The Evolution of Cooperation: Insights from Experimental Populations of *Pseudomonas fluorescens*

Peter L. Meintjes

A thesis submitted in partial fulfilment of the
requirements of the degree of
Doctor of Philosophy in Biological Sciences
The University of Auckland
New Zealand
March 2009

0.1 ABSTRACT

The field of experimental evolution is burgeoning under the power of microbial systems. Our ability to manipulate experimental design for use with microbes is only limited by our imagination. This thesis is a study that uses *Pseudomonas fluorescens*, a soil dwelling bacterium, as an experimental tool for understanding evolutionary processes. The evolution of cooperation has been a thorny issue for many years, because it initially seems to contradict the intrinsically selfish concepts established in Darwin's theory of evolution by natural selection. Advances in microbiology and the ability to test important evolutionary theories using microbes, provides an exciting opportunity for those working in the field of experimental evolution.

This thesis uses *P. fluorescens* to investigate four aspects of the evolution of cooperative behaviour organised into four results chapters (Chapters 2-5). The first describes the genotypic and phenotypic diversity of 26 independently derived 'wrinkly spreader' genotypes in order to analyse the genetic and phenotypic variation among morphotypes. Mutations were identified in 25 of the 26 wrinkly spreaders including a new locus *mws* and three new genes of known loci *wspE*, *awsR* and *awsO*. This new genetic information provided additional insight into the molecular causes of the wrinkly spreader phenotype. Multivariate analysis of the phenotypic traits revealed that *wspF* mutants were phenotypically distinct from other morphotypes at a level below the ecological niche. The second chapter extended existing studies on the evolution of wrinkly spreader genotypes within the well-known Haystack model for evolution in group-structured populations, by studying the population dynamics of cooperative genotypes with and without group structure, in a multi-level selection one framework. It was shown that the time spent in a haystack affects the fitness of cooperators, because the longer group-generation treatment conformed to the predictions of the Haystack model, while the shorter group-generation treatment did not. The third chapter was an investigation into how the fitness of the emergent group-level phenotype formed by cooperating wrinkly spreader cells was dependent on the density of wrinkly spreader cells. Contrary to prediction, no density dependence was observed when calculated in a multi-level selection one framework, but rather it was determined that the emergent fitness was dependent on time, implicating a role for a development-like process. The final

results chapter of this thesis incorporated the hypothesised role for a development-like process into a novel theoretical model for the evolution of multicellularity in which fitness would be determined in a multi-level selection two framework. Novel apparatus and experimental design were developed to determine if it were possible to observe a response to a selective regime that selected simultaneously at the level of the individual cell and the level of the group of cells. A significant response was shown after only six group-generation cycles.

In summary, this thesis exploits *P. fluorescens* as an experimental tool to gain insight into complex ecological and evolutionary phenomena such as cooperation, biofilm formation and the evolution of multicellularity, and provides insight into the molecular causes of the cooperation among wrinkly spreader genotypes.

0.2 ACKNOWLEDGEMENTS

This thesis is dedicated to those who have assisted me in achieving my goal of completing a PhD thesis, for it bares no thought as to how such an undertaking may have played out without it.

For supervising me in the official capacity of completing this thesis, I thank Professor Paul B. Rainey, Dr. Tim F. Cooper and Professor Allen G. Rodrigo.

For supervising me in an unofficial capacity by providing me with endless professional and social support, and improving me as an academic and a person, I thank Dr. Hubertus J. E. Beaumont and Dr. Christian Kost.

For providing the ongoing parental support and unyielding belief in my ability, I thank my parents, Elsie and Liston Meintjes.

For providing me with extra statistical assistance, I thank Professor Allen G. Rodrigo and Dr. Tim F. Cooper.

For providing assistance in achieving wet-lab experimental success, I thank, Dr. Bertus Beaumont, Dr. Xue-Xian Zhang, Dr. Dominik Refardt, Dr. Darby Brown, Dr. Christian Kost, Michael McDonald, Jenna Gallie, Jeremy Gray, Annabel Gunn, Dr. Tim Cooper and Professor Paul B. Rainey.

For reading this thesis before it was good enough to hand to my supervisors, I thank Mara Mulrooney, Jenna Gallie, Lisa Meintjes and particularly Christian Kost whose constructive criticism was highly significant in allowing me to complete ($F_{1,1}$ = 42.23, $p < 0.0001$).

For the supply of photographic equipment, I thank Iain McDonald.

For making the flatting environment a sanctuary of good times at “Carlton L. Gore” and No.1 Ireland Street, I thank Nicholas Mabey, James Coe, Tim Elliot, Mara Mulrooney and Nick Gow.

For being part of my on campus distraction team, I thank Matt Barrett, Sean Tobin, Stu Preece, Hayden Smith, James Coe, Nicholas Mabey, Lincoln Colling, Alistair Law, Joshua Bradley, Richard Bunker, Aashish Patel, Manea Sweeney, Michelle Fox and Mara Mulrooney.

For organising netball games, I thank Michelle Fox, Julie Harper, Rachael Fullard, Jessica Chaston and Richard Bunker.

For teaching me to juggle, I thank Bertus Beaumont.

For making post-graduate functions the most successful functions at the University of Auckland, I thank Hayden Smith, Aashish Patel and Sean Tobin.

For providing me with a Bright Futures Top Achiever Doctoral Scholarship, I thank the Tertiary Education Commission (TEC).

For providing me with a University of Auckland Doctoral Scholarship extension, I thank the University of Auckland Research Committee.

For taking excellent and efficient care of all my scholarship-associated finances, I thank Angela Pearse.

For providing me with extra funding to attend conferences and workshops I thank the committee of the SBS contestable travel fund, the graduate research fund of the UARC and the PReSS funding scheme.

For providing me with the opportunity to tutor and demonstrate various undergraduate courses, I thank Professor Allen Rodrigo, Associate Professor Don Love, Dr. Howard Ross, Dr. Marti Anderson and Dr. Mat Goddard.

0.3 TABLE OF CONTENTS

0.4 GLOSSARY AND TERMS

Actor: The focal individual that performs a behaviour (West *et al.*, 2007b).

Adaptation: A process of genetic change of a population, owing to natural selection, whereby the average state of a character becomes improved with reference to a specific function, or whereby a population is thought to have become better suited to some feature of its environment (Futuyma, 1998).

Adaptive Radiation: The evolution of ecological diversity within a rapidly multiplying lineage characterised by phenotypic divergence and speciation due to the availability of environments, resources and resource competition (Schluter, 2000).

Altruism: A behaviour that is costly to the actor and beneficial to the recipient(s) where these costs and benefits are defined by the consequences on the lifetime fitness of the actor and the recipient, *i.e.* –/+ (West *et al.*, 2007b).

Cheater: An individual that does not cooperate (or contributes less than its 'fair share'), and gains the benefit of others cooperating (Velicer, 2003; West *et al.*, 2007b).

Cheating: Obtaining benefits from a collectively produced public good that are disproportionately large relative to a cheater's own contribution to that good (Velicer, 2003).

Coevolution: Evolution in two or more species in which the evolutionary changes of each species influence the evolution of the other species (Ridley, 2004).

Cooperation: Any action *selected to enhance* the fitness of others at a relative cost to the acting individual (West *et al.*, 2006). Cooperation includes all behaviours that are altruistic (–/+) and some mutually beneficial (+/+) behaviours.

Cooperator: An individual that provides a benefit that increases the fitness of another individual (other individuals) at a relative cost to itself (West *et al.*, 2006).

Defector: An individual that does not cooperate, but may or may not benefit from the cooperative benefits of others. Therefore, biologically not all defectors are cheats (Velicer, 2003).

Deme: A group of individuals that readily intermix during some point in their life cycle, giving any two an equal probability of becoming neighbours (Wilson, 1977).

Direct fitness: The component of fitness gained through the impact of an individual's behaviour on the production of [their own] offspring (Hamilton, 1964a; West *et al.*, 2007b).

Emergence: The appearance of emergent properties. *See also* Emergent property.

Emergent character: An emergent property in which any character of groups originating from non-additive interaction among lower-level units that do not exist at the lower level (Gould, 2002 p.657).

Emergent fitness: Any trait that characterises or influences the differential rate of proliferation of groups in interaction with the environment (Gould, 2002 p.659).

Emergent property: Any property arising from an interaction among individuals that is 'not otherwise attainable' (Corning, 2002), because the effect is qualitatively or quantitatively 'greater than the sum of the parts' (Gould, 2002).

Exaptation: A character previously evolved for one reason (whether or not as an adaptation), and then coopted for utility in another role (Gould and Vrba, 1982).

Fitness: The average number of offspring produced by individuals with a certain genotype relative to the number produced by individuals with other genotypes (Ridley, 2004).

Focal individual (FI): The focal individual is the individual for whom fitness is evaluated in mathematical calculation and is also the conceptual individual for the description of the model.

Gene flow: The movement of genes into, or through, a population by interbreeding or by migration and interbreeding (Ridley, 2004).

Genetic drift: Random changes in gene frequencies in a population (Ridley, 2004).

Group-generation: In MLS 1, the period of time between successive formations of groups, *i.e.* the period of time spent within a group. In MLS 2, the period of time between the reproduction of successive group offspring.

Group selection: The process of genetic change caused by the differential proliferation and extinction of groups of organisms (Wright, 1945; Wynne-Edwards, 1962; Maynard Smith, 1964; Williams, 1966; Lewontin, 1970; Wade, 1977). *See also* trait-group selection.

Kin selection: The process of selection by which traits are favoured because of their beneficial effects on the fitness of relatives (West *et al.*, 2007b).

Inclusive fitness: The combination of an individual's direct fitness plus the indirect fitness of that individual's effect on all other individuals (Grafen, 1984).

Indirect fitness: The component of fitness gained from aiding the reproduction of related individuals (Hamilton, 1964a; West *et al.*, 2007b).

Individual: A physiologically discrete organism (Buss, 1987) that satisfies Lewontin's conditions for a unit of selection (Lewontin, 1970). *See also* Unit of selection.

Individual Selection: Selection that favours the spread of a trait through a population based only on the number of offspring left by that individual (Grafen, 1984).

Individuality: The properties of an individual. *See also* Individual.

Interdemic group selection: Selection between groups where group membership is defined by interactions between individuals and not by all individuals in the vicinity (Wilson, D.S., 1975).

Iterated Prisoner's Dilemma: A successively repeated version of the Prisoner's Dilemma.

Local group: A subset of the population that interact with one another; the local group may vary from the perspective of different behaviours or traits (West *et al.*, 2007b).

Malthusian parameter: The ratio of final to initial population density after bacterial growth (Lenski *et al.*, 1991).

Malthusian ratio: The relative fitness of a strain compared to a competitor calculated by taking the ratio of Malthusian parameters (Lenski *et al.*, 1991).

Multi-level selection 1: Multi-level selection 1 is said to occur whenever an individual's expected viability, mating success, and/or fertility cannot be accounted for solely on the basis of that individual's phenotype, but rather additional information is required about properties of the group or groups of which the individual is a member (Heisler and Damuth, 1987).

Multi-level selection 2: Multi-level selection 2 occurs whenever any group properties co-vary with group-level fitness, implying that the proportions of different kinds of gorups will change in the population (and noting that group characters may change as a result of lower-level selection among the individuals that the groups comprise) (Heisler and Damuth, 1987).

Multicellularity: The property of an individual that spends part of its life cycle as an entity consisting of more than a single cell (Michod and Roze, 1997; Michod and Roze, 2001; Michod *et al.*, 2005).

Mutual benefit: A behaviour which is beneficial to both the actor and the recipient, *i.e.* +/+ (West *et al.*, 2007b).

Mutualism: Cooperation between species (West *et al.*, 2007b).

Natural selection: The differential reproductive success of evolutionary individuals based on the fitnesses of their traits in interaction with the environment (Gould, 2002).

Phenotype: Any observable physical manifestation of an organism, such as its morphology, development, biochemical or physiological properties, function or behaviour.

Prisoner's Dilemma (PD): A popular non-zero sum game theory approach to analysing cooperation in which cooperation by two individuals generates the highest average payoff, but unilateral defection gives the greatest individual advantage regardless of the choice of the other prisoner.

Public good: Any fitness-enhancing resource that is accessible to multiple individuals within a local group (Velicer, 2003).

Recipient: Any individual receiving the benefit of a cooperative behaviour (West *et al.*, 2007b). Recipients may be cooperators or defectors.

Trait-group selection: *See* interdemic group selection.

Trait-groups: Populations enclosed in areas smaller than the boundaries of the deme (Wilson 1975).

Unit of selection: Any physical entity in nature that has variation, reproduction, and heritability (Lewontin, 1970).

Weak altruism: any action that provides a benefit to others that leads to a decrease in the fitness of the focal individual, relative to the other members of its group (Wilson, D.S., 1975; Wilson, 1977; West *et al.*, 2007b).

Wild-type: The normal or ancestral form of members of a species, as distinct from derived mutant forms (Ridley, 2004).

0.5 ABBREVIATIONS

aa: Amino acid

ANOSIM: Analysis of similarity

ANOVA: Analysis of variance

BLAST: Basic local alignment search tool

c-di-GMP: Cyclic-di-guanosine monophosphate

CDD: Conserved domain database

CR: Congo red

DGC: Di-guanylate cyclase

DMF: Dimethyl Formamide

EPS: Exopolysaccharide

FI: Focal individual

GLS: Group level selection

IPD: Iterated Prisoner's Dilemma

IWS: Independent wrinkly spreader isolates

KB: King's medium B (King *et al.*, 1954)

LB: Lysogeny broth (Bertani, 1951; Bertani, 2004)

LSWS: Large spreading wrinkly spreader

MLS 1: Multi-level selection 1 (Heisler and Damuth, 1987; Damuth and Heisler, 1988)

MLS 2: Multi-level selection 2 (Heisler and Damuth, 1987; Damuth and Heisler, 1988)

MPD: Modified Petri dish

NCBI: National Centre for Biotechnology Information

NF: Nitrofurontoin

OD: Optical density

PDE: Phosphodiesterase

PTFE: Poly-tetrafluoroethene

REC: Signal receiver domain

SM: Smooth colony morphotype of *P. fluorescens* (Rainey and Travisano, 1998)

SNP: Single nucleotide polymorphism

SRC: Selection rate constant (Lenski *et al.*, 1991)

TMHHM: Transmembrane hidden Markov model

WS: Wrinkly spreader colony morphotype of *P.fluorescens* (Rainey and Travisano, 1998)

0.6 LIST OF TABLES

0.7 List of Figures

1 INTRODUCTION

The major tenets of the evolutionary synthesis, then, were that populations contain genetic variation that arises by random (i.e. not adaptively directed) mutation and recombination; that populations evolve by changes in gene frequency brought about by random genetic drift, gene flow, and especially natural selection; that most adaptive genetic variants have individually slight phenotypic effects so that phenotypic changes are gradual (although some alleles with discrete effects may be advantageous, as in certain colour polymorphisms); that diversification comes about by speciation, which normally entails the gradual evolution of reproductive isolation among populations; and that these processes, continued for sufficiently long, give rise to changes of such great magnitude as to warrant the designation of higher taxonomic levels (genera, families, and so forth).

-- Futuyma, D. J. 1986

1.1 Evolution: The Unity of Life

One of the most commonly quoted statements about evolutionary biology is that "nothing makes sense in biology except in the light of evolution" (Dobzhansky, 1964). The strength of this statement underpins a now almost ubiquitous acknowledgement among biologists and indeed the greater scientific community that the entire living world has arisen from common ancestors by means of an evolutionary process. The fact that every living organism on the planet shares a common ancestor estimated to have lived between 3 and 3.5 billion years ago (Cavalier-Smith, 2006) is entirely remarkable. In this section, I introduce the concepts of evolution by natural selection as pioneered by Darwin and as extended by the contributors to the Modern Synthesis, and then illustrate how the core features of evolution by natural selection have a generality that allows evolution to be applied simultaneously to multiple levels of the biological hierarchy observed in nature.

1.1.1 Darwin's Legacy

Darwin's original principle of evolution by natural selection as laid out in *The Origin of Species* (Darwin, 1859) is frequently considered one of the great contributions to science. Its monumental impact on science and society has seen it put under the harshest scrutiny, yet many of the concepts developed by Darwin have withstood a barrage of scepticism and continue to form the core of the way biologists think about the evolution and maintenance of the diversity in nature. Onboard the HMS Beagle as a companion of Captain Robert Fitzroy, Darwin pursued his skills as a naturalist, and after returning in 1836 from five years abroad, began to formulate his theory of evolution by natural selection. The core of Darwinian evolution is descent with modification; that is that offspring resemble their parents, but that small changes in what is inherited leads to new variation, and that the finite nature of available

resources provides a 'struggle for existence' that leads to only the 'fittest' surviving[1]. Darwin (1859 p.5) writes:

> *As many more individuals of each species are born than can possibly survive; and as, consequently, there is a frequently recurring struggle for existence, it follows that any being, if it vary however slightly in any manner profitable to itself, under the complex and sometimes varying conditions of life, will have a better chance of surviving, and thus be naturally selected. From the strong principle of inheritance, any selected variety will tend to propagate its new and modified form.*

One aspect of evolution of which Darwin was unaware was the particulate nature of inheritance. As a result, he was troubled by how diversity could be maintained in the face of relentless 'blending inheritance'. Although the evidence for discrete inheritance of genetic information was published by Gregor Mendel (1865) while Darwin was alive, it was not until the work of R. A. Fisher, J. B. S. Haldane and others in the 1920s and 30s that this was properly contextualised. The progress of evolutionary thought was further hampered by a division between geneticists and naturalists (systematists), due largely to the lack of familiarity with the advances in each others' work (Mayr, 1980). The 1930s and 40s saw a younger generation of biologists that brought about the amalgamation of the concepts developed by the geneticists and the naturalists to form what is now termed the Modern Synthesis, a phrase coined by Julian Huxley in his book *Evolution: The Modern Synthesis* (Huxley, 1942). The major contributors (T. Dobzhansky, J. Huxley, E. Mayr, B. Rensch, G. G. Simpson, and G. L. Stebbins)[2] abandoned the divisions of their

---

[1] The great philosopher Karl Popper was extremely dismissive of the concept "survival of the fittest" as it appeared to him to be a tautology. "Who are the fittest?' he asked, "Those who survive, of course". However, this attack is defensible, because fitness, although a truly abstract concept unique to biology, is usually defined as a *property of organisms* (Preface to *Darwinian Dynamics* – Michod, 1999).

[2] Mayr (1982 p. 568) notes that the six authors mentioned above were "those authors who in major publications actually constructed bridges among various fields." Other authors that made major contributions to the knowledge base of the Modern Synthesis include S. S. Chetverikov, R. A. Fisher, T. H. Morgan, J. B. S. Haldane, C. D. Darlington, E. B. Ford, F. Sumner, A. H. Sturtevant and S. Wright.

predecessors and extended the understanding of evolution by natural selection to include the concepts of mutation, recombination, genetic drift, adaptation, gene flow and speciation. The main tenets of the Modern Synthesis (Mayr, 1982; Futuyma, 1998) are:

1 Genetic variation in a population arises *de novo* through random mutation.

2 Variation is inherited in discrete units (genes), but can be re-organised through recombination and sex.

3 Populations evolve by changes in gene frequency brought about by random genetic drift, gene flow (migration) and natural selection.

4 Speciation arises from a gradual accumulation of genetic change and can ultimately give rise to differences that warrant the designation of higher taxonomic levels.

Since the formation of the Modern Synthesis, most debate about evolutionary theory has focussed on the degree to which each of these forces influence the diversity and the convergence observed in nature, with the core of the original theory remaining remarkably intact. The prevailing view of evolutionary theory is that natural selection is the most powerful of the forces and acts on the variation within populations to drive the appearance of adaptations. One important aspect of natural selection is that it ensures that the majority of biological features are adaptations, or have been adaptations at some point in their history (Maynard Smith, 1982). Some features may be exaptations (Gould and Vrba, 1982), some may be vestigial, some may be the result of chance (often called 'frozen accidents'), some may be by-products of other adaptations and a minority may indeed be the result of random processes such as drift. However, conventional understanding about the tempo and mode of evolution suggests that most features are (or have been) adaptations and that chance and history play smaller, but significant roles (Fisher, 1930; Wright, 1945; Williams, 1966; Mayr, 1982; Maynard Smith *et al.*, 1985; Williams, 1992). Therefore, it is not only justified, but the primary task of evolutionary biologists to engage in adaptationist reasoning (Dennett, 1995) to confirm either that a feature is adapted for its current function (an adaptation), that a feature was adapted for one function and has subsequently been co-

opted for another function (exaptation), that a feature was adapted for one function and is no-longer needed (vestigial), that a feature's function is orthogonal with respect to natural selection (by-products and chance) or that a feature has an unknown function.

The Modern Synthesis crystallised the principles of evolution by natural selection at the level of the individual; however, life on Earth is hierarchically organised, with the entities at higher levels composed of lower-level entities (Table 1-1). The generality of evolutionary theory to apply to all levels of the biological hierarchy was recognised by Fisher, but simultaneously dismissed as unimportant (Fisher, 1958). However, the importance of selection at multiple levels of the biological hierarchy has become increasingly appreciated. Lewontin (1970 p.1) writes that "any entities in nature that have variation, reproduction, and heritability may evolve." Many of the hierarchically organised levels of biological systems satisfy these three conditions (Gould, 2002), hence, there is potential for selection to operate on entities at multiple levels simultaneously[1].

Table 1-1: The levels of the biological hierarchy.

	Description
Ecosystems	Includes all living organisms and abiotic elements
Communities	All species in an ecosystem
Species	All individuals of a single species
Populations	All individuals of a single population
Organisms	Single individual
Organ system	Specialised functional system (*e.g.* nervous system)
Organ	Specialised structural system (*e.g.* brain)
Tissue	Specialised substructure of an organ (*e.g.* nervous tissue)
Cell	Individual cell (*e.g.* bacteria, yeast, nerve cell)
Molecule	Biological molecules (*e.g.* DNA)

This table is a summary of the hierarchies in nature adapted from Campbell and Reece (2005).

A multi-level perspective implies a diachronic view that acknowledges two important aspects of evolution. Firstly, selection can act at multiple levels of the biological hierarchy simultaneously and secondly, the hierarchically organised levels observed

[1] The generality of evolution to act on entities that have variation, reproduction, and heritability has notable parallels with linguistics (reviewed in Atkinson and Gray, 2005) and has been extended to among others, memes (Dawkins, 1976) and even galaxies (Smolin, 1992).

in nature are themselves the product of evolution (Okasha, 2006a). One important consequence of a diachronic view is that there is potential for the existence of conflict between higher and lower levels, because natural selection may operate in different directions at each level. Examples of conflict between levels of selection include meiotic drive (Cosmides and Tooby, 1981), mitochondria induced degenerative conditions (Taylor *et al.*, 2002), germ-line sequestration (Michod and Roze, 2001), within- and between-colony conflicts in social insects (Tarpy *et al.*, 2004) and cancer (Nowell, 1976). Moreover, before conflicts can exist between levels, a higher level must itself have evolved as the product of evolution. Expansions of the biological hierarchy to include higher levels have been termed the major evolutionary transitions. Two concepts are particularly important in understanding these transitions: first, the cooperation among the lower-level units, and second, reproduction of higher-level entities as individuals in their own right (Buss, 1987). The first of these, cooperation, and its importance in the expansion of the biological hierarchy begins in the next section.

1.2 COOPERATION

Cooperation is interesting to evolutionary biologists for two different, but closely related reasons. Firstly, in the orthodox paradigm of adaptationism, it is difficult to explain how cooperation can be both costly to an individual and an adaptation. Secondly, cooperation among lower-level units drives the expansion of the biological hierarchy (*i.e.* genes coming together to form chromosomes or the first cells coming together to form multicellular individuals). Before difficulties associated with cooperation and the importance of the expansion of the biological hierarchy can be discussed in more detail, the history of cooperation in evolutionary biology is reviewed (Section 1.2.1) and an explicit working definition of cooperation is established for use in this thesis (Section 1.2.2). Then, the existing explanations of cooperation are contextualised in a multi-level selection framework with a novel approach using the natural framework provided by the biological hierarchy to address the issue of how the cost of cooperation to an individual is offset (Section 1.2.3). Finally, this is integrated into a fully diachronic view of evolution provided by the multi-level perspective (Section 1.2.4).

1.2.1 A BRIEF HISTORY OF COOPERATION

Numerous behaviours exist in all kingdoms of life that enhance the fitness of other individuals both within a single species and between different species. However, with the acceptance of evolutionary theory, such cooperative behaviours seem paradoxical, because the cost imposed by the act of cooperating renders cooperative individuals less fit relative to non-cooperators, and natural selection favours types that are relatively more fit than others (Wilson, D.S., 1975; Wilson, 2008). Darwin recognised the need for his theory to account for cooperative behaviours in *The Descent of Man, and Selection in Relation to Sex* (1871 p.166) with this passage of how cooperative behaviours might evolve via selection among groups:

> *There can be no doubt that a tribe including many members who, from possessing in a high degree the spirit of patriotism, fidelity, obedience, courage, and sympathy, were always ready to aid one another, and to sacrifice themselves for the common good, would be victorious over most other tribes; and this would be natural selection.*

Similarly, Wright (1945 p.417) concluded, "It is indeed difficult to see how socially advantageous, but individually disadvantageous mutation can be fixed without some form of inter-group selection." Despite this, few early twentieth century authors appreciated the paradox posed by cooperation, because casual attitudes towards evolutionary theory saw cooperative behaviours routinely assigned as group-level adaptations. This uncritical 'for the good of the group' thinking masked the true complexity of the problem that cooperation posed for evolutionary theory, because cooperation was always considered adaptive at the group level.

The 1960s saw 'for the good of the group' thinking come to an abrupt halt, due largely to the work of Hamilton (1963; 1964a; 1964b), Maynard Smith (1964) and Williams (1966). Their work was a response to, what is now the most frequently criticised publication with regard to 'for the good of the group' thinking, *Animal Dispersion in Relation to Social Behaviour* by V.C. Wynne-Edwards (1962). Wynne-Edwards observed that many species of bird do not lay as many eggs as they are capable of physiologically laying. He argued that this decreased reproduction would not be predicted by individual selection, but that it would be predicted by group selection. He proposed that the observed regulation of population size was a property of the group that existed to maximise the use of the available resources in a habitat (*i.e.* a group-level adaptation). Therefore, groups in which individuals did not practise reproductive restraint had become extinct with their habitats colonised by offshoots from groups that did contain such individuals.

It is not difficult to imagine a scenario in which this group-level selection argument would break down. For example, any bird that gained a genetic mutation that made it unconcerned by the existing conventions of regulation would reproduce without discern. Hamilton (1963; 1964b; 1964a), Maynard Smith (1964) and Williams (1966) lead the charge of protagonists who argued for the primacy of individual-level selection in determining the success of traits, a view that coloured evolutionary biology in the 1960s. They viewed selection as intrinsically selfish and acting more strongly at the level of the individual than at the level of the group, because individuals were more numerous, possessed more variation and reproduced more

rapidly than groups. In support of his conjecture, Hamilton developed the theory of inclusive fitness, which was effectively relabelled as kin selection by Maynard Smith[1] (Section 1.2.3.1). These theories are often taught and understood from what is termed the 'gene's eye' view. This outlook pioneered by Williams (1966) and popularised by Dawkins in *The Selfish Gene*[2] (1976) focuses on the gene as the unit of selection and is often a helpful heuristic for looking at the problem of cooperation. The gene's eye view focuses on what a gene might 'want' and one way for a gene to maximise the number of copies of itself found in the next generation is to cause its host organism to behave cooperatively towards other bearers of the same gene. Therefore, while cooperation may seem paradoxical when looked at from the individual organism's point of view, if the behaviour is directed towards relatives it makes perfect sense when looked at from the gene's point of view. Most people find the gene's eye view approach to kin selection heuristically simpler than the inclusive fitness approach, but they are mathematically equivalent (Michod, 1982).

While kin selection proved a powerful explanation for selection of cooperative interactions among related individuals, cooperative interactions between individuals of different species could not be explained by this approach (by definition). However, one theory, termed reciprocal altruism (Trivers, 1971), emerged as another alternative to the out-of-favour group selection. This approach has been used to study numerous interactions (not only cooperation/defection) in nature with great success (Section 1.2.3.2).

Despite these major contributions to our understanding of how cooperation might

---

[1] The usage of inclusive fitness and kin selection is often synonymous in the literature, however, kin selection is a process, and inclusive fitness is the property maximised by kin selection therefore, it is considered most appropriate to say that kin selection is the process by which organisms maximise their inclusive fitness (West *et al.*, 2007b).

[2] Dawkins clarifies a misinterpretation of his title in the introduction to the 30th anniversary edition of *The Selfish Gene*. "The correct word in the title to stress is *gene*" which clears up a common misconception about the content of the book, which actually "devotes more attention to altruism" (Dawkins, 2006 p.vii).

evolve, group selection continued to be maintained as a plausible theory. The 1970s saw the re-invigoration of group selection with a theoretical model of interdemic group selection (also termed trait-group selection, Wilson, D.S., 1975)[1] and empirical studies of the flour beetle *Triboleum casteneum* (Wade, 1977). Wade's experiments illustrated a response to selection when selecting on the properties of groups, while Wilson's interdemic group selection model explicitly showed how differential productivity of groups could account for the lower individual fitnesses within the groups (Wilson, 1983). However, Wilson's model also altered the definition of group selection, leading to widespread misunderstanding (an issue dealt with in Section 1.2.3.5). Progressively, group-level selection has developed into the more encapsulating multi-level selection theory, (MLS, Heisler and Damuth, 1987; Damuth and Heisler, 1988; Sober and Wilson, 1998; Okasha, 2006a), which successfully delineates between different types of group selection (Section 1.2.4). There is some disagreement between authors about the history surrounding the transition of 'old group selection' (*sensu* Wynne-Edwards) to 'new group selection' (the interdemic group selection models of Wilson) and on the further transition to MLS (Heisler and Damuth, 1987; Damuth and Heisler, 1988; Sober and Wilson, 1998; Okasha, 2006a; West *et al.*, 2007b; West *et al.*, 2008; Wilson, 2008). However, MLS provides the framework for extending group selection to a fully diachronic view of evolution (Okasha, 2006a), where multiple levels of the biological hierarchy are simultaneously under selection.

1.2.2 Definitions of Cooperation

From an evolutionary perspective, social behaviours are those that have fitness consequences for the acting individual and another individual (West *et al.*, 2007b). Therefore, it is common to examine such behaviours by deconstructing the fitness for each individual separately to describe the behaviour more precisely (Sachs *et al.*, 2004). Hamilton (1964a) classified social behaviours according to the consequences the interaction entailed for the *absolute fitness* of the actor and the recipient

[1] D.S. Wilson's initials are included in his 1975 citation to avoid confusion with E.O. Wilson's publication in the same year that is also cited in this thesis.

(actor/recipient). For example, in altruism the actor has a decrease in absolute fitness, while the recipient has an increase, (Figure 1-1). This relationship can be described as –/+, where the first symbol represents the sign of the fitness change to the actor, and the second represents the sign of the fitness change to the recipient(s), as a result of their interaction through the behaviour of the actor. In contrast, mutually beneficial interactions are +/+, selfish interactions are +/–, and spiteful interactions are –/– (Hamilton, 1964a; Hamilton, 1970; Maynard Smith, 1982; West *et al.*, 2007b). However, cooperative interactions only require that the actor is less fit *relative* to a non-actor and it is not implicit in cooperation that the absolute fitness of the actor should increase or decrease as a consequence of their actions. Therefore cooperative behaviours include both –/+ (altruistic) and +/+ (mutually beneficial) interactions where the actor has a lower fitness relative to a non-actor (Axelrod and Hamilton, 1981; Frank, 1995; Maynard Smith and Szathmáry, 1995; Sachs *et al.*, 2004; West *et al.*, 2007b).

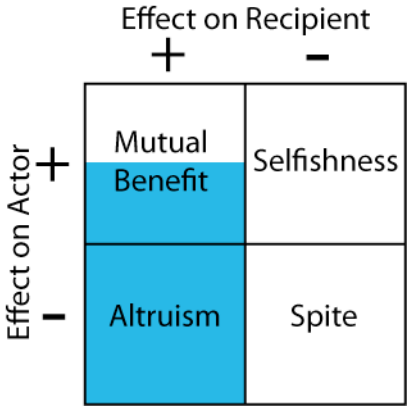

Figure 1-1: Classification of social interactions in terms of absolute fitness.
The light-blue shaded area represents those interactions that are considered cooperative. All cases of altruism are cooperation, but only those cases of mutual benefit in which the actor has a relative fitness cost are considered cooperative.

Cooperation is most modestly defined as an action by one individual that benefits one or more other individuals (Sachs *et al.*, 2004). The individual performing the action is termed the actor and any individual receiving the benefit is termed the recipient (actors may also be recipients). However, this interpretation requires further refinement, because without modification, it can be applicable to a myriad of biological phenomena that are not examples of cooperation at all (West *et al.*, 2006; West *et al.*, 2007b). West *et al.* (2007b p.429) emphatically point this out with a simple example:

> *...when an elephant produces dung, this is beneficial to the elephant (emptying waste), but also beneficial to a dung beetle that comes along and uses that dung. It does not seem useful to term behaviours such as this, which provide a one-way by-product benefit, as cooperation. Consequently, we prefer that a behaviour is only classed as cooperation if that behaviour is selected for because of its beneficial effect on the recipient.*

The point illustrated here is that the utilisation of elephant dung by a dung beetle has no adaptive significance for the elephant, *i.e.* elephant defecation *is not an adaptation for* increasing the fitness of dung beetles, but simply a by-product of elephant metabolism. In this example, the effect of elephant dung on the fitness of dung beetles is indiscriminate with respect to the fitness of the elephant (West *et al.*, 2007b p.429), because there is no fitness feedback from the dung beetle (recipient) to the elephant (actor). The issue at stake here is one of adaptation: identifying a behaviour as cooperative implies that it is an adaptation for cooperation, because the paradox associated with cooperative behaviours does not exist if the behaviour is an adaptation for another function or if the real paradox of cooperation was associated with a different behaviour. If it is determined that a behaviour is adapted for some other function, or if it only appears cooperative because of its association with another behaviour that is cooperative, this would not be an example of cooperation (West *et al.*, 2007b). Thus, cooperation is defined as a trait *selected to enhance* the fitness of others at a relative cost to the acting individual (West *et al.*, 2006) and can only be a target of selection when there is a fitness feedback from the recipient to the actor.

General characteristics for identifying cooperative behaviours have never been clearly stated in the literature, presumably because cooperation comes in many different varieties that have subtly different expectations. However, here I present a summary of four expectations that are commonly used as sufficient criteria to identify a behaviour as cooperative:

1 There is an interaction between two or more individuals, one of whom is the actor and the other the recipient (actors may also be recipients).

2 There are relative fitness consequences for both the actor and the recipient, *i.e.* a fitness cost for the actor relative to a defector (non-actor) and a fitness benefit to the recipient.

3 The fitness of the actor(s) is negatively impacted by the number of defectors and positively affected by the number of actors.

4 The number of defectors (non-actors) capable of being sustained by a population is proportional to the number of actors.

These criteria are typically used as a checklist to determine whether a behaviour satisfies the requirements to be identified as cooperative. The first criterion is a general requirement of all social behaviours (interactions) that is satisfied when an action has fitness consequences for another individual (not specific to cooperation). The second criterion requires demonstration of reduced fitness to actors and increased fitness to recipients or a relative fitness cost of actors to non-actors (defectors[1]) as a result of the actor/recipient interaction. This fitness cost can be demonstrated by measuring the fitness of actors relative to defectors in an environment where acting is not required (*e.g.* the benefit is artificially added to the environment by the researcher). A benefit to recipients can be demonstrated by increasing the number of actors (it does not matter whether or not actors are also recipients) and the fitness of recipients with fewer actors will be lower than the fitness of recipients with more actors. The third and fourth criteria are more precise demonstrations of the dynamical relationship between the number of actors and the number of defectors. Demonstrating the appropriate dynamics between actors and defectors is typically considered sufficient to identify a behaviour as cooperative. However, it has been suggested that demonstrating these dynamics may not unambiguously lead to the conclusion that the behaviour is cooperative, as it includes by-products of self-interested behaviours and examples of diffusion sensing (Redfield, 2002) or a response to a cue or chemical manipulation (Keller and Surette, 2006; and see also the discussion in Section 4.3.2; Meintjes and Rainey, manuscript in preparation).

[1] Defectors are individuals that do not perform the cooperative act, but differ from cheaters in that they do not necessarily benefit from the cooperative actions of others. Therefore, biologically not all defectors are cheaters, but all cheaters are defectors (Velicer, 2003).

1.2.2.1 BY-PRODUCTS

By-products have become increasingly emphasized as a type of cooperation (Sachs *et al.*, 2004). For example, a one-way by-product benefit that a recipient, Y, receives from an actor, X, is described as "an automatic consequence of the otherwise selfish act in which individual X does something to benefit itself" (Sachs *et al.*, 2004 p.145). However, to identify one-way by-product benefit as a type of cooperation is misleading, because there is no feedback from the recipient to the actor, thus the effect on the recipient has no adaptive significance for the actor (Figure 1-2). When the absolute fitness effects of the actor and recipient are deconstructed separately, it is clear that the absolute fitness effects are mutually beneficial, +/+. However, one-way by-product benefit describes a situation identical to that of the elephant and the dung beetle quoted earlier in this section that was rejected precisely because it was non-adaptive. This has been clearly illustrated by extending the same example in Sachs *et al.* (2004 p.145):

> *The faeces from large ungulates are food for dung beetles; vultures and carrion-feeding insects benefit from abandoned lion kills. Following Connor (1995), there has been* ***no evolution of cooperation*** **per se** *in these cases. That is, lion behavior has not been evolutionarily modified to benefit vultures or other carrion feeders, and vulture behavior has not evolved to increase the chance of a kill. Whatever evolves in the case of this byproducts model****, it is not selected to offer a cooperative act****. In byproduct models, there is* ***no potential Darwinian dilemma****, because the basic cooperative trait directly benefits its bearer and only incidentally benefits others*[1].

