helping resolve the debate(Weisse, 2008; Adl *et al.*, 2013).

Microscopy-based soil protist diversity studies far outnumber soil environmental DNA (eDNA) studies, and both lag far behind their marine and freshwater counterparts. Diversity assessments that rely on culturing and/or visual identification have revealed a large diversity of cell forms and taxa dominated by bacterivores, predators, and some autotrophs (Stout, 1984; Bamforth, 2007; Domonell *et al.*, 2013) These studies often rely on protists capable of growing in culture medium supplemented with bacteria and recognition of visually distinctive (and relatively large) forms, e.g. via liquid aliquot isolation techniques (Domonell *et al.*, 2013). One consequence of this bias is that naked and testate amoebae, ciliates, some flagellates, diatoms, and green algae dominate the results. In some cases fungi are reported and/or the focus is specifically on heterotrophs/bacterivores. Environmental polymerase chain-reaction (PCR) and sequencing studies are not taxonomically, ecologically, or visually constrained in the same ways and reveal many novel lineages including parasites (Geisen, Laros, *et al.*, 2015). Indeed, culture independent studies are an evidence of resting-stage diversity: the large number of operational taxonomic studies (OTUs) occurring in low densities defines a 'rare biosphere' of organisms. These can possibly represent species dispersed form all over the world but not able to thrive in the locality they were found in. if this rare biosphere is thus constrained solely by ecological and environmental conditions, organisms would be expected to activate in favourable laboratory conditions.

Non-PCR based metagenome sequencing studies rarely feature protists, mostly focusing on bacteria (Pearce *et al.*, 2012; Fierer *et al.*, 2013) and/or aspects of metabolism. Metatranscriptomic studies (Urich *et al.*, 2008; Lehembre *et al.*, 2013; Turner *et al.*, 2013) interestingly reveal diversity profiles of active soil biota that differ in some important respects to amplicon studies, particularly demonstrating higher diversity and abundance of genetically divergent lineages (including many parasitic lineages) that are underrepresented in amplicon studies at least partly due to negative PCR biases resulting from mismatches

between commonly used primer sequences and divergent templates, and amplicon length variation. However, there are so few molecular studies on soil for comparison that it is too early to generalise about soil protistan diversity, particularly because soils are so heterogeneous. Often soil-based molecular studies are primarily concerned with specific ecological situations and focus on broad changes in total eukaryote community structure and rarely look in detail at the validity of protist hits, so usually illuminate protistan diversity and distribution at relatively low taxonomic resolution (Murase and Frenzel, 2008; Turner *et al.*, 2013). Furthermore, most studies use the standard SILVA 18S (Quast *et al.*, 2013) database for eukaryotic taxonomies, while other highly curated ones such as the Protist Ribosomal Reference database (PR2; Guillou *et al.*, 2013)remain fairly unknown or unused.

1.2.2 Small big players

Protists are key components of microbial communities in both aquatic and terrestrial environments. They represent 10^4-10^7 individuals per gram of dry soil (Adl and Coleman, 2005; Adl and Gupta, 2006; Bamforth, 2007) and over 50% of total aquatic biomass (Sherr and Sherr, 2002, 2007). The diversity of functional groups (trophic status, free-living vs symbiotic, etc.) makes them major participants of the microbial loop and regulators of biogeochemical flows (Calbet and Landry, 2004). With the continuing development and growing capability of molecular techniques, protist diversity is increasingly revealed as orders of magnitude greater than morphological or even earlier sequence-based assays suggested (e.g. Bates *et al.*, 2013). The past ten years have seen extraordinary advances in our knowledge of microbial eukaryotic diversity, primarily through the adoption of molecular tools for phylogenetically based classification which provides a coherent evolutionary framework to explore diversity. Additionally the routine use of environmental sequencing utilising high throughput sequencing (HTS) technologies has permitted the discovery of many new groups and novel eukaryotic lineages in many different biomes (Takishita *et al.*, 2007; Jones *et al.*, 2011; Vargas *et al.*, 2015). However, the challenge of

overlaying ecological and biogeographical insight onto this diversity still remains, particularly in the complex and heterogeneous soil environment.

Free-living protist forms – opposed to parasitic ones – are highly abundant in natural aquatic systems, and may occur in the water column as well as in the benthos. Along with other species such as bacteria, micro-algae and micro-metazoans (rotifers and crustaceans), they represent the major interacting components of microbial food webs in aquatic environments. When employing phagotrophy, free-living protists predate on similar-sized particles, such as bacteria, micro-algae and other protists. Bacterivorous protists were shown to be major regulators of bacterial densities (Finlay 2004), and predation in aquatic microbial food webs is dominated by phagotrophic protists (Sherr and Sherr, 2002). When actively grazing on bacteria, in both aquatic and terrestrial environments, protists become part of the “microbial loop” (Azam *et al.*, 1983; Clarholm, 1985), and their pressure on bacterial biomass and nutrient mineralisation thus determines the rhizosphere’s use of water and nutrient (Bonkowski, 2004) as well as bacterial community composition (Kreuzer *et al.*, 2006; Glücksman *et al.*, 2010).
Many species are also commonly found in various terrestrial habitats, and their biomass equals or exceeds that of all other soil animals (but for earthworms) in most soils (Schröter *et al.*, 2003). A major constituent of the protozoan fauna in soils throughout the world consists of free-living heterotrophic biflagellates (Howe *et al.*, 2009), such as Cercozoa, along with Ciliophora, as well as groups like Apicomplexa and Dinophycea (Bates *et al.* 2013).

1.3 A bigger game

Although prey-predator interactions have a strong theoretical basis, these remain minimal representations of predator-prey systems (Abrams, 2000). Indeed, when assuming no density changes in prey or predator populations, Lotka's model (1925) leads to a stable periodic dynamic system where predator density peaks with a certain lag after the prey’s population maximum, while extinction is the final outcome in a host-parasite situation (Nicholson and Bailey,

1935). Natural systems are not, however, static in time in terms so f population density, nor composed of only two interacting species. While parasitic systems may represent 'simple' one predator – one-prey systems, free-living organisms are usually coexisting with more than one species involving a variety of trophic levels. Community dynamics is thus directly related to food-web complexity and the plethora of interactions – between and within trophic levels – of the system. Furthermore, variation of abiotic factors such as light, temperature or pH equally impact predator-prey interactions, both direct and indirectly (Moore *et al.*, 2004; Hiltunen *et al.*, 2015).

Controlled experiments being virtually impossible with large organisms, predator-prey interactions and community dynamics have been empirically investigated using microbial species (Holyoak and Lawler, 2005). With the advent of new technologies such as high throughput sequencing, both protistan and bacterial species that had been previously well studied in terms of laboratory experiments have since had their genomes studied. Microbial eukaryotes and prokaryotes not only require little maintenance, have shorter generation times (when compared to typical model organisms), they also represent a variety of domains. This implies that even "simple" two- or three-species systems can actually represent complex food webs (Hiltunen *et al.*, 2013). The conjunction of the many bacterial and protozoan 'omics and 'classical' microcosms studies allowed those organisms to be considered as models for

Protist-bacteria microcosm studies involving physiologically and phylogenetically diverse species can thus be highly informative when it comes to interactions in complex systems, while being easily controllable.

1.3.1 Predation

Predation is a major mode of interaction in natural environments, and is defined as 'the consumption of all or part of a living organism by another' (Morin, 2011). It is the type of interaction where the predator will benefit from the prey demise. Predators have a significant role in shaping prey populations, communities and

their ecology. By changing the balance between mortality and growth rates, predation defines the prey community structure, which reflects on the composition of different trophic levels and the food web complexity (Corno and Jürgens, 2008).

Interactions between protists and bacteria are particularly interesting as they are both key players in biogeochemical cycles in aquatic and terrestrial environments (Burgin *et al.*, 2011; Madsen, 2011). Bacterivorous protists are, along with the availability of nutritional resources, major drivers of bacterial diversity and community composition, and may control the proportion of active bacteria and biomass (Jürgens and Matz, 2002; Sherr and Sherr, 2002; Pernthaler, 2005). When grazing, protists exert a strong top-down control on bacterial populations, setting the amount of available organic matter to higher trophic levels (Zöllner *et al.*, 2009). In addition, generalist predation, or predation on selected taxa differently affect intraspecific prey interactions, impacting the microbial net production (biomass). On the other hand, the ability of bacteria to develop physiological and morphological defence mechanisms under eukaryotic grazing pressure affects both bacterial diversity and predatory eukaryotes' grazing ability and survival (Jousset, 2012).

Heterotrophic phagotrophy is considered the primitive protistan nutrition mode, and is widespread amongst free-living organisms. But a wide range of feeding strategies has been observed and described. Jones (1997) categorised four mixotroph feeding behaviours, where organisms are able to employ either autotrophy or heterotrophy according to environmental conditions and prey availability. Therefore, group A protists are primarily heterotrophic and only use phototrophy when prey concentration is too low to ensure heterotrophic grow; group B is mainly phototrophic and, inversely to group A, ingest prey when light is limiting. Group C phagotrophy is inversely proportional to light intensity. Finally, group D protists, mainly autotrophic, ingest prey rarely.

At the lowest level of trophic interactions, microscopic organisms represent one of the major actors in food webs and their interactions with bacteria and primary producers are one of the main factors regulating biogeochemical flows (Calbet

and Landry, 2004). Indeed, bacterial metabolism is responsible for the transfer and recycling of elements within the microbial loop, as well as for the export towards higher trophic levels, thus generating essential ecosystem functioning fluxes from bottom to top trophic levels (Calbet and Landry, 2004; Jürgens and Massana, 2008). Fluxes heading from bottom to top levels play a major role in ecosystem functioning, which makes the large number of protists dwelling in soils (10^4-10^7 active individuals m^{-2}, (Finlay, 2004) an important component of biogeochemical cycles (Adl and Gupta, 2006). Understanding the general ecological principles that determine how predators affect prey biodiversity might be especially important for conservation biology, but also for community and applied ecology.

1.3.2 Competition

Interspecific competition most likely affects population dynamics of competing species, which in turn affects their growth, distribution and survival rate (Morin, 2011). Broadly, competition is the interaction where a species finds itself deprived of resources (due to lack of space, limited nutritional sources, etc) as a consequence of the consumption of that same resource by another species. While predation represents interactions with a negative side (for the prey) and a positive side (for the predator), competition is characterised by a double negative effect, where both organisms suffer from their competitors' mutual action, although one species does not consume the other. However, not every competitive interaction is necessarily detrimental to both parties. Many different mechanisms can define competition, and can either lead to stable coexistence of both (or all) species or to the extinction of one or more of the competing species (Schoener, 1983; Morin, 2011). This classification accounts for interspecific – between different species – competition of sessile and motile organisms. In this study, however, only motile protist species have been used.

While there have been numerous studies of the effect of single protist predators on simple and complex, natural bacterial communities, the effect of multiple predators on diverse prey communities remains little studied. Even though species pairs may occur in natural systems, they are seldom independent of the

other players in it, and community dynamics depart from simple predictions based on pairwise interactions (Friman and Buckling, 2013). Predator-prey interactions thus imply not only pairs, but also all the combinations between (predation) and within (competition) present species. Furthermore, most ecological interactions analysed in complex systems have a theoretical basis, and few recent studies have been empirically testing ecological models (Rainey and Rainey, 2003; Heger *et al.*, 2014; Quintana *et al.*, 2014).

1.3.3 Co-evolution

Adaptations to biotic and abiotic pressures allow survival or death, and define population and species dynamics (Doebeli and Dieckmann, 2000; Hiltunen and Becks, 2014). Prey must adapt to avoid predation: for instance, bacterial resistance to protist predation results in bacteria flocking behaviour, biofilm formation, size differentiation, etc. Indeed, bacteria observe 'rapid evolution' mechanisms in experimental conditions that produce distinctive population dynamics and behaviours (Ellner and Becks, 2011; Friman *et al.*, 2014), but showed inconclusive in theoretical predictions (Yoshida *et al.*, 2003).

In parallel, predators must overcome anti-predatory defences in order to thrive. While rapid evolution and defence adaptation have been widely described in microbial systems, the focus has always been on bacteria – easy to maintain and follow in laboratory conditions, bacteria-phage systems can easily simulate predator-prey interactions (Friman and Buckling, 2013, 2014; Scanlan *et al.*, 2015). The effect of protist predation on bacterial communities also has mostly been described from the prokaryotic point of view, very little studies based on the eukaryotic predator evolutionary aspects (Pernthaler *et al.*, 1997; Becks *et al.*, 2005; Hiltunen and Becks, 2014). Furthermore, predator-prey co-existence triggers co-evolution, which in turn feedbacks onto population dynamics. One of the aspects of the use of microbial systems in laboratory studies is the rapidity at which each species develop, a new bacterial generation being started at times as short at 30min depending on the species. Similarly, protist generation times – although slightly longer (personal observations: from 5 to 10 days) – make them easy-to-use organisms. But experiments rarely consider much more

than hundreds of bacterial generations, and longer-term dynamics of microbial systems remain understudied.

Although co-evolution has been showed to impact population dynamics and species coexistence (Yoshida *et al.*, 2003; Becks *et al.*, 2005; Friman *et al.*, 2014), few have considered those changes on the genome and gene expression level, notably when considering the whole transcriptome of a microorganism (Gissot *et al.*, 2009; Passalacqua *et al.*, 2009). Many studies have investigated the transcriptome of a whole system, complementing new environmental DNA analyses. Indeed, while sequencing the DNA allows to observe the presence (and therefore absence) of species (and discover so many new ones!), RNA sequencing gives a more dynamic image. Because the latter technique gathers the collection of all transcribed sequences in a cell, we actually see what is going on: given a controlled modification of abiotic factors, can we identify the genes driving the observed morphological and behavioural changes?

With the unfolding of the 21st century, global (climate) change is accelerating (IPCC, 2013) It is therefore essential to deepen the understanding of our planet and its natural mechanisms – may they be biological, physical, chemical, etc. When combining that knowledge to all the years of previous scientific research, one can predict future responses to climate change (among other phenomena, natural and anthropogenic) more and more accurately.

1.4 Aims of thesis

The main aim of this study is to understand the effect of predator – prey interactions in diversity, structuring community and the mechanisms underlying those interactions in a variety of experimental conditions.

At first, I analysed a eukaryotic 454 sequencing 18S rDNA dataset, generated from a subset of the 2007 Countryside Survey samples. The Countryside Survey (www.countrysidesurvey.org.uk), a multi-sample assessment of bacterial communities across the full spectrum of UK soil types showed that

bacterial community structure was strongly determined by soil variables such as soil pH (Griffiths *et al*., 2011). I seek to compare community structures across the three soil pH classes (low, medium, high), and contrast with patterns observed in bacterial communities. I also explored which taxonomic groups differ between different soil types, and at what level of taxonomic hierarchy differences are manifest. Finally I examined the reliability of protist taxonomic assignments by comparing the performance of different databases. I used the databases to provide an in depth evaluation of some novel groups, which are highly represented in soil 18S libraries but whose evolutionary affiliations and relationships are yet to be resolved.

Cercomonads and glissomonads (=sarcomonads) are highly diverse groups of heterotrophic flagellates, abundant in both aquatic and terrestrial environments. The first morphological observations of those groups by Dujardin (1841) and others have recently been revised by the use of more precise morphological and molecular techniques (Bass, Chao, *et al.*, 2009; Howe *et al*., 2009). Sarcomonads include a variety of morphotypes both within and between taxa, although little is known about the potential functional diversity independent of morphological similarity or ecological preferences. Glücksman *et al.* (2010) analysed the effect of various phenotypic traits, as well as genetic relatedness, of a variety of (single) cercomonad and glissomonad strains on natural bacterial communities. Grazed bacterial communities tRFLP fingerprinting of post-grazing bacterial communities showed that protist cell size and “amoeboid-ness” (morphological cell-shape plasticity), and more weakly phylogenetic distance between the predator protist strains influenced prey community composition, thus prey selection. However, tRFLP profiles provide only a fingerprint of differences between communities, not sequence-based taxonomic information about which bacterial lineages are affected by protist grazing.

Since Glücksman *et al*. (2010) other studies have investigated microbial predator-prey interactions, although we have not found any that consider the impact of protist grazing on bacterial communities at the level of individual lineages (operational taxonomic units; OTUs). Therefore this study was designed to extend the system studied by Glücksman *et al*. (2010) to

incorporate this level of resolution, using high throughput sequencing (HTS) to characterise grazed bacterial communities.

In the following chapters of this thesis, the effects of predators in experimental communities composed of bacteria and bacterivorous protists were investigated.

To assess evolution, the bacterium *Pseudomonas fluorescens* SBW25 was submitted to more or less complex predator communities and the prey's rapid evolution analysed as phenotypic changes according to physical defence adaptation. Specific phenotypes correlated with protist predation (Friman, Dupont, *et al.*, 2015). In a complementary study, I analysed genotypic adaptation of *P. fluorescens* in terms of gene expression, sequencing the bacterial transcriptome of prey grown under predation pressure and gene expression.

In parallel, predators also evolve along with their prey: in the race for survival, how do protist predators fare in comparison of their bacterial prey? A variety of ecologically and physiologically diverse bacterivorous protist species were individually fed with *P. fluorescens* SBW25 over thousands of prey generation, and their ability to survive after longer exposure to the bacterium was analysed.

Chapter 2 METHODS

In order to investigate the effect of predator type and diversity in structuring the prey community, and in eliciting response from bacterial prey, we used laboratory microcosms involving the bacterium *Pseudomonas fluorescens* SBW25 as prey to a variety of bacterivorous protistan predators.

General methods encountered throughout the following chapters are described here. Techniques and methodologies unique to a specific chapter are detailed at the said chapter.

2.1 Protozoa

Pure cercozoan cultures *Cercomonas pigra* (strain code CaSphII), *Cercomonas effusa* (Beaver-Creek), *Cercomonas paraglobosa* (19-3E), *Cercomonas braziliensis* (B13), *Paracercomonas saepenatans* (CA5HKv), *Paracercomonas minima* (SW2), *Paracercomonas oxoniensis* (Wa8), *Paracercomonas vonderheydeni* (NZ1-5c), *Eocercomonas uvella* (11-7E) and *Allapsa scotia* (Kv-Hf) were obtained from the Microbiology Laboratory collection at the Natural History Museum (London), where most experimental work has been carried for this project. Strains were stored in clear Parafilm-sealed Petri dishes with their natural bacterial communities in freshwater medium made of Volvic water (Danone) and a sterile wheat grain previously boiled in Volvic, serving as carbon source for the bacteria.

The bodonid *Bodo saltans* was first observed in cow faecal matter collected in October 2012 in Langton Matravers (Swanage, Dorset). In order to isolate the organism, a faecal-derived medium ("poo-tea") was created by incubating raw faecal material for 3 days in Volvic water. Solid material was removed and the medium autoclaved in a Boxer 400/150l at 121°C for 15min. A small amount of faecal matter was inoculated in 30ml of said medium in Petri dishes. Once micro-eukaryotes peak density was obtained, 10μl of culture were serially

diluted in 200μl, on a 96well plate. Visually apparent pure cultures were firstly transferred to “poo-tea” and sub-cultured into 0.1% LB medium.

Axenic cultures of the ciliates *Tetrahymena pyriformis* (CCAP #1630/1W), *Tetrahymena vorax* (CCAP #1630/3C), the flagellate *Chilomonas paramecium* (CCAP #977/2A), and the amoeba *Acanthamoeba polyphaga* (CCAP #1501/18) were obtained from the Culture Collection for Algae and Protozoa (CCAP). *Tetrahymena* species were maintained in Proteose Peptone Yeast Extract medium (PPY medium; 20g.L^{-1} Oxoid L85 proteose peptone, 2.5g.L^{-1} Oxoid L21 yeast extract); *Chilomonas paramecium* in Chilomonas Medium (1.0g.L^{-1} sodium acetate, 1.0g.L^{-1} Oxoid L29 “Lab Lemco” powder). *Acanthamoeba polyphaga* was maintained in Proteose Peptone Glucose medium (PPG medium; 15g.L^{-1} Oxoid L85 proteose peptone, 18g.L^{-1} D-glucose in 1L Page’s Amoeba saline medium (CCAP)).

Prior to experimental work, the selected flagellates were transferred to sterile 0.1% Luria-Bertani medium (15g.L^{-1} Oxoid L24 LB powder) and acclimated for seven days or until they reached visually high density. To each protist-containing microcosm, 1μl of high density *Pseudomonas fluorescens* SBW25 bacterium was added. In order to eliminate all other bacterial strains, this procedure was repeated every second day for two weeks.

In order to verify that protists were indeed growing in a large majority of *P. fluorescens* SBW25, aliquots of every culture were taken prior to experiments (chapters 5 and 6). DNA extraction was carried as described in section 2.3, and Sanger sequencing was carried at Natural History Museum sequencing facility. Clean sequences were BLASTed against GenBank and when returned *P. fluorescens,* protist cultures were maintained ready for experimental work. When sequences were not clean, or when BLAST returned strains other than SBW25, samples from which strains were originating were acclimated for seven extra days.

Protist cultures observations were made under a Nikon eclipse TS100 inverted trinocular microscope at magnification 200x unless specified otherwise.

Details for every strain used are specified table 2.1.

Table 2.1 Protist strains used across this study.

Protist	Strain Code	Description	Bacterial grazing (chapter 4)	Long-term co-evolution (chapter 5)	Competition (chapter 6)
Allapsa fimicola	G9	Length: 4-5.5μm. Cells travel very little, often remaining in single location.	✓	✗	✗
Allapsa scotia	Kv-Hf	Length: 3.5-5μm. Gliding rapid and fairly smooth; slight nodding of cell, and vibrations from flickering anterior flagellum (AF).	✓	✗	✗
Cercomonas braziliensis	B13	Length: 13-23μm. Cell movement rapid and direct when spindle-shaped, slower and probing otherwise. Metabolic, but less so when spindle-shaped.	✓	✗	✗
Cercomonas magna	IVY8c	Length: 18-38μm. Cell movement fast and direct. Not very metabolic.	✓	✗	✗
Cercomonas mtoleri	BuffaloH5	Length: 8-11μm. Cell movement medium to slow, often remaining in one location. Very metabolic.	✓	✗	✗
Cercomonas parincurva	IVY7a	Length: 18-24μm. Cell movement slow, occasionally direct, occasionally with many changes of direction. Very metabolic.	✓	✗	✗
Cercomonas ricae	IB3	Length: 10-15μm. Cell movement occasionally rapid. Extremely metabolic.	✓	✗	✗
Cercomonas sphagnicola	CASphI	Length: 13-35μm. Progress of gliding flagellate sedate rather than sluggish.	✓	✗	✗

Protist	Strain Code	Description	Bacterial grazing (chapter 4)	Long-term co-evolution (chapter 5)	Competition (chapter 6)
Cercomonas volcana	C18	Length: 6-10*μ*m. Cell movements slow, often stationery. Somewhat metabolic.	✓	✗	✗
Eocercomonas ramosa	C-80	Length: 5-15*μ*m. When not slow motile phase, stays in one location. Very metabolic.	✓	✗	✗
Eocercomonas uvella	11-7E	Length: 5-7*μ*m. Cell movements very slow. Not metabolic.	✓	✓	✗
Paracercomonas ambulans	W80	Length: 3-7*μ*m. Cell movement 'walking' motion of extending and retracting posterior pseudopodia, or remains in one location probing and turning. Very metabolic.	✓	✗	✗
Paracercomonas minima	SW2	Length: 5-9*μ*m. Cell often still and metabolic.. Very metabolic.	✓	✓	✓
Paracercomonas oxoniensis	WA8	Length: 8-16*μ*m. Cell movement occasionally relatively direct and rapid, though often remains in one location probing. Very metabolic.	✓	✗	✗
Paracercomonas producta	WA42	Length: 4-9*μ*m. Cell movement slow, progress slow, frequent interruptions to pause, probe and change directions. Very metabolic.	✓	✗	✗
Paracercomonas virgaria	C-71	Length: 6-14*μ*m. Usually remains in one location. Very metabolic.	✓	✗	✗

Protist	Strain Code	Description	Bacterial grazing (chapter 4)	Long-term co-evolution (chapter 5)	Competition (chapter 6)
Paracercomonas vonderheydeni	NZ1-5c	Length: 5-9*μ*m. Cell movement often quite rapid and direct, cell usually travelling. Very metabolic.	✓	✗	✗
Sandona dimutans	G11	Length: 2.5-6.5*μ*m. Jiggles in regular nodding motion; usually travelling although seen to stop, remain still and slowly sweep.	✓	✗	✗
Cercomonas pigra	CASphII	Length: 13-60*μ*m. Movement sluggish, slow, direct.	✗	✓	✓
Cercomonas effusa	Beaver-Creek	Length: 13-25*μ*m. Changes shape frequently.	✗	✓	✗
Cercomonas paraglobosa	19-3E	Length: 5-15*μ*m. Slow progression; very metabolic.	✗	✓	✓
Bodo saltans	-	Length: 4-5*μ*m. Attached to substratum with posterior flagellum, or free-swimming in helical path.	✗	✓	✓
Poterioocrhomonas sp.	-	Length: 5-15*μ*m. Cells appear either as sessile or swimming in circular paths. Can form dense cell clusters that derive.	✗	✓	✓
Tetrahymena pyriformis	TP	Length: 50-75*μ*m. Free-living, active fast swimmers. Becomes motionless but with flickering cilia in older cultures.	✗	✓	✓
Tetrahymena vorax	TV	Length: up to 120*μ*m. Free-living, swimming. Polymorphic, 'microstome' types feed on bacteria and detritus while 'macrostome' ones feed on other smaller eukaryotes.	✗	✓	✗

Protist	Strain Code	Description	Bacterial grazing (chapter 4)	Long-term co-evolution (chapter 5)	Competition (chapter 6)
Chilomonas paramecium	CP	Length: 20-40μm. Rapid free swimming.	✗	✓	✗
Acanthamoeba polyphaga	AP	Length: 15-35μm. Amoeboid with spiny projections. Rapidly develops cysts.	✗	✓	✓
Paraphysomonas sp.	PML5D	Length: 5-10μm. Mostly round, with helical swimming motion.	✗	✓	✗
Paracercomonas saepenatans	Ca5HKv	Length: 8-14μm. Sedate gliding, but can detach and become free swimming while keeping directional locomotion.	✗	✓	✓

2.2 Bacteria

All experiments were performed with the bacterium *Pseudomonas fluorescens* strain SBW25 as prey to protists.

Pseudomonas fluorescens (Flügge 1886, Migula 1895) is a genomically and physiologically diverse species capable to colonise soils, water and plant surfaces. Most strains are obligate aerobes and, except for some that can use NO^-_3 as electron acceptor, unable to grow under anaerobic conditions. When exposed to low wave UV radiation (ca 260nm), iron-depleted colonies fluoresce, giving the species its name (Palleroni, 1994).

The strain SBW25 is a saprophytic, gram-negative rod-shaped and plant growth promoter bacterium, firstly isolated in 1989 from the leaf surface of a sugar beet plant grown in Oxford (Rainey *et al.*, 1994). It harbours a 6,622,539 bp-long circular genome with a 60.5% GC content. Of the three *P. fluorescens* strains, SBW25 is the only one carrying a plasmid (pQBR103), acquired during a field release experiment (Rainey *et al.*, 1994; Bailey *et al.*, 1995).

Pseudomonas fluorescens SBW25 used for experiments was obtained from laboratory cultures maintained at Imperial College Silwood campus (Dr Thomas Bell Microbial Ecology laboratory). Stock cultures are kept at -80ºC in 20% glycerol, and an aliquot was thawed and suspended in fresh medium prior to every experiment.

2.3 DNA extractions and sequencing

In order to verify the nature and purity of protist strains, cercozoan-specific primers were used to amplify a 1,200-bp region of the 18S rRNA gene, which was then Sanger-sequenced to check that the sequence obtained matched that given with the species description (Bass et al 2009; Howe et al 2009) by Blastn searches of NCBI GenBank, and that no other sequence was obtained. Sequencing was done on an ABI 37XX sequencer at the Natural History Museum, London.

Sub-cultures of every protist strain were grown for five days in Petri dishes until reaching visual high-density levels. Cultures were then filtered on GF/F filters (Whatman). The filters, holding the eukaryote cells, were cut into pieces and put into 2ml PowerBead® tubes from the PowerSoil DNA isolation Kit (MO BIO) and then vigorously shaken three times for 45seconds, with 15s breaks, on a Precellys 24 (Bertin Technologies). DNA was extracted according to the kit's protocol.

PCR amplification was performed with the GoTaq DNA polymerse kit (M300, Promega), at 55°C over 35 cycles, primers V4F (5' CTGCCAGCMGCCGCCGCGTAA 3)' and V4R (3' TATTCTACTTAGTATCTT 5'). Sanger sequencing was held at the Sequencing Facility in the Natural History Museum (London) on a 3730xl DNA Analyser (Applied Biosystems).

2.4 Data treatment

Chapters 3 and 4 analyse data generated by collaborators. Sequencing data treatment generalities apply to both chapters, and more specific details are given in the Methods section of each chapter. Similarly, statistical analyses outlined here are generalised to all chapters, and specific relevant applications are explained within each chapter.

2.4.1 Sequences processing and taxonomic affiliation

Data obtained by 454 Sequencing with a Roche 454 FLX instrument for chapters 1&2 were analysed with the Quantitative Insight Into Microbial Ecology (QIIME) package (Caporaso *et al.*, 2010). The Roche 454 FLX generates a FASTA file (454reads.fna) containing the sequence for each read, and a quality-scores file (454reas.qual) with a quality-score (Q-score) for each base in each sequence in the FASTA file. A mapping file is created by the user, and contains all the information about the samples, such as the names of each sample, barcode sequences for each sample and a description of each sample. Those three files are either obtained independently, or can be recovered from a Standard Flowgram Format (SFF) file.

Quality filtering of barcode-based demultiplexed reads removed sequences with means quality score under 25, no primer or primer mismatches, as well as those with 6 or more nucleotide homopolymers. Initial length filtering excluded sequences under 200bp and over 1000bp. Analysis of output files histograms.txt – that contains the counts of sequences with a particular length – and the log file produced by the script determined whether filtering was coherent, and assessed accordingly (see specificities at chapters' methods).

The subsequent filtered library assigned sequences to de novo Operational Taxonomic Units (OTUs) with Uclust (Edgar, 2010); seed sequences (centroids) were selected according to Edgar (2010) and OTUs picked at a 97% similarity threshold and no reverse strand matching. For each OTU obtained, the most abundant sequence was selected as representative sequence and used for subsequent analysis. The OTU table generated, where columns correspond to samples and rows to OTUs, computes the number of times a sample appears in a particular OTU, for all OTUs.

Taxonomy assignments were obtained with BLASTn (Altschul *et al.*, 1990) searches of the representative sequences set against the database specified in each chapter - the Protist Ribosomal Reference database (PR2, Guillou *et al.*, 2013) and SILVA119 (Quast *et al.*, 2013) for 18S data; Greengenes v13-5 (DeSantis *et al.*, 2006) for 16S data.

2.4.2 Statistics

All downstream statistical analyses were performed using the R software version 2.15.1 or above (R Core development Team, 2005). Community ecology analyses were run under the package Vegan 2.0-8 (Oksanen *et al.*, 2007) in R, for all relevant chapters.

Analysis of similarity (ANOSIM) is a non-parametrical statistical method that tests for differences between two or more pre-defined groups (Clarke, 1993; Ramette, 2007). Based on permutations, it takes ranks of distances and compares it between and within groups, generating a R statistic that identifies

total dissimilarity between groups (R=1) – indicating grouping by category to be significant when the calculated p-value ≤ 0.05 – or no separation between groups (R=0). However, the method requires testing of within-group dispersion in order to check for the validity of significant results (Legendre and Legendre, 1998; Anderson and Walsh, 2013). ANOSIM analyses were based on Bray-Curtis community dissimilarity distances that take relative abundance into account, and differences considered statistically significant when p-value obtained was equal or inferior to 0.05. Analyses were carried with the vegan function anosim(), and within-group dispersion tested with function betadisp().

Similarity percentage analysis (SIMPER) identifies the variables that explain the similarity/dissimilarity observed between groups and that drive the observed pattern, based on Bray-Curtis dissimilarity matrix. The vegan function in R, simper(), gives the cumulative sum of the contribution percentage for each variable, in decreasing order of contribution.

Chapter 3 DIFFERENCES IN SOIL MICRO-EUKARYOTIC COMMUNITIES OVER SOIL PH GRADIENTS

The following chapter is the work of collaboration. R. I. Griffiths and colleagues designed the experiment, collected and treated the samples as indicated in the methods of this chapter. I obtained the eukaryotic dataset and performed all bioinformatics, statistical analyses and interpretation of results.

3.1 Overview

The Countryside Survey (www.countrysidesurvey.org.uk), a recent multi-sample assessment of bacterial communities across the full spectrum of UK soil types (fig 2.1) showed that bacterial community structure was strongly determined by soil variables such as soil pH (Griffiths *et al.*, 2011). Alpha diversity was positively related to pH, with greater diversity in soils of decreasing acidity. However, beta diversity was higher in acidic soils, possibly reflecting greater habitat heterogeneity in those samples. Here we produced and analysed a eukaryotic 454 sequencing 18S rDNA dataset, generated from a subset of the 2007 Countryside Survey samples. We seek to compare community structures across the three soil pH classes (low, medium, high), and contrast with patterns observed in bacterial communities. We also explored which taxonomic groups differ between different soil types, and at what level of taxonomic hierarchy differences are manifest.

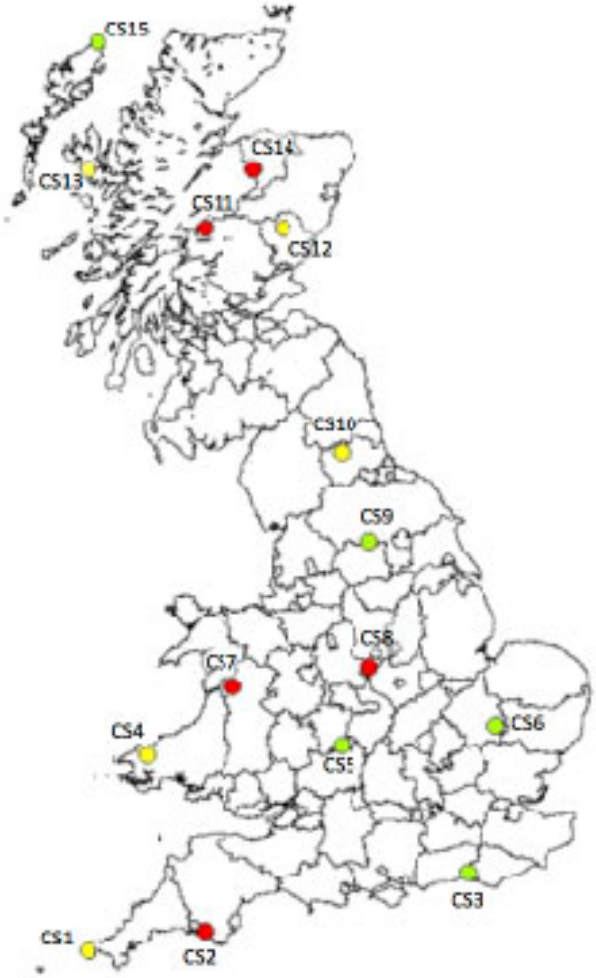


Figure 3.1 The countryside Survey 2007 sampling strategy across the UK. Yellow markers represent low pH samples, red ones indicate medium pH soil samples and green ones high pH samples.

Finally we examined the reliability of protist taxonomic assignments by comparing the performance of different databases. We used the databases to provide an in depth evaluation of some novel groups, which are highly represented in soil 18S libraries but whose evolutionary affiliations and relationships are yet to be resolved.

3.2 Methods

3.2.1 Sample details; DNA amplification and sequencing

Fifteen soil DNA samples were selected from the 2007 Countryside Survey (Griffiths *et al.*, 2011) representing 5 replicates each of low (pH 4.23 ±0.23), medium (pH 6.15 ±0.08) and high (pH 8.28 ±0.16) soil pH categories (fig 2.1). Primer sets EukA7F 5'-AACCTGGTTGATCCTGCCAGT-3' (Medlin *et al.*, 1988) and Euk570R 5'-GCTATTGGAGCTGGAATTAC-3' (Weekers *et al.*, 1994) were used to amplify a ~600bp product covering the V1 to V3 region of the 18S rRNA gene. Bacterial 16S rRNA genes were assessed using the primer sets 28F (GAGTTTGATCNTGGCTCAG) and 519R (GTNTTACNGCGGCKGCTG) as described in Dowd *et al.* (2008). Amplicons were sequenced in the forward direction by microbial tag-encoded pyrosequencing utilising a Roche 454 FLX instrument (Roche 454 Life Sciences, Branford, CT, USA).

3.2.2 Sequence processing and taxonomic affiliation

The resulting sequences obtained from 454 pyrosequencing were analysed using the QIIME software (Caporaso *et al.*, 2010). Data quality filtering removed sequences with length under 150bp, mean quality score lower than 25, those with no primer or with primer mismatches and with homopolymers over 6 nucleotides. Sample sequences were then de-multiplexed based on their barcode sequences. The subsequent library was assigned into Operational Taxonomic Units (OTU) with Uclust at 97% pairwise sequence similarity and no reverse strand matching. Representative sequences were picked up as the most abundant sequences in each OTU, and an OTU table was generated. Rarefaction of the OTU table was obtained with rarefy() function from the vegan package in R. Samples were rarefied to the level representing the lowest number of sequences across all samples, for both bacterial and eukaryotic OUT tables. Taxonomic assignments were obtained by BLASTn (Altschul *et al.*, 1990) searches of the representative set against the PR2 database (Protist Ribosomal Reference database (Guillou *et al.*, 2013), and the SILVA 119 database for 18S data (Quast *et al.*, 2013).

3.2.3 Phylogenetic analyses

Sequence alignments were generated using the e-ins-i algorithm of MAFFT alignment online (Katoh and Standley, 2013). Phylogenetic trees were built using RAxML-BlackBox (Stamatakis *et al.*, 2008) on the Cipres Science Gateway Portal (Miller *et al.*, 2010). The ML analyses used the generalised time-reversible (GTR) model with site-heterogeneous mixture model (CAT) approximation (all parameters estimated from the data); bootstrap values were mapped onto the tree with the highest likelihood value. After taxonomic affiliation, OTUs corresponding to metazoans and plants species were removed prior to further analyses. Where BLAST matches were below the thresholds specified (e-value <1e-30 and percentage identity 90%) a "No Blast Hit" report was produced. These were blasted separated against the NCBI GenBank nr/nt database and analysed phylogenetically in a RAxML tree of a selection of 500 eukaryotic 18S sequences including representatives of all supergroups as well as phylogenetically poorly resolved lineages, downloaded from GenBank and aligned (results not shown). Where taxonomic affiliation was then possible at some level of the taxonomic hierarchy the taxonomic affiliation results were amended. Highly divergent and/or taxonomically unresolved OTUs are shown in Table 3.1. In other cases the sequences were clearly not 18S rRNA genes, or were putatively chimeric/artefactual and were therefore removed.

Some OTUs were unassignable using the QIIME pipeline and returned "none" or "no blast hit". Manual re-blasting showed some of these to be closely related to characterised lineages in well-established groups and the taxon assignments duly amended.

3.2.4 Statistical analyses

Statistical analyses were carried on the R software version 2.15.1 (R Core development Team, 2005), under the Vegan 2.0-8 (Oksanen *et al.*, 2007) and FactoMineR 1.25 (Lê *et al.*, 2008) packages. Similarity percentage (SIMPER) and analysis of similarity (ANOSIM) analyses, using Bray-Curtis dissimilarities, were carried out in the R software, within Vegan.

3.3 Results

3.3.1 Data processing and OTU calling

45,505 quality-filtered sequences were analysed using the QIIME pipeline. After removing singleton and chimeric sequences these were clustered into 2566 operational taxonomic units (OTUs) across all 15 samples (fig 3.1). Following taxonomic affiliation based on the PR2 database, sample CS11 was found to be dominated by fungi (two OTUs accounting for >75% of sequence reads) so this sample was omitted from subsequent analyses. Metazoan and Streptophyta OTUs were also removed, leaving 2284 OTUs representing ten high level protistan taxa (at taxonomic level 2 – see below) (Fig. 3.2). Highly divergent OTUs are summarised in Table 3.1.

Taxonomy assignment outputs are presented as an informal taxonomic hierarchy of six or seven levels depending on the reference database used (SILVA119 and PR2 respectively). Level one (L1) specifies the eukaryotic domain and is not discussed further. Subsequent levels range from L2 (approximates to supergroup/phylum) to L6. Our analysis defines OTUs at a higher resolution than this; therefore, a single taxonomic profile may apply to more than one OTU. The most highly represented high ranking taxa were opisthokonts (mostly fungi), alveolates (mostly apicomplexans), and rhizarians (subphylum Filosa; Bass and Cavalier-Smith, 2004) (fig 2.2). The ten most abundant OTUs included five fungi (the coprophile *Lasidiobolus, Taphrina/Cryptococcus* (possible pathogen), *Bannoa/Sporobolomyces* (yeast associated with plant leaf surfaces), *Penicillium* (common soil saprotroph; sometimes plant pathogens), a divergent possibly parasitic apicomplexan (see below), the common soil flagellates *Eocercomonas*, *Sandona*, and *Oikomonas*, and an uncharacterised divergent variosean amoebozoan.

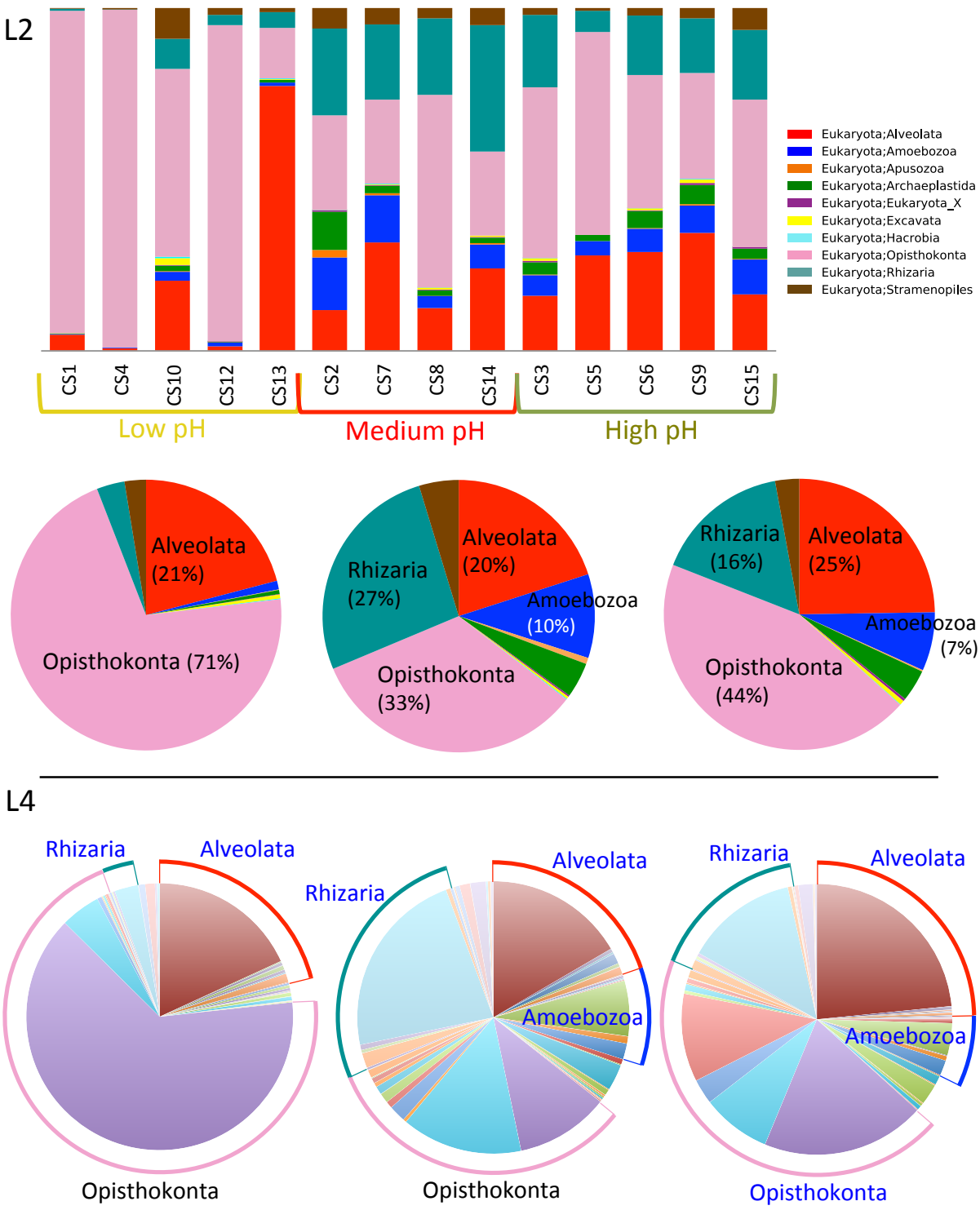


Figure 3.2 Soil microbial diversity comparisons according to pH, per sample (bars) and pH category (pie-charts), for both supergroup/phylum (L2) and class/order (L4) levels.

Table 3.1 The most divergent 18S rDNA sequences detected in this study.

Most of these were unassigned to any taxon by the QIIME procedure. The sequences are too short to be robustly resolved phylogenetically, however assignations in the Group column were estimated by their branching positions in a pan-eukaryote tree (see Methods). OTU 526 is probably chimeric. Most sequences in this table had 85% or less similarity to taxonomically characterised sequences in GenBank. In cases where this value is >85% the corresponding match to the most probable hit (in most cases an environmental sequence) was 90% or less. In one case (OTU 947) the best match was to a named specimen in GenBank

			Greatest similarity % match to GenBank sequences			
OTU	**Group**	**Closest named match on Genbank**	**Env.**	**Accession**	**Charact.**	**Accession**
1528	Cercozoa; Filosa	Placocista	96	FO181529	85	GQ144680
2308	Cercozoa; Filosa	Paulinella	93	JX456225	82	X81811
945	Cercozoa; Filosa	Gynmophrys (= Limnofila)	89	EU567223	88	FJ973365
920	Cercozoa; Endomyxa	Clathrina (env = Opisthokonta)	82	GQ844577	83	AM180960
1878	Cercozoa; Endomyxa	Metabolomonas	86	AB526173	85	HM536167
1190	Alveolata (see Fig. 4)	Gregarina	87	JN846840	84	JQ970325
334	Alveolata (see Fig. 4)	Gregarina	87	JN846840	86	JQ970325
1002	Alveolata (see Fig. 4)	Gregarina	78	JN846839	88	JQ970325
2298	Alveolata (see Fig. 4)	Gregarina	76	JN846839	75	JQ970325
529	Alveolata (see Fig. 4)	Apicomplex sp. 1	88	JN846840	87	KC890798
2360	Alveolata (see Fig. 4)	Apicomplex sp. 1	88	JN846840	88	KC890798
1689	Alveolata (see Fig. 4)	Diophrys	83	EF024740	82	EU267930

947	Alveolata (see Fig. 4)	Eimeria			89	GU479633
2554	Alveolata (see Fig. 4)	Colpodella	89	AB970393	88	AY234843
1031	Diplophrys/stramenopil	Amphifilidae sp.	78	EF023442	72	AB856528
1297	Diplophrys/stramenopil	Amphifilidae sp.	78	EF023442	72	AB856528
2291	Diplophrys/stramenopil	Amphifilidae sp.	78	EF023658	72	AB856528
328	Diplophrys/stramenopil	Amphifilidae sp.	76	KC454889	73	AB856528
1179	Diplophrys/stramenopil	Amphifilidae sp.	76	KC454889	73	AB856528
829	?	Pilobolus	91	AB970383	72	DQ211050
526	Excavata?	(Petalomonas)	77	JX069065	78	AF386635
459	Excavata?	Ichthyobodo	86	EU860484	79	KC208028
518	Excavata	Notosolenus	81	FO181403	81	KC990930
1021	?	Halichondria	87	HQ910364	81	KC899029
450	?	Halichondria	91	HQ910364	84	KC899029
630	?	Halichondria	90	HQ910364	84	KC899029
1510	Fungi	Alternaria	88	EF023366	87	KJ489375
1645	Fungi	Schizangiella	88	JX003447	88	AF368523
2122	Mesomycetozoea	Fabomonas	94	AB510393	82	JQ340335
505	Amoebozoa/Fungi	Monoblepharis	87	EF023424	88	KJ668082

51	Amoebozoa	Ceratiomyxella	88	AM409569	87	FJ544419
1824	Amoebozoa?	Glaucocystis	89	AM409569	87	X70803

3.3.2 Relationship between community structure and pH

At all taxonomic levels from L2 to OTU, there were significant differences in micro-eukaryote composition between low and high pH soils ($p < 0.05$). This was also the case for low and medium pH soils from levels L3 to OTU ($p < 0.05$; Table 3.2). There was no significant difference between medium and high pH soils at any taxonomic level. Even at phylum level (L2) the low pH soils have a distinct community structure, being dominated by opisthokonts (with a high representation of fungi), with markedly fewer rhizarian and amoebozoan OTUs than medium and high pH soils (fig 3.2). Lower in the taxonomic hierarchy (L4) differences in other groups in addition to fungi become more apparent. The low pH soils had a significantly lower total OTU count (447; average 146/sample) than medium and high pH (1247 (avg. 478) and 1314 (avg. 398) respectively), although note that high and low pH were represented by five samples and medium by four only. However, beta-diversity of the low pH soils was the highest (3.06) compared to medium (2.61) and high (2.64). Only 11 OTUs (2.5%) were detected in all low pH samples.

Low pH samples correlated positively with the first axis of a principal component analysis (fig 3.3), while medium and high pH ones correlated mostly with the negative first axis, so that samples belonging to low pH cluster together and apart from the rest. The first two axes of the analysis explained over 63% of the variance, although the projection of some samples is rather poor on those axes. Indeed, low pH samples were the strongest contributors for defining the first axis. High and medium pH samples correlate positively to different environmental variables, the strongest being bulk density (BD) and pH (ph_class); low pH samples were positively correlated to moisture and the first axis from a plant detrended correspondence analysis DCA1_2007, see Griffiths *et al.* (2011). All other variables, although significant, were more weakly correlated ($r^2 < 0.7$).

Table 3.2 ANOSIM comparisons between pH levels at different taxonomic levels (according to PR2 database

	pH comparison					
	L-M		M-H		L-H	
Level	R	p	R	p	R	p
L1						
L2	0.4375	0.053	0.1313	0.195	0.444	0.019
L3	0.45	0.044	0.2313	0.148	0.452	0.035
L4	0.45	0.0288	0.2313	0.114	0.452	0.024
L5	0.5438	0.021	0.3438	0.052	0.504	0.024
L6	0.5625	0.032	0.275	0.065	0.62	0.01
OTU	0.7188	0.021	0.05625	0.719	0.57	0.022

R-statistic (R) and p-values (p) for each pH level comparison are given (L: low pH; M: medium pH; H: high pH); micro-eukaryotic community composition between pH levels is significantly different when $p \leq 0.05$ (L-H all levels, L-M from taxonomic level 3).

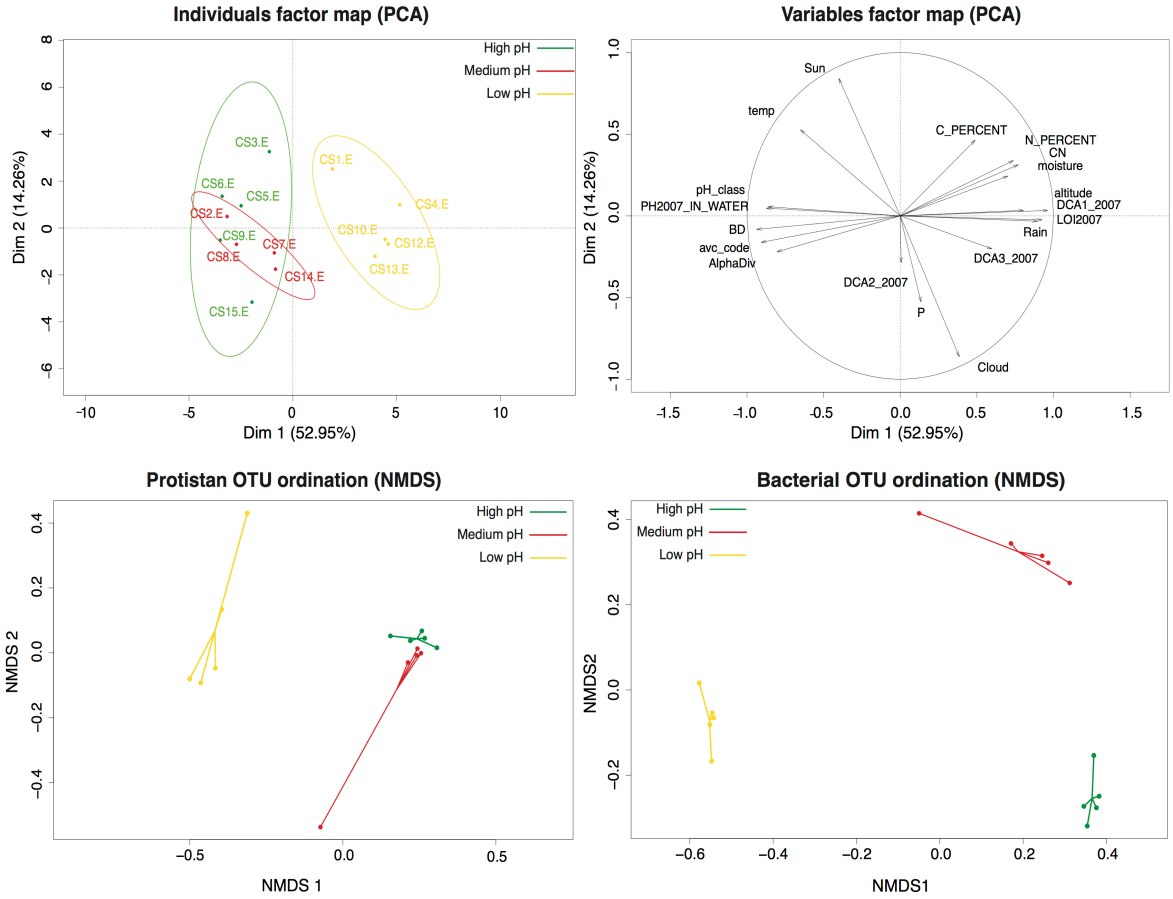


Figure 3.3 Relationships between soil variables and microbial communities:

A3. Individuals' factor map of a principal component analysis (PCA) groups samples belonging to high pH and medium pH soils together, but apart from low pH ones.

3b. The variables' factor map of the PCA correlates low pH samples positively to moisture (first axis), while medium and high pH ones correlate mostly with bulk density (BD).

3c,d. Bacterial and protistan OTUs ordination (respectively) according to pH groups. Although protestant OTUs cluster together according to the group they belong – high, medium or low – this is much clearer for the bacterial ones. Indeed, the latter separate clearly according to pH groups, while medium and high pH protist OTUs do not separate as clearly from each other

The SIMPER results in Table 3.3 show the 30 OTUs contributing most strongly to protistan community differences between each of the different pH levels. These explained 61% of the differences between medium and high pH and low to medium pH, and 54% of the differences between low and high pH. Of these 41 OTUs 41% are related to organisms with parasitic lifestyles, 20% related to those with pathogenic/symbiotic lifestyles associated with living plants, 20% to known saprotrophs, 17% bacterivores, and 5% photosynthetic autotrophs. The

(putatively) parasitic lineages were dominated by fungi and Apicomplexa (which together accounted for 31 of the 41 OTUs) plus one mesomycetozoean. Other high SIMPER-ranking taxa included Cercozoa (2 OTUs), chlorophytes (2), Amoebozoa (2) and one stramenopile OTU. Other parasites in the taxonomic assignments in addition to those shown in fig. 3.3 included other mesomycetozoeans, plasmodiophorids (Neuhauser *et al.*, 2014), and kinetoplastids (Ichthyobodo-relative).

Some OTUs near the top of the SIMPER table (Table 3.3) showed striking differences in occurrence between pH levels (i.e. contributing most strongly to community differences). For example, OTU 2542 (most closely matching *Archaeorhizomyces finlayi*, 98% identity) was strongly present in medium and high pH soils, but absent from all but one low pH sample, in which it was represented by only four sequence reads. Conversely, OTU 2440, also matching *Archaeorhizomyces finlayi* (92% identity) but with a different genotype, was more strongly represented in low pH samples. The sequences from the bacterivores *Sandona*, *Eocercomonas*, and the variosean amoeba lineage Mb5C were markedly more abundant in medium and high than low pH samples. The apicomplexan putative parasite OTUs 2376 and 2342 were also markedly more frequent in medium and high pH soils; 1787 was only found in high pH.

The taxonomic assignments showed a large number of OTUs (311) belonging to Alveolata. 59% of these grouped with parasitic Apicomplexa in a phylogenetic analysis, many of which were phylogenetically divergent (fig 2.4). The majority of the apicomplexan OTUs branched with terrestrial gregarines, but also included deep-branching relatives of lecudinid gregagrines, rhytidocystids, *Selenidium*, apicomonads, *Colponema*, and novel lineages. The rest of the alveolate OTUs grouped with perkinsids and ciliates.

Protist community differences across samples correlated with those of bacteria (Mantel test; r= 0.509, p=0.001). To visualise this we plotted the bacterial ordinations (non-metric multideimensional scaling (NMDS; fig 3.3) as well as the

pairwise correlations between the prokaryotic and eukaryotic OTUs (fig 3.5). The result showed blocks of positive and negative associations between bacterial and eukaryotic OTUs. Many of these likely reflect the shared constraints of soil pH. The figure also provides candidates for ecological interactions, including potential specialised parasite/host and predator/prey relationships.

Table 3.3 Similarity percentages analyses (SIMPER) of micro-eukaryote community differences between soil pH levels (Low-High (LH), Low-Medium (LM), Medium-High (MH)) and ranking of most influential species in the difference of compositions between pH levels.

The number following the pH level comparison code is the ranking of that OTU relevant to that comparison, e.g. LH1 is the OTU contributing most strongly to the community difference between low and high pH soils.

Soil type comparisons			OTU No.	Taxonomic affiliation	% identity to database sequence	Accession No.
LH1	LM4	MH2	2376	Alveolata, Apicomplexa, Gregarines_XX	98	EF024723
	LM2	MH3	2542	Opisthokonta, Fungi, Ascomycota, Archaeorhizomyces finlayi	98	JF836020
LH8	LM1	MH1	280	Opisthokonta, Fungi, Ascomycota, Pezizomycetes, Lasiobolus ciliatus	100	DQ646532
LH2	LM5	MH4	962	Rhizaria, Cercozoa, Glissomonadida, Sandonidae_X	100	EU646934
LH3	LM3	MH6	1801	Opisthokonta, Fungi, uncharacterised	100	EF023474
LH6	LM7		1787	Alveolata, Apicomplexa, Gregarines_XX	90	EF024723
LH4	LM6	MH11	147	Opisthokonta, Fungi, Ascomycota, Pezizomycotina, Penicillium sp.	100	GU190185
LH7	LM8	MH7	2342	Alveolata, Apicomplexa, Gregarines_XX	95	GQ462637
LH5	LM13	MH5	1052	Opisthokonta, Fungi, Basidiomycota, Agaricomycotina, Mrakia frigida	100	AB032665
LH9	LM9	MH9	38	Rhizaria, Cercozoa, Cercomonadida, Eocercomonas sp.	100	EF023536
LH10	LM11	MH8	612	Amoebozoa, Variosea, Mb5C-lineage	100	AB425950
	LM10	MH10	2197	Opisthokonta, Fungi, Ascomycota, Taphrinomycotina, Taphrina	92	AJ495835
LH11	LM12		163	Opisthokonta, Fungi, Chytridiomycota, Rhyzophidiales_X	99	GQ995433
LH13		MH13	1691	Opisthokonta, Fungi, Basidiomycota, Agaricomycotina, Cryptococcus	100	AB032627
LH14	LM15		2135	Opisthokonta, Fungi, Basidiomycota, Agaricomycotina, Catathelasma	98	DQ435811
	LM16	MH14	2440	Opisthokonta, Fungi, Ascomycota, Archaeorhizomyces finlayi	95	GQ404765

Soil comparisons		type	OTU No.	Taxonomic affiliation	% identity to database sequence	Accession No.
LH12	LM22	MH12	342	Alveolata, Apicomplexa, Gregarines_XX	98	EF024723
LH18		MH15	809	Rhizaria, Cercozoa, Plasmodiophorida, Polymyxa graminis	100	AF310898
LH15	LM14	MH22	2539	Alveolata, Apicomplexa, Gregarines_XX	100	EF024926
	LM20	MH17	216	Opisthokonta, Fungi, Ascomycota, Archaeorhizomyces finlayi	98	JF836020
LH16	LM23		1353	Alveolata, Apicomplexa, Gregarines_XX	93	EF024926
	LM19	MH20	2157	Opisthokonta, Fungi, Basidiomycota, Agaricomycotina, Camarophyllopsis	99	DQ444862
	LM21	MH19	554	Opisthokonta, Fungi, Ascomycota, Archaeorhizomyces finlayi	95	JF836020
LH23	LM18	MH21	2501	Rhizaria, Cercozoa, Cercomonadida, Paracercomonas sp.	100	AM114800
LH29	LM17	MH16	738	Amoebozoa,Tubulinea, Nolandellidae_X	99	EF023499
LH17	LM29	MH18	2412	Archaeplastida, Chlorophyceae, Oedocladium prescottii	100	DQ078298
LH19	LM24		1850	Alveolata, Apicomplexa, Gregarines_XX	97	EF024723
LH20		MH23	777	Opisthokonta, Fungi, Mortierellales, Mortierella sp.	100	EF023700
LH21		MH25	2187	Opisthokonta, Fungi, Basidiomycota, Agaricomycotina, Asterotremella	97	AB035586
LH22		MH24	2565	Alveolata, Ciliophora, Litostomatea, Enchelys polynucleata	99	DQ411861
LH24			2024	Alveolata, Apicomplexa, Gregarines, Ascogregarina taiwanensis	90	DQ462455
LH25			2194	Alveolata, Apicomplexa, Coccidia, Cryptosporidium serpentis	94	AF093500
LH26	LM25		1039	Opisthokonta, Fungi, Ascomycota, Pezizomycotina, Verticillium albo-	100	ABPE010014
	LM26		2069	Opisthokonta, Fungi, Cryptomycota_X	100	AB695466
	LM28	MH26	2321	Opisthokonta, Mesomycetozoa, Ichthyosporea, Ichthyophonida sp.	100	AJ130859

Soil comparisons		type	OTU No.	Taxonomic affiliation	% identity to database sequence	Accession No.
LH28	LM27		283	Opisthokonta, Fungi, Chytridiomycota, Chytridiomycotina,	98	DQ244005
		MH28	2276	Stramenopiles, Chrysophyceae-Synurophyceae, Clade-C_X	100	EF023425
LH27	LM30		2360	Alveolata, Apicomplexa, Gregarines_XX	88	KC890798
LH30		MH27	970	Archaeplastida, Chlorophyceae, Sphaeropleales_X	100	EF023843
		MH29	448	Opisthokonta, Fungi, Basidiomycota, Pucciniomycotina, Bannoa sp.	98	DQ631899
		MH30	422	Opisthokonta, Fungi, Basidiomycota, Agaricomycotina, Austropaxillus sp.	99	DQ534673

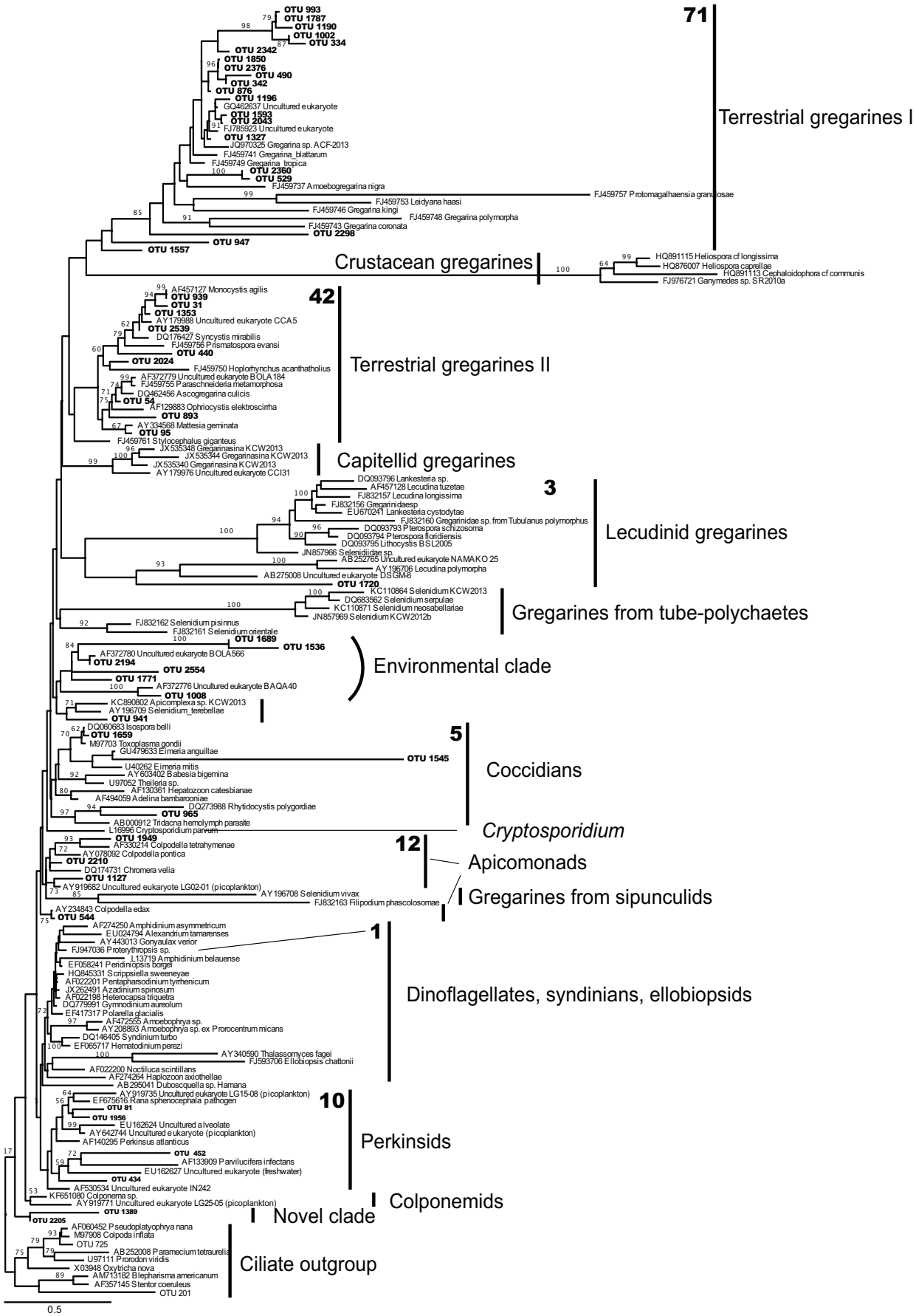

71
Terrestrial gregarines I
OTU 993
OTU 1787
OTU 1190
OTU 1002
OTU 334
OTU 2342
OTU 1850
OTU 2376
OTU 490
OTU 342
OTU 876
OTU 1196
GQ462637 Uncultured eukaryote
OTU 1593
OTU 2043
FJ785923 Uncultured eukaryote
OTU 1327
JQ970325 Gregarina sp. ACF-2013
FJ459741 Gregarina_blattarum
FJ459749 Gregarina_tropica
OTU 2360
OTU 529
FJ459737 Amoebogregarina nigra
FJ459757 Protomagalhaensia granulosae
FJ459753 Leidyana haasi
FJ459746 Gregarina kingi
FJ459748 Gregarina polymorpha
FJ459743 Gregarina coronata
OTU 2298
OTU 947
OTU 1557
Crustacean gregarines
HQ891115 Heliospora cf longissima
HQ876007 Heliospora caprellae
HQ891113 Cephaloidophora cf communis
FJ976721 Ganymedes sp. SR2010a
42
Terrestrial gregarines II
AF457127 Monocystis agilis
OTU 939
OTU 31
OTU 1353
AY179988 Uncultured eukaryote CCA5
OTU 2539
DQ176427 Syncystis mirabilis
FJ459756 Prismatospora evansi
OTU 440
OTU 2024
FJ459750 Hoplorhynchus acanthatholius
AF372779 Uncultured eukaryote BOLA184
FJ459755 Paraschneideria metamorphosa
DQ462456 Ascogregarina culicis
OTU 54
AF129883 Ophriocystis elektroscirrha
OTU 893
AY334568 Mattesia geminata
OTU 95
FJ459761 Stylocephalus giganteus
JX535348 Gregarinasina KCW2013
JX535344 Gregarinasina KCW2013
JX535340 Gregarinasina KCW2013
AY179976 Uncultured eukaryote CCI31
Capitellid gregarines
3
Lecudinid gregarines
DQ093796 Lankesteria sp.
AF457128 Lecudina tuzetae
FJ832157 Lecudina longissima
FJ832156 Gregarinidaesp
EU670241 Lankesteria cystodytae
FJ832160 Gregarinidae sp. from Tubulanus polymorphus
DQ093793 Pterospora schizosoma
DQ093794 Pterospora floridiensis
DQ093795 Lithocystis BSL2005
JN857966 Selenidiidae sp.
AB252755 Uncultured eukaryote NAMAKO 25
AY196706 Lecudina polymorpha
AB275008 Uncultured eukaryote DSGM-8
OTU 1720
KC110864 Selenidium KCW2013
DQ683562 Selenidium serpulae
KC110871 Selenidium neosabellariae
JN857969 Selenidium KCW2012b
Gregarines from tube-polychaetes
FJ832162 Selenidium pisinnus
FJ832161 Selenidium orientale
OTU 1689
OTU 1536
AF372780 Uncultured eukaryote BOLA566
OTU 2194
OTU 2554
OTU 1771
AF372776 Uncultured eukaryote BAQA40
OTU 1008
Environmental clade
KC890802 Apicomplexa sp. KCW2013
AY196709 Selenidium_terebellae
OTU 941
DQ060683 Isospora belli
OTU 1659
M97703 Toxoplasma gondii
GU479633 Eimeria anguillae
OTU 1545
5
Coccidians
U40262 Eimeria mitis
AY603402 Babesia bigemina
U97052 Theileria sp.
AF130361 Hepatozoon catesbianae
AF494059 Adelina bambarooniae
DQ273988 Rhytidocystis polygordiae
OTU 965
AB000912 Tridacna hemolymph parasite
L16996 Cryptosporidium parvum
Cryptosporidium
OTU 1949
AF330214 Colpodella tetrahymenae
12
Apicomonads
AY078092 Colpodella pontica
OTU 2210
DQ174731 Chromera velia
OTU 1127
AY919682 Uncultured eukaryote LG02-01 (picoplankton)
AY196708 Selenidium vivax
FJ832163 Filipodium phascolosomae
Gregarines from sipunculids
AY234843 Colpodella edax
OTU 544
AF274250 Amphidinium asymmetricum
EU024794 Alexandrium tamarenses
AY443013 Gonyaulax verior
FJ947036 Proterythropsis sp.
1
L13719 Amphidinium belauense
EF058241 Peridiniopsis borgei
HQ845331 Scrippsiella sweeneyae
AF022201 Pentapharsodinium tyrrhenicum
JX262491 Azadinium spinosum
AF022198 Heterocapsa triquetra
DQ779991 Gymnodinium aureolum
EF417317 Polarella glacialis
AF472555 Amoebophrya sp.
AY208893 Amoebophrya sp. ex Prorocentrum micans
DQ146405 Syndinium turbo
EF065717 Hematodinium perezi
AY340590 Thalassomyces fagei
FJ593706 Ellobiopsis chattonii
AF022200 Noctiluca scintillans
AF274264 Haplozoon axiothellae
AB295041 Duboscquella sp. Hamana
Dinoflagellates, syndinians, ellobiopsids
AY919735 Uncultured eukaryote LG15-08 (picoplankton)
EF675616 Rana sphenocephala pathogen
OTU 81
10
OTU 1956
EU162624 Uncultured alveolate
AY642744 Uncultured eukaryote (picoplankton)
AF140295 Perkinsus atlanticus
OTU 452
AF133909 Parvilucifera infectans
EU162627 Uncultured eukaryote (freshwater)
OTU 434
Perkinsids
AF530534 Uncultured eukaryote IN242
KF651080 Colponema sp.
AY919771 Uncultured eukaryote LG25-05 (picoplankton)
Colponemids
OTU 1389
Novel clade
OTU 2205
AF060452 Pseudoplatyophrya nana
M97908 Colpoda inflata
OTU 725
AB252008 Paramecium tetraurelia
U97111 Prorodon viridis
X03948 Oxytricha nova
AM713182 Blepharisma americanum
AF357145 Stentor coeruleus
OTU 201
Ciliate outgroup
0.5

Figure 3.1 Maximum Likelihood SSU rDNA phylogeny showing phylogenetic position of non-ciliate alveolates detected in this study. The parasitic apicomplexans occupy all branches above the dinoflagellates, syndinians, and ellobiopsids clade. Maximum Likelihood bootstrap values given where >60%. OTUs produced by this study shown in bold. Numbers associated with vertical lines marking groups to the right of the tree indicate the total number of OTUs called by the taxonomic annotation pipeline (see Methods); those with < 2% sequence from another OTU were omitted from the tree

3.3.3 Comparison of PR2 and SILVA taxonomy

We compared the taxonomic assignments produced using the same QIIME pipeline on the whole dataset with two 18S rDNA databases – SILVA 119 (Quast *et al.*, 2013) and PR2 (Guillou *et al.*, 2013). At taxon level 2, which should give the most informative high-level taxonomic overview, the profiles appeared quite different (fig 3.6). This partly resulted from different composition of high-level taxa between databases – for example Stramenopiles (3%), Rhizaria (16%), and Alveolata (24%) were shown separately in the PR2 analysis, but as the supergroup SAR (38%, grouping Stramenopiles, Alveolata and Rhizaria) in SILVA. However, the proportions of SAR and Opisthokonta in our results were different, depending on the database used, as some OTUs were accounted for in other groupings. Other differences result from some single lineages being represented at several taxonomic levels in Silva (e.g. BW-dinoclone28, *Colponema* sp. Peru, LG5-05, RT5iin25) because they are incompletely annotated across levels in the database.

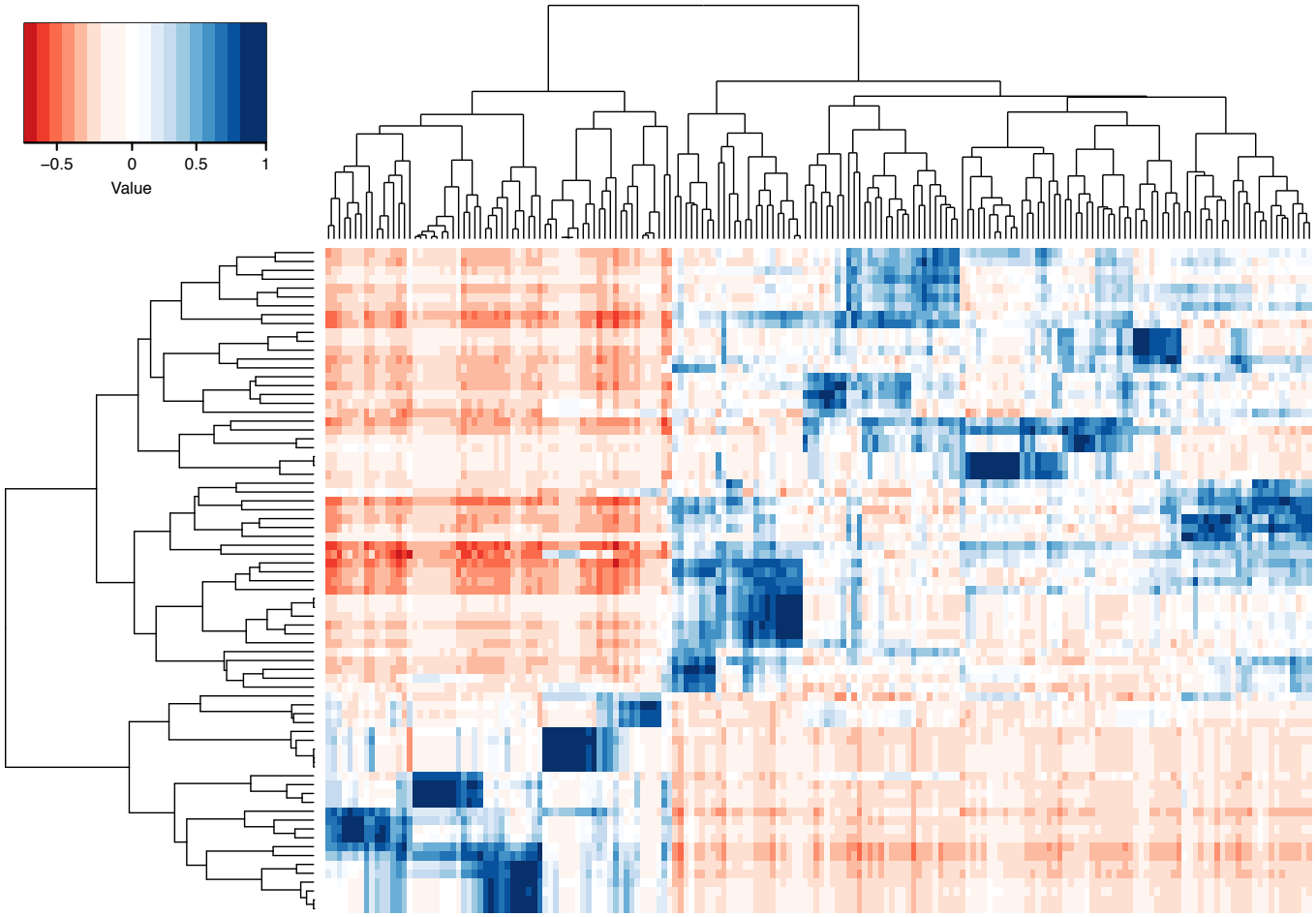


Figure 3.2 Bacterial-eukaryote correlation matrix. Shades of blue squares indicate positive correlation between bacterial (columns) and eukaryote (rows) OTUs, while red ones indicate negative correlations.

3.4 Discussion

We show that soil protist communities differ significantly between soils of different pH classes but to a lesser extent than bacterial communities analysed from the same samples. Low pH soils had markedly different micro-eukaryote assemblages from medium and high pH soils, whereas the latter categories were much more similar to each other. As for bacteria, protistan beta-diversity was also highest at low pH (Griffiths *et al.*, 2011). This might be a trivial expectation if protists were interacting solely with bacteria. However, only a small proportion of the protist taxa most characteristic of protist assemblage differences between the different pH levels were related to bacterivores, such as many cercozoan flagellates (Bass, Howe, *et al.*, 2009; Howe *et al.*, 2009; Howe, Bass, Scoble, *et al.*, 2011); the majority were related to parasites (of animals, plants, and other eukaryotic microbes), and protist and fungi otherwise known to interact with plant rhizospheres or phyllospheres (e.g. *Taphrina*, *Polymyxa*, *Archaeorhizomyces*; Table 3.2). Therefore, the ecological distribution of both above- and below-ground larger

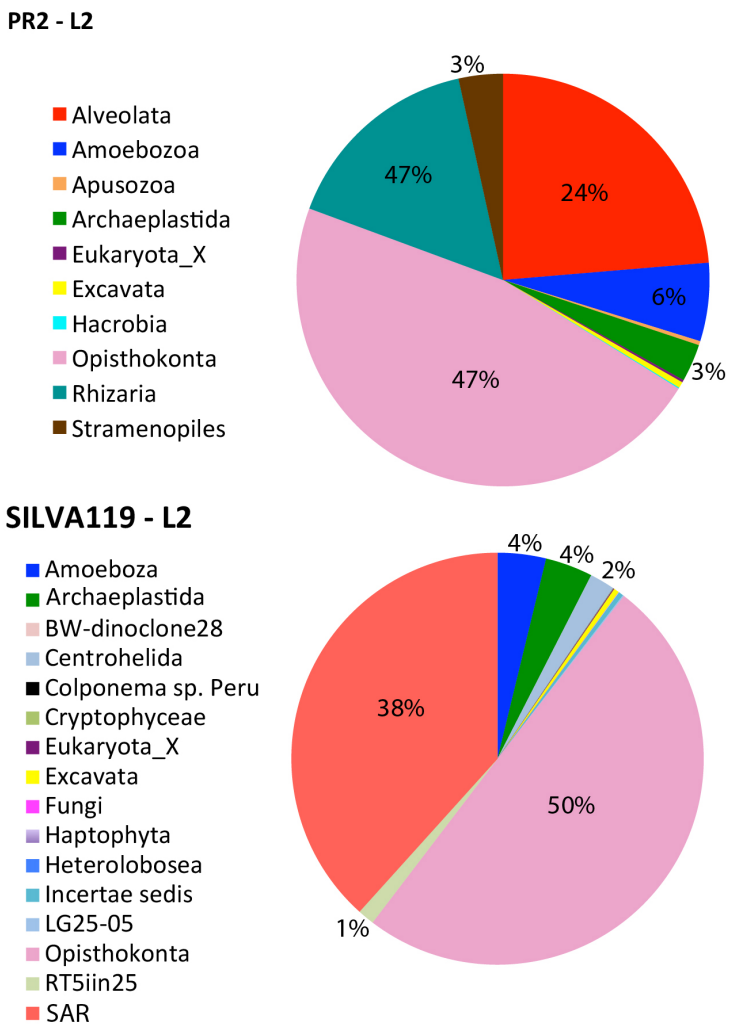


Figure 3.3 Taxonomic assignment comparisons between PR2 and Silva119 SSU rDNA databases

organisms appear to play strong roles in the determination of soil protist community structure, articulated by saprotrophy, coprophily, parasitism, and symbiosis (e.g. ectomycorrhizal fungi and rhizosphere-associated protists). Correlation analyses showed strong variation in co-occurrence between protistan and bacterial OTUs. Negative or positive correlations might simply be explained by shared preference of members of each domain for certain environmental conditions. However other interactions, for example preferential grazing of bacteria by protists (Chrzanowski and Simek, 1990; Glücksman *et al.*, 2010), antagonistic interactions such as chemical and morphological defence (Jürgens and Matz, 2002), pathogenicity, competition, etc., and synergistic interactions such as trophic cascades (Brussaard, 1977; Corno *et al.*, 2013) offer more biologically complex and powerful explanations for the related responses of both domains to pH level differences in their environment.

Detailed taxonomic interpretation of the OTUs revealed an interesting diversity of novel and recently characterized lineages, many of which appear to be soil specialists, perhaps important in biological processes specific to this habitat. For example, Archaeorhizomycetes, a recently described class of soil fungi (Rosling *et al.*, 2011), was represented by 29 OTUs, some of which contributed relatively strongly to micro-eukaryote assemblage differences between pH classes. At least some Archaeorhizomycetes are associated with plant roots (Rosling *et al.*, 2011). Our data suggests that distribution of members of this group is also influenced by pH, perhaps by being associated with plants characteristic of different soil types.

The summary of the most divergent valid OTUs in Table 2.1 shows that these belong to Cercozoa, many members of which are known to be important in soils (Bass, Chao, *et al.*, 2009; Bass, Howe, *et al.*, 2009; Howe, Bass, Chao, *et al.*, 2011), Alveolata – most of which are Apicomplexa, shown on fig. 3.4 and discussed more below, Stramenopiles (fig 3.7), a novel mesomycetozoan and putative kinetoplastid, fungi (unsurprisingly; Richards and Bass, 2005; Bass and Richards, 2011) and amoebozoans, which harbour a large and most uncharacterized diversity in soils (Berney *et al.*, 2015). One amoebozoan OTU,

affiliated to the lineage Mb-5c, is most closely related to *Arboramoeba*, a very recently described genus of large, network-forming variosean amoebae (Berney *et al.*, 2015), and which was a high-ranking discriminator between low and other pH categories in the SIMPER analysis (Table 3.3). Thirty other OTUs were also affiliated with *Arboramoeba*. When BLASTn-searched against the nt database in GenBank, many sequences in Table 2.1 and other taxonomically uncertain OTUs from this study returned environmental sequences generated by other soil eDNA studies, particularly Lehembre and colleagues (2013) and the taxonomically unfortunately mis-annotated study by Lesaulnier *et al.* (2008), strongly indicating that many protist lineages found preferentially or exclusively in soils, often phylogenetically distinct from currently characterized lineages, await discovery.

Particularly interesting are five mutually related OTUs, which our eukaryote-wide analysis (see Methods) show branch within Labyrinthulea, a class of often fungal-like stramenopiles, many of which are decomposers or parasites. More specifically they are related to two more environmental clades – one from soil, the other soil and freshwater, clustering at the base of the Amphifilidae clade, which apart from the marine *Amphifila marina* comprises all freshwater environmental sequences (Anderson and Cavalier-Smith, 2012; Takahashi *et al.*, 2014). The phylogenetic position of a representative three OTUs from this clade are shown on fig. 3.7; although the branch leading to these does not look that long Table 3.1 shows that these have only 76-78% sequence similarity with the next most closely related sequences in GenBank. This phylogenetic analysis suggests that these organisms may also be filopodial thecate amoebae but their actual phenotype and ecology can only be confirmed when they are directly observed. Other notable highly divergent OTUs in Table 1 include several with no discernable affiliation, some novel putative excavate sequences (OTUs 459, 518, 526?), and endomyxans (OTUs 920 & 1878), which may be plant or animal parasites or free-living filose/reticulose amoebae (Bass, Chao, *et al.*, 2009).

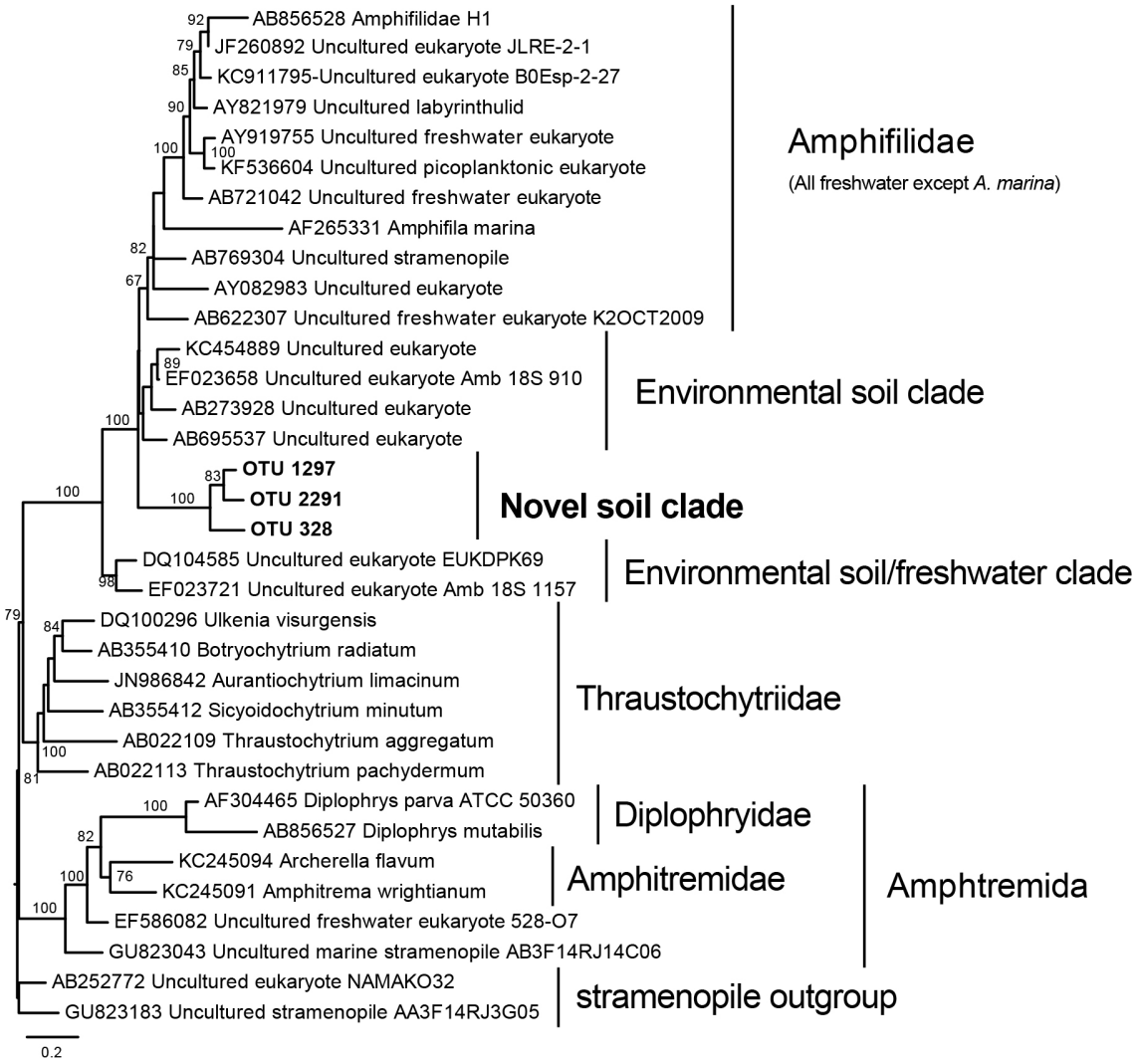


Figure 3.4 Maximum Likelihood SSU rDNA phylogeny of Amphifilidae, Thraustochytriidae, and Amphitremida (Labyrinthulea, Stramenopiles), showing novel divergent soil clade detected in this study (shown in bold). This clade contains two more sequences that were omitted from the analyses as they were significantly shorter than the others. Maximum Likelihood bootstrap values given where >75% or useful for interpretation.

Another group of interest that also accounted for many highly divergent OTUs was Apicomplexa (Table 2.1; fig 4.1), a phylum including a vast diversity of obligate parasites, including the causative agents of malaria, coccidiosis, cryptosporidiosis, and toxoplasmosis. Within Apicomplexa are the Gregarines, unicellular parasites of terrestrial, freshwater, and marine habitats, which form very widely distributed and resistant cysts (Rueckert *et al.*, 2011) and have the largest variation of rDNA evolution rates of any eukaryote group (Cavalier-Smith, 2014). Most apicomplexan diversity is thought to be marine (Rueckert *et*

al., 2010), but there is increasing evidence of their extreme (and often separate) diversity in soils (Bates *et al.*, 2013). We detected 147 gregarine OTUs, the majority of which grouped with (but often highly distinctly from) known terrestrial gregarines, which cluster in two clades (Rueckert *et al.*, 2011; Wakeman and Leander, 2013) that in some phylogenetic trees group together (Wakeman and Leander, 2012). Notably, apicomplexan OTUs dominate the diversity detected in sample CS13, including a high representation of OTU 2376, which fig 3.4 shows branches in the Terrestrial Gregarines I clade. Local concentrations of host individuals/material may account for the dominance of gregarines in this sample, which may also be the case to varying extents in other samples.

Apicomplexans provide a good illustration of cases where databases are very incomplete and/or taxonomic marker genes very divergent; for these a taxonomic annotation based on phylogenetic inference is far more informative than sequence affinity measures, and often essential. However, it is important to remember that the resolution of such analyses is limited due to the HTS read lengths. Nonetheless, to our knowledge fig. 3.4 is the first phylogenetic analysis of apicomplexan diversity detected as part of a soil HTS study.

Other OTUs putatively from parasites included plant root-infecting plasmodiophorids (27 OTUs), a group that includes the causative agents of clubroot in *Brassica* spp, powdery potato scab, and virus-vectoring parasites (Neuhauser *et al.*, 2014), labyrinthulids other than the divergent group discussed above (87 OTUs), Mesomycetozoea including 24 ichthyosporean OTUs, many fungi including 105 cryptomycotan and 106 chytrid OTUs, oomycetes and hyphochytrids (17 OTUs), and single-figure numbers of perkinsid relatives, metamonad gut symbionts, and kinetoplastids. Some further OTUs grouped within or were related to parasitic groups that could not be clearly affiliated, e.g. Holozoa (del Campo *et al.*, 2013) and Endomyxa (including the highly divergent OTUs 920 and 1878; Table 3.1), which includes predatory and parasitic amoebae (Hess *et al.*, 2012; Berney *et al.*, 2013) and ascetosporean invertebrate parasites (Hartikainen *et al.*, 2014) in addition to plasmodiophorids and their relatives. We also detected and expanded the

known diversity of an uncharacterised apicomplexan clade, predatory colpodellids, and novel diversity within perkinsids, which were also earlier thought to be exclusively marine but environmental diversity sequencing studies have also shown to be diverse in freshwater habitats (Bråte *et al.*, 2010). Our evidence suggests that these putative parasites are also frequent in soils, perhaps with small invertebrate or micro-eukaryote hosts. It is clear that parasite/symbiont diversity in soils is highly undersampled and its potential role as a reservoir of pathogens relevant to agriculture, silviculture and aquaculture understudied. The majority of the 'parasitic' OTUs sequenced were clearly distinct from named organisms, and often also from environmental sequences in GenBank (even if they didn't meet the criteria for inclusion in Table 3.1), and therefore inferring lifestyles of these novel and otherwise unknown organisms should remain tentative until more information is available.

In general, we cannot assume that all members of clades including known parasites are also parasitic, and inferring function based on environmental sequence data/phylogenetic position alone is risky unless the sequence identity of thoroughly characterised lineages is high and appropriately resolving. Groups partly comprising parasites may also include symbionts for which detrimental parasitism (pathology) has not been demonstrated (e.g. some plasmodiophorids), and other trophic strategies – saprotrophism being a frequent example (e.g. oomycetes, fungi, labyrinthulids). Similarly, groups known to be generally bacterivorous based on evidence from culture isolation studies (e.g. cercomonads and glissomonads; Bass, Howe, *et al.*, 2009; Howe *et al.*, 2009) may also contain lineages with quite different lifestyles (e.g. the algivorous viridiraptorid glissomonads; Hess and Melkonian, 2013).

In terms of general micro-eukaryotic soil diversity our results are in agreement with previous sequencing-based studies, showing a high proportion of fungi, alveolates, and rhizarians. Recent studies (Urich *et al.*, 2008; Geisen, Tveit, *et al.*, 2015) showed a similar diversity profile by sequencing the soil metatranscriptome, (a good indicator of active cells as opposed to dormant or dead forms), and also that parasitic lineages are more abundant than many had

assumed. For instance, strongly represented in Urich *et al.* (2008) data were the plasmodiophorid plant parasites, which are not conducive to culturing or cell isolation diversity studies and whose environmental diversity is much greater than host-oriented studies and those of economically important taxa would suggest (Neuhauser *et al.*, 2014). Alveolates were also well represented in all sequence based studies; Bates *et al.* (2013) noted that a significant proportion of their OTUs affiliated with Apicomplexa. Comparison of DNA and RNA-derived studies of soil apicomplexans will be important to distinguish between encysted and actively infecting forms (Rueckert *et al.*, 2011).

Even though short HTS-generated sequences have inherently low phylogenetic resolution, a combined approach to their taxonomic affiliation using both sequence similarity matching and phylogenetic analyses can provide more resolution and accuracy than blast-based methods alone. Further biological interpretation is possible via functional inference based on the resulting taxon profiles. We emphasise the need for phylogenetic moderation of raw taxon assignment outputs. It is important to acknowledge the significance of the percentage similarity between query and subject sequences. An 18S rDNA match of 95% or less (which dominate most HTS protistan diversity analyses) to a named database sequence is almost certainly not the species specified in the subject ID (if one is given) and may well not be the same genus. Below 85-90% assignments in the lower half of the taxonomic hierarchy become very doubtful. Here phylogenetic analyses can help, but are limited by both the signal carried by the OTU sequence fragment and database representation of related sequences. Databases themselves also powerfully influence perception of community structures. Their different outputs might misleadingly suggest strong biological differences between communities. The enduring lack of a generally adopted, comprehensive, and uniformly high quality taxonomic database for protists hinders the emergence of a body of data that can be consistently compared across studies.

Chapter 4 PROTIST GRAZING IMPACT ON BACTERIAL COMMUNITY COMPOSITION.

This chapter is the work of collaboration. Experiments were designed by Drs David Bass and Thomas Bell. Bacterial communities were obtained by Katja Lehman, who collected and treated all samples to my current knowledge.

I obtained the data from Thomas Bell and performed all bioinformatic and statistical analyses, as well as the interpretation of results.