It is difficult to illustrate the point more clearly, yet the conclusion of these authors is that one-way by-product benefit constitutes a meaningful type of cooperation. While agreeing that they are interesting in their own right, one-way by-product benefits do not constitute cooperation, because by definition, they arose to perform a function that is not that of cooperation, *i.e.* it arose to increase the actor's fitness. West *et al.* (2007b p.419) clarify this position, by focusing on the fitness effect on the recipient:

> *Consequently, we prefer that a behaviour is only classed as cooperation if that behaviour is selected for because of its beneficial effect on the recipient. We do not wish to imply that the behaviour is selected for purely because of its beneficial effect on the recipient****, just that it has at least partially done so***[1].

[1] Boldface emphasis added.

Identifying one-way by-product benefits as a type of cooperation fails to distinguish cause from consequence, because the behaviour of X cannot be *selected to enhance* the benefit of individual Y, because there is no feedback from Y to X. However, if it can be demonstrated that a by-product of X has in part been selected for its beneficial effect on individual Y, due to a feedback of Y on X, then the behaviour will be considered cooperative (West *et al.*, 2006).

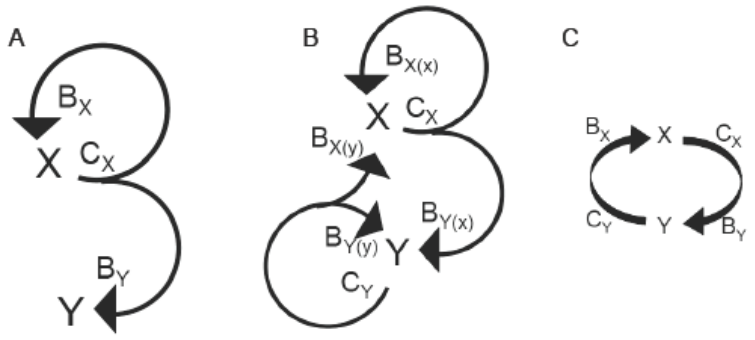

Figure 1-2: Types of by-products.
Three types of by-products have been identified by Sachs *et al.* (2004) for two individuals X and Y. A) describes a one-way by-product benefit with a cost only to X (C_X), but with the opportunity for benefit to both X and Y (B_X and B_Y). B) describes two-way by-product benefit in which X receives a benefit from itself $B_{X(x)}$ and from Y ($B_{X(y)}$). C) describes by-product reciprocity.

Considering the diagram (Figure 1-2), feedback from Y to X can occur in one of two ways, that both reduce to the principle of reciprocity (Section 1.2.3.2). For example, in two-way by-product benefit, X provides an indirect benefit to Y ($B_{Y(x)}$), but also receives an indirect benefit from Y ($B_{X(y)}$). Although both benefits are produced indirectly as by-products of self-interested actions, this feedback loop provides potential for adaptive significance for the interaction that may lead to the coupling of fitness observed in by-product reciprocity. In by-product reciprocity, individual Y evolves to enhance its benefit to X, which in turn increases the by-products it receives from X. The by-product from X does not necessarily evolve, but the effect of Y on X does (Connor, 1986; Sachs *et al.*, 2004; West *et al.*, 2006). In this example the fitnesses of X and Y have become coupled, a feature of reciprocal cooperation described in Section 1.2.3.2.

1.2.2.2 Public Goods

The related issue of 'public goods' is vulnerable to similar criticisms as by-products, because the definition fails to recognise the need for a trait to have been *selected to*

enhance the fitness of others. Public goods are defined as "any fitness-enhancing resource that is accessible to multiple individuals within a local group" (Velicer, 2003 p.330). In this paradigm, individuals within the group that produce the public good are commonly viewed as cooperators, because they incur a metabolic cost for producing, and individuals that do not produce but maintain the ability to utilise the public good are said to 'sequester unfairly', so are considered cheats. However, this paradigm is not sufficient to consider all public goods an *a priori* example of cooperation, because the benefit to others may be a by-product and not an adaptation. As noted above, evidence of a fitness feedback from the recipient to the actor needs to be determined, *e.g.* a by-product from X is only considered a type of cooperation if the fitness of X is increased by a reciprocal by-product from Y or some other reciprocal action of Y. Crucially, in public goods, all actors are by definition also recipients, because the public good is equally available to all individuals in the vicinity. Therefore, the effect of the actor on the recipient is intrinsically linked to the fitness of the actor. There are two consequences of this view. First, that it is possible for selection to favour cooperation, because a fitness feedback exists, and second, that the effect on other recipients may simply be "an automatic consequence of the otherwise selfish act in which individual X does something to benefit itself" *i.e.* a by-product.

Distinguishing between these two effects is difficult, but important, because only if the behaviour has been *selected to enhance* the fitness of others should it be considered cooperation. If actors are also recipients, how can one determine that the benefit to others was at least in part selected for? One expectation is that if selection has acted to enhance the fitness of others, then the fitness benefit of the trait will increase when there are higher numbers of other individuals, *i.e.* there will be an emergent density-dependent fitness/phenotype that is quantitatively (emergent fitness) or qualitatively (emergent phenotype[1]) 'greater than a sum of the parts' (Vrba, 1983; Corning, 2002; Gould, 2002; Velicer, 2003; Greig and Travisano, 2004;

[1] The emergent phenotype are best case scenarios of emergent fitnesses in which emergence can be detected under the emergent character criterion (Vrba, 1983).

Griffin *et al.*, 2004; West *et al.*, 2006). Density dependent fitness and emergence from cooperative interactions are examined empirically for the wrinkly spreader of *P. fluorescens* in Chapter 4.

A second expectation comes from examining the following inequality:

$$\frac{b}{n} - c < 0 \qquad \text{Eq 1-1}$$

where b is the shared benefit produced by an actor, n is the total number of recipients (actors and non-actors) in the vicinity and c is the metabolic cost. Whenever this inequality is satisfied, the model of public goods is an example of cooperation, because defection always has the higher payoff. The expectation from this inequality is that if the mechanisms for offsetting this cost are removed (*i.e.* benefits no longer directed at relatives) then the non-actors should drive the actors extinct and then the non-actors should themselves go extinct, a tragedy of the commons (Hardin, 1968). Importantly, this expectation is testable, but has never been shown for any bacterial examples of a public good.

The view advanced in this thesis is that to assert *a priori* that the 'public goods' paradigm is an example of cooperation is to ignore the need to show that the effect on recipients must have adaptive significance for the actor (*i.e.* feedback from the recipient to the actor), and is to uncritically assign the cooperative behaviour at the level of the group *sensu* Wynne-Edwards. Therefore, it is imperative to distinguish between situations of 'public goods' and situations in which 'public goods' meet the requirements of cooperation. In addition, if it is necessary to use loaded social terms such as cheater and cooperator, this should be restricted to situations in which cooperation has been verified.

1.2.2.3 Summary – Cooperation and Microbes

In this thesis, the term cooperation is used as previously defined: any action *selected to enhance* the fitness of others at a relative cost to the acting individual as identified by the four criteria outlined earlier in this section. One-way by-product benefits that are commonly confused as a type of cooperation are not considered an example of cooperation, because there is no adaptive significance for the actor. Two-way by-

product benefits, by-product reciprocity and public goods all have the potential to be cooperative, but this should be tested empirically and not assumed *a priori*.

Cooperation among microbes always involves the extracellular secretion of compounds, because this is the sole mechanism of social interaction among bacteria. Therefore, if cooperation is identified in microbes by the criteria described above, the scenario is either one of public goods (all actors are also recipients, and the benefit of the action is equally available to both actors and recipients) or by-product reciprocity (two actions are linked by a fitness feedback loop). This thesis examines a niche-specific public goods[1] example for *Pseudomonas fluorescens* in which the costly secretion of cellulose is offset by the emergent group-level benefit of being able to occupy a novel niche at the air-liquid interface (Rainey and Rainey, 2003). The precise nature of this cooperation is dissected further in Chapter 4.

1.2.3 Explanations of Cooperation

Cooperation has been one of the most difficult phenomena requiring explanation in evolutionary biology for the greater part of the last 100 years (Wright, 1945; Hamilton, 1963; Maynard Smith, 1964; Trivers, 1971; Maynard Smith and Szathmáry, 1995; Frank, 1998; Sachs *et al.*, 2004). The difficulty arises because behaviours that solely increase the fitness of other individuals cannot be tolerated by natural selection. Could these behaviours simply be the result of an imperfect match between organisms and their environment? This suggestion is interesting because organisms are never perfectly adapted to their environment, but implies that natural selection is unable to efficiently remove a large number of maladaptive traits that are expected to be eradicated in the long run (Sterelny, 2001). Therefore, the continued existence of cooperative behaviours implies that the overwhelming majority of these behaviours have adaptive significance (Maynard Smith *et al.*, 1985).

Various methods have been proposed to account for the manifest evidence of

[1] It will become clear in the description of wrinkly spreader cooperation that the niche-specific public good that is available to all individuals at the air-liquid interface, but not to individuals in the broth phase.

cooperation (Hamilton, 1963; Maynard Smith, 1964; Trivers, 1971; Wilson, D.S., 1975; Axelrod and Hamilton, 1981; Sober and Wilson, 1998; Kerr *et al.*, 2004; Sachs *et al.*, 2004; Lehmann and Keller, 2006; Nowak, 2006; West *et al.*, 2007b). The central question that each method is designed to answer is essentially always the same: how can individually costly behaviours continue to exist? In many cases, different types of explanation are required for the different types of cooperation that exist, but in all cases, a mechanism for offsetting the cost of the cooperation must be identified.

This thesis aims to synthesise existing explanations for cooperation in a novel perspective within the natural framework provided by the biological hierarchy. Within a multi-level framework there are three principal ways in which costs to cooperative behaviours are offset: at a level below the individual (*e.g.* genes); at the level of the individual (*e.g.* reciprocal interactions between individuals); or at the level above the individual (*e.g.* groups, species or populations). These three categories provide the framework for discussing the explanations of cooperation through the next three subsections.

1.2.3.1 SHARED GENES

Explanations for cooperation that fall into the category of shared genes have flourished in the literature under many names. Selfish gene theory, inclusive fitness and kin selection can all be reduced to the concept of shared genes. Inclusive fitness theory was formalised by Hamilton (1963) to describe how, if cooperative actions are directed towards relatives, an individual increases his/her own inclusive fitness because other copies of the gene in his/her relatives receive the benefit of the cooperative action. This concept is elegantly defined by the inequality $rb - c > 0$, where r is the coefficient of relatedness, b is the benefit associated with the cooperative act, and c is the cost to the cooperating individual, with costs and benefits measured in terms of reproductive success. Kin selection (Maynard Smith, 1964) was the re-branding of inclusive fitness theory as a type of selection that favoured the evolution of characteristics that enhance the survival of close relatives. Although kin selection is technically a slightly restricted version of inclusive fitness, because kin selection requires identity by descent, in practice, the concepts are usually synonymous, because it is largely groups of related individuals that fit the model

(Sachs *et al.*, 2004). The predictions from Hamilton's rule are that cooperation is more likely to be directed towards relatives, and that cooperation will be favoured by natural selection whenever the inequality is satisfied. From the organism's point of view, although an altruistic behaviour reduces the organism's direct fitness (by definition), it increases its inclusive fitness through indirect effects on relatives. The explanatory power of kin selection/inclusive fitness theory has contributed overwhelmingly to our understanding of self-sacrificial behaviour (Axelrod and Hamilton, 1981), parental investment (Trivers, 1972), as well as sterility and sex ratios in social insects (Sundstrom *et al.*, 1996; Queller and Strassmann, 1998).

Sachs *et al.* (2004) further delineate this shared gene framework to illustrate a difference between kin fidelity and kin choice. Both kin fidelity and kin choice require the inequality, $rb - c > 0$ to be true before they will evolve, but each is subtly different mechanistically. Kin fidelity relies on the natural spatial distribution of related individuals (*e.g.* related birds being in the same nest). Cooperative actions can be directed towards relatives simply by directing cooperative actions towards neighbours, because on average, individuals in close proximity are more likely to be relatives (the raising of young birds in nests is a common example). This is considered a passive strategy for directing actions towards kin, because there is no identification of kin *per se*. Alternatively, kin choice directs benefits to related individuals, not by their serendipitous spatial distribution, but by active recognition of the individual based on their phenotype. Evidence for this is common in birds where individuals can be identified by characteristic plumage (Wilson, E.O., 1975; Beecher, 1982; Whitfield, 1986; Whitfield, 1987; Dale *et al.*, 2001) or even vocal contact cues, which young learn from adults during the nesting period (Sharp *et al.*, 2005). Alternatively, it has been hypothesised that kin choice might also be achieved through the 'greenbeard' effect (Dawkins, 1976), which requires a single gene (or a number of tightly linked genes) to cause the cooperative behaviour and to promote recognition by individuals. In addition, the hypothesis of the greenbeard effect will also hold between non-related organisms that share the gene(s) of the greenbeard locus.

1.2.3.2 RECIPROCITY

Many examples of cooperation in biology can be explained by kin selection, because the benefits are directed towards kin. However, there are also many examples where

individuals enhance the fitness of another individual from a completely different species. Therefore, kin selection is unable to account for all types of cooperation. Reciprocal altruism[1] (Trivers, 1971) is a mechanism that offsets the cost of cooperation at the same level where the cost is incurred (*i.e.* the level of the individual). Unlike in kin selection, there is no requirement for the other individual to be related, and this type of directed reciprocity can evolve if there are repeated interactions or an individual has the ability to vary its behaviour in each interaction according to a partner's previous action. If these two requirements are met, the interaction between the two individuals can usually be analysed within the framework of evolutionary game theory (Lewontin, 1961) using the iterated Prisoner's Dilemma (IPD, Trivers, 1971). One of the most successful strategies in evolutionary game theory is 'tit-for-tat' where individuals respond to their partner with the same action (Axelrod and Hamilton, 1981). Many empirical examples of cooperation have been analysed using the IPD (Axelrod and Hamilton, 1981), however, while there is no doubt that the IPD describes conditions that favour cooperation, few biological examples are thought to adequately satisfy its assumptions (Sachs *et al.*, 2004).

Sachs *et al.* (2004) suggest distinguishing between types of directed reciprocity along the lines of the two requirements of the IPD. The first, termed partner fidelity feedback (Sachs *et al.*, 2004), requires an extended series of repeated interactions between any pair of individuals, such that the fitnesses of two individuals X and Y become coupled (*e.g.* mitochondria-host cell endosymbioses, Rand *et al.*, 2004). If Y cooperates with X, the fitness of X increases, which in turn increases the benefit Y receives from X. Conversely, if Y does not cooperate, then the fitness of X will decrease and the return benefit to Y will decrease as well. The feedback is automatic, and therefore, does not require a conditional response. The second, termed partner choice (Sachs *et al.*, 2004), requires that each individual has the ability to vary its behaviour between cooperation and defection using a decision-rule to choose more

[1] It is commonly noted that reciprocal altruism is poorly named, because it isn't *real altruism*, but rather an enlightened form of self-interest. However, this criticism misses the point, because such behaviours cannot evolve in the absence of mechanisms to identify one's 'debtors', 'collectors' or cheats (Dennett, 1995).

cooperative partners (*e.g.* legume-rhizobium symbioses, Simms and Taylor, 2002). Partner choice may be either direct, by choosing to cooperate with individuals that reciprocate, or indirect by choosing to cooperate with individuals that cooperate with others (West *et al.*, 2007a). The indirect effect of individual X choosing the more cooperative partner Y_1 over Y_2, is that Y_1 will also be more fit within its own population, so the evolution of cooperation is also promoted within population Y_n (Sachs *et al.*, 2004). In this example, repeated interaction is not required, but when permitted, increases the propensity of cooperation to evolve. Choice of partners by individuals can drive the assortative interactions of cooperators leading to increased between-group variance and potentially allows selection to favour some groups over others. Selection of this nature (*i.e.* between groups) is dealt with in the next section.

1.2.3.3 GROUP SELECTION

Another proposed mechanism for understanding the evolution of cooperative behaviours is selection at the level of groups. Group selection is the process of genetic change caused by the differential proliferation and extinction of groups of organisms (Wright, 1945; Wynne-Edwards, 1962; Maynard Smith, 1964; Williams, 1966; Lewontin, 1970; Maynard Smith, 1976; Wade, 1977). The explicit requirement for proliferation and extinction of groups is argued to be fundamental to group selection because "in a finite universe multiplication [proliferation] implies death… in the long run, group selection requires group death just as evolution by individual selection requires individual death" (Maynard Smith, 1976 p.279). For many years the consensus was that the conditions for the continued existence of cooperation in these models were too stringent to be realistic (Grafen, 1984), and were almost universally agreed to be unable to explain ecological adaptations (Wright, 1945; Maynard Smith, 1964; Williams, 1966; Maynard Smith, 1976).

The revival of group selection in the 1970s redefined groups to include only those individuals that interacted with one another and not all individuals within the deme (Wilson, D.S., 1975). This eliminated the need to define boundaries for groups, because only interacting individuals were considered. Importantly, differential extinction of groups was also relaxed as a criterion for defining this interdemic or 'trait-group' selection (Wilson, D.S., 1975). Interestingly, the first of these models had been mentioned by Wright (1945), but had gone largely unnoticed. Wright

described groups in which cooperators provide a benefit for everyone in the group (including themselves), but pay only a personal cost (a relative cost compared to other individuals within the group). The single locus model for gene 'A', describes the fitness for three types of individuals: AA (homozygous cooperator), Aa (heterozygous) and aa (homozygous defector) as follows:

$$\omega_{AA} = (1 + pb)(1 - 2s) \quad \text{Eq 1-2}$$

$$\omega_{Aa} = (1 + pb)(1 - s) \quad \text{Eq 1-3}$$

$$\omega_{aa} = (1 + pb) \quad \text{Eq 1-4}$$

where p is the proportion of cooperators, b is the benefit to the group and s is the cost to the individual. It is easy to see that AA has the lowest fitness within the group, but gene A can increase its absolute fitness whenever b is large in relation to s. However, Wilson argues that the distinction between types of cooperation in which absolute fitness increases and types in which it does not, is largely irrelevant, because natural selection is based on relative fitness (Wilson, D.S., 1975; Wilson, 2008). Therefore, the relative cost associated with cooperating must be offset for gene A to continue to evolve. This comes from the differences in the frequency of A between the groups, such that groups with the most cooperators contribute more individuals to the total population. This type of group selection treats the structure of the population as part of the environment of an individual, and as such explains only the changing frequency of an individual-level behaviour (*e.g.* a cooperative trait) in the overall population (Okasha, 2006a). The primary conclusion of this theoretical work is that viability differences between groups based on the genetic composition within the groups can lead to a form of interdemic group selection in which individuals are the units of selection and groups are considered part of the environment (Wilson, D.S., 1975; Damuth and Heisler, 1988; Okasha, 2006a).

1.2.3.4 Empirical Evidence of Group Selection

The first evidence from empirical studies involving selection at the level of groups came from an experiment on the flour beetle *Tribolium casteneum* (Wade, 1977; Wade, 1978). The purpose of this study was to investigate the genetic effects of

differential extinction and re-colonisation of populations. Four treatments (A-D) that differed in their selection regime were allowed to evolve for eight generations. Treatment A selected the group of beetles that had produced the most offspring after 37 days of incubation to found the next generation of beetles. Treatment B selected the group of beetles that had produced the fewest offspring after 37 days of incubation. Treatment C was a control treatment that selected at the level of individuals only, and treatment D selected groups randomly to found the next generation. Both treatments C and D were expected to have no selection with respect to groups. Each selective regime was repeated a further eight times before the results were analysed.

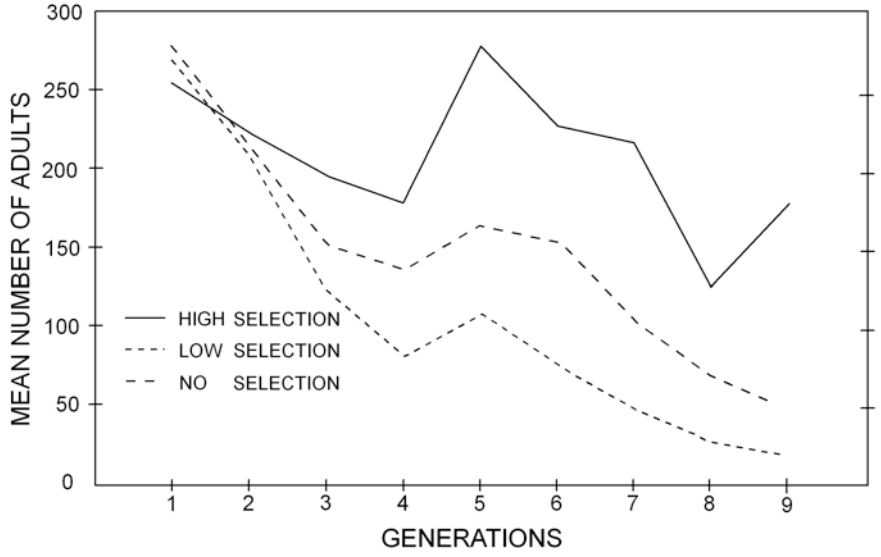

Figure 1-3: The mean number of adult *T. casteneum* from three group selection treatments.
The graph shows the divergence of the treatment A (solid line) towards higher numbers of adults and the divergence of treatment B (shortest dashed line) towards lower numbers of adults from the control treatment C (dashed line). The experiment was terminated after nine group-generations because it reached the lower limit of the experimental design. Treatment D was not represented because it was statistically identical to treatment C. The figure has been retraced from the original Figure 2 in Wade (1977).

The results showed a divergence between treatments for the differential extinction and re-colonisation of groups. Treatment A gave rise to groups that produced the highest number of adults, treatment B gave rise to groups with the lowest numbers of adults, and treatments C and D (controls) produced intermediate numbers of adults. The primary conclusion from this study is that it is experimentally possible to observe a response to selection at the level of groups.

1.2.3.5 CRITICISM OF GROUP SELECTION

There are four major criticisms with which proponents of group selection are typically confronted. Firstly, group selection before the 1960s is criticised for its naivety, because adaptations were routinely assigned at the level of the group, *e.g.* Wynne-Edwards (1962). This has been criticised as 'for the good of the group' thinking (Maynard Smith, 1964; Williams, 1966; Dawkins, 1976; Grafen, 1984) often under the title of 'old group selection' (West *et al.*, 2007b; West *et al.*, 2008). Secondly, the redefinition of groups to include only those individuals that interact with one another (Wilson, D.S., 1975; Sober and Wilson, 1998) which has become known as 'new group selection' (West *et al.*, 2007b; West *et al.*, 2008) has also drawn criticism. This approach leads to the misunderstanding that 'within-group selection' is the same as 'individual selection' and that 'between-group selection' is 'group selection' (Grafen, 1984). However, individual selection is defined as selection favouring the spread of a trait by its effect on the actor's number of offspring alone, which is incompatible with defining it based on selection within the group only (Grafen, 1984). A further consequence is that altruism becomes redefined as only costly within a group, but not at the population level. Since altruism was originally defined at the population level, if a trait is only locally altruistic and not at the level of the population, the paradox of reduced fitness as a consequence the action no longer exists, and thus an explanation involving altruism for the evolution of the trait is no longer required (Grafen, 1984). Thirdly, some models do not conform to the original definition of group selection (Wilson, D.S., 1975), because they make no explicit statement about the differential proliferation and extinction of groups of organisms (Maynard Smith, 1964; Maynard Smith, 1976). Okasha (2006a p.57) notes that, "Wilson's model is designed to explain the changing frequency of an individual trait in the overall population and although the explanation makes essential appeal to group structure, and treats groups as fitness-bearing entities, it permits no inference about the frequency of different types of group." Wade's experimental approach also comes under fire from this criticism, because ultimately individuals, and not groups, are the units of selection. Selecting groups with the highest/lowest reproduction of individuals is confounded by the fact that the highest/lowest reproducing individuals are always selected: because they are nested within the highest/lowest reproducing groups, and the groups themselves do not reproduce. Finally, using Hamilton's model

of local mate competition (Hamilton 1967), the mathematical equivalency of interdemic group selection with kin selection has been proven (Maynard Smith, 1976; Frank, 1986; West *et al.*, 2008) and it is often claimed that this has lead to unnecessary confusion at the expense of insight (West *et al.*, 2007b; West *et al.*, 2008).

Despite these criticisms, multi-level selection theory (Heisler and Damuth, 1987) has emerged to unite the group selection literature by distinguishing between types of group selection in which groups are part of the environment of individuals and group selection in which groups are units of selection. These are the topic of the next section.

1.2.4 MULTI-LEVEL SELECTION THEORY

Multi-level selection acknowledges two important aspects of evolution. Firstly, selection can act at multiple levels of the biological hierarchy simultaneously, because many of those levels qualify as units of selection under the definition of interactors (Hull, 1980) and secondly, the hierarchically organised levels observed in nature are themselves the product of evolution (Okasha, 2006a). The replicator/interactor distinction is important for understanding what constitutes a unit of selection. Replicators are units of which copies are made (*i.e.* genes) and interactors are units that interact directly with the environment as a cohesive whole, in such a way that replication is differential. Evolution by natural selection is "a process in which the differential extinction and proliferation of interactors cause the differential perpetuation of the replicators that produced them" (Hull, 1980 p.318). Therefore, many levels of the biological hierarchy participate meaningfully in the process of evolution by natural selection as interactors, *e.g.* cells, organisms and species.

Acknowledging that selection can act simultaneously at multiple levels also acknowledges the potential for conflict between higher and lower levels, because adjacent levels may interact in the full range of conceivable ways – synergistically, orthogonally or in opposition (Gould, 2002). For example, under certain conditions, cooperation among lower-level units can lead to group-level benefit, and is thus favoured by selection at the level of groups. However, because cooperation is costly,

defection is favoured by selection at the level of the individual. Therefore, although cooperation is fundamental to the emergence of higher levels of the biological hierarchy, it is not sufficient, because types that do not cooperate (*i.e.* defectors) have a fitness advantage over types that do cooperate at the lower level. Consistent with theoretical predictions, as defectors increase in frequency, they undermine the benefit obtained from being part of the group and the higher-level unit is lost. Therefore, although cooperation provides the drive for the emergence of higher-level entities, it is not sufficient to explain the evolution and maintenance of the higher-level entity (Michod, 1999).

The next three subsections identify and contrast two distinct types of multi-level selection to show how selection can act at multiple levels of the biological hierarchy simultaneously. A fourth subsection deals with emergence and the expansion of the biological hierarchy.

1.2.4.1 MULTI-LEVEL SELECTION 1

Multi-level selection 1 (MLS 1) occurs "whenever an individual's expected viability, mating success, and/or fertility cannot be accounted for solely on the basis of that individual's phenotype, but rather additional information is required about properties of the group or groups of which the individual is a member" (Heisler and Damuth, 1987 p.584). This definition clearly illustrates how MLS 1 is focused on addressing questions that relate to the evolution of phenotypic characteristics of *individuals*. This definition does not require that groups have differential probabilities of giving rise to new groups, or of becoming extinct, but does allow for such group-level processes to affect the fitness of individuals (Heisler and Damuth, 1987). Therefore, most theoretical work on 'group selection', as discussed in Section 1.2.3.3, would be included in MLS 1 including not only models that explicitly include differential extinction of entire groups (Maynard Smith, 1964; Levins, 1970; Gilpin, 1975; Maynard Smith, 1976), but also models that do not (Wright, 1945; Wilson, D.S., 1975).

1.2.4.2 MULTI-LEVEL SELECTION 2

Multi-level selection 2 (MLS 2) occurs "whenever any group properties co-vary with group-level fitness, implying that the proportions of different kinds of *groups* will

change in the population (and noting that group characters may change as a result of lower-level selection among the individuals that the groups comprise)" (Heisler and Damuth, 1987 p.584). It is clear from this definition that MLS 2 requires that groups are units of selection (*i.e.* that they have the properties of variation, heredity and reproduction). Most models of species-level selection would be included in MLS 2 (Vrba, 1984; Eldredge, 1985; Gould, 2002).

1.2.4.3 USING MLS 1 AND MLS 2

The MLS 1 and MLS 2 distinction is a powerful, but largely unappreciated framework in the literature. MLS 1 and MLS 2 are commonly referred to as a single concept, multi-level selection theory, but this phrase under-utilises the value of the two concepts that are grouped by that term. Understanding the distinction, between the MLS 1 and MLS 2 is the principal value of multi-level selection theory, because it removes the dogmatic associations and confusion that continue to surround the ambiguous use of the term group selection (West *et al.*, 2007b; West *et al.*, 2008; Wilson, 2008), and provides a clear philosophically supported framework in which to analyse selection at multiple levels of a hierarchy (Heisler and Damuth, 1987; Damuth and Heisler, 1988; Griesemer, 2001; Okasha, 2006a). Failure to distinguish between MLS 1 and MLS 2 has generated needless controversy within biological sub-disciplines and impeded unification of the multi-level selection traditions (Damuth and Heisler, 1988). The principal differences between MLS 1 and MLS 2 are outlined in Table 1-2. Much of MLS 1 has been couched in the language of MLS 2 particularly in the 'group selection' literature that relates more specifically to interdemic group selection (Wilson, D.S., 1975). The crucial difference is that MLS 1 focuses on the fitnesses of *individuals* organised into groups, while MLS 2 focuses on the fitnesses of *groups* composed of individuals. Some authors argue that examples of MLS 1 do not meet the requirements of group selection, because the groups are not units of selection in their own right (Maynard Smith, 1976). Instead, while acknowledging that there is a "group effect on gene or organismal fitness" (Gould, 2002 p.653), the traditional definition of selection should be used to define selection at the level of groups, and that has "always been based on causal plurification and not mere effect" (Gould, 2002 p.653).

Table 1-2: Conceptual differences between MLS 1 and MLS 2.

	MLS 1	MLS 2
Description	Populations consist of individuals, organised into groups	Populations consist of groups, composed of individuals.
Fitness	Properties of individuals	Properties of groups
Group Selection	Effects of group membership on individual fitnesses	Change in the frequencies of different groups
Characters	Values attributed to individuals (including both individual and contextual characters)	Values attributed to groups (including aggregate and emergent characters)

This table summarises the main points developed by Damuth and Heisler (1988 p.410). Individual characters are measured directly on individuals while contextual characters are measured on groups, but assigned to individuals. Aggregate characters are assigned to groups but are a combination of measurements on individuals while emergent characters are not obtainable by measuring individuals (Lazarsfeld and Menzel, 1961; Damuth and Heisler, 1988).

Okasha (2006b) argues that taking Maynard Smith's view "has the unwelcome consequence that most models of group selection do not deal with real group selection." This is a very weak position, because if Maynard Smith is correct, it would be wise to disregard MLS 1 as a type of group selection, indeed good scientific practice should *welcome* it. However, a second and more substantial point is that simply because MLS 1 incorporates groups into the environment and then focuses on individuals as the units of selection, does not mean that it can not tell us anything about groups, because "*groups* are composed of *individuals*, explaining the evolution of an *individual* character could help explain salient features of *groups* too." (Okasha, 2006a p.59)[1]. Including MLS 1 as part of a pluralistic approach for understanding groups may be especially helpful when trying to understand what properties of individuals are important for their ability to cooperate or police their selfish tendencies. Therefore, abandoning the confusing group selection labels and distinguishing between MLS 1 and MLS 2 may be sufficient to remedy the misunderstanding surrounding 'interdemic group selection' and 'differential proliferation and extinction' of groups. Once it is agreed that higher-level selection of the MLS 2 form is defined by differential proliferation of higher-level units based on

[1] Italicised words are substituted from the original for consistency in this thesis. Collectives has been changed to *groups* and particles has been changed to *individuals*. In this context, the words are synonymous.

interaction between their traits and the environment, clear criteria for the definition and recognition of higher-level traits must be developed.

1.2.4.4 EMERGENCE AND THE EXPANSION OF THE BIOLOGICAL HIERARCHY

An essential feature of the biological hierarchy that exists today is that the hierarchy itself is also the product of evolution (Okasha, 2006b). Therefore, a complete evolutionary theory will need to explain how new, higher units of selection arise and not simply take their existence for granted. Higher order entities arise and participate meaningfully in evolution when they are units of selection. To become units of selection in their own right, the higher level requires emergent[1] properties that are 'not otherwise attainable' (Corning, 2002), because the effect is qualitatively or quantitatively 'greater than the sum of the parts' (Gould, 2002). Emergent properties are the consequence of lower-level interactions and do not exist at the lower level. They are observed as either emergent fitnesses (Lloyd, 1988) or emergent characters (Vrba, 1983). Emergent characters are characters of groups originating from "non-additive interaction among lower-level units", while emergent fitnesses are any trait that "characterises or influences the differential rate of proliferation in interaction with the environment" (Gould, 2002 p.657 and 659).

The strength of the emergent character approach lies in its ability to clearly identify emergent properties. If a trait does not exist at the lower level, or cannot even be represented by a combination of lower-level units, then the trait is unambiguously emergent under the emergent character approach. However, if a trait can be represented by a combination of lower-level units (*e.g.* variability) and it can impact

---

[1] The term emergent was coined by G. H. Lewes who wrote: "Every resultant is either a sum or a difference of the co-operant forces; their sum, when their directions are the same -- their difference, when their directions are contrary. Further, every resultant is clearly traceable in its components, because these are homogeneous and commensurable. It is otherwise with *emergents*, when, instead of adding measurable motion to measurable motion, or things of one kind to other individuals of their kind, there is a co-operation of things of unlike kinds. The *emergent* is unlike its components insofar as these are incommensurable, and it cannot be reduced to their sum or their difference." (Lewes, 1875 p.412)

the differential survival and proliferation of groups, then this would be an emergent property of the group, but would only qualify under the definition of emergent fitness and not emergent character. Importantly, using the emergent fitness approach to define emergence does not disprove the emergent character approach, because emergent characters are a subset of best cases of emergent fitness (Gould, 2002). Therefore, the current orthodox view is to define emergent properties based on the minimum criterion of emergent fitness.

Expansion of the levels of the biological hierarchy by the emergence of fitness and characteristics at the higher level, have been called 'changes in the level of complexity' (Maynard Smith, 1988), 'transitions in the units of fitness' (Michod, 1999), 'evolutionary transitions in individuality' (Buss, 1987; Michod and Roze, 1997) and the 'major evolutionary transitions' (Maynard Smith and Szathmáry, 1995). Examples of such transitions include genes coming together to form chromosomes, the endosymbiosis of mitochondria within eukaryotic cells, the evolution of multicellular organisms from free-living single cells, and the evolution of groups, species and societies.

The major evolutionary transitions share two common themes: the emergence of cooperation among lower-level units in the functioning of the new higher-level unit; and regulation of conflict among the lower-level units. Cooperation among the lower-level units drive the emergence of higher-level entities, but this is not sufficient to maintain the higher-level entity, because it is expected that among lower-level units, types that do not cooperate would have a fitness advantage over types that do. As these types increase in frequency, they would undermine the benefit obtained from being part of the group and the higher-level unit would be lost. Therefore, although cooperation is necessary for the emergence of higher levels of biological organisation, it is not sufficient, because regulation of the conflict among the lower-level units will also be required to maintain the individuality of the higher-level entity. The challenge remains for evolutionary theory to explain *why* lower-level selection did not disrupt the initial formation of the higher-level entities (Buss, 1987; Maynard Smith and Szathmáry, 1995; Michod, 1999).

1.3 THE POWER OF MICROBIAL MODEL SYSTEMS

Darwin's formulation of the theory of evolution by natural selection was based on observations of macroorganisms. Since then, most hypotheses and empirical evidence about the mode of evolution have continued to focus on larger organisms; however, there are various restrictions to studying evolution in macroorganisms that leave certain experiments intractable. Microbial systems provide the opportunity to complement the insights developed by macroevolutionists through direct testing of hypotheses of the mode of evolution in the laboratory (Dykhuizen, 1990; Lenski, 1992). Microbial systems are ideal for studying evolutionary processes because bacteria grow rapidly under laboratory conditions and quickly reach large population sizes, due to their short mean generation time. Microbes (usually) reproduce asexually, which affords the researcher clonal replicate genotypes for parallel experimentation. Bacterial cultures can be stored at -80˚C in a state of suspended animation almost indefinitely, which allows comparisons among derived lines or between derived and ancestral lines. By carefully controlling the environmental conditions of replicate clonal bacterial populations, experimenters are able to distinguish between chance and repeatable outcomes, and thoroughly test the mechanistic processes underlying evolution.

There are a number of other reasons why bacteria are a suitable model system for evolutionary experimentation. For example, this thesis uses strain SBW25-*lacZ*, a strain marked at the defective prophage region, attTn7. This marked strain has been determined to have no measurable effect on fitness during 24 h of incubation for multiple substrates, *in planta* (Zhang and Rainey, 2007), and in long-term experiments (Appendix 9.1.1). Neutrally marked strains are central to experimental evolution, because they provide the ability to simply and reliably measure the relative fitness of specific genotypes by calculating either the ratio or difference of Malthusian parameters between the marked and unmarked competitor strains (Lenski *et al.*, 1991; Travisano *et al.*, 1995b). In this regard SBW25-*lacZ* has already proved a powerful marker in published experimental evolution studies (Fukami *et al.*, 2007). In addition, there is extensive knowledge of bacterial physiology that enables researchers to generate testable hypotheses for protein interaction to determine detailed molecular pathways or networks. Finally, advancements in computer

software have boosted analysis of genetic data with powerful visualisation software, the ability to predict function from sequence based on homology, and the rapid identification of mutations, primer sites or transposon insertion sites. Because of their simplicity and effectiveness, such techniques are employed routinely in the laboratory on microbes with high efficiency. For these reasons, bacterial populations are a major driving force in the current explosion of research into experimental evolution.