4.1 Introduction

Natural communities involve many levels of organismal size, ranging from the smallest virus all the way to the biggest of elephants, with a variety of physiological properties. Meanwhile, every single species belonging to said ecosystem is in direct or indirect interaction with all others species composing the community. Even though one might say that all those species are different in all aspects, they still share a similar set of environmental conditions.
Environmental conditions are variable over time and space, and the local adaptation of every species; each with their own set of characteristics, to a specific location at a certain time identifies similarities between a diversity of organisms sharing a common environment. For every single species belonging to a complex adaptation to variation, by means of acquired traits inherited by the next generation, evolution at the level of the community eventually increases the level of diversity. However, understanding the processes that generate, but also maintain levels of diversity in a population is also one of ecological studies major objectives.

Diversity per se implies interactions between organisms, such as mutualism, competition or predation. A lot of effort has been put in the ecological effects of predation and competition, for example, in both theoretical and empirical studies (Murrell, 2005, 2010). Indeed, predation is among major drivers of population diversity: by affecting the mortality rate at a specific trophic level, predators define prey structure, which indirectly impacts on resource availability for co-existing prey as well as for competing predators. More recently, evolutionary processes – such as defence mechanisms developed to escape predation – have been shown to play an important role in population dynamics (Abrams and Matsuda, 1997; Abrams, 2005; Mougi, 2012a, 2012b).
Microbial systems given their small size, fast generation and rapid evolution represent the ideal controllable setup for the study of ecological interactions (Holyoak and Lawler, 2005). Therefore, experiments involving bacteria as prey to bacterivorous protists allow the validation of theory through observational studies. When grazing, protists exert a strong top-down control on bacterial

populations. In addition, generalist predation, or predation on selected taxa differently affect intraspecific prey interactions, impacting the microbial net production (biomass). On the other hand, the ability of bacteria to develop physiological and morphological defence mechanisms under eukaryotic grazing pressure affects both bacterial diversity and predatory eukaryotes' grazing ability and survival (Jousset, 2012).

Cercomonads and glissomonads (=sarcomonads) are highly diverse groups of heterotrophic flagellates, abundant in both aquatic and terrestrial environments. The first morphological observations of those groups by Dujardin (1841) and others have recently been revised by the use of more precise morphological and molecular techniques (Bass, Chao, *et al.*, 2009; Howe *et al.*, 2009). Sarcomonads include a variety of morphotypes both within and between taxa, although little is known about the potential functional diversity independent of morphological similarity or ecological preferences. Glücksman *et al.* (2010) analysed the effect of various phenotypic traits, as well as genetic relatedness, of a variety of (single) cercomonad and glissomonad strains on natural bacterial communities. Grazed bacterial communities tRFLP fingerprinting of post-grazing bacterial communities showed that protist cell size and "amoeboid-ness" (morphological cell-shape plasticity), and more weakly phylogenetic distance between the predator protist strains influenced prey community composition, thus prey selection. However, tRFLP profiles provide only a fingerprint of differences between communities, not sequence-based taxonomic information about which bacterial lineages are affected by protist grazing.

Since Glücksman et al. (2010) other studies have investigated microbial predator-prey interactions with respect to their importance in trophic network structure (Sintes and del Giorgio, 2014), top-down/bottom-up controls of bacteria by protists (Chow *et al.*, 2014), prey selectivity of bacteria by protists (Thurman *et al.*, 2010), or added environmental impact on the communities (Corno *et al.*, 2008; Ren *et al.*, 2013; Julia *et al.*, 2014). However, we have not found any that consider the impact of protist grazing on bacterial communities at the level of individual lineages (operational taxonomic units; OTUs). Therefore

this study was designed to extend the system studied by Glücksman et al (2010) to incorporate this level of resolution, using high throughput sequencing (HTS) to characterise grazed bacterial communities.

A variety of individual protist strains – cercomonads and glissomonads (sarcomonads) - were inoculated into bacterial communities from the rivers Lambourn, Wye, Pang and Kennet from the River Thames Basin in southern England, UK. We analysed the taxonomic consequences of protist predation on bacterial diversity, based on 454 sequencing of 16S rRNA genes, with or without protists, and in all four rivers, and compared these methods to tRFLP fingerprinting of bacterial communities.

4.2 Material and methods

Eighteen bacterivorous protist strains were incubated for ten days with diverse, natural bacterial communities from four different British rivers: the Pang, Lambourn, Kennet and Wye. The protists were cercozoan flagellates in the orders Cercomonadida and Glissomonadida: six species of *Paracercomonas*: *P. ambulans* (W80), *P. vonderheydeni* (NZ1-5c), *P. minima* (SW2), *P. oxoniensis* (Wa8), *P. producta* (WA42) and *P. virgaria* (C71); seven *Cercomonas* species: *C. braziliensis* (B13), *C. mtoleri* (BuffaloH5), *C. magna* (IVY8c), *C. volcana* (C18), *C. sphagnicola* (CASphI), *C. ricae* (IB3) and *C. parincurva* (IVY7A), and the *Eocercomonas* species *E. uvella* (11-7E) and *E. ramosa* (C80). Members of Glissomonadida were *Allapsa scotia* (Kv-Hf), *Allapsa fimicola* (G9) and *Sandona dimutans* (G11). Each protist strain and a control (free of eukaryotes) were incubated with every bacterial community in triplicates in 24-well plates at a constant temperature of 20^{o}C, with natural lighting. Of the eighteen initial protist strains, all glissomonads except *Allapsa scotia* (Kv-Hf); and the cercomonads *C. mtoleri* (BuffaloH5), *C. magna* (IVY8c), *C. volcana* (C18), *C. sphagnicola* (CASphI), *C. ricae* (IB3), *C. parincurva* (IVY7A), *E. uvella* (11-7E), *E. ramosa* (C80) and *P. virgaria* (C71) did not survive the length of the experiment or showed a contamination level making

them unusable (table 1). The remaining ten treatments were sequenced and analysed as follows.

4.2.1 Bacterial communities

The bacterial communities were obtained from the Rivers Lambourn, Kennet, Pang and Wye. In the lab, 2L of water from each river was filtered through a 1µ filter (Millipore, Watford, UK) and spun down at 3000 rpm for 15 minutes. The bacterial communities were then suspended in a growth medium that had been prepared by autoclaving water from the River Lambourn at Boxford (OS grid SU 42977 72065) augmented by autoclaved organic matter from Seacourt Stream at Wytham (OS grid SP 47189 09991). The communities were left to acclimatize at 16 °C until the start of the experiment. At the start of the experiment, the water samples were filtered through 0.22µ filters (Millipore, Watford, UK). These filters, holding the acclimatised bacterial communities, were immediately re-suspended in 25 µl of growth medium. The resuspended communities, together with the protozoan grazer strains were filled into 24-well plates with a capacity to hold 1ml per well (Millipore, Watford, UK). The overall amount of substrate was made up to 1ml per well by adding autoclaved water from the Lambourn. At the end of the experiment, the whole sample (1ml) was harvested in an Eppendorf tube and spun down at 14000rpm for 20 minutes. Thereafter, the supernatant was discarded and replaced with lysis buffer for extraction.

4.2.2 DNA amplification and sequencing

To extract DNA we added 300 µl of lysis buffer (100 mM NaCl, 500 mM Tris (pH 8), 10% (w/v) sodium dodecyl sulfate, 2 mg ml-1 proteinase K, 2 mg ml-1 lysing enzyme mix (both Sigma-Genosys, Gillingham, UK)) and 300 µl of NaH2PO4 (pH 8.0) to the pooled sample, incubated the DNA in a 55°C water bath for 30 min and mixed every 10 min., added 80 µl of prewarmed 10% CTAB solution (65°C), incubated in 65°C for 10 minutes, added 680µl chloroform:isoamyl alcohol (24:1 vol/vol). The tubes were centrifuged for 5 minutes at 14000 rpm. The aqueous top layer was aspirated into a new tube and the DNA precipitated by adding 300% (w/v) TE Buffer, pH 8.0 (10 mM TRIS-HCl, 1 mM EDTA, pH

8.0) and 200% (w/v) PEG/MgCl2 mix (30% (w/v) PEG 8000, 30 mM MgCl2), leaving the samples overnight at 5°C (Paithankar and Prasad 1991). We then centrifuged the replicates (12 per treatment) for 10 min at 14000 rpm, discarded the supernatant and washed the DNA pellets by adding 300 µl 70% chilled ethanol. We centrifuged the tubes again, discarded the ethanol and left the tubes to dry in a laminar flow cabinet until the ethanol had evaporated. We added 50 µl ultrapure water and left the DNA to resuspend for 1 h on the bench.

454 fusion PCR primers were constructed with A and B adaptors, with a 5-10 nucleotide identity tag (MID) incorporated into the forward primer. Forward primer: 454 forward adaptor [/5BioTEG/CCATCTCATCCCTGCGTGTCTCCGACTCAG]-MID-bacterial specific forward primer [GGTTAAGTCCSGYAACGA]. Reverse primer: bacterial specific reverse primer [AAGTCGTAACAAGGTANC]-454 reverse adaptor [CTGAGACTGCCAAGGCACACAGGGGATAGG/5BioTEG/]. Cycling conditions were 35 x [5 s at 95 C, 30 s at 55 C, 90 s at 72 C], 10 mins at 72 C. Equal numbers of each tag length were used across the sequencing run to minimise sequence homogeneity at equilvalent sites in the amplicon, which leads to low numbers of poor quality sequence reads due to over exposure of an particular nucleotide being read simultaneously. PCR reactions were carried out in quadruplicate on each replicate DNA sample, to minimise biases and aretfacts associated with individual reactions. The four reactions per sample were then pooled, quantified prior to sequencing. Sequencing was carried out on one 454 plate using Titanium chemistry by Eurofins (Germany). Sequences without both primer sequences (no mistmatches) were removed.

4.2.3 Sequence processing and bioinformatics

The resulting sequences obtained from 454 pyrosequencing were analysed using QIIME (Quantitative Insights Into Microbial Ecology; Caporaso *et al.*, 2010). Data quality filtering removed sequences with length under 200bp and over 1000bp, mean quality score lower than 25, no primer or with primer mismatches and with homopolymers over 6 nucleotides. Sample sequences were then de-multiplexed based on their barcode sequences. The subsequent

library was assigned into Operational Taxonomic Units (OTU) with Uclust at 97% pairwise sequence similarity and no reverse strand matching following data denoising. Representative sequences were picked up as the most abundant sequences in each OTU, and an OTU table was generated. Taxonomy assignments were obtained by BLASTn (Altschul *et al.*, 1990) searches against the Greengenes v13-5 database (DeSantis et al., 2006). Sequences from OTUs identified as "None" were separately blasted against the NCBI GenBank nr/nt database, in order to confirm their identities. Most of these didn't Blast to 16S rDNA genes; only one (OTU1902) was kept in the dataset, and was identified as an uncultured bacterium, possibly from the phylum Bacteroidetes.

4.2.4 Statistical analyses

Statistical analyses were carried out in R version 2.15.1 (R Core development Team, 2005), using the Vegan 2.0-8 (Oksanen *et al.*, 2007) packages. Similarity percentage (SIMPER) and ADONIS (PERMANOVA) analyses, using Bray-Curtis dissimilarities, were carried out in the R software, within Vegan.
Diversity indexes were calculated with the Vegan diversity function as Shannon entropy (natural base), and transformed into Effective Number of Species (ENS) according to Jost (2006).

Samples representing protist strains that didn't survive the whole length of the experiment were omitted from the analyses, as well as samples/replicates with fewer than 1000 sequences – although an exception was made for Control 2 in the river Kennet (230 sequences), in order to keep all three replicates in the analysis. On this basis the River Lambourn samples were not analysed further. Similarly, singletons (defined as OTUs represented by a total of five or fewer sequences within a single sample) were filtered out. This also removed any OTUs detected in only one library. In order to maximise statistical robustness, only protist treatments with three replicates were analysed. The protist strains differed strongly in their ability to persist on the different bacterial communities (Table 1). The final dataset consisted of 410 OTUs distributed across 36 samples belonging to rivers Kennet, Pang and Wye. We separated those into 3

datasets representing each river individually and constituted of four treatments: control, cercomonad strains SW2, B13 and NZ1-5c for both Wye and Kennet. Treatments analysed for the Pang were control, SW2, B13 and the glissomonad Kv-Hf (Table 1).

4.3 Results

Over 363,000 raw sequences covering 106 samples were initially quality-filtered with QIIME. The final analysed dataset consisted of 109,703 sequences contained in 410 OTUs distributed across 36 samples. Of those, 26,601 sequences were assigned to bacterial lines originated from the river Kennet, 31,679 to those obtained from the Pang and 51,423 to communities from the Wye, underlying the unequal sequence distribution between treatments.

Of the initial set of eighteen treatments, only nine (plus controls) where effectively sequenced, as a variety of predators died before the end of the experiment. *C. braziliensis* (B13) was the only *Cercomonas* representative that survived the whole length of the experiment (Fig 4.1). Of the three Glissomonads used, only *Allapsa scotia* (Kv-Hf) provided enough sequences in all treatments and replicates to be analysed. Amongst the *Paracercomonas* species only *P. virgaria* (C71) did not survive the whole length of the experiment, while *P. producta* (WA42), *P. ambulans* (W80) and *P. oxoniensis* (WA8), despite good survival rates and density, sequencing of bacteria grown with those strains didn't provide enough sequences in all replicates of all treatments, so they couldn't be included in the final analyses.

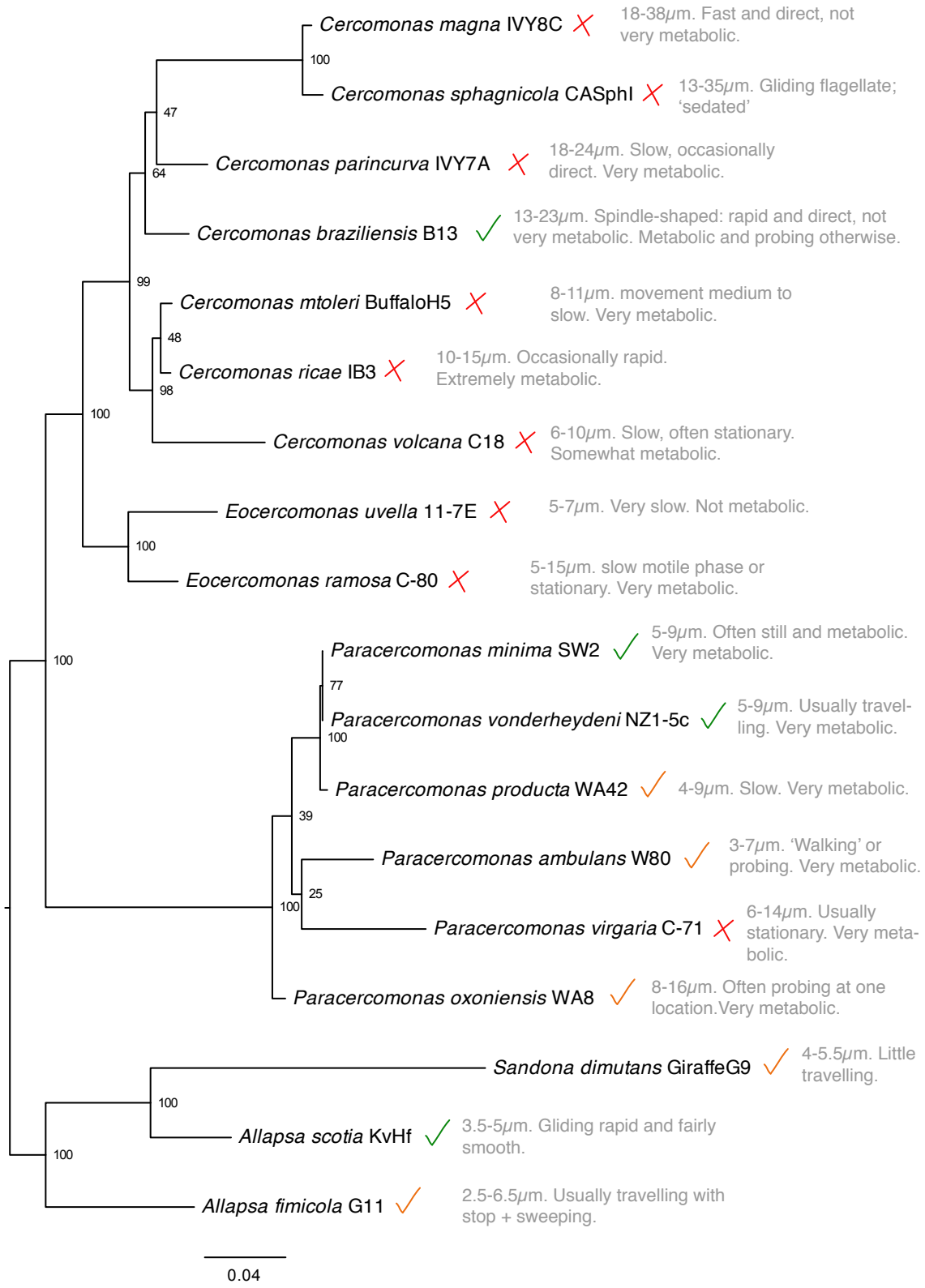


Figure 4 1 Phylogenetic tree of protist species used in the experiment. Red crosses indicate species that didn't survive the whole length of the experiment; green checks indicate sequenced and analysed species while orange ones indicate sequenced but not analysed species

Taxonomy assignments via QIIME are obtained across 'levels' that approximate a taxonomic rank; Level 1 (L1) corresponds to the prokaryotic domain, Level 2 (L2) approximates the supergroup/phylum organisation and so on until L6 that approximates the genus level. OTU levels indicate the species when possible. Across all samples, most abundant phyla (L2) were Proteobacteria (89%), Bacteroidetes (7%) and Actinobacteria (2%). Verrucomicrobia and Cyanobacteria equally represented 1% of all sequences (fig 4.2). At the order level of taxonomy (L4), most abundant taxa were represented by Burkholderiales (57%), Pseudomonadales (13%), Rhodocyclales (6%) Legionellales and Cytophagales (4%), Bdellovibrionales (3%) and Actinomycetales (3%). With the exception of Cytophagales (Bacteroidetes) and Actinomycetales (Actinobacteria), all other taxa belonged to the Proteobacteria. OTUs 2298 (34.79%), 337 (11.2%), 1130 (4.97%), 1560 (3.8%) and 343 (3.2%), corresponding respectively to *Limnohabitans* sp., *Pseudomonas stutzeri*, *Polynucleobacter* sp., an unknown Oxallobacteraceae and *Limnohabitans curvus*, were the overall most abundant OTUs.

4.3.1 Bacterial community composition and diversity

Proteobacteria and Bacteroidetes were the most abundant phyla in all three rivers. Proteobacteria represented 86% of all sequences in the Pang, 85% of sequences in the Kennet and up to 96% in the Wye (Fig 4.2). Overall, Bacteroidetes ranged from an average of 3% in the Wye to 9% in the Pang, and 7% in the Kennet. Actinobacteria represented 4% of all sequences in the Pang, but only 2% in the Kennet and 1% in the Wye. Verrucomicrobia and Cyanobacteria also represented 2% of all sequences from the Kennet, but only 1% or less in the Pang and Wye.

Detailed observation of bacterial communities in each river, and according to the protist grazer they were submitted to, revealed varied levels of diversity. Bacterial phyla distribution was more similar across samples in the river Wye, but also less diverse. It is interesting to note that communities incubated in the Kennet have ~1% of Firmicutes, while they represented less than 0.5% of the communities in all other treatments. At the lowest level of taxonomy obtained

with QIIME (L6 ~genus), it is clear to observe the dominance by *Limnohabitans* sp. in all samples, and more homogeneous distribution of genera in the Kennet – retracing the higher alpha-diversity in the latter (Fig 4.3).

Transformation of Shannon's entropy into effective number of species (ENS) allows comparing the difference in species numbers potentially present in each sample, as unities (Jost, 2006). Measured by ENS, samples derived from the Kennet were the most diverse, followed by the Wye, then Pang (Fig. 4.3). While the difference between Shannon indices of two samples can be difficult to interpret, ENS values are directly comparable as they are represented in a non-logarithmic scale (suppl. Table S4.1).

In every river treatment, community composition of controls was more similar to grazed than between themselves: the same bacterial lineages appeared at similar abundance (fig 4.2) levels between protist treatments of a same river than between controls of different rivers. Diversity indices, however, varied significantly between control replicates of a same river, for all three rivers (Hutcheson's t-test: calculated |t| > t=1.960 at 5% - supplementary Table S4.4) Interestingly though, average ENS of control communities in the Kennet and in the Wye were lower than SW2-grazed ones, while in the Pang, average ENS of non-grazed bacterial communities was higher than the SW2-grazed ones (Hutcheson's t-test for comparing Shannon indices; table S4.4). Within the Kennet, diversity (measured by ENS) of communities grazed by SW2 was the highest, and more similar to controls than to other treatments, while diversity of *P. braziliensis* (B13) grazed ones was the lowest. Similarly, bacterial diversity under *P. minima* (SW2) grazing pressure is more important in the river Wye, although variation between treatments and controls is lesser than observed in either Kennet or Pang: diversity levels are very similar for all treatments in the Wye. As for the Pang, bacterial communities grown under control and B13 conditions presented ~11 potential species, in contrast with the 8 for SW2 and *A. scotia* (Kv-Hf) treatments.

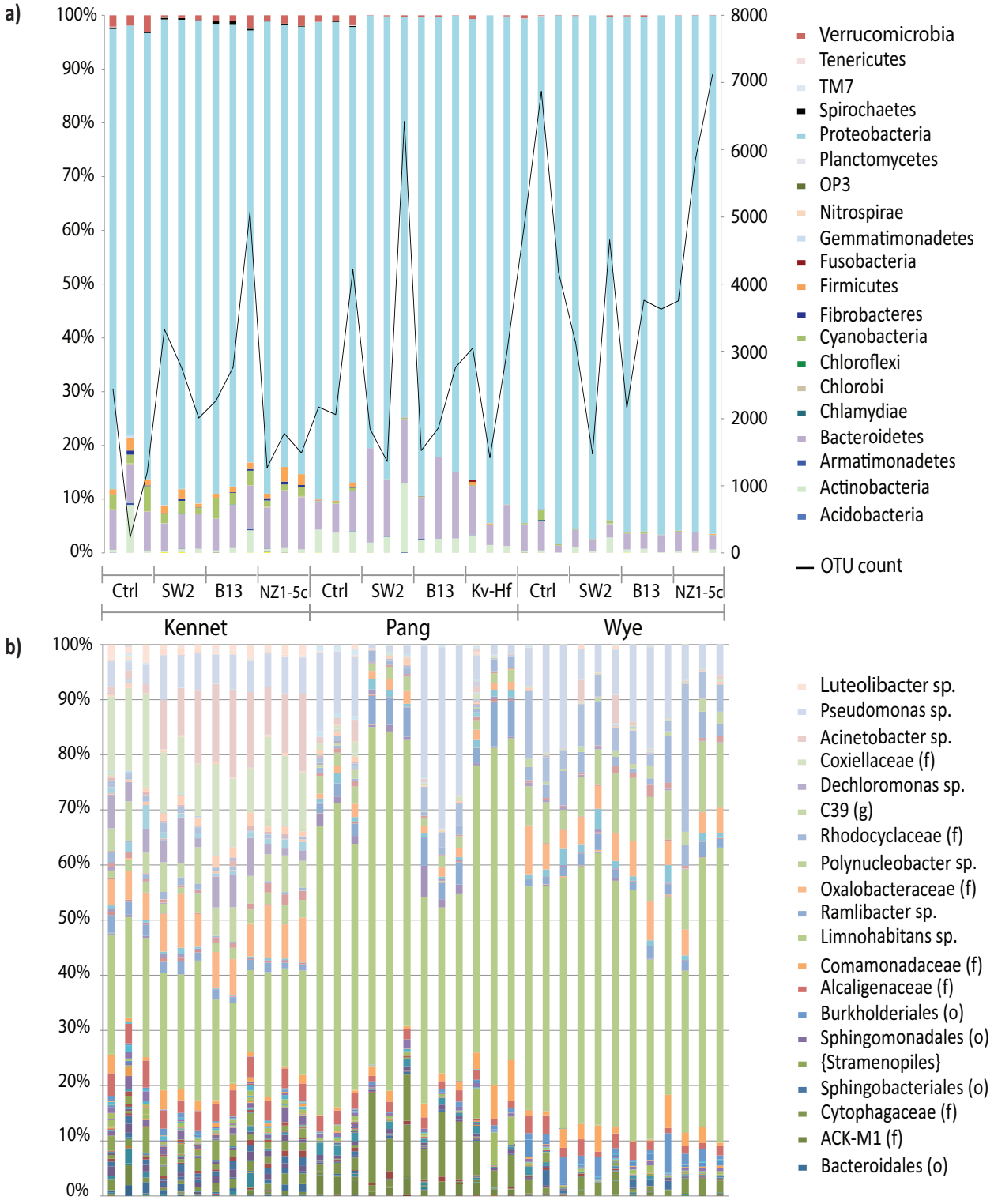


Figure 4.2 Taxonomic profiles for all samples in rivers Kennet, Pang and Wye.

a) Phyla (L2) distribution (bars, left axis) and OTU counts (right axis) for every sample; b) L6 (genus/species) level taxonomy composition in all samples. Genus or species levels (L6) are represented when possible. Lower levels are indicated otherwise (o: order, f: family, g: genus).

Note: Legend in fig 2.b) only illustrates most abundant and distinguishable taxa. Complete legend is provided in supplementary material.

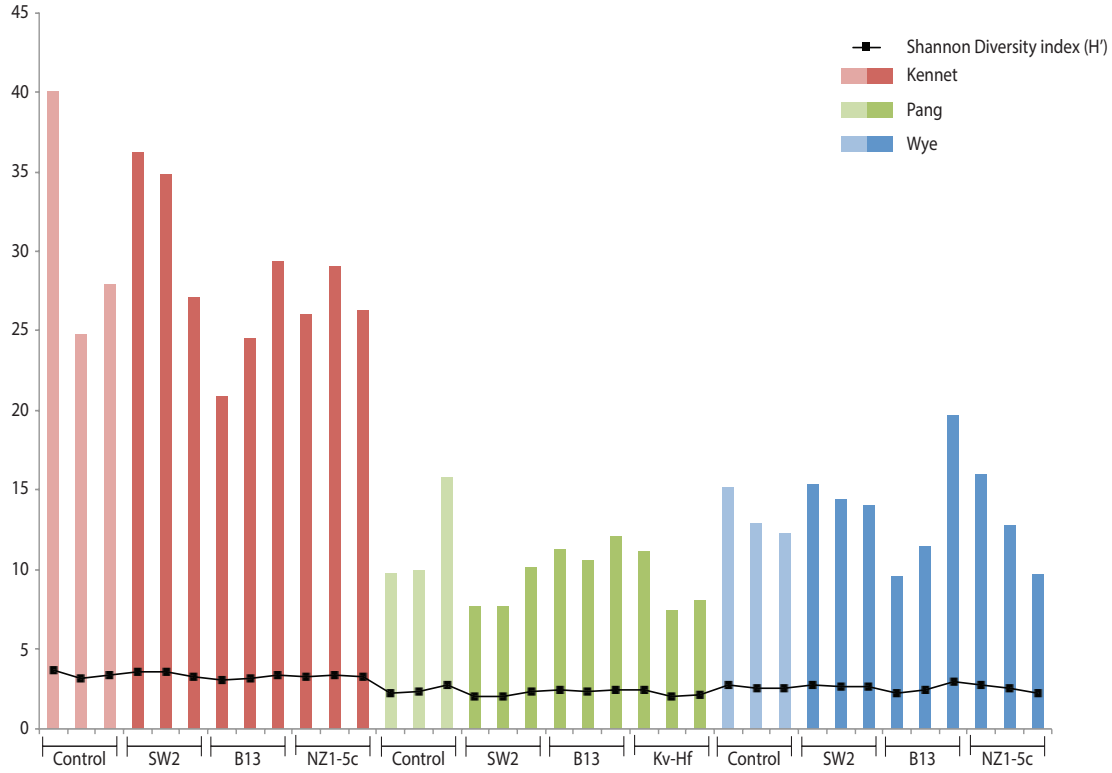


Figure 4.3 Effective number of species (ENS – bars) and Shannon diversity indices (black line) for every sample replicate in all three rivers Kennet (red), Pang (green) and Wye (blue). Lighter colour shades indicate control treatments, while darker shades indicate protist grazer treatments.

4.3.2 Community structure in relation to protistan predation

Non-metric multidimensional scaling (NMDS) ordination (Bray-Curtis similarities) of post-grazing bacterial communities and controls according to treatment as well as analysis of similarity plots – indicating the distance between similarities of replicates of every sample – for each river, are illustrated in figure 4.4; ellipses represent the 95% confidence interval around each group's (treatment) centroid. In the Kennet and Wye, triplicate replicates for each predator treatment clustered more closely together according to treatment, indicating higher similarity between them. But no clear pattern was observed in the Pang (Fig 1c). Analysis of similarity (ANOSIM; Table 4.3) indicated that only the grazer *Paracercomonas minima* (SW2) had a significant impact on the Kennet's bacterial community composition (ANOSIM Fig 4.4b), meaning that

communities grown under control conditions and under predators *Cercomonas braziliensis* (B13) and *Paracercomonas vonderheydeni* (NZ1-5c) have the same composition. Similarly, in the Pang, only bacterial communities grown in the presence *C. braziliensis* (B13) had a significantly different composition (ANOSIM p=0.033; Fig4.4c, d). Protist grazer strains had no significant impact on bacterial communities from the Wye. Interestingly, only in the Pang the control treatment (non-grazed bacterial community) had a significantly different composition (ANOSIM p=0.019) compared to other treatments. In both Kennet and Wye control treatments did not differ in composition (p=0.626 and p=0.831 respectively) compared to grazer treatments.

Table 4.1 Analysis of similarity (ANOSIM) of control bacterial community compositions from rivers Kennet (K), Pang (P) and Wye (W).

	R2	*p-value*
Controls$River	0.42742,	0.003
K vs P	0.69741,	0.102
K vs W	0.72845,	0.094
P vs W	0.66734,	0.113

Table 4.2 ANOSIM of bacterial community composition from rivers Kennet, Pang and Wye according to protist predator (strain). Every protist treatment was compared (pairwise) to all other treatments available within each river independently

	R2	*p-value*
Kennet$strain	0.2654	0.024
Pang$strain	0.3302	0.027
Wye$strain	0.2191	0.066

Table 4.3 ANOSIM of bacterial community composition from rivers Kennet, Pang and Wye, within each river between specific protist treatments (i.e. in the Kennet, SW2 bacterial communities vs bacterial communities grazed from all other treatments).

	KENNET		PANG		WYE	
Treatment	R2	*p-value*	R2	*p-value*	R2	*p-value*
SW2	0.33387	0.043	0.19859	0.084	0.22639	0.132
B13	0.10829	0.211	0.30596	0.033	0.11865	0.268
Kv-Hf*	-	-	0.00545	0.798	0.00238	0.904
NZ1-5c*	0.01986	0.656	-	-	-	-

Analysis of similarity (ANOSIM) at the OTU level indicated significant community composition differences according to predator (protist strain; p=0.001), and for the interaction river-strain (p=0.004). The river provenance of the bacterial communities on which the protists grazed did not significantly impact community composition (ns, p>0.05). However, due to the unequal protist representation per river – Kv-Hf absent from rivers Kennet and Wye, and NZ1-5c absent from Pang - those results must be carefully considered.

Non-grazed bacterial populations (controls) did not differ in composition between them at the OTU taxonomic level (ns, p>0.05) according to river provenance.

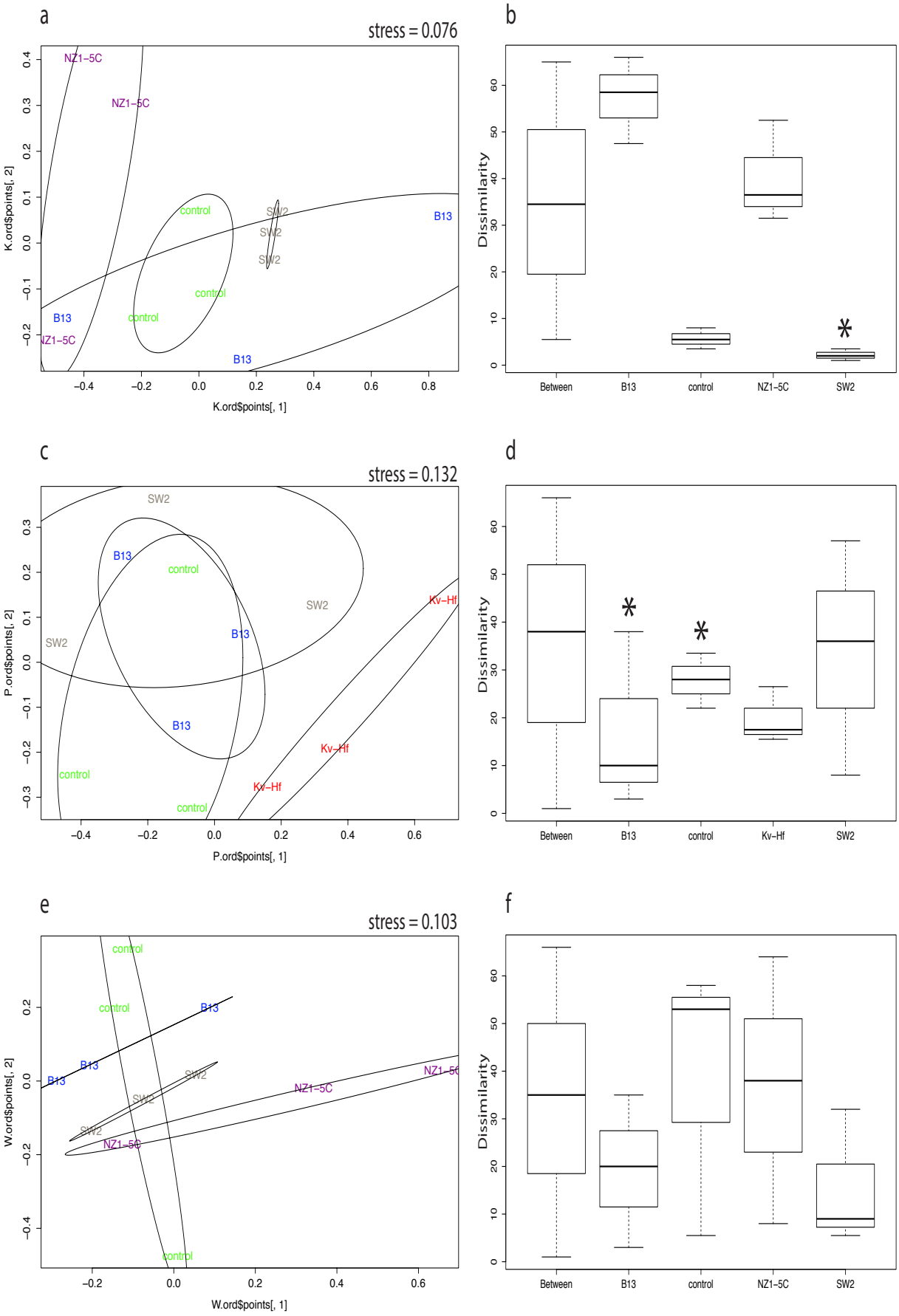

a
stress = 0.076
NZ1−5C
NZ1−5C
control
SW2
SW2
SW2
B13
control
B13
control
NZ1−5C
B13
K.ord$points[, 2]
K.ord$points[, 1]
b
Dissimilarity
Between
B13
control
NZ1−5C
SW2
c
stress = 0.132
SW2
B13
control
SW2
Kv−Hf
SW2
B13
B13
Kv−Hf
control
Kv−Hf
control
P.ord$points[, 2]
P.ord$points[, 1]
d
Dissimilarity
Between
B13
control
Kv−Hf
SW2
e
stress = 0.103
control
control
B13
B13
B13
SW2
SW2
SW2
NZ1−5C
NZ1−5C
control
W.ord$points[, 2]
W.ord$points[, 1]
f
Dissimilarity
Between
B13
control
NZ1−5C
SW2

Figure 4 4 Non-metric multidimensional scaling (NMDS) ordination and analyses of similarity (ANOSIM) boxplots of post-grazed bacterial communities originating from rivers Kennet (a, b), Pang (c, d) and Wye (e, f). Letters represent protist predator treatment and colours represent river provenance. Ellipses indicate 95% intervals around each group (treatment) centroid.

Stars indicate significant difference in bacterial community composition; 'between' indicates the dispersion of dissimilarities between classes (i.e. treatments), while the others are dissimilarities within classes (i.e. SW2 treatments in Kennet have very low dispersion)

4.3.3 Drivers of bacterial community composition

Similarity percentage analyses (SIMPER, table 3) revealed the ranking of the most important bacterial OTUs accounting for differences between the grazed bacterial communities in pairwise comparison of treatments (protist grazer). In the Pang, OTUs 2998, 337, 1240, 343 and 260 explained 70% of the differences caused by B13 as predator, and up to 12 OTUs explained differences between communities grown under Kv-Hf and SW2 (suppl. Table1). Similarly, 14 OTUs made up 70% of differences between control and NZ1-5c bacterial communities, while only 9 OTUs drove differences between control and SW2-grazed communities. In the Kennet, 16 OTUs explained 70% of the differences between control treatments and *C. braziliensis* (B13) treatments, and up to 29 OTUs were necessary to explain 70% of the difference between bacterial communities grown in control conditions vs. *P. vonderheydeni* (NZ1-5c). In all three rivers, OTU2998 (*Limnhoabitans* sp.) was the highest contributor to differences observed in the bacterial communities.

Frequency detection of OTUs in each treatment is also given table 3. This indicates the relative abundance of the OTU analysed, for a specific protist grazer treatment, amongst all OTUs in the considered river. In the Kennet, the first five most important OTUs (2998 – *Limnohabitans*, 59 – *Acinetobacter*, 463 – Coxiellaceae, 1560 Oxalobacteraceae, and 337 – *Pseudomonas stutzeri*) in driving the differences between communities subjected to different grazers were generally more abundant after grazing by *P. minima* (SW2) than by *C. braziliensis* (B13). When comparing communities grazed by B13 vs. those grazed by SW2, OTU2998 was almost four times more abundant after grazing by SW2 than by B13, and twice as abundant after grazing by *A. scotia* (NZ1-5c)

than B13. The corresponding post-grazing proportions of OTU 2998 in the Wye and Pang showed different relationships (e.g. OTU 2998 is nearly twice as highly represented after grazing by NZ1-5c in the Wye medium than by the other strains).

Table 4.4 Similarity percentage analyses (SIMPER) of post-grazing bacterial communities originating from the rivers Kennet, Pang and Wye according to protist predator pairwise comparison.

Pairwise comparisons indicate the ranking of OTUs' contributions in differences amongst the two compared communities (protist treatments). Detection frequencies illustrate under which protist predator the specified OTU is most frequently sequenced in each river.

KENNET		Detection frequency (%)				SIMPER pairwise comparisons		
		Control	SW2	B13	NZ1-5c	B13_NZ1-5c	B13_SW2	NZ1-5c_SW2
denovo2998	*Limnohabitans sp.*	3.58	10.13	3.09	7.13	1	2	1
denovo59	*Acinetobacter sp.*	0.20	4.37	2.85	3.58	2	1	2
denovo463	Coxiellaceae (genus unknown)	2.47	4.24	2.42	4.51	3	3	3
denovo1560	Oxalobacteraceae (genus unknown)	0.76	2.82	1.29	2.83	4	6	4
denovo337	*Pseudomonas stutzeri*	0.64	2.64	1.19	2.28	5	5	5
denovo2393	Rhodocyclaceae C39	0.71	2.40	1.15	2.00	6	7	6
denovo1130	*Polynucleobacter sp.*	0.26	1.12	0.43	1.04	9	12	8
denovo2261	Alcaligenaceae (genus unknown)	0.78	1.40	0.80	1.28	8	8	9
denovo1332	*Dechloromonas sp.*	0.94	1.75	1.27	0.61	7	4	7
denovo1956	Methylophilaceae (genus unknown)	0.10	0.18	0.21	0.62	-	20	-
denovo350	PSB-M-3 (Erysipelotrichaceae)	0.18	0.50	0.17	0.57	14	-	14
denovo1240	Cytophagaceae (genus unknown)	0.21	0.18	0.32	0.54	-	13	20
denovo2031	*Ramlibacter sp.*	0.49	0.87	0.32	0.49	10	10	10
denovo192	*Luteolibacter sp.*	0.00	0.25	0.24	0.26	-	14	23
denovo106	Bacteriovoracaceae (genus unknown)	0.13	0.30	0.15	0.35	-	-	-
denovo795	*Leadbetterella sp.*	0.00	0.14	0.01	0.25	-	-	-
denovo2752	*Pedobacter sp.*	0.05	0.52	0.11	0.27	11	-	12

denovo343	*Limnohabitans curvus*	0.25	0.62	0.15	0.29	12	16	11
denovo2312	HTCC2188	0.02	0.43	0.31	0.25	13	9	-
denovo257	Oxalobacteraceae (genus unknown)	0.04	0.36	0.03	0.09	15	-	15
denovo2892	*Bdellovibrio sp.*	0.13	0.39	0.04	0.12	16	-	13
denovo172	Comamonadaceae (genus unknown)	0.05	0.29	0.12	0.12	18	-	-
denovo2190	*Acinetobacter sp.*	0.01	0.28	0.10	0.16	19	-	-
denovo1868	*Propionibacterium acnes*	0.06	0.06	0.23	0.00	-	15	17
denovo1325	Bacteriovoracaceae (genus unknown)	0.20	0.05	0.02	0.07	-	17	-
denovo2082	Sphingomonadales family unknown)	0.23	0.09	0.05	0.05	-	18	-
denovo471	*Luteolibacter sp.*	0.20	0.02	0.05	0.08	-	19	-
denovo2179	*Dyadobacter sp.*	0.00	0.24	0.00	0.00	-	-	18
denovo136	*Rheinheimeria sp.*	0.08	0.07	0.04	0.26	-	-	-
denovo377	*Variovorax paradoxus*	0.07	0.22	0.02	0.05	-	-	16
denovo2599	Bacteroidales (family unknown)	0.15	0.21	0.04	0.29	-	-	19
denovo2406	Rhodocyclaceae (genus unknwon)	0.04	0.18	0.00	0.03	-	-	22
denovo701	*Luteolibacter sp.*	0.18	0.02	0.01	0.16	-	-	-
denovo351	Burkholderiales (family unknown)	0.04	0.17	0.01	0.03	-	-	-
denovo1769	*Rhodoferax sp.*	0.05	0.20	0.01	0.09	-	-	21
denovo627	*Fluviicola sp.*	0.15	0.03	0.19	0.10	-	-	24
denovo1482	Coxiellaceae (genus unknown)	0.10	0.22	0.05	0.18	-	-	25

PANG		Detection frequency (%)				SIMPER pairwise comparisons		
		Control	SW2	B13	Kv-Hf	B13_Kv-Hf	B13_SW2	Kv-Hf_SW2
denovo2998	*Limnohabitans sp.*	10.68	15.97	10.73	14.99	1	1	1
denovo337	*Pseudomonas stutzeri*	2.64	0.33	9.42	0.69	2	8	2
denovo1240	Cytophagaceae (genus unknown)	0.42	2.71	3.09	1.36	3	2	6
denovo343	*Limonhabitans curvus*	0.66	2.05	1.41	0.68	4	3	-
denovo260	*Rhodococcus sp.*	0.00	1.46	0.00	0.00	5	5	-
denovo2031	*Ramlibacter sp.*	0.49	1.63	1.46	1.47	6	7	5
denovo821	Actinomycetales (family unknown)	0.60	0.53	0.59	0.36	7	9	8
denovo59	*Acinetobacter sp.*	0.49	0.02	0.03	0.12	8	-	7
denovo1560	Oxalobacteraceae (genus unknown)	0.36	0.61	0.54	0.43	9	-	12
denovo1130	*Polynucleobacter sp.*	0.66	0.69	0.70	0.51	10	-	11
denovo727	*Agrobacterium sp.*	0.01	0.07	0.06	1.36	-	4	3
denovo172	Comamonadaceae (genus unknown)	0.03	0.22	0.41	1.19	-	6	4
denovo2052	*Bdellovibrio sp.*	0.00	0.00	1.24	0.01	-	-	-
denovo2772	*Methylibium sp.*	0.36	0.21	0.83	0.10	-	-	10
denovo2261	Alcaligenaceae (genus unknown)	0.63	0.52	0.70	0.42	-	-	9

WYE		Detection frequency (%)				SIMPER pairwise comparisons		
		Control	SW2	B13	NZ1-5c	B13_NZ1-5c	B13_SW2	NZ1-5c_SW2
denovo2998	*Limnohabitans sp.*	15.40	9.14	7.46	15.13	1	1	1
denovo2892	*Bdellovibrio sp.*	1.62	0.88	0.54	3.34	2	3	4
denovo1130	*Polynucleobacter sp.*	2.14	1.87	2.94	4.29	3	4	3
denovo337	*Pseudomonas stutzeri*	6.01	1.34	3.24	1.78	4	2	2
denovo1270	*Bdellovibrio sp.*	0.45	0.18	0.08	0.77	5	8	-
denovo1560	Oxalobacteraceae (genus unknown)	1.63	0.89	1.12	1.36	-	-	6
denovo2393	Rhodocyclaceae C39	0.20	0.21	0.30	0.80	9	7	-
denovo59	*Acinetobacter sp.*	0.00	0.69	0.03	0.07	-	-	-
denovo1240	Cytophagaceae (genus unknown)	0.32	0.31	0.14	0.82	6	11	-
denovo1629	Rhodocyclaceae (genus unknown)	0.29	0.29	0.38	0.67	7	-	-
denovo343	*Limnohabitans curvus*	1.33	0.82	0.68	1.20	8	6	8
denovo2406	Rhodocyclaceae (genus unknown)	0.87	0.16	0.27	0.15	-	5	5
denovo351	Burkholderiales (family unknown)	0.74	0.14	0.43	0.21	-	9	7
denovo2261	Alcaligenaceae (genus unknown)	1.02	0.53	0.39	0.50	-	10	9
denovo1416	Sphingobacteriales (family unknown)	0.46	0.00	0.28	0.00	-	-	10
denovo2665	Comamondaceae (genus unknown)	0.17	0.16	0.17	0.20	-	-	-

4.4 Discussion

We analysed the impact of protistan grazing on the taxonomical composition of different bacterial communities isolated from three different rivers. The presence of protistan predators partially influenced composition of communities by altering proportions and absence/presence of a few major bacterial strains. However, factors such as riverine source and bacterial and community life history also seem to drive diversity. Bacterial communities were obtained from four different rivers belonging to the Thames basin, although only three were analysed (see methods). All three sites are located within the Thames catchment, underlain by a chalk aquifer. The presence of major urban centres such as Swindon, Oxford, Slough, Reading and London has an important impact on water quality, from the number of sewage treatment works, as well as agriculture wash-off from the upstream area (Bowes *et al.*, 2012). All three rivers are categorised into either 'poor' (Pang, Wye) or 'moderate' (Kennet) overall environmental conditions by the Environment Agency in 2009, having suffered from different types of pollution, and implying higher nutrient content available for prokaryotes. Higher resource availability benefits faster-growing bacterial strains, although equally enhances predation (Corno and Jürgens, 2008). Furthermore, Bell *et al.*, (2010) showed that, across a productivity gradient, bacterial abundance increases while diversity increases then decreases slightly with increasing productivity. This demonstrates the prey's trade-off between predator resistance and competitive ability (Jürgens and Matz, 2002).

It is interesting to notice that from an initial set of eighteen treatments, only nine (plus controls) where effectively sequenced, as a variety of different predators died before the end of the experiment. As it appears, those that died were all members of the genus Cercomonas; *C. braziliensis* (B13) was the only representative that survived the whole length of the experiment (Fig 4.1). Similarly, of the three Glissomonads used, only *Allapsa scotia* (Kv-Hf) survived. Amongst the *Paracercomonas* species only *P. virgaria* (C71) did not survive the whole length of the experiment. What drove the extinction of certain protistan

predators and not others is difficult to tell, although a vague pattern can be observed. Some of the Cercomonas species tested have a large cell size, which are possibly more likely to feed on medium-sized bacteria. However, such bacteria are more likely to disappear when defensive adaptions evolve, shifting the global community size towards bigger and/or smaller cells (Salcher *et al.*, 2005; Corno *et al.*, 2008). With the exception of *C. magna*, a rapid moving but with little morphological plasticity ("amoeboidness"), all *Cercomonas* spp tested displayed slow directed movement but were more amoeboid then *C. magna*. Contrastingly, most surviving *Paracercomonas* spp had high morphological plasticity in addition to faster directed movement. It may be significant that *P. virgaria* (C-71), that didn't survive the experiment, 'usually remains in one location' (Bass et al 2009). In parallel, the most abundant bacteria are either motile, thus able to escape slow predators; or have fast growing rates leading to high bacterial densities that easily overtake the experimental system. However, both *P. oxoniensis* (WA8) and *P. producta* (WA42) harbour slow or almost static behaviours, and survived well, although sequencing of bacteria grown under those strains didn't provide enough sequences, so that the treatments couldn't be included in the final analyses.

Of all the bacterial communities we studied, those from the Kennet were the most diverse, even after grazing treatments, with an effective number of species (ENS) ranging from ~24 to 40 potential different species. This was approximately four times more than in the Pang ($7.4<ENS<15.7$) and 1.5X times more than in the Wye. A relatively equal proportion of major bacterial taxa represent diversity in the Kennet: even though *Limnohabitans* sp. dominated in all communities, the unidentified Coxiellaceae and Oxalobacteraceae, *Acinetobacter* sp., *Pseudomonas* sp. were equally abundant in every treatment but the control. Indeed, in all protist treatments, *Acinetobacter* sp. represented 8% to 15% of all species, but only 1.5% in controls. A pattern is less clear in the Pang, and differences between bacterial communities seem to be driven by different protist grazers according to treatment.

In contrast, control and grazed bacterial communities from the Pang – the ones with lowest Shannon indices/ENS numbers – were all clearly dominated by *Limnohabitans* sp. (33% under B13 and 59% under Kv-Hf), and except for *Pseudomonas* sp., *Ramlibacter* sp. and an unidentified Comamonadaceae, all other strains represent 2% or less of all present taxa. Interestingly, while absent in control treatments, an unidentified Cytophagaceae accounted for up to 16% of all taxa: a particular genus of Cytophagaceae, Sporocytophaga, is known to form resting microcysts, possibly more effective in avoiding protistan grazing than it's active counterpart. Likewise, *Pseudomonas* sp. accounted, in average, for 12% and 27% in control and B13-grazed communities respectively, but less than 2% in SW2 and Kv-Hf treatments, reinforcing the similarity in diversity between control & *P. braziliensis* (B13) treatments on one side, and *P. minima* (SW2) & *A. scotia* (Kv-Hf) on the other. In the river Wye, but for *Limnohabitans* sp. and *Pseudomonas* sp., unknown Rhodocyclaceae, *Polynucleobacter* sp., unknown Oxalobacteraceae, *Bdellovibrio* sp. – a predator of other bacteria – and unknown Oxalobacteraceae, other bacterial taxa represented a minimum proportion of the whole community. Contrary to what was observed in the other two rivers, bacterial communities originating from the Pang had comparable levels of diversity, with the same major taxa present in similar proportions across all treatments.

But for Coxiellaceae and Cytophaceae, all dominant taxa belong to the betaproteobacteria, a physiologically diverse group of gram-negative bacteria usually abundant in freshwater ecosystems (Glöckner *et al.*, 1999). These aspects indicate that differences between bacterial communities are strongly determined by the site/conditions from which they were sampled, and although the grazing protists in our experiment drove changes in diversity profiles by shifting relative detection frequencies of a range of OTUs, these differences were much less profound than the original differences between sites. It should be remembered that only a small proportion of the protist strains initially grown on the different bacterial communities survived in sufficient replicates for the analyses reported here. It is therefore possible that these are generalist, r-selected lineages with similar grazing preferences, and that more

specialist/selective feeders, which potentially could have exerted a wider range of effects on the bacterial communities were not considered because our methodological approach selected against them.

Protists are known to select bacteria within a limited size range, easy to ingest and containing enough nutrient to sustain growth. This results in the shifting of the cell size in prokaryotic communities towards much larger, difficult to ingest cells and very small, with lower encounter rates and nutrition value ones. Interestingly, one of the most abundant species (OTU2998, *Limnohabitans* sp.), not only has an important growth rate, it also ranges in cell size from very small 0.4*μ*m cocci for certain species, to 5*μ*m long curved bacilli (Kasalický *et al.*, 2013; Šimek *et al.*, 2013). Latter sizes are comparable to those of smaller protists such as *P. vonderheydeni*, *P. minima* or *A. scotia*. Furthermore, high grazing pressure triggers defence mechanisms from the prey, such as the formation of clumps or biofilms – a strategy adapted by *Pseudomonas* and *Acinetobacter* species - too large to be ingested whole while making the access to individual bacterial cells difficult (Hahn and Höfle, 2001; Blom *et al.*, 2010); or the development of grazing-resistant morphotypes such as flocks and filaments. In addition to community size structure, protist taxonomic relatedness, predator nutritional state, prey motility, cell surface physicochemical properties and even toxicity have been shown to impact microbial predator-prey interactions (Montagnes *et al.*, 2008; Bell *et al.*, 2010; Glücksman *et al.*, 2010; Meunier *et al.*, 2012).

Communities originating from all three rivers presented a few common major OTUs (table 4.3.), although none of those driving 70% of differences between Kennet, Pang and Wye were in common between both The Pang and the Wye (Fig 4.5). As a matter of act, the latter observed only four unique major OTUs, underlying how divergent bacterial communities are, most likely as result of both original characteristics and microcosm-related differentiation. Indeed, it is important to notice that microcosms were all set with water obtained from the Lambourn (see this chapter's methods). Even though the medium created from that was autoclaved, nutrients already existing could have affected the bacterial

communities in all three microcosms in similar ways. Even if bacterial communities from all four rivers were originally (very) different, providing them independently with the same set of nutrients (medium created from the Lambourn) could have strongly profit the same bacterial species in all different microcosms, thus rendering diversity differences lesser than what expected in their original environments. In order to test this normalisation of the communities, incubations of each bacterial community should be carried with media obtained from the rivers they were taken from. The effect of river nutrient selection versus microcosm differentiation can than be tested.

Furthermore, protist strains were obtained from laboratory cultures grown with their own bacterial communities. Protists cultures are usually isolated from a certain environment, and then kept in microcosms with their own set of bacteria, which they ingest regularly (the bacterial level is maintained thanks to the nutrient release of a boiled grain – see chapter 2). The addition of protists to the microcosms of study most certainly results in the addition of foreign bacteria: the ones present in the medium from which eukaryotes were sampled as well as the ones composing the protists bacterial consortium (gut, eventually on cilia, etc). These could represent highly competitive species, impacting the original community we wish to study. In addition, the fact that these bacteria have been exposed to protist grazing for over thousands of generations could also mean that they are highly specialised in terms of defence. So when predators are presented with a new bacterial community, they likely would graze on naïve bacteria, unable to defend as well as the ones from the protist cultures. This could have an important impact on the final bacterial community (sequenced): is it really an image of the community isolated from the river or a portrait of highly defensive species living in protist coexistence?

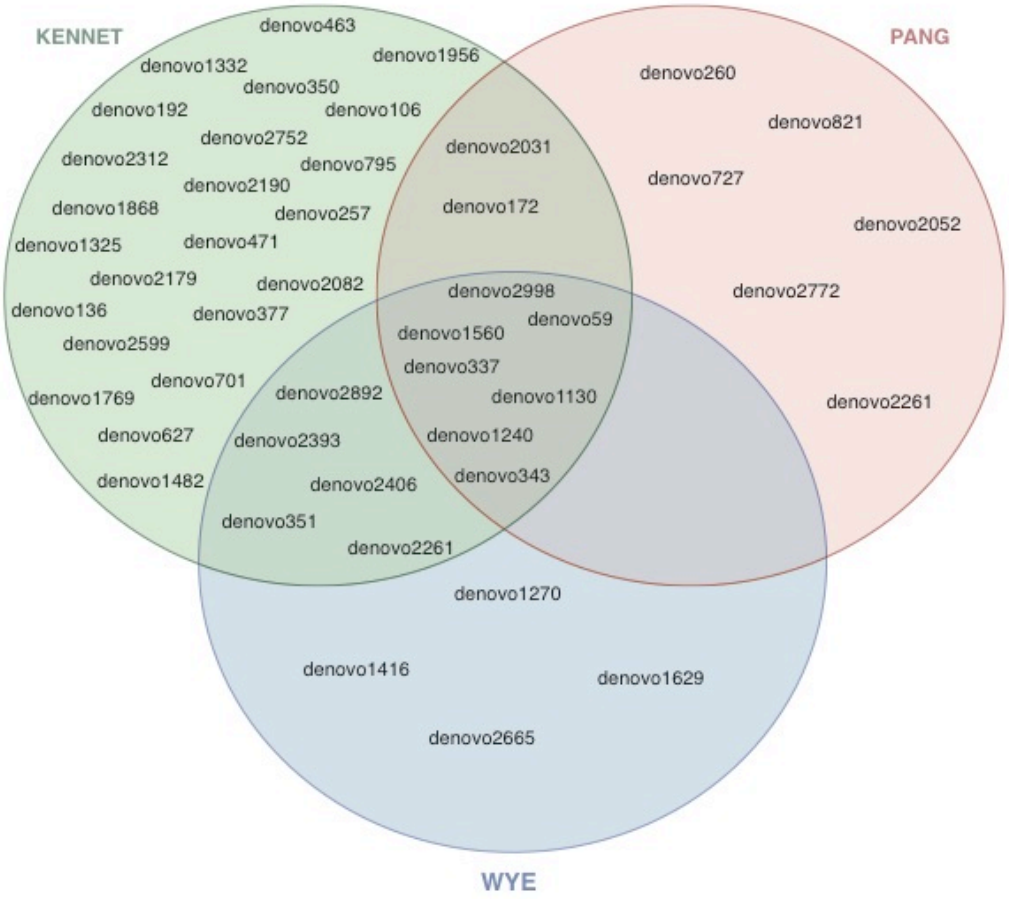


Figure 4.5 Distribution of OTUs driving 70% of the differences between communities in the Kennet, Pang and Wye (SIMPER).

Only the small cercozoan *Paracercomonas minima* SW2 significantly impacted the composition of communities originated from the river Kennet, where communities grown under SW2 have a significantly higher diversity when compared to all other treatments (Hutcheson's t-test). No other predator had a significant effect in shaping the overall bacterial community although shifts in frequency of detection of individual OTUs were strongly suggestive of strain-specific effects. Differences are firstly driven by the highly abundant OTU2998, representing *Limnohabitans sp.*, a fast growing genus of bacteria, also able to support predators' important growth rates (Šimek et al., 2013; Kasalický et al., 2013), and with similar abundances in all treatments (Fig 2). Interestingly, the next highest contributor to the differences in the Kennet bacterial community was an unknown *Acinetobacter* species, with similar frequencies under *P. minima* and *P. vonderheydeni* (NZ1-5c) and very low abundance in control

treatments – i.e. it was up-regulated in grazed communities. *Acinetobacter* is a genus of Gammaproteobacteria commonly found in water, soil and living organisms, and able to use a variety of substrates as energy sources (Baumann, 1968; Barbe *et al.*, 2004). Their higher abundance in the grazed communities compared to non-grazed controls indicates that they are able to proliferate when other bacteria are being suppressed, being a better competitor under those conditions. Indeed, many *Acinetobacter* species produce surface polysaccharides, which usually play a role in biofilm formation, rendering ingestion difficult (Abdel-el-haleem, 2003). Similarly, the genus *Pseudomonas* harbours species with high morphological plasticity, thus able to develop grazing-resistant morphotypes, as well as nutritionally versatile species (Palleroni, 2010; Silby *et al.*, 2011).

Communities from the Pang differed significantly in composition both when grazed by the medium-sized *Cercomonas braziliensis* (B13) and between the non-grazed controls and grazed treatments. This is driven mostly by the high but varying abundance of OTU 2998 (*Limnohabitans* sp.), but equally by *Pseudomonas stutzeri*, which was up-regulated when grazed by B13 in comparison with SW2 and NZ1-5c. *Pseudomonas* sp. occurs at relatively high frequencies (27%) in bacterial communities grazed by B13, while *Limnohabitans* sp. abundances are much lower than in other treatments: by grazing on the most abundant, accessible bacterial strain, *C. braziliensis* enables the development of other bacterial strains, such as *Pseudomonas stutzeri*. This bacterium is a more nutritionally versatile and competitive species than many pseudomonads, is motile and able to form biofilms by swarming (Lalucat *et al.*, 2006). Interestingly, *Acinetobacter* sp. – abundant in bacterial communities selected from the Kennet and promoted under grazing by *Paracercomonas minima* – has very low concentrations in communities from the Pang or Wye.

Our results suggest that interactions within microbial systems are far more complex than what one can see. However, this study would require higher replication in order to reduce within-treatment and between replicate variation and provide more robust, comparable results.

Grazer's characteristics – such as shape plasticity, mobility, feeding mode and phylogenetic relation – affected the growth of a variety of bacterial strains, which affected bacterial community diversity. But the latter also appears to be linked to life history and origin (river). Indeed, prokaryotes are highly dependent on available resources, and adapted to thrive in the environment they were taken from. But at the same time, prey inter- and intraspecific interactions have repercussions on the environment as a whole, and environmental perturbations as well as coexistence (in natural environments) with other competitive predatory species most likely alter protist behaviour. Furthermore, the sarcomonads used in this study harbour high genetic diversity and convergent morphologies (Bass, Howe, *et al.*, 2009), possible driver of a variety of ecological functions, which. In addition to that, little is known about protist adaptations to bacterial rapid evolution and defence mechanisms.

Table S4.1 SIMPER cumulative sums (treatments pairwise comparisons) of OTUs driving 70% of the differences between bacterial communities obtained from the Wye

Control vs. NZ1-5C`		Control vs. SW2		Control vs. B13	
denovo2998	0.2193907	denovo2998	0.2852034	denovo2998	0.3121408
denovo337	0.3502901	denovo337	0.4703136	denovo2892	0.428939
denovo1130	0.4624633	denovo1130	0.5200915	denovo1130	0.5414275
denovo59	0.4980531	denovo2892	0.5636184	denovo337	0.585333
denovo1560	0.5333825	denovo1560	0.6009086	denovo1270	0.6135883
denovo2892	0.5655502	denovo2406	0.626509	denovo1560	0.6417494
denovo351	0.5901803	denovo59	0.6512448	denovo2393	0.6683693
denovo1629	0.6136473	denovo2261	0.6738398	denovo59	0.6945774
denovo2261	0.6344058	denovo343	0.6962973	denovo1240	0.7171602
denovo1416	0.6516999	denovo351	0.7176144		
denovo2406	0.667991				
denovo343	0.6837366				
denovo2665	0.6977434				
denovo2393	0.71106				

NZ1-5c vs SW2		B13 vs. NZ1-5C		B13 vs. SW2	
denovo2998	0.3059995	denovo2998	0.3146289	denovo2998	0.2106126
denovo337	0.4543098	denovo2892	0.4370612	denovo337	0.3708182
denovo1130	0.5057592	denovo1130	0.5214395	denovo2892	0.4645958
denovo2892	0.5556067	denovo337	0.5812534	denovo1130	0.54804
denovo2406	0.5846	denovo1270	0.6115349	denovo2406	0.5748914
denovo1560	0.6128559	denovo1240	0.6397449	denovo343	0.5983579
denovo351	0.6408479	denovo1629	0.6632387	denovo2393	0.621268
denovo343	0.6671696	denovo343	0.686352	denovo1270	0.6438954
denovo2261	0.6930334	denovo2393	0.7075403	denovo351	0.6656596
denovo1416	0.7091573			denovo2261	0.6864928
				denovo1240	0.7050762

Table S4.2 SIMPER cumulative sums (treatments pairwise comparisons) of OTUs driving 70% of the differences between bacterial communities obtained from the Pang.

B13_control		Control vs. SW2		Control vs. Kv-Hf`	
denovo2998	0.3152359	denovo2998	0.308344	denovo2998	0.2717163
denovo337	0.5596164	denovo337	0.5527326	denovo337	0.4769234
denovo1240	0.6245532	denovo1240	0.6084671	denovo1240	0.5593009
denovo343	0.6678994	denovo727	0.65389	denovo2052	0.5996554
denovo260	0.7075453	denovo2052	0.6873679	denovo2031	0.630987
		denovo2031	0.718696	denovo343	0.6552276
				denovo2772	0.6777334
				denovo59	0.6959442
				denovo1130	0.7140243

B13 vs. Kv-Hf		B13 vs. SW2		Kv-Hf vs. SW2	
denovo2998	0.2374648	denovo2998	0.2949104	denovo2998	0.2788023
denovo337	0.3500748	denovo1240	0.3801441	denovo337	0.3856606
denovo1240	0.4565278	denovo343	0.4559399	denovo727	0.4554721
denovo343	0.5273356	denovo727	0.5220968	denovo172	0.5146332
denovo260	0.5831595	denovo260	0.5836671	denovo2031	0.5673906
denovo2031	0.6339926	denovo172	0.6309473	denovo1240	0.6158687
denovo821	0.6563561	denovo2031	0.6601297	denovo59	0.6379764
denovo59	0.6774051	denovo337	0.6845256	denovo821	0.6549079
denovo1560	0.6944352	denovo821	0.7040515	denovo2261	0.6696762
denovo1130	0.7100735			denovo2772	0.6841748
				denovo1130	0.6980098
				denovo1560	0.7114289

Table S4.3 SIMPER cumulative sums (treatments pairwise comparisons) of OTUs driving 70% of the differences between bacterial communities obtained from the Kennet.

Control vs. NZ1-5C		Control vs. B13		Control vs. SW2	
denovo2998	0.1496688	denovo2998	0.1491817	denovo2998	0.189992
denovo59	0.2230976	denovo59	0.2696039	denovo463	0.290269
denovo463	0.2879874	denovo463	0.3567591	denovo1560	0.362776
denovo1332	0.3359368	denovo1560	0.4296863	denovo337	0.4158268
denovo2393	0.3835197	denovo337	0.4898632	denovo2393	0.4585428
denovo1560	0.4252733	denovo2393	0.5365536	denovo59	0.4985876
denovo1130	0.4568444	denovo1130	0.5648619	denovo1332	0.5320574
denovo337	0.4879125	denovo2261	0.5888435	denovo1130	0.5619762
denovo2261	0.5117768	denovo1332	0.6101748	denovo2261	0.5869709
denovo1956	0.5295178	denovo1956	0.6283874	denovo350	0.6072029
denovo2031	0.5459433	denovo350	0.6431582	denovo1956	0.6269598
denovo1240	0.5608917	denovo1240	0.6563957	denovo2599	0.6397704
denovo343	0.5748588	denovo2031	0.6693415	denovo795	0.6511717
denovo2752	0.5869508	denovo192	0.6785956	denovo1868	0.662185
denovo257	0.5974877	denovo106	0.6871616	denovo1240	0.6730697
denovo2892	0.6077248	denovo795	0.6955634	denovo106	0.6833838
denovo350	0.6177925	denovo1008	0.7037986	denovo136	0.6935543
denovo2179	0.6258721			denovo2031	0.7027108
denovo192	0.6333519				
denovo136	0.640548				
denovo377	0.6476474				
denovo172	0.6544731				
denovo2312	0.6611417				
denovo2599	0.6671036				
denovo795	0.6729887				
denovo2406	0.6787946				
denovo701	0.6845962				
denovo351	0.6900658				
denovo2190	0.695489				
denovo106	0.7008253				

B13 vs. NZ1-5c		B13 vs. SW2		NZ1-5C vs. SW2`	
denovo2998	0.1844602	denovo59	0.1523311	denovo2998	0.2249798
denovo59	0.2928691	denovo2998	0.2716111	denovo59	0.2827947
denovo463	0.357227	denovo463	0.3556389	denovo463	0.3398688
denovo1560	0.4174374	denovo1332	0.4052579	denovo1560	0.3918559
denovo337	0.4736011	denovo337	0.450707	denovo337	0.4382178
denovo2393	0.5188636	denovo1560	0.4930076	denovo2393	0.4788513
denovo1332	0.5583406	denovo2393	0.5306344	denovo1332	0.5087809
denovo2261	0.5816154	denovo2261	0.55976	denovo1130	0.5298995
denovo1130	0.6024089	denovo2312	0.5766195	denovo2261	0.5509129
denovo2031	0.6190481	denovo2031	0.5934293	denovo2031	0.5702489
denovo2752	0.6311173	denovo1008	0.608037	denovo343	0.5857045
denovo343	0.643062	denovo1130	0.621958	denovo2752	0.5985543
denovo2312	0.654239	denovo1240	0.6353742	denovo2892	0.6110223
denovo350	0.6645523	denovo192	0.6486843	denovo350	0.6234734
denovo257	0.6738705	denovo1868	0.6600304	denovo257	0.6352553
denovo2892	0.6817505	denovo343	0.6694965	denovo377	0.6431495
denovo1008	0.6888065	denovo1325	0.6789586	denovo1868	0.6505513
denovo172	0.6955406	denovo2082	0.6873527	denovo2179	0.6579173
denovo2190	0.702189	denovo471	0.6954449	denovo2599	0.6649315
		denovo1956	0.7032588	denovo1240	0.6715367
				denovo1769	0.6781312
				denovo2406	0.6845303
				denovo192	0.6908781
				denovo627	0.6969693
				denovo1482	0.7028993

Table S4.4 Hutcheson's t-test for comparing Shannon's diversity indices. For each sample (Ctrl, SW2, B13 and NZ1-5c) a) replicates are compared to each other in each river, and b) between treatments. Values in red indicate a calculated t-value greater than the t-value at a=0.05 and df<.

	Kennet	Pang	Wye
a)			
Ctrl1-2	31.54106247	-11.07334291	0.000460705
Ctrl2-3	-7.777147138	-298.5339011	0.000139254
Ctrl1-3	103.2614513	-363.060389	0.000582716
SW2.1-2	22.84003509	0	-0.001506064
SW2.2-3	75.2500032	-110.7062694	0.001891353
SW2.1-3	92.02715131	-167.2160957	0.000525308
B13.1-2	-105.5055777	24.21643061	0.000727197
B13.2-3	-164.4048832	-60.98561378	-0.000250704
B13.1-3	-286.3814665	-26.13957447	0.000492374
NZ15c.1-2	-20.50363631	180.9906691	0.00028637
NZ15c.2-3	26.00858806	-36.67700123	-0.000336172
NZ15c.1-3	-1.828719819	226.9137573	-5.21863E-05
b)			
Ctrl-SW2	-22.29367335	170.7636037	-68.39633992
Ctrl-B13	64.75781306	13.06788844	32.49614581
Ctrl-Nz15c	34.17746369		80.1521664
Ctrl-KvHf		161.9631815	
SW2-B13	104.8973916	-212.7805905	82.93639124
SW2-NZ15c	100.393312		121.981142
SW2-KvHf		-23.60930081	
B13-NZ15c	-34.5691014		26.30329065
B13-KvHf		56716.38401	

Chapter 5 *PSEUDOMONAS FLUORESCENS* SBW25 TRANSCRIPTOMICS

The following work is the result of collaboration with Dr Ville Friman, at the time at Imperial College, Silwood Campus.

Preliminary results were obtained from experiments held at Silwood Campus, in the laboratory of Dr Thomas Bell.

Bacterial RNA samples from lines resulting from said experiment were extracted and quantified by sequencing; multivariate and me analyses were performed at the Centre for Genomic Research at the University of Liverpool. Further statistical analyses and results interpretation are my personal work.

5.1 Introduction

The bacterium *Pseudomonas fluorescens* (Flügge 1886, Migula 1895) is a versatile species, capable of colonising multiple habitat types and of great physiological and morphological plasticity (Rainey and Travisano, 1998; Palleroni, 2010). Members of this genus are present in most freshwater, marine and terrestrial environments, free-living or associated to other organisms, and can be beneficial or pathogenic. *P. aeruginosa*, for example, happens to be an opportunistic pathogen of humans with a weakened immune system, and is often associated to nosocomial infections (Jarvis and Martone, 1992). It is thus one of the most studied species among pseudomonads. The ability of *Pseudomonas* to thrive in such varied conditions indicates an important physiological and genetic plasticity. Early biochemical studies described the genus' incredible physiological diversity, and consequently its taxonomic diversity (Stanier *et al.*, 1966). Later investigations demonstrated similarly important genetic diversity (Rainey and Bailey, 1996; Ginard *et al.*, 1997; De Ita *et al.*, 1998).

The strain SBW25 is a saprophytic, gram-negative rod-shaped and plant growth promoter bacterium, firstly isolated in 1989 from the leaf surface of a sugar beet plant grown in Oxford (Rainey *et al.*, 1994). It harbours a 6,622,539 bp-long circular genome with 60.5% GC content. *P. fluorescens* SBW25 exhibits a phenotypic plasticity directly linked to niche preferences, which defines "the ability of a single genotype to produce more than one alternative form of morphology, physiological state, and/or behaviour in response to environmental conditions" (West-Eberhard, 1989). In spatially heterogeneous environments, SBW25, originally defined as 'smooth' (SM) ancestral strains, are able to develop into three principal morph classes: {SM}, wrinkly spreader {WS} and fuzzy spreader {FS} (Rainey and Travisano, 1998). These colonies have different fitness advantages when rare, and occupy different ecological niches in culture, where smooth types {SM} appear in suspension in liquid media, while fuzzy-spreader {FS} develop at the bottom of culture vessels and wrinkly-spreaders {WS} grow at the air-liquid interface. Given the ease with which *P.*

fluorescens SBW25 morphological changes can be observed, the strain has been extensively studied and is a model organism in studies of adaptive radiation. The molecular mechanisms triggering adaptive radiation are by now well known. WS morphs have been shown to over-produce a cellulose-like polymer (CLP) responsible for biofilm formation and the typical colony morphology of that phenotype, due to a mutation of the *wss* operon (Spiers *et al.*, 2002; Rainey and Rainey, 2003). Similarly, the FS morphotypes has been shown to arise from a loss-of-function mutation in the gene *fuzY*: this causes the bacterium to create unstable biofilms that collapse under their own weight and sink to the bottom of the microcosm. Interestingly, this means that fuzzy-spreaders start as wrinkly-spreaders with weak biofilms, and not as occupants of anaerobic niches (Ferguson *et al.*, 2013).
While diversification has been extensively studied in the context of environmental change, little is known about the effect of predators on the bacterium's phenotypic plasticity or its link to gene expression regulation.

Every living organism is influenced by the biotic and abiotic factors of their surrounding environment, as well as the interactions between those factors. Just like all other participants of any (natural) ecosystem, bacteria are subjected to variations in temperature, moisture, etc., as well as competition and predation. Protist predation is one of the major sources of bacterial mortality in most terrestrial and aquatic environments (Fenchel, 1987). In order to survive and succeed, prokaryotes have developed different strategies to ensure not only successful growth and reproducibility (when facing strong competitors), but also ways to avoid predators and/or defend against grazing. Bacteria can avoid predators before or after ingestion (phagocytosis; Matz and Kjelleberg, 2005; Montagnes *et al.*, 2008). For example microcolonies and biofilms, while too large to be ingested whole make individual cells inaccessible, and allow better cell-to-cell communication (De Kievit, 2009); increased speed allows avoiding predators more effectively. Post-ingestion strategies usually require the excretion of harmful metabolites and toxins, or rending the cell indigestible (Jousset *et al.*, 2006; Mazzola *et al.*, 2009; Song *et al.*, 2015).

When facing predation, *P. fluorescens* developed effective defence mechanisms of both specialist and generalist properties (Friman, Dupont, *et al.*, 2015). The gram-negative bacteria were incubated with four protist predators – the ciliates *Tetrahymena pyriformis* (TP) and *Tetrahymena vorax* (TV), the cryptophyte *Chilomonas paramecium* (CP) and the amoebozoan *Acanthamoeba polyphaga* (AP) – in all possible single- predator, two- and four-predator communities for 24 days at 22°C without shaking in a bacterial defence selection experiment (Fig. 5.1, blue boxes). Defence was measured as bacterial biofilm biomass, indicative of growth capacities of lines having evolved in either presence or absence of predators. Phenotypic diversification was calculated as the frequency of different colony morphologies observed in bacterial communities at the end of the experiment. The bacterium was shown to develop defence strategies when exposed to *T. pyriformis* only as unique predator. However, defence was reduced when other predators were added to the system, by means of better growth rates and community stability relative to the control treatment. Other predators alone did not drive any defence mechanism evolution. When grazed by both the amoeba and the ciliate (TPAP), the bacteria evolved defence against *A. polyphaga* by developing grazing-resistant morphotypes. *T. pyriformis* was the main driver of bacterial phenotypic diversification, while *T. vorax* completely repressed it; *A. polyphaga* and *C. paramecium* enhanced diversification. Evolved colonies (i.e. grown in the presence of predators) were either generalist defenders – initially selected by *T. pyriformis*, they were good in resisting grazing of all predators – or specialist defenders – poor at defending against *T. pyriformis* but effective against *C. paramecium* and *A. polyphaga*.

Ecological studies of organisms submitted to specific environmental conditions have been carried for decades, in both natural and controlled environments. Therefore, behaviour and phenotypic responses of organisms to biotic and abiotic variations between and within populations and individuals have been extensively observed and described in many ecological studies. With the advent of molecular tools, notably RNA sequencing, the patterns dictating observed ecological changes could be observed at the most basic level of expression in

an organism, i.e. at the genotypic level (Rowe *et al.*, 2004; Alvarez *et al.*, 2015). RNA sequencing, or transcriptomics analysis, indicates which genes are currently 'active', i.e. being transcribed (and translated into proteins). Differential expression thus allows analysing which genes are actually being up-regulated – meaning a higher degree of activation, thus a higher production of specific proteins – or down-regulated – genes being 'shut-down'. Measuring gene expression in an organism submitted to specific environmental conditions allows determining the molecular regulation response of that organism to the ecological variation.

Protist predation on *P. fluorescens* SBW25 drove morphological diversification, presumably as a defence mechanism to protist grazing. In order to analyse responses of *Pseudomonas fluorescens* to protist grazing at the gene expression level, we performed RNA extraction and sequencing of bacteria submitted to different predators types. To do so, we selected bacterial lines grown with protists that drove observable evolutionary changes in the prey. Since *T. vorax* did not select for any morphological diversification or detectable defence mechanism, it was not included in the transcriptome analyses. Similarly, the two-predator system *C. paramecium – A. polyphaga* was omitted from the gene expression study. Finally, the four-predator treatment was not considered for sequencing since ecological interactions within this system would be difficult to interpret within this experimental design. Triplicates of all remaining treatments – single-predators TP, CP and AP as well as two-predator systems TPAP and TPCP – were sequenced and analysed. This study aims to link observed morphological changes – related to predation defence – to differentially expressed genes, and identify patterns of expression to ecological variation.

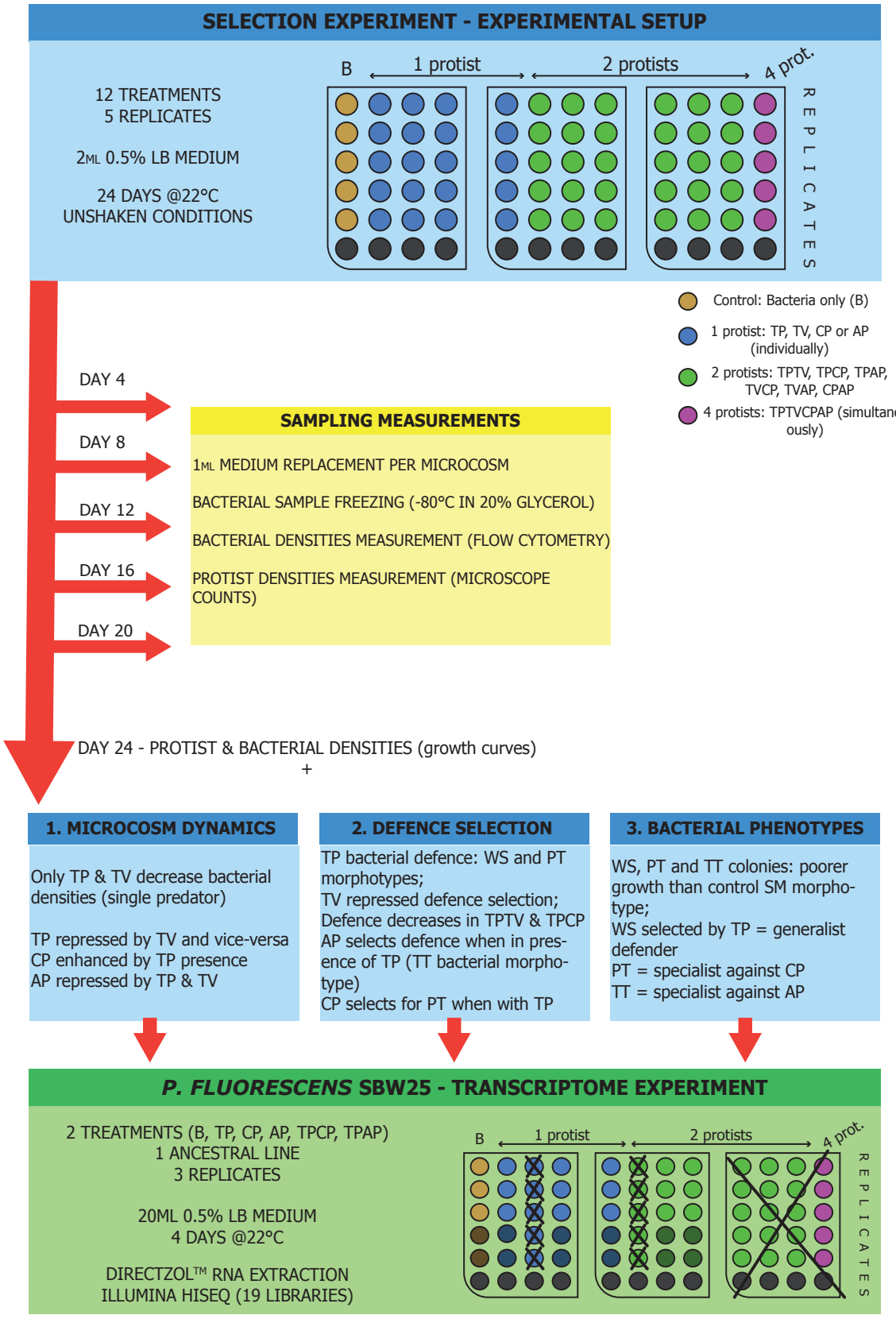


Figure 5.1 *P. fluorescens*' selection experiment.

TP: *Tetrahymena pyriformis*; TV: *Tetrahymena vorax*; CP: *Chilomonas paramecium*; AP: *Acanthamoeba polyphaga*

5.2 Methods

5.2.1 Selection experiment and RNA extraction

Based on the selection experiment results (Fig.1), we selected *P. fluorescens* SBW25 lines that showed the most important and interesting evolution patterns when submitted to protist grazing, such as the development of resistant colony morphotypes or better fitness (growth rate) amongst bacteria grown with or without protist grazers. Selected treatments were:

Control (B treatments): bacteria only, i.e. never exposed to any predator over the course of the selection experiment, indicative of adaptation to environmental conditions only - two replicates;

Tetrahymena pyriformis as single predator of *P. fluorescens* SBW25 (TP treatments), three replicates;

Chilomonas paramecium as single predator (CP treatments), three replicates;

Acanthamoeba polyphaga as single predator (AP treatments), three replicates;

T. pyriformis and *C. paramecium* in a two-predator community, grazing on *P. fluorescens* (TPCP treatments). Three replicates;

T. pyriformis and *A. polyphaga* in two-predator communities (TPAP treatments). Three replicates.

One 'ancestral' non-evolved (Ancestor) line was added to the selection. The final set consisted of control duplicates, triplicates for each treatment and a single non-evolved strain (Table 5.1). Prior to RNA sequencing, bacterial strains corresponding to the selected treatments (Fig 5.1) were thawed from stock conditions (-80°C) and grown for 4 days at 22°C without shaking, in the absence of predators and in the same media as in the previous selection experiment. In order to ensure enough material for RNA sequencing, bacteria were incubated in 20ml of 0.5% LB medium, in 50ml Falcon tubes.

Table 5.1 Experimental design and sequencing libraries of *P. fluorescens'* SBW25 transcriptome experiment.

Samples (CGR annotation)	Treatment group	Protist treatment
Sample 1-1	B1	Bacteria only (control 1)
Sample 2-2	B2	Bacteria only (control 2)
Sample 3-4	TP1	*T. pyriformis 1*
Sample 4-5	TP2	*T. pyriformis 2*
Sample 5-6	TP3	*T. pyriformis 3*
Sample 6-7	CP1	*C. paramecium 1*
Sample 7-8	CP2	*C. paramecium2*
Sample 8-9	CP4	*C. paramecium 4*
Sample 9-10	AP1	*A. polyphaga 1*
Sample 10-11	AP4	*A. polyphaga 4*
Sample 11-12	AP5	*A. polyphaga 5*
Sample 12-13	TPCP1	*T. pyriformis - C. paramecium 1*
Sample 13-14	TPCP3	*T. pyriformis - C. paramecium 3*
Sample 14-15	TPCP5	*T. pyriformis - C. paramecium 5*
Sample 15-16	TPAP2	*T. pyriformis - A. polyphaga 2*
Sample 16-17	TPAP3	*T. pyriformis - A. polyphaga 3*
Sample 17-18	TPAP5	*T. pyriformis - A. polyphaga 5*
Sample 18 - ANCESTRAL	Ancestor (non evolved)	Ancestral *P. fluorescens* SBW25

5.2.2 RNA extraction

RNA was extracted with the DirectZolTM RNA miniprep kit (Zymo Research, The Epigenetics Company®). Sample preparation requires lysing the cells in

suspension. To do so, 2ml aliquots were transferred in 2ml eppendorf tubes, and centrifuged at 10,000g for 1 minute (Desaga MC2); 1.7ml of supernatant were discarded and the bacterial pellet kept. This step was repeated until the total volume was used from the Falcon tubes. After the last step, the supernatant was discarded and cells lysed in the same tube with 1ml of TRI Reagent (Zymo Research, The Epigenetics Company®) and mixed. Next steps followed the kit's protocol. At the end of the last step, RNA was eluted in DNAse/RNAse-free water and concentration was assessed with NanoDrop 1000 Spectrophotometer (Thermo Fisher Scientific) and Qubit® 2.0 Fluorometer. RNA samples were stored at -80°C.

5.2.3 Library preparation and sequencing

Library preparation and RNA sequencing of P. fluorescens samples were carried out at the Centre for Genomic Research (CGR), University of Liverpool, UK. Methods and bioinformatic analyses are taken from the report provided by the CGR (annex).

After quality check, samples containing 1μg/40μl or more of total RNA were depleted of ribosomal RNA with the ScriptSeq Complete Bacteria Low Input kit (Epicentre), according to the Low Input protocol. Eighteen libraries (Table 5.1) were prepared according to the ScriptSeq v2 protocol, and amplified for 15 cycles before being purified using AMPure XP beads. Each library was quantified with Qubit and size distribution assessed using the Agilent 2100 Bioanalyser. Amplified libraries were multiplexed as 2 individual pools with nine libraries each. Quality and quantity of pools were assessed with Qubit and the Agilent 2100 Bioanalyser, and subsequentely by qPCR using the Illumina Library Quantification Kit (Kapa) on a Roche Light Cycler LC480II.

Template DNA was denatured according to the protocol described in the Illumina cBot User guide and loaded at 12.5 pM concentration. Sequencing was carried out on two lanes of an Illumina HiSeq2500 at 2x125bp paired-ended sequencing with v4 chemistry.

5.2.4 Bioinformatic analyses

5.2.4.1 Initial processing and quality assessment of sequence data

Dr Richard Gregory, University of Liverpool, developed the following data processing and quality assessment.

Basecalling and de-multiplexing of indexed reads was performed by CASAVA version 1.8.2 (Illumina) to produce 18 samples from the 2 lanes of sequencing data, in fastq format. Raw fastq files were trimmed to remove the Illumina adapter sequences with Cutadapt v1.2.1 (Martin, 2011). Reads that matched the adapter sequence for 3 or more base pairs were trimmed off at the 3' end. Further trimming removed low quality bases with a minimum window quality score of 20 (Sickle version 1.200). Reads shorter than 10bp were removed after trimming. When both reads from a pair passed quality filtering, they were each included in the R1 (forward reads) or R2 (reverse reads) file. If only one read of a pair passed the filter, it was assigned to the R0 file (unpaired reads).

5.2.4.2 Alignment of reads to reference sequences

Pseudomonas fluorescens genome assembly SBW25 was used as reference for the alignment. Reference sequence and annotations were downloaded from

http://www.pseudomonas.com/downloads/pseudomonas/pgd_r_15_2/Pseudomonas_fluorescens_SBW25_116/Pseudomonas_fluorescens_SBW25_116_seq.gff

R1/R2 read pairs were aligned to the reference using short read mapper Tophat v2.1.0 (Langmead and Salzberg, 2012). Reads were aligned using option "-g 1", which instructs Tophat to report at most one alignment to the reference for a given read. If there are multiple hits for a single read, the alignment with the best score, or a randomly selected one amongst alignments with equally best scores is selected.

5.2.4.3 Differential expression analysis

Bam files generated from the alignments were processed with HTseq-count (Anders *et al.*, 2014) to obtain read counts for genes. Data variation assessment was performed in order to identify and removed potential outlier samples. Differential gene expression (DE) was performed with packages edgeR (Robinson *et al.*, 2009) in the R environment (the R core development team 2008).

Variation in the count data was analysed with pairwise scatter-plot of mean log10 counts for each sample group (Fig. 5.2a). Samples with similar gene counts appear as plots with low dispersion of points, while higher dispersion illustrates genes that are not equally expressed in two different samples. Correlations between gene expressions (number of counts per gene) of all samples were visualised with a heatmap: the more similar two samples are in terms of gene expression, the highest their correlation (Fig. 5.2b). Highly correlated samples appear in dark red, while blue squares indicate lower correlation between samples.

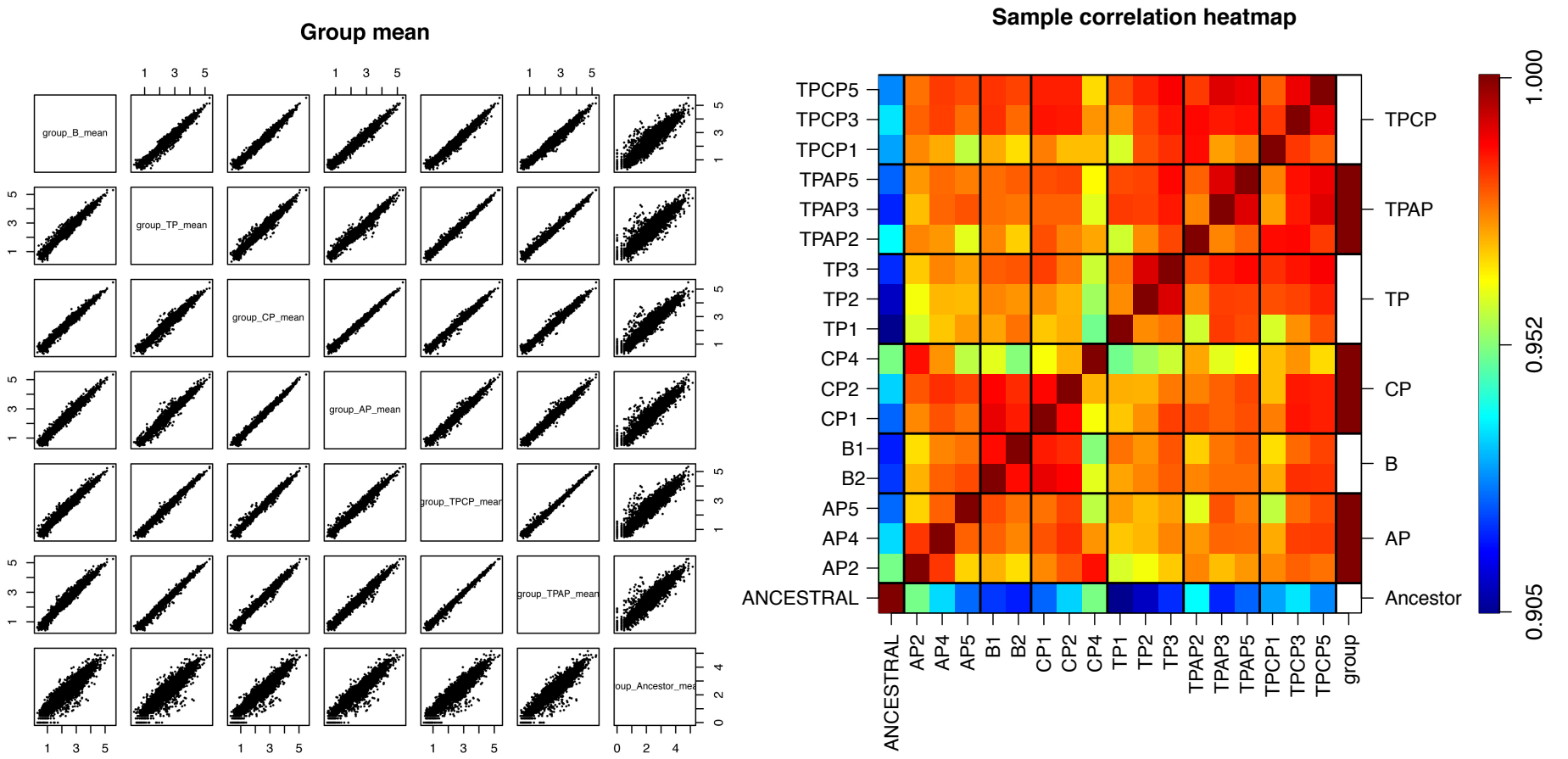


Figure 5.2 Pairwise scatter-plots (a) for each sample group: low dispersion indicates similar gene count for both groups (axes) in each pairwise plot. Gene expression correlation (b): dark red indicates highly correlated samples (i.e. similar gene expression), blue squares indicate low correlation between samples.

Metric dimensional scaling (MDS) and hierarchical clustering of samples read counts based on Euclidean distance measures were used to group together samples that were similar to one another (Fig. 5.3).

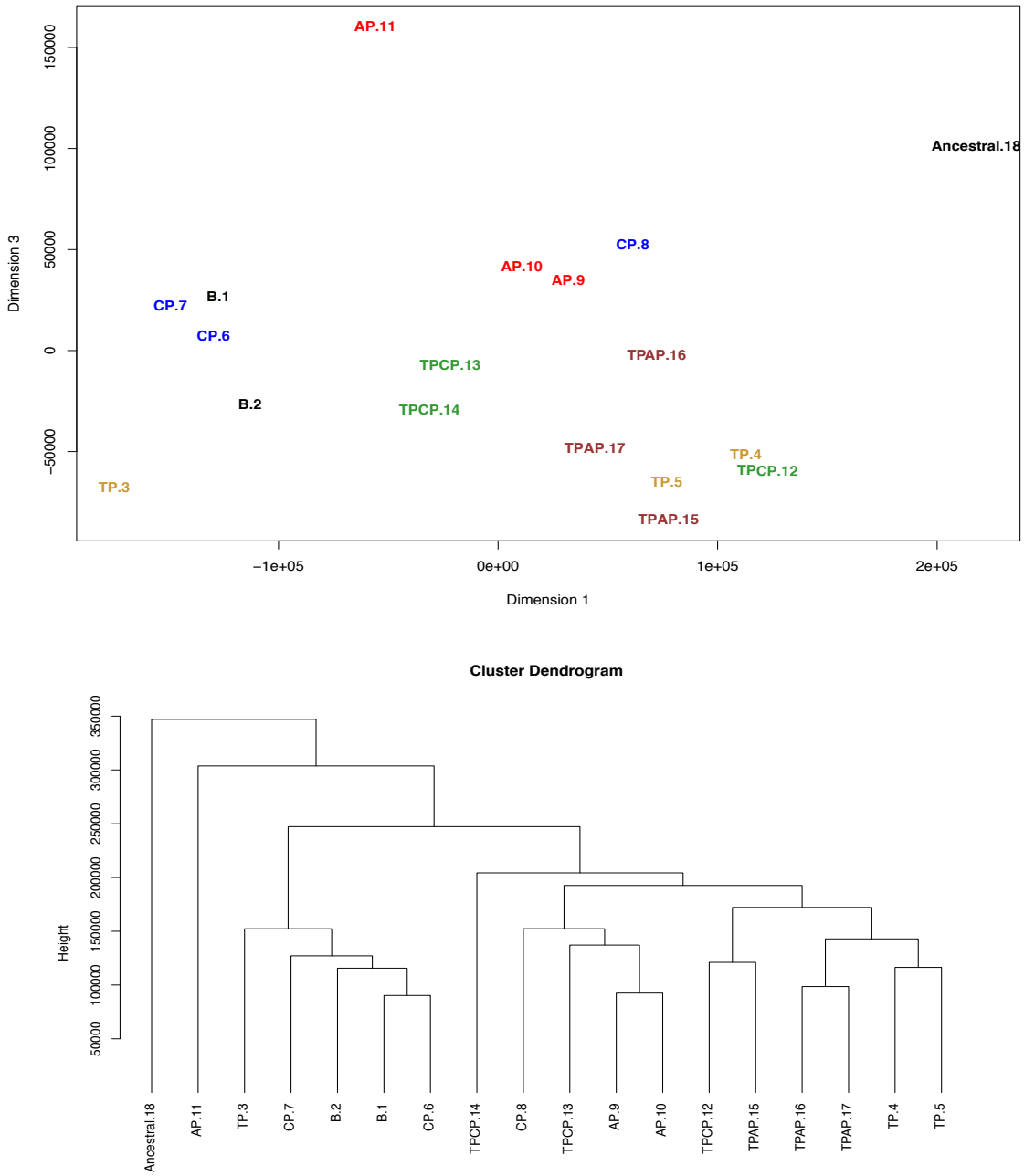


Figure 5.3 MDS ordination (left) and hierarchical clustering dendogram (right) of samples' read counts. Colours indicate protist predator treatment. Most similar observations cluster together in the MDS plot and are combined in the first level of the dendogram; the more distant the samples (as Euclidean distance), the higher the level of fusion in the dendogram (height).