1.4 THE *P. FLUORESCENS* EXPERIMENTAL SYSTEM

Pseudomonas fluorescens is an aerobic, Gram-negative, rod-shaped, saprophytic, γ-proteo-bacterium commonly found in the rhizosphere – the area of soil and water surrounding and including the plant root (Palleroni, 1984). Its name is derived from the fluorescent protein pyoverdine (formerly fluorescein), a type of siderophore (iron scavenging protein) that it produces in low-iron environments. *P. fluorescens* encompasses a genetically and metabolically diverse group of bacteria capable of colonising many environments. These organisms are important in biotechnology because they can enhance plant growth and protect against disease (O'Sullivan and O'Gara, 1992). *P. fluorescens* are free-living and motile through the use of multiple flagella (Palleroni, 1984), but have also been extensively studied for their ability to form biofilms (Costerton *et al.*, 1995; Rainey and Travisano, 1998; Drenkard and Ausubel, 2002; Rainey and Rainey, 2003; Spiers and Rainey, 2005).

In this thesis I use SBW25, a strain of *P. fluorescens* isolated from field-grown sugar beet at University Farm, Wytham, Oxford in 1989 for which the physical map, genetic map (Rainey and Bailey, 1996) and full genome sequence are available (Sanger Institute, 2007). The collaborative effort that has gone in to providing this resource has proved extremely valuable for much of the work in this thesis.

1.4.1 DIVERSIFICATION IN A MICROCOSM

The wild-type or ancestral strain of *P. fluorescens* SBW25 is termed smooth (SM) based on a description of its colony morphology when grown on agar plates. When incubated statically in laboratory culture of King's Medium B (KB, King *et al.*, 1954) *P. fluorescens* SBW25 undergoes a rapid adaptive radiation to form a variety of distinct morphotypes (Rainey and Travisano, 1998). This adaptive radiation is fuelled by the ecological opportunity afforded by spatial structure and a strong vertical oxygen gradient that develops rapidly within the medium (Rainey and Travisano, 1998). There are three major derived classes of colony morphotype the wrinkly spreader (WS), the fuzzy spreader (FS) and those that resemble the ancestor, also SM (Rainey and Travisano, 1998; Rainey and Rainey, 2003). In each case the phenotype is heritable, and various analyses have shown that the phenotype correlates with both

genotype and niche specificity (Rainey and Travisano, 1998; Rainey *et al.*, 2000; Spiers *et al.*, 2000; Spiers *et al.*, 2002; Rainey and Rainey, 2003; Gehrig, 2005; Bantinaki *et al.*, 2007). In addition, the phenotypic variation within WS colony morphology provides some indication of the locus at which the mutation is present (Section 2.2.2). Visual differences among colony types allow rapid identification of the major morphotypes (Figure 1-4).

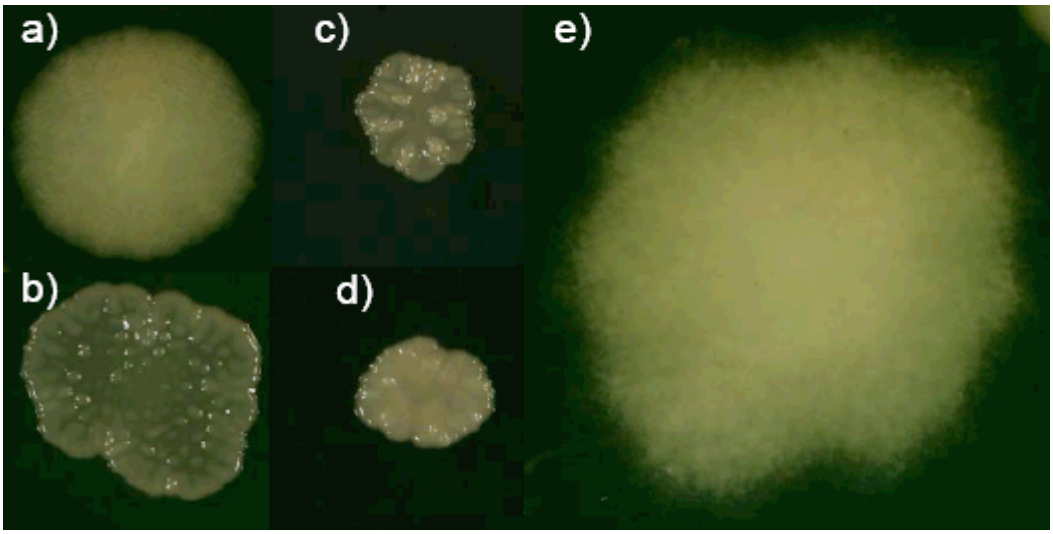

Figure 1-4: The major morphotypes of *P. fluorescens* on KB agar.
a) The ancestral or smooth (SM) type. b) A wrinkly spreader (WS) with a mutation in the *wsp* operon. c) A WS with a mutation in *aws* operon. d) A WS with a mutation in the *mws* gene. e) A fuzzy spreader (FS). All photographs taken with an eight second exposure at F11 with a 65mm macro. Photos a)-d) taken with a 4x lens and e) taken with a 2x lens and scaled.

1.4.1.1 WS MAT FORMATION IS COOPERATIVE

WS types are a diverse group of genotypes so named because of their characteristic wrinkly morphology on agar plates (Rainey and Travisano, 1998). Many independent single mutations have been shown to cause the WS colony morphology (Bantinaki *et al.*, 2007). However, although WS types are named for their colony morphology, the defining biological characteristic of WS types is that they exhibit specificity for the niche at the air-liquid interface suggesting that all WS genotypes share a consistent ecology (Rainey and Travisano, 1998; Rainey and Rainey, 2003). When incubated statically, WS cells adhere to one another and to the glass wall of the microcosm (Figure 1-5) due to the over production of bacterial cellulose and a proteinaceous adhesin (Spiers *et al.*, 2003). These groups of WS cells at the air-liquid interface are commonly referred to as mats, and as the mats become established, oxygen rapidly becomes depleted in the nutrient rich KB media (Figure 1-6). WS cells within the mat

are able to occupy a niche that SM cells cannot, thus there is a niche-specific advantage for WS genotypes, because they obtain the benefits of exposure to oxygen at the air-liquid interface.

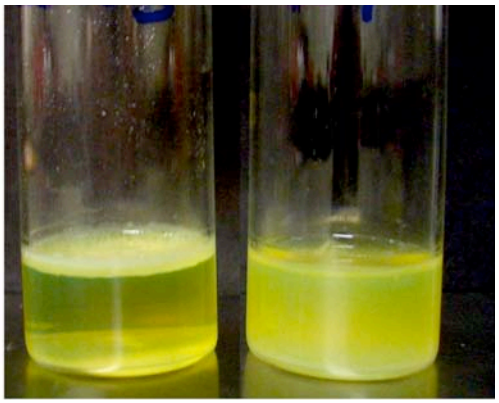

Figure 1-5: Niche specificity of the wrinkly spreader (WS) and the ancestral smooth (SM).
A) WS grows at the air-liquid interface. B) SM grows in the broth phase. The fluorescent pigment is stronger at the air-liquid interface. Photos were taken for LSWS and SBW25 after 60 h of static incubation at 28˚C.

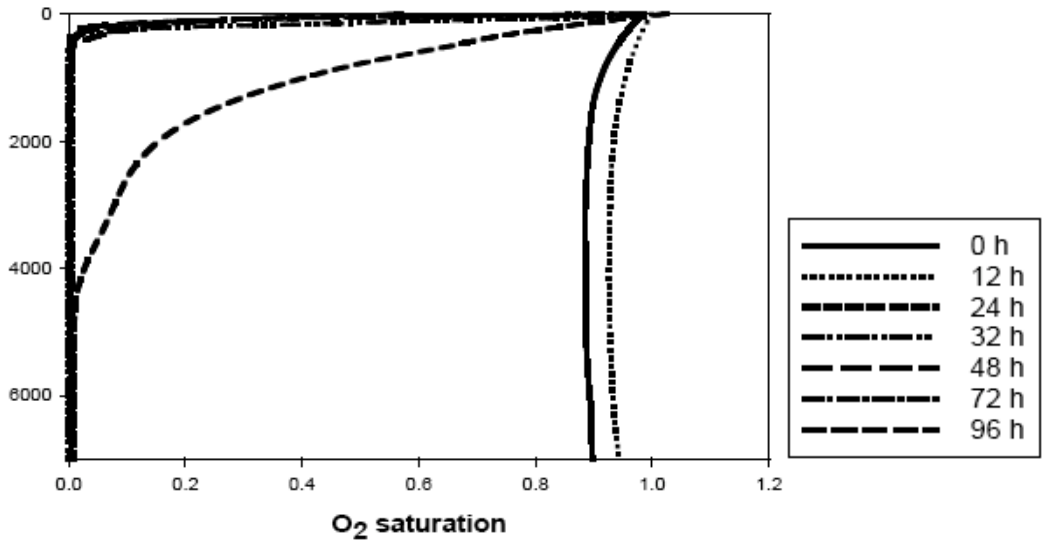

Figure 1-6: Depletion of oxygen in a statically incubated microcosm.
Oxygen saturation was measured at multiple depths for multiple time points. Each time point is plotted on axes of oxygen saturation and depth below the air-liquid interface (units are mm). The figure and the data were created by Bas Ibelings (EAWAG, Zurich).

Crucially, WS mat formation meets the criteria of a cooperative trait. The evidence for this was established by Rainey and Rainey (2003) for one particular WS genotype, the large spreading wrinkly spreader (LSWS, isogenic with the ancestor except for a point mutation in *wspF*). Firstly, they established that there was a cost to

cooperation, because LSWS had a relative fitness of 0.8[1] (95% CI: 0.69 – 0.91) compared to ancestral SBW25 when incubated statically with abundant resources (Rainey and Rainey, 2003). Secondly, this cost is offset by benefits at the group level (*i.e.* the mat), because WS types are able to increase in frequency and become the dominant morph by occupying a novel niche at the air-liquid interface. Thirdly, defecting ancestral-like SM types arise and increase in frequency in microcosms founded by WS types. These SM types are fitter in the presence of WS types than in their absence. Finally, the emergent group-level phenotype is destroyed (*i.e.* the mat collapses) if the number of defecting SM types increases to too high a frequency (Figure 1-7). Therefore, WS cells obtain a group-level benefit (exposure to oxygen) from being part of the mat despite an intrinsic metabolic cost to individual WS cells of synthesising bacterial cellulose (Rainey and Rainey, 2003).

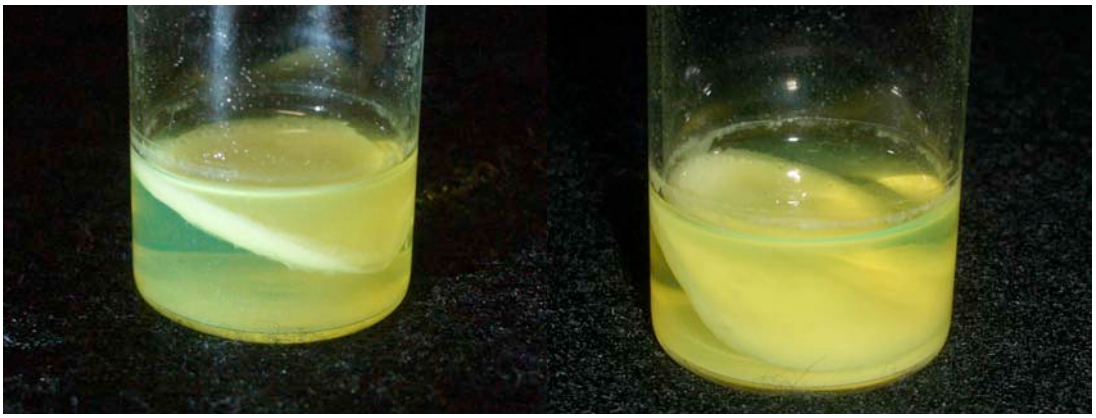

Figure 1-7: Partially collapsed mats.
Both pictures show mats formed from 60 h growth in microcosms inoculated with LSWS and were shaken gently to detach the mats from the sides of the glass microcosm.

1.4.2 The Independent Wrinkly Spreaders

It was previously reported that a single non-synonymous point mutation in *wspF* was responsible for the LSWS phenotype (Gehrig, 2005). However, the distribution of possible mutations that gives rise to the wider diversity of WS types was unknown. To investigate the extent of genetic and phenotypic diversity of the WS types, 26 independent wrinkly spreaders (IWS) were collected after incubation in separate microcosms for seven days (Bantinaki *et al.*, 2007). Each of the 26 IWS was

[1] This has been verified by other work included in this thesis, (Appendix 9.1.2.1).

sequenced at the *wspF* locus to determine how many contained an independent *wspF* mutation or to determine if mutations at other loci were involved in generating the WS phenotype. Mutations in *wspF* were identified in 13 of the 26 IWS (Bantinaki *et al.*, 2007). To identify other genes involved in the WS phenotype, a strain with the entire *wsp* operon deleted (SBW25-Δ*wsp*) was engineered so that WS types with *wsp* mutations were impossible. This strain was inoculated into a microcosm and incubated statically to allow WS types that did not have a mutation in *wsp* to evolve. WS types still evolved from the SBW25-Δ*wsp* strain, indicating that at least one other locus could give rise to WS via a single mutation. A transposon mutagenesis screen identified the alternate wrinkly spreader locus (*aws*, Section 1.5.3) and when the *aws* locus was sequenced, a mutation was found in *awsX* (Gehrig, 2005). Subsequently, *awsX* was sequenced for each of the remaining 13 IWS genotypes, but only one of those, IWS_T, was shown to have a mutation at that locus. As a result of work in this thesis, all but one of the IWS genotypes (IWS_X) have now been identified (Section 2.2).

1.5 *P. FLUORESCENS* EVOLUTIONARY GENETICS

The evolutionary genetics of *P. fluorescens* is the subject of ongoing investigation. Of particular relevance to this thesis is the biosynthesis of bacterial cellulose, the ultimate cause of the WS phenotype (Spiers *et al.*, 2003). The *wss* locus has been identified as the primary cellulose biosynthesis operon and, due to results of previous and current work in the Rainey lab, it is now known that three other loci (*wsp, aws* and *mws*) have been implicated in regulation of the WS phenotype by post-transcriptionally activating Wss enzymes.

1.5.1 THE *WSS* LOCUS

The *wss* locus is a 15.6 kb cluster of ten genes primarily involved in the biosynthesis of bacterial cellulose (Gal *et al.*, 2003; Spiers *et al.*, 2003). The *wss* operon is unique in its gene composition (Figure 1-8), and to date no other bacterium has been reported to possess a cellulose biosynthetic operon as complex as that of *P. fluorescens* SBW25 (Gehrig, 2005). Table 1-3 summarises the relevant information in Spiers *et al.* (2002) showing the role of each of the genes. The flanking regions *wssA* and *wssJ* each contain a MinD domain (commonly required for the formation of a septum at the mid-cell) and show homology with Soj ATPases (commonly involved in chromosome partitioning) detected using the conserved domain database (CDD, NCBI Conserved Domain Search, 2007) on the NCBI website (Marchler-Bauer *et al.*, 2005). From this information, it is predicted that WssA and WssJ are membrane-associated cell-cycle ATPases involved in ensuring cell-division at the mid-cell through non-specific inhibition of the septum protein FtsZ (Spiers *et al.*, 2002).

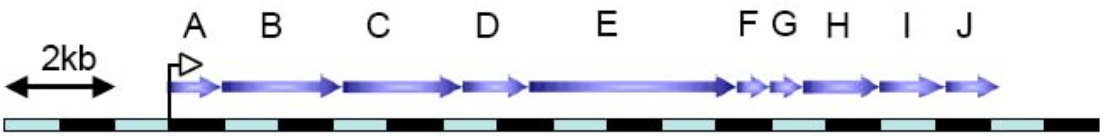

Figure 1-8: The ten genes of the *wss* locus.
wssA-J make up the *wss* locus. Their predicted functions are described in Table 1-3.

Four *wss* genes (*wssB, wssC, wssD* and *wssE*) encode proteins with significant amino acid identity to the known *Gluconacetobacter xylinus* (formerly *Acetobacter xylinum*) proteins BcsA (75%), BcsB (65%), CMCase (64%) and BcsC (75%) respectively (Spiers *et al.*, 2002). These proteins have been identified as being involved in the

biosynthesis of bacterial cellulose and act to polymerise UDP-glucose (Wong *et al.*, 1990). Using the CDD, WssF is identified as a serine hydrolase and is predicted by homology with BcsX to act as an acetyl-transferase (Spiers *et al.*, 2003). *WssG*, *wssH* and *wssI* encode proteins with significant amino acid identity to known *P. aeruginosa* proteins AlgF (63%), AlgI (70%) and AlgJ (67%) respectively. In *P. aeruginosa*, these proteins have been identified as playing a role in the acetylation of the bacterial cellulose polymer.

Table 1-3: Predicted function of the proteins in the *wss* operon.

Gene	Predicted Function	CDD Match	Homology	Genetic Analysis
wssA	Spatial localisation	MinD	YhjQ (*E. coli)*	(Spiers *et al.*, 2002)
wssB	Cellulose synthase subunit	Cellulose Synthase	BcsA (*G. Xylinus*) YhjO (*E. coli*)	(Spiers *et al.*, 2002)
wssC	Cellulose synthase subunit	BcsB	BcsB (*G. Xylinus*) YhjN (*E. coli*)	(Spiers *et al.*, 2002)
wssD	Endogluconase	Glycosyl Hydrolase	CMCase (*G. xylinus*) YhjM (*E. coli*)	(Spiers *et al.*, 2002)
wssE	Cellulose synthase subunit	BcsC	BcsC (*G.xylinus*) YhjL (*E. coli*)	(Spiers *et al.*, 2002)
wssF	Serine hydrolase	SGNH hydrolase	BcsX (*G. xylinus*)	n/a
wssG	Polymer modification	n/a	AlgF (*P.aeruginosa*)	n/a
wssH	Polymer modification	DltB	AlgI (*P. aeruginosa*)	(Spiers *et al.*, 2002)
wssI	Polymer modification	n/a	AlgJ (*P. aeruginosa*)	n/a
wssJ	Spatial localisation	MinD	YhjQ (*E. coli*)	(Spiers *et al.*, 2002)

CDD matching (Marchler-Bauer *et al.*, 2005) was done at the NCBI website (NCBI Conserved Domain Search, 2007). Homology to known proteins is a summary of the data presented in Spiers *et al.* (2002). In some cases, mutation studies have already been done to determine the effect of each of the genes adding to the evidence for the predicted role of the proteins encoded by each of these genes.

Studies of the regulation of bacterial cellulose synthesis in *G. xylinus* showed that bis-(3'-5'-)-cyclic dimeric guanosine monophosphate (c-di-GMP) is a positive allosteric activator of cellulose synthase (Ross *et al.*, 1990). Similarly, the Wss proteins are positively affected by the presence of c-di-GMP (Goymer *et al.*, 2006; Bantinaki *et al.*, 2007; Malone *et al.*, 2007). Each of the three identified regulatory loci (*wsp*, *aws* and *mws*) negatively regulate di-guanlyate cyclases (McDonald *et al.*, 2008). Mutations in each of these loci cause the constitutive over-production of c-di-GMP, leading to the over-stimulation of Wss proteins. The activity of the Wss proteins, in conjunction with the production of a proteinaceous adhesin (Spiers *et al.*, 2003), causes daughter cells to remain attached to one another after cell division. This cell-to-cell adhesion is required for the emergent group-level phenotype, *i.e.* mat formation at the air-liquid interface (Spiers *et al.*, 2002; Rainey and Rainey, 2003).

1.5.2 THE *WSP* LOCUS

The *wsp* locus is an 8.4 kb cluster of seven genes that is involved in the regulation of cellulose biosynthesis and the WS phenotype (Spiers *et al.*, 2002; Goymer *et al.*, 2006; Bantinaki *et al.*, 2007). Figure 1-9 shows the arrangement of the seven genes that form a single transcriptional unit (Bantinaki *et al.*, 2007) and the functions of their protein products are summarised in Table 1-4.

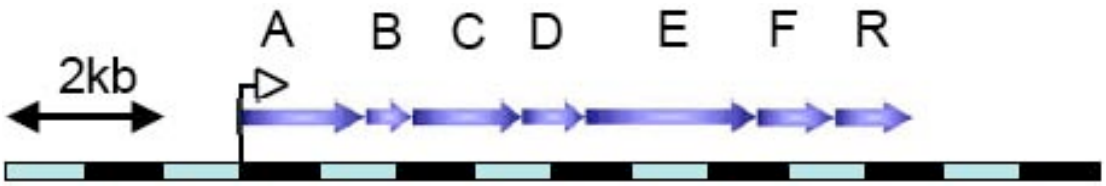

Figure 1-9: The seven genes of the *wsp* locus.
WspA-R are all transcribed as a single unit.

WspABDE of the *wsp* locus form a membrane-bound receptor-signalling complex with significant similarity to the thoroughly characterised chemosensory pathway (Che pathway) described for *E. coli* (Amsler and Matsumura, 1995; Bantinaki *et al.*, 2007). WspA is a methyl-accepting chemotaxis protein (MCP). MCPs are membrane-spanning proteins that sense chemical gradients in the environment with an N-terminal, periplasmic ligand-binding domain (Falke *et al.*, 1997). They bind to their target chemical inducing a conformational change in the C-terminal cytoplasmic domain. WspB and WspD are scaffold proteins similar to CheW from *E. coli* (Levit *et al.*, 2002). The last of the complex is WspE, which has four conserved domains (NCBI Conserved Domain Search, 2007). The first three domains are a histidine-phosphotransfer domain, a histidine kinase domain and a CheA regulatory domain. Collectively these have homology with all of CheA. The final domain of WspE is a REC domain that has homology with CheY response regulator. Interestingly, this combination of domains from CheA and CheY is observed in the FrzE protein of *Myxococcus xanthus* (McCleary and Zusman, 1990b; McCleary and Zusman, 1990a).

The cytoplasmic proteins of the *wsp* locus that do not form part of the membrane-bound receptor-signalling complex are WspC, WspF and WspR. WspC contains a conserved N-terminal CheR domain that has methyltransferase activity (Djordjevic and Stock, 1997). CDD matching (NCBI Conserved Domain Search, 2007) also revealed that WspC contains a tetratrico peptide repeat region downstream of the

CheR domain in an arrangement that is similar to that of the FrzF protein of *M. xanthus* (McCleary *et al.*, 1990). WspF has two predicted domains: the N-terminus has a CheY response regulator domain and a C-terminal CheB methylesterase domain with overall similarity to CheB from *E. coli* (Bantinaki *et al.*, 2007). WspR is a diguanylate cyclase (DGC) response regulator that has been described in detail (Goymer *et al.*, 2006). It has two domains: an N-terminal CheY response regulator domain and a GGDEF domain that confers its DGC activity upon phosphorylation of Asp67 (Goymer *et al.*, 2006).

Table 1-4: Predicted function of proteins in the *wsp* locus.

Gene	Genbank Acession#	Predicted Function	CDD Match	Che Homology (*E. coli*)
wspA	AA092333	Methyl-accepting chemotaxis protein (MCP)	Tar (MCP)	None
wspB	AA092334	Scaffold protein	CheW	CheW
wspC	AA092335	Methyl transferase	CheR	CheR
wspD	AA092336	Scaffold protein	CheW	CheW
wspE	AA092337	Histidine kinase	HPT, HATPase, CheA and REC (CheY)	CheA and CheY
wspF	AA092338	Response regulator methylesterase	CheB	CheB
wspR	AAl71852	Post-translational response regulator	GGDEF and REC (CheY)	CheY

The genes of the *wsp* operon are all transcribed as a single transcriptional unit. CDD matching was performed at the NCBI website (NCBI Conserved Domain Search, 2007).

1.5.2.1 THE WSP PATHWAY

WspABCDEF govern a signal-dependent pathway of the response regulator WspR (WS phenotype response regulator) for an as yet unidentified stimulus (Figure 1-10). Under normal conditions, a stimulus acting on WspA causes a conformational change in WspA that triggers the autophosphorylation of WspE. WspE, which is attached to WspA in part by the scaffold proteins WspB and WspD, phosphorylates WspF, which negatively regulates WspR by removing methanol groups from WspA. WspC works in opposition to WspF by positively regulating WspR by adding methyl groups to WspA. Ultimately WspR regulates cellulose production at a post-transcriptional level by overproduction of c-di-GMP (Bohannon 2002, Goymer 2006). Mutations affecting the operation of the Wsp pathway are predominantly found in *wspF*. WspF is a negative regulator, so mutations in *wspF* lead to constitutive expression of WspR and all downstream effects culminating in the biosynthesis of cellulose.

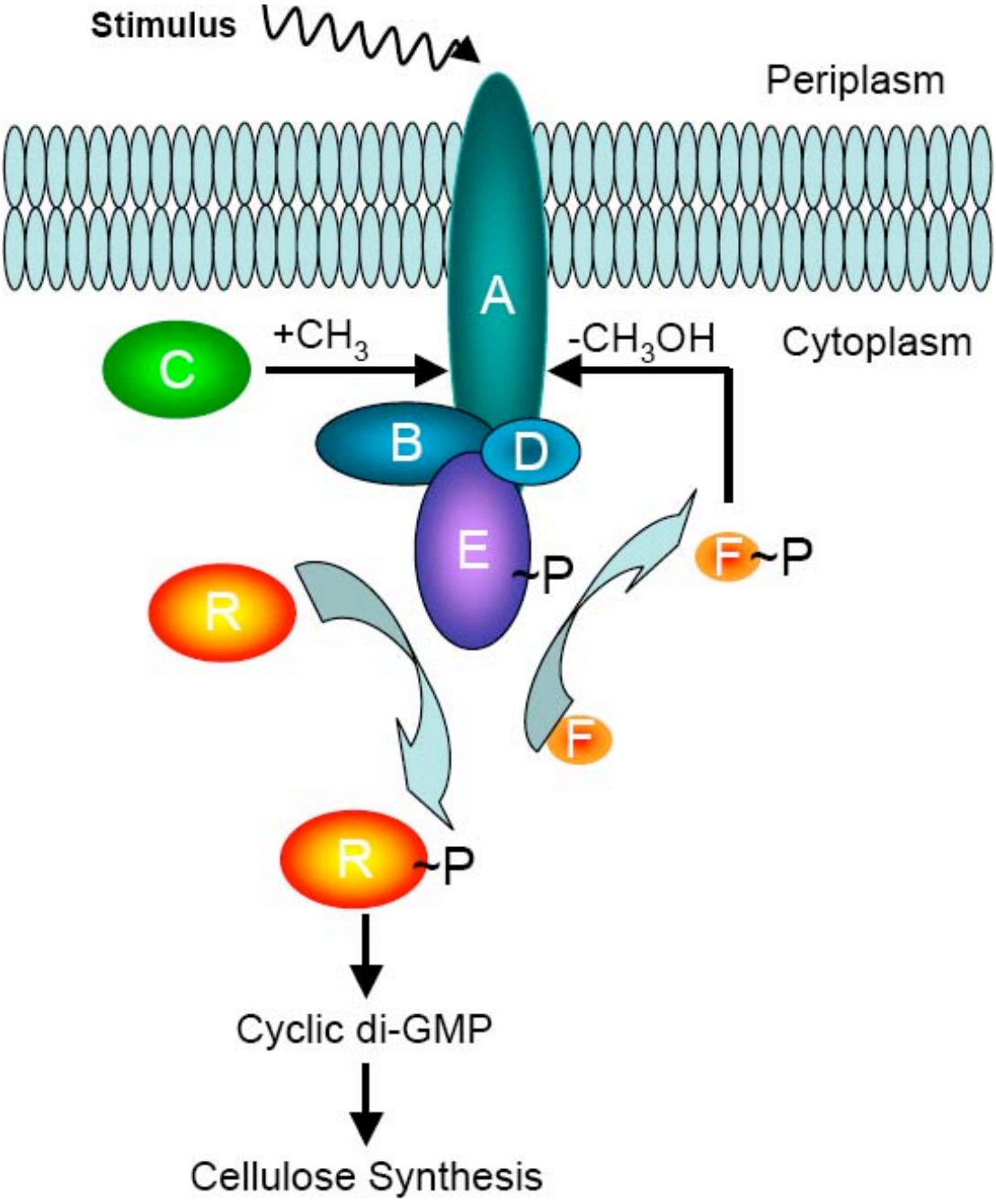

Figure 1-10: The Wsp pathway.
WspA (the MCP) receives a stimulus that causes WspE to become autophosphorylated. In turn, WspE causes the phosphorylation of WspF and WspR. WspR allosterically activates c-diGMP leading to cellulose biosynthsis while WspF negatively regulates activity in the Wsp pathway by removing methanol from WspA ending the signal. WspC acts antagonistically to WspF by adding methyl groups to WspA.

1.5.3 THE *AWS* LOCUS

The *aws* locus is a 2.3 kb cluster of three genes (*awsX, awsR* and *awsO*) that is involved in the regulation of cellulose biosynthesis and the WS phenotype (Gehrig, 2005). Figure 1-11 shows the arrangement of the three genes that are predicted to

form a single transcriptional unit (Gehrig, 2005). It was identified through transposon mutagenesis screening of a WS derived from an engineered strain, SBW25-Δ*wsp*. The first mutations were identified in *awsX*, which has no homology to any known protein domains, but BLAST searches revealed that it is conserved throughout *Pseudomonas spp* (Appendix 9.1.3). It contains a putative proteolytic cleavage site, indicating that it is likely a periplasmic protein. Downstream of *awsX* is *awsR*, which encodes a protein that contains a HAMP (histidine kinase) domain and a GGDEF (di-guanylate cyclase) domain with two predicted membrane-spanning domains at the N-terminus. In addition to the predicted di-guanlyate cyclase domain, is a like "I site" recognised by the motif RxxD (Christen *et al.*, 2006), indicating that c-di-GMP binds allosterically to AwsR. The third gene, *awsO*, encodes a protein that has homology with OmpA (a common outer membrane protein domain) and is predicted to form a pore, most likely in the outer membrane (Gehrig, 2005). The GGDEF domain in AwsR suggests that it synthesises the secondary messenger c-di-GMP similarly to WspR. In addition, AwsX has been confirmed by mutational studies to negatively regulate AwsR, similar in function to WspF acting as a negative regulator of WspR (McDonald *et al.*, 2008). Mutations that are sufficient to cause the WS phenotype have been identified in each of the three *aws* genes (Section 2.2.1).

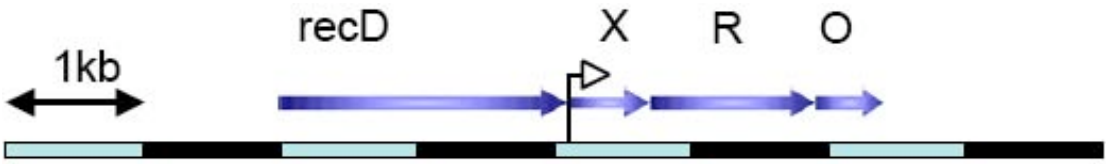

Figure 1-11: The *aws* locus.
awsXRO lies upstream of the *recD* locus.

AwsXRO form the putative signal-dependent pathway that leads to the WS phenotype (Figure 1-12). AwsO is predicted to form a pore in the outer membrane and to be associated with the periplasmic protein AwsX in an unstimulated state. The working model predicts that the association between AwsO and AwsX blocks signal transduction to AwsR, which is bound to the inner membrane. An environmental cue provides the stimulus for AwsX to be released from AwsO. AwsX is then free to bind to AwsR and induce a conformational change that leads to the dimerisation and activation of AwsR. AwsR acts to synthesise c-di-GMP, which ultimately leads to the up-regulation of cellulose expression.

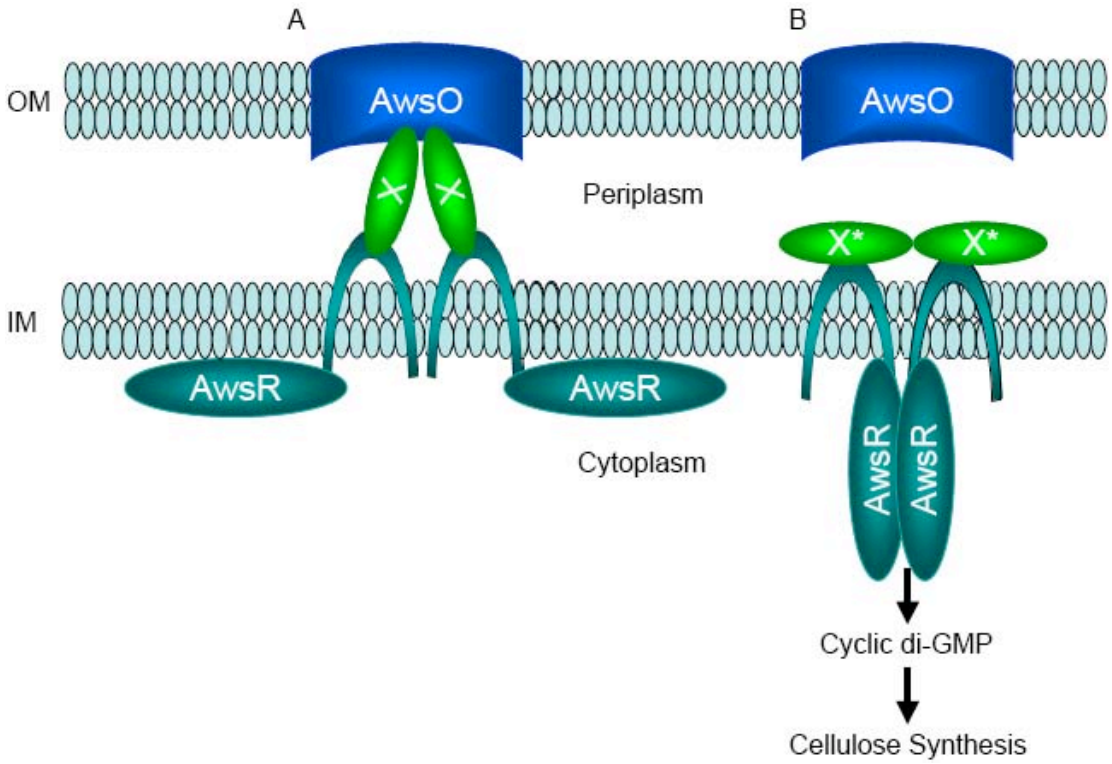

Figure 1-12: A model for the Aws protein interaction.
A) The unstimulated state of Aws Proteins. The response regulator AwsR is prevented from dimerising due to an interaction between AwsX and AwsO. B) The dimersation of AwsR in the presence of mutated AwsX (indicated by X*).

1.5.4 THE *MWS* LOCUS

The *mws* locus, is a single 3.8 kb gene, *mwsR*, that is involved in the regulation of cellulose biosynthesis and the WS phenotype (McDonald *et al.*, 2008). It is located next to *glyA* in the genomes of all sequenced *Pseudomonas* species, although GlyA has no predicted association with the activity of MwsR or the WS phenotype. *mwsR* that was identified by transposon mutagenesis screening of a WS type derived from an engineered strain of SBW25 that had the *wsp* and *aws* operons deleted (SBW25-Δ*wsp*Δ*aws*). Three domains have been identified within the *mws* locus using CDD (Marchler-Bauer *et al.*, 2005) at NCBI website (NCBI Conserved Domain Search, 2007). The MwsR protein contains a PAS (signal transduction) domain, a GGDEF domain with di-guanlyate cylcase (DGC) activity (also observed in the *aws* and *wsp* operons) and an EAL domain with phosphodiesterase (PDE) activity that is implicated in the degradation of secondary messengers like c-di-GMP (D'Argenio and Miller, 2004). This suggests a possible mechanism for *mws* as both an activator and a regulator of Wss protein activity through c-di-GMP, though previously, all enzymes

with DGC/PDE domains have been shown to have only PDE activity (Jenal and Malone, 2006).

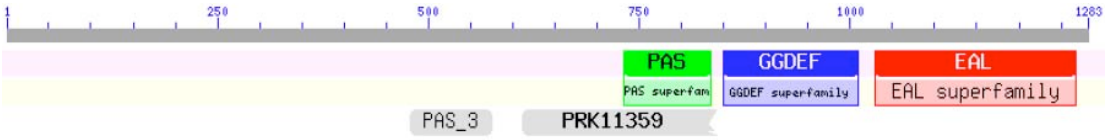

Figure 1-13: The predicted protein domain structures of MwsR (predicted by CDD).
All predicted domains occur after approximately 500 amino acid residues with the N-terminal sequences predicted to be either transmembrane spanning or inside the periplasm (Figure 1-14).

The N-terminal region of the protein showed no match to any domains in the CDD, so the entire protein sequence was analysed for membrane association. Analysis of the protein sequence with a hidden Markov model for transmembrane helix prediction (TMHMM; Figure 1-14), suggested that the N-terminus of MwsR is bound to the inner membrane and that each of the predicted domains is found on the cytoplasmic side.