Differential expression analysis was applied to gene count values. Variations of the count data were modelled by a negative binomial distribution and the data modelled using a generalized linear model (GLM) (Nelder and Wedderburn, 1972). For each sub-set of data, a GLM containing parameters representing the mean expression of corresponding sample groups was used. Pair-wise comparisons of 7 samples groups were conducted based on the model fitting results. 21 contrasts were denoted as: AP/Ancestor, B/Ancestor, CP/Ancestor, TP/Ancestor, TPAP/Ancestor, PCP/Ancestor, B/AP, CP/AP, TP/AP, TPAP/AP. TPCP/AP, CP/B, TP/B, TPAP/B, ATPCP/B, TP/CP, TPAP/CP, TPCP/CP, TPAP/TP, TPCP/TP, TOCP/TPAP. The group tag positioned after "/" is the baseline group in a contrast. For example AP group is the baseline group in comparison "B/AP".

The GLM was parameterised using the count data to obtain the logFC values for the required comparison. The estimated log2 Fold Changes (logFC) were tested in edgeR using a Likelihood-Ratios (LR) test (Wilks, 1938). P-values associated with logFC were adjusted for multiple testing using the false discovery rate (FDR) approach (Benjamini and Hochberg, 1995). Significantly differentially expressed genes were defined as those with FDR-adjusted P-value < 5%.

MA plots describe log-intensity ratios as a function of log-intensity averages (log2 of calculated fold change (logFC) against log2 of counts per million mapped reads (log2CPM)) (Ritchie et al., 2015). 'Volcano'plots retrace –log10 p-values against logFC. The package 'calibrate' function 'textxy()' provided the gene names on the plots. Values in red indicate differentially expressed genes (FDR<0.05).

5.3 Results

On average, 98% of raw reads were kept after quality and adapter trimming (suppl. table S5.1), with replicate 1 of *Tetrahymena pyriformis* (TP1) showing the most accepted sequences post-trimming (98.96%) and replicate 5 of treatment *Tetrahymena pyriformis* - *Acanthamoeba polyphaga* (TPAP5) the

lowest percentage of reads (92.84%). However, treatment TPCP5 had the maximum number of reads while replicate 2 of *Chilomonas paramecium* the lowest. Reads allocated to the R0 file (unpaired reads) appeared more heavily trimmed than those in the R1/R2 files (suppl fig. F5.1), indicating that, as expected, paired reads had an overall better quality than unpaired reads.

Mapping of the libraries' reads to the reference genome of *P. fluorescens* SBW25 ranged from 86.56% of aligned sequences for the ancestral strain to 73.56% for replicate 5 of *A. polyphaga* grazed-bacteria.

5.3.1 Count data variation assessment

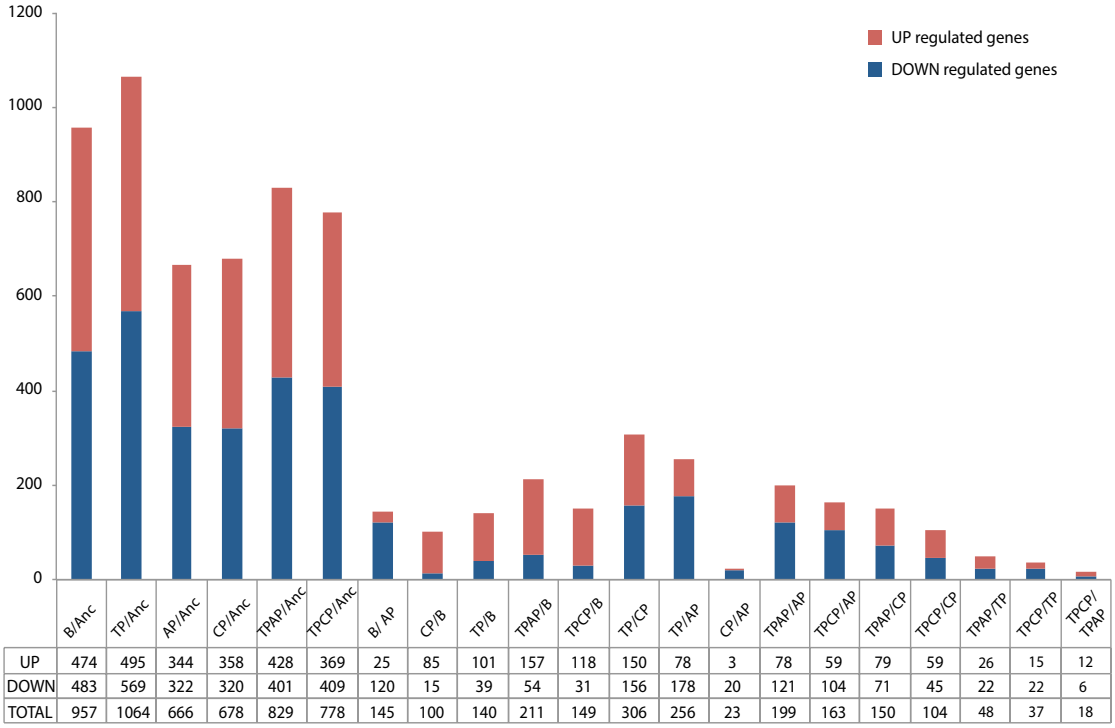
Count data variation analysis indicates whether differences between samples in a same treatment group (Table 5.1) are due to technical and biological variation only (within-group variation) or to treatment effect (overall variation). When the treatment effect is the dominant contributor to variation, sample groups can be clearly separated by statistical tools. However, when the treatment effect is weak compared to biological or technical variation, grouping of samples can be difficult to determine based on the data. Thus, if a sample appears extremely far from the other samples from the same group, it is most likely an outlier sample, which can be singled out by statistical tools.

The pairwise scatter plots of mean counts (log10) for each sample group indicated a higher variation in counts for the ancestral *P. fluorescens* SBW25 strain (non-evolved) only (fig. 5.2a). Indeed, the correlation between ancestral strains and all other strains appeared low (0.905) when compared to the ones of all other groups (fig. 5.2b). Of those, higher correlations happened between treatments B and CP, CP and AP, TP and TPCP, TP and TPAP and finally between TPCP and TPAP, indicating that bacteria submitted to those predator combinations evolved in more similar ways than the ones in other treatments. Metric dimensional scaling (MDS, fig. 5.3) of all samples isolated the non-evolved strain (Ancestral), while all other treatments clustered more or less together. Hierarchical clustering of samples' Euclidean distances inferred the difference of the ancestor bacterial line from other treatments, as well as the

correlations between samples according to samples. Despite variation assessment results, the ancestral *P. fluorescens* SBW25 sample could not be ruled as an outlier as it is the only replicate of this treatment.

5.3.2 Differential expression

Differential expression analysis was performed on gene count values for comparison of all 7 treatments, based on 5921 currently identified genes in P. fluorescens SBW25.



	B/Anc	TP/Anc	AP/Anc	CP/Anc	TPAP/Anc	TPCP/Anc	B/ AP	CP/B	TP/B	TPAP/B	TPCP/B	TP/CP	TP/AP	CP/AP	TPAP/AP	TPCP/AP	TPAP/CP	TPCP/CP	TPAP/TP	TPCP/TP	TPCP/ TPAP
UP	474	495	344	358	428	369	25	85	101	157	118	150	78	3	78	59	79	59	26	15	12
DOWN	483	569	322	320	401	409	120	15	39	54	31	156	178	20	121	104	71	45	22	22	6
TOTAL	957	1064	666	678	829	778	145	100	140	211	149	306	256	23	199	163	150	104	48	37	18

Figure 5.4 Significantly differentially expressed genes (FDR<0.05) in every group contrast.

The number of significantly differentially expressed (DE) genes (FDR<0.5%) for each of those comparisons is reported figure 5.4. The contrast TP/Ancestor had the higher number of DE genes (1064), with approximately half of those being up regulated (46.5%) and the other half down-regulated (53.5%). Conversely, the comparison between treatments TPCP and TPAP only showed 18 DE genes, which confirms the similarity between the bacterial strains grown under *T. pyriformis-C. paramecium* and *T. pyriformis-A. polyphaga* predators. Up

regulated genes were twice more abundant than down-regulated ones (12 and 6 respectively). Number of DE genes in all other comparisons ranged between the two previous contrasts, as well as the proportions of up- and down-regulated genes. They indicate bacterial adaptation to environmental and predator pressures.

Highly expressed genes in an organism submitted to specific experimental conditions record high numbers of reads mapped on a reference genome. MA plots (Figs. 5.5, 5.8 & 5.9) report the relationship between the intensity of expressed genes (log2CPM: log2 counts per million (mapped) reads) and the difference between 2 treatments (logFC.contrast: log fold-change between the two treatments in the contrast). For every contrast considered, points with high logCPM and high absolute logFC values represent genes with a high number of mapped reads and that are very different – up (positive logFC values) or down-regulated (negative logFC) – from each other. Volcano plots (Figs. 5.6, 5.7 & 5.10), on the other hand, describe whether the observed change between two treatments (logFC.contrast) is significant or not (-log10 (PV); the lower the original p-value, the larger it will appear on the x-axis).

Significantly differentially expressed genes for all contrasts are described table 5.3 (digital material). LogFC values with red font indicate significance of differential expression, while cell filling colour indicate whether the gene is up (positive values, red cell highlight) or down regulated (negative values, blue cell highlight).

5.3.3 Ancestral non-evolved versus evolved P. fluorescens SBW25 strains

Ancestor contrasts detected the highest numbers of DE genes in comparison to all other treatment contrasts. 957 significantly differentially expressed genes, of which 49.53% up-regulated genes and 50.47% down-regulated, were detected between ancestral non-evolved and control lines (treatment B). Since 5921 genes are currently identified in P. fluorescens SBW25, this indicated that 16.16% of all genes from the alone-evolved strains were significantly different to the ancestral strain, affected solely by experimental conditions. Single-protist

predation generated 17.97% DE genes between the ancestral and *T. pyriformis* grazed strains, 11.45% DE genes between non-evolved and *C. paramecium* predated strains and 11.25% for Ancestral versus *A. polyphaga*. Both two-predator treatments generated 13.14% and 14% of differences for TPCP and TPAP grown strains, respectively. Interestingly, most of DE genes for the above treatments were down regulated, with the exception of CP/Ancestor.

Mostly the same genes were highly expressed (logCPM) as well as significantly differentially expressed (FDR<0.05, red points) in contrasts B/Ancestor, AP/Ancestor, CP/Ancestor and TP/Ancestor (fig 5.1).

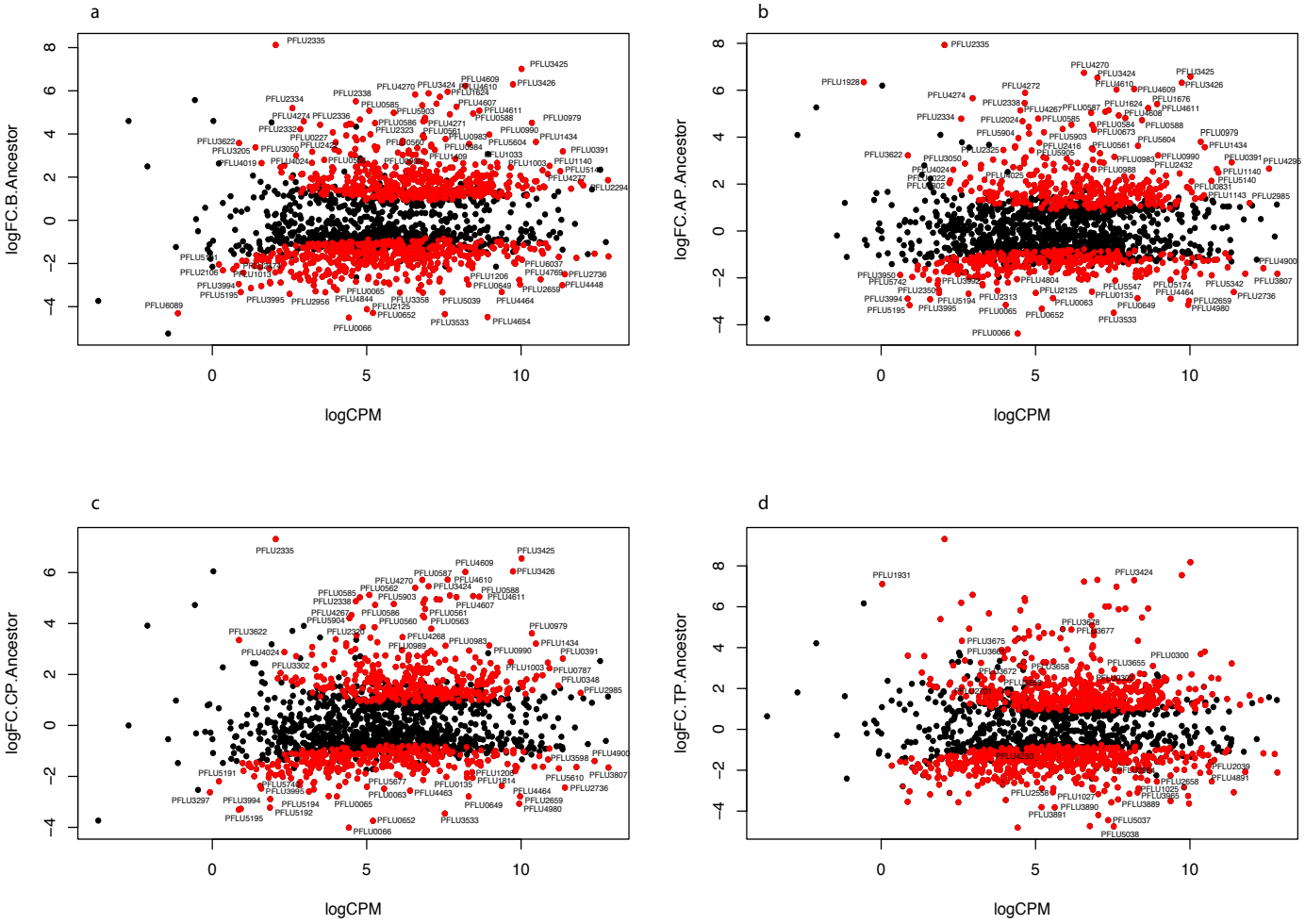


Figure 5.5 intensity of expressed genes (log2CPM) against treatment difference (logFC) in single predator/ancestor comparisons for (a) B/ancestor, (b) AP/ancestor, (c) CP/ancestor and (d) TP/ancestor.

Genes that differed more between ancestor and B-treatments or between ancestor and single-protist treatments were globally the same as well (fig 5.2): *P. fluorescens* SBW25 response to environmental conditions only (B treatment) or to a single protistan predator seems to be driven by the same global set of genes. From all down-regulated genes in the B/Ancestor contrast, 121 were highly differentially expressed (|logFC| ≥ 2; table 5.3). Of those, most of them were located in the cytoplasmic membrane and involved in molecule (active or passive) transport between the bacteria's inner and outer-membrane spaces; or signal transduction, often linked to bacterial chemotaxis. Within the cell's cytoplasm, most genes encoded for metabolic processes, more precisely catalytic ones (PFLU4654, PFLU0361, PFLU5192) indicating an apparent reduction of compounds' breakdown and energy production. Similarly, a few genes responsible for sugar transport or metabolism (PFLU4845, PFLU3995, PFLU 5038, PFLU5040) or linked to cell surface structures such as flagella (PFLU1155, PFLU4448), fimbriae (PFLU0649, PFLU0638) and lipoproteins (PFLU0159, PFLU3403, PFLU0163) were also down regulated. In parallel, importantly up-regulated genes in the alone-evolved bacterial strains (173 genes) were directly or indirectly involved in nitrogen metabolic processes (PFLU3425, PFLU0562, PFLU4607), often linked to N assimilation and synthesis of amino acid, building blocks of proteins. Different types of active and passive transport elements also observed a higher level of transcription. Interestingly, alginate – a biopolymer usually present in bacterial biofilms – had its biosynthesis enhanced in alone-evolved bacterial treatments.

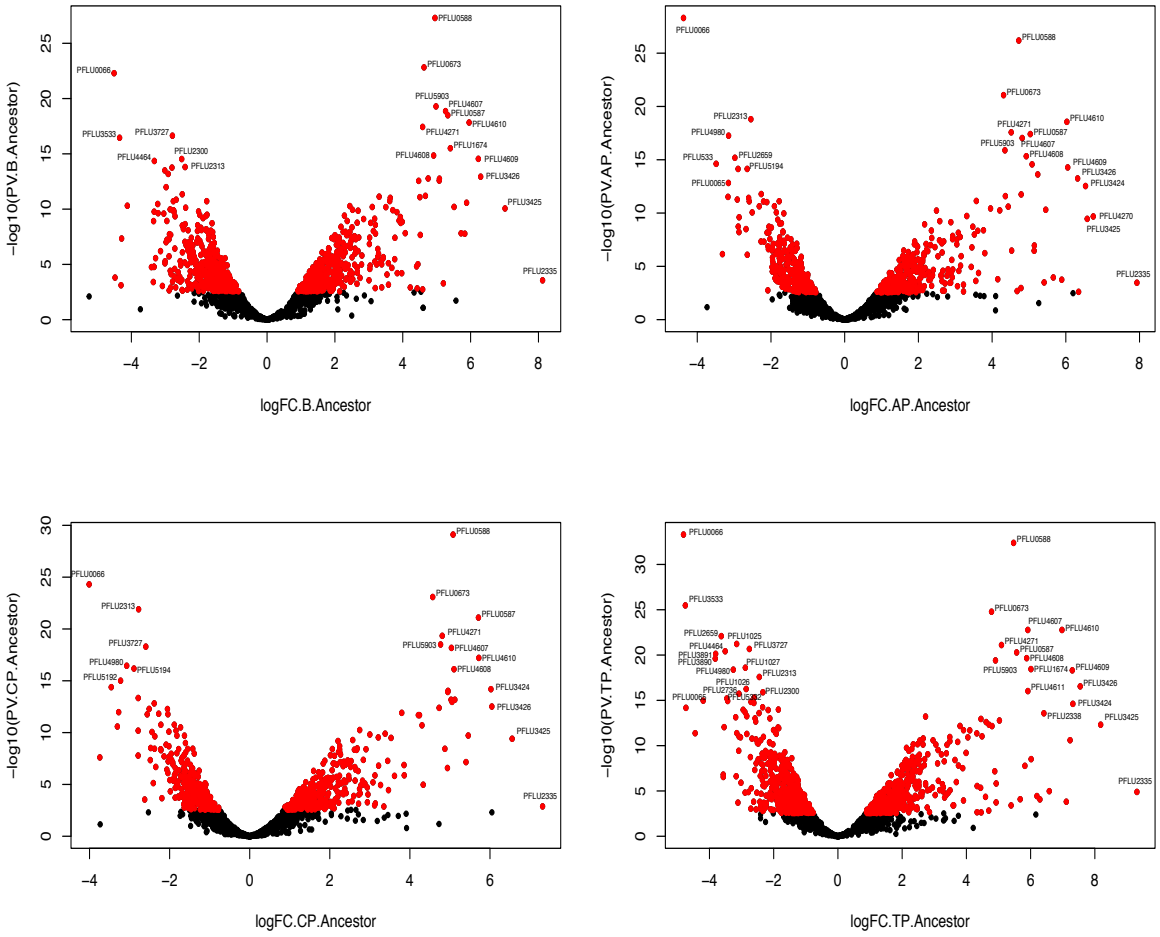


Figure 5.6 Significance levels (-log10 p-values) as a function of treatment difference (logFC) in ancestor/single predator comparisons for (a) B/ancestor, (b) AP/ancestor, (c) CP/ancestor and (d) TP/Ancestor contrasts.

5.3.4 Alone-evolved bacterium versus single-predator evolved

Comparison of bacterial strains grown without (B) and with a single protist (*T. pyriformis* TP, *C. paramecium* CP or *A. polyphaga* AP) describes specific adaptation to one particular type of predation. It is important to notice that contrasts were done using the treatment behind "/" as a reference (baseline), and comparisons were made as follows: B/AP, CP/B and TP/B. The contrast B/AP used *A. polyphaga* as baseline group, while CP/B and TP/B used the bacteria-only treatment as baseline group. Because of that, in fig 6.2, genes were up regulated in the alone-evolved bacteria, thus down regulated in the AP-grazed strains, and vice-versa. In parallel, up regulated genes in figs 6.2c and 6.2d were effectively so in CP and TP treatments respectively.

T. pyriformis-evolved *P. fluorescens* presented 140 DE genes when compared to alone-evolved bacteria, with most genes up regulated (72.12%). Similarly, CP-evolved *P. fluorescens* had 100 DE genes compared to the bacteria-only treatment, with 85% of those up regulated. As for the comparison B/AP, 145 genes were differentially expressed with only 17.24% up regulated in the bacteria-only treatment, the majority (82.76%) thus down regulated. This represented a higher percentage of up regulated genes in the *A. polyphaga* microcosm. Adaptation is thus mostly driven by an increase in transcription for all grazed bacterial lines when compared to those adapting to environmental conditions only.

Bacterial strains grown under *A. polyphaga* and *C. paramecium* presented 13 (out of 15 and 25 for CP and AP respectively) common genes with reduced expression, mostly related to molecule transport within and around the cell and catabolism. The remaining down regulated genes in AP (indicated as up regulated in the B/AP treatment) presented the same global functions, although some were included in signal transduction - which usually requires two ATP-dependent proteins (one in the cytoplasmic membrane and one in the cytoplasm) in order to relay the signal – and or in relation with flagellar motility (PFLU4456, PFLU4440). Similarly, all of up regulated genes in CP were the same as in AP. Many are involved in signal transduction/chemotaxis or membrane lipoprotein attachment sites, meaning an active monitoring of the external environment. Interestingly, when genes were down regulated in *T. pyriformis*, the few common ones with the other treatments were actually mostly up regulated in AP and CP (when compared to B) (table 5.3). Furthermore, genes with lower expression in the ciliate treatment are only significantly so at lower levels of expression, more precisely at $-1.85 < \log FC < -0.800$ (B/AP: $-5.05 < \log FC < -.42$; CP/B: $-4.60 < \log FC < -0.94$). As for up regulated TP genes, the 20 most different ones (high logFC) were unique to the ciliate-grazed bacteria; other highly different genes were also up-regulated in both AP and CP-grazed *P. fluorescens* in comparison to the alone-evolved strains. Down regulated genes in *T. pyriformis* include metabolic pathways, apparently mostly catalytic activities - responsible for providing the cell with energy and

synthesising essential elements such as amino acids, building blocks of proteins. Meanwhile, increased transcription of genes regulating signal transduction, chemotaxis ‘(PFLU2358, PFLU3655, PFLU3358) and protein phosphorylation/dephosphorylation (PFLU3980, PFLU3678) indicated regulation of physiologic processes or post-translational enzymes/proteins modification as well as cell communication and aggregation, in line with enhanced biofilm formation observed in evolved bacteria (Friman et al., 2015).

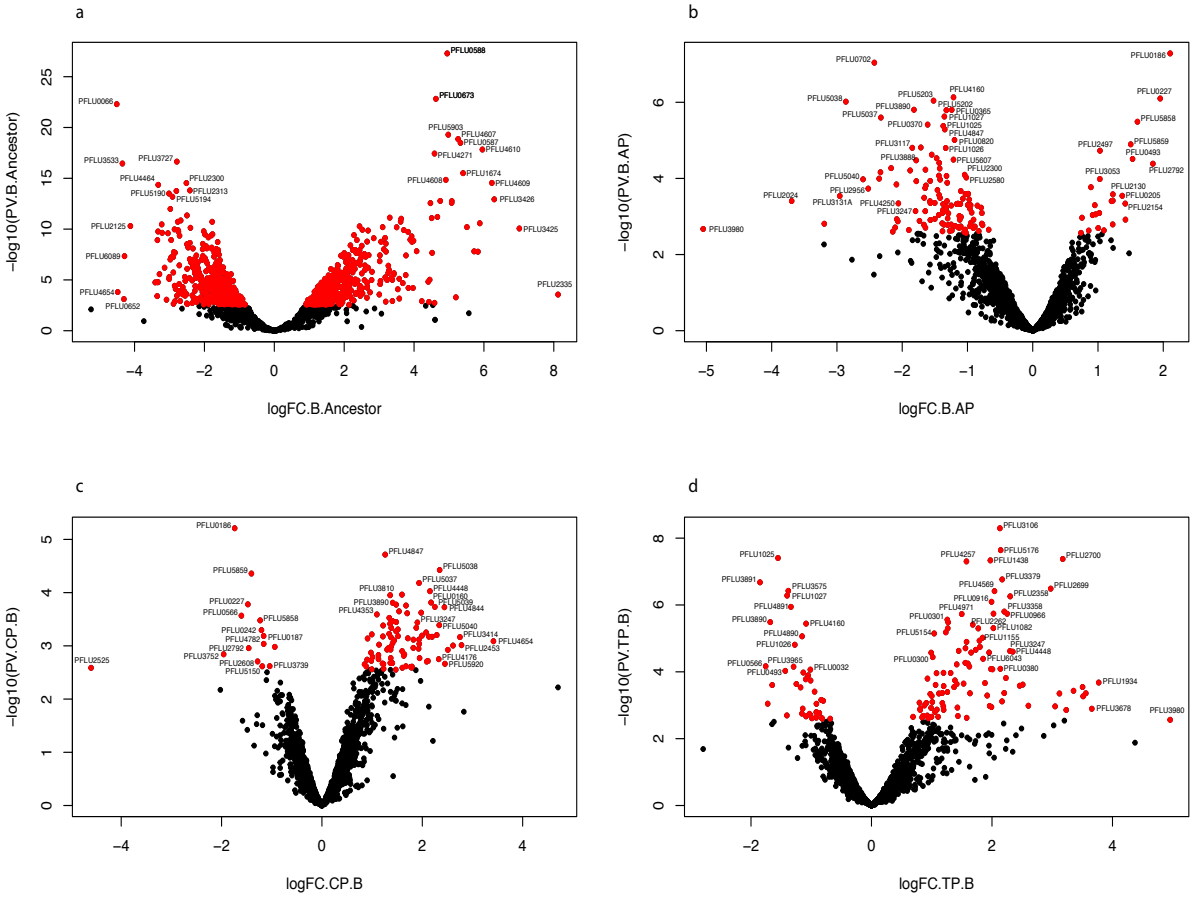


Figure 5.7 intensity of expressed genes (log2CPM) against treatment difference (logFC) in bacteria alone/single predator comparisons for (a) B/ancestor, (b) B/AP, (c) CP/B and (d) TP/B.

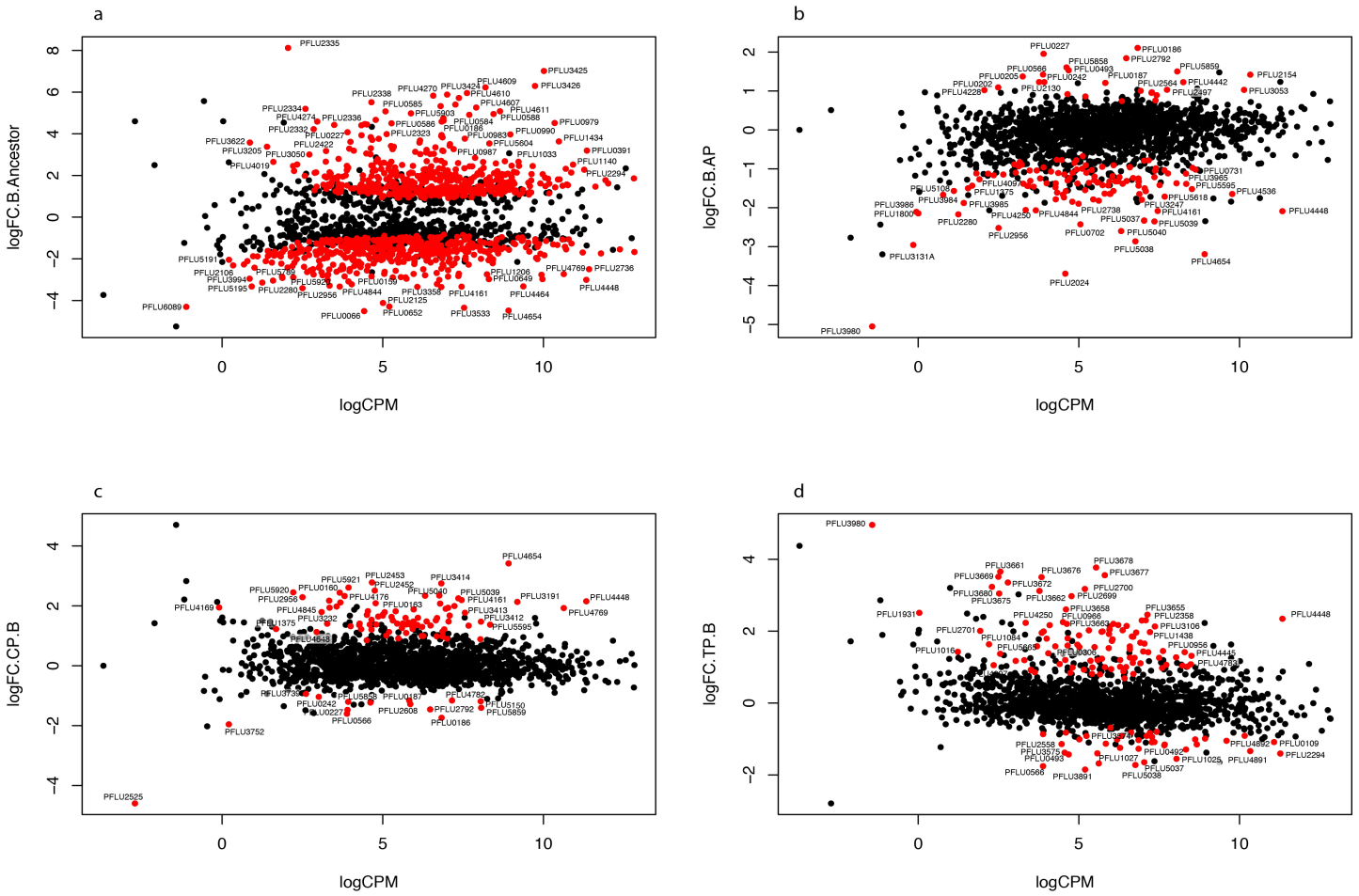


Figure 5.8 significance levels (-log10 p-values) as a function of treatment difference (logFC) in alone-grown bacteria/ancestor or alone-grown/single-predator comparisons for (a) B/ancestor, (b) B/AP, (c) CP/B and (d) TP/B contrasts.

5.3.5 Single-predator evolved bacterium versus two-predator evolved

Differential gene expression analysed for single protist and two-protist treatments indicates how the presence of an additional predator might impact bacterial evolution. Interestingly, the contrast TPCP/TP only detected 37 DE genes (40.45% up regulated/59.46% down regulated), and the contrast TPAP/TP 48 DE genes (54.17% up/45.83% down). The relatively low number of differentially expressed genes in treatments TPCP and TPAP when compared to TP only indicate that *P. fluorescens*, when adapted to *T. pyriformis* grazing pressure, is already well defended against the extra *C. paramecium* or *A. polyphaga* grazers, thus needs less transcriptional regulation. Indeed, Friman *et al.* (2015) showed that the ciliate selects for bacterial mutation that confers generalist resistance. *C. paramecium* as an extra predator drives adaptation by discreetly reducing transcription levels in the bacterial cells (higher level of down DE genes). Conversely, *A. polyphaga* regulated gene expression by somewhat increasing transcription (fig. 5.10). Concurrently, two-protist predator treatments - TPCP and TPAP - compared to CP-only demonstrated differential expression of 104 and 150 genes, with most of them up regulated, meaning that bacteria undergoing *T. pyriformis* + *C. paramecium* or *T. pyriformis* + *A. polyphaga* grazing pressure adapt by increasing cells' transcription levels. Contrast with *A. polyphaga* treatments indicate slightly higher DE levels (163 DE genes for TPCP/AP and 199 DE genes for TPAP/AP), mostly down regulated: more genes are actively transcribed under *A. polyphaga* grazing alone then under joint TPCP or TPAP pressure.

Despite the apparent weak gene expression regulation in two-predator grazed bacteria compared to the single *T. pyriformis* treatment (TPCP/TP and TPAP/TP), very different set of genes were involved in the contrasts (figs 5.9 & 5.10). Of all significantly down regulated genes in the TPCP/TP contrast, three of them were also down regulated in the TPAP/TP contrast: PFLU0186 (protein with unknown function), PFLU5859 (putative peroxidase – usually involved in the degradation of toxic oxygen forms in the cell) and PFLU3584 (putative active transport protein). Interestingly, these were up regulated in the bacteria-

only (B) treatments when compared to the ancestor, CP and AP strains. Similarly, only two of all TPCP/TP up regulated genes were also so in the TPAP/TP contrast: PFLU4353 (putative aminotransferase) and PFLU4847, an unknown protein supposed to be responsible of catalytic activities. Conversely, most genes with lowered expression in the TPAP/AP comparison were equally down regulated in contrasts TPCP/AP, TPAP/CP and TPCP/CP. A similar pattern was observed for up regulated genes in those contrasts, and only a few were common with genes significantly differentially expressed in all other contrasts. Interestingly, many appeared to be related to (bacterio)phages or encoding for phage-like proteins.

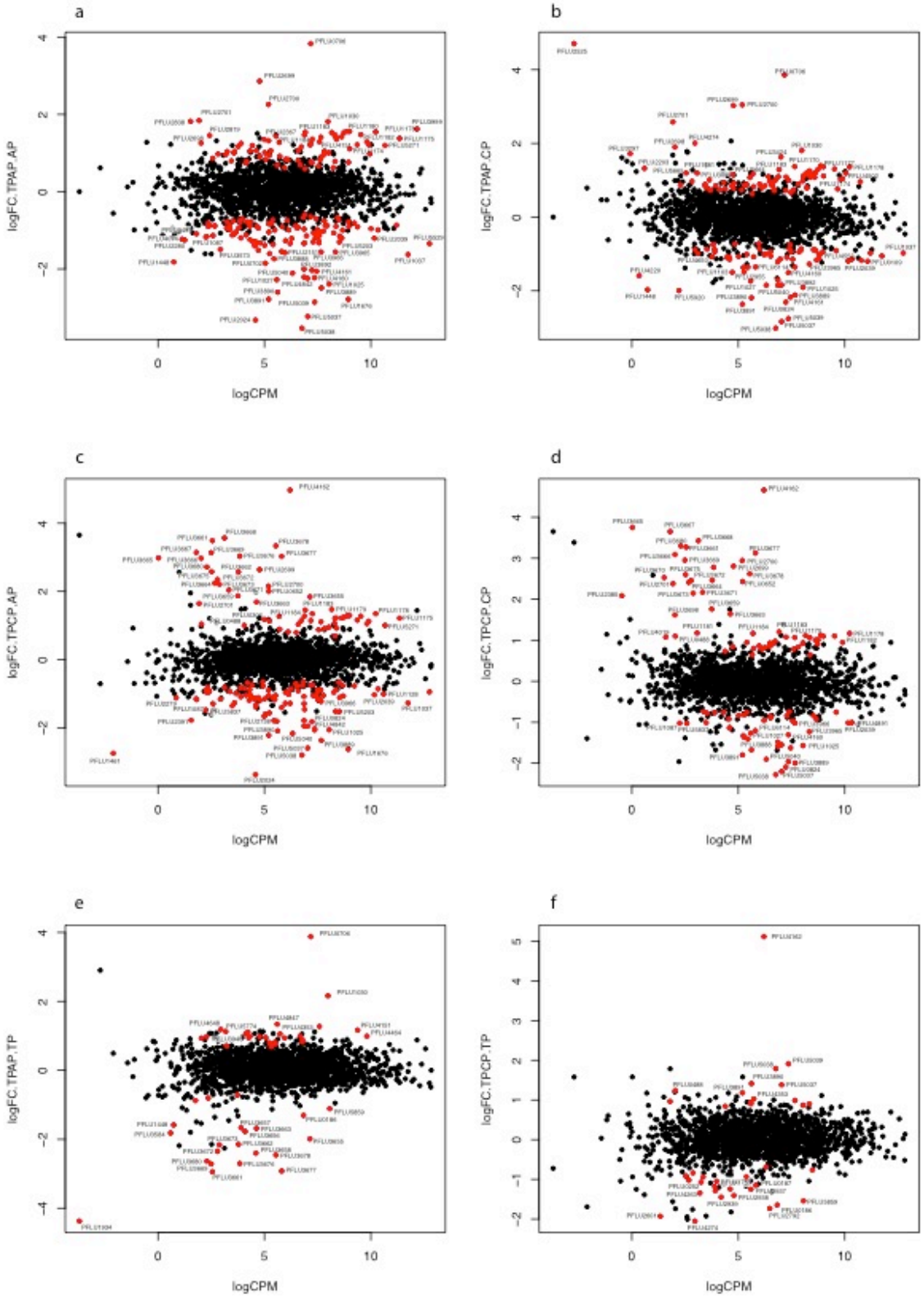


Figure 5.9 Intensity of expressed genes (log2CPM) against treatment difference (logFC) in single predator/two predator comparisons for (a) TPAP/AP, (b) TPAP/CP, (c) TPCP/AP (d) TPCP/CP, (e) TPAP/TP and (f) TPCP/TP contrasts.

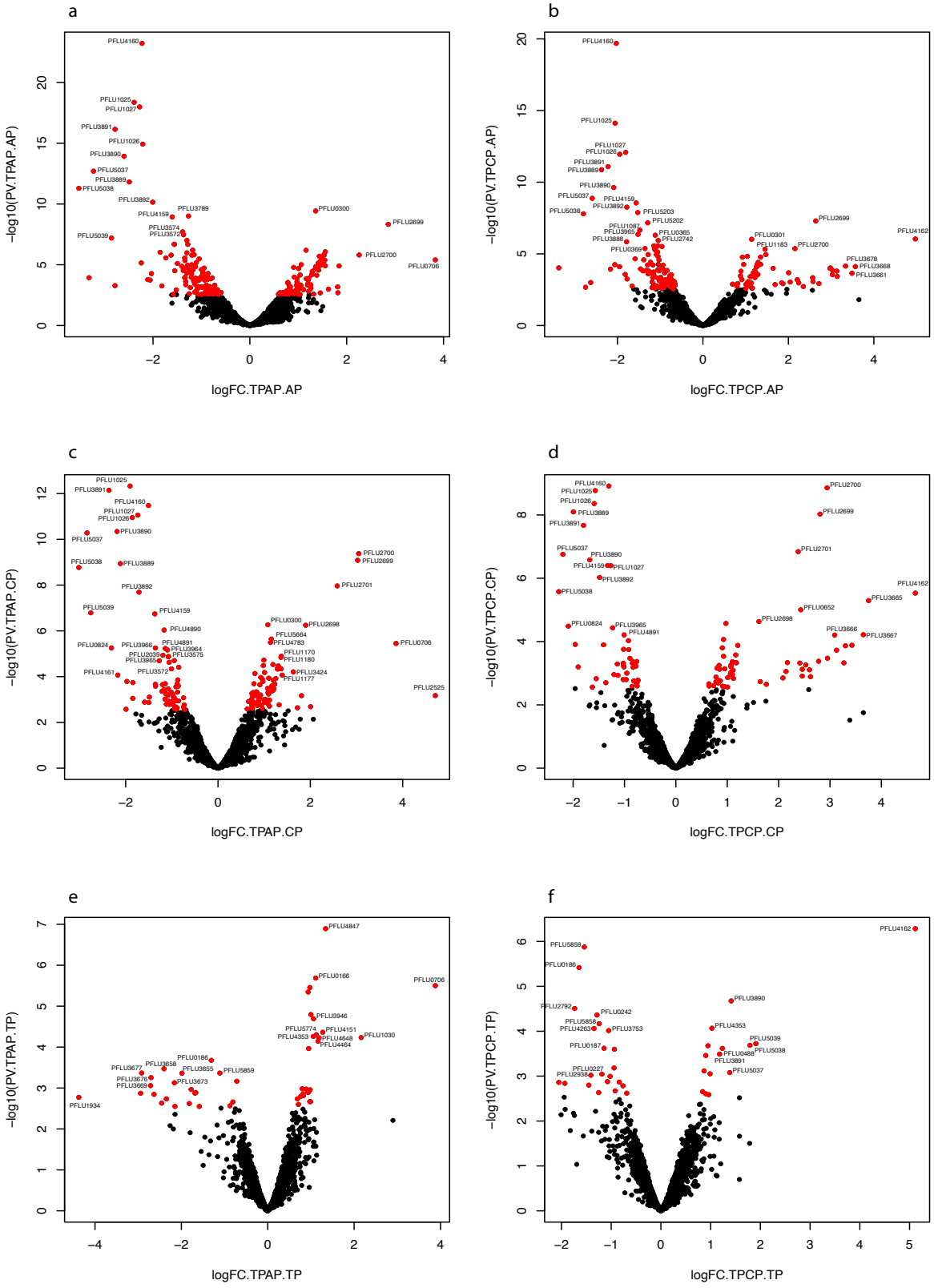


Figure 5.10 Significance levels (-log10 p-values) as a function of treatment difference (logFC) in single-predator/two-predator comparisons for (a) TPAP/AP (b) TPCP/AP, (c) TPAP/CP, (d) TPTPCP/CP, (e) TPAP/TP and (f) TPCP/TP contrasts.

5.4 Summary of main conclusions

When submitted to protistan grazing pressure, *Pseudomonas fluorescens* SBW25 developed specific colony morphologies linked to predation resistance. Transcriptome analyses of bacteria showing those characteristics illustrated how gene expression was regulated, and if we could relate phenotypic observations obtained from Friman *et al.* (2015) to genotypic changes. Overall, significantly differentially expressed genes were involved in metabolism and solute transport through the cell membranes as well as inside cell compartments. Those functions, while coded by different genes, were common to all contrasts studied. However, specific genes were involved for every individual contrast.

While no morphological difference was observed in the control lines throughout the selection experiment (B treatment lines compared to ancestral ones) adaptation to a novel environment only triggered important gene expression. As it happens, the original *P. fluorescens* SBW25, firstly isolated from soil, was grown in an aquatic environment for the experiment. Furthermore, the selection experiment that allowed the observation of bacterial phenotypic diversification, involved growing every protist predator in the presence of the prey. Bacterial changes were therefore related to the presence of predators. In order to perform RNA sequencing, enough material had to be obtained. To do so, bacterial lines obtained from the selection experiment were grown alone for 5 days prior to extraction. This situation most likely impact observed gene expression regulation in terms of bacterial (rapid) adaptation. Indeed, bacteria previously submitted to grazing pressure were suddenly free of predators. It is necessary to consider that, while heritable changes were actually present in the bacteria grown prior to RNA extraction, expression of genes regulated due to the conditions of incubation, i.e. protist presence, was observed instead of those regulated because of predatory pressure.

Signal transduction and chemotaxis-related receptors and proteins, as well as external membrane components were amongst most down regulated genes in

the bacteria only treatment. Usually, soil environments are more heterogeneous than aquatic ones, which requires the bacteria constant and regular sensing and monitoring of its environment, in order to move away or towards specific chemicals. When an attractant or repellent substance is identified – via receptors on the external cell membrane (lipoproteins), transported through the internal membrane (porins, permeases, ABC transporters) – the cell must relay the information to the motility systems (flagella and pili), which in turn requires phosphorylation/dephosphorylation or methylation/demethylation of enzymes. In a liquid medium gradients are less likely to happen, and environment probing need not be so frequent, reducing the necessity of large numbers of receptors, transport, signal transduction and all the intermediate 'participants'. Catabolic processes, such as proteolysis and amino acid degradation and catalysed by enzymes (dioxygenases, hydrolases, isomerases), were also amongst the most down-regulated genes in comparison to ancestral non-evolved lines. Those are typically energy yielding reactions, necessary to the maintenance of basic cell metabolism, in line with the reduction of active transport of molecules (ABC transporters) and signal transduction, but also responsible of breaking-down elements obtained from the outside. In parallel, up regulated genes were also involved in solute transport – necessary to import nutrients from the outside, that can be used in biosynthetic pathways (algD, purU2 fold), export waste or for maintaining homeostasis.

Single protist treatments changed SBW25 regulation differently relative to B treatment: while bacteria grown with all three predator protists individually had mostly up-regulated genes (fig. 5.4), expression driven by TP predation involved different ones than from CP and AP (fig 5.7 and 5.8). In fact, bacterial phenotypic defence resulting from the selection experiment differentiated colonies grazed by the ciliate or by either AP or CP. Indeed, in single-predator communities, TP was the only organism shown to have an impact on bacterial densities, triggering the development of the wrinkly-spreader (WS) generalist defensive colony morphotypes (as well as few 'petite' colony types (PT), but not very good against *T. pyriformis*). While the WS new phenotype was more effective against grazing, it also harboured lower growth rates compared to the

smooth colony types (SM) isolated from bacteria-only treatments: down regulation of catabolic traits coincides with growth cost data and the appearance of small colony variants observed in the selection experiment (Fig. 5.1). Meanwhile, up regulation of chemotactic functions indicates active environment monitoring, possibly in sensing the presence of predators. Similarly, cell-to-cell communication and alginate biosynthesis indicate biofilm formation and cell aggregation (membrane attachment sites, adhesion). Higher transcription levels of organisational elements such as cell membrane and flagella are also in accordance with chemotaxis/signal transduction systems, suggesting that the bacteria are regulating their movement, possibly in order to avoid predator or seek nutrients in unexplored areas. Interestingly, common up regulated genes in bacteria grown under all three predators (individually, compared to bacteria grown alone) are mostly the same down regulated ones in B-treatments compared to ancestral lines (figs 5.5 & 5.6, table 5.3), and are related to chemotactic functions and its related signal cascade. This suggests that while bacteria adapting to environmental changes only decrease sensing functions, those adapting to predation (TP, CP and AP treatments) 're-activate' those same functions.

Adding a second protist had a small effect on *P. fluorescens*' base level of expression (figs 5.10); however, it significantly changed which genes were regulated leading to unique bacterial populations. One and two predator-evolved bacteria differed mostly when compared to bacteria grown under TP as single predator (TPCP/TP and TPAP/TP), and fewer genes were differentially expressed than in contrasts involving CP and AP (TPCP/CP, TPAP/CP and TPCP/AP, TPAP/AP respectively). This indicates that, as generalist defenders, WS colony morphotypes present in both TP and TPCP require less regulation of gene expression. In addition to the WS morphotypes emergence (selection experiment), simultaneous predation by TP and CP selected for a 'petite' colony type (PT) that was highly defensive against CP but not against TP, and with lower growth capacity than the SM colonies from B treatments. Reduced growth rates are in line with reduced regulation of gene expression observed in TPCP/TP contrasts, including amino acid transport, catabolic process (less

energy needed) or sulphur acquisition (alkanesulfonate-related processes). Amongst up regulated genes, degradation of branched-chain amino acids (valine, leucine and isoleucine) were the most abundant. This usually leads to the biosynthesis of polyketides, common secondary metabolites usually involved in survival advantage (i.e. antibiotics, antifungals, etc), providing better grazing resistance to PT colony types. Similarly, when grown under both the ciliate and the amoeba, P. fluorescens developed a 'transparent' morphotype (TT), somewhat defensive against AP but susceptible to TP. Most down regulated genes involved biosynthesis of polypeptides and related transporters. This suggests reduced needs in cell (membrane) structure formation, in line with reduced growth abilities of defensive TT colony types. Conversely, initiation of transcription, amino acid and lipoproteins biosynthesis as well as secretion (transporter) systems have enhanced activities, indicating the bacteria is investing in possible defence or biofilm formation products.