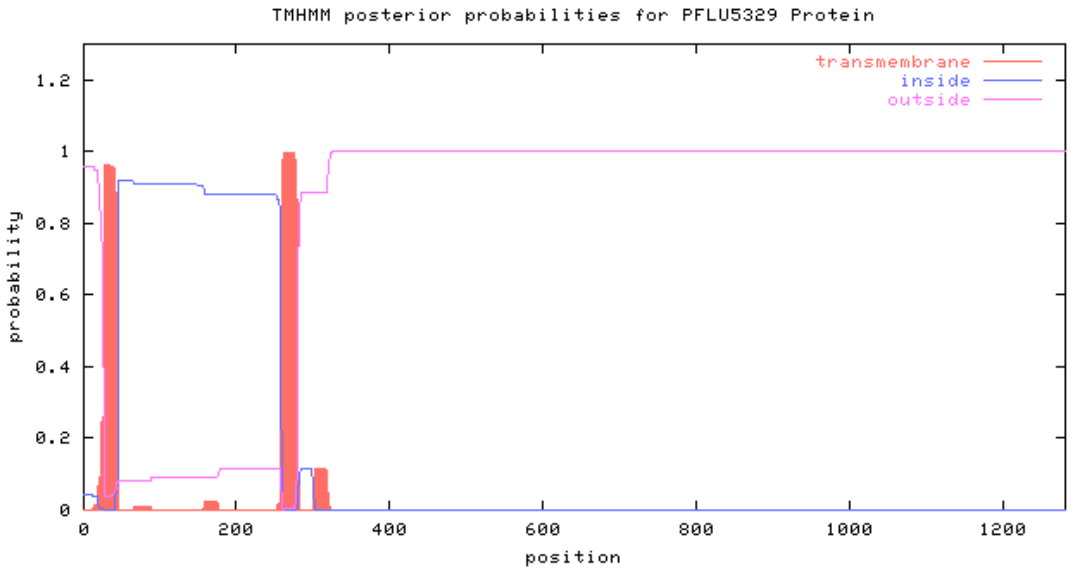

Figure 1-14: TMHMM posterior probabilities of Sequence.
Transmembrane helix prediction was performed on the Mws peptide using the hidden Markov model available online at the Centre of Biological Sequence Analysis at Denmark Technical University (DTU, 2008). The posterior probabilities of sequence analysis show two putative N-terminal transmembrane domains with probability exceeding 0.99.

Determining a working model for MwsR activity posed new challenges, because previous work had shown that DGC/PDE domain proteins had on PDE activity (Jenal and Malone, 2006), and there were no other candidate DGC proteins adjacent to *mwsR*. The mutation causing the WS phenotype from SBW25-Δ*wsp*Δ*aws* was found

in $MwsR_{EAL}$ (aa 1086) and deletion of $MwsR_{EAL}$ confirmed that the loss of PDE activity causes the WS phenotype (McDonald *et al.*, 2008). In addition, deletion of *mwsR* in the WS strain derived from SBW25-Δ*wsp*Δ*aws* and the ancestor resulted in SM colony morphology in both cases, indicating that both the DGC function of $MwsR_{EAL}$ and the PDE function of $MwsR_{EAL}$ are active (McDonald *et al.*, 2008). Therefore, the working model for the function of MwsR suggests that $MwsR_{GGDEF}$ acts to increase the levels of c-di-GMP and that $MwsR_{EAL}$ acts to lower the levels by breaking down the c-di-GMP (Figure 1-15). One might expect that mutations that lead to the WS phenotype will occur predominantly in $MwsR_{EAL}$, because abolishing c-di-GMP degradation would lead to activation of cellulose biosynthesis. However, although the majority of mutations in *mws* are in $MwsR_{EAL}$, causal mutations have been detected in $MwsR_{PAS}$, $MwsR_{PAS}$ and $MwsR_{GGDEF}$ (Section 2.2.1). This suggests that ability of $MwsR_{GGDEF}$ to up-regulate c-di-GMP and the ability of $MwsR_{EAL}$ to degrade c-di-GMP are most likely not equal (McDonald *et al.*, 2008).

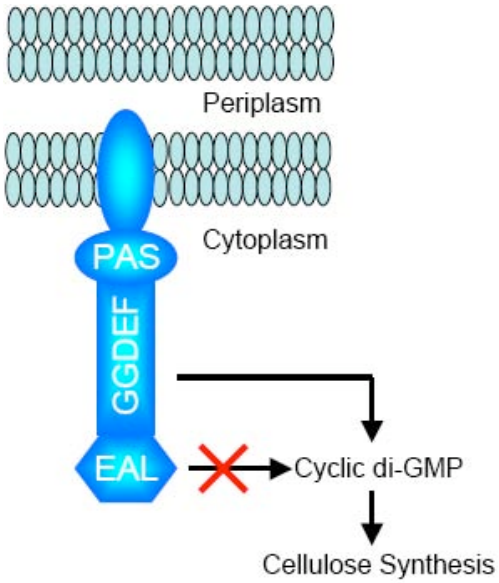

Figure 1-15: Working model for Mws function.
The GGDEF domain is predicted to up-regulate cyclic di-GMP and the EAL domain is predicted to degrade cyclic di-GMP (McDonald *et al.*, 2008).

1.6 SUMMARY

The *Pseudomonas fluorescens* system presents a unique opportunity to contribute to an enormous and controversial field in evolutionary biology by generating empirical evidence for the evolution of cooperation. It is particularly significant, because the cooperative WS phenotype in this experimental system has evolved *de novo* in the laboratory. The underlying genetic and molecular mechanisms for the adaptive radiation observed in the laboratory already have the support of a substantial body of experimental evidence and a complete genome sequence is also available. The well documented ecology and the fortuitous correlation between genotype, phenotype on an agar plate and niche specificity affords the opportunity to design experiments capable of testing hypotheses about cooperation in a multi-level selection framework. This thesis uses the ecology of *P. fluorescens* system as an experimental tool to investigate four hypotheses relating to cooperation in bacteria within a multi-level selection framework.

1 The set of 26 IWS were characterised for variation among the genotypes and phenotypes to estimate the distribution of naturally occurring WS mutants and to develop a genotype-phenotype map for phenotypic identification of genotype (Chapter 2).

2 Replicate populations of SBW25 were evolved in a long-term experiment within an MLS 1 framework to test the prediction of a modified version of Maynard Smith's Haystack model that population substructure (group structure) can favour the evolution of the cooperative WS trait (Chapter 3).

3 The ecology of the emergent group-level phenotype was contrasted with theoretical predictions of density dependent emergence to determine the nature of the unambiguously emergent group property of the mat at the air-liquid interface (Chapter 4).

4 A novel theoretical model for the evolution of multicellularity was developed that incorporated a development-like process, and a novel apparatus and

experimental design were used to determine if it were possible to observe a response to a selective regime that selected simultaneously at the level of the individual cell and the level of the group of cells (Chapter 5).

2 VARIATION AMONG INDEPENDENT WRINKLY SPREADER GENOTYPES

The theory of evolution is based on the struggle for life and the survival of the fittest. Yet cooperation is common between members of the same species and even between members of different species.

-- Axelrod, R. and Hamilton, W.D. 1981

2.1 INTRODUCTION

Rainey and Travisano (1998) showed that when *P. fluorescens* SBW25 is inoculated into a static microcosm, the broth colonising ancestral, smooth (SM) genotype evolves into a variety of niche-specialist genotypes. There are three derived types that can be broadly separated into three niche-specialist types: the wrinkly spreader (WS), the fuzzy spreader (FS) and those that resemble the ancestor (also SM). This chapter focuses on a set of 26 independent wrinkly spreaders (IWS) previously evolved in the Rainey lab by sampling 26 separate statically incubated microcosms after seven days (Bantinaki *et al.*, 2007). The IWS genotypes were characterised genetically and phenotypically to meet three main objectives. The first objective was to identify the number of mutational routes to the WS phenotype by determining the precise genetic mutations responsible for each of the IWS genotypes (Section 2.2.1). The second objective was to determine the phenotypic variation among the IWS genotypes by measuring six phenotypic characteristics: colony size, colony shape, cellulose production, mat strength, and fitness in two environments (Section 2.2.2). Finally, the genetic and phenotypic data were combined to determine if genetic changes were associated with any suite of observable phenotypes (Section 2.2.3). These objectives were motivated by the desire to have a complete understanding of the genetic mechanisms that cause the WS phenotype, the differences among different WS phenotypes, and to identify a possible correlation between the genotype and the phenotype.

2.2 RESULTS

2.2.1 IDENTIFYING THE MUTATIONAL ROUTES TO WS

At a molecular level, the WS phenotype is characterised by the over-production of cellulose and a proteinaceous adhesin (Spiers *et al.*, 2002; Spiers *et al.*, 2003). Mutations leading to the WS phenotype had previously been identified in two genes, *wspF* and *awsX* (Gehrig, 2005; Bantinaki *et al.*, 2007). The 26 IWS genotypes had been sequenced for both of these genes; however, only 14 of 26 genotypes had mutations identified within either of these genes – 13 had mutations in *wspF* and one in *awsX* (Gehrig, 2005; Bantinaki *et al.*, 2007). Therefore, it was anticipated that the causal mutations in the remaining 12 IWS would be at other loci within the genome.

WspF and AwsX are both negative regulators of GGDEF proteins, WspR and AwsR respectively (Bantinaki *et al.*, 2007; McDonald *et al.*, 2008). Loss-of-function mutations in either *wspF* or *awsX* are predicted to cause the corresponding target GGDEF proteins to increase the production of c-di-GMP in turn increasing the production of cellulose leading to the WS phenotype. A scan of the annotated open reading frames of the full genome sequence using Artemis (Rutherford *et al.*, 2000) revealed 38 GGDEF/GGEEF proteins (PF009900[1]) in *P. fluorescens* SBW25 (Appendix 9.2.1). Most of these 38 proteins have not been studied and their functions are unknown, however, given the role of GGDEF proteins in altering c-di-GMP levels, and the role of c-di-GMP in the regulation of cellulose biosynthesis, each represented a potential mutational route to the WS genotype.

To identify mutational routes to the WS phenotype for each of the 12 unknown IWS mutants, a transposon-mediated mutagenesis strategy similar to that used to identify the first *aws* mutation was employed (Gehrig, 2005; Methods 7.2.14). The transposon integrates randomly into the genome, and by chance occasionally disrupts genes involved in the WS phenotype. Disruption of the WS phenotype was observed by a

[1] PF0090 is the pfam annotation used in Artemis. It is a few amino acids longer than the conserved domain match cd01949 found using the conserved domain database at NCBI. An alignment of the two sequences is included in Appendix 9.2.1.

reversal in colony morphology from WS to SM. For 11/12[1] unknown IWS genotypes, approximately 2000 colonies from four independent conjugations were screened for a change from WS to SM colony morphology. Twenty SM colonies (WS phenotype disrupted) were selected for sequencing of the transposon insertion site. The insertion site was determined by sequencing DNA flanking the site of integration, and matching a short but unique sequence[2] from this region to the annotated sequence of the full genome using Artemis (Rutherford *et al.*, 2000). The results of the transposon-mediated mutagenesis are displayed in Table 2-1.

Table 2-1: Results of the transposon-mediated mutagenesis of 11/12 unknown IWS genotypes.

IWS	# successfully sequenced transconjugant colonies	Insertion locus	Support
D	16	*wsp*	14/16
H	14	*wsp*	13/14
I	15	*wsp*	13/15
K	10	*aws*	4/10
M	18	*aws*	6/18
P	14	*aws*	4/14
Q	19	*mws*	4/19
R	20	*mws*	7/20
S	17	*aws*	7/17
V	17	*aws*	1/17
X	17	*wsp*	17/17

Transposon insertion sites were identified by matching transposon sequence data to the annotated genome of *P. fluorescens* SBW25 using Artemis (Rutherford *et al.*, 2000). The support for the insertion site is the number of times that locus was identified from independent transconjugants.

The result of this transposon-mediated mutagenesis strategy identified another locus, *mws*, as a mutational route to the WS phenotype. This locus had previously been identified (McDonald *et al.*, 2008) as a mutational route to WS phenotype from a double knock-out mutant (SBW25-Δ*wsp*Δ*aws*), but it was not known whether or not *mws* mutants could arise *de novo* from ancestral SBW25 to become the dominant morph in statically incubated microcosms when *wsp* and *aws* pathways were available targets of mutation. In addition, the *wsp* operon was identified four times by the transposon-mediated mutagenesis screen, despite each of these strains already

[1] Only 11/12 unknown IWS were subjected to Tn-mediated mutagenesis, because IWS_Z was found to have a mutation in *mws* before Tn-mediated mutagenesis was completed.

[2] The short but unique sequence used for comparison with the whole genome and the results of this comparison are available in Appendix 9.2.1.

being sequenced for *wspF* with no mutation found. This suggested that mutations within other genes of the *wsp* operon could also contribute to the WS phenotype.

To confirm the precise genetic location of each mutation, each IWS was sequenced at the operon identified by the transposon-mediated mutagenesis screen (*e.g.* the *wsp* operon for IWS_D or the *aws* operon for IWS_V)[1]. The results displayed in Table 2-2 show the precise mutation for each of the IWS genotypes. The precise mutation was identified at each locus determined by the transposon-mediated mutagenesis strategy for all but one genotype of the IWS (IWS_X), and despite repeated sequencing of the entire *wsp* operon and the ~50 bp intergenic sequence between *wspF* and *wspR* no mutation has yet been identified.

[1] Relevent portions of the alignment are available in Appendix 9.2.1.

Table 2-2: All IWS mutations.

IWS	Gene	Mutation Type	Nucleotide change	Amino acid change	Source/Reference
A	*wspF*	SNP (transversion)	T 14 G	Residue change I 5 S	(Bantinaki *et al.*, 2007)
B	*wspF*	Deletion	Δ620-674	Truncation APTTTFVY*	(Bantinaki *et al.*, 2007)
C	*wspF*	SNP (transversion)	G 823 T	Residue change G 275 C	(Bantinaki *et al.*, 2007)
D	*wspE*	SNP (transition)	A 1916 G	Residue change D 638 G	This study
E	*wspF*	SNP (transversion)	G 658 T	Residue change V 220 L	(Bantinaki *et al.*, 2007)
F	*wspF*	SNP (transition)	C 821 T	Residue change T 274 I	(Bantinaki *et al.*, 2007)
G	*wspF*	SNP (transition)	C 556 T	Residue change H 186 Y	(Bantinaki *et al.*, 2007)
H	*wspE*	SNP (transversion)	A 2202 C	Residue change K 734 N	This study
I	*wspE*	SNP (transversion)	G 1915 T	Residue change D 638 Y	This study
J	*wspF*	Deletion	Δ865-868	Truncation KVI*	(Bantinaki *et al.*, 2007)
K	*awsO*	SNP (transversion)	G 125 T	Residue change G 41 V	This study
L	*wspF*	SNP (transversion)	G 482 A	Residue change G 161 D	(Bantinaki *et al.*, 2007)
M	*awsR*	SNP (transition)	C 164 T	Residue change S 54 F	This study
N	*wspF*	SNP (transversion)	A 901 C	Residue change S 301 R	(Bantinaki *et al.*, 2007)
O	*wspF*	Deletion	Δ235–248	Truncation VIARPT*	(Bantinaki *et al.*, 2007)
P	*awsR*	Insertion	ins(9) after 225	3 residue duplication ATE → ATEATE	This study
Q	*mwsR*	Insertion	ins(6) after 3271	2 residue insertion AVG → AVAVG	This study
R	*mwsR*	SNP (transition)	T2183 C	Residue change V 727 A	This study
S	*awsX*	SNP (transition)	C 472 T	Truncation *	This study
T	*awsX*	Deletion	Δ92–130	In-frame deletion AQAPDPADLADQR	(Gehrig, 2005)
U	*wspF*	Deletion	Δ823–824	Truncation HGARRGPGPQIAT*	(Bantinaki *et al.*, 2007)
V	*awsX*	SNP (transversion)	T 74 G	L 24 R	This study
W	*wspF*	Deletion	Δ149	Truncation S*	(Bantinaki *et al.*, 2007)
X	*?*	?	?	?	This study
Y	*wspF*	Deletion	Δ166-180	Inframe deletion IMPVM	(Bantinaki *et al.*, 2007)
Z	*mwsR*	SNP (transition)	G 3055 A	Residue change A 1018 T	This study

Relevant mutations are shown in Appendix 9.2.1.

One interesting result is the distribution of the three mutations in *mws*. One mutation was observed just outside the PAS domain, but within the multi-domain PRK11359 (aa 727), one in the EAL domain (aa 1090), and one in the region separating the GGDEF domain from the EAL domain (aa 1018). The mutations of IWS_R and IWS_Z have been introduced into the background of the ancestor as well as a different EAL domain mutation as part of another study to show that mutations in all domains of MwsR are causal of the WS phenotype (McDonald *et al.*, 2008). Mutations were generally expected to be found in $MwsR_{EAL}$, because such mutations are more likely to abolish the phosphodiesterase (PDE) function of MwsR causing an increase of c-di-GMP leading to over-expression of bacterial cellulose culminating in the WS phenotype. Although mutations elsewhere within *mwsR* could have a similar effect on protein function, both mutations outside of $MwsR_{EAL}$ (IWS_R and IWS_Z) were amino acid substitutions, which suggests either a gain of function or a site-specific inactivation of the PDE function from outside $MwsR_{EAL}$. This complex relationship between the activities of $MwsR_{GGDEF}$ and $MwsR_{EAL}$ to increase the DGC activity of in $MwsR_{GGDEF}$ are explored later in the discussion (Section 2.3.3).

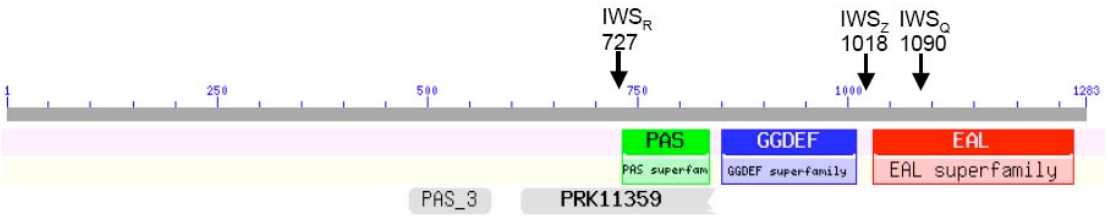

Figure 2-1: Domain structure of *mwsR* showing known mutations.
Illustration was obtained using the CDD (Marchler-Bauer *et al.*, 2005) at NCBI. The mutation in IWS_Z is outside the predicted GGDEF domain (ends at aa 1009) and EAL domain (begins at 1029), while IWS_R are IWS_Q are within the PAS and EAL domains respectively.

Another interesting result is the identification of three independent mutations in *wspE*, observed in IWS_D, IWS_H and IWS_I. The working model for the Wsp pathway has WspE autophosphorylating WspF and WspR and, in turn, WspE is negatively regulated by WspF as part of a feedback loop that controls WspE activity (see Section 1.5.2.1 for a detailed description of the Wsp pathway). This result suggests that mutations in *wspE* can cause WspE to become constitutively active. Each of the three mutations was located at the C-terminal end of the protein in the signal receiver (REC) domain, which spans from amino acid 635 to 755 (Figure 2-2).

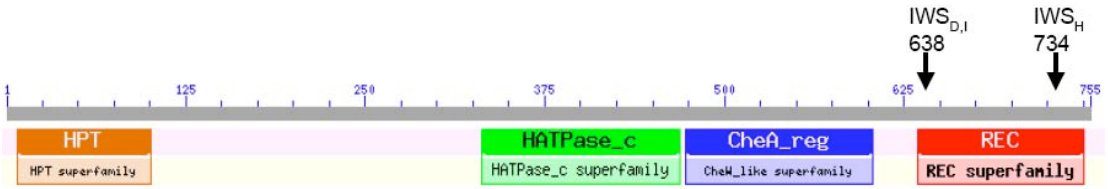

Figure 2-2: Domain structure of WspE showing the known mutations.
Illustration was obtained using the CDD (Marchler-Bauer *et al.*, 2005) at NCBI. The mutations for IWS_D and IWS_I are both located in amino acid codon 638, but are independent mutations at different nucleotide residues. The amino acid change is also different for IWS_D and IWS_I.

To identify conserved residues within the REC domain of WspE (denoted as $WspE_{CheY}$, because of its similarity with CheY from *E. coli*), the 121 amino acids of $WspE_{CheY}$ was BLAST searched against the Genbank database to obtain the four most closely related protein sequences. The $WspE_{CheY}$ and the four most closely related sequences were aligned with CheY from *E. coli* (the most well studied of all the proteins with a REC domain), the REC domain of FrzE ($FrzE_{CheY}$) from *M. xanthus* (the protein with the highest overall similarity to WspE and the protein with the most significant biochemical data) and the REC domain of WspR from *P. fluorescens* (Figure 2-3). The observed mutations at aa 638 and 734 were both at highly conserved residues indicating that they are of functional significance to $WspE_{CheY}$.

Label	Genbank Accession	Organism
WspE	29612000	*P. fluorescens* SBW25
FrzE	120546	*M. xanthus*
CheY	145534	*E. coli*
WspR	18535672	*P. fluorescens* SBW25
Blast1	77457283	*P. fluorescens* Pf-01
Blast2	70728515	*P. flourescens* Pf-5
Blast3	170723287	*P. putida* W619
Blast4	104780455	*P. entomophila* L48

Figure 2-3: Alignment of REC domains.
Mutation sites are marked with red boxes and are highly conserved. IWS_D and IWS_I have mutations at residue 4 in this alignment, while IWS_H has a mutation at residue 100.

2.2.1.1 Genetic Reconstruction of a *wspE* Mutation

The novelty of finding mutations in *wspE* prompted the need for confirmation that mutations in *wspE* were necessary and sufficient to cause the WS phenotype. IWS_H was selected for genetic reconstruction of its mutation in the ancestral wild-type background. Primers were designed one kb upstream and downstream of the mutation site in *wspE* and the resulting two kb fragment was ligated into a pCR8TOPO cloning vector. Chemically competent *E. coli* cells were transformed and grown up overnight. The insert was cleaved from a pCR8TOPO vector and ligated into the suicide vector pUIC3. Chemically competent *E. coli* cells were transformed with the pUIC3+insert before it was introduced into wild-type SBW25 using a tri-parental mating method (Methods 7.2.12). The presence of the mutation in the wild-type background was confirmed by sequencing (Appendix 9.2.2). Fitness assays (Methods 7.2.15.1) confirmed that there was no difference in relative fitness between the reconstructed genotype and IWS_H in a 24-h competition with SBW25-*lacZ* in a shaken environment ($F_{1,6}$ = 0.4582, p = 0.5237, Figure 2-4). In addition, the colony morphologies appeared similar to one another (Figure 2-5).

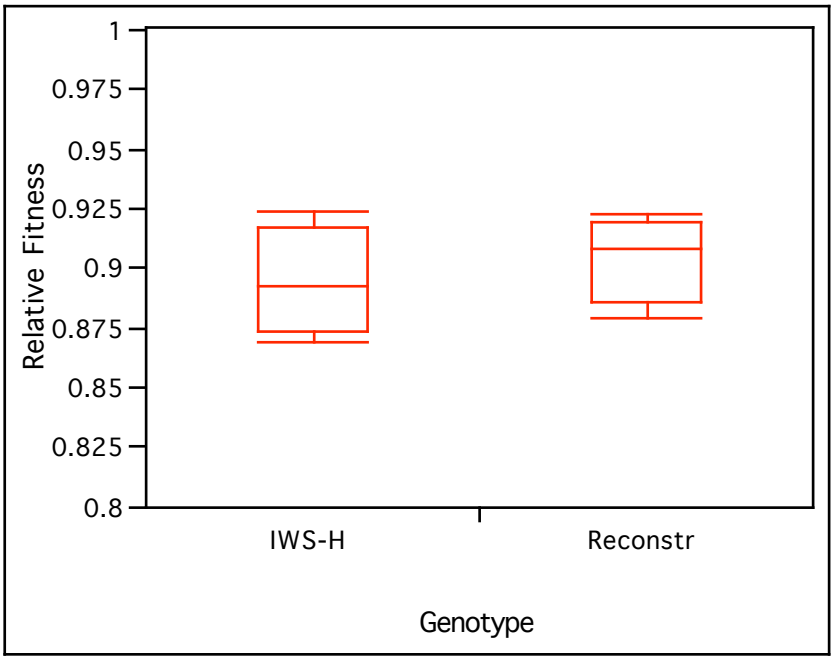

Figure 2-4: Relative fitness of IWS_H and the reconstructed genotype.
Relative fitness was measured as a Malthusian ratio in competition with SBW25-*lacZ* over 24 h in a shaking environment. There is no significant difference between the reconstructed genotype and IWS_H. The grey line represents the population average.

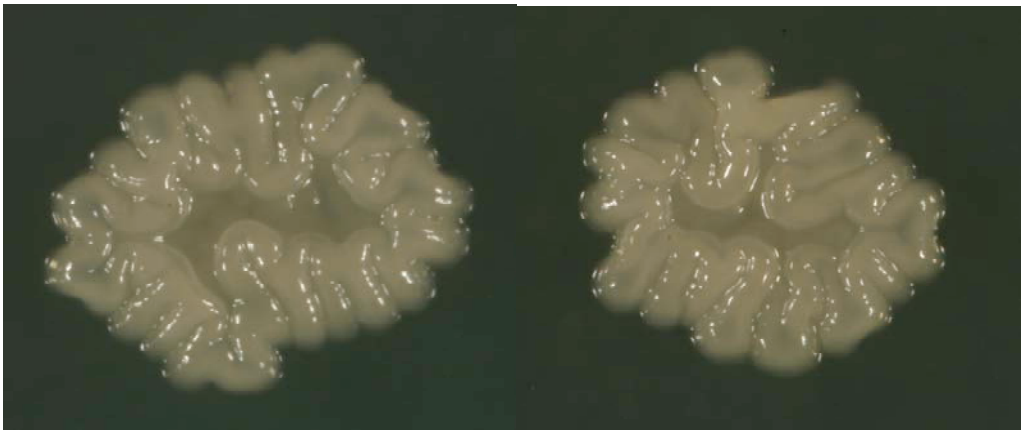

Figure 2-5: Comparison of the colony morphology of IWS_H and the reconstructed genotype. The morphology for IWS_H is displayed on the left and the morphology for the wild-type with the reconstructed mutation in *wspE* identical to that of IWS_H is on the right. Both photos were taken after 48 h of incubation at 28˚C on KB agar plates.

2.2.2 PHENOTYPIC VARIATION OF WS

Wrinkly spreaders are a phenotypically diverse group of genotypes unified by their ecological specificity for the air-liquid interface. All 26 of the IWS exhibit the definitive characteristics of niche specificity for the air-liquid interface, appear 'wrinkly' on agar plates and exhibit 'clumping' in shaken cultures. The cause of these phenotypes at the molecular level is over-production of cellulose in conjunction with a similarly overproduced proteinaceous adhesin (Spiers *et al.*, 2003). Several assays were employed to generate quantitative measurements of phenotypic characteristics of WS in three environments, agar plates, shaken cultures and static cultures. The six phenotypic characteristics chosen for this analysis were colony size, colony shape (more specifically colony 'wrinkliness'), cellulose production, mat strength, fitness over 24 h in shaking cultures and fitness over 72 h in static cultures. While all WS genotypes share these characteristics, quantification of the extent to which WS genotypes vary phenotypically was unknown.

2.2.2.1 VARIATION IN COLONY SIZE AND WRINKLINESS

To assess the variation in colony size and 'wrinkliness', overnight cultures of each IWS genotype and the *wspF* deletion mutant (SBW25-Δ*wspF*) were grown, diluted appropriately and plated on agar plates. Five replicate digital photographs were taken of randomly chosen colonies of the same genotype after two days of incubation on KB agar plates at 28˚C for each IWS genotype and the *wspF* deletion mutant

(SBW25-Δ*wspF*) (Methods 7.2.18). Each of the five digital photographs for each IWS genotype and SBW25-Δ*wspF* were analysed using Sigmascan Pro (Systat Software Inc.) by manipulating built-in macros[1]. To identify the boundaries of objects within the photographic image, the 'intensity threshold' was altered to distinguish the colony from the background. An overlayed object was defined for each colony image based on intensity threshold (*e.g.* Figure 2-6). Overlayed layers can be subjected to measurements or further refined with additional overlaying functions.

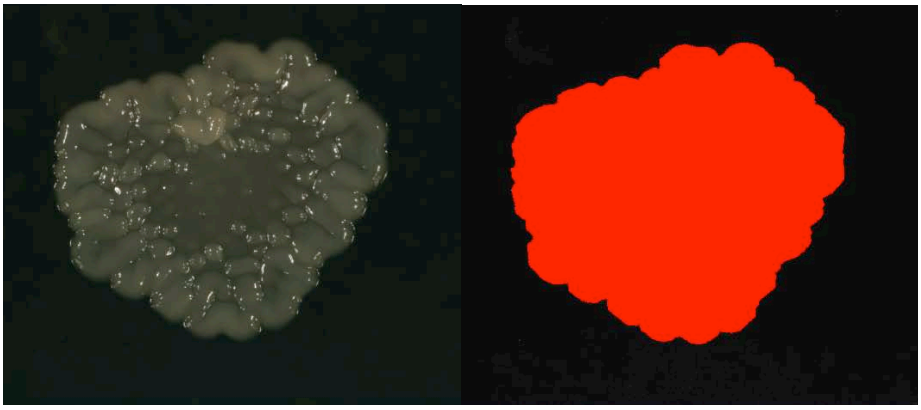

Figure 2-6: Image overlaying in Sigmascan Pro for an SBW25-Δ*wspF* colony.
The red overlayed object can be subjected to measurements or further overlaying. Photo taken after 48 h of incubation at 28˚C on KB agar plates.

The area of each of the overlayed objects was measured by counting the number of pixels within the overlayed object. An ANOVA of the colony area was used to determine the variation among IWS genotypes for five replicates (Figure 2-7).

[1] Example macros available in Appendix 9.2.3.

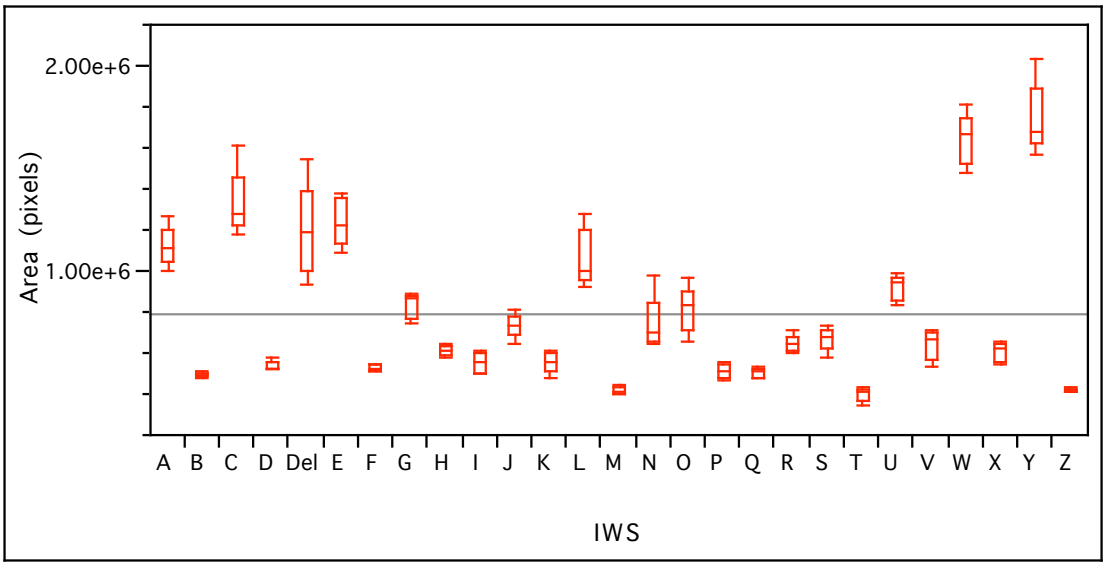

Figure 2-7: Variation in colony area for all IWS genotypes.
The box plots presented are for all 26 IWS genotypes plus SBW25-Δ*wspF* (labelled Del) for five replicates. The grey line indicates the population mean. A Tukey's pair-wise comparison shows that there are 11 significantly different groups, and one unique group (IWS_W and IWS_Y) at the 95% level (Appendix Table 9-15)

The results of a one-way ANOVA showed that there was significant variation in the colony area among the IWS genotypes ($F_{26,108} = 78.57$, $p = < 0.0001$). Although the variances were unequal under a Brown-Forsythe test ($F_{26,108} = 1.668$, $p = 0.0363$), the variation in colony area among the IWS genotypes was still significant under a Welch ANOVA allowing for unequal variances ($F_{26,108} = 71.68$, $p = < 0.0001$). The photographs in Figure 2-8 illustrate the large size range among the IWS genotypes between the smallest (IWS_T) and the largest (IWS_Y).

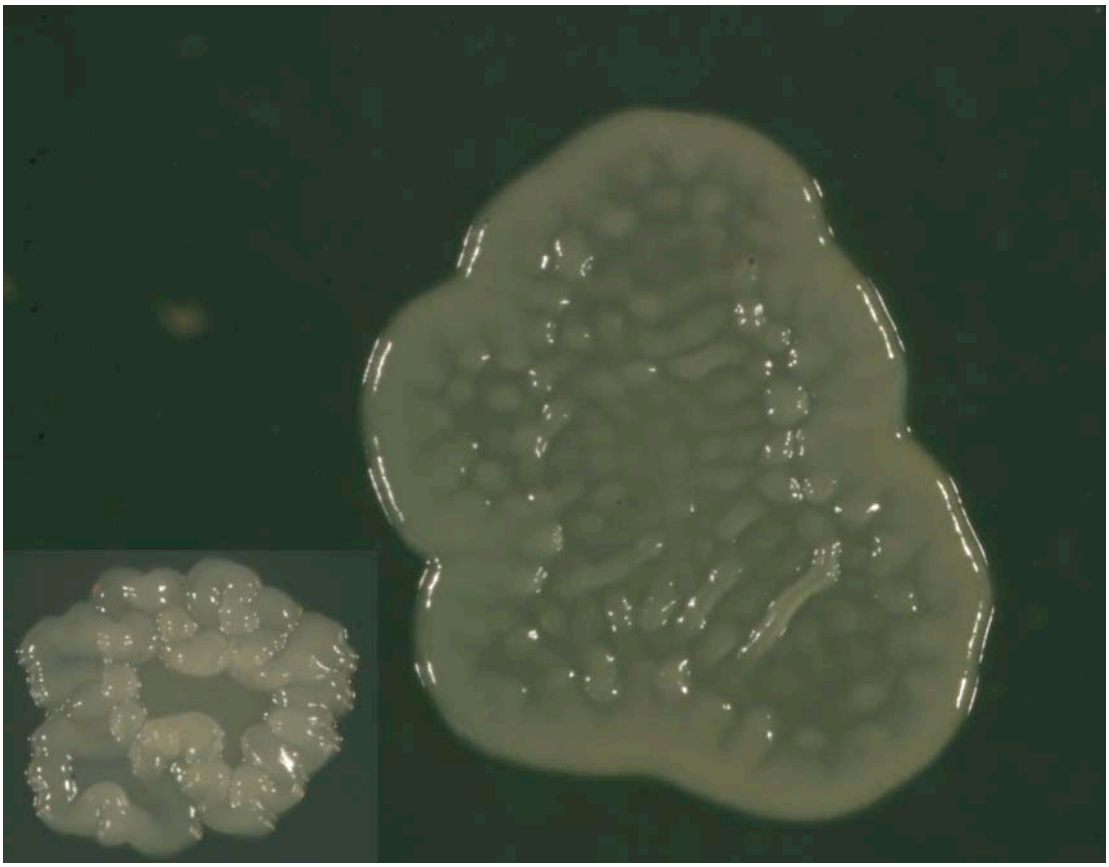

Figure 2-8: Size difference in wrinkly spreader colonies.
These photographs are of the smallest (IWS_T) and largest (IWS_Y) wrinkly spreaders among the IWS genotypes. Both photos were taken after incubation for 48 h at 28˚C on KB agar plates at the same magnification and scale.

Another way in which WS colonies vary is the degree to which they appear wrinkled. SM colonies are circular on KB agar plates, but over-production of cellulose among WS genotypes, leads to the wrinkled morphology and a departure from a circular shape. Therefore, in this thesis I use the term 'circularity' as a measure of how wrinkly WS colonies are. The circularity of a colony is determined by comparing the perimeter of the colony (measured on the overlayed object using Sigmascan Pro) to the 'estimated perimeter'. The estimated perimeter is calculated from the area of the colony assuming that the object is a circle (Eq 2-1).

$$estimated\ perimeter = \frac{2\,area}{\sqrt{\frac{area}{\pi}}} \qquad \text{Eq 2-1}$$

The colony circularity was calculated as the ratio of the measured perimeter to the estimated perimeter (Figure 2-9) and deviates above one as a positive function of the

wrinkliness of the colony. Circularity is expected to be approximately equal to one for circular objects (*e.g.* a computer generated image of a circle was tested to have a ratio of 1.094, 95% CI: 1.063 – 1.125).

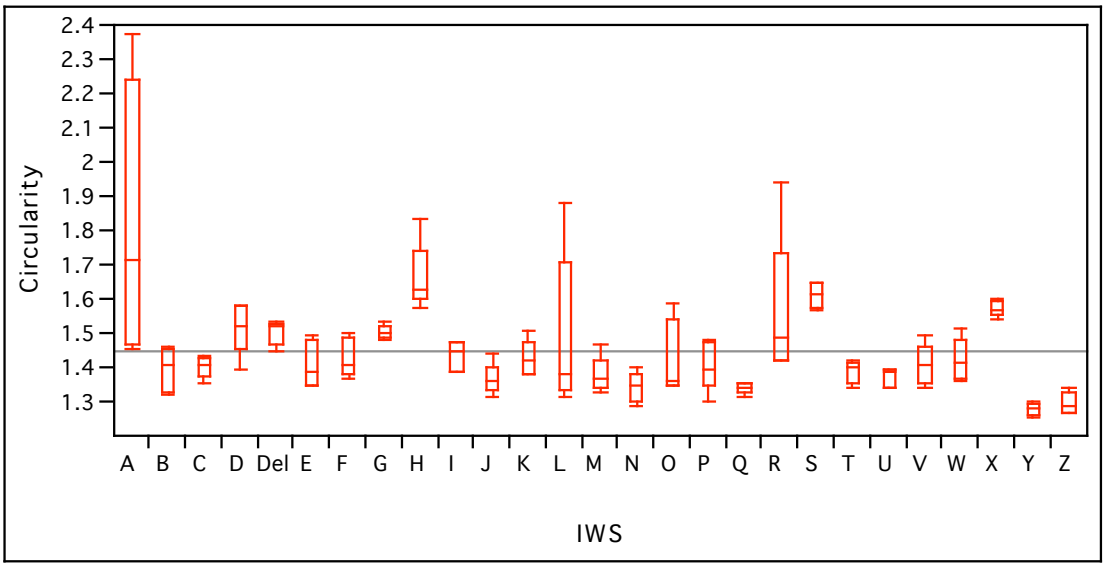

Figure 2-9: Variation in colony circularity for all IWS genotypes.
The box plots presented are for all 26 IWS genotypes plus SBW25-Δ*wspF* (labelled Del) for five replicates. The grey line indicates the population mean. A Tukey's pair-wise comparison of the samples showed that there were six significantly different groups, but no unique groups at the 95% significant level (Appendix Table 9-16).

A Brown-Forsythe test showed that variances for colony circularity across the IWS genotypes were not equal ($F_{26,108} = 4.5765$, $p < 0.0001$). However, a Welch ANOVA allowing for unequal variances showed that there was significant variation in colony circularity among the IWS genotypes ($F_{26,108} = 38.681$, $p < 0.0001$).

2.2.2.2 Variation in Mat Strength

The definitive emergent character for groups of cooperating WS cells is the mat formed at the air-liquid interface. Eight replicate microcosms for each of the IWS genotypes were grown statically for 72 h at 28˚C to allow mats to form. The strength of each of the mats was measured as the speed of rotation at which the mat was destroyed by rotating on a platform shaker (Methods 7.2.19). The mat was deemed destroyed when either partial detachment or fragmentation of the mat was observed. Figure 2-10 shows the variation in mat strength among the IWS genotypes. The results of a one-way ANOVA show that there is significant variation in mat strength

among the IWS genotypes ($F_{26,189}$ = 26.31, p = < 0.0001) with no significant difference among the variances (Brown-Forsythe test: $F_{26,189}$ =1.097, p = < 0.348).

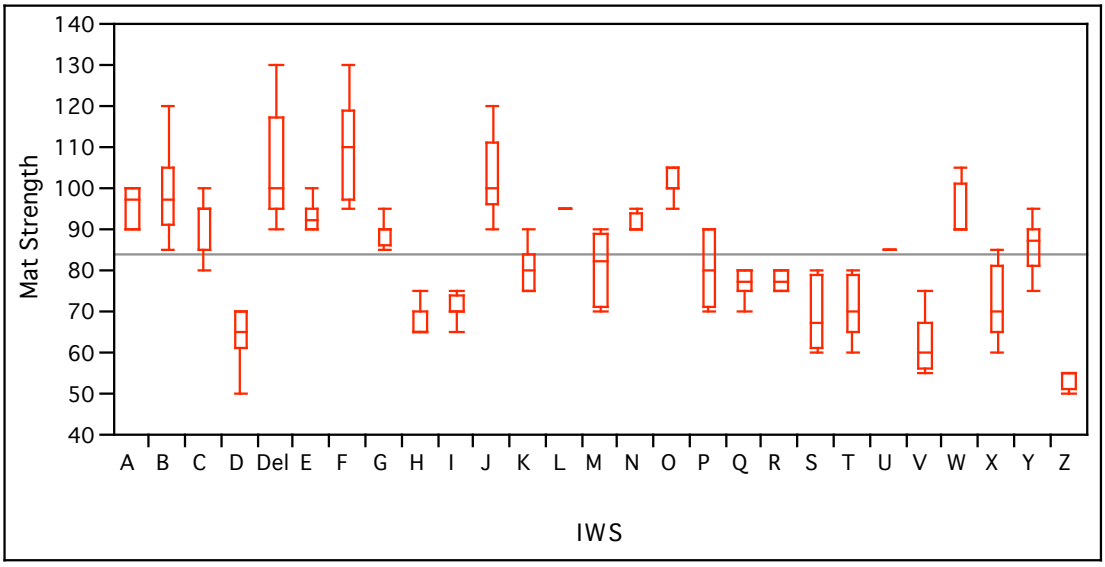

Figure 2-10: Variation in mat strength for all IWS genotypes.
The box plots presented are for all 26 IWS genotypes plus SBW25-Δ*wspF* (labelled Del) for eight replicates. The grey line indicates the population mean. A Tukey's pair-wise comparison showed that there were 11 significantly different groups, but no unique groups at the 95% level (Appendix Table 9-17).

2.2.2.3 Variation in Congo Red Binding

An increase in cellulose production is the downstream consequence and molecular level phenotype for all known WS mutations. Congo Red (CR) is a dye with a strong non-covalent affinity for cellulose and other adhesive substrates secreted by WS cells (Weiner *et al.*, 1999; Spiers *et al.*, 2003). Ten replicate microcosms for each of the IWS genotypes were grown statically for 72 h at 28˚C, then homogenised by vortexing. The homogenate was mixed with CR and incubated to allow the CR to bind. The cells were pelleted and the supernatant transferred to a microtitre plate to measure the absorbance at 490 nm in a microtitre-plate reader. The variation in absorbance among the IWS genotypes is displayed in Figure 2-11 (lower absorbance means that more CR has been bound). The results of a one-way ANOVA showed that there was significant variation in the production of CR binding substrates among the IWS genotypes ($F_{26,243}$ = 9.8498, p = < 0.0001) and that the variances can be assumed to be equal (Brown-Forsythe test: $F_{26,243}$ = 1.0523, p = 0.3998).

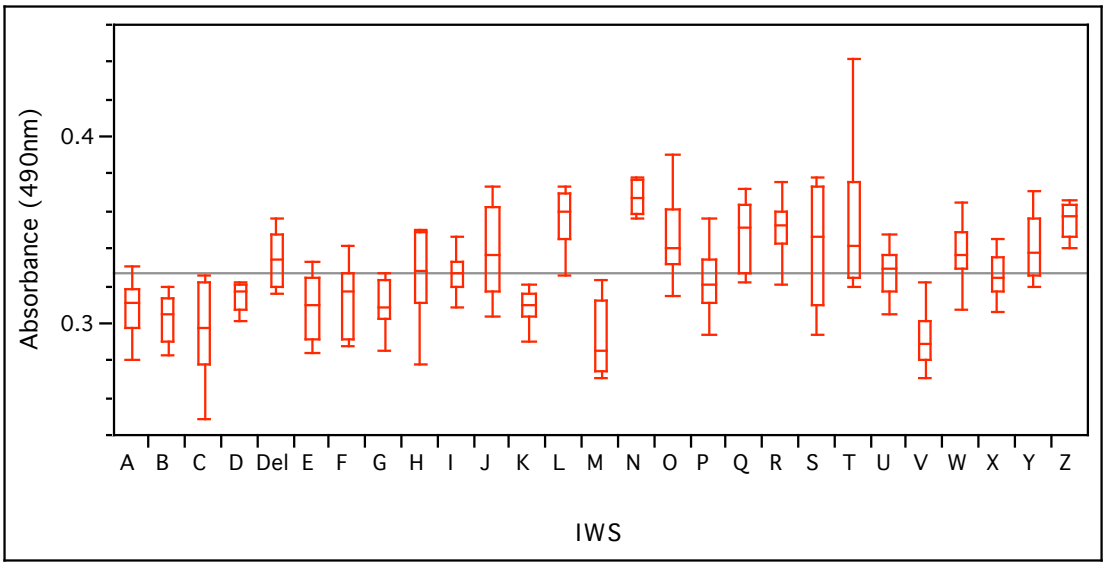

Figure 2-11: Variation in Congo red binding for all IWS genotypes.
The box plots presented are for all 26 IWS genotypes plus SBW25-Δ*wspF* (labelled Del) for eight replicates. A Tukey's pair-wise comparison showed that there were seven significantly different groups, but no unique groups at the 95% level (Appendix Table 9-18).

2.2.2.4 Variation in Fitness

The variation in fitness for the IWS genotypes was measured in two environments with three different aims. The first aim was to measure the relative fitness of each WS genotype against ancestral SBW25 in a shaking environment with abundant resources and oxygen supply to demonstrate the intrinsic metabolic cost associated with the WS phenotype. The second aim was to calculate the relative fitness of each WS genotype against LSWS at the air-liquid interface by including a third competitor, SM, to occupy the broth phase. The third aim was to correlate the fitness data from the two environments to determine if there was a fitness trade-off associated with the niche-specific adaptation to the air-liquid interface among WS types.

The relative fitness of WS types compared to the ancestral SM type over 24 h in a shaking microcosm had not previously been studied. However, the fitness of LSWS relative to SM was previously measured at 0.8 (95% CI: 0.69 – 0.91) during the first 24 h of competition in a static microcosm, when growth is exponential and all resources other than oxygen are abundant (Rainey and Rainey, 2003). Growth during exponential phase with abundant resources is the ideal environment to illustrate the intrinsic metabolic cost of over-production of cellulose (*i.e.* of the WS phenotype), although it is not the environment in which it arose and increased in frequency. Therefore, to measure this cost, eight replicate 24-h fitness assays were performed in

a shaking environment for each of the IWS or SBW25-Δ*wspF* against a neutrally marked version of the ancestor (SBW25-*lacZ*)[1]. The relative fitness over 24 h of exponential growth was calculated as the ratio of the Malthusian parameters for the strain and the competitor (Lenski *et al.*, 1991). The results of a one-way ANOVA showed that there was significant variation in relative fitness in a shaking environment among the IWS genotype ($F_{26,189}$ = 28.9174, p = < 0.0001) and that the variances could be assumed to be equal (Brown-Forsythe test: $F_{26,189}$ = 1.4275, p = 0.092).

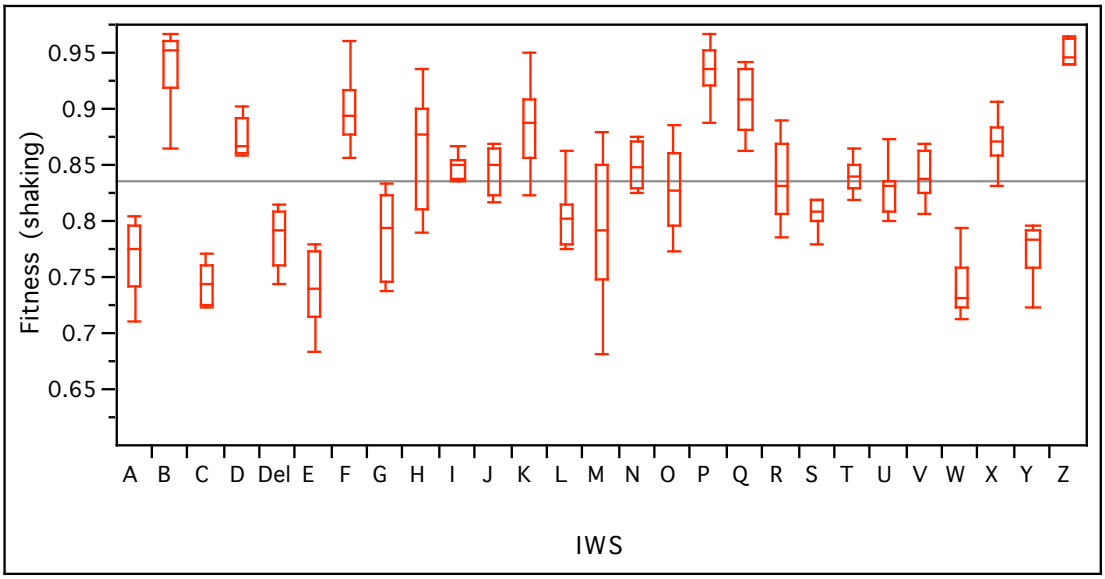

Figure 2-12: Variation in the fitness of each IWS genotype in a shaking environment.
The box plots presented are for all 26 IWS genotypes plus SBW25-Δ*wspF* (labelled Del) for eight replicate fitness measurements. A Tukey's pair-wise comparison shows that there are 12 significantly different groups, but no unique groups at the 95% level (Appendix Table 9-19).

The variation in WS fitness in a three-way competition in a static environment had previously been reported for each of the IWS genotypes from four-fold replicated data (Bantinaki *et al.*, 2007). These experiments were designed to focus the competition between WS types on the air-liquid interface by including an SM competitor strain (wild-type SBW25) to occupy the broth phase (Methods 7.2.15.2). Eight replicate microcosms for each of the IWS genotypes or SBW25-Δ*wspF* were

[1] The validity of the SBW25-*lacZ* as a neutrally marked strain of SBW25 for short-term experimentation has been established (Zhang and Rainey, 2007), but see also Appendix 9.1.1 for validation of long-term neutrality of the marker.

grown in competition with LSWS-*lacZ* and SBW25 inoculated at approximately equal densities and incubated statically for 72 h at 28˚C. The relative fitness of each IWS genotypes compared to LSWS-*lacZ* during 72 h of three-way competition growth was calculated as the ratio of the Malthusian parameters for the strain and the competitor (Lenski *et al.*, 1991). The results of a one-way ANOVA showed that there was significant variation in the relative fitness in a static environment among the IWS genotypes ($F_{26,189} = 42.509$, $p = < 0.0001$) and that the variances could be assumed to be equal (Brown-Forsythe test: $F_{26,189} = 1.199$, $p = 0.242$).

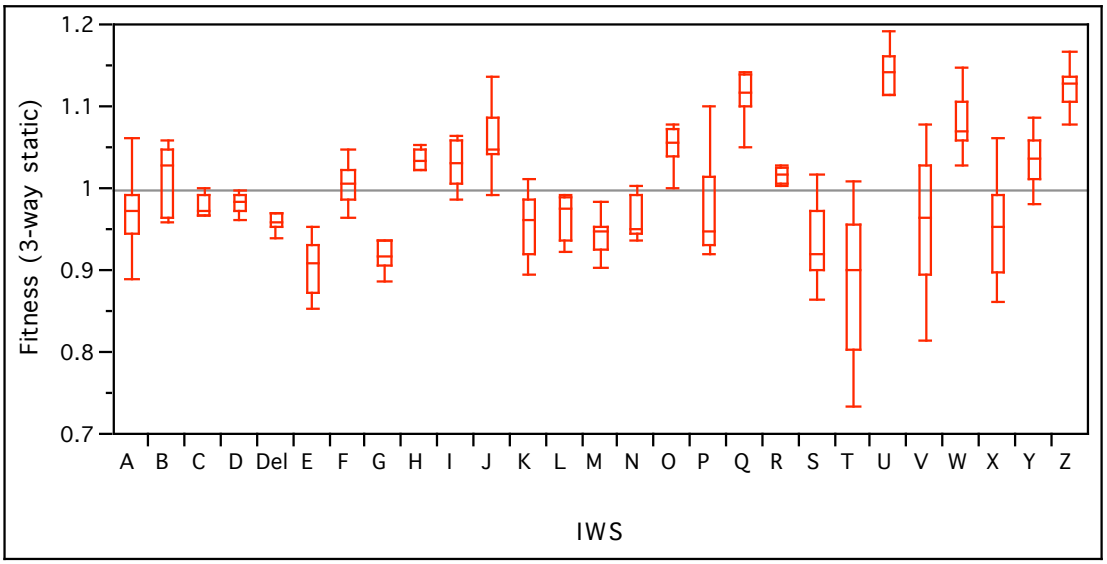

Figure 2-13: Variation in fitness of each IWS genotypes in a static environment.
The box plots presented are for all 26 IWS genotypes plus SBW25-Δ*wspF* (labelled Del) for eight replicates fitness measurements. A Tukey's pair-wise comparison shows that there are nine significantly different groups, but no unique groups at the 95% level (Appendix

Table 9-20).

The WS phenotype is a niche-specific adaptation to the air-liquid interface in a static culture that has a trade-off in a shaking culture for all WS types compared to SM types. The magnitude of this trade-off was expected to differ among WS types, but this had never been shown. With reference to the diagram (Figure 2-14), two hypotheses were examined about the fitness trade-offs among WS mutants. In the first hypothesis, WS mutants that have a higher fitness in the static environment will have a higher fitness in a shaking environment, *i.e* the causal WS mutation has conferred a general increase in fitness. Second, WS mutants that have a higher fitness in the static environment will have a lower fitness in a shaking environment, *i.e.* the causal WS mutation has conferred a niche-specific increase in fitness that is traded-off between the environments even within WS types[1].

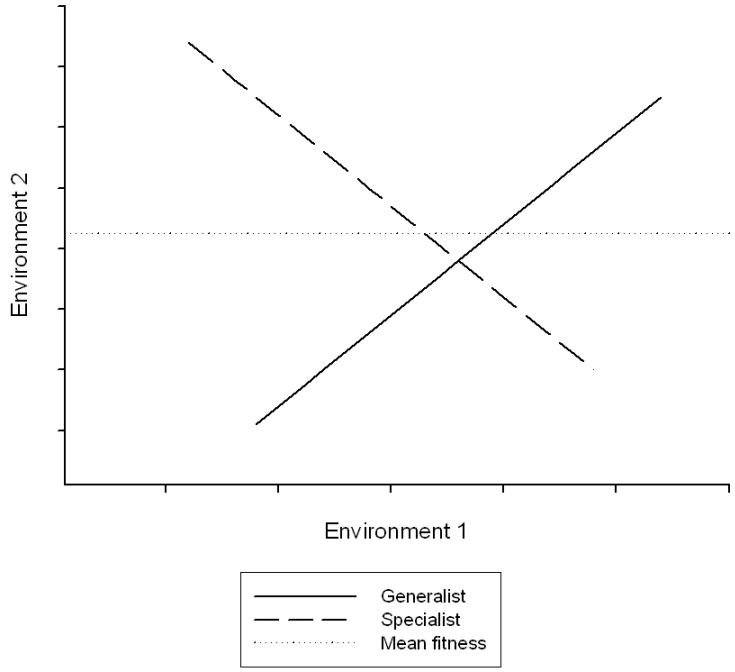

Figure 2-14: Predictions for fitness in different environments.
A positive correlation is predicted if mutations confer general increases in fitness, while a negative correlation is predicted if mutations confer environment specific increases in fitness.

[1] A trade-off can be observed when benefits in one environment are traded by sacrificing benefits in the other, but will also be observed if there is an accumulation of mutations that are neutral in one environment, that are deleterious in the other.

The two hypotheses for the trade-off among WS types were investigated by correlating the fitness data from the shaking and the static environments (Figure 2-15). A straight line was fit to the data to determine the relationship of fitnesses between the two environments among WS mutants; however, there was no correlation between the fitness in a shaking and the fitness in a static environment, (r^2 = 0.037) and the fit was not significant ($F_{1,25}$ = 0.974, p = 0.333). Therefore, counter to expectations, neither of these hypotheses was supported.

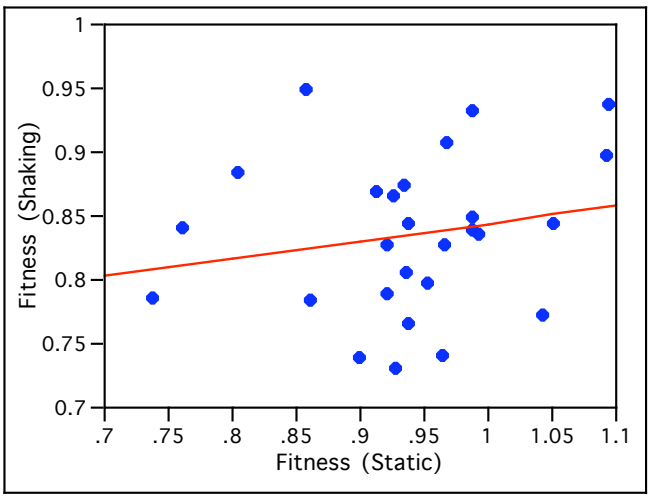

Figure 2-15: Contrast of Shaking Fitness and 3-way Static Fitness for the IWS genotypes. The fitnesses in the shaking and static environments are on different scales because they were measured relative to different genotypes. A straight line was fit to the data with r^2 = 0.037. Genotypes that are further from the line of best fit are relatively fitter in one environment than the other.

2.2.3 Correlations between Genotype and Phenotype

With the knowledge of the mutational origins for all 26 IWS genotypes and six replicated measurements of phenotypic characteristics, an exciting and obvious question arises. Is there a correlation between genetic variation and phenotypic effect? A prelude to this question was examined in Bantinaki *et al.* (2007) by investigation of the correlation between sequence variation within only *wspF* mutations and fitness (in a three-way static competition). The conclusion from this work was that the substantial variation in fitness could not be predicted based on the particular kind of mutation, nor on the predicted activity of the resulting protein (Bantinaki *et al.*, 2007). However, when the genetic data from Bantinaki *et al.* (2007) was applied to photographs of the colonies from the IWS genotypes, it appeared initially that there were similarities among genotypes with mutations in *wspF* that distinguished them from genotypes with mutations at other loci. As more mutations

became known within the IWS genotypes visually intuitive similarities and differences appeared to group similar types with one another. Therefore, it was suggested that appropriate measures of phenotypic variation might encapsulate these visually intuitive differences and similarities, and that this could be use to develop a model capable of predicting phenotype based on genotype.

To determine the significant genotypic factors for this phenotypic variation, univariate analyses for the effect of gene, locus and mutation type were performed independently for each of the six phenotypic characteristics (Table 2-3). The results show some suggestive evidence that both gene and locus can affect the phenotypic characteristics of area and mat strength. However, the only significant result under a Bonferroni adjustment was the effect of gene on mat strength.

Table 2-3: Multiple univariate analyses of effects on phenotypic characteristics.

Characteristic	Locus (*p*-value)	Gene (*p*-value)	Mutation Type (*p*-value)
Area	0.0229*	0.0189*	0.0779
Wrinkliness	0.6534	0.6001	0.2444
Mat strength	0.0078*	<.0001**	0.1173
Cellulose Binding	0.1233	0.3413	0.4314
Fitness (shaking)	0.0955	0.1404	0.1008
Fitness (3-way static)	0.4350	0.6099	0.5037

An asterisk* denotes a *p*-value less than 0.05. However, after Bonferroni adjustment ($\alpha = 0.05/18 = 0.00278$) only mat strength was significant, indicated with a double asterisk**.

2.2.3.1 Principal Components Analysis

Principal components analysis (PCA) is a method that reduces data with multiple dimensions to fewer dimensions by generating a linear combination of the descriptor variables that best summarise the variation in the data. This data set has six descriptor variables (colony area, colony circularity, cellulose production, mat strength, fitness in a shaking environment and fitness in a static environment) that describe 26 independent WS genotypes and SBW25-Δ*wspF*. Each of the six descriptor variables were weighted by their respective means and PCA was performed on those weighted variables in JMP 5.0.1 (S.A.S. Institute) using the covariance method to generate three principal components (PC1, PC2 and PC3). The Eigenvalues associated with the first three principal components show that 95.88% of the total variation is explained by the first three components. PC1, PC2 and PC3 were used to generate the axes of a three dimensional spinning plot and the values associated with each sequence plotted

in three dimensions for a visual analysis. Two dimensional scatter plots were constructed from the first two principal components (Eigenvalue = 92.574%) coloured by locus (Figure 2-16) and by gene (Figure 2-17).

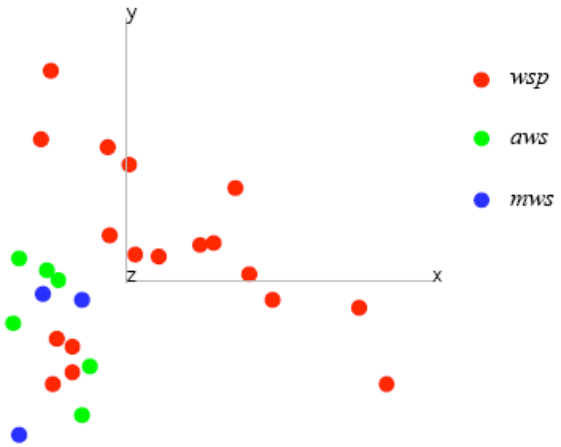

Figure 2-16: Scatter plot of PC1 against PC2 coloured by locus.
No clusters of a single colour can be determined.

The patterning observed in the first scatter plot (Figure 2-16) shows a very large spread for all *wsp* mutations with smaller clusters of *aws* and *mws*. In addition, each of the three clusters overlap with one another, which illustrates that the effect of locus is unlikely to be a powerful predictor of phenotype. However, the clustering in the second scatter plot (Figure 2-17) shows that *wspF* mutations can be differentiated from all other mutations based on the first two principal components, because they form a non-overlapping cluster (red circles). In addition, it is clear from the figure that if the data were collapsed in either dimension to look at either PC1 or PC2 on their own, that the distinct red cluster would no longer be observed.

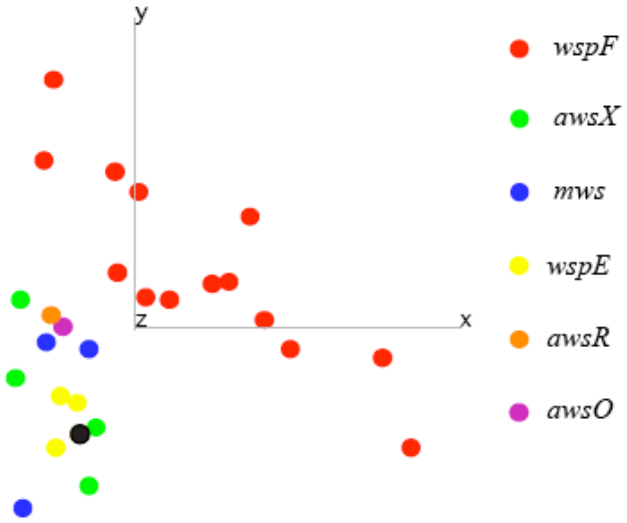

Figure 2-17: Scatter plot of PC1 against PC2, coloured by gene.
A non-overlapping, but variable cluster of red dots (*wspF* mutants) can be determined when coloured by gene. The black dot represents the unknown *wsp* mutation.

To understand which descriptor variables contributed to each of the principal components, the coefficients of each of the linear combinations of the scaled descriptor variables[1] in PC1 and PC2 and PC3 were examined (Table 2-4). These data showed that colony area has the largest coefficient for PC1, illustrating that colony area explains the largest proportion of the variation in PC1 while mat strength explains the largest portion of the variation in PC2.

Table 2-4: Coefficients of the first three principal components.

PC	Area	Circularity	Mat Strength	Cellulose Production	Fitness (shaking)	Fitness (3-way static)
1	0.9753	0.0006	0.1861	0.0006	-0.1192	0.0023
2	-0.1738	-0.042	0.9395	0.0016	0.0502	0.2876
3	0.0719	-0.4934	-0.2328	0.4114	0.2379	0.6865

All coefficients have been weighted by their respective means. Therefore, their value reflects a proportional contribution to that principal component.

To analyse for any significant effects among combinations of descriptor variables in PC1, PC2 or PC3, univariate analyses for the same factors as tested earlier (gene, locus, and mutation type) were performed, and the results of the ANOVA are shown

[1] Scaling the descriptor variables does not affect the PCA. However, if each descriptor variable were not equally weighted, the coefficients would have to be scaled after the PCA.

in Table 2-5. These data illustrate that none of the effects of locus, gene or mutation type were able to explain the variation in the first three principal components.

Table 2-5: Multiple univariate analyses of principal components.

Characteristic	Locus (*p*-value)	Gene (*p*-value)	Mutation Type (*p*-value)
PC 1	0.0159*	0.0083*	0.0660
PC 2	0.1991	0.0365*	0.4021
PC 3	0.2947	0.6608	0.2092

An asterisk* denotes a *p*-value less than 0.05. However, no *p*-values were significant after a Bonferroni adjustment ($\alpha = 0.05/9 = 0.0056$).

2.2.3.2 Multivariate Analyses of Similarity

The visual inspection of the spinning plots for the PCA data coloured by gene suggested that there might be significant association between certain combinations of the phenotypic effects and the underlying genetic changes. ANOSIM (analysis of similarity) is a multivariate method that tests whether genotypes of one group are on average more phenotypically similar than genotypes of another group (Clarke, 1993). The groupings are specified *a priori* to test a specific hypothesis, *e.g.* genotypes grouped by locus to test whether genotypes with a mutation at the same locus are significantly more similar to one another than those genotypes with mutations within other loci. ANOSIM generates a distance matrix by ranking all pairs of genotypes from most similar (lowest rank) to least similar (highest rank), and then calculates the difference in the mean rank of the between group comparisons to that of the within-group comparisons. Statistical significance was determined using a randomisation test with 1000 permutations. The implementation of ANOSIM was performed using the statistical software package, R (Ihaka and Gentleman, 1996), using the R-package, Vegan (Oksanen *et al.*, 2008)[1].

Three groupings of IWS genotypes were analysed by ANOSIM: gene, locus and '*wspF* vs non-*wspF*'. The results of the first ANOSIM (Figure 2-18) showed that there was no significant difference for the global comparison among all genotypes grouped by gene ($R = 0.095$, $p = 0.19$). In addition, none of the pair-wise comparisons

[1] Example R-code is shown in Appendix 9.2.4.

among genotypes showed statistical significance after Bonferroni correction (Table 2-6).

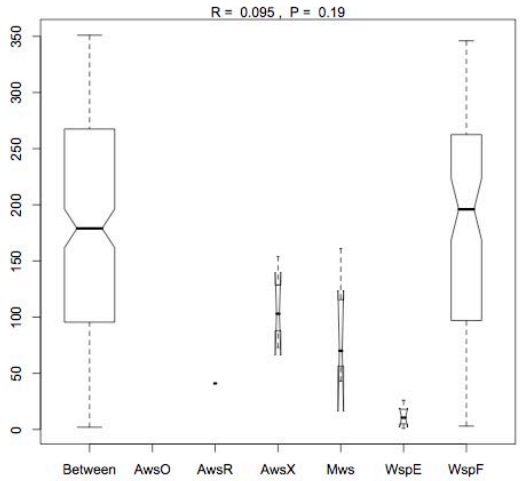

Figure 2-18: ANOSIM for the IWS genotypes grouped by gene.
There are no groups of two or more genotypes with mutations in *awsO*, so there is no ANOSIM value to be plotted for *awsO*. For *awsR* there is a one group of two or more genotypes, resulting in a single point and not a distribution. The width of the boxes in the plot represents the number in each group.

Table 2-6: Pair-wise ANOSIM comparisons for gene.

Comparison	*R* value	*p*-value
wspF* vs *awsX	0.274	0.037*
wspF* vs *wspE	0.135	0.222
wspF* vs *mws	0.309	0.035*
wspE* vs *awsX	0.241	0.104
wspE* vs *mws	0.148	0.377

No meaningful comparisons can be made for *awsO* and *awsR*, because *awsO* has zero groups of two or more genotypes, and *awsR* has just one group of two or more for comparison. * indicates a *p* value of < 0.05; however, the Bonferroni adjusted value of significance for multiple pair-wise comparisons is $\alpha = 0.05/5 = 0.01$.

The results of the ANOSIM for genotypes grouped by locus are displayed in Figure 2-19 and showed that there was no significant effect of locus for the global comparison ($R = -0.011$, $p = 0.19$). In addition, the pair-wise comparisons among loci showed no statistical significance after Bonferroni correction (Table 2-7).

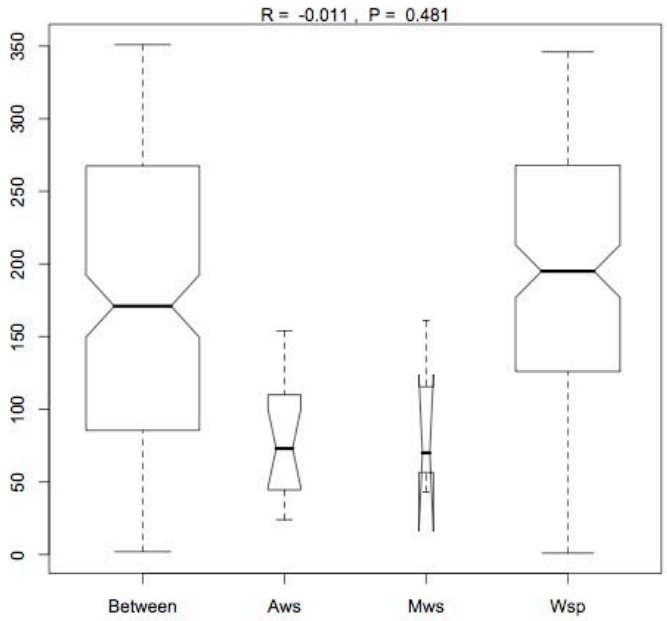

Figure 2-19: ANOSIM for the IWS genotypes grouped by locus.
There is no statistically significant effect of locus on the similarities among the IWS phenotypes. The width of the boxes in the plot represents the number in each group.

Table 2-7: Pair-wise ANOSIM comparisons for locus.

Comparison	*R* value	*p* value
aws* vs *mws	-0.056	0.596
aws* vs *wsp	0.043	0.286
mws* vs *wsp	0.015	0.413

There is no statistically significant effect of locus on the similarities among the IWS phenotypes. Bonferroni adjusted significance value is $\alpha = 0.05/3 = 0.0167$.

The final ANOSIM analysis was suggested from the results of the PCA labelled by gene, because it appeared that genotypes with a mutation in *wspF* formed a cluster on their own that may be significantly different from non-*wspF* mutants. An ANOSIM was performed on the data grouped by '*wspF* and non-*wspF*' (Figure 2-20) and the results showed that this effect was highly significant ($R = 0.566$, $p < 0.001$).

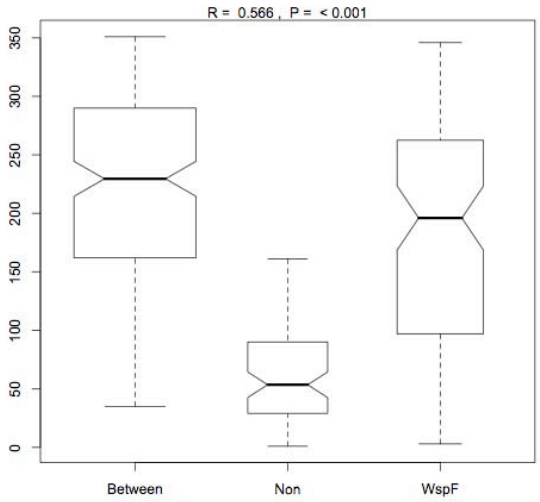

Figure 2-20: ANOSIM for the IWS genotypes grouped as *wspF* and non-*wspF*.
The differences between *wspF* and non-*wspF* mutants are significant. The width of the boxes in the plot represents the number in each group.

2.3 DISCUSSION

2.3.1 COMPLEMENTARY STRATEGIES CONFIRM THREE GENETIC ROUTES TO WS

Mutational effects can be examined using two complementary strategies. These are commonly called the 'bottom up' and 'top down' approach. The 'bottom up' approach involves the use of genetic manipulation techniques to deliberately introduce genetic changes into a population to determine function. These might range from subtle single nucleotide changes through to more blunt strategies involving transposon insertion to disrupt gene function. The consequences of those manipulations can typically be measured by phenotypic and fitness changes (Dean *et al.*, 1986). For example, the deliberate introduction of mutations into the various genes of the Wsp pathway have lead to numerous advances in our understanding of the functions of their protein products (Goymer *et al.*, 2006; Bantinaki *et al.*, 2007; Malone *et al.*, 2007). However, these approaches elucidate little of whether such mutations are likely to be observed spontaneously, in the lab or in the wild.

The 'top down' approach complements the 'bottom up' approach, by surveying the variants that occur in nature or the spontaneous mutants in the laboratory environment to determine causal relationships between genotype and phenotype. The set of 26 independently derived wrinkly spreaders (IWS) represents such a survey of the possible mutational routes to the WS phenotype during the adaptive radiation of wild-type *P. fluorescens* SBW25 in the laboratory. Transposon mutagenesis of the spontaneous mutants identified three of the 38 GGDEF/GGEEF domain proteins *wsp*, *aws* and *mws*. Therefore, the 'top down' approach suggests that only *wsp*, *aws* and *mws*, are involved in the WS phenotype.

To determine that no other loci were involved in generating the WS phenotype, this 'top down' survey of 26 spontaneous mutants in the laboratory was complemented by 'bottom up' approach of another study in which the wild-type SBW25, three single knock-out mutants (SBW25-Δ*wsp*, SBW25-Δ*aws*, SBW25-Δ*mws*), all combinations of double knock-out mutants (SBW25-Δ*wsp*Δ*aws*, SBW25-Δ*wsp*Δ*mws*, SBW25-Δ*aws*Δ*mws*) and the triple knock-out mutant (SBW25-Δ*wsp*Δ*aws*Δ*mws*),

were grown for up to a week at 28˚C to assess their genetic capability of evolving WS types (McDonald *et al.*, 2008). In the wild-type, the single knock-out mutants and all double knock-out mutants, WS phenotypes were detected after two days of incubation, but in the triple knock-out mutant, WS phenotypes failed to arise and increase to a detectable frequency until five days of incubation (McDonald *et al.*, 2008). This 'top down' approach provided additional support for the 'bottom up' data of the 26 IWS genotypes presented in this chapter, thus, in combination these data suggest that *wsp*, *aws* and *mws* operons represent the only single-step, spontaneous mutational routes to the WS phenotype available to wild-type SBW25.