Meanwhile, TPCP and TPAP treatments compared to AP had similar gene expression (up and down) than TPCP and TPAP compared to CP. Many of the down regulated genes in both TPAP and TPCP (compared to either CP or AP) bacterial strains were common to B/ancestor. They involved transporter activities, within the cell as well as transmembrane transport, catabolic processes and transcription regulation– indicative of reduced growth rates of different phenotypes observed in TPCP and TPAP microscosms. Conversely, most up regulated genes in TPCP/AP, TPAP/CP, TPCP/CP and TPAP/AP are unique to those contrasts. They include a number of genes with unknown function, as well as a few related to bacteriophage – possible indication of phage genome transduction. Otherwise, metabolic processes and biosynthesis of defence elements (aureothin antibiotic, secretion) show that bacteria grown with two grazers respond to predation pressure more actively and effectively than when only one protist is present.

In summary, despite the absence of observable phenotypic changes in control bacterial lines (non-grazed), transcriptome analysis indicated that adaptation to environmental conditions only triggered significant changes in gene transcription. This has been observed for longer, in studies seeking to understand single species adaptation to biotic and abiotic factors. Predator pressure, on the other had, also affects gene expression regulation, but at a much lower level in this experimental setting. Adaptation to one or two predators involved mostly gene down-regulation in line with results obtained on the selection experiment carried by Friman et al. (2014). Protists *C. paramecium* and *A. polyphaga* had similar patterns of transcription regulation, but *T. pyriformis* appeared to regulate gene expression in an opposite way compared to the other two predators. Indeed, CP and AP also observed similar phenotypic changes different from those in TP, in the selection experiment.

SUPPLEMENTARY MATERIAL

Table S5.1 Summary of raw and trimmed reads (before ad after adapter and quality trimming).

Sample	Barcode	Number of Raw Reads	Number (%) of Post trimming Reads	Number of Post trimming Read pairs	Number (%) of Post trimming Single reads
Sample_1-1	ATCACG	42,680,904	42,221,037 (98.92)	20,905,247	410,543 (0.97)
Sample_18-ANCESTRAL	GTGAAA	41,817,624	40,878,305 (97.75)	20,196,084	486,137 (1.19)
Sample_2-2	CGATGT	49,844,522	49,260,702 (98.83)	24,378,035	504,632 (1.02)
Sample_3-4	TTAGGC	43,616,386	43,162,596 (98.96)	21,380,605	401,386 (0.93)
Sample_4-5	TGACCA	50,578,860	49,937,177 (98.73)	24,704,188	528,801 (1.06)
Sample_5-6	ACAGTG	48,452,912	47,737,788 (98.52)	23,627,774	482,240 (1.01)
Sample_6-7	GCCAAT	46,712,418	46,148,324 (98.79)	22,845,507	457,310 (0.99)
Sample_7-8	CAGATC	39,711,710	39,238,643 (98.81)	19,433,684	371,275 (0.95)
Sample_8-9	ACTTGA	40,324,004	39,781,352 (98.65)	19,675,296	430,760 (1.08)
Sample_9-10	GATCAG	39,918,746	39,373,690 (98.63)	19,486,546	400,598 (1.02)
Sample_10-11	TAGCTT	44,055,860	43,365,729 (98.43)	21,382,769	600,191 (1.38)
Sample_11-12	GGCTAC	49,854,284	49,198,861 (98.69)	24,319,283	560,295 (1.14)
Sample_12-13	CTTGTA	47,822,660	46,744,300 (97.75)	23,084,434	575,432 (1.23)
Sample_13-14	AGTCAA	43,815,046	43,147,910 (98.48)	21,333,498	480,914 (1.11)
Sample_14-15	AGTTCC	56,857,730	55,995,565 (98.48)	27,684,745	626,075 (1.12)
Sample_15-16	ATGTCA	46,141,690	44,563,201 (96.58)	22,016,156	530,889 (1.19)
Sample_16-17	CCGTCC	45,781,528	43,981,618 (96.07)	21,722,702	536,214 (1.22)
Sample_17-18	GTAGAG	49,788,372	46,274,668 (92.94)	22,862,458	549,752 (1.19)

Table S5.2 Summary of sequence alignment to Pseudomonas fluorescens SBW25 genome

Samples	Total Reads	Mapped Reads	% of Mapped Reads	Reads Mapped as A Pair	% of Reads Mapped as a Pair
Sample_1-1	42,680,904	34,709,982	81.32	32,885,164	77.05
Sample_18-ANCESTRAL	41,817,624	36,195,280	86.56	34,185,880	81.75
Sample_2-2	49,844,522	38,398,206	77.04	36,417,314	73.06
Sample_3-4	43,616,386	37,207,610	85.31	35,080,406	80.43
Sample_4-5	50,578,860	41,507,180	82.06	39,086,922	77.28
Sample_5-6	48,452,912	39,082,056	80.66	36,821,714	75.99
Sample_6-7	46,712,418	38,231,583	81.84	36,224,164	77.55
Sample_7-8	39,711,710	32,331,282	81.41	30,531,636	76.88
Sample_8-9	40,324,004	33,400,399	82.83	31,434,812	77.96
Sample_9-10	39,918,746	34,331,620	86	32,469,162	81.34
Sample_10-11	44,055,860	33,835,326	76.8	31,627,962	71.79
Sample_11-12	49,854,284	36,674,146	73.56	34,376,554	68.95
Sample_12-13	47,822,660	40,165,241	83.99	37,485,480	78.38
Sample_13-14	43,815,046	35,406,146	80.81	33,193,456	75.76
Sample_14-15	56,857,730	47,228,359	83.06	44,420,976	78.13
Sample_15-16	46,141,690	37,281,670	80.8	35,186,880	76.26
Sample_16-17	45,781,528	36,004,022	78.64	33,906,884	74.06
Sample_17-18	49,788,372	36,868,685	74.05	34,833,934	69.96

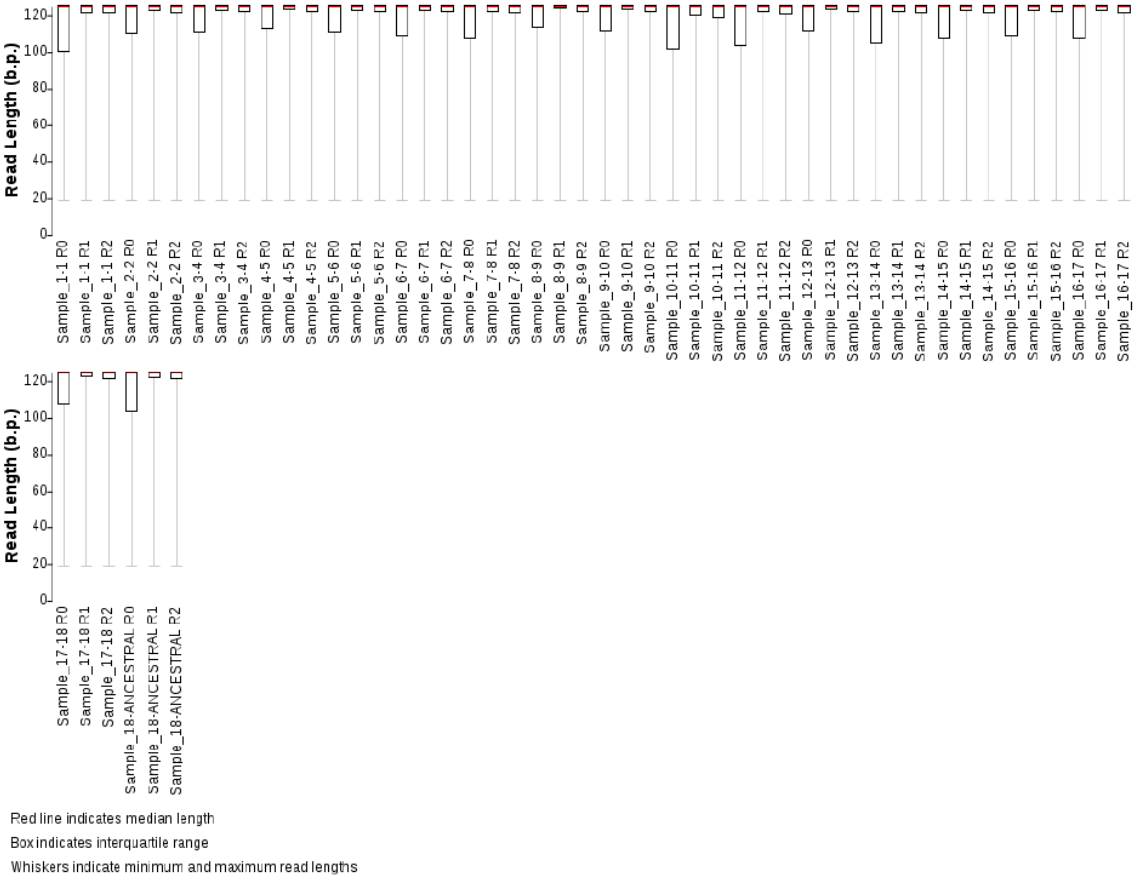


Figure S5.1 Read length distributions for all samples after adapter and quality trimming. R1 stands for paired forward reads and R2 for paired reverse reads; R0 indicates non-paired reads (singletons).

Chapter 6 PREDATOR-PREY DYNAMICS OVER LONG-TERM CO-EXISTENCE

This chapter was discussed with Dr Friman and my supervisors. Experimental design, data collection and treatment, as well as statistical analyses and results interpretation are my own work

6.1 Introduction

Ecosystem functioning is related to a variety of factors, biotic and abiotic. Variations of these factors directly and indirectly affect organisms that thrive in the environment. For example, an increase in temperature can directly impact a species with low tolerance to temperature change, and likely induce high mortality of a significant percentage of that species population. The indirect effect of temperature change is then mostly observed where the original species should have been: empty niches are quickly occupied by a new species. In this example, the ratio change (but not only!) of the original species to the new occupier influences the interactions not only between the new occupier and the already existing species, but also the interactions between all the other organisms previously present.

Understanding the dynamics of complex communities is one of the major goals of ecology. Therefore, organismal interactions have been studied for longer in both theoretical and empirical studies. Many mechanisms such as predation and competition have been well characterised over the years of ecological research, and now, evolutionary changes resulting in better 'resistance' (to environmental variations and/or predation and competitive pressure) – i.e. natural selection – have been shown to also influence the inter- and intraspecific interactions. The heritability of such traits over time requires adaptation of the co-existing species which have a new effect on the surrounding environment, creating a constant feedback between ecological and evolutionary processes governing natural communities, thus driving 'eco-evolutionary dynamics' (Pelletier *et al.*, 2009).

Somewhat recent studies have shown that many species observe rapid changes of heritable traits when faced with short-term environmental fluctuations. This rapid evolution can greatly impact other species in a community (Yoshida *et al.*, 2003; Ezard *et al.*, 2009). In particular prey rapid evolution has been showed to affect the dynamics of predator-prey interactions both in theory and experimental studies (Friman *et al.*, 2008; Hiltunen and Becks, 2014). And again, changing interactions between species ultimately

affect the surrounding environment, creating a new environment, to which species must adapt again. Evolution in this context can drive a total change of the community.

In the presence of bacterivorous predators, bacteria have been shown to develop defence mechanisms. Those can be achieved by a variety of strategies, which include morphological changes – increased cell size renders ingestion by the predator more difficult, while smaller cells have reduced encounter rates with other cells, including predatory ones – chemical cues, increased swimming speed or swarming and clumping. Selection for competitive traits, however, requires a trade-off between resource allocation – for the production of toxins for example, or larger cells – and growth rate, but also when it comes to investing in specific defence effective against one single grazer, or generalist defence. Friman and colleagues (2015) showed that the bacterium *P. fluorescens* SBW25, when in presence of the ciliate *T. pyriformis* and the cryptophyte *C. paramecium*, developed colonies effectively protected against *T. pyriformis* but vulnerable to *C. paramecium*. In parallel, when in the presence of the amoeba *A. polyphaga* and the cryptophyte *C paramecium*, *P. fluorescens* SBW25 colonies were well defended against both protists but observed a much lower growth rate compared to the ancestral bacterial strain. The diversification of the bacterial colonies then affected the whole community, by rendering it less stable and less productive.

In a system where prey productivity and stability is reduced due to fitness trade offs, predators are then faced with prey that are not only better defended, but also potentially less edible or nutritionally poor. In order to thrive, the predator needs to develop its own mechanisms to overcome bacterial defence. In microbial systems, predator adaptations have been shown in parasite-host (Friman and Buckling, 2014) and bacteria-phage (Scanlan *et al.*, 2015) situations, where strong pairwise interactions drove evolution through an arms race scenario. In protist-bacteria trophic interactions, protist can recognise their prey with chemical recognition for example. But when bacteria can inhibit surface recognition – by down-regulating the expression of cell surface

elements, as observed in the previous chapter – predators adapt to evolving prey.

In the previous chapter, bacterial rapid evolution previously observed in terms of phenotypic changes and community dynamics were further investigated. By comparing *Pseudomonas fluorescens*' SBW25 morphological adaptation when grown under varied predation pressures to the transcriptome of the bacteria obtained from the same experimental conditions, we aimed to related genotypic changes to morphological adaptation. Initially, *P. fluorescens* presented a great number of differentially expressed genes when cultured in a new environment. But more interesting yet, changes in gene expression of the different bacterial lines analysed – chosen as the ones presenting an observable defensive morphology – were very much in line with strategies of defence observed in other studies. The presence of additional predators in this system, however, did not have a major impact in gene expression. While it is clear that in this case the prokaryotic prey rapidly developed defence mechanisms possibly in order to avoid predation, little is known about the protistan predator strategy. Co-evolution in interacting species has been extensively studied, as well as the effects of interactions in populations involved. However, in systems where rapid evolution is observed, studies generally focus on prey evolution or the outcome of whole system (Friman *et al.*, 2008; Ellner and Becks, 2011; Lawrence *et al.*, 2012). Only one study considering the point of view of the (protist) predator had been found at the time this study was held (Hiltunen and Becks, 2014).

Understanding the mechanisms of co-evolution of protists submitted to rapidly evolving bacterial prey in stable environmental conditions adds one more level to the still growing knowledge of eco-evolutionary dynamics. The better complex ecosystems are studied, the better outcomes can be predicted in situations of drastic environmental changes, more and more common as the 21st century unfolds in the era of climate change (IPCC, 2013).

To test the co-evolution of protist predators and their bacterial prey, one predator – one prey microbial microcosms were set over thousand of bacterial generations. A variety of unicellular eukaryotes, representing different ecological

niches and feeding strategies were fed with gram-negative *Pseudomonas fluorescens* strain SBW25 bacteria for over 6 months. Growth rates of evolved and naïve protist predators on both evolved and non-evolved (ancestral) SBW25 was tested at the end of the experiment.

6.2 Methods

Predator-prey systems – fourteen pure protist strains and a bacteria-only control line (table 5.4) – were maintained over hundreds of protist generations, from 20/04/2015 to 19/11/2015 (213 days), with bacterial prey *P. fluorescens* SBW25.

6.2.1 Microbial microcosms

Initially, 1μl of dense SBW25 bacterial suspension were added to four 24-well cell culture plates (84 wells) with 2ml 0.1% Luria-Bertani medium (15g.L-1 Oxoid L24 LB broth powder) and allowed to grow for 24h in 24°C, unshaken conditions. In order to maintain the 2ml initial volume, 300μl were removed from each well, and 700μl of dense protist strains (table 5.3.1) were added to the microcosms (fig 5.8), except for the extremely dense A. polyphaga stock culture of which only 500μl were used.

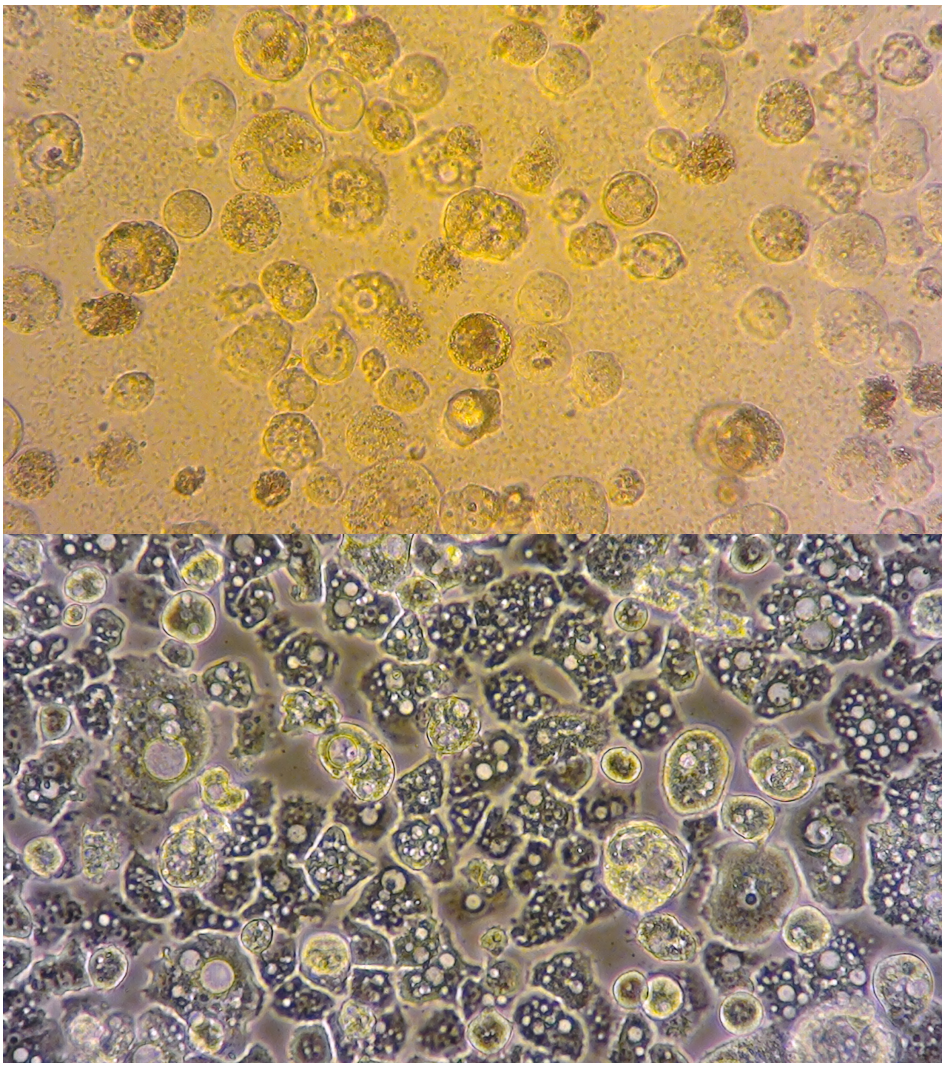


Figure 6.1 *A. polyphaga* cultures at the time of cross-feeding experiments incubation (above) and at the first time-point counting (below). Photo area 0.916 mm2

Bacterial and protist visual density assessments were held every 7th day, and ranked on a scale from 1 (very few bacteria or protists) to 4 (high density). After each inspection, 1ml of medium from every microcosm was collected, of which 200μl were transferred into 96-well plates and frozen in 20% glycerol at -80°C at the Natural History Museum's Molecular Collection facility. 1μl of the collected medium was plated onto LB agar (LB agar Oxoid powder) for bacterial morphological diversification analysis (fig 5.9). 1ml of fresh 0.1%LB medium was replaced into every well.

Originally, 1ml protist samples were taken every 10th day for freezing in 7.5% DMSO according to Product Information Sheet for ATCC® 50366TM , and the

medium replaced in the microcosms. However most of the eukaryotes did not survive either the freezing or thawing process, so this procedure was terminated.

6.2.2 Predator and prey co-evolution measurements

At the end of the co-evolution experiment, protist strains that were pure (uncontaminated) and alive across three or more replicates were selected for a cross feeding experiment with both ancestral and evolved bacterial lines. Seven protist strains out of the initial fourteen were used: *Cercomonas effusa* (Beaver-Creek), *Cercomonas paraglobosa* (19-3E), *Bodo saltans*, *Poterioochromonas* sp., *Tetrahymena pyriformis*, *Tetrahymena vorax*, *Acanthamoeba polyphaga* and *Paracercomonas saepenatans* (Ca5HKv). Bacterial strains were selected on a morphological basis, observed at the last collection point: smooth (SM) bacterial lines were isolated from one control (non-grazed) culture, wrinkly spreader (WS) were isolated from a *E. uvella* culture and fuzzy spreader (FS) colony types from a *Bodo saltans* microcosm. Ancestral bacterial and protist lines were recovered from stock cultures (see chapter 2).

Every selected protist strain, ancestral and evolved, was incubated with all three selected evolved as well as ancestral bacterial strains in 24-well plates with 0.1% LB medium, for seven days at 24°C, unshaken conditions. Protists were counted under an inverted microscope at 20h, 93h, 115h and 127h (evolved protists only) in average1. Bacterial growth was measured after 24h and 48h growth (in 96-well plate with 200μl 0.1%LB, 24°C unshaken conditions) with the FLUOstar OPTIMA microplate reader (BMG LABTECH) as absorbance at 600nm.

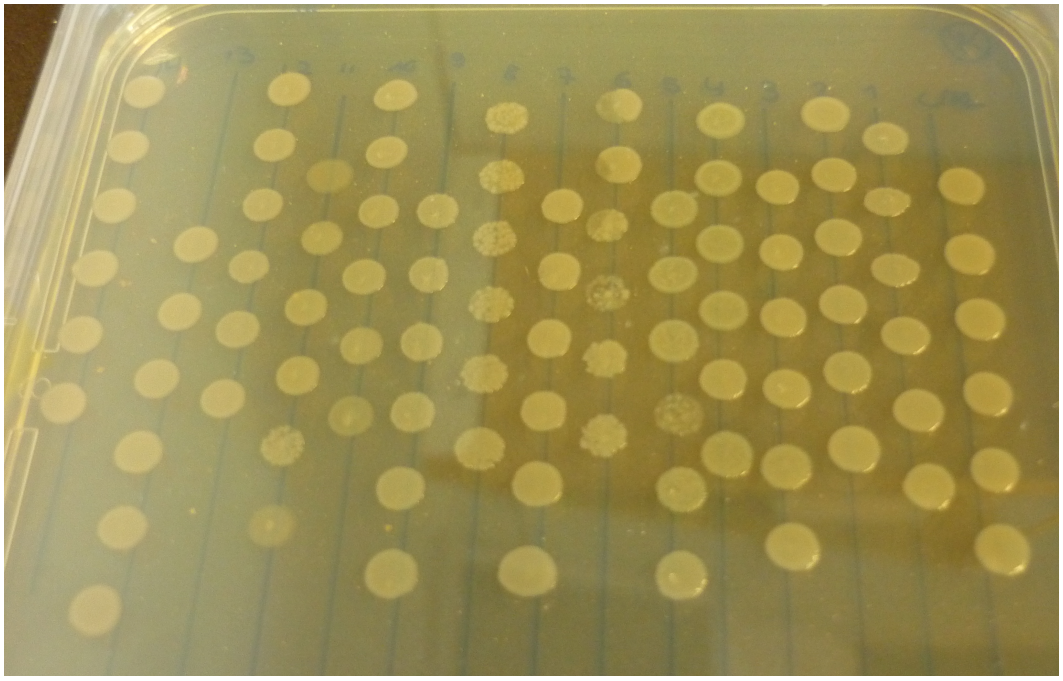


Figure 6.2 Bacterial morphological diversification assessment (13/06/15) for all six replicates for all treatments (left to right: *P. oxoniensis*, *C. saepenatans*, PML5D, *A. polyphaga*, *C. paramecium*, *T. vorax*, *T. pyriformis*, *Poterioochromonas* sp., *B. saltans*, *P. minima*, *E. uvella*, *C. paraglobosa*, *C. effusa*, *C. pigra* and control).

6.2.3 Statistical analyses

Bacterial growth was analysed as variance differences between the different phenotypes at every time point with one-way ANOVA.

Protists mean density for every bacterial treatment (ancestral, SM, WS and FS) was analysed at every time point with one-way ANOVA for all protist strains, ancestral and evolved. Protist bulk density differences between bacterial treatments were analysed as differences of area under the curve (AUC) for every protist strain, between bacterial treatment and protist evolutionary state (evolved versus ancestral) Difference of AUCs for different treatments was tested with one-way ANOVA.

In addition, growth rate – as the slope of density curves – of naïve and evolved predators was calculated. Growth rate comparisons between the two lines (naïve vs. evolved) of every protist species, fed with each evolved bacterial line (SM, WS and FS) were done with one-way ANOVA.

6.3 Results & discussion

6.3.1 Long-term protist dynamics

While all other protists harboured typical growth dynamics (table 5.5): starting at low to medium densities, reaching a peak density towards the end of the second month and keeping it stable for a couple weeks before declining in numbers, *C. paraglobosa* – a medium to large sized cercomonad (5-15μm) – maintained the highest density levels throughout the experiment, declining slightly towards the end. *Poterioochromonas*, on the other hand, described a slower initial growth phase, reaching higher density levels later in time when compared to all other protists. However, the stramenopile has the tendency to form very dense large clumps of cells when growing, often separated by areas almost devoid of cells, indicative of locally high densities but relatively rare over the whole microcosm. This could lead to underestimating the real density of *Poterioochromonas* when screening microcosms.

The ciliate *T. pyriformis*, a very active pelagic species, declined in activity and density levels. Cells that were initially in constant movement became immobile, although actively moving cilia and creating particle flow. Furthermore, TP microcosms, and *T. vorax* ones to a lesser extent, harboured much detritus, very difficult to determine whether faecal particles or dense bacterial clumps. *T. vorax* developed mostly pointed-tail cells and a few macrostome morphologies from the second week of the experiment (first observed on the 18/05/15; table 5.5); the ancestral oval shape was very rare at the end of the incubation time.

Table 6.1 Long-term and cross feeding experiment treatments. Numbers indicate the protists used as predators of P. fluorescens SBW25, ✗ indicates strains not used in the cross feeding experiment.

Protist	Protist code	Long term evolution treatment	Cross-feeding experiment treatment
None (bacteria only)	-	B	✗
Cercomonas pigra	CaSphII	1	✗
Cercomonas effuse	Beaver-Creek	2	✗
Cercomonas paraglobosa	19-3E	3	1
Eocercomonas uvella	11-7E (Spain)	4	✗
Paracercomonas minima	SW2	5	✗
Bodo saltans	-	6	2
Poterioochromonas sp.	-	7	3
Tetrahmena pyriformis	Tp	8	4
Tetrahymena vorax	Tv	9	5
Chilomonas paramecium	Cp	10	✗
Acanthamoeba polyphaga	Ap	11	6
Chrysophyte	PML5D	12	✗
Paracercomonas saepenatans	Ca5HKv	13	7
Paracercomonas oxoniensis	WA8	14	✗

Of all 14 initial protists used, *Chilomonas paramecium* (CP) and the chrysophyte PML5D did not survive the whole length of the experiment (table 5.5; appendix). *Cercomonas pigra* (CaSphII) only survived in microcosms cross-contaminated with *Cercomonas effusa* (Beaver-Creek), but develop cysts that died in pure cultures. Similarly, *Paracercomonas minima* (SW2) cultures were cross-contaminated by *Bodo saltans* – which did not influence the sarcomonad survival – or with *Poterioochromonas*, which drove extinction of SW2. *E. uvella* also grew and maintained 'medium' density levels when contaminated with *C. paraglobosa*, and developed cysts (only) in pure strains. Cross-contamination of *C. effusa* and *Poterioochromonas* by other strains did not affect their growth, as both strains survived well in both contaminated and clean replicates. Similarly *A. polyphaga* equally rapidly declined in density then developed cysts in both clean and contaminated replicates.

Table 6.2 Protist densities and survival in microcosm conditions: very dense (++++), dense (+++), dense/medium dense (++(+)), medium dense (++), medium/low dense (+(+)), low (+), very (low). X indicates death of all protist cells; ! indicates cross-contamination of at least on replicate; * indicates density of protist in presence of contaminant; c indicates the presence of cysts only in all replicates.

Date	*C. pigra*	*C. effus*	*C. paraglobosa*	*E. uvella*	*P. minim*	*Bodo saltan*	*Poteriooch romonas*	*T. pyriformis*	*T. vorax*	*Chilomonas paramecium*	*A. polyphag*	PML5 D	*C. saepenat*	*P. oxoniensi*
24.04	++	++	c +++	+++	+	++	+	++	++	+++	++++	+	+	+
01.05	++	++	c +++	+++	+(+)	++(+)	++	++	++	+++	+++	+	+(-)	+
06.05	++(+	++(+	+++	++(+)	++(+)	++	++		++	++(+)	! +++	+(+)	+(-)	+
13.05	++(+	++(+	+++	++	+++(+	+(+)	+(+)	++	+(+)	++	! ++(+)	+(+)	+(-)	+(-)
18.05	+++	++(+	+++(+)	++(+)	+++	++	+(+)	++	++	+(-)	! ++(+)	++	++	+(+)
27.05	+++	+++	+++(+)	+++	+++	++(+)	++	+	++(+)	X	! +++	++	! ++	++
05.06	++(+	++(+	+++(+)	+++	+++	++	++	++(+)	++(+)	X	! +(+)	++	! +++	+++
15.06	! +++	++	+++(+)	++(+)	+++	++	++	++	++	X	! +(+)	++	! +++	+++
19.06	! +++	++(+	+++(+)	++(+)	+++	++(+)	+++	++(+)	+++	X	! +(+)	+	! +++	+++
29.06	!	+(+)	++++	++	+(+)	++(+)	! +++(+)	++++	+++	X	! ++	+	! +++	+++
09.07	!	++	+++(+)	+(+)	! +	+++(+	! ++++	+++	+++(+	X	! ++	X	! +++	+++
18.07	!	! ++	+++(+)	! ++	! +	++	! +++	+++	+++	X	! ++	X	! +++	+++
22.07	! +++	! ++	+++	! +	! +(-)	+	! +++	++	+++	X	! c ++	X	! +++	+++
28.07	!	! ++	+++	! +	! +(-)	+(+)	! +++	++	++(+)	X	! c +(+)	X	! ++	+
02.08	! ++	!	+++(+)	! +(-)	! -	++	! +++	++	++	X	! c +(+)	X	! ++	+(-)
06.08	!	!	+++	! +(-)	! -	++	! +++	++	++	X	! c +(+)	X	! ++	+(-)
21.08	!	!	++++	! ++	! -	++	! +++	++	++(+)	X	! c ++	X	! ++(+)	+(-)

Date	*C. pigra*	*C. effus*	*C. paraglobosa*	*E. uvella*	*P. minim*	*Bodo saltan*	*Poteriooch romonas*	*T. pyriformis*	*T. vorax*	*Chilomonas paramecium*	*A. polyphag*	PML5 D	*C. saepenat*	*P. oxoniensi*
26.08	! +	! ++	+++(+)	!	! -	+(+)	! +++	++	++	X	! c +(+)	X	! ++	-
01.09	! +	! +	+++(+)	! ++	++*	+(+)	! ++	++	++(+)	X	! c +	X	! ++	+
15.09	+/++	!	++++	! ++	++*	+(+)	! +(+)	++	++(+)	X	! c +	X	! ++	++(+)
06.10	-/++*	!	+++(+)	++/++	+(+)*	++	! +(+)	+(+)	++(+)	X	! c +(+)	X	! ++	+++
26.10	X/++*	! ++	+++	++/++*	++*	++	! +(+)	++	++(+)	X	! c +	X	! ++	++(+)
15.11	X/++	! ++	++(+)	+(+)*	++*	++	! +(+)	++	++(+)	X	! c +	X	! ++	++(+)

6.3.2 Protist post-evolution adaptation

Protist adaptation to bacterial defence evolution was measured by comparing growth of both ancestral non-evolved and 'bacteria-evolved' protists fed with either ancestral non-evolved bacteria and one of the three selected evolved bacterial morphotypes SM, WS and FS. Evolved and ancestral protist strains observed different growth dynamics according to which bacterial phenotype they were exposed to (figs 6.3, 6.4, 6.5, 6.6 and 6.7). In order to test for possible co-evolution, growth rate was compared between ancestral and 'evolved' protistan lines for time intervals of effective growth.

Cercomonas paraglobosa (19-3E)

Ancestral *C. paraglobosa* (19-3E) reached a maximum of 307 cells.mm^{-2} when grazing on WS bacteria but the evolved line only grew up to 74 cells.mm^{-2} when fed with ancestral SBW25 (fig 6.3). Despite the ancestral *C. paraglobosa* maximum density being almost four times higher than those reached by evolved ones, growth (as area under the curve) between ancestral and evolved protist lines was not significantly different. Furthermore, ancestral *C. paraglobosa* growth did not differ when feeding on different bacterial ancestral and evolved phenotypes (non-significant protist density or AUC differences for every bacterial treatment), despite dynamics for SM-fed protists presenting decline from 95h of incubation. Evolved *C. paraglobosa* presented an initial growth phase for all bacterial treatments (Ancestral, SM, WS and FS) followed by a rapid decrease in numbers for lines grown with ancestral (most important decline) and smooth (SM) phenotypes, but lower for those fed with WS and FS phenotypes, although not significantly different between any feeding treatment at any time.

In terms of growth rate, ancestral 19-3E fed with evolved bacterial morphotypes {SM} observed a significant better growth than its evolved counter after 24h of co-existence with the bacteria (ANOVA $F_{1,4}$=7.989, p=0.0475). The same scenario was observed for growth rates of 19-3E fed with evolved bacterial phenotype {WS} (ANOVA $F_{1,4}$=7.741, p=0.0497) and morphotypes {FS} ($F_{1,4}$= 50.12, p= 0.0021).

The protozoan *C. paraglobosa* (19-3E) appears to be hindered by the bacterial diversification after co-existence, while naïve predators seem to feed efficiently on both evolved and ancestral bacterial types. Therefore, in the absence of apparent co-evolution in this experimental, protist grazers would most likely go extinct.

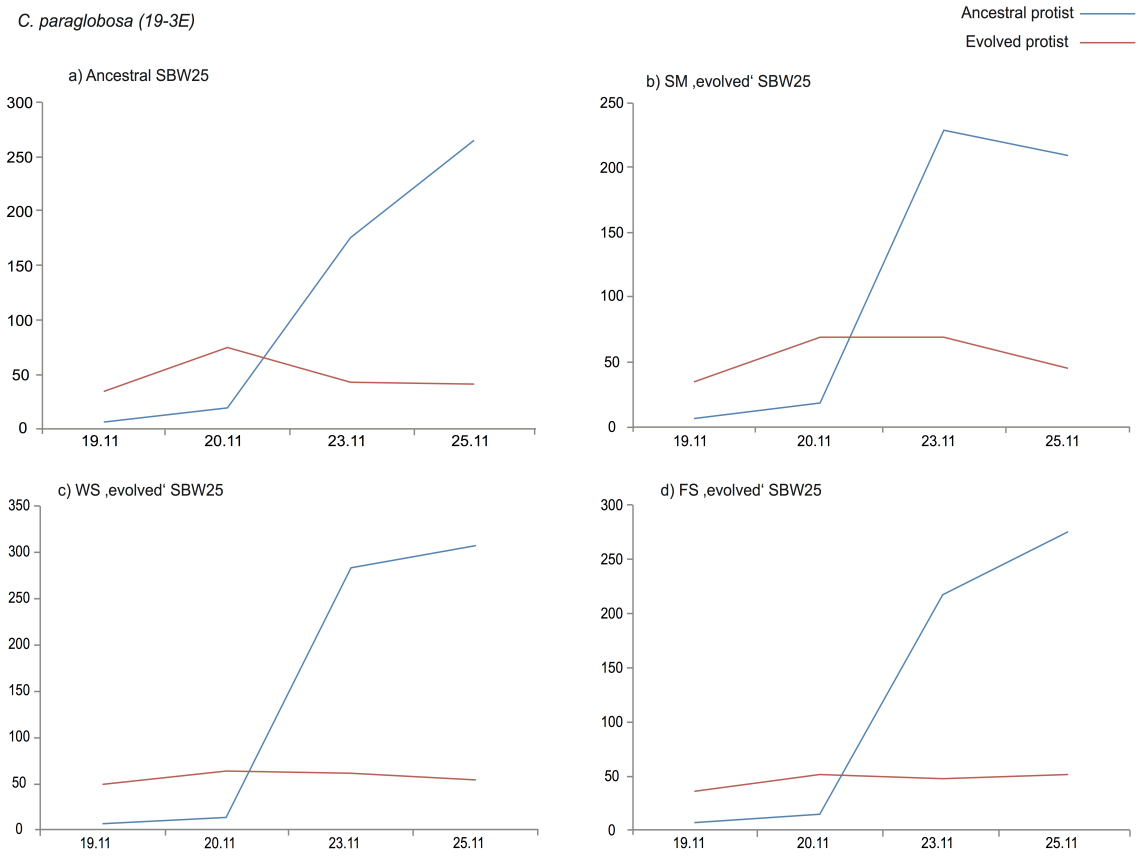


Figure 6.3 *C. paraglobosa* (19-3E) growth before co-existence with *P. fluorescens* SBW25 (blue lines) and after co-existence with SBW25 (red lines) in the presence of a) ancestral SBW25, b) SM phenotype of 'evolved' SBW25, c) WS phenotype of 'evolved' SBW25 and d) FS phenotype of 'evolved' SBW25.

Poterioochromonas sp.

Ancestral *Poterioochromonas* lines did not grow significantly differently when feeding on different bacterial ancestral and evolved phenotypes. Furthermore, comparison of growth rates (slope) between ancestral and evolved lines of *Poterioochromoas* sp. fed with all three bacterial evolved morphotypes individually (fig 6.4) did not differ significantly.

Evolved lines grew well at first and started slowly to decline after 20h, for all feeding treatments (fig. S6.2), at significantly different densities at 90h of incubation (ANOVA $F_{3,32}$=6.569, p=0.0014). Protists grown with FS and WS morphotypes reached highest population densities at 112h of incubation, significantly different from both ancestral and SM treatments (ANOVA {FS} $F_{3,32}$=4.162, p=2.22e-4 and {WS} $F_{3,32}$=2.759, p=0.0095). Evolved *Poterioochromonas* sp. growth was affected by bacterial treatment (ANOVA $F_{3,8}$=181.6, p=1.07e-7): comparison of growth (AUC) between protists incubated with different SBW25 phenotypes indicated the lowest growth for predators fed with ancestral bacterial lines ($F_{3,8}$=3.573, p=0.00726). Evolved *Poterioochromonas* fed with the FS phenotype had highest growth rate when compared to all other treatments ($F_{3,8}$=19.260, p=5.48e-8). Similarly, the evolved protist grew significantly better than the ancestral line only when incubated with fuzzy-spreader bacterial morphotypes (FS; $F_{1,3}$=116, p-value=0.001713). All other bacterial treatments (ancestral, SM and WS) did not affect growth between ancestral and evolved protists. The fuzzy-spreader bacterial morphotype, although not specifically defensive against ancestral *Poterioochromonas*, appeared more sensitive to the evolved microeukaryote, which is in line with the rare development of this phenotype in *Poterioochromonas* microcosms.

While *Poterioochromonas* total growth (area under the curve) over the cross-feeding experiment was significantly better after co-existence, when fed with FS bacterial lines, growth rate of either protist line wasn't. However, due to a missing data point for ancestral protist feeding experiments, this result must be considered carefully.

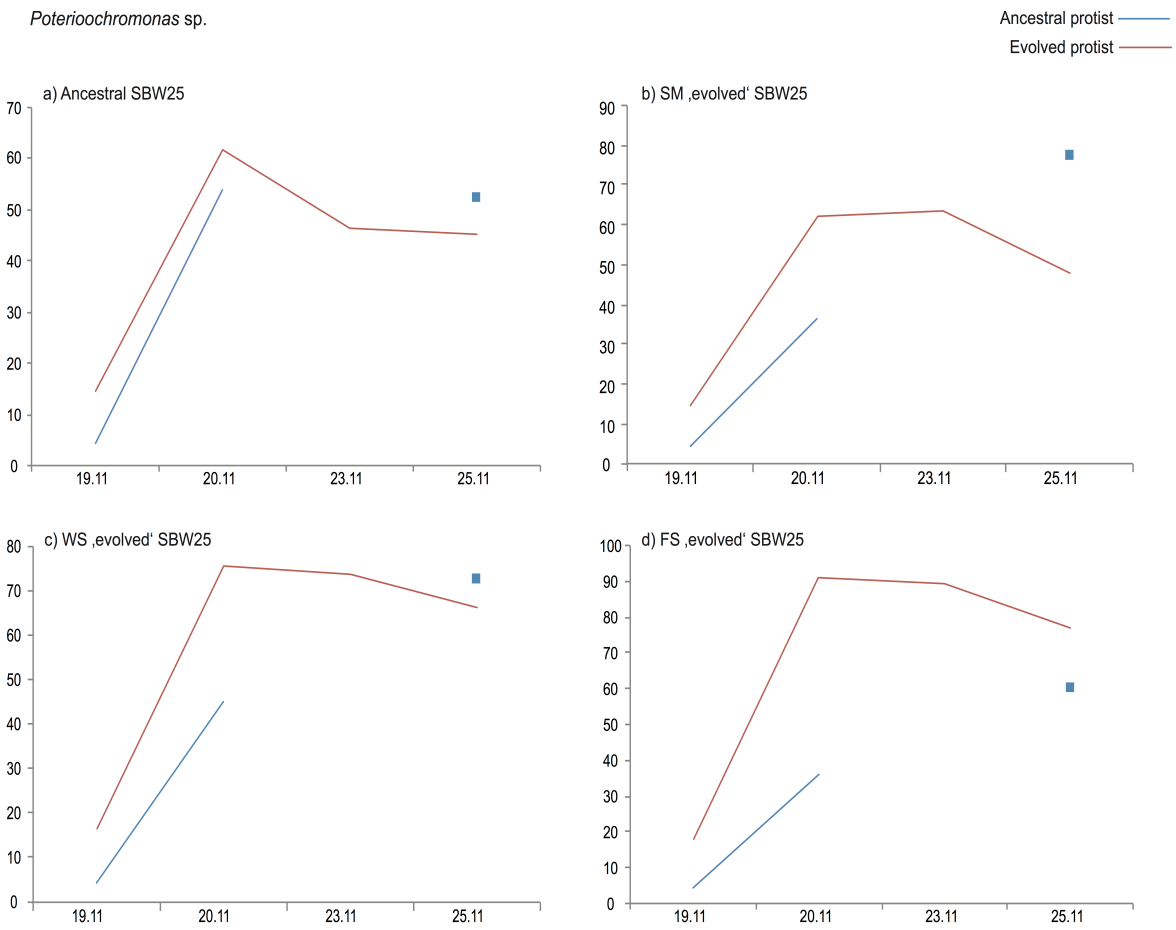


Figure 6 4 *Poterioochromonas* sp. growth before co-existence with *P. fluorescens* SBW25 (blue lines) and after co-existence with SBW25 (red lines) in the presence of a) ancestral SBW25, b) SM phenotype of 'evolved' SBW25, c) WS phenotype of 'evolved' SBW25 and d) FS phenotype of 'evolved' SBW25.

Bodo saltans (BS)

Ancestral *B. saltans* (BS) started declining after 95h of incubation for lines fed with ancestral, SM and FS bacterial types, but maintained growth for the line feeding on WS bacteria (fig. S6.1). SBW25 morphotypes impacted ancestral BS growth differently according to time after incubation ($F_{3,32}$=5.404, p=4.62e-3): after 23h protist grown with SM presented the lowest concentrations ($F_{1,4}$=, p=0.0165), although these could be partially due to unbalanced spread variability of residuals between the different bacterial treatments. Similarly, ancestral BS grown with FS bacterial lines (isolated from one *B. saltans* microcosms for the cross-feeding experiments) reached highest concentrations after 95h of incubation, and those grown with WS lines at 117h, but not significantly different from those grown with other bacterial lines (p≥0.05).

Evolved BS observed similar growth profiles when feeding on bacterial phenotypes SM, WS and FS; protists fed with ancestral bacteria harboured small differences in density from the other treatments after 20h (ANOVA $F_{3,32}$=2.979, p=0.046), starting to decline at that point while BS fed on evolved bacteria continued to grow. BS fed with FS bacterial phenotypes showed their highest density after 90h of incubation but declined after that; SM and WS continued to grow. Some bacterial phenotypes significantly impacted evolved protist density after 90h incubation (ANOVA $F_{3,32}$=15.39, p-value=2.254e-7). Evolved BS fed with SM and WS bacteria reached significantly higher densities after 112h of incubation (SM: $F_{3,32}$=4.714, p=4.56e-5 and WS: $F_{3,32}$=4.770, p=3.87e-5). Growth rate between evolved BS incubated with the every different bacterial line however did not differ. In parallel, ancestral BS fed with ancestral *P. fluorescens* showed better growth rates than the evolved protist with the same bacterial line (ANOVA $F_{1,4}$=55.59, p=0.001729).
Otherwise, there was no significant difference in growth rate between evolved and naïve predators, when fed with one of the three bacterial evolved phenotypes (S, WS and FS).

These results indicate possible specific protist adaptation to bacterial defence that is not effective against the original bacterial line. Evolved protists reached different concentrations according to bacterial treatment: *B. saltans* fared better when feeding on WS and SM morphotypes, lines with which the protist reached lowest density levels. Ancestral *B. saltans* observed lowest population density at 23h of incubation with SM bacterial types, and higher density with FS at 95h of incubation, reflecting that the bacterial line effective against the evolved protist was not so when exposed to the ancestral line, and vice-versa.

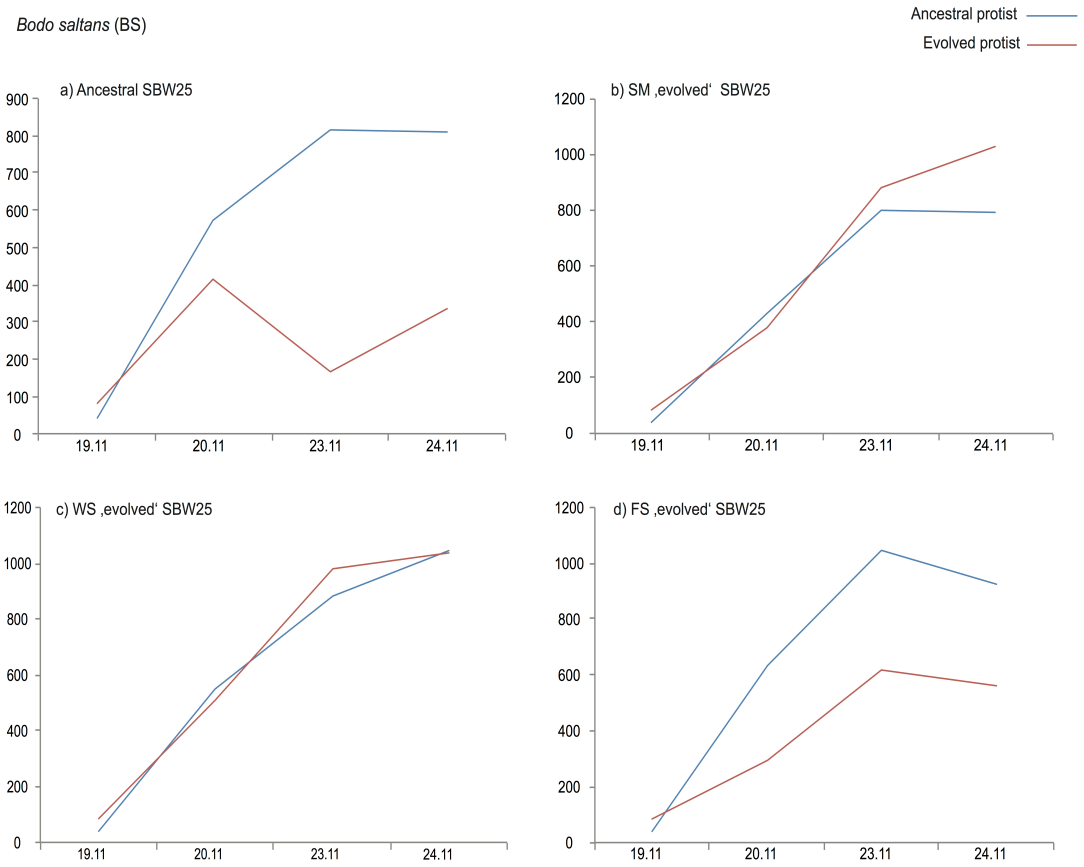


Figure 6.5 *Bodo saltans* (BS) growth before co-existence with *P. fluorescens* SBW25 (blue lines) and after co-existence with SBW25 (red lines) in the presence of a) ancestral SBW25, b) SM phenotype of 'evolved' SBW25, c) WS phenotype of 'evolved' SBW25 and d) FS phenotype of 'evolved' SBW25.