2.3.2 Alternate Routes to WS

After the discovery that mutations in *wspF* were causal of the WS phenotype, a model of the Wsp pathway was developed drawing heavily from homology with the chemosensory system in *E. coli* (Section 1.5.2). This model has continued to be supported by further studies (Spiers *et al.*, 2002; Gehrig, 2005; Goymer *et al.*, 2006; Bantinaki *et al.*, 2007; Malone *et al.*, 2007; McDonald *et al.*, 2008) and confirms the role of WspR in increasing the c-di-GMP levels, by virtue of a GGDEF/GGEEF domain. Thirty-seven other open reading frames share this domain, however, only three have been implicated in the WS phenotype. One suggestion as to why only these three loci may be capable of generating the WS phenotype is that they may be the only loci that also contain a negative regulator of the GGDEF/GGEEF domain protein (a hypothesis for the role of adaptation is examined later in the discussion, Section 2.3.5). It is has been determined that WspR, AwsR and MwsR are all negatively regulated (Gehrig, 2005; McDonald *et al.*, 2008), and in the case of MwsR it is suggested that its DGC activity of $MwsR_{GGDEF}$ is regulated by the PDE activity of $MwsR_{EAL}$. Sequence analysis of the open reading frames has identified 16 other GGDEF/GGEEF domain proteins that also have an EAL domain (Appendix 9.2.1) indicating that they could function similarly to MwsR, however none of these genes have ever been implicated in WS phenotype. Combined with the evidence from Jenal and Malone (2006) that DGC/PDE combinations are not known to have DGC activity, these data suggest that MwsR may be unique in possessing both DGC and PDE activity. Future biochemical studies investigating DGC activity for the 16 DGC/PDE proteins in *P. fluorescens* would provide additional evidence in favour of

this hypothesis. Therefore, to conclusively determine why these three loci alone account for all the spontaneous mutations to the WS phenotype observed in the laboratory still requires extensive investigation.

2.3.3 LOCATION OF MUTATIONS IN *MWSR*

Based on conserved domain matching (Section 1.5.4) and preliminary genetic studies (McDonald *et al.*, 2008) the working model for the function of MwsR suggests that $MwsR_{GGDEF}$ acts to increase the levels of c-di-GMP and that $MwsR_{EAL}$ acts to lower the levels by breaking down the c-di-GMP. This model predicts that the WS phenotype will be caused by a loss of the PDE activity of $MwsR_{EAL}$, but still requires functional DCG activity of $MwsR_{GGDEF}$ to generate the WS phenotype. Both of these predictions of the model have been confirmed using a series of deletion mutants; deletion of $MwsR_{EAL}$ in the wild-type leads to the WS phenotype, deletion of the entirety of MwsR in the wild-type leads to no change of the SM morphology, and deletion of MwsR in the original *mws* WS mutant reverted the phenotype from WS to SM (McDonald *et al.*, 2008). These data suggest that MwsR may be unique in possessing both DGC and PDE activity.

Mutations that disrupt PDE activity were expected to occur more frequently within $MwsR_{EAL}$, though mutations elsewhere within the gene could have a similar effect on protein function. Of the three mutations found in MwsR, one was outside the PAS domain, but inside the multi-domain PRK11359, one was in the EAL domain, and the third was between GGDEF and EAL domains. Another analysis of the distribution of mutations in *mwsR* has shown 7/20 in $MwsR_{GGDEF}$, 2/20 in $MwsR_{EAL}$ and 11/20 between the GGDEF and EAL domains (Figure 2-21; McDonald *et al*., 2008)[1].

[1] In this analysis, only the GGDEF and EAL domains of 28 of which 20 had mutations in either the GGDEF or EAL domains. The mutational background was a mixture of either wild-type or the triple mutant, SBW25-Δ*wsp*Δ*aws*Δ*mws*.

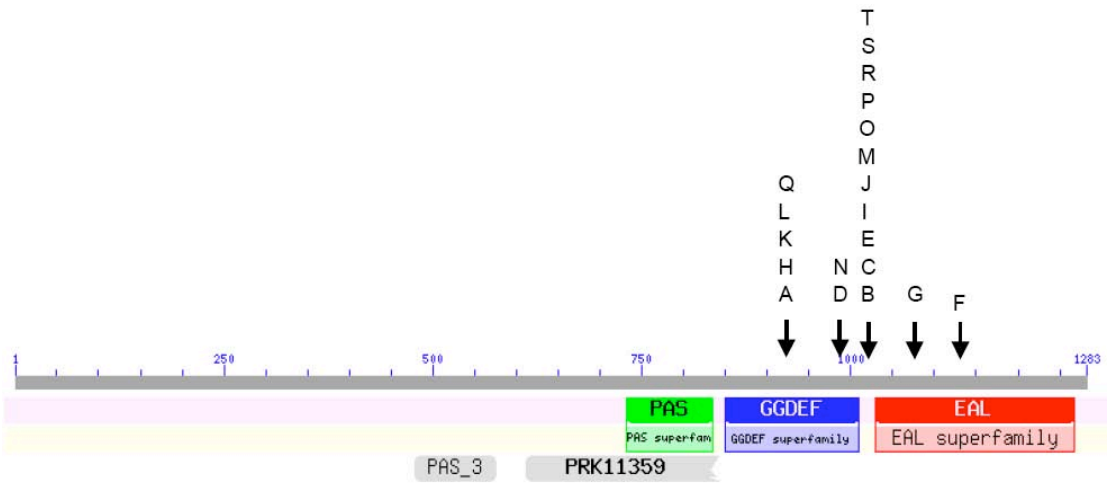

Figure 2-21: Position of the *mws* mutations from McDonald *et al.* (2008).
A-T represent the 20 mws mutations in the McDonald *et al.* (2008) study.

There are two interesting aspects to these results. Firstly, although not part of the EAL conserved domain, those mutations found between the GGDEF and EAL domains are all at amino acid residues that are conserved among GGDEF/EAL domain proteins with PDE activity (McDonald, unpublished results). This emphasises the problem that functionally important loci may not always be captured by *in silico* analyses. Secondly, it is easily conceded that the mutations between the GGDEF and EAL domains are likely abolishing the PDE function of $MwsR_{EAL}$; however, it is harder to reconcile that this is the case for the eight[1] mutations found in the GGDEF domain. Given the working model for Mws that has been developed (McDonald *et al.*, 2008), one of two hypotheses must be correct for the mutations found in the GGDEF domain: either the mutations in the GGDEF domain abolish the activity of $MwsR_{EAL}$ and do not abolish the activity of $MwsR_{GGDEF}$, or they increase the DGC activity of $MwsR_{GGDEF}$ relative to the PDE activity of $MwsR_{EAL}$. The eight mutations within the GGDEF domain are all single nucleotide changes that occur at only three nucleotide positions *i.e* independent identical mutations at specific nucleotide positions. Unfortunately, these hypotheses cannot be distinguished from one another given the current data, because both predict that only few nucleotide positions could be altered to generate the WS phenotype. Further empirical molecular and genetic studies into the balance between the DCG activity of $MwsR_{GGDEF}$ and the PDE

[1] One from the IWS genotypes in this chapter and seven from McDonald *et al.* (2008)

activity of $MwsR_{EAL}$ are required to fully understand the role of this apparently unique dual functioning DCG/PDE protein.

2.3.4 The Significance of Mutations in *wspE*

Most mutations are loss-of-function rather than gain-of-function mutations, because there are significantly more ways in which a mutation can abolish the function of a protein than there are to confer new or enhanced activity to it. Previous work on the IWS genotypes revealed that 13 out of 26 of the IWS genotypes had mutations in *wspF*, that lowered or abolished WspF activity (Bantinaki *et al.*, 2007). Deleterious mutations in *wspF* represent a large mutational target for generating the WS phenotype, because abolishing the function of a negative regulator acts to increase WspR activity. There are no other negative regulators in the Wsp pathway, so it was anticipated that the remaining mutations among the IWS genotypes would be found at loci other than the *wsp* operon. Counter to that prediction, in addition to identifying *aws* and *mws*, the results of the transposon-mediated mutagenesis strategy identified the *wsp* operon four times. Subsequent sequencing of the entire *wsp* operon revealed that the precise mutations for three of the genotypes were in *wspE*, while one was unaccounted for. The novelty of finding *wspE* mutations demanded a more detailed look at the Wsp pathway to determine whether such mutations were compatible with the working model (Figure 2-22).

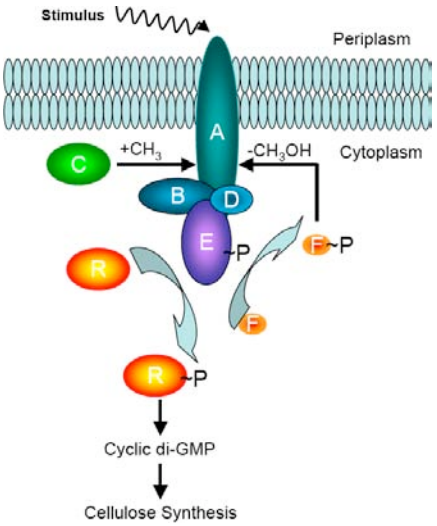

Figure 2-22: The Wsp pathway.
WspA (the MCP) receives a stimulus that causes WspE to become autophosphorylated. In turn, WspE causes the phosphorylation of WspF and WspR. WspR allosterically activates c-diGMP leading to cellulose biosynthsis while WspF negatively regulates activity in the Wsp pathway by removing methanol from WspA ending the signal. WspC acts antagonistically to WspF by adding methyl groups to WspA. This is a smaller reproduction of the Wsp pathway pictured in Figure 1-10.

WspE has highest overall similarity with FrzE from *M. xanthus*. Both are approximately 750 amino acids long, and have CheA N-terminal domains and CheY C-terminal domains. Fortunately, significant biochemical studies have been done on both the Che chemosensory system in *E. coli* and on the Frz pathway that controls motility in *M. xanthus*. CheA and CheY belong to a family of two-component regulatory systems, where CheA is a histidine kinase that autophosphorylates in the presence of ATP and divalent cations, and then transfers the phosphate group to an aspartate residue of CheY (Parkinson and Kofoid, 1992). Phospho-CheY then interacts with the components of the flagellar motor resulting in a change of rotation of the flagella. In *M. xanthus* biochemical and genetic studies have confirmed that the two domains of FrzE are part of a single 83 kD protein, and is not translated as two smaller peptides or processed post-translationally into two smaller peptides (McCleary and Zusman, 1990b). In addition, it was shown that $FrzE_{CheA}$ can autophosphorylate, and that the phosphate can subsequently be transferred to the $FrzE_{CheY}$, though it can also be transferred to other proteins with a two-component receiver domain (Acuña *et al.*, 1995; Li *et al.*, 2005). Therefore, the working model suggests that $FrzE_{CheY}$ effectively competes with the CheY domains of other proteins (FrzZ is the likely candidate) for the phosphate present on $FrzE_{CheA}$ even though the downstream effects of the pathway are not yet known (Li *et al.*, 2005).

It is anticipated that given the sequence identity with FrzE, WspE will function similarly, *i.e.* that intramolecular phosphotransfer occurs between $WspE_{CheA}$ and $WspE_{CheY}$ and that $WspE_{CheY}$ competes with the CheY-like domains of WspF and WspR for the phosphate present on $WspE_{CheA}$. Though the mutations did not affect the principle phosphate-accepting residue (Asp67 in CheY, Asp709 in FrzE and Asp682 in WspE) and it is unclear how these mutations lead to the WS phenotype, they were all at highly conserved residues within WspE. However, one possible explanation is that if $WspE_{CheY}$ competes with WspR and WspF for the phosphate on $WspE_{CheA}$ as suggested by the FrzE working model, then a loss of ability to accept phosphate could alter the available phosphate for WspR and WspF leading to the downstream consequences characteristic of the WS phenotype. This emphasises a role for $WspE_{CheY}$ in balancing the available phosphate for WspR and WspF.

2.3.5 PHENOTYPIC PARALLELISM AMONG THE WS GENOTYPES

The diversity in nature is the product of three fundamental evolutionary forces: adaptation, chance and history. The relative influences of these forces have often been the subject of intense debate in evolutionary biology (Kimura, 1968; Gould and Lewontin, 1979; Maynard Smith *et al.*, 1985; Williams, 1992). The prevailing view is that adaptation is the most dominant among these forces (Williams, 1966; Maynard Smith *et al.*, 1985; Williams, 1992), though this has prompted other authors to emphasize the roles of chance and history (Kimura, 1968; Gould and Lewontin, 1979; Gould, 2002). Chance effects are usually invoked at the molecular level, *e.g.* mutation and genetic drift, but are also important at the phenotypic level, because beneficial mutations that arise at random may be lost soon-after they appear, even in large populations (Travisano *et al.*, 1995a). In addition, the role of history is crucial, because the ancestral background can constrain or promote particular evolutionary outcomes, limiting potential adaptations. To address the roles of these various forces, Gould presented a thought experiment about 'rewinding life's tape' (Gould, 1989) in which the repeatability of evolution might be tested. Although one cannot actually do such an experiment, rigorous experiments of shorter duration in simple environments can be performed in parallel to investigate the roles of adaptation, chance and history in evolution.

Parallel experiments in static microcosms inoculated with *P. fluorescens* SBW25 repeatedly and reliably lead to the evolution of the WS phenotype, which occupies the niche and the air-liquid interface. Such parallel phenotypic evolution is a hallmark of adaptive evolution (Simpson, 1953; Stewart *et al.*, 1987; Travisano *et al.*, 1995a; Wichman *et al.*, 1999) and suggests that the phenotypic outcome of adaptive radiation is broadly reproducible for genetically identical organisms capable of growing to large population sizes, under strong selection and inhabiting similar environments (Rainey and Travisano, 1998; MacLean and Bell, 2003). The extent of this parallelism may continue further towards the genotypic level, but this is determined by the complexity of the relationship between genotype and phenotype. If the adaptive niche-specialist phenotype (WS) is produced only by a limited number of genotypes, then adaptive radiation should be reproducible at the genotypic level;

however, if many genotypes can produce the same phenotype, then chance and history may play a dominant role (MacLean and Bell, 2003).

The WS phenotype is caused by mutations at any one of three loci – of the 26 WS genotypes, 17 were in *wsp*, six in *aws* and three in *mws*. This indicates parallelism at the level of the operon, because the same loci were observed repeatedly. There were lower levels of parallelism at the level of the gene with 13 mutations in *wspF*, three in *wspE*, two in *awsR*, three in *awsX*, and three in *mwsR* and even one instance of parallelism at the level of the amino acid codon for IWS_D and IWS_I at position 638. Does natural selection acting on chance variation (adaptation) drive this parallelism or is it the result of constraints imposed by the history of the ancestor?

An adaptive explanation for parallelism of the WS phenotype would predict that specific mutations or mutations within certain genes or loci that confer greater fitness benefits are observed more frequently than those that confer lower fitness. There are two ways of using the available data to look for a possible role of adaptation in the WS phenotype, which will be examined separately. The first looks for a role of adaptation in the observed over-representation of *wspF* mutants among the IWS genotypes. The second looks for a role of adaptation in explaining why types other than *wsp*, *aws* and *mws* mutants are not observed.

Firstly, among the IWS genotypes *wspF* mutations accounted for half of the mutations; therefore, if *wspF* mutations were adaptive, their over-representation in the distribution of mutations that lead to the WS phenotype would be as a consequence of having higher fitness than genotypes with mutations at other loci. However, an ANOVA[1] of the static fitness data showed that *wspF* mutants do not have higher fitness than other genotypes ($F_{1,188}$ = 0.5892, p = 0.4437). Therefore, the high proportion of *wspF* mutants is not likely to be a consequence of adaptation.

Secondly, among the IWS genotypes, mutations are observed at only three loci; therefore, if other possible adaptations to the air-liquid interface exist, they are not observed because they are out-competed by any of the better-adapted WS genotypes.

[1] 'IWS genotype' was specified as a random factor in the ANOVA.

Though there are undoubtedly many, at least two such genotypes are known, the slow wrinkly spreader (SWS, McDonald *et al.*, 2008) and a bistable, switching phenotype (Beaumont *et al.*, 2008). The SWS was evolved from the strain that had all spontaneous mutational routes to WS deleted (SBW25-Δ*wsp*Δ*aws*Δ*mws*). It arose and increased in frequency to detectable levels in a static environment and occupies the niche at the air-liquid interface (McDonald *et al.*, 2008); however, it required five days of incubation before it was detected compared to two for WS genotypes. This delay indicates that SWS likely requires multiple mutations before giving rise to a phenotype capable of occupying the air-liquid interface, and is evidence for why it was never observed in the wild-type genetic background. Similarly, the bistable, switching phenotype arose in a genetic background other than the ancestor – a genotype derived after four repeats of a selection regime that was altered between a static and shaken environment (Beaumont *et al.*, 2008). Against this genetic background, the bistable, switching phenotype increased in frequency to occupy the niche at the air-liquid interface, but when the causal mutation was reconstructed in the wild-type background this type was less fit than the ancestor (Beaumont *et al.*, 2008), which explains why this type does not evolve in the wild-type background. Therefore, adaptation appears to play a role in the success of genotypes with a mutation at any of *wsp*, *aws* and *mws*, because they are consistently favoured by selection in a static environment.

2.3.6 Refining the Genotype-phenotype Map of the WS Phenotype

The principle phenotype that defines all WS genotypes is their ability to occupy the niche at the air-liquid interface. Other phenotypes of all known WS genotypes include 'clumping' in shaken cultures and the characteristic wrinkly appearance on agar plates; therefore, in a naïve or global sense, the genotype-phenotype map of WS was known, *i.e.* introduce a mutation to SBW25 that appropriately disrupts the Wsp, Aws or Mws pathways and you will generate the WS phenotype; and it will be able to form mats at the air-liquid interface, clump in shaking cultures and appear wrinkly on agar plates. However, in a deeper sense, there are at least three different pathways at the molecular level that give rise to the 'mat forming, clumping, wrinkly' phenotype

and there can conceivably be differences among those types and theoretically also within those types.

Each of the six phenotypic variables showed significant variation among the set of IWS genotypes; however, individually they were not good predictors of genotype, and with the exception of mat strength, there was no phenotype that made statistically significant predictions of genotype. Despite the failure of univariate analyses on these variables to encapsulate the differences statistically that were observable intuitively, the PCA analysis suggested that the large cluster of genotypes with mutations in *wspF* might be significantly different from other types. This was confirmed using an ANOSIM, which showed that while the effect of gene and locus were non-significant, the delineation between *wspF* and non-*wspF* mutants was statistically significant.

Being able to show measurable phenotypic differences between *wspF* and non-*wspF* mutants shows a partial success of the attempt to determine the genotype-phenotype map at a level below ecological niche-preference. Although mutations within certain genes were phenotypically indistinguishable from others, showing a phenotypic distinction between *wspF* and non-*wspF* is a valuable contribution. This distinction provides evidence that molecular level differences can be unravelled by observing subtle differences in phenotypes, while simultaneously showing how challenging such a task is in practice. In addition, it suggests that, of the molecular level similarities and differences that can be determined phenotypically, *wspF* mutants form the most unified group.

Two biological explanations have been made to account for the observed grouping of *wspF* and non-*wspF*. The first explanation suggests that the pathway in *wspF* mutants interact with the c-di-GMP regulation network and the subsequent cellulose machinery differently from other mutant genotypes. Controlled synthesis and degradation of c-di-GMP has come to be recognised as a ubiquitous secondary messenger system in bacteria. The c-di-GMP-signalling network links environmental stimuli to the specific adaptive responses, primarily cell motility and biofilm formation (Jenal and Malone, 2006). The c-di-GMP signalling appears highly complex, and although it is known to operate at the transcriptional, translational and the posttranslational levels (Christen *et al.*, 2006; Weber *et al.*, 2006), almost every

major aspect of c-di-GMP signalling remains largely uncharacterised (Jenal and Malone, 2006). Of particular importance to the suggestion that WspR/c-di-GMP signalling is different from other c-di-GMP signalling in the WS phenotype, is understanding how a molecule that is so ubiquitous can successfully affect certain c-di-GMP-dependent functions and not others.

One suggestion is sequestration, in which distinct c-di-GMP circuits within a cell could be separated either in time through differential expression, *e.g.* tightly regulated expression of low number of molecules (Jenal, 2004) or in space through compartmentalised signalling, *e.g.* separation into microcompartments of larger complexes (Paul *et al.*, 2004). The model for spatial separation into microcompartments has the advantage of coupling specific signalling input directly with downstream targets. In addition, microcompartments are consistent with the observation that only three out of 38 GGDEF proteins are capable of generating the WS phenotype, because increasing expression is only sufficient for generating the WS phenotype for circuits that are appropriately localised.

The second biological explanation suggests that pleiotropic consequences of *wspF* mutations might be responsible for a suite of similar phenotypic effects that are not present for other WS mutants. Pleiotropic effects of niche specialisation have been shown to reduce fitness of WS in terms of reduced carbon catabolism on Biolog[1] plates (MacLean *et al.*, 2004). In this study, catabolic defects as a result of the WS mutations were not randomly distributed across independent WS types, but rather tended to occur on a narrow range of substrates. In contrast, the pleiotropic effects of six independent WS genotypes have been characterised using protein expression levels and showed variation among independent evolutions of the same phenotype for 69% of proteins (Knight *et al.*, 2006). This suggested that the protein expression changes were specific to the genotype, even among genotypes with different

[1] Biolog plates are 96-well microtitre plates in which 95 wells contains a unique carbon source and an indicator dye for growth with one well as a carbon-free control.

mutations in the same gene, *wspF*, which were the most divergent[1]. It is unclear from these contrasting results whether or not the effects of pleiotropy associated with the WS phenotype could be responsible for affecting a suite of phenotypic effects leading to the distinction of *wspF* genotypes from non-*wspF*. However, for both studies, but particularly the MacLean *et al.* (2004) study, where 20 WS genotypes were examined, there is significant potential to extend the analyses to examine genotype-phenotype mapping among WS genotypes, because only three loci would require sequencing to determine all the exact mutations.

2.3.6.1 Problems with Measuring Wrinkliness

Without any form of magnification, it is clear that independent WS colonies differ qualitatively in their degree of wrinkliness, and measuring circularity was an attempt to quantify this observation. The circularity was a comparison of the true perimeter (measured) to the perimeter estimated as if the object were a circle. During the circularity analysis of the digital photos for the IWS genotypes, some outliers were identified that appeared from visual inspection not to be that different from the others. After inspecting the outliers and reconsidering the measure of wrinkliness, there appear to be two confounding factors that may lead to a lack of reliability when using this measure of circularity to quantify wrinkliness. The measurement of expected perimeter correctly increases positively as a function of wrinkliness of the edges; however, it can also increase positively as a result of elongation of the overall shape of the colony (*i.e.* if the colony is more elliptical rather than circular). This appears to be the source of the variation among the outliers. Furthermore, the circularity of colonies is affected by the efficacy of analysing the wrinkliness of colonies that are three-dimensional in two dimensions. Colonies that are wrinkled in the vertical plane rather than on the horizontal plane will be scored as having low circularity when in fact they are highly wrinkled. Taking this limitation into consideration, it may be that circularity of colonies in digital photos is not appropriate as a measure that quantifies wrinkliness in three dimensions.

---

[1] Three of the independent genotypes were known to have mutations in *wspF* (IWS_A, IWS_B and IWS_Y) and three were unknown (IWS_D, IWS_V, IWS_X). Two of the wsp

2.3.6.2 Fitness in Two Environments

The fitness of WS types was compared in two different environments, a shaking environment and a static environment. All WS types have a fitness trade-off compared to SM between these two environments due to the niche-specialisation of WS types to the air-liquid interface in the static environment, and the magnitude of this trade-off was expected to differ among WS types. Two hypotheses were investigated: first, that there was a positive correlation for fitness between the two environments, because mutations to the WS phenotype that conferred the greatest fitness in the static environment also conferred the greatest fitness in the shaking environment, and second, that there was a negative correlation for fitness between the two environments, because mutations to the WS phenotype that conferred the greatest fitness in the static environment conferred the lowest fitness increase in the shaking environment due to a trade-off. The results from a regression analysis of the fitness between the two environments supported neither of these hypotheses, but showed instead that the fitnesses were randomly distributed with neither a positive nor a negative correlation.

Knowing that fitness in each environment on its own is not systematically affected by either locus or gene (Section 2.2.3), three biological factors have been suggested to contribute to the observed lack of fit to either hypothesis. The first contributing factor simply might be that the two environments are not actually sufficiently different from one another for the anticipated trade-off to exist within WS genotypes. The second contributing factor is the role of pleiotropic effects (MacLean *et al.*, 2004; Knight *et al.*, 2006), which has been discussed earlier as a potentially significant effect in a slightly different context. The third factor is that SBW25 is effectively in a 'fitness valley' of an adaptive landscape (Wright, 1932), because the nutrient rich microcosm is a completely novel environment. This fitness valley is expected to be surrounded by multiple local peaks in a rugged environment (Rainey and Travisano, 1998; Rozen and Lenski, 2000), and thus significant trade-offs between adaptive peaks may not exist until further adaptation has occurred over longer incubation periods than seven days.

3 A TEST OF WS COOPERATION USING THE HAYSTACK MODEL

We are beginning to see that the awesome wonder of the evolution from amoeba to man - for it is without a doubt an awesome wonder - was not the result of a mighty word from a creator, but of a combination of small, apparently insignificant processes. The structural change occurring in a molecule within a chromosome, the result of a struggle over food between two animals, the reproduction and feeding of young - such are the simple elements that together, in the course of millions of years, created the great wonder. This is nothing separate from ordinary life. The wonder is in our everyday world, if only we have the ability to see it.

-- Alfvén, H. 1969

3.1 INTRODUCTION

Previous work by Rainey and Rainey (2003) showed that the WS phenotype satisfies the criteria for being considered a cooperative trait, a trait that evolved *de novo* in the laboratory (Section 1.4.1.1). Firstly, there is a cost to cooperation, because WS types have a relative fitness of 0.8 relative to ancestral SBW25 when incubated statically with abundant resources (Rainey and Rainey, 2003). Secondly, this cost is offset by benefits provided by an emergent group-level phenotype – the mat occupying a novel niche at the air-liquid interface, as evidenced by the increase in frequency of WS types in a structured environment as oxygen becomes limiting. Thirdly, selfish ancestral-like SM types arise in microcosms founded by WS types that are fitter in the presence of WS types than in their absence. Finally, the emergent group-level phenotype (*i.e.* the mat) is destroyed as the number of defecting SM types increases too high in frequency – a 'tragedy of the commons' (Hardin, 1968).

MLS 1 theory predicts that under certain conditions, group structure can favour the evolution and maintenance of cooperative *individuals* (Wright, 1945; Wilson, D.S., 1975; Damuth and Heisler, 1988; Sober and Wilson, 1998; Okasha, 2006a). The Haystack model (Maynard Smith, 1964), has commonly been used to describe the procedure for experimental studies in an MLS 1 framework. In the Haystack model, it is supposed that a fictitious population of mice live entirely in haystacks. A haystack is always colonised by a single fertilised female and her offspring live in the haystack until the next year (one group-generation[1]). At the next year, new haystacks become available, and the mice disperse, mate and compete to colonise the newly available haystacks. Maynard Smith introduced the idea to the literature for the purpose of distinguishing 'group selection' (*i.e.* differential proliferation and extinction of groups) from kin selection to show that 'group selection' of this sort was "too artificial to be worth pursuing further" (Maynard Smith, 1964, p.1146). However, his original model in 1964 was a mixture of MLS 1 and MLS 2 concepts, and assumed a

[1] A group-generation in an MLS 1 context is the period of time between successive formations of groups, *i.e.* the amount of time spent within a haystack.

worst-case scenario that is biologically restrictive – that cooperative types would always become extinct within mixed groups (haystacks) of cooperators and selfish types before the end of the group-generation. This assumption has been relaxed in future examples of the Haystack model to be more biologically realistic and to fit precisely into an MLS 1 framework, such that cooperators only decrease in proportion within a group during a group-generation (Wilson, 1987). Therefore, the experimental procedure that is described by the Haystack model, in an MLS 1 framework predicts that, while cooperators are less fit within a group, they may increase in frequency within the overall population if groups with more cooperators contribute a disproportionately high number of individuals to the next group-generation (Damuth and Heisler, 1988; Sober and Wilson, 1998; Gould, 2002; Okasha, 2006a). In essence, MLS 1 theory predicts that group structure will increase the spread of a cooperative trait through a population whenever there is a positive relationship between the number of cooperators and the number of individuals within a group.

The standard experimental microcosm can be exploited as an investigative tool to test the Haystack model, because it is a simple and effective way of dividing populations of cooperators and selfish types into groups[1] (the haystacks). To test the Haystack model in an MLS 1 framework, the phenotypic characteristics of *individuals* must be measured, and the fitness of those *individuals* must depend on their organisation into groups. Individuals within each group (microcosm) may be either cooperators (WS types) or selfish types (SM types). All groups within a population are founded with ancestral SBW25 (SM type), and therefore all WS types arise *de novo*. Crucially, stochastic processes guarantee that, by chance, different WS genotypes will arise in each microcosm and, that different microcosms will vary in their proportion of WS types. Thus, group membership will differentially affect the fitness of individuals within different groups. As in the Haystack model, group structure can periodically be

[1] There is potential for confusion here, because two entities, mats at the air-liquid interface and microcosms are commonly referred to as groups. In this chapter a group refers to a microcosm containing both WS and SM types, or a haystack in the context of in the Haystack model, and groups of WS cells at the air-liquid interface will only be referred to as mats.

removed to allow the mixing of individuals from different groups (microcosms). This affords an opportunity to *individuals* within more successful groups to contribute disproportionately to the next generation of groups (MLS 1). The proportion of WS *individuals* in a population with this group structure and dispersal (experimental treatment) can be contrasted with the proportion of WS *individuals* in a population in which individuals from different groups are never afforded the opportunity to mix (control treatment). This provides us with an experiment to investigate the hypothesis that group-structure enhances cooperation among the WS types of the *P. fluorescens* system.

This work extends the results of Rainey and Rainey (2003) by studying the dynamics of the *P. fluorescens* system to achieve three main aims. Firstly, SBW25 and the neutrally marked SBW25-*lacZ* strain were grown statically for ten days to verify that the diversification of genotypes via adaptive radiation is conserved for the marked strain (Section 3.2.1). The diversification pattern of genotypes was used to determine the appropriate durations for the group-generations of the selection regime (72 h and 24 h) to use in the multi-level selection experiments. Secondly, the WS system was used to test the Haystack model using a 72-h group-generation time, which predicts that the cooperative WS types will be favoured when groups are afforded the opportunity to mix (Section 3.2.2). Thirdly, the group-generation time was shortened to 24 h so that SM types were favoured within each microcosm to test the Haystack model under conditions that were expected to disfavour the evolution of cooperative WS types (Section 3.2.3).

3.2 RESULTS

3.2.1 THE DIVERSIFICATION OF NEUTRALLY MARKED SBW25-*LACZ* IS CONGRUENT WITH WILD-TYPE SBW25

P. fluorescens SBW25 rapidly diversifies in spatially structured microcosms into a variety of niche-specialist genotypes (Rainey and Travisano, 1998). A marked variant, SBW25-*lacZ*, had been shown to have no measurable cost associated with the marker on multiple substrates, and *in planta* (Zhang and Rainey, 2007). In addition, the marked strain SBW25-*lacZ* was grown under continuous, shaking, log-phase growth conditions for ten days to verify its long-term neutrality with respect to fitness (Appendix 9.1.1). Imperative to long-term experiments in which strains have time to evolve and adapt is that the marked strain shows no deviation in its ecology and although SBW25-*lacZ* was expected to undergo a similar adaptive radiation in a spatially structured environment, this had never been shown. Therefore, by way of a preliminary experiment, it was crucial to show that the ecology of this adaptive radiation was the same in both SBW25 and SBW25-*lacZ*. Each strain was inoculated separately into eight replicate lines of static microcosms and left to grow for between one and ten days. At each 24-h time point, the microcosms were sampled destructively by vortexing, diluted and plated on KB plates (or KB+Xgal). The comparison between the adaptive radiations of SBW25 (white colonies) and SBW25-*lacZ* (blue colonies) is shown in Figure 3-1.

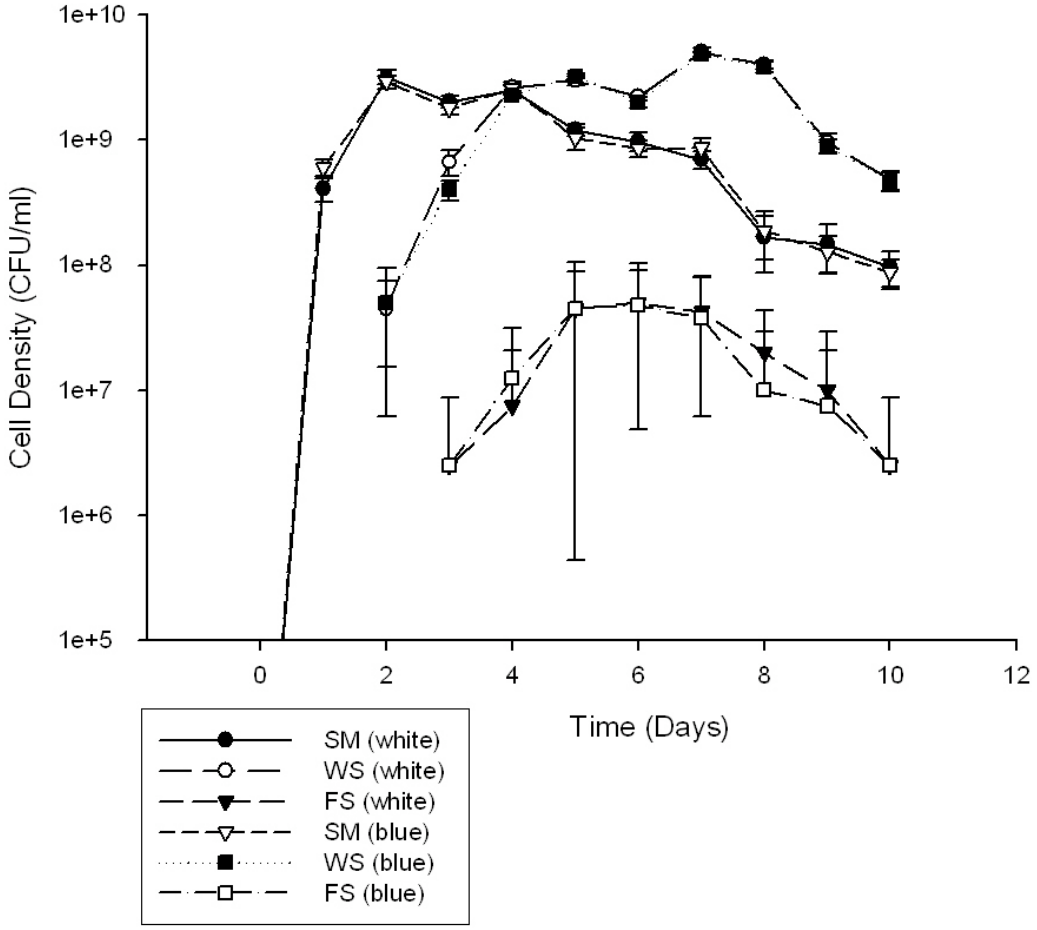

Figure 3-1: Diversification of SBW25 and SBW25-*lacZ* in a spatially structured microcosm. SBW25 (white) and SBW25-*lacZ* (blue) were grown statically for ten days. WS types arise at day two and increase in frequency to become the dominant type after four days. FS types arise on day three and are only occasionally detected above the 0.5% level (counting >200 colonies). Error bars represent 95% confidence limits.

To confirm that there was no statistical difference in the adaptive radiation between the marked and unmarked strains, a variant of a randomisation test[1] (Methods 7.2.23.1) was developed to test for a difference between the curves of SBW25 and SBW25-*lacZ*. The data for SBW25 and SBW25-*lacZ* were grouped together for both the SM and the WS morphotypes[2]. A sixth degree polynomial function was fit to each

[1] Thanks to Allen Rodrigo for the suggestion of this randomisation test.

[2] The FS data was ignored because it was commonly below the detection threshold during the experiment leading to erratic and untestable data.

of the combined data sets (Figure 3-2). The proportion of variability explained by the sixth degree polynomials for SM and WS were $r^2 = 0.861$ and $r^2 = 0.859$ respectively. The residuals from the fit of the polynomial to the data were calculated for further analysis. An ANOVA of the residuals with model effects "blue/white", "time" and "blue/white x time" revealed that there was no significant difference between marked and unmarked SM types during the ten-day adaptive radiation in static microcosms ($F_{1,156} = 0.4321$, $p = 0.5119$) and that the interaction effect of "blue/white x time" was also not significant for SM types ($F_{1,156} = 0.1826$, $p = 0.6698$). An ANOVA of the derived WS types with the same model effects as for SM also revealed that there was no significant difference between the marked and the unmarked derived WS types ($F_{1,156} = 1.0704$, $p = 0.3025$) and that the interaction effect of "blue/white x time" was again not significant ($F_{1,156} = 0.0019$, $p = 0.965$).

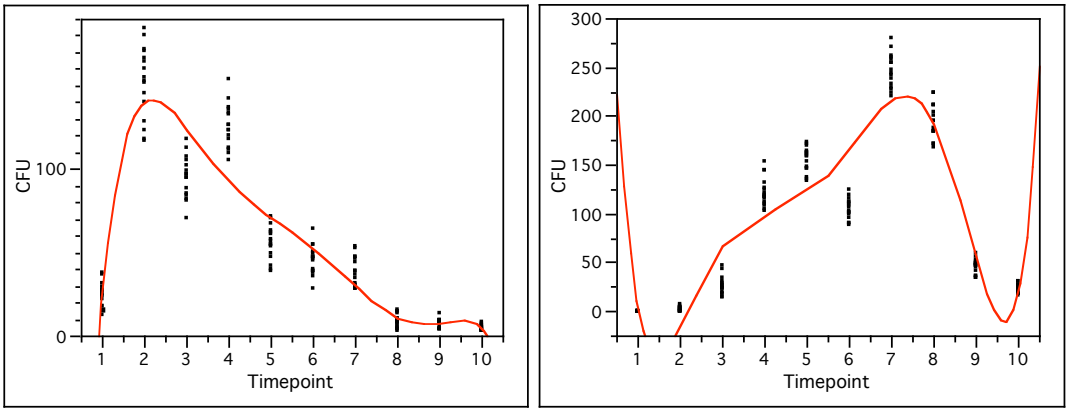

Figure 3-2: Sixth degree polynomial fit to the data.
The graph on the left is for the combined (blue and white) SM data, and the graph on the right is for the combined WS data. Red lines are the sixth degree polynomial fit to the data.