Tetrahymena pyriformis (TP)

T. pyriformis reached higher population density before coexistence with SBW25 (fig 6.6): ancestral TP observed significantly different rates when grown with different bacterial lines (ANOVA $F_{1,4}$=12.21, p-value=0.002352); FS and WS bacterial phenotype sustained higher protist population density while ancestral and smooth (SM) SBW25 did not impact protist growth. After 23h of incubation with ancestral, SM and WS bacterial types and after 95h when fed with the evolved fuzzy-spreader (FS), ancestral *T. pyriformis* density decreased. The ciliate reached highest densities at 95h (ANOVA $F_{3,32}$=6.3071, p=0.001745) and 117h (ANOVA $F_{3,32}$=20.564, p-value=1.304e-7) of incubation when in presence of WS and FS bacterial morphotypes. In contrast, evolved TP decreased in density following inoculation until 23h (most markedly when fed on SM), then

increased for all feeding treatments similarly, but grew less well than ancestral TP in all bacterial treatments.

Only SM bacterial types significantly affected growth of evolved TP when compared to the ancestral protist line (ANOVA $F_{3,32}$=12.58, p-value=0.0239). Growth rate comparison of ancestral and evolved protists fed with evolved *P. fluorescens* morphotypes SM does not indicate significant difference. This is however most likely due to the large spread of data for that treatment. As for naïve and evolved TP growth rates, there was a significance difference for both remaining feeding treatment, with the ancestral protist lines observing a much more important growth rate initially (WS: $F_{1,4}$=12.47, p=0-0242 and FS: $F_{1,4}$=163.8, p=2.15e-4).

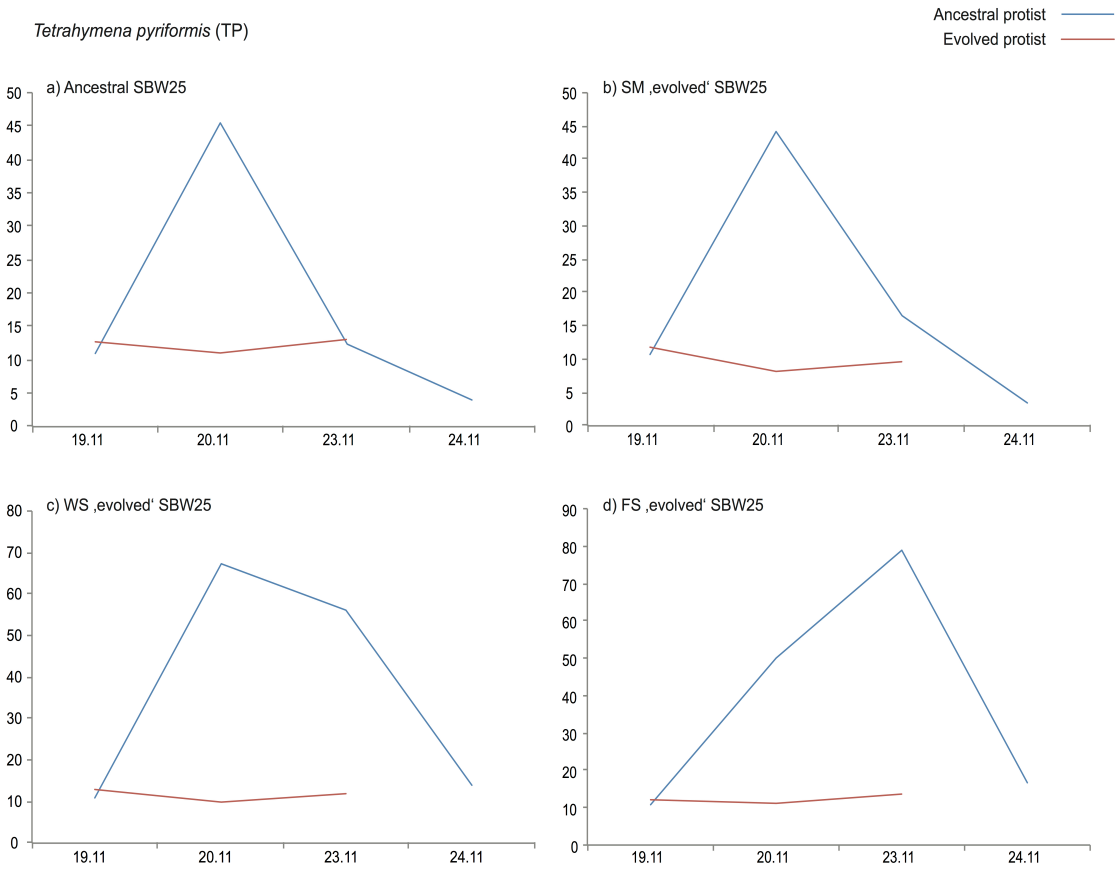


Figure 6.6 *T. pyriformis* (TP) growth before co-existence with *P. fluorescens* SBW25 (blue lines) and after co-existence with SBW25 (red lines) in the presence of a) ancestral SBW25, b) SM phenotype of 'evolved' SBW25, c) WS phenotype of 'evolved' SBW25 and d) FS phenotype of 'evolved' SBW25.

This shows that bacteria and protist evolved in an arms race scenario, and when defensive SBW25 were submitted to non-evolved grazer, it was not capable of defending itself anymore.

Cercomonas saepenatans (Ca5HKv; CS)

Ancestral *C. saepenatans* (CS) demonstrated enhanced growth after 23h of incubation when fed with ancestral, SM and WS bacteria, and after 95h only for FS treatments (fig 6.7). CS grown with evolved bacterial type WS only presented different density at 23h and 95h after incubation ($F_{3,32}$=3.069, p=0.004 and $F_{3,32}$=2.804, p=0.008 respectively) from the other bacterial treatments. Evolved CS showed similar initial growth to its ancestral form (up to 20h) followed by a stabilisation in density after 20h of incubation when fed with WS and FS, declined on ancestral SBW25 and increased importantly – but not significantly – on SM. The WS bacterial phenotype used for the cross-feeding experiments was isolated from one C. saepenatans microcosm, but the protist didn't appear to be better adapted to it after co-existence (fig. 5.11), although the WS bacteria seemed more vulnerable to the ancestral CS (fig. 5.10).

Growth (in terms of AUC) of both ancestral and evolved *C. saepenatans* was not affected by bacterial treatment, nor between protist lines (ancestral or evolved) fed with either ancestral or evolved bacterial lines. In a similar way to *C. paraglobosa*, coexistence of *P. fluorescens* and *C. saepenatans* did not affect the apparent predator grazing ability. Indeed, growth rates of naïve predator lines (fed with either bacterial SM, WS and FS morphotypes) did not significantly differ from growth rates of evolved predator lines fed with the same bacterial lines independently.

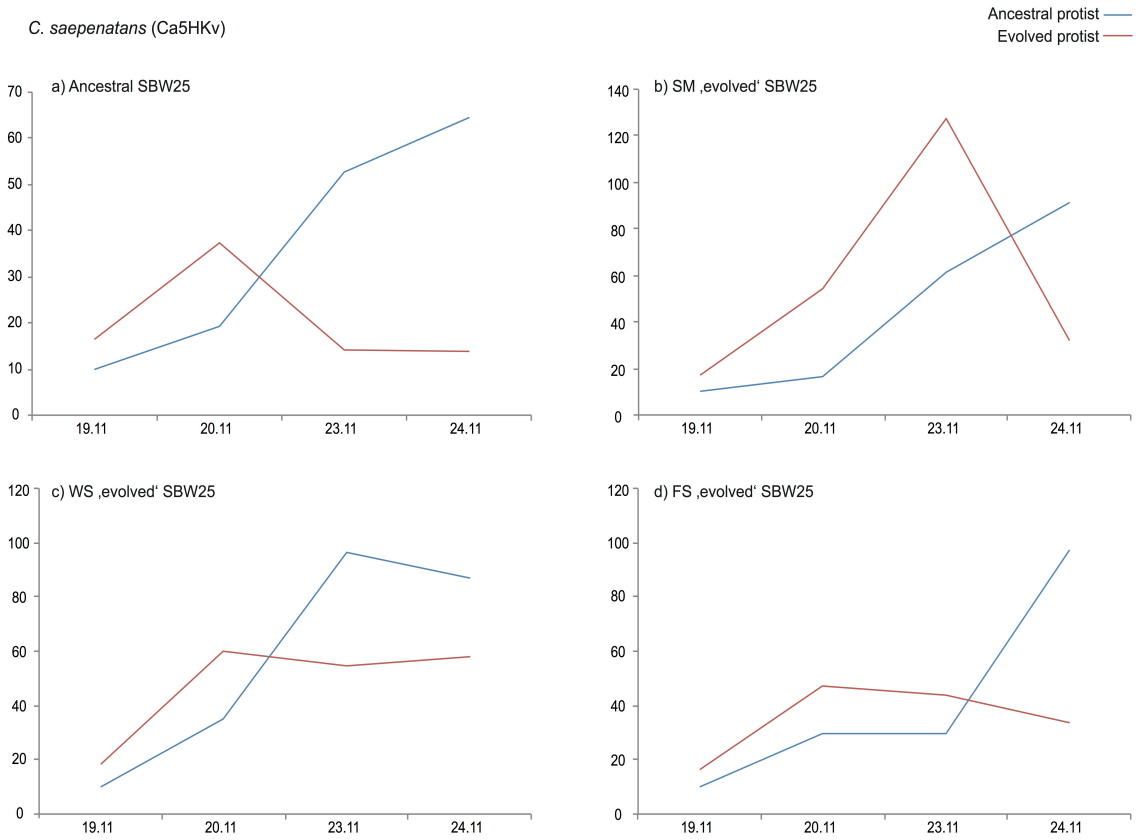


Figure 6.7 *C. saepenatans* (Ca5HKv) growth before co-existence with *P. fluorescens* SBW25 (blue lines) and after co-existence with SBW25 (red lines) in the presence of a) ancestral SBW25, b) SM phenotype of 'evolved' SBW25, c) WS phenotype of 'evolved' SBW25 and d) FS phenotype of 'evolved' SBW25.

Amongst all seven ancestral eukaryotes strains used, the ancestral ciliate *T. vorax* (TV) did not survive past 24h when fed with ancestral and evolved P. fluorescens, while all other strains did. It's evolved counterpart, however, experienced immediate and stable growth for all treatments alike (no difference in protist density at any time point or of area under the curve (AUC) when fed with ancestral, SM, WS or FS bacterial phenotypes). Why ancestral TV did not survive more than 24h when fed with both ancestral and evolved SBW25 is unclear, since co-existence with at least ancestral SBW25 was proven possible at the beginning of the experiment. it is possible, although unlikely, that TV drove its own extinction after exhausting all resources. In the presence of evolved bacteria, TV might have been unable to feed, thus declining until extinction.

Similarly, evolved *A. polyphaga* (AP) survived in all replicates but at very low numbers (one cyst observed in one out of three microcosm replicates) of biologically inactive cysts, and was thus not analysed. In parallel, the ancestral amoeba line was the only species that did not present any growth in all feeding treatments (no significant difference in protist density or AUC for all bacterial treatments). It is important to notice that ancestral AP started as very dense microcosms at the cross-feeding experiments (fig. 5.8), making it unlikely for the amoeba to get any denser either by excessive resources consumption, triggering death as opposed to growth, or by simple spatial distribution.

Of all 14 initial protists used, *Chilomonas paramecium* (CP) and the chrysophyte PML5D did not survive the whole length of the experiment (table 5.5). While the latter observed good survival rates in mixed-bacteria laboratory microcosms, its adaptation to SBW25 as single prey was also more difficult to achieve than with other strains. It could be that on the longer term, the bacteria developed more effective defence mechanisms against this predator. Or, in an opposite scenario, the predator observed better growth than its prey and consumes all the resources, driving its own demise. Indeed, PML5D survived for over a month in presence of constantly evolving SBW25. As for CP, population densities but mostly cell motility declined fairly fast: usually fast-swimming protists rapidly became slow then static, with only their cilia moving. Once the majority of cells became static, the bacteria could swarm around and degrade them. Survival patterns differed for all other protist strains. *Cercomonas pigra* (CaSphII), a large slow but metabolic protist (Bass, Howe, *et al.*, 2009), was unfortunately cross-contaminated by *Cercomonas effusa* (Beaver-Creek) or *Cercomonas paraglobosa* (19-3E) in three replicates out of six. Interestingly though, replicates of CaSphII that remained pure did not survive the whole length of the experiment, as the declining cells turned into cysts that ended up swarmed by bacteria; contaminated microcosms presented both original protist and contaminant thriving, the former with medium to high density and activity levels. A similar situation was observed with *Paracercomonas minima* (SW2): five replicates were cross-contaminated with either *Bodo saltans* or *Poterioochromonas*. SW2 did well when in presence of

B. saltans (medium density levels) but not with *Poterioochromonas*. The remaining uncontaminated replicate contained very small elements at the bottom, which were interpreted as cysts. *E. uvella* also grew and maintained 'medium' density levels when cross-contaminated by C. paraglobosa, and developed cysts (only) in pure strains. Cross-contamination of *C. effusa* and *Poterioochromonas* by other strains did not affect their growth, as both strains survived well in both contaminated and clean replicates. Similarly A. polyphaga rapidly declined in density then developed cysts in equally clean and contaminated replicate.

6.3.3 Bacterial evolution

Bacterial evolution was assessed as colony morphology observed on LB agar throughout the experiment (fig. 5.9). Amongst the variety of detected morphotypes, many were recorded as an overlap of more than one colony type, rendering difficult to effectively define the bacterial type. From those, three major bacterial morphologies were chosen based on their previously described characteristics: the smooth ancestral-like phenotype (SM), the wrinkly-spreader (WS) and the fuzzy-spreader (FS) (Rainey and Travisano, 1998). Types that appeared as an overlap of two or more divergent bacterial lines were categorised into one of the three chosen when possible (based on the classification used for description) or disregarded due to their rare presence and difficulty of classification.

P. fluorescens grown alone (control) globally presented the same stable smooth (SM) phenotype for almost all replicates during the experiment (fig 5.12); occasionally, replicates 1, 2 and 3 developed a WS phenotype on different individual times, but switched back into SM by the next sampling point. Similarly, replicate 5 diverged into a WS morphotype on last time point, although it appeared as SM on every previous sampling. Conversely, replicate 4 appeared as WS on the first time point only, indicating fast differentiation solely in the first four days of the experiment. Although P. fluorescens is known for differentiating into niche specialist morphotypes, no previous record of similar

variation was found. But one can must consider that differentiated morphotypes are fittest when rare (Rainey and Rainey, 2003; Ferguson *et al.*, 2013): over-population of the microcosm air-liquid interface might be one reason of "back-differentiation" into SM phenotypes. Bacterial differentiation was more variable according to the protist grazer. The large surface dwelling C. pigra (CaSphII) induced bacterial differentiation into WS phenotypes in all replicates, with re-emergence of SM, although less frequently than the former. On the other hand, *C. effusa* (Beaver-Creek) initially caused the bacteria to switch between SM and WS phenotypes, to finally select for WS only but reverting twice to SM in replicate 6. A similar profile of high alternation between both SM and WS phenotypes for all replicates was observed for bacteria grown under P. oxoniensis (WA8) or E. uvella (11-7E), the latter with selection for FS morphotypes on two occasions. All other protist treatments selected mostly for WS bacterial phenotypes, with divergences towards FS morphotypes for P. minima, Bodo saltans, Poterioochromonas sp. and C. saepenatans treatments, as well as a few SM types when grazed by T. pyriformis, T. vorax or A. polyphaga.

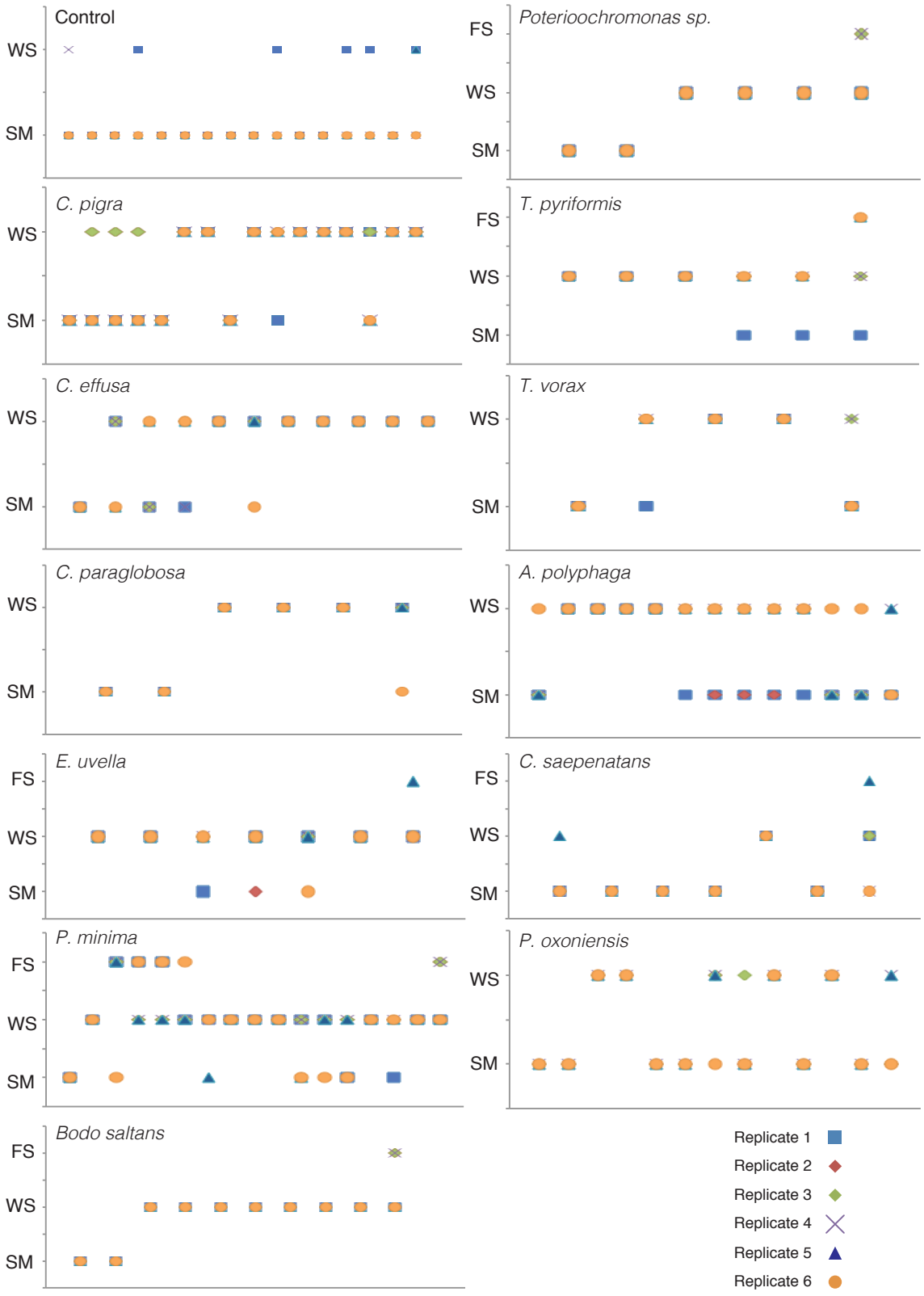


Figure 6.8 Bacterial phenotypic transition (smooth SM, wrinkly-spreader WS and fuzzy-spreader FS) over time (days) for all six replicates when grazed by a single protist predator.

At the end of the experiment, OD measurements at 600nm of evolved P. fluorescens lines SM, WS and FS growth at 0, 24 and 48 hours after inoculation were compared to those of ancestral non-evolved ones (fig 5.13). Absorbance was highest for the evolved SM, WS and FS bacterial types, with that for SM increasing throughout the growth period, but lower and decreasing for ancestral SBW25. Initial absorbance did not differ between ancestral, SM and WS lines; FS values were significantly higher at 0h ($F_{3,16}$=2.126, p-value=0.04). However, absorbance at 24h and 48h is significantly different for all four bacterial lines.

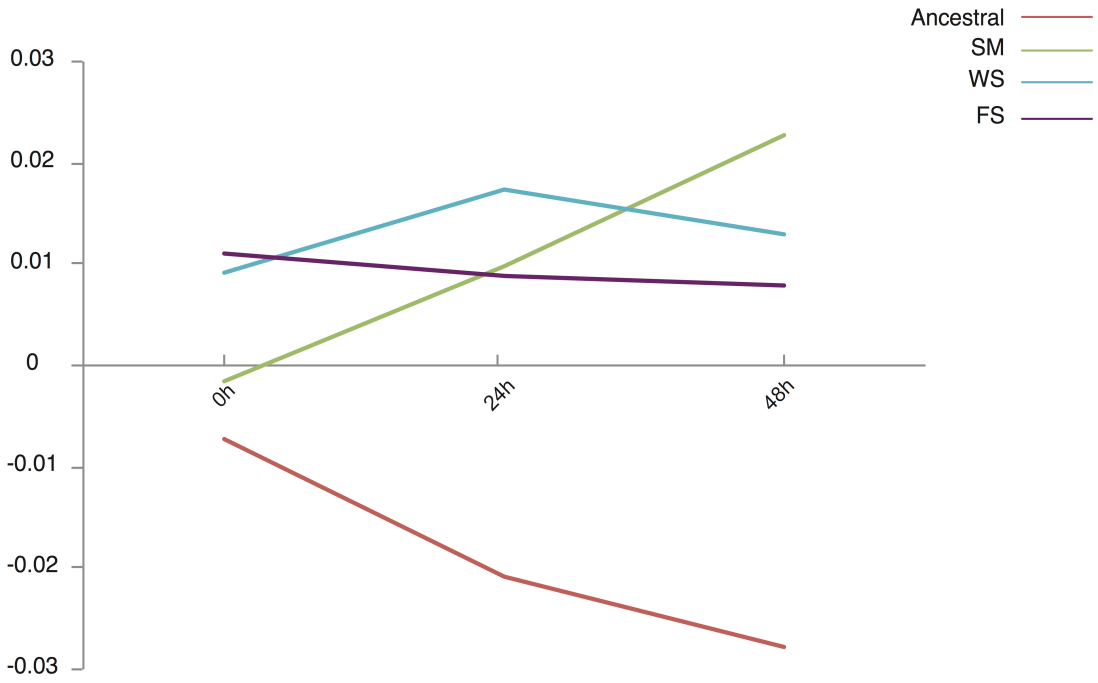


Figure 6.9 *P. fluorescens* ancestral and evolved (SM, WS and FS) phenotype community absorbance at 0h, 24h and 48h of incubation.

While absorbance can indicate bacterial growth, any bacterial exudate and element in suspension potentially rejected by active bacteria, able to absorb the light, will also contribute in the increase of observed OD.

6.4 Summary of conclusions

Fourteen morphologically and ecologically diverse protist strains were initially incubated with the same bacterial prey *P. fluorescens* SBW25, and maintained for 213 days.

In terms of predator-prey coexistence, seven protist strains were analysed. Except for *Bodo saltans* and *Poterioochromonas* sp., evolved protist lines achieved lower population densities when compared to their ancestral counterparts, when fed both ancestral and evolved SBW25. Only the evolved *Poterioochromonas* grew better on one evolved *P. fluorescens* line: the fuzzy-spreader (FS) previously developed in the presence of *E. uvella* (see methods). Conversely, ancestral *T. pyriformis* fed on predator-free evolved smooth (SM) bacteria grew better than evolved TP: the ciliate became more susceptible to evolved (SM) bacterial phenotype after coexistence. However, ancestral TP preferred FS and WS bacterial lines to SM and ancestral ones, while evolved TP did not differentiate. Ancestral *B. saltans* grew better (AUC) than the evolved BS on ancestral SBW25, although it fed better on FS and WS evolved bacteria. Evolved BS fed better on WS and SM evolved bacteria. *C. paraglobosa* fed and grew equally well on all bacterial lines indifferently. *C. saepenatans* grew equally well on all bacterial lines as well, but the ancestral CS preferred evolved bacterial type WS.

Comparison between evolved and ancestral *A. polyphaga* lines indicated that long-term exposure to a single bacterial line did not improve AP fitness. The evolved but not ancestral ciliate *T. vorax* survived exposure to ancestral and evolved bacterial lines, and its growth was not affected by SBW25's evolutionary state.

Only two protist species – the cryptophyte *Chilomonas paramecium* and the chrysophyte *Paraphysomonas* sp. PML5D – did not survive the whole length of the experiment. Among the remaining twelve strains that survived, seven of them became cross-contaminated with a 'plate-neighbour' species in at least

one microcosm (table 5.5). This unexpected situation allowed however the emergence of more complex interactions not initially predicted in the scope of the experiment.

Bacterial evolution appeared somewhat more difficult to assess. Co-existence with all protist species drove diversification into wrinkly-spreader phenotypes. Only *P. minima* led differentiation of *P. fluoresecens* into fuzzy-spreader morphotypes more consistently than other protist strains. Ancestral SBW25 lines observed apparent population decline over 48h, while the only the evolved SM line grew. Both evolved FS and WS bacterial lines showed slight decline in OD: niche-specific phenotypes present higher fitness costs, and do better when rare in diverse bacterial populations.

In summary, protists usually preferred evolved bacterial lines over ancestral ones, and when so, lines that evolved with another protist (FS and WS) rather than alone (SM). Only *Poterioochromonas* – and to a certain extent, *T. vorax* - fared better after coexistence with *P. fluorescens* SBW25, while *T. pyriformis*, *B. saltans* and *A. polyphaga* observed better fitness before contact with the bacteria. Only *C. paraglobosa* presented no observable differences before and after bacterial co-evolution.

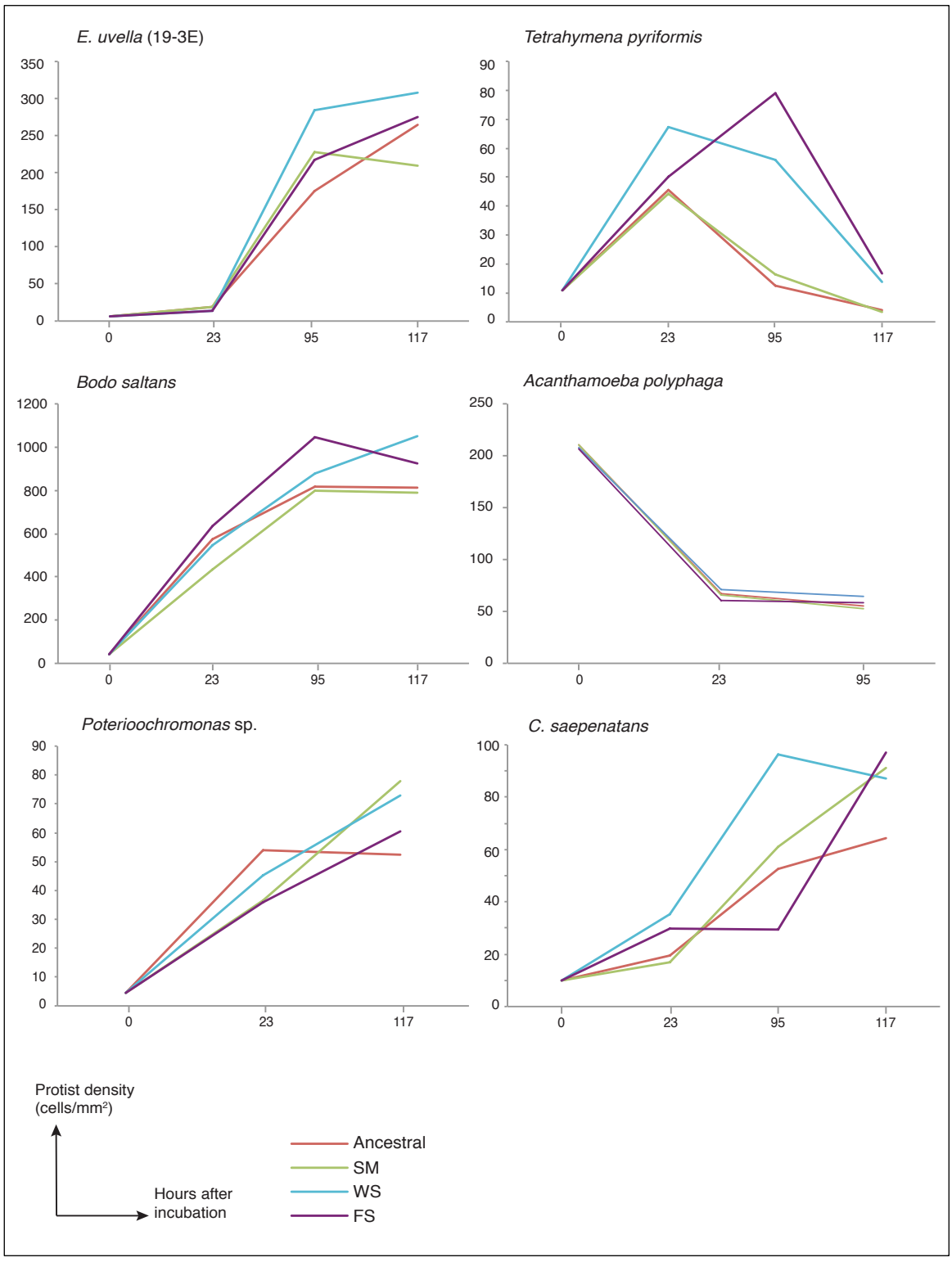


Figure S6.1 Ancestral protist growth (cells/mm2) over time (hours) after incubation with non-evolved (ancestral) and evolved (SM, WS and FS) bacterial lines.

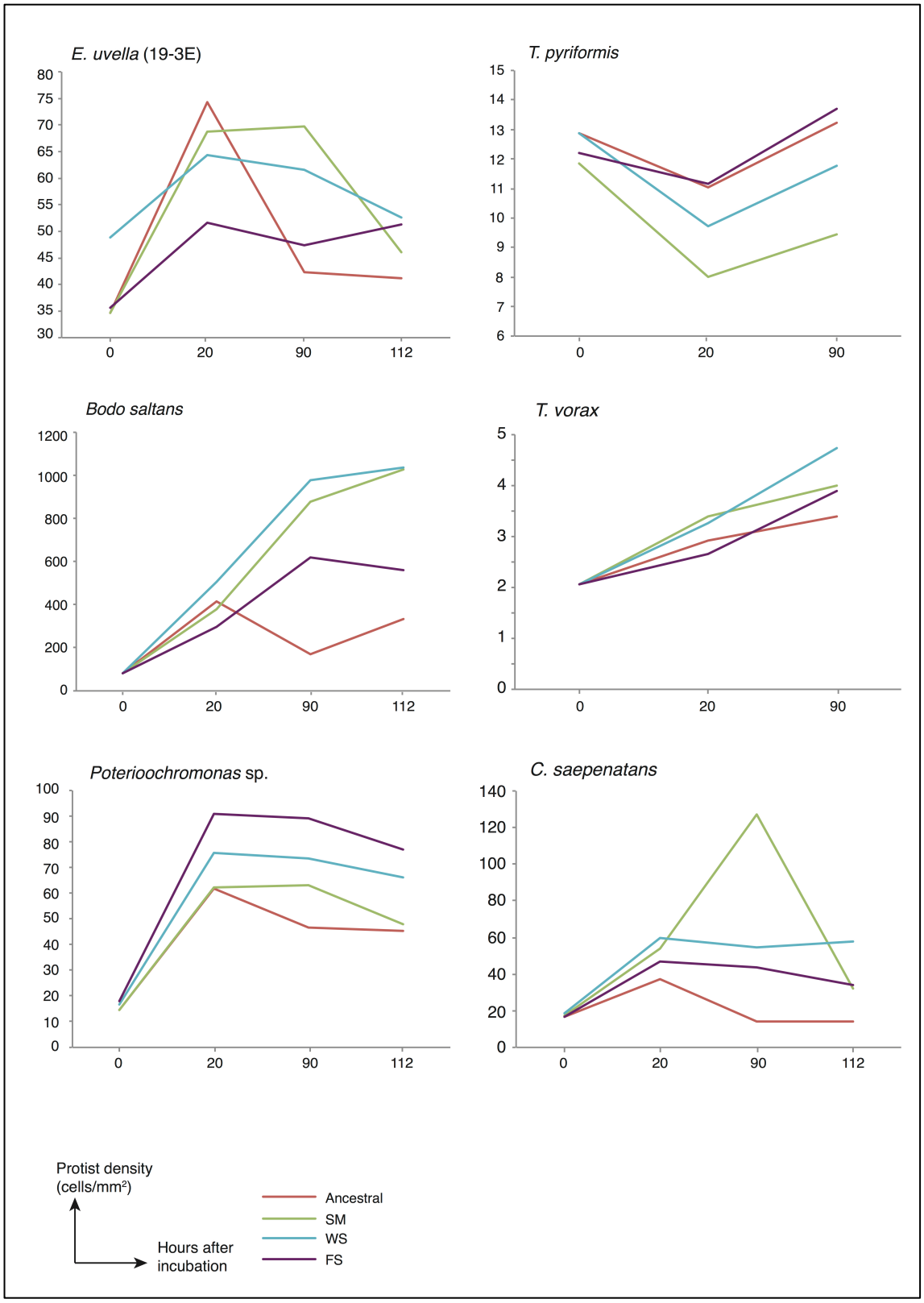


Figure S6 2 Evolved protist growth (cells/mm2) over time (hours) after incubation with non-evolved (ancestral) and evolved (SM, WS and FS) bacterial lines.

Chapter 7 SUMMARY AND DISCUSSION

The study, and most of all the (total) understanding of biological systems and what regulates them is an ambitious task. Natural environments are highly complex multileveled systems, made of an incredible number of players interacting at all times. Predator and prey relationships are defined by a multitude of factors including organism characteristics and environment specificities.

7.1 Where is everything? Who is everything?

Soil protist communities were shown to differ significantly between soils of different pH classes but to a lesser extent than bacterial communities analysed from the same samples (chapter 3). Low pH soils had markedly different micro-eukaryote assemblages from medium and high pH soils, whereas the latter categories were much more similar to each other. As for bacteria, protistan beta-diversity was also highest at low pH (Griffiths *et al.*, 2011). This might be a trivial expectation if protists were interacting solely with bacteria. However, only a small proportion of the protist taxa most characteristic of protist assemblage differences between the different pH levels were related to bacterivores, such as many cercozoan flagellates (Bass, Howe, *et al.*, 2009; Howe *et al.*, 2009; Howe, Bass, Scoble, *et al.*, 2011); the majority were related to parasites (of animals, plants, and other eukaryotic microbes), and protist and fungi otherwise known to interact with plant rhizospheres or phyllospheres (e.g. *Taphrina*, *Polymyxa*, *Archaeorhizomyces*; Table 3.2). Therefore, the ecological distribution of both above- and below-ground larger organisms appear to play strong roles in the determination of soil protist community structure, articulated by saprotrophy, coprophily, parasitism, and symbiosis (e.g. ectomycorrhizal fungi and rhizosphere-associated protists).

Correlation analyses showed strong variation in co-occurrence between protistan and bacterial OTUs indeed, the presence of protistan predators partially influenced composition of bacterial communities by altering proportions and absence/presence of a few major bacterial strains (chapter 4). However,

environmental factors such as riverine source as well bacterial community life history also seem to drive diversity changes. Negative or positive correlations between prey and predator might simply be explained by shared preference of members of each domain for certain environmental conditions. Although processes that structure communities are, to a certain extent, bound to major habitat types (Hairston, 1989), emergent community properties appear that can generalise over habitats (Begon *et al.*, 2007; Morin, 2011). Therefore, other interactions, for example preferential grazing of bacteria by protists (Chrzanowski and Simek, 1990; Glücksman *et al.*, 2010), antagonistic interactions such as chemical and morphological defence (Jürgens and Matz, 2002), pathogenicity, competition, etc., and synergistic interactions such as trophic cascades (Brussaard, 1977; Corno *et al.*, 2013) offer more biologically complex and powerful explanations for the related responses of both domains to pH level differences in their environment.

In terms of general micro-eukaryotic soil diversity our results are in agreement with previous sequencing-based studies, showing a high proportion of fungi, alveolates, and rhizarians. Recent studies (Urich *et al.*, 2008; Geisen, Tveit, *et al.*, 2015) showed a similar diversity profile by sequencing the soil metatranscriptome, (a good indicator of active cells as opposed to dormant or dead forms), and also that parasitic lineages are more abundant than many had assumed. For instance, strongly represented in Urich *et al.* (2008) data were the plasmodiophorid plant parasites, which are not conducive to culturing or cell isolation diversity studies and whose environmental diversity is much greater than host-oriented studies and those of economically important taxa would suggest (Neuhauser *et al.*, 2014). Alveolates were also well represented in all sequence based studies; Bates *et al.* (2013) noted that a significant proportion of their OTUs affiliated with Apicomplexa. Comparison of DNA and RNA-derived studies of soil apicomplexans will be important to distinguish between encysted and actively infecting forms (Rueckert *et al.*, 2011).

Even though short HTS-generated sequences have inherently low phylogenetic resolution, a combined approach to their taxonomic affiliation using both sequence similarity matching and phylogenetic analyses can provide more resolution and accuracy than blast-based methods alone. Further biological interpretation is possible via functional inference based on the resulting taxon profiles. We emphasise the need for phylogenetic moderation of raw taxon assignment outputs. It is important to acknowledge the significance of the percentage similarity between query and subject sequences. An 18S rDNA match of 95% or less (which dominate most HTS protistan diversity analyses) to a named database sequence is almost certainly not the species specified in the subject ID (if one is given) and may well not be the same genus. Below 85-90% assignments in the lower half of the taxonomic hierarchy become very doubtful. Here phylogenetic analyses can help, but are limited by both the signal carried by the OTU sequence fragment and database representation of related sequences. Databases themselves also powerfully influence perception of community structures. Their different outputs might misleadingly suggest strong biological differences between communities. The enduring lack of a generally adopted, comprehensive, and uniformly high quality taxonomic database for protists hinders the emergence of a body of data that can be consistently compared across studies.

Results suggest that interactions within microbial systems are far more complex than what one can see. Grazer's characteristics – such as shape plasticity, mobility, feeding mode and phylogenetic relation – affected the growth of a variety of bacterial strains, which affected bacterial community diversity. But the latter also appears to be linked to life history and origin. Indeed, prokaryotes are highly dependent on available resources, and adapted to thrive in the environment they were taken from. It is important to note, in the scope of experimental procedures held in this study, the limitations of working with non-axenic cultures. Protist strains were, for the great majority, obtained from cultures containing natural bacterial communities originating from the first isolation material. This implies that those bacteria most certainly possess a long-term adaptation to protist grazing, and could possibly overcome new

bacterial communities – as highlighted in chapter 4. But at the same time, prey inter- and intraspecific interactions have repercussions on the environment as a whole, and environmental perturbations as well as coexistence (in natural environments) with other competitive predatory species affect protist behaviour. For example, long-term coexistence of protist and their bacterial prey clearly induced changes in both players. In general, co-existence appeared to decrease protist growth abilities for most of the studied species, but this can be directly related to enhanced or specific defensive prey capture or digestion development as a trade-off to the number of ingested prey as well as to improve competitiveness between predators. Indeed, protist species showcase a variety of strategies when it comes to resource competition: some prefer to hide (encyst) and wait for better times, while others escape competition by playing on two trophic levels (intraguild predation) Furthermore, many of the species used in this study, particularly sarcomonads, harbour high genetic diversity and convergent morphologies (Bass, Howe, *et al.*, 2009), possible driver of a variety of ecological functions awaiting discovery.

7.2 Players in a bigger game

7.2.1 Overview

Every species of a community influences composition in a variety of ways, each one providing resources to its neighbour, be it as prey to a predator, host to a parasite or available matter to primary producers. But every player developed its own tricks in order to survive and thrive in such a dangerous world. While predator strategies rely on being the best at capturing its prey, the latter must adapt in order to escape predation.

Protists are known to select bacteria within a limited size range, easy to ingest and containing enough nutrient to sustain growth. This results in the shifting of cell size in prokaryotic communities towards much larger, difficult to ingest cells and very small ones, which represent lower nutrition value as well as being more difficult to find (lower encounter rates)(Simek *et al.*, 1999; Montagnes *et al.*, 2008). When submitted to protist grazing, bacterial community composition

was partially modulated by the predator species (chapter 3). For example, one of the most abundant bacterial species in all grazed communities, *Limnohabitans* sp. (OTU2998), not only has an important growth rate, it also ranges in cell size from very small 0.4*μ*m cocci for certain species, to 5*μ*m long curved bacilli (Kasalický *et al.*, 2013; Šimek *et al.*, 2013). Latter sizes are comparable to those of smaller protists such as *P. vonderheydeni*, *P. minima* or *A. scotia* that will most likely not be able to feed on that evolved prey. Protist survival relies then on its ability to either feed on more accessible prey – in either size or numbers - or to develop strategies to counter the bacterial adaptation. Furthermore, generalist or selective predation alters the balance between competing prey species, favouring r-strategy bacterial strains (high growth rate in response to importance grazing rates). In parallel, high grazing pressure triggers defence mechanisms from the prey, such as the formation of clumps or biofilms – a strategy adapted by *Pseudomonas* and *Acinetobacter* species, abundant in grazed communities - too large to be ingested whole while making the access to individual bacterial cells difficult (Hahn and Höfle, 2001; Blom *et al.*, 2010); or the development of grazing-resistant morphotypes such as flocks and filaments.

In addition to community size structure, protist taxonomic relatedness, predator nutritional state, prey motility, cell surface physicochemical properties and even toxicity have been shown to impact microbial predator-prey interactions (Montagnes *et al.*, 2008; Bell *et al.*, 2010; Glücksman *et al.*, 2010; Meunier *et al.*, 2012).

7.2.2 Bacterial response to protist predation

P. fluorescens SBW25, a gram-negative bacterium common in aquatic and terrestrial environments, is a model in the study of adaptive radiation: when in structured environments, it quickly diversifies into niche specialist phenotypes (Rainey and Travisano, 1998). Interestingly, similar morphotypes develop when *P. fluorescens* is submitted to protist predation (Friman, Dupont, *et al.*, 2015), and genotypic changes appeared in line with observed phenotypic adaptation

(chapter 5.1). Although adaptation to environmental (experimental) conditions triggered the most changes in P. fluorescens gene expression, predator presence further shaped regulation. Even though we selected bacterial lines that demonstrated phenotypic changes in the previous selection experiment, we can't confirm the same phenotypes were effectively sequenced. However, clear differences between bacterial lines grazed by protist *C. paramecium* (CP) and *A. polyphaga* (AP) – which drove unique specialist bacterial adaptation – and bacterial lines grazed by the generalist *T. pyriformis* – that drove generalist defensive differentiation – were observed. In response to the fast swimming ciliate (TP), SBW25 developed the ecologically successful 'wrinkly-spreader' phenotype (WS), which involves mutations of a small number of loci over a limited number of mutational pathways (McDonald *et al.*, 2009) that enhance secretion of cellulose in order to form biofilms able to attach to the air-surface interface. Indeed, grazing by TP mostly activated regulation of genes involved in pathway regulation, excretion and lipopolysaccharide biosynthesis – elements of biolfim. In parallel, morphotypes observed in the presence of AP and CP were not observed in the presence of TP, and genes involved in those mutations were regulated in opposite directions than the same genes driving adaptation to TP grazing. Although no information could be found about specific transparent (TT) and petite (PT) *P. fluorescens* colony types observed, the bacterium is know to diversify into more than one phenotype. Both TT and PT morphotypes were observed in much smaller proportions than the WS one, as well as much lower relative fitness when compared to ancestral smooth types (SM). It is likely that gene expression regulation in AP and CP, by developing highly defensive morphotypes, also involves deleterious mutations, conferring better predator resistance but lower fitness in rarely encountered specialists (MacLean *et al.*, 2004) – as observed for the loss of function of gene fuzY, responsible of FS phenotypes (Ferguson *et al.*, 2013).

In simple one protist – one bacterium (*P. fluorescens*) systems, bacterial response over time revealed somehow expected changes, although their exact nature was not clear due mainly to experimental conditions (chapter 5.2). The bacterial SM lines, effective against evolved *B. saltans* (BS), was not so when

exposed to the ancestral BS line, and vice-versa. In the presence of the ciliate *T. vorax* (TV), very stable bacterial differentiation was also observed (figure 5.12). However, when previously exposed to TV in the selection experiment, SBW25 defence development was repressed (Friman, Dupont, *et al.*, 2015), which indicates that over the long term predation pressure over SBW25 is lesser: *T. vorax* is a large polymorphic active pelagic protist that can develop different morphotypes in response to environmental conditions. By increasing both body and oral apparatus sizes, it can engulf large preys such as bacterial clumps, biofilms and other smaller protists (Gronlien *et al.*, 2011). Both microstome (small buccal opening) and macrostome (large buccal opening) were observed in evolved *T. vorax* populations, but no macrostome was observed in the ancestral populations used for inoculation, which could imply preferential intraguild predation in evolved populations over to bacterial grazing. In addition, the evolved but not ancestral ciliate *T. vorax* survived exposure to ancestral and evolved bacterial lines. Although why ancestral TV did not survive when fed with ancestral P. fluorescens, it can be however that evolved bacteria were too defensive against ancestral TV. Similarly, Poterioochromonas drove stable bacterial diversification over long-term predator-prey co-existence, while reaching the highest cell concentrations observed amongst all protist population and a better fitness after co-evolution. As it happens, Poterioochromonas are able to ingest smaller protists, including smaller (Poterio)Ochromonas species (Landry *et al.*, 1991; Ishigaki and Sleigh, 2001). The fuzzy-spreader bacterial morphotype, although not specifically defensive against ancestral *Poterioochromonas*, appeared more sensitive to the evolved microeukaryote, which is in line with the rare development of this phenotype in *Poterioochromonas* microcosms.

Predator-prey coexistence adaptation was dependent on the organism considered (chapter 5.2). All protists, whether ‘evolved’ – i.e. after being in contact with *P. fluorescens* over many generations of both domains – or ‘ancestral – never in contact with that bacterium before – showed a preference for prokaryotes that had previously been exposed to protist predation over bacteria grown alone. Except for *Bodo saltans* and *Poterioochromonas*, evolved

protist lines achieved lower population densities when compared to their ancestral counterparts. Only *Poterioochromonas* effectively observed better fitness after exposure to *P. fluorescens.* However, co-evolution of sarcomonads *E. uvella* or *C. saepenatans* (CS) with *P. fluorescens* individually did not affect the predator's performance. Both protists are plastic 'slow' species, and although mostly bound to the substrate, they can access bacteria in suspension (SM morphotypes) as well as break through loose bacterial flocculate (FS types). Furthermore, CS can differentiate into a swimming morphology that can eventually predate on liquid-air interface colonisers such as WS bacterial morphotypes.

Tetrahymena pyriformis (TP) became more susceptible to the evolved smooth (SM) bacterial phenotype after coexistence, but the ancestral line grew better on evolved bacterial lines. Although the previous selection experiment indicated TP as having a strong impact over *P. fluorescens*, over the long term it could well be that bacteria and protist evolved in an arms race scenario, and when defensive SBW25 were submitted to non-evolved grazer, it was not capable of defending itself anymore.

Non-evolved *Poterioochromonas* sp. population density levels were independent of *P. fluorescens* phenotypes, but its evolved counterpart fared better in the presence of FS and WS types. Indeed, growth rates of the evolved paraphysomonad were significantly higher than those of the ancestral ones when fed with FS bacteria.