3.2.2 MLS 1 WITH 72-H GROUP-GENERATION TIME

Two populations, each consisting of 12 groups (microcosms), were studied in a long-term experiment totalling 36 days of experimentation (Section 7.2.16). Each group (microcosm) was founded with an SM type (one with SBW25 and the other with SBW25-*lacZ*) and incubated statically for 72 h. One of the populations was evolved under the experimental treatment (Figure 3-3) and the other population was evolved as a control experiment (Figure 3-4). The difference between the experimental and the control treatments is the opportunity for competition among groups. In the experimental treatment, equal volumes of each the 12 groups were mixed together following the 72-h static incubation. Samples were taken to estimate the proportion of

WS types in the mixed population and 12 new groups were founded from the mixed population to begin the next group-generation. The ten-day adaptive radiation experiment showed that 72 h of static incubation in a microcosm is sufficient to favour the evolution of WS (Figure 3-1). Therefore, WS types were expected to evolve within each microcosm within the first group-generation.

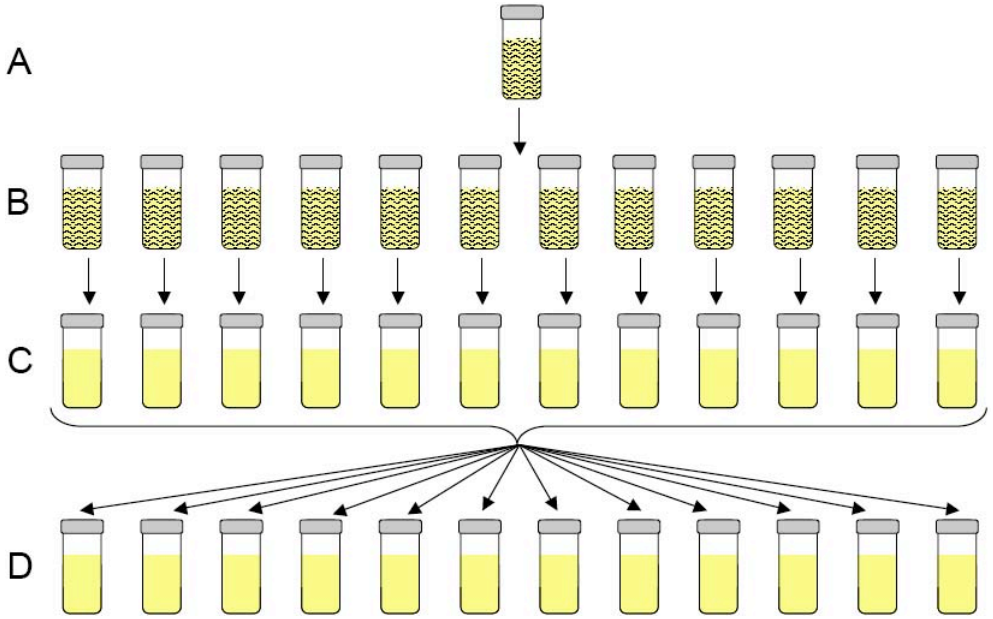

Figure 3-3: The four stages in the MLS 1 experimental treatment.
At step A, SBW25 is pre-cultured from the -80˚C stock. At step B each of the 12 groups of the population are pre-cultured shaking overnight. At step C, each of the 12 groups is inoculated to start the first group-generation, and left to incubate statically for 72 h. After thee days, each group was vortexed thoroughly and equal volumes of each group were mixed together before founding the new groups. Step D is the second group-generation of 72 h. A total of 12 group-generations were completed, totalling 36 days.

The control treatment (Figure 3-4) is different from the experimental treatment, because it does not allow any opportunity for mixing of individuals from different groups. The groups remain permanently isolated from one another for the duration of the experiment. Therefore, there is no opportunity for individuals within groups that reach higher densities to become over-represented within the population in future group-generations.

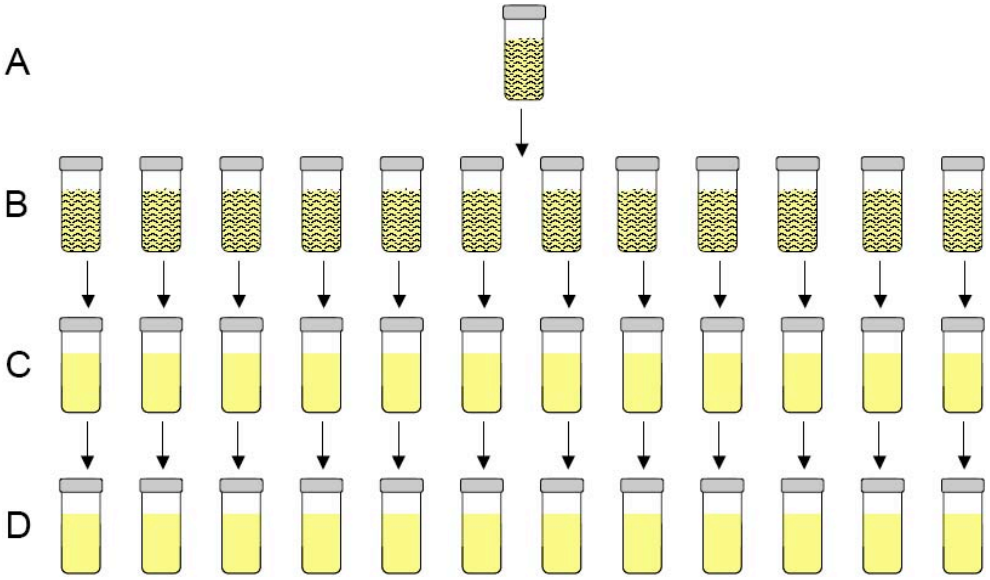

Figure 3-4: The four stages in the control treatment.
At step A, SBW25 is pre-cultured from the -80˚C stock. At step B each of the 12 groups of the population are pre-cultured shaking overnight. At step C, each of the 12 groups is inoculated to start the first group-generation, and left to incubate statically for 72 h. After 72 h, each group was vortexed thoroughly, diluted and directly inoculated into fresh microcosms to found the new groups. Step D is the second group-generation of 72 h. A total of 12 group-generations were completed, totalling 36 days.

For the control treatment, each group was homogenised by vortexing, and the cells were diluted and inoculated directly into fresh microcosms to found the next group-generation. At each transfer, samples were plated to count the proportion of WS types in each group (Figure 3-5).

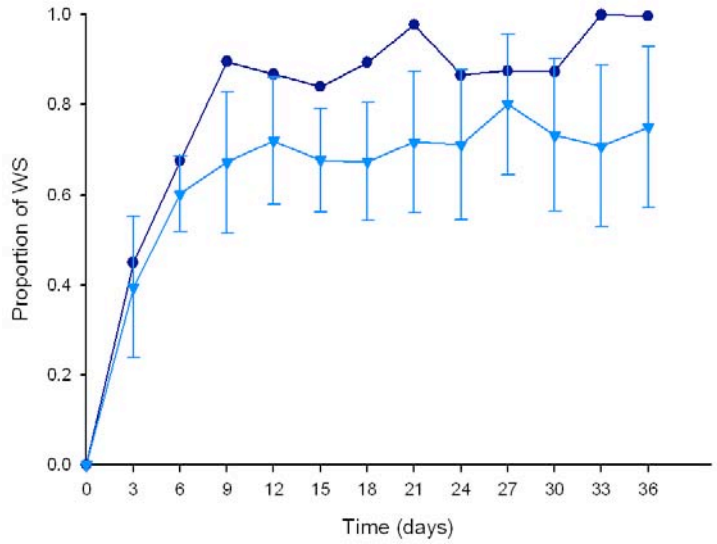

Figure 3-5: MLS 1 experiment with 72-h transfer.
Dark blue circles illustrate the group-structured treatment and light-blue triangles illustrate the control treatment. Error bars represent 95% confidence intervals from 12 independent replicates for the control. The group-structured treatment is an average of 12 measurements, but they are pseudoreplicates in this experimental design; therefore standard error cannot be calculated.

The proportion of WS types in the experimental treatment falls within the range of the control treatment at the first time point (day three). This is a trivial, but important result, because it illustrates that the average of the 12 independent control samples is not significantly different from the mixed sample of the experimental treatment using a one-sample two-sided *T*-test, (T_{11}, $p = 0.666$). To test for a difference between the curves associated with the experimental and the control treatments, a randomisation test identical to that developed in Section 3.2.1 was employed. The data for the experimental and the control treatment were pooled together and a sixth degree polynomial function was fit to the combined data set (Figure 3-6). The proportion of variability explained by the sixth degree polynomial was $r^2 = 0.5953$. The residuals were calculated and an ANOVA of the residuals revealed that the proportion of WS types in the experimental treatment was significantly higher than the control ($F_{1,20} = 93.5929$, $p < 0.0001$). These results demonstrate that cooperative WS types reached a significantly higher proportion, and thus were more successful, in populations in which groups competed with one another to contribute to each new population.

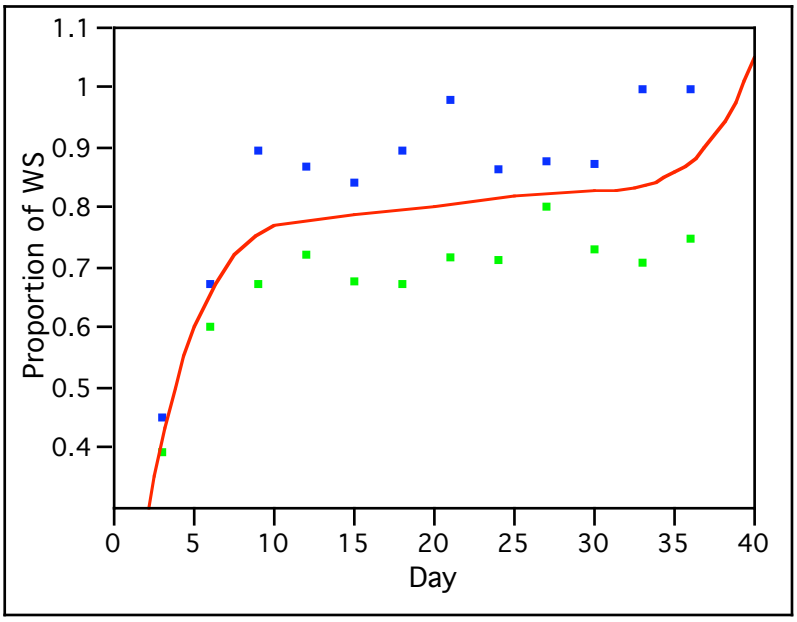

Figure 3-6: Sixth degree polynomial fit to the combined experimental and control data.
The red line is a sixth degree polynomial fit to the data. The blue data points are the experimental treatment and the green data points are the control treatment. An ANOVA of the residuals showed that the difference between the experimental and control is statistically significant.

3.2.2.1 Fitness of Each Derived WS after 36 days

The group-structured treatment was designed so that selection favoured cooperators more strongly than selfish types. However, a further consequence of the group

structure is that strong selection for cooperators also favours selection for 'better' cooperators. Therefore, one expectation from these observations is that the dominant WS type in the experimental treatment will be fitter in a static environment than the other WS types that arise in each of the control lines, because the competition between WS types from different groups allows the fittest WS type to spread between all groups of the population. In addition, given the duration of the experiment, there was significant opportunity for niche-specific adaptation to the air-liquid interface that might elucidate a trade-off among 36-day derived WS that did not exist among the seven-day IWS genotypes from Chapter 2. Therefore, a second expectation is that the dominant WS type from the experimental treatment would have the highest fitness in a static environment, the lowest fitness in a shaking environment and the largest trade-off in the shaking environment.

To assess the fitness of WS types at the termination of the experiment, the dominant WS type from the experimental treatment, 3GWS, and the dominant WS type from each of the control lines, 3WS01 – 3WS12, were selected for fitness assays. For each of these 13 types, the fitness was measured in two environments (three-way static and 24 h shaking) to determine if the fitness of 3GWS was higher than the fitnesses of 3WS01 – 3WS12 in the static environment, lower in the shaking environment, and to investigate the nature of the trade-off between the two environments.

In the three-way static competition each of the 13 WS types was competed against LSWS-*lacZ* with wild-type SBW25 as a broth phase competitor at 28˚C for 72 h. The relative fitness of each of the genotypes compared to LSWS-*lacZ* during the 72 h of competition was calculated as the ratio of the Malthusian parameters for the strain and the competitor (Lenski *et al.*, 1991). The results of a one-way ANOVA (Figure 3-7) showed that there was significant variation in the relative fitness in a static environment among the WS types ($F_{12,91}$ = 9.909, p = < 0.0001) and that the variances can be assumed to be equal (Brown-Forsythe test: $F_{12,91}$ = 1.5221, p = 0.131).

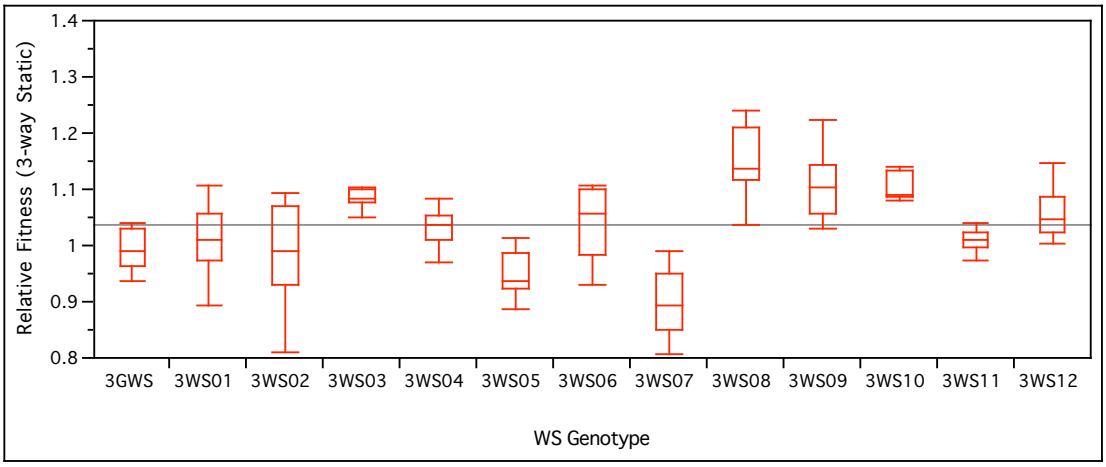

Figure 3-7: Relative fitness of day-36 WS types for a 3-way static competition.
The relative fitness of each of the 13 WS types was measured as a ratio of Malthusian parameters in a three-way competition with LSWS-*lacZ* and SBW25 as a broth phase competitor. 3GWS is the group-selected experimental treatment and the others are the 12 control treatments. A Tukey's pair-wise comparison revealed no unique groupings at the 95% significance level. The grey line indicates the average for all samples.

In the shaken environment each of the 13 WS types was competed against SBW25-*lacZ* at 28˚C for 24 h to measure fitness as the ratio of the Malthusian parameters for the two competitors. The results of a one-way ANOVA (Figure 3-8) showed that there was significant variation in the relative fitness in a shaken environment among the WS types ($F_{12,91}$ = 30.600, p = < 0.0001) and that the variances can be assumed to be equal (Brown-Forsythe test: $F_{12,91}$ = 1.1397, p = 0.3389).

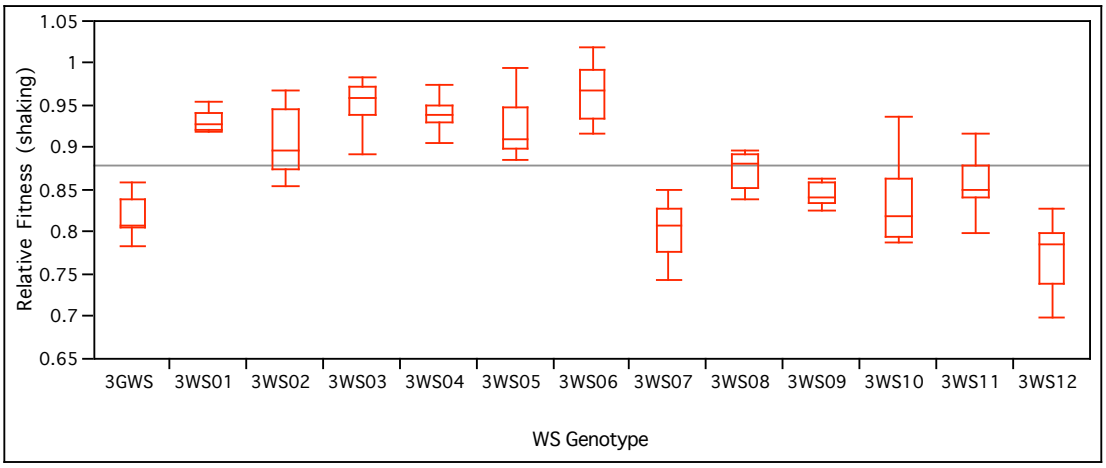

Figure 3-8: Relative fitness of day 36 WS types for a 24 h competition under shaken conditions.
The relative fitness of each of the 13 WS genotypes was measured as a ratio of Malthusian parameters in a shaken environment competed against SBW25-*lacZ*. 3GWS is the group-selected experimental treatment and the others are the 12 control treatments. A Tukey's pair-wise comparison revealed no unique groupings at the 95% significance level. The grey line indicates the average for all samples.

Counter to expectation, the relative fitness of 3GWS was not the highest among the 13 genotypes sampled from the end of the experiment in the three-way static environment, nor was it the least fit in or the 24 h shaking competition. Despite this, these data still provide the opportunity to examine whether or not a trade-off exists among the fitness of the derived WS genotypes in each of the two environments. Therefore, to determine the relationship between the fitnesses of the derived WS types in different environments, the two measurements for fitnesses were correlated with one another (Figure 3-9) as performed previously for the IWS Genotypes (Section 2.2.2.4). A straight line was fit to the data and showed that there was no correlation between the fitness in a shaking and the fitness in a static environment (r^2 = 0.003) and that the fit was not significant ($F_{1,11}$ = 0.032 , p = 0.861). This result agreed with the observation in Section 2.2.2.4 for the correlation between shaking and three-way static fitnesses of the IWS genotypes, and indicated that a lack of a defined relationship between the fitness of WS genotypes in the different environments remained true for long-term experiments.

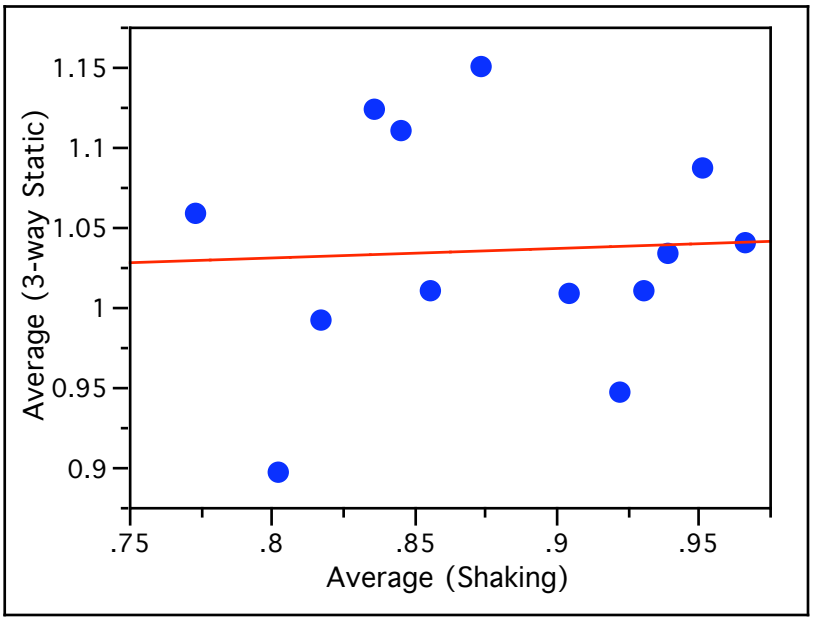

Figure 3-9: Contrast of Shaking Fitness and 3-way Static Fitness for day-36 WS types.
The fitnesses in the shaking and static environments are on different scales because they were measured relative to different genotypes. A straight line was fit to the data with r^2 = 0.003. Genotypes that are further from the line of best fit are relatively fitter in one environment than the other.

3.2.3 MLS 1 WITH A 24-HOUR GROUP-GENERATION TIME

In the 72-h group-generation experiment there was sufficient time for WS types to arise *de novo* and increase in frequency within each group (microcosm), because WS

types are more fit than the ancestral SM type over 72 h. The higher fitness of WS comes from the benefits of being part of the mat, which offsets the cost of cellulose production. However, in traditional MLS 1 models of cooperation, cooperators (WS) have a lower fitness than selfish types (SM) based on their 'within group' fitness. To lower the relative fitness of cooperators within a group, the same experimental design was employed, but with the group-generation time reduced from 72 h to 24 h, because the fitness of LSWS relative to ancestral SBW25 during the first 24 h is known to be 0.8 (Rainey and Rainey, 2003). This effectively resets the growth conditions to that of the first 24 h after every group-generation. Given that the design of the experiment requires that WS types evolve *de novo*, and that WS types are less fit than the ancestral SM types during the first 24 h, one possibility was that the WS types might be selected against so strongly that they would not arise and increase to detectable frequencies within a microcosm. However, in six out of the twelve control lines WS types were detected, and in one of those lines WS types increased to almost 40% (Figure 3-10).

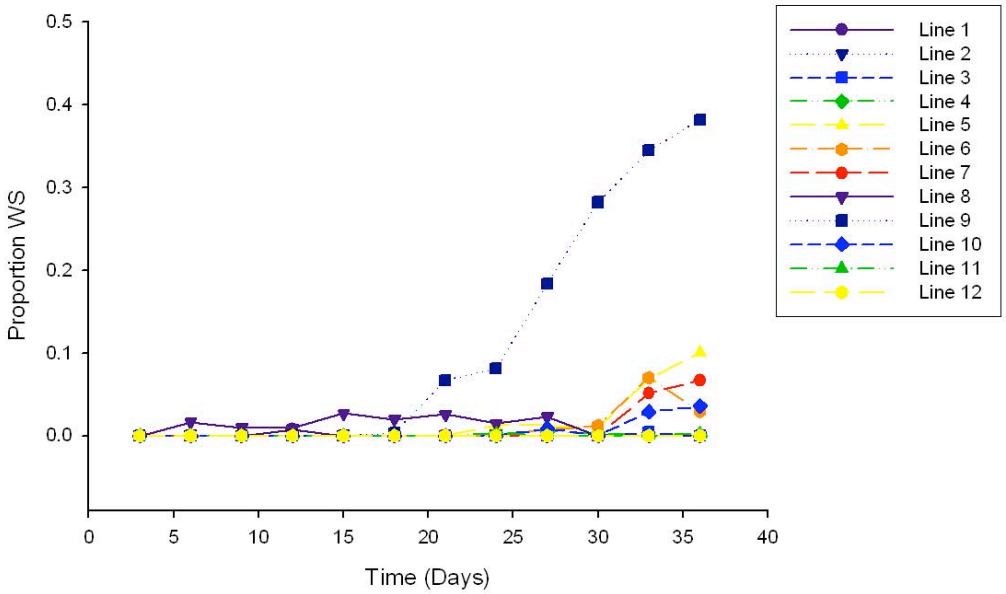

Figure 3-10: Proportion of WS types in each of the 24-h control lines.
The proportion of WS types was measured every 24 h. In line 9 WS types arose and increased to approximately 40%.

The appearance of WS types in the 24-h group-generation control lines indicated that despite a lower relative fitness for the entire 24-h period, some WS types evolved during the course of the experiment (notable in line 9 at day 18) that are more fit than

SM types. Therefore, given that such WS types could arise within a microcosm, it was expected that in the experimental treatment, such WS types would be capable of spreading and increasing their representation across the population when afforded the opportunity of competition among groups. However, counter to this prediction, no WS types were ever detected in the experimental treatment. Figure 3-11 shows a plot of these data for the experimental treatment (groups are mixed, dark blue circles) and for the average of the control treatment (no mixing, light blue triangles).

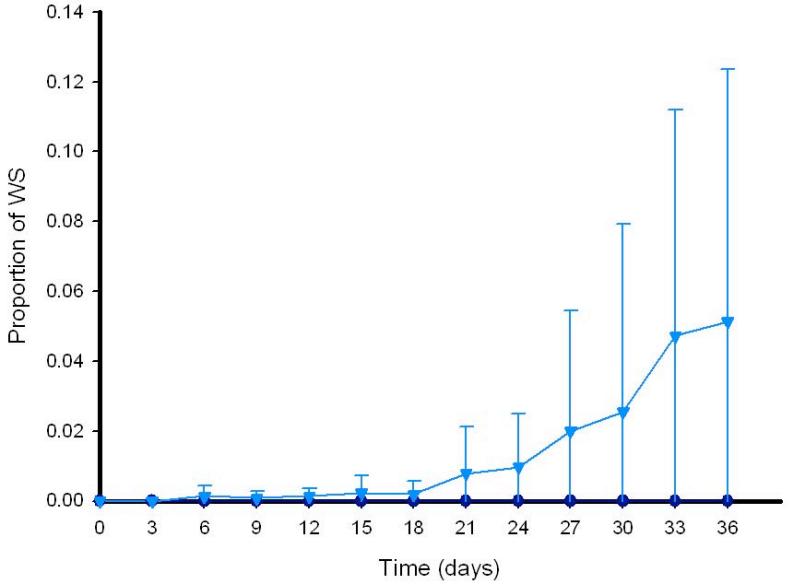

Figure 3-11: MLS 1 Experiment with 24-h transfer.
Dark blue circles illustrate the group-structured treatment and light-blue triangles illustrate the control treatment. Error bars represent 95% confidence intervals from 12 independent replicates for the control. The group-structured treatment is an average of 12 measurements, but they are pseudoreplicates in this experimental design; therefore standard error cannot be calculated.

One possible explanation for never observing WS types in the group-structured treatment is the evolution of a 'super smooth' (SSM) type, an SM type with an unusually high fitness such that WS types in the SSM genetic background can not increase in frequency to detectable levels, because they are of lower fitness than the SSM type. Two complementary approaches were used to test this hypothesis. Firstly, the dominant SM types from days 18, 27 and 36 (designated SSM18, SSM27 and SSM36) were grown statically for five days to determine if WS types could arise when provided with a longer incubation period (Figure 3-12).

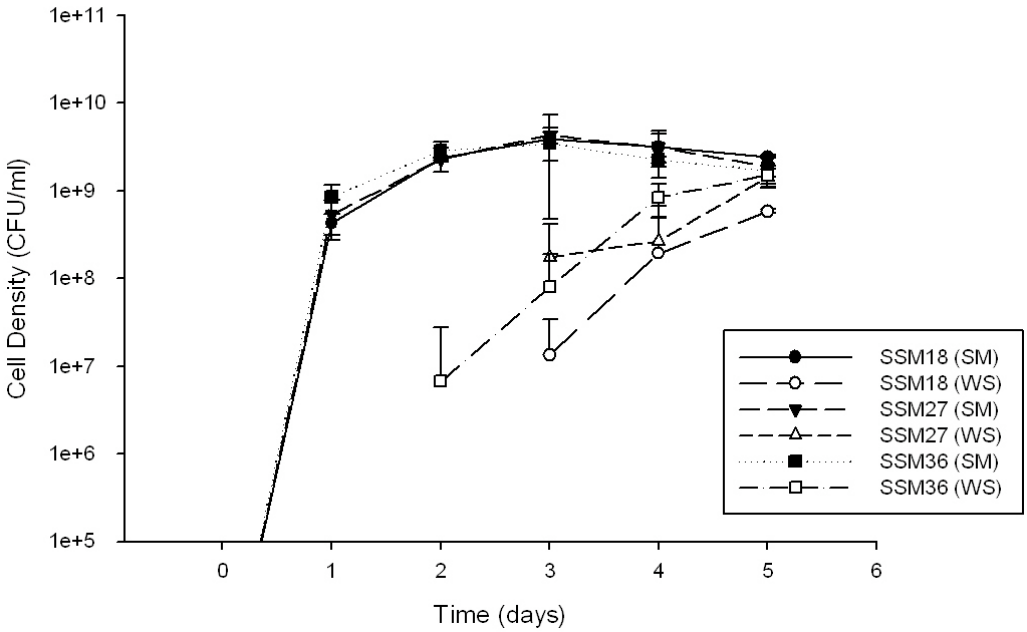

Figure 3-12: Evolution of WS types from SSM genotypes during five days of static incubation. WS types were capable of evolution from each of the selected SSM genotypes. The filled in shapes are the cell densities for the SM morph and the unfilled shapes are for the derived WS types. Shapes are paired for a given ancestral genotype *i.e.* SSM18 (SM) and SSM18 (WS) are both circles.

In all cases WS types arose and increased to detectable frequency within three days, indicating that WS types can arise in the genetic backgrounds designated SSM18, SSM27 and SSM36. However, the WS types derived from each of the SSM strains were observed at lower proportions when compared to the derived types of the adaptive radiation of wild-type SBW25. Multiple Tukey's pair-wise comparisons revealed that the differences in the proportion of WS types among each the SSM genotypes and ancestral SBW25 were statistically significant for days three, four and five (Table 3-1) suggesting possible constraints on the evolvability of the SSM strains.

Table 3-1: Tukey's Test for differences among proportions of WS types for during SSM evolution.

	Tukey's Test grouping at $\alpha = 0.05/4 = 0.0125$			
Day	**SBW25**	**SSM18**	**SSM27**	**SSM36**
2	A	A	A	A
3	A	B	B	B
4	A	C	C	B
5	A	C	B	B

SBW25 was always defined as group A and more distant letters reflect greater difference from SBW25 in terms of the analysis.

In a second approach to testing the hypothesis that an SSM had evolved, SSM18, SSM27 and SSM36 were competed against SBW25-*lacZ* in a static environment for 24 h (Figure 3-13). An ANOVA revealed that the average relative fitness for each of the SSM genotypes was not significantly different from the fitness of the ancestor ($F_{3,36}$ = 1.3219, p = < 0.2823) and that the variances could be assumed to be equal (Brown-Forsythe test: $F_{3,36}$ = 0.9147, p = 0.4435).

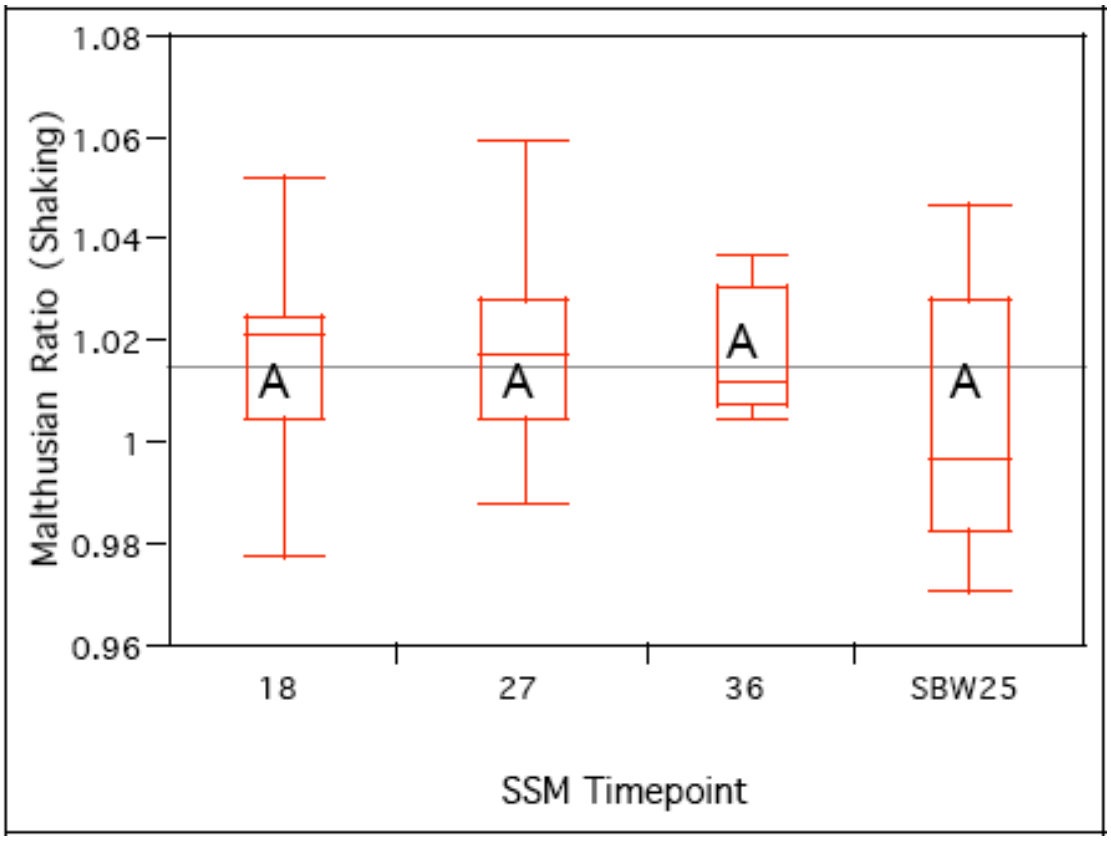

Figure 3-13: SSM competitions in a static environment against the ancestor.
The dominant SM types from three time points (18, 27 and 36 days) were competed against SBW25 to determine their relative fitness. A control of SBW25 against SBW25-*lacZ* was used to generate the null distribution of fitnesses for the ancestor. A Tukey's pair-wise comparison shows no significantly different groups indicated by identical letters ($\alpha < 0.05$).

3.3 DISCUSSION

3.3.1 ADAPTIVE RADIATION IS CONSERVED IN THE MARKED STRAIN SBW25-*LACZ*

The principal feature of the biology of SBW25 is the well-characterised adaptive radiation in a static microcosm (Rainey and Travisano, 1998; Rainey and Rainey, 2003). The engineered strain SBW25-*lacZ* was known to have no measurable difference in fitness compared to SBW25 on multiple substrates and *in planta* (Zhang and Rainey, 2007), and further experiments showed that there was no measurable change in fitness for longer-term experiments (Appendix 9.1.1). This feature was exploited to determine if the characteristic biology of the system was conserved as well. The adaptive radiation data of the major morphotypes were compared to one another, and statistical tests confirmed that the proportions of SM, WS and FS types from the marked and unmarked strains followed the same trajectories. The congruency between the marked and unmarked strains shows considerable opportunity for SBW25-*lacZ* as a marked variant of SBW25 in evolutionary experiments. For example, Fukami *et al.* (2007) developed an experimental design to investigate how migration history affects diversification and extinction using the *P. fluorescens* system. The experiment used derived niche-specialists of both strains as migrants and as founder populations to alter the 'boom-and-bust' dynamics of the adaptive radiation. Importantly, the conclusions depend on the marked and unmarked strain sharing a consistent biology of adaptive radiation. These results confirm the suitability of the defective prophage locus as an ideal site for the insertion of other coloured or resistance markers, *e.g.* gentamicin as used later in this thesis (Chapter 4). In addition, this evidence suggests that such strains might also be conducive to expansion of their use from batch cultures to continuous-culture experiments in chemostats in which fitness differences between types are magnified if present.

3.3.2 72-H GROUP-GENERATION SUPPORTS THE HAYSTACK MODEL

MLS 1 theory predicts that competition between groups, afforded by the group structure of a population will benefit those individuals that are members of more productive groups. The results for the 72-h group-generation experimental treatment

agreed with the prediction that competition between groups increases the benefit to individuals within more successful groups. The proportion of WS cooperators was consistently higher in the experimental treatment than in the control treatment. Similar results have been observed by other authors studying siderophore secretion as a cooperative trait in populations of *P. aeruginosa* (Griffin *et al.*, 2004; Ross-Gillespie *et al.*, 2007). Ross-Gillespie and co-workers showed empirically how negative frequency dependence among cooperators and cheaters could explain the maintenance of cooperation only in situations of 'intermediate structuring'. In completely unstructured (homogenised) or fully structured (groups are never mixed) environments frequency dependence is never observed, so no relative benefits of cooperation can be selected (Ross-Gillespie *et al.*, 2007). The other study (Griffin *et al.*, 2004), investigated how cooperation depends on population structure, relatedness and the interaction between population structure and relatedness (Queller, 1994; Frank, 1998; West *et al.*, 2002). Although couched in a different vernacular, their experimental design describes a Haystack model within an MLS 1 framework, and their results show how cooperation depends on population structure with and without the interaction with relatedness. However, one point of difference between the WS system of *P. fluorescens* and traditional Haystack models of group structure is that cooperative genotypes are usually less fit within the group. This is not true for the WS system when the group-generation time is 72 h, because 72 h is sufficient time for the costs of cooperation to be offset by the benefits of occupying the niche at the air-liquid interface (Rainey and Rainey, 2003). Importantly, this distinction is irrelevant with respect to the hypothesis that competition between groups afforded by group structure will benefit those individuals from more successful groups, because group membership still affects the fitness of individuals within the group. Therefore, these results confirm that group structure increases the benefit of cooperation for WS types with a 72-h group-generation.

One expectation from these results is that the most productive WS type that spreads through the population in the experimental treatment, 3GWS, should be fitter than the WS types in the control lines, and this is not observed. One possible explanation is the destabilisation of synergistic coevolutionary interactions among WS types by a high migration rate in the experimental treatment. The destabilisation of these interactions is a direct consequence of the experimental regime, due to the

homogenisation of the genotypes among the 12 groups at the end of each group-generation. From a coevolutionary view point, at the time of mixing, 92.7% (11/12) of the population is replaced through migration, and this occurs repeatedly every 72 h destabilising existing combinations. This contrasts strongly with the control treatment in which coevolution among derived genotypes is continually maintained for the duration of the experiment.