Only two protist species – the cryptophyte *Chilomonas paramecium* and the chrysophyte *Paraphysomonas* sp. PML5D – did not survive the whole length of the experiment. It is interesting to note that *C. paramecium*, a small pelagic and highly active species, while it survives well in microcosms in presence of *P. fluorescens*, its growth dynamics seem affected by the presence of other protists (Mucibabic, 1957; Friman, Dupont, *et al.*, 2015; Friman, Guzman, *et al.*, 2015). Furthermore, CP harbours an important growth rate, rapidly achieving high population density before declining (table 5.5). It is possible that CP alone could not maintain its population levels after depleting the environment of all resources because of its fast growth and high activity levels. But we must also

consider that new resources were added regularly to the (closed) system, ensuring protistan survival, rendering resources burn out unlikely. In parallel, in a previous selection experiment, the bacteria did not develop any observable morphological changes when predated by CP alone, but only when in presence of both TP and CP predators (Friman, Dupont, *et al.*, 2015). However, no specific bacterial phenotypes (transparent TT or petite PT colonies) were observed over the course of the long-term incubation in any of the CP microcosms, contrarily to what was observed in the selection experiment. It is plausible that, in the absence of a complementary predator species, the cryptophyte was unable to feed on the defensive bacteria. Similar scenarios can be considered for PML5D, although further studies are necessary to see whether the protist fares better in the presence of one or more additional microeukaryotes. Furthermore, the genus Paraphysomonas can feed not only on bacteria, but also smaller eukaryotes such as microalgae and eventually smaller protists by specifically selecting its prey, partially based on its nutritional quality (Landry *et al.*, 1991; Ishigaki and Sleigh, 2001). It is possible that, in the presence of a unique, maybe not so well nutritionally suited and defensive prey, PML5D could not adapt.

Interestingly, *Cercomonas pigra*, a large metabolic species only survived to the end of the experiment when in presence of another protist. Lines that remained pure were extinguished as active cells turned into cysts that were swarmed by bacteria. *E. uvella*, however, develop cysts in microcosms that remained un-contaminated, while cells stayed active when in the presence of *C. paraglobosa*. *Acanthamoeba polyphaga* (AP) survival was hindered after long-term exposure to a single bacterial line; AP developed cysts in both pure and contaminated microcosms, and cysts were unable to thrive when in presence of ancestral and evolved (SM, FS and WS) bacterial lines. Encystment in free-living protozoa is still under-studied in comparison to parasitic ones, but it is globally accepted that cyst biogenesis allows microeukaryotes to survive unfavourable conditions (Corliss and Esser, 1974; Bamforth, 1988). Excystment might happen when better conditions are met, either after local environmental change or by dispersal of the cyst. In close microcosm situations, however,

neither could be achieved, and in most extreme cases the protist did not survive (*C. pigra*). For AP, cysts transferred to a new environment in presence of either adapted or ancestral bacterial prey did not find conditions to excyst and multiply.

P. minima (SW2), a small metabolic sarcomonad, was depleted by the presence of *Poterioochromonas* sp. but not in the presence of *Bodo saltans*. While *B. saltans* is bacterivorous only, *Poterioochromonas* species are toxic and able to ingest smaller protists (Boxhorn *et al.*, 1998; Ishigaki and Sleigh, 2001; Moser and Weisse, 2011), which could explain *P. minima* mortality. On the other hand, *Bodo saltans* and SW2 occupy different ecological niches; the latter dwelling on the bottom while the former can swim and occupy the whole water column. Both species are thus able to coexist by grazing on different bacterial phenotypes: WS bacteria usually develop at the air-liquid interface, are out of reach for SW2 but accessible to *B. saltans*, SM types appear in suspension thus more susceptible to the bodonid. FS bacterial phenotypes tend to sink to the bottom of the culture vessel, and appear only initially in the presence of SW2. In parallel, survival of *P. minima* alone revealed difficult to determine. *C. effusa* and *Poterioochromonas* were not affected by cross-contamination, and even out-competed the intruder. Both protist species can form very dense populations in microcosm, and probably outgrew the contaminant. Furthermore, *Poterioochromonas* could have fed on smaller protists as well as harmed the competitor. In addition to protist-protist interactions, *P.fluorescens* apparently diverged regularly between morphotypes (fig 5.12) in a similar way for all replicates in microcosms that were cross-contaminated, indicative of rapid evolution in response of predation but possibly also of environmental conditions changes, which equally affect predators.

Natural communities are defined by the presence of a multitude of species, and interaction is the only option when individuals sharing the same area, resource or even both. The impact of the players upon each other will mostly depend on

the nature of their interaction, whether detrimental (predation, competition) or beneficial (mutualism). When resource is a limiting factor, competition will favour the species with the best foraging capacity, while the least able one will be deprived, thus growing slower, reproduce at lower levels and even be at risk of extinction. Under non-limiting conditions, the fastest growing species can multiply faster, thus observing an important increase in abundance that can lead to resource limitation and fall into a resource-limiting scenario. At this point, over-exploitation of resource equally leads to extinction. Life history and growth strategies coexisting species dictate survival rates.

In a laboratory made of eight ecologically and morphologically distinct protist community (chapter 6), medium concentration mimic resource availability – higher medium concentration allowed the maintenance of higher bacterial growth rates – while microcosm size allowed species dispersion and possible environment patchiness. Smaller microcosms drove important populations fluctuation, representative of dynamics systems: in smaller spatial scales, local interspecific trade-offs such as differential use of resource and fitness in variable environments (Chesson and Huntly, 1989). Indeed, in this community, all predators had varied feeding strategies: substrate-dwelling sarcomonads engulf their prey, and *C. saepenatans* differentiated into swimming types; A. polyphaga developed cysts as a way to avoid predation (Bamforth, 1988). Large microcosms presented flattened dynamics: when resource is spread, smaller and less motile species – such as *P. minima* – are at higher risk of extinction since their relative colonisation ability is lower than highly motile species such as *T. pyriformis* or *B. saltans*. However, *P. minima* appeared more resistant to stress in view of the fast recovery after enrichment (fig 6.2). Furthermore, in larger spatial scales, good competitors – organisms with higher growth rates are more affected than their slower-growing counterpart (Kneitel and Chase, 2004). On a longer term, protists strains such as *A. polyphaga*, encysted until less competitive times show up, or *E. uvella*, that is not particularly affected by prey diversification and defence evolution, despite apparent low productivity, would most likely thrive for longer than fast growing species such as T. pyriformis.

7.3 Future work

In this study, I extended the simple one predator – one prey system and analysed the effect of a single protist on natural bacterial communities in terms of taxonomic community composition. In addition, the interactions of eight competitive ecologically and physiologically diverse protists in varied environmental conditions were studied. Although the effect of protists on bacteria (Pernthaler *et al.*, 2001; Šimek *et al.*, 2013; Salcher, 2014), few have concentrated on the effects of more than one or two protists on their prey. The challenge now is to investigate the interactions of a variety of eukaryotic predators on (natural) bacterial communities.

Another interesting aspect in predator – prey interactions resides in both players' adaptation and evolution. Rapid evolution in microbial systems has been studied more and more, but mostly analyse the prey adaptation (Cortez and Ellner, 2010; Ellner and Becks, 2011; Friman *et al.*, 2014; Scanlan *et al.*, 2015). Complementary future experiments would focus on the adaptation – over long and short terms – of protist predator.

The analyses and experiments done throughout this thesis have extended the knowledge of what drives protist and bacteria interactions in natural and experimental conditions, but much remains to be done in the realms of the biggest microscopic predators.

References

Abdel-el-haleem, D. (2003) Minireview Acinetobacter : environmental and biotechnological applications. *African J. Biotechnol.* **2**: 71–74.

Abrams, P.A. (2005) “Adaptive Dynamics” vs. “adaptive dynamics.” *J. Evol. Biol.* **18**: 1162–1165.

Abrams, P.A. (2000) The evolution of predator-prey interactions: theory and evidence. *Annu. Rev. Ecol. Syst.* **31**: 79–105.

Abrams, P.A. and Matsuda, H. (1997) Prey adaption as a cause of predator-prey cycles. *Evolution (N. Y).* **51**: 1742–1750.

Adl, S.M. and Coleman, D.C. (2005) Dynamics of soil protozoa using a direct count method. *Biol. Fertil. Soils* **42**: 168–171.

Adl, S.M. and Gupta, V.V.S.R. (2006) Protists in soil ecology and forest nutrient cycling. *Can. J. For. Res.* **36**: 1805–1817.

Adl, S.M., Habura, A., and Eglit, Y. (2013) Amplification primers of SSU rDNA for soil protists. *Soil Biol. Biochem.* **69**: 328–342.

Adl, S.M., Simpson, A.G.B., Farmer, M.A., Andersen, R.A., Anderson, O.R., Barta, J.R., et al. (2005) The new higher level classification of eukaryotes with emphasis on the taxonomy of protists. *J. Eukaryot. Microbiol.* **52**: 399–451.

Adl, S.M., Simpson, A.G.B., Lane, C.E., Luke??, J., Bass, D., Bowser, S.S., et al. (2012) The revised classification of eukaryotes. *J. Eukaryot. Microbiol.* **59**: 429–493.

Altschul, S.F., Gish, W., Miller, W., Myers, E.W., and Lipman, D.J. (1990) Basic local alignment search tool. *J. Mol. Biol.* **215**: 403–10.

Alvarez, M., Schrey, A.W., and Richards, C.L. (2015) Ten years of transcriptomics in wild populations: What have we learned about their ecology and evolution? *Mol. Ecol.* **24**: 710–725.

Anders, S., Pyl, P.T., and Huber, W. (2014) HTSeq – A Python framework to work with high-throughput sequencing data. *Bioinformatics.* **31**: 0–5.

Anderson, M.J. and Walsh, D.C.I. (2013) PERMANOVA , ANOSIM , and the Mantel test in the face of heterogeneous dispersions : What null hypothesis are you

testing ? Published by : Ecological Society of America Linked references are available on JSTOR for this article : Your use of the JSTOR archi. *Ecol. Monogr.* **83**: 557–574.

Anderson, O.R. and Cavalier-Smith, T. (2012) Ultrastructure of Diplophrys parva, a new small freshwater species, and a revised analysis of Labyrinthulea (Heterokonta). *Acta Protozool.* **51**: 291–304.

Azam, F., Fenchel, T., Field, J., Gray, J., Meyer-Reil, L., and Thingstad, F. (1983) The Ecological Role of Water-Column Microbes in the Sea. *Mar. Ecol. Prog. Ser.* **10**: 257–263.

Bailey, M.J., Lilley, A.K., Thompson, I.P., Rainey, P.B., and Ellis, R.J. (1995) Site directed chromosomal marking of a fluorescent pseudomonad isolated from the phytosphere of sugar beet; stability and potential for marker gene transfer. *Mol. Ecol.* **4**: 755–63.

Bamforth, S.S. (1988) Interactions between protozoa and other organisms. *Agric. Ecosyst. Environ.* **24**: 229–234.

Bamforth, S.S. (2007) Protozoa from aboveground and ground soils of a tropical rain forest in Puerto Rico. *Pedobiologia (Jena).* **50**: 515–525.

Barbe, V., Vallenet, D., Fonknechten, N., Kreimeyer, A., Oztas, S., Labarre, L., et al. (2004) Unique features revealed by the genome sequence of Acinetobacter sp. ADP1, a versatile and naturally transformation competent bacterium. *Nucleic Acids Res.* **32**: 5766–5779.

Bass, D. and Cavalier-Smith, T. (2004) Phylum-specific environmental DNA analysis reveals remarkably high global biodiversity of Cercozoa (Protozoa). *Int. J. Syst. Evol. Microbiol.* **54**: 2393–2404.

Bass, D., Chao, E.E.Y., Nikolaev, S., Yabuki, A., Ishida, K. ichiro, Berney, C., et al. (2009) Phylogeny of Novel Naked Filose and Reticulose Cercozoa: Granofilosea cl. n. and Proteomyxidea Revised. *Protist* **160**: 75–109.

Bass, D., Howe, A.T., Mylnikov, A.P., Vickerman, K., Chao, E.E., Edwards Smallbone, J., et al. (2009) Phylogeny and Classification of Cercomonadida (Protozoa, Cercozoa): Cercomonas, Eocercomonas, Paracercomonas, and Cavernomonas gen. nov. *Protist* **160**: 483–521.

Bass, D. and Richards, T.A. (2011) Three reasons to re-evaluate fungal diversity “on Earth and in the ocean.” *Fungal Biol. Rev.* **25**: 159–164.

Bates, S.T., Clemente, J.C., Flores, G.E., Walters, W.A., Parfrey, L.W., Knight, R., and Fierer, N. (2013) Global biogeography of highly diverse protistan communities in soil. *ISME J.* **7**: 652–659.

Baumann, P. (1968) Isolation of Acinetobacter from soil and water. *J. Bacteriol.* **96**: 39–42.

Becks, L., Hilker, F.M., Malchow, H., Jürgens, K., and Arndt, H. (2005) Experimental demonstration of chaos in a microbial food web. *Nature* **435**: 1226–1229.

Begon, M., Townsed, C.R., and Harper, J.L. (2007) Ecology – From Individuals to Ecosystems, 4th ed. Blackwell Publishing.

Bell, T., Bonsall, M.B., Buckling, A., Whiteley, A.S., Goodall, T., and Griffiths, R.I. (2010) Protists have divergent effects on bacterial diversity along a productivity gradient. *Biol. Lett.* **6**: 639–642.

Benjamini, Y. and Hochberg, Y. (1995) Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J. R. Stat. Soc.* **57**: 289–300.

Berney, C., Geisen, S., Van Wichelen, J., Nitsche, F., Vanormelingen, P., Bonkowski, M., and Bass, D. (2015) Expansion of the “Reticulosphere”: Diversity of Novel Branching and Network-forming Amoebae Helps to Define Variosea (Amoebozoa). *Protist* **166**: 271–295.

Berney, C., Romac, S., Mahé, F., Santini, S., Siano, R., and Bass, D. (2013) Vampires in the oceans: predatory cercozoan amoebae in marine habitats. *ISME J.* **7**: 2387–2399.

Blom, J.F., Horňák, K., Šimek, K., and Pernthaler, J. (2010) Aggregate formation in a freshwater bacterial strain induced by growth state and conspecific chemical cues. *Environ. Microbiol.* **12**: 2486–2495.

Bonkowski, M. (2004) Protozoa and plant growth: The microbial loop in soil revisited. *New Phytol.* **162**: 617–631.

Bowes, M.J., Ings, N.L., McCall, S.J., Warwick, A., Barrett, C., Wickham, H.D., et al. (2012) Nutrient and light limitation of periphyton in the River Thames: Implications for catchment management. *Sci. Total Environ.* **434**: 201–212.

Boxhorn, J.E., Holen, D.A., and Boraas, M.E. (1998) Toxicity of the Chrysophyte flagellate Poterioochromonas malhamensis to the rotifer Brachionus angularis. *Hydrobiologia* **387**: 283–287.

Bråte, J., Logares, R., Berney, C., Ree, D.K., Klaveness, D., Jakobsen, K.S., and Shalchian-Tabrizi, K. (2010) Freshwater Perkinsea and marine-freshwater colonizations revealed by pyrosequencing and phylogeny of environmental rDNA. *ISME J.* **4**: 1144–1153.

Brussaard, L. (1977) Biodiversity and Ecosystem Functioning in Soil. *Ambio* **26**: 563–570.

Burgin, A.J., Yang, W.H., Hamilton, S.K., and Silver, W.L. (2011) Beyond carbon and nitrogen: How the microbial energy economy couples elemental cycles in diverse ecosystems. *Front. Ecol. Environ.* **9**: 44–52.

Calbet, A. and Landry, M.R. (2004) Phytoplankton growth, microzooplankton grazing, and carbon cycling in marine systems. *Limnol. Ocean.* **49**: 51–57.

del Campo, J., Not, F., Forn, I., Sieracki, M.E., and Massana, R. (2013) Taming the smallest predators of the oceans. *ISME J.* **7**: 351–358.

Caporaso, J.G., Kuczynski, J., Stombaugh, J., Bittinger, K., Bushman, F.D., Costello, E.K., et al. (2010) QIIME allows analysis of high-throughput community sequencing data. *Nat. Methods* **7**: 335–6.

Cavalier-Smith, T. (2014) Gregarine site-heterogeneous 18S rDNA trees, revision of gregarine higher classification, and the evolutionary diversification of Sporozoa. *Eur. J. Protistol.* **50**: 472–495.

Cavalier-Smith, T. and Chao, E.E.Y. (2003) Phylogeny and classification of phylum Cercozoa (Protozoa). *Protist* **154**: 341–358.

Chesson, P. and Huntly, N. (1989) Short-term instabilities and long-term community dynamics. *Trends Ecol. Evol.* **4**: 293–298.

Chow, C.-E.T., Kim, D.Y., Sachdeva, R., Caron, D.A., and Fuhrman, J.A. (2014) Top-down controls on bacterial community structure: microbial network analysis of bacteria, T4-like viruses and protists. *ISME J.* **8**: 816–29.

Chrzanowski, T.H. and Simek, K. (1990) Prey-size selection by freshwater flagellated protozoa. *Limnol. Oceanogr.* **35**: 1429–1436.

Clarholm, M. (1985) Interactions of bacteria, protozoa and plants leading to mineralization of soil nitrogen. *Soil Biol. Biochem.* **17**: 181–187.

Clarke, K.R. (1993) Non-parametric multivariate analyses of changes in community structure. *Aust. J. Ecol.* **18**: 117–143.

Corliss, J.O. (1998) Haeckel's Kingdom Protista and Current Concepts in Systematic Protistology. *Stapfia* **131**: 85–104.

Corliss, J.O. and Esser, S.C. (1974) Comments on the Role of the Cyst in the Life Cycle and Survival of Free-Living Protozoa. *Trans. Am. Microsc. Soc.* **93**: 578–593.

Corno, G., Caravati, E., Callieri, C., and Bertoni, R. (2008) Effects of predation pressure on bacterial abundance, diversity, and size-structure distribution in an oligotrophic system. *J. Limnol.* **67**: 107–119.

Corno, G. and Jürgens, K. (2008) Structural and functional patterns of bacterial communities in response to protist predation along an experimental productivity gradient. *Environ. Microbiol.* **10**: 2857–2871.

Corno, G., Villiger, J., and Pernthaler, J. (2013) Coaggregation in a microbial predator-prey system affects competition and trophic transfer efficiency. *Ecology* **94**: 870–881.

Cortez, M.H. and Ellner, S.P. (2010) Understanding rapid evolution in predator-prey interactions using the theory of fast-slow dynamical systems. *Am. Nat.* **176**: E109-27.

DeSantis, T.Z., Hugenholtz, P., Larsen, N., Rojas, M., Brodie, E.L., Keller, K., et al. (2006) Greengenes, a chimera-checked 16S rRNA gene database and workbench compatible with ARB. *Appl. Environ. Microbiol.* **72**: 5069–5072.

Doebeli, M. and Dieckmann, U. (2000) Evolutionary Branching and Sympatric Speciation Caused by Different Types of Ecological Interactions. *Am. Nat.* **156**: S77–S101.

Domonell, A., Brabender, M., Nitsche, F., Bonkowski, M., and Arndt, H. (2013) Community structure of cultivable protists in different grassland and forest soils of Thuringia. *Pedobiologia (Jena).* **56**: 1–7.

Dowd, S.E., Callaway, T.R., Wolcott, R.D., Sun, Y., McKeehan, T., Hagevoort, R.G., and Edrington, T.S. (2008) Evaluation of the bacterial diversity in the feces of cattle using 16S rDNA bacterial tag-encoded FLX amplicon pyrosequencing (bTEFAP). *BMC Microbiol.* **8**: 125.

Edgar, R.C. (2010) Search and clustering orders of magnitude faster than BLAST. *Bioinformatics* **26**: 2460–2461.

Ellner, S.P. and Becks, L. (2011) Rapid prey evolution and the dynamics of two-

predator food webs. *Theor. Ecol.* **4**: 133–152.

Ezard, T.H.G., Côté, S.D., and Pelletier, F. (2009) Eco-evolutionary dynamics: disentangling phenotypic, environmental and population fluctuations. *Philos. Trans. R. Soc. B Biol. Sci.* **364**: 1491–1498.

Fenchel, T. (1987) Ecology of protozoa: the biology of free-living phagotrophic protists. *Ecol. protozoa Biol. Free. phagotrophic protists.*

Fenchel, T. and Finlay, B.J. (2004) The Ubiquity of Small Species: Patterns of Local and Global Diversity. *Bioscience* **54**: 777.

Ferguson, G.C., Bertels, F., and Rainey, P.B. (2013) Adaptive divergence in experimental populations of Pseudomonas fluorescens. V. insight into the niche specialist fuzzy spreader compels revision of the model Pseudomonas radiation. *Genetics* **195**: 1319–1335.

Fierer, N., Ladau, J., Clemente, J.C., Leff, J.W., Owens, S.M., Pollard, K.S., et al. (2013) Reconstructing the microbial diversity and function of pre-agricultural tallgrass prairie soils in the United States. *Science* **342**: 621–4.

Finlay, B.J. (2002) Global dispersal of free-living microbial eukaryote species. *Science (80-.).* **296**: 1061–1063.

Finlay, B.J. (2004) Protist Diversity is Different? *Protist* **155**: 15–22.

Foissner, W. (2006) Biogeography and dispersal of micro-organisms: A review emphasizing protists. *Acta Protozool.* **45**: 111–136.

Foissner, W. (1999) Protist Diversity: Estimates of the Near-Imponderable. *Protist* **150**: 363–368.

Friman, V.-P. and Buckling, A. (2013) Effects of predation on real-time host-parasite coevolutionary dynamics. *Ecol. Lett.* **16**: 39–46.

Friman, V.-P. and Buckling, A. (2014) Phages can constrain protist predation-driven attenuation of Pseudomonas aeruginosa virulence in multienemy communities. *ISME J.* **8**: 1820–1830.

Friman, V.-P., Dupont, A.Ö.C., Bass, D., Murrell, D.J., and Bell, T. (2015) Relative importance of evolutionary dynamics depends on the composition of microbial predator–prey community. *ISME J.* 1–11.

Friman, V.-P., Guzman, L.M., Reuman, D.C., and Bell, T. (2015) Bacterial adaptation to

sublethal antibiotic gradients can change the ecological properties of multitrophic microbial communities. *Proc. Biol. Sci.* **282**: 20142920.

Friman, V.-P., Hiltunen, T., Laakso, J., and Kaitala, V. (2008) Availability of prey resources drives evolution of predator-prey interaction. *Proc. Biol. Sci.* **275**: 1625–33.

Friman, V.-P., Jousset, A., and Buckling, A. (2014) Rapid prey evolution can alter the structure of predator-prey communities. *J. Evol. Biol.* **27**: 374–380.

Geisen, S., Laros, I., Vizcaíno, A., Bonkowski, M., and De Groot, G.A. (2015) Not all are free-living: High-throughput DNA metabarcoding reveals a diverse community of protists parasitizing soil metazoa. *Mol. Ecol.* **24**: 4556–4569.

Geisen, S., Tveit, A.T., Clark, I.M., Richter, A., Svenning, M.M., Bonkowski, M., and Urich, T. (2015) Metatranscriptomic census of active protists in soils. *ISME J.* **9**: 1–13.

Ginard, M., Lalucat, J., Tummler, B., and Romling, U. (1997) Genome Organization of Pseudomonas stutzeri and Resulting Taxonomic and Evolutionary Considerations. *Int. J. Syst. Bacteriol.* **47**: 132–143.

Gissot, M., Kim, K., Schaap, D., and Ajioka, J.W. (2009) New eukaryotic systematics: A phylogenetic perspective of developmental gene expression in the Apicomplexa. *Int. J. Parasitol.* **39**: 145–151.

Glöckner, F.O., Fuchs, B.M., and Amann, R. (1999) Bacterioplankton compositions of lakes and oceans: A first comparison based on fluorescence in situ hybridization. *Appl. Environ. Microbiol.* **65**: 3721–3726.

Glücksman, E., Bell, T., Griffiths, R.I., and Bass, D. (2010) Closely related protist strains have different grazing impacts on natural bacterial communities. *Environ. Microbiol.* **12**: 3105–3113.

Griffiths, R.I., Thomson, B.C., James, P., Bell, T., Bailey, M.J., and Whiteley, A.S. (2011) The bacterial biogeography of British soils. *Environ. Microbiol.* **13**: 1642–1654.

Gronlien, H.K., Hagen, B., and Sand, O. (2011) Microstome-macrostome transformation in the polymorphic ciliate Tetrahymena vorax leads to mechanosensitivity associated with prey-capture behaviour. *J. Exp. Biol.* **214**: 2258–2266.

Guillou, L., Bachar, D., Audic, S., Bass, D., Berney, C., Bittner, L., et al. (2013) The

Protist Ribosomal Reference database (PR2): A catalog of unicellular eukaryote Small Sub-Unit rRNA sequences with curated taxonomy. *Nucleic Acids Res.* **41**: 597–604.

Hahn, M.W. and Höfle, M.G. (2001) Grazing of protozoa and its effect on populations of aquatic bacteria. *FEMS Microbiol. Ecol.* **35**: 113–121.

Hairston, N.G. (1989) Ecological Experiments: Purpose, Design and Execution.

Hartikainen, H., Ashford, O.S., Berney, C., Okamura, B., Feist, S.W., Baker-Austin, C., et al. (2014) Lineage-specific molecular probing reveals novel diversity and ecological partitioning of haplosporidians. *ISME J.* **8**: 177–186.

Heger, T.J., Edgcomb, V.P., Kim, E., Luke??, J., Leander, B.S., and Yubuki, N. (2014) A resurgence in field research is essential to better understand the diversity, ecology, and evolution of microbial eukaryotes. *J. Eukaryot. Microbiol.* **61**: 214–223.

Hess, S. and Melkonian, M. (2013) The mystery of clade X: Orciraptor gen. nov. and viridiraptor gen. nov. are highly specialised, algivorous amoeboflagellates (Glissomonadida, Cercozoa). *Protist* **164**: 706–747.

Hess, S., Sausen, N., and Melkonian, M. (2012) Shedding light on vampires: The phylogeny of vampyrellid amoebae revisited. *PLoS One* **7**:.

Hiltunen, T., Ayan, G.., Becks, L., Ayan, B., and Becks, L. (2015) Environmental fluctuations restrict eco-evolutionary dynamics in predator – prey system. *Proc. R. Soc. B* **282**: 20150013.

Hiltunen, T. and Becks, L. (2014) Consumer co-evolution as an important component of the eco-evolutionary feedback. *Nat. Commun.* **5**: 5226.

Hiltunen, T., Jones, L.E., Ellner, S.P., and Jr, N.G.H. (2013) Temporal dynamics of a simple community with intraguild predation: an experimental test. *Ecology* **94**: 773–779.

Holyoak, M. and Lawler, S.P. (2005) The Contribution of Laboratory Experiments on Protists to Understanding Population and Metapopulation Dynamics. *Adv. Ecol. Res.* **37**: 245–271.

Howe, A.T., Bass, D., Chao, E.E., and Cavalier-Smith, T. (2011) New Genera, Species, and Improved Phylogeny of Glissomonadida (Cercozoa). *Protist* **162**: 710–722.

Howe, A.T., Bass, D., Scoble, J.M., Lewis, R., Vickerman, K., Arndt, H., and Cavalier-

Smith, T. (2011) Novel Cultured Protists Identify Deep-branching Environmental DNA Clades of Cercozoa: New Genera Tremula, Micrometopion, Minimassisteria, Nudifila, Peregrinia. *Protist* **162**: 332–372.

Howe, A.T., Bass, D., Vickerman, K., Chao, E.E., and Cavalier-Smith, T. (2009) Phylogeny, Taxonomy, and Astounding Genetic Diversity of Glissomonadida ord. nov., The Dominant Gliding Zooflagellates in Soil (Protozoa: Cercozoa). *Protist* **160**: 159–189.

IPCC (2013) IPCC Fifth Assessment Report (AR5). *IPCC*.

Ishigaki, T. and Sleigh, M.A. (2001) Grazing Characteristics and Growth Efficiencies at Two Different Temperatures for Three Nanoflagellates Fed with Vibrio Bacteria at Three Different Concentrations. *Microb. Ecol.* **41**: 264–271.

De Ita, M.E., Marsch-Moreno, R., Guzman, P., and Alvarez-Morales, A. (1998) Physical map of the chromosome of the phytopathogenic bacterium Pseudomonas syringae pv. phaseolicola. *Microbiology* **144**: 493–501.

Jarvis, W.R. and Martone, W.J. (1992) Predominant pathogens in hospital infections. *J. Antimicrob. Chemother.* **29**: 19–24.

Jones, H.L.J. (1997) A classification of mixotrophic protists based on their behaviour. *Freshw. Biol.* **37**: 35–43.

Jones, M.D.M., Forn, I., Gadelha, C., Egan, M.J., Bass, D., Massana, R., and Richards, T.A. (2011) Discovery of novel intermediate forms redefines the fungal tree of life. *Nature* **474**: 200–203.

Jost, L. (2006) Entropy and di v ersity. *Oikos* **113**: 363–375.

Jousset, A. (2012) Ecological and evolutive implications of bacterial defences against predators. *Environ. Microbiol.* **14**: 1830–1843.

Jousset, A., Lara, E., Wall, L.G., and Valverde, C. (2006) Secondary metabolites help biocontrol strain Pseudomonas fluorescens CHA0 to escape protozoan grazing. *Appl. Environ. Microbiol.* **72**: 7083–7090.

Julia, J., Yossef, C., Marina, de L., Ariel, K., Edouard, J., and Antonis, C. (2014) Multiple micro-predators controlling bacterial communities in the environment. *Curr. Opin. Biotechnol.* **27**: 185–190.

Jürgens, K. and Massana, R. (2008) Protist grazing on marine bacterioplankton. *Microb. Ecol. Ocean.* 383–441.

Jürgens, K. and Matz, C. (2002) Predation as a shaping force for the phenotypic and genotypic composition of planktonic bacteria. *Antonie van Leeuwenhoek, Int. J. Gen. Mol. Microbiol.* **81**: 413–434.

Kasalický, V., Jezbera, J., Hahn, M.W., and Šimek, K. (2013) The Diversity of the Limnohabitans Genus, an Important Group of Freshwater Bacterioplankton, by Characterization of 35 Isolated Strains. *PLoS One* **8**:.

Katoh, K. and Standley, D.M. (2013) MAFFT multiple sequence alignment software version 7: Improvements in performance and usability. *Mol. Biol. Evol.* **30**: 772–780.

De Kievit, T.R. (2009) Quorum sensing in Pseudomonas aeruginosa biofilms. *Environ. Microbiol.* **11**: 279–288.

Kneitel, J.M. and Chase, J.M. (2004) Trade-offs in community ecology: Linking spatial scales and species coexistence. *Ecol. Lett.* **7**: 69–80.

Kreuzer, K., Adamczyk, J., Iijima, M., Wagner, M., Scheu, S., and Bonkowski, M. (2006) Grazing of a common species of soil protozoa (Acanthamoeba castellanii) affects rhizosphere bacterial community composition and root architecture of rice (Oryza sativa L.). *Soil Biol. Biochem.* **38**: 1665–1672.

Lalucat, J., Bennasar, A., Bosch, R., García-Valdés, E., and Palleroni, N.J. (2006) Biology of Pseudomonas stutzeri. *Microbiol. Mol. Biol. Rev.* **70**: 510–547.

Landry, M.R., Lehner-Fournier, J.M., Sundstrom, J.A., Fagerness, V.L., and Selph, K.E. (1991) Discrimination between living and heat-killed prey by a marine zooflagellate, Paraphysomonas vestita (Stokes). *J. Exp. Mar. Bio. Ecol.* **146**: 139–151.

Langmead, B. and Salzberg, S.L. (2012) Fast gapped-read alignment with Bowtie 2. *Nat Methods* **9**: 357–359.

Lawrence, D., Fiegna, F., Behrends, V., Bundy, J.G., Phillimore, A.B., Bell, T., and Barraclough, T.G. (2012) Species interactions alter evolutionary responses to a novel environment. *PLoS Biol.* **10**:.

Lê, S., Josse, J., and Husson, F. (2008) FactoMineR : An R package for multivariate analysis. *J Stat Softw* **25**: 1–18.

Legendre, P. and Legendre, L. (1998) Numerical Ecology, Volume 24, Second Edition (Developments in Environmental Modelling) Elsevier Science.

Lehembre, F., Doillon, D., David, E., Perrotto, S., Baude, J., Foulon, J., et al. (2013) Soil metatranscriptomics for mining eukaryotic heavy metal resistance genes. *Environ. Microbiol.* **15**: 2829–2840.

Lesaulnier, C., Papamichail, D., McCorkle, S., Ollivier, B., Skiena, S., Taghavi, S., et al. (2008) Elevated atmospheric CO2 affects soil microbial diversity associated with trembling aspen. *Environ. Microbiol.* **10**: 926–941.

Lotka, A.J. (1925) Elements of Physical Biology. *Williams Wilkins Co.* 435.

MacLean, R.C., Bell, G., and Rainey, P.B. (2004) The evolution of a pleiotropic fitness tradeoff in Pseudomonas fluorescens. *Proc. Natl. Acad. Sci. U. S. A.* **101**: 8072–8077.

Madsen, E.L. (2011) Microorganisms and their roles in fundamental biogeochemical cycles. *Curr. Opin. Biotechnol.* **22**: 456–464.

Martin, M. (2011) Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet.journal* **17**: 10.

Matz, C. and Kjelleberg, S. (2005) Off the hook - How bacteria survive protozoan grazing. *Trends Microbiol.* **13**: 302–307.

Mazzola, M., De Bruijn, I., Cohen, M.F., and Raaijmakers, J.M. (2009) Protozoan-induced regulation of cyclic lipopeptide biosynthesis is an effective predation defense mechanism for Pseudomonas fluorescens. *Appl. Environ. Microbiol.* **75**: 6804–6811.

McDonald, M.J., Gehrig, S.M., Meintjes, P.L., Zhang, X.X., and Rainey, P.B. (2009) Adaptive divergence in experimental populations of Pseudomonas fluorescens. IV. Genetic constraints guide evolutionary trajectories in a parallel adaptive radiation. *Genetics* **183**: 1041–1053.

Medlin, L., Elwood, H.J., Stickel, S., and Sogin, M.L. (1988) The characterization of enzymatically amplified eukaryotes 16S-like ribosomal RNA coding regions. *Gene* **71**: 491–500.

Meunier, C.L., Hantzsche, F.M., Dupont, A.Ö.C., Haafke, J., Oppermann, B., Malzahn, A.M., and Boersma, M. (2012) Intraspecific selectivity, compensatory feeding and flexible homeostasis in the phagotrophic flagellate Oxyrrhis marina: Three ways to handle food quality fluctuations. *Hydrobiologia* **680**: 53–62.

Miller, M.A., Pfeiffer, W., and Schwartz, T. (2010) Sc2010_Paper.

Montagnes, D.J.S., Barbosa, A.B., Boenigk, J., Davidson, K., Jürgens, K., Macek, M., et al. (2008) Selective feeding behaviour of key free-living protists: Avenues for continued study. *Aquat. Microb. Ecol.* **53**: 83–98.

Moore, J.C., Berlow, E.L., Coleman, D.C., De Suiter, P.C., Dong, Q., Hastings, A., et al. (2004) Detritus, trophic dynamics and biodiversity. *Ecol. Lett.* **7**: 584–600.

Morin, P.J. (2011) Community Ecology.

Moser, M. and Weisse, T. (2011) The outcome of competition between the two chrysomonads Ochromonas sp. and Poterioochromonas malhamensis depends on pH. *Eur. J. Protistol.* **47**: 79–85.

Mougi, A. (2012a) Predator-prey coevolution driven by size selective predation can cause anti-synchronized and cryptic population dynamics. *Theor. Popul. Biol.* **81**: 113–118.

Mougi, A. (2012b) Unusual predator-prey dynamics under reciprocal phenotypic plasticity. *J. Theor. Biol.* **305**: 96–102.

Mucibabic, S. (1957) The growth of mixed populations of Chilomonas paramecium and Tetrahymena pyriformis. *J. Gen. Microbiol.* **16**: 561–571.

Murase, J. and Frenzel, P. (2008) Selective grazing of methanotrophs by protozoa in a rice field soil. *FEMS Microbiol. Ecol.* **65**: 408–414.

Murrell, D.J. (2005) Local spatial structure and predator-prey dynamics: counterintuitive effects of prey enrichment. *Am. Nat.* **166**: 354–367.

Murrell, D.J. (2010) When does local spatial structure hinder competitive coexistence and reverse competitive hierarchies? *Ecology* **91**: 1605–1616.

Nelder, J.A. and Wedderburn, R.W.M. (1972) Generalized Linear Models. *J. R. Stat. Soc. A.* **135**: 370–384.

Neuhauser, S., Kirchmair, M., Bulman, S., and Bass, D. (2014) Cross-kingdom host shifts of phytomyxid parasites. *BMC Evol. Biol.* **14**: 33.

Nicholson, A.J. and Bailey, V.A. (1935) The balance of animal populations. *J. Anim. Ecol.* **2**: 132–178.

Oksanen, J., Kindt, R., Legendre, P., O'Hara, B., Henry, M., Stevens, H., and Suggests, M.J.O.M.A.S.S. (2007) The vegan package. *Community Ecol. Packag.* **10**: 631–637.

Palleroni, N.J. (1994) *Pseudomonas* classification. *Antonie Van Leeuwenhoek* **64**: 231–251.

Palleroni, N.J. (2010) The pseudomonas story. *Environ. Microbiol.* **12**: 1377–1383.

Passalacqua, K.D., Varadarajan, A., Ondov, B.D., Okou, D.T., Zwick, M.E., and Bergman, N.H. (2009) Structure and complexity of a bacterial transcriptome. *J. Bacteriol.* **191**: 3203–3211.

Pawlowski, J. (2014) Protist Evolution and Phylogeny. *eLS* 1–9.

Pearce, D.A., Newsham, K.K., Thorne, M.A.S., Calvo-Bado, L., Krsek, M., Laskaris, P., et al. (2012) Metagenomic analysis of a southern maritime Antarctic soil. *Front. Microbiol.* **3**: 1–13.

Pelletier, F., Garant, D., and Hendry, A.P. (2009) Eco-evolutionary dynamics. *Philos. Trans. R. Soc. B Biol. Sci.* **364**: 1483–1489.

Pernthaler, J. (2005) Predation on prokaryotes in the water column and its ecological implications. *Nat. Rev. Microbiol.* **3**: 537–546.

Pernthaler, J., Posch, T., Simek, K., Vrba, J., Amann, R., and Psenner, R. (1997) Contrasting bacterial strategies to coexist with a flagellate predator in an experimental microbial assemblage. *Appl. Environ. Microbiol.* **63**: 596–601.

Pernthaler, J., Posch, T., Simek, K., Vrba, J., Pernthaler, A., Glöckner, F.O., et al. (2001) Predator-Specific Enrichment of Actinobacteria from a Cosmopolitan Freshwater Clade in Mixed Continuous Culture. *Appl. Environ. Microbiol.* **67**: 2145–2155.

Quast, C., Pruesse, E., Yilmaz, P., Gerken, J., Schweer, T., Yarza, P., et al. (2013) The SILVA ribosomal RNA gene database project: Improved data processing and web-based tools. *Nucleic Acids Res.* **41**: 590–596.

Quintana, X.D., Arim, M., Badosa, A., Blanco, J.M., Boix, D., Brucet, S., et al. (2014) Predation and competition effects on the size diversity of aquatic communities. *Aquat. Sci.* **77**: 45–57.

Rainey, P.B. and Bailey, M.J. (1996) Physical and genetic map of the Pseudomonas fluorescens SBW25 chromosome. *Mol. Microbiol.* **19**: 521–33.

Rainey, P.B., Bailey, M.J., and Thompson, I.P. (1994) Phenotypic and genotypic diversity of fluorescent pseudomonads isolated f rom field- grown sugar beet.

Rainey, P.B. and Rainey, K. (2003) Evolution of cooperation and conflict in experimental bacterial populations. *Nature* **425**: 72–74.

Rainey, P.B. and Travisano, M. (1998) Adaptive radiation in a heterogeneous environment. *Nature* **394**: 69–72.

Ramette, A. (2007) Multivariate analyses in microbial ecology. *FEMS Microbiol. Ecol.* **62**: 142–160.

Ren, L., He, D., Zeng, J., and Wu, Q.L. (2013) Bacterioplankton communities turn unstable and become small under increased temperature and nutrient-enriched conditions. *FEMS Microbiol. Ecol.* **84**: 614–624.

Richards, T.A. and Bass, D. (2005) Molecular screening of free-living microbial eukaryotes: Diversity and distribution using a meta-analysis. *Curr. Opin. Microbiol.* **8**: 240–252.

Robinson, M.D., McCarthy, D.J., and Smyth, G.K. (2009) edgeR: A Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* **26**: 139–140.

Rosling, A., Cox, F., Cruz-Martinez, K., Ihrmark, K., Grelet, G.-A., Lindahl, B.D., et al. (2011) Archaeorhizomycetes: unearthing an ancient class of ubiquitous soil fungi. *Science* **333**: 876–879.

Rowe, G., Sweet, M., and Beebee, T. (2004) An Introduction to Molecular Ecology.

Rueckert, S., Chantangsi, C., and Leander, B.S. (2010) Molecular systematics of marine gregarines (Apicomplexa) from North-eastern Pacific polychaetes and nemerteans, with descriptions of three novel species: Lecudina phyllochaetopteri sp. nov., Difficilina tubulani sp. nov. and Difficilina paranemertis sp. n. *Int. J. Syst. Evol. Microbiol.* **60**: 2681–2690.

Rueckert, S., Simdyanov, T.G., Aleoshin, V. V., and Leander, B.S. (2011) Identification of a divergent environmental DNA sequence clade using the phylogeny of gregarine parasites (apicomplexa) from crustacean hosts. *PLoS One* **6**:.

Salcher, M.M. (2014) Same same but different: Ecological niche partitioning of planktonic freshwater prokaryotes. *J. Limnol.* **73**: 74–87.

Salcher, M.M., Pernthaler, J., Psenner, R., and Posch, T. (2005) Succession of bacterial grazing defense mechanisms against protistan predators in an experimental microbial community. *Aquat. Microb. Ecol.* **38**: 215–229.

Scanlan, P.D., Hall, A.R., Blackshields, G., Friman, V.-P., Davis, M.R., Goldberg, J.B., and Buckling, A. (2015) Coevolution with bacteriophages drives genome-wide host evolution and constrains the acquisition of abiotic-beneficial mutations. *Mol. Biol. Evol.* **32**: 1425–35.

Schoener, T.W. (1983) Field Experiments on Interspecific Competition. *Am. Nat.* **122**: 240–285.

Schröter, D., Wolters, V., and De Ruiter, P.C. (2003) C and N mineralisation in the decomposer food webs of a European forest transect. *Oikos* **102**: 294–308.

Sherr, E.B. and Sherr, B.F. (2007) Heterotrophic dinoflagellates: A significant component of microzooplankton biomass and major grazers of diatoms in the sea. *Mar. Ecol. Prog. Ser.* **352**: 187–197.

Sherr, E.B. and Sherr, B.F. (2002) Significance of predation by protists. *Antonie Van Leeuwenhoek* **81**: 293–308.

Silby, M.W., Winstanley, C., Godfrey, S.A.C., Levy, S.B., and Jackson, R.W. (2011) Pseudomonas genomes: Diverse and adaptable. *FEMS Microbiol. Rev.* **35**: 652–680.

Šimek, K., Kasalický, V., Jezbera, J., Horňák, K., Nedoma, J., Hahn, M.W., et al. (2013) Differential freshwater flagellate community response to bacterial food quality with a focus on Limnohabitans bacteria. *ISME J.* **7**: 1519–30.

Simek, K., Kojecká, P., Nedoma, J., Hartman, P., Vrba, J., and Dolan, J.R. (1999) Shifts in bacterial community composition associated with different microzooplankton size fractions in a eutrophic reservoir. *Limnol. Oceanogr.* **44**: 1634–1644.

Sintes, E. and del Giorgio, P.A. (2014) Feedbacks between protistan single-cell activity and bacterial physiological structure reinforce the predator/prey link in microbial foodwebs. *Front. Microbiol.* **5**: 1–11.

Song, C., Mazzola, M., Cheng, X., Oetjen, J., Alexandrov, T., Dorrestein, P., et al. (2015) Molecular and chemical dialogues in bacteria-protozoa interactions. *Sci. Rep.* **5**: 12837.

Spiers, A.J., Kahn, S.G., Bohannon, J., Travisano, M., and Rainey, P.B. (2002) Adaptive divergence in experimental populations of Pseudomonas fluorescens. I. Genetic and phenotypic bases of wrinkly spreader fitness. *Genetics* **161**: 33–46.

Stamatakis, A., Hoover, P., and Rougemont, J. (2008) A rapid bootstrap algorithm for

the RAxML Web servers. *Syst. Biol.* **57**: 758–71.

Stanier, R.Y., Palleroni, N.J., and Doudoroff, M. (1966) The aerobic pseudomonads: a taxonomic study. *J. Gen. Microbiol.* **43**: 159–271.

Stout, J.D. (1984) The protozoan fauna of a seasonally inundated soil under grassland. *Soil Biol. Biochem.* **16**: 121–125.

Takahashi, Y., Yoshida, M., Inouye, I., and Watanabe, M.M. (2014) Diplophrys mutabilis sp. nov., a New Member of Labyrinthulomycetes from Freshwater Habitats. *Protist* **165**: 50–65.

Takishita, K., Tsuchiya, M., Kawato, M., Oguri, K., Kitazato, H., and Maruyama, T. (2007) Genetic Diversity of Microbial Eukaryotes in Anoxic Sediment of the Saline Meromictic Lake Namako-ike (Japan): On the Detection of Anaerobic or Anoxic-tolerant Lineages of Eukaryotes. *Protist* **158**: 51–64.

Thurman, J., Parry, J.D., Hill, P.J., and Laybourn-Parry, J. (2010) The Filter-Feeding Ciliates Colpidium striatum and Tetrahymena pyriformis Display Selective Feeding Behaviours in the Presence of Mixed, Equally-Sized, Bacterial Prey. *Protist* **161**: 577–588.

Turner, T.R., Ramakrishnan, K., Walshaw, J., Heavens, D., Alston, M., Swarbreck, D., et al. (2013) Comparative metatranscriptomics reveals kingdom level changes in the rhizosphere microbiome of plants. *ISME J.* **7**: 2248–58.

Urich, T., Lanzén, A., Qi, J., Huson, D.H., Schleper, C., and Schuster, S.C. (2008) Simultaneous assessment of soil microbial community structure and function through analysis of the meta-transcriptome. *PLoS One* **3**:.

Vargas, C. de, Audic, S., Henry, N., Decelle, J., Mahé, F., Logares, R., et al. (2015) Eukaryotic plankton diversity in the sunlit ocean. *Science (80-.).* **348**:.

Wakeman, K.C. and Leander, B.S. (2013) Molecular phylogeny of marine gregarine parasites (Apicomplexa) from tube-forming polychaetes (sabellariidae, cirratulidae and serpulidae), including descriptions of two new species of selenidium. *J. Eukaryot. Microbiol.* **60**: 514–525.

Wakeman, K.C. and Leander, B.S. (2012) Molecular phylogeny of pacific archigregarines (Apicomplexa), including descriptions of Veloxidium leptosynaptae n. gen., n. sp., from the sea cucumber Leptosynapta clarki (Echinodermata), and two new species of Selenidium. *J. Eukaryot. Microbiol.* **59**: 232–245.

Weekers, P.H., Gast, R.J., Fuerst, P. a, and Byers, T.J. (1994) Sequence variations in

small-subunit ribosomal RNAs of Hartmannella vermiformis and their phylogenetic implications. *Mol. Biol. Evol.* **11**: 684–690.

Weisse, T. (2008) Distribution and diversity of aquatic protists: An evolutionary and ecological perspective. *Biodivers. Conserv.* **17**: 243–259.

West-Eberhard, M.J. (1989) Phenotypic Plasticity and the Origins of Diversity. *Annu. Rev. Ecol. Syst.* **20**: 249–278.

Wilks, S.S. (1938) The large sample distribution of the likelihood ratio for testing composite hypotheses. *Ann. Math. Stat.* **231**: 60–62.

Woese, C.R. (1987) Bacterial Evolution. *Microbiology* **51**: 221–271.

Yoshida, T., Jones, L.E., Ellner, S.P., Fussmann, G.F., and Hairston, N.G. (2003) Rapid evolution drives ecological dynamics in a predator-prey system. *Nature* **424**: 303–6.

Zöllner, E., Hoppe, H.-G., Sommer, U., and Jürgens, K. (2009) Effect of zooplankton-mediated trophic cascades on marine microbial food web components (bacteria, nanoflagellates, ciliates). *Limnol. Oceanogr.* **54**: 262–275.