In a study on T4 bacteriophage, migration pattern alone has been shown empirically to favour 'restraint' (cooperation) leading to a resolution of the 'tragedy of the commons' (Kerr *et al.*, 2006). Under conditions of restricted migration, the 'prudent' cooperative strategy was able to avert the tragedy of the commons; however, when migration was unrestricted, the 'rapacious' (selfish) phage displaced the 'prudent' cooperative phage resulting in lower phage production overall. Furthermore, in *P. fluorescens*, coevolution among divergent types has been shown to promote cooperation in biofilms, and that productivity is likely to be higher in groups with more diverged genotypes resulting in increased complementary resource use (Brockhurst *et al.*, 2006). Cooperation has also been shown to be favoured by smaller bottleneck size at the point of transfer among coevolved genotypes for the range 10^3 – 10^8 cells/ml (Brockhurst, 2007). The bottleneck size employed in the experiments in this chapter was ~10^4 cells/ml, which suggests that selection might favour cooperation among the coevolved genotypes in the control lines. Finally, coevolution among WS genotypes in a group-selected experimental regime has shown the evolution of strong synergies (McDonald, unpublished results). Groups of WS cells were selected on mat strength, and in one experimental line, the mat always consisted of two WS genotypes. When grown statically in monoculture, the mat strength of each of the two WS genotypes was significantly lower than when the two were grown together, exciting theories of potential division of labour among different WS types (McDonald, unpublished results). This preliminary evidence supports the idea that coevolutionary dynamics can potentially lead to the ascent of adaptive peaks through synergistic interactions among divergent genotypes that are not possible when migration is high. The complexity of coevolutionary interactions illustrates that confounding factors may have contributed to the unexpected result that 3GWS does not have the highest fitness among the WS types sampled at the termination of the experiment.

3.3.3 24-H GROUP-GENERATION AND THE SUPER-SMOOTH HYPOTHESIS

The length of the group-generation, the period of time between successive formations of groups, is theorised to influence MLS 1 models of group-selection in two ways (Frank, 1998). Firstly, increasing the group-generation time increases the advantage to groups with more cooperators (favours cooperation), but, secondly, this advantage is traded-off against increased competition within the group (favours selfishness). The nature of this trade-off determines whether cooperation or selfishness is favoured empirically by the shortening of the group-generation time from 72 h to 24 h. To determine this empirically, the group-generations from the 24 h and 72 h group-generation experiments must be 'synchronised' so that the intrinsic difference in turnover rate is exactly compensated for (Okasha, 2006a). Fortunately, this is simple for this experimental design, because three 24-h group-generations are exactly equal to a one 72-h group-generation[1]. A comparison of the data from the 72-h group-generation experiment to the 24-h group-generation experiment showed that after three days (one 72-h group-generation or three 24-h group-generations) 40% of cells were WS types in the 72-h group-generation, while there were no WS types observed for the 24-h group-generation experiment. This indicates that the trade-off between group-advantage of cooperation and individual within-group selfishness is dominated by the group advantage of cooperation in the WS system.

Although SM types were favoured by the shortened 24-h group-generation time, WS types did arise and increase to detectable frequencies in 6/12 lines of the control treatment, indicating that some WS types can be favoured by selection within a microcosm. In each case, the WS types arise much later than the usual two to three days for uninterrupted static incubation, suggesting either a strong bottlenecking effect of dilution to ~10^4 cells or that multiple mutations during a longer period of evolution are required before WS types of high enough fitness are discovered by

[1] This assumes a constant growth rate between treatments, which is unlikely to be true given that after 24 h of growth cells are in static microcosms are usually in stationary phase. However, this effect is likely to be dwarfed by the differences in selective pressures between the two treatments.

natural selection. However, given that such WS types could arise within a microcosm in some of the control lines of the 24-h group-generation, it was surprising that WS types were never observed in the experimental treatment. One possible explanation comes from re-examining the properties of the model. Although the Haystack model and other models that fit into the paradigm of MLS 1 theory were introduced to explain the evolution and maintenance of cooperative traits, the models actually do not require types to be cooperative or selfish. Instead, the models only require that group membership differentially affects the fitness of the individuals within that group. Therefore, if group membership has an effect, the only guarantee from the experimental design is that the types from the most successful groups should be over-represented, regardless of whether or not the group success is due to cooperation.

One type that may have arisen within successful groups without cooperation is a 'super smooth', SSM. An SSM type would be any SM genotype that were to have unusually high fitness, such that WS types in the SSM genetic background could not increase in frequency to detectable levels, because they were of lower fitness. No evidence for the existence of an SSM was found in competitive fitness assays, because each of the potential SSM genotypes had no measurable increase in fitness relative to the ancestor. However, when left to evolve in a static environment, WS types derived from each of the SSM strains were observed at lower proportions when compared to the derived types of the adaptive radiation of wild-type SBW25.

One possible explanation for this observation is that the derived SSM genotypes have a reduced evolvability, an observation consistent with a previous study in *P. fluorescens* (Buckling *et al.*, 2003). In this study it was demonstrated that adaptation by SM to the static environment limited the ability of SM populations to diversify when the fitness landscape was 'rugged'. A fitness landscape is a conceptual tool created by plotting fitness against all possible genotypes within the population (Wright, 1932), and is described as rugged when there are multiple local peaks and valleys. The fitness landscape for SBW25 in a static microcosm is considered rugged, because there is niche-specialisation to multiple niches of the structured environment among the various genotypes (Rainey and Travisano, 1998; Buckling and Rainey, 2002; Rainey and Rainey, 2003). In addition, the ancestor is expected to be in a fitness valley allowing for the ascent of multiple local peaks, because the laboratory

microcosm is a novel environment (Rainey and Travisano, 1998; Rozen and Lenski, 2000). Such a reduction in evolvability as a consequence of adaptation, may account for the lower proportion of WS types observed in the SSM lines, and the unexpected observation of no WS types in the group-structured treatment.

3.3.4 The Limitations of MLS 1 with Respect to the Emergence of Higher Levels

MLS 1 is a framework that analyses the fitness of *individuals* when the fitness of that individual is affected by group membership (Damuth and Heisler, 1988; Okasha, 2006a). It emphasizes how the total success of individuals can be partitioned into selection within and among groups (Hamilton, 1975; Wilson, 1980; Frank, 1998). Crucially, although MLS 1 makes essential appeal to group structure, groups are part of the environment, and not fitness-baring entities, though it is tempting to describe MLS 1 phenomena using MLS 2 language by thinking of groups as fitness-baring entities. As Okasha (2006a p. 91) notes,

> *For in MLS 1, collective fitness is a logical construct of particle fitness. So any causal factors that affect the fitness of a collective can* only *do so by affecting the fitnesses of the particles it contains; there is no other way to affect a collective's fitness.*

Therefore, MLS 1 permits no inference about the fitness of different types of groups in any capacity exceeding its definition of 'average individual fitness', because selection within and among groups contributes to changes in fitness at the level of the individual only.

Groups as evolutionary parameters are dealt with exclusively in an MLS 2 framework, which requires the additional assumption that groups are units of selection. Units of selection, or 'Darwinian individuals' must have the properties of variation, reproduction and heredity (Lewontin, 1970). Groups in an MLS 1 framework possess variation, but the variation is not inherited at the level of the group, because the reproducing unit is the individual. Although groups in an MLS 1 framework are occasionally broken down and reformed, any analogy with reproduction would be unfounded, because the hereditary particle is the individual (Okasha, 2006a). By contrast, groups in MLS 2 are units of selection and by definition possess variation, reproduction and heredity. MLS 2 suffices for the study

of situations in which groups have already evolved to be autonomous units of selection, however, it cannot elucidate anything about the *de novo* emergence of the group-level unit. To assume that MLS 2 can explain how groups evolved is to commit a logical fallacy that conflates product with process, because to explain the emergence of groups from individuals requires the *a priori* assumption that groups are already units of selection. Therefore, the resolution of this fallacy necessitates the conclusion that groups must be reproduced through individuals, and implicates a role for a primordial life cycle or development-like process in the evolution of multicellularity. Further theory and novel experimental work associated with hypothesis are continued in Chapter 5.

4 THE ECOLOGY OF WS COOPERATION: AN EMERGENT GROUP-LEVEL PHENOTYPE

The theory of emergence points to a gradual, inexorable evolution of the cosmos, from atoms to galaxies to planets to life. Each emergent step arises from the interactions of numerous agents and yields an outcome much greater than the sum of its parts. Each emergent step increases the degree of order and complexity, and each step follows logically, sequentially from its predecessor.

A comprehensive theory of emergence will foster applications to myriad problems in everyday technology: long-range weather prediction, computer network design, traffic control, the stabilization of ecosystems, the control of epidemics, perhaps even the prevention of war. Armed with such a law, we will acquire a deeper understanding of any system of many interacting agents — indeed, even of the origin of life itself.

-- Hazen, R. M. 2005

4.1 INTRODUCTION

The product of any ecological interaction must theoretically be described as either synergistic or antagonistic, while non-interactions are orthogonal (Thompson, 1999). For example, when an actor, X, interacts with a recipient, Y, the fitness of X as a consequence of the interaction, ω'_X and the fitness of Y as a consequence of the interaction, ω'_Y are dependent on a baseline fitness ω_X (or ω_Y) and the product of the interaction (Eq 4-1 and Eq 4-2 respectively).

$$\omega'_X = \omega_X + \omega_{X*Y} \qquad \text{Eq 4-1}$$

$$\omega'_Y = \omega_Y + \omega_{Y*X} \qquad \text{Eq 4-2}$$

Although, each interaction term must by definition always exist, either or both may in some circumstances be neutral with respect to fitness (*e.g.* neutralism, commensalism and amensalism), and will not necessarily be the same size or sign. Among all types of interactions, synergistic interactions produce emergent properties – effects that are 'not otherwise attainable' (Corning, 2002), because the effects are quantitatively (emergent fitness) or qualitatively (emergent phenotype/character) 'greater than a sum of the parts' (Gould, 2002). Emergent properties depend on the interaction of the individuals to produce the synergism; therefore, ecological theory predicts that increasing the number of acting individuals should increase the synergistic effect.

Cooperative interactions among microbes are a subset of synergistic interactions that typically fit the public goods model, because all actors are recipients, but some recipients may be defectors (non-actors). Therefore, if the cooperative interaction has been selected for its benefit to others, a fitness increase for all individuals in the population depending positively on the number of actors is predicted (*i.e.* positive density dependence). This density-dependent increase in fitness on the number of actors is an emergent property of the cooperative interaction, because the product of the interaction is qualitatively or quantitatively 'greater than a sum of the parts'. However, the precise relationship, *e.g.* exponential or stepwise, between the density of actors and the fitness of the recipients remains to be determined experimentally,

and is expected to dovetail with the distinction of emergent fitness and emergent character.

Three types of cooperation have been identified for communities of microorganisms (Velicer, 2003), though one of these will be disputed. The first two are called *simple density-dependent cooperation*[1] and *group-limited cooperation*. Simple density-dependent cooperation is described as "positively synergistic, in that individuals obtain greater benefit at higher densities than at low densities" (Velicer, 2003 p.332). Group-limited cooperation is more complex, because a sufficiently large quorum of cells is required before a synergistic benefit is observed. The third type, called *minimal cooperation* (Velicer, 2003), is described as "restraint from unfairly sequestering a publicly available benefit when the magnitude or mere existence of that benefit is not density dependent" (Velicer, 2003 p.332). However, minimal cooperation does not result in an emergent property of any sort, because there is no interaction among individuals (*i.e.* the action is orthogonal with respect to other individuals), which violates the first criterion outlined in Section 1.2.2. Moreover, from an evolutionary standpoint, minimal cooperation cannot be favoured (or disfavoured) by selection for its effect on others, because it is inadvertent or non-adaptive. The inclusion of minimal cooperation as a type of cooperation is in conflict with the definition of cooperation advocated by West *et al.* (2007b) and as used in this thesis (Section 1.2.2), because cooperation is defined as any action *selected to enhance* the fitness of others at a relative cost to the acting individual. Therefore, minimal cooperation does not meet the requirements to be considered a type of cooperation, because there is no feedback from the recipient to the actor, thus selection cannot act on the benefit to the recipient by definition. Despite this, the expectations of minimal cooperation as described by Velicer (2003) will be

[1] Simple density-dependent cooperation was originally termed 'density-dependent cooperation' by Velicer (2003); however, this is misleading, because both are types of density-dependent cooperation, but are mechanistically different.

considered for comparison as a null model for the absence of cooperation in microbial systems.

Simple density-dependent cooperation, group-limited cooperation and minimal cooperation each have different predictions for the way that cooperator fitness is affected by initial cooperator density (Figure 4-1). Minimal cooperation is "equally beneficial in isolation or in pure groups of [producers][1]" (Velicer, 2003 p.332). Therefore, the expectation for minimal cooperation is that relative fitness will not increase with initial cooperator density, *i.e.* there is no emergent fitness. Simple density-dependent cooperation occurs if "a trait that is beneficial in isolation provides an even greater benefit to its bearers in a high density group of cooperators than at low density" (Velicer, 2003 p.332). Therefore, relative fitness should increase as some function of initial cell density, *i.e.* there is an emergent fitness. Group-limited cooperation requires a quorum to be sufficiently large before a fitness benefit is obtained and "such a trait is disadvantageous if expressed at low density" (Velicer, 2003 p.332). Therefore, the relative fitness of cooperators will be below one at low densities and above one at high densities, and the quorum required for the benefit at high densities is expected to be identifiable as an emergent phenotype.

[1] Producers are called cooperators in the original of this quote, but given the refutation of minimal cooperation as a type of cooperation, producers is a preferable term when referring to the actors in a scenario described by minimal cooperation.

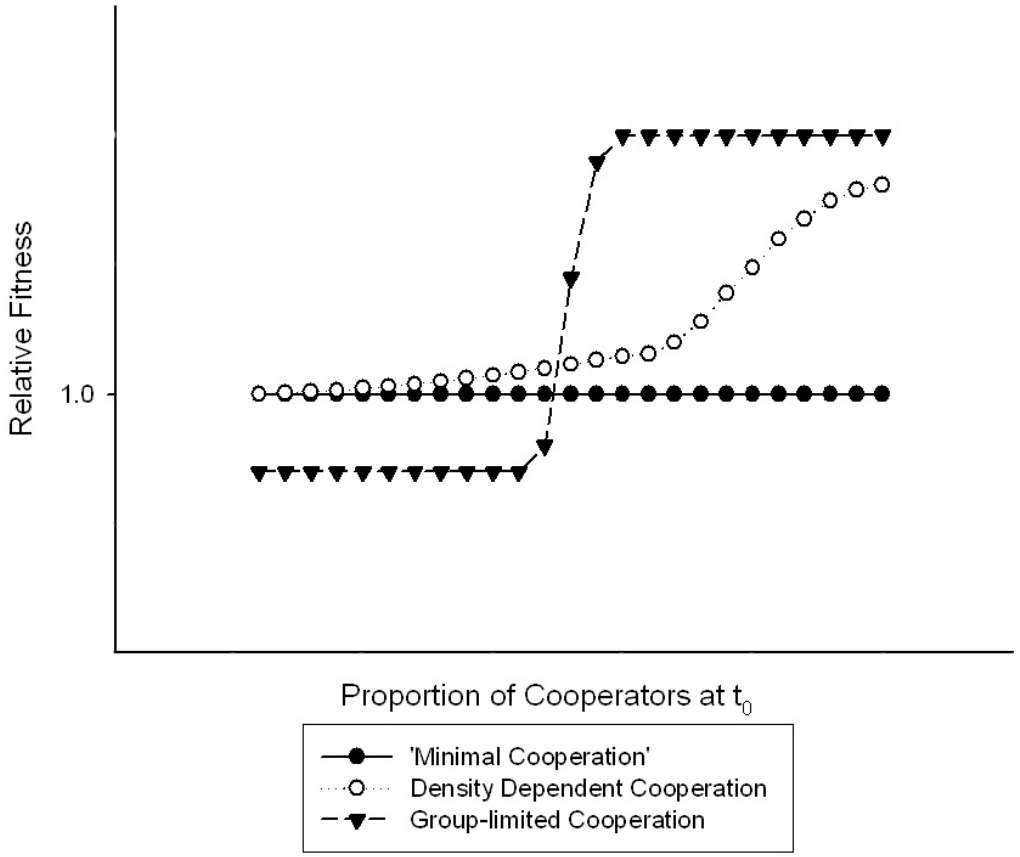

Figure 4-1: Expectations of relative fitness as a function of initial cooperator density for different types of cooperation.
The black circles show the expectation for non-adaptive or minimal cooperation. The open circles show the expectation for simple density-dependent cooperation (the shape of the function may vary). The black triangles show the expectation for the group-limited cooperation.

The *a priori* expectation for WS cooperation was that it is an example of group-limited cooperation, because of the unambiguous emergent group-level phenotype – the mat at the air-liquid interface. However, as will be shown in this chapter, although the mat is certainly an emergent group-level phenotype, WS cooperation does not meet the density-dependent expectations of group-limited cooperation outlined by Velicer (2003), but rather the emergent phenotype appears to be the product of cells growing together over time in a development-like process.

In this chapter, the emergent group-level phenotype in the WS system was examined to determine the ecological nature of the WS mat at the air-liquid interface. Two pre-experiments were required before commencing with the main experiment. First, the marked strains to be used for the main experiment were tested to verify that the

inclusion of the marker had no measurable affect on fitness (Section 4.2.1). Second, to investigate the fitness benefit attributable to the mat at the air-liquid interface, the broth phase of the microcosm was excluded as a potential niche in which cells could grow. This was achieved by saturating the broth phase of microcosms with an SM strain that could not evolve the WS phenotype, and verifying two predictions. First, a different SM type with identical fitness added to the broth-saturated microcosm should not be able to increase in frequency, and second, a WS type added to the broth-saturated microcosm should be able to increase in frequency (Section 4.2.2). Finally, having verified that the broth phase was unavailable for growth, the broth-saturated microcosm was used as an investigative tool in the main experiment to determine the relationship between WS fitness with density and time, and compared to the theoretical expectations outlined above (Section 4.2.3). Contrary to prediction, no density dependence was observed, but rather it was determined that the emergent fitness was dependent on time, implicating a role for a development-like process.

4.2 RESULTS

4.2.1 NEUTRAL MARKING OF STRAINS

It has been previously reported that the *lacZ* marker used to create the SBW25-*lacZ* strain was neutral in short-term fitness assays on multiple substrates and *in planta* (Zhang and Rainey, 2007). In addition, work in this thesis has shown the marker has neither a measurable long-term effect on fitness (Appendix 9.1.1) nor a difference in the characteristic adaptive radiation (Section 3.2.1). Two of the strains used for experiments in this section carried markers integrated at the same locus, SBW25-Gm and LSWS-Gm. The third strain, SBW25-Δ*wss* had a kanamycin resistance cassette inserted into the *wss* operon that abolishes cellulose production (Gehrig, 2005). Each of these strains was competed in ten-fold replicated fitness assays against SBW25-*lacZ* (Methods 7.2.15.1) and relative fitness was calculated as a ratio of the Malthusian parameters to test for fitness differences among genotypes (Figure 4-2).

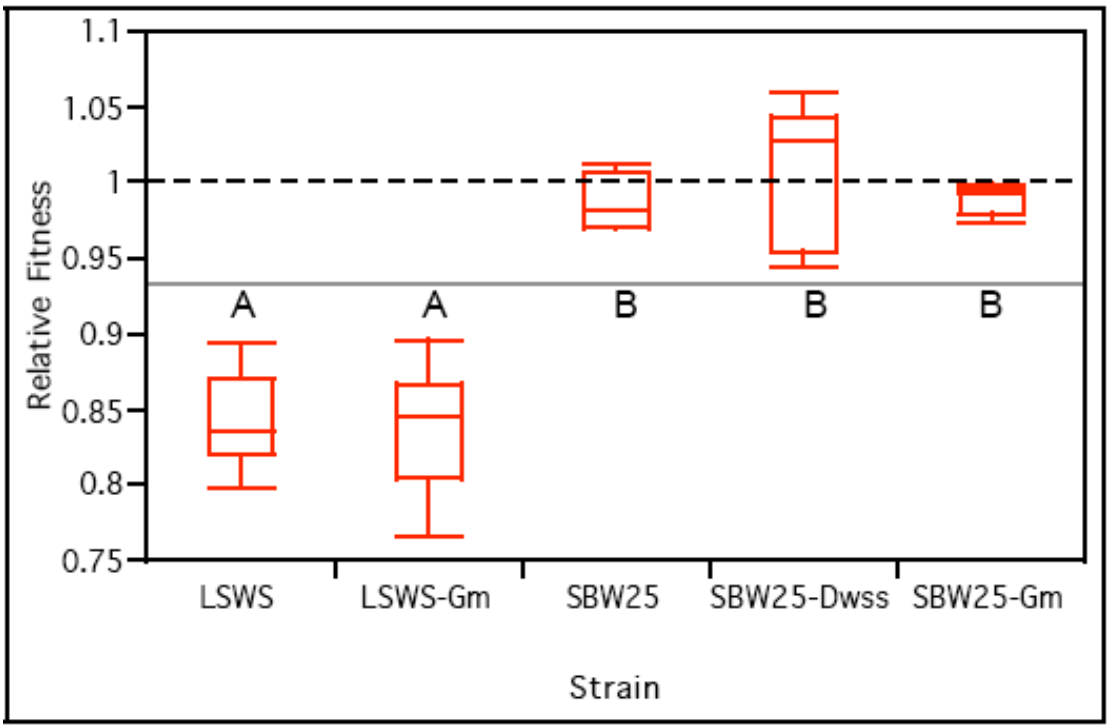

Figure 4-2: Results of the fitness assays among the genotypes used in Chapter 4 relative to SBW25-*lacZ*.
The relative fitness of each genotype was measured against SBW25-*lacZ*. The dashed line indicates a relative fitness of 1. The grey line indicates the average for all samples. Genotypes not connect by the same letter are significantly different (Tukey's HSD $\alpha < 0.05$).

An ANOVA of the relative fitnesses revealed a significant effect of strain ($F_{4,45}$ =

49.075, $p < 0.0001$) and that the variances can be assumed to be equal (Brown-Forsythe test: $F_{4,45} = 0.9582$, $p = 0.4397$). A Tukey's pair-wise comparison showed that the fitnesses of both WS strains were not significantly different from one another and that the fitness of both SM strains were not significantly different from one another at the 95% significance level. Additionally, both WS strains were significantly different from both of the SM strains at the 95% significance level. Therefore, each of the markers is neutral with respect to fitness and the observed differences in fitness were solely attributable to the LSWS genotype.

4.2.2 Restricting Growth to the Niche at the Air-Liquid Interface

The multiple ecological niches of the standard microcosm can be generalised to two primary niches, the broth phase in which SM types are favoured, and the air-liquid interface at which WS types are favoured (Section 1.4). To investigate how WS types gain their advantage at the air-liquid interface, the broth phase of the microcosm had to be excluded as a potential niche in which cells could grow. To achieve this, it was suggested that the broth phase could be removed as an opportunity for growth by saturation with SBW25-Δ*wss*, a strain incapable of producing cellulose and the WS phenotype due to the deletion of the entire *wss* operon[1]. Broth-saturated microcosms (Methods 7.2.15.4) could be considered to have successfully excluded the broth phase as a niche for growth if the following two predictions were met. First, a different SM type with identical fitness added to the broth-saturated microcosm should not be able to increase in frequency, and second, a WS type added to the broth-saturated microcosm should be able to increase in frequency.

To verify the prediction that SM types could not increase in frequency in the broth-saturated microcosms, a gentamicin-resistant SM strain (SBW25-Gm) was added to microcosms saturated with SBW25-Δ*wss* at multiple starting densities. The use of a gentamicin resistance marker was necessary, because the density of the added strain,

[1] Three-way static competitions as in Bantinaki *et al.* (2007) and as used elsewhere in this thesis have a broth phase competitor, but the broth phase is not completely excluded as an opportunity for growth.

SBW25-Gm, was orders of magnitude lower than SBW25-Δ*wss* and was only detectable on agar plates when growth of SBW25-Δ*wss* was inhibited by the presence of the antibiotic. The starting proportion of SBW25-Gm relative to SBW25-Δ*wss* was varied from 10^{-3} to 10^{-8} and a negative control in which no cells were added was also included[1]. Microcosms were incubated for either 24 h or 48 h and sampled destructively by vortexing, and the proportions of SBW25-Gm relative to SBW25-Δ*wss* were counted (Figure 4-3). The results of a Tukey's pair-wise comparison showed that, for all starting densities, there was no increase in the proportion of SBW25-Gm, because the proportion of SBW25-Gm at 24 h and 48 h did not differ statistically from the proportion at t_0 ($\alpha = 0.05$, $df = 18$).

[1] No colonies were ever observed after incubation for 48 h in the control treatment, indicating the efficacy of gentamicin as an antibiotic against sensitive strains of *Pseudomonas*.

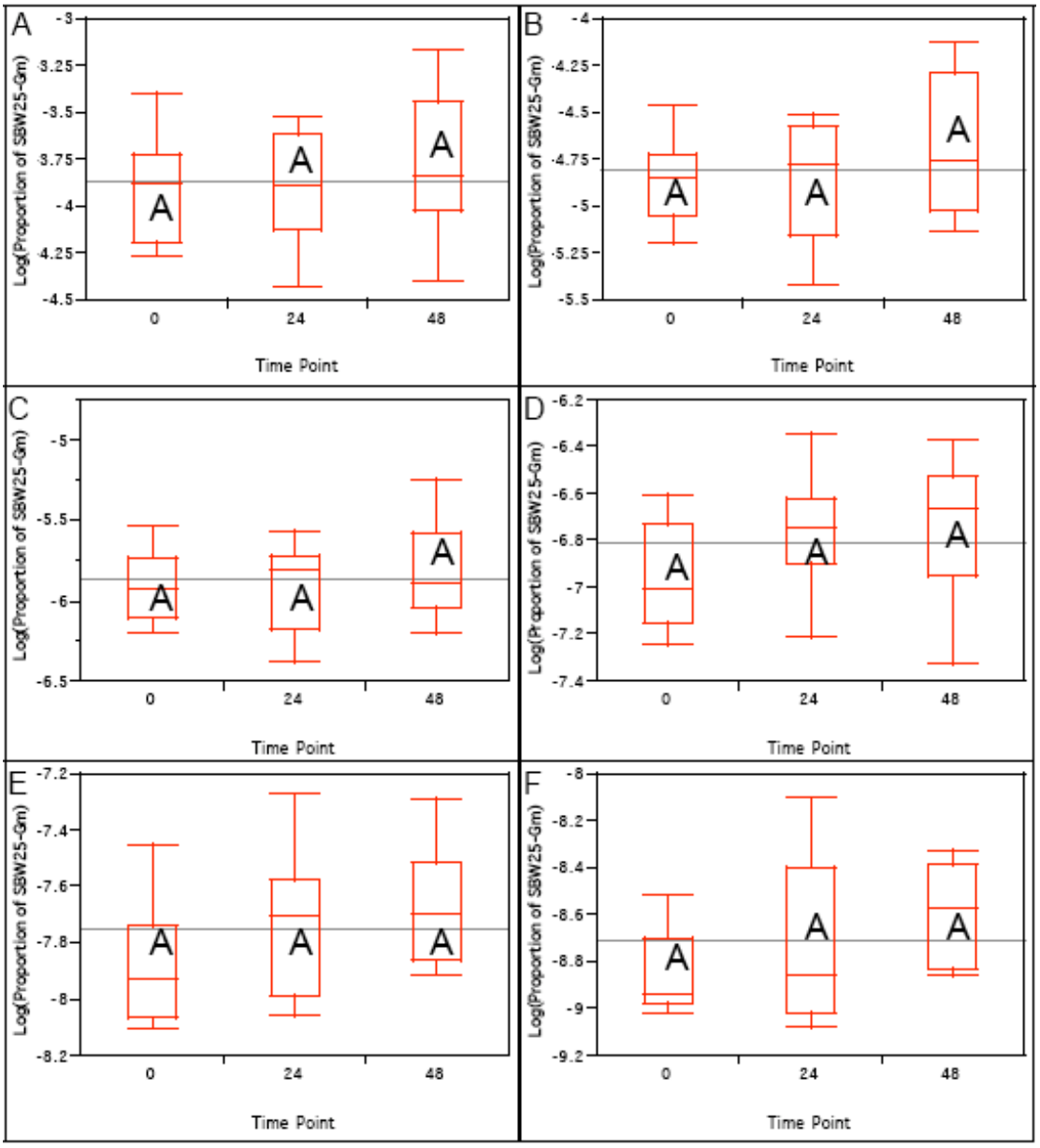

Figure 4-3: Change in the proportion of SBW25-Gm with respect to time at multiple densities. The box plots for log(A) – log(F) are for densities ranging from 10^{-3} to 10^{-8}. The results of a Tukey's pair-wise comparison are indicated by the group letter ($\alpha = 0.05$, $df = 18$).

In contrast to SBW25-Gm, it was expected that LSWS-Gm would be able to increase in frequency in a broth-saturated microcosm, because it is capable of occupying the niche at the air-liquid interface. To test this hypothesis, LSWS-Gm was added to microcosms saturated with SBW25-Δ*wss* as previously performed for SBW25-Gm. The starting density of LSWS-Gm relative to SBW25-Δ*wss* was varied from 10^{-3} for to 10^{-8} and a negative control in which no cells were added was also included. Microcosms were incubated for either 24 h or 48 h and sampled destructively by

vortexing, and the proportions of LSWS-Gm relative to SBW25-Δ*wss* were counted (Figure 4-4). The results of a Tukey's pair-wise comparison showed that the proportion of LSWS-Gm at 24 h was not significantly different from the proportion at t_0; however the proportion of LSWS-Gm at 48 h was significantly different from the proportions at both t_0 and 24 h across all starting densities ($\alpha = 0.05$, $df = 18$).

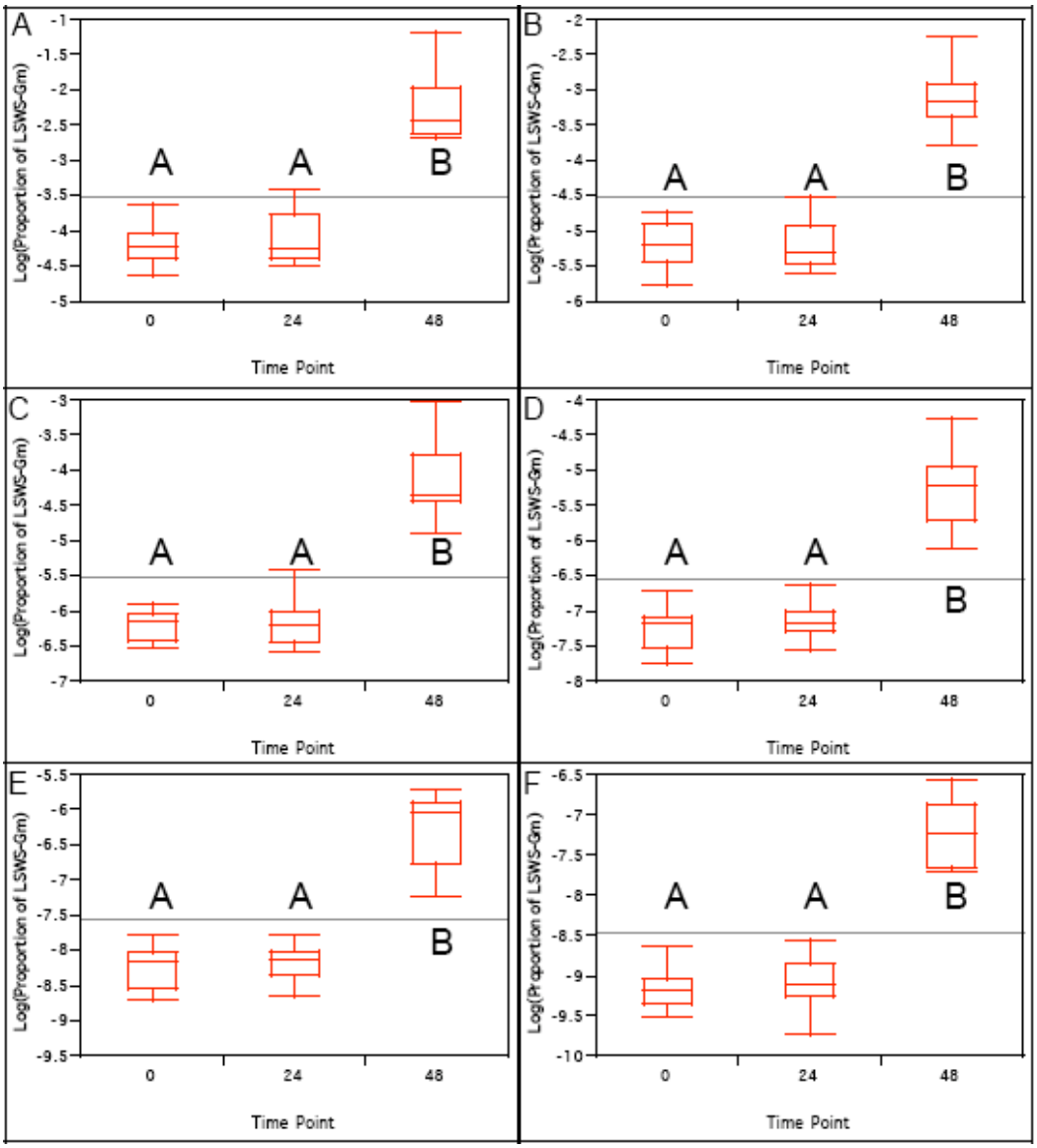

Figure 4-4: Change in the proportion of LSWS-Gm with respect to time at multiple densities. The box plots for log(A) – log(F) are for densities ranging from 10^{-3} to 10^{-8}. The results of a Tukey's pair-wise comparison are indicated by the group letter ($\alpha = 0.05$, $df = 18$).

Combining the results for SBW25-Gm and LSWS-Gm in broth-saturated microcosms justifies the conclusion that growth was possible only at the air-liquid interface, because no further growth of SBW25-Gm was observed while LSWS-Gm was able to

increase in frequency at the air-liquid interface between 24 h and 48 h. Therefore, growth was successfully restricted to the niche at the air-liquid interface, a requirement for the main experiment.

4.2.3 The Fitness Dynamics of LSWS with Respect to Density and Time

The results in the previous section showed that broth-saturated microcosms allowed growth only at the air-liquid interface by WS cells, because the proportion of LSWS Gm cells increased whereas the proportion of SBW25-Gm did not. Importantly, the increase in proportion of LSWS-Gm was not immediate, but rather was only observable during the second 24 h time interval. To analyse the differences in fitness at the air-liquid interface between SBW25-Gm and LSWS-Gm for the intervals 0-24 h and 24-48 h, the selection rate constant (SRC, Lenski *et al.*, 1991) was calculated for all six starting densities (Figure 4-5)[1]. The SRC data are not independent across densities; therefore, to test for a difference between the first and second time intervals a variant of a randomisation test was performed on the experimental and control treatments separately (Methods 7.2.23.2). A straight line was fit to the SRC values from both time intervals combined, and the residuals calculated for each. An ANOVA of the residuals revealed that there was no significant difference in fitness between the first and second time intervals for SBW25-Gm ($F_{1,119}$ = 0.3821, p = 0.5377), but that the fitness for the second time interval was significantly higher than that of the first time interval for LSWS-Gm ($F_{1,119}$ = 800.4, p < 0.0001).

[1] The selection rate constant was preferred over the Malthusian ratio as a measure of fitness, because it is less sensitive to low counts of one competitor and allows for cell death.

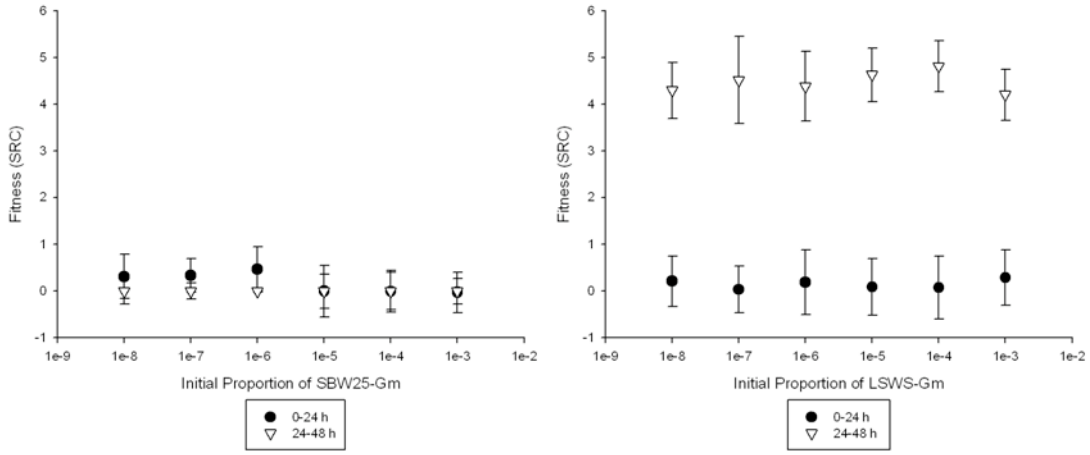

Figure 4-5: Fitness of SBW25-Gm control and LSWS-Gm treatment in broth-saturated microcosms.
The fitnesses (SRC) of SBW25-Gm (left) and LSWS-Gm (right) during the first 24 h growth (black circles) and the second 24 h growth (clear triangles). Error bars are 95% confidence intervals from ten replicates.

For direct comparison with the expectations outlined in Figure 4-1, a linear regression was performed on the fitness values (SRC) of LSWS-Gm against initial density from each time interval separately (Figure 4-6). For both the 0-24 h and 24-48 h time intervals, the regression line was not significantly different from zero ($T_{58} = 0.20$, $p = 0.840$ and $T_{58} = 0.29$, $p = 0.773$ respectively). This did not match the expectation for group-limited cooperation, but rather appeared to fit the null model of 'minimal cooperation' but with a clear effect of time. Therefore, despite the obvious emergent phenotype, these data show that the fitness benefit to WS cells within the mat was not dependent on the density of WS cells.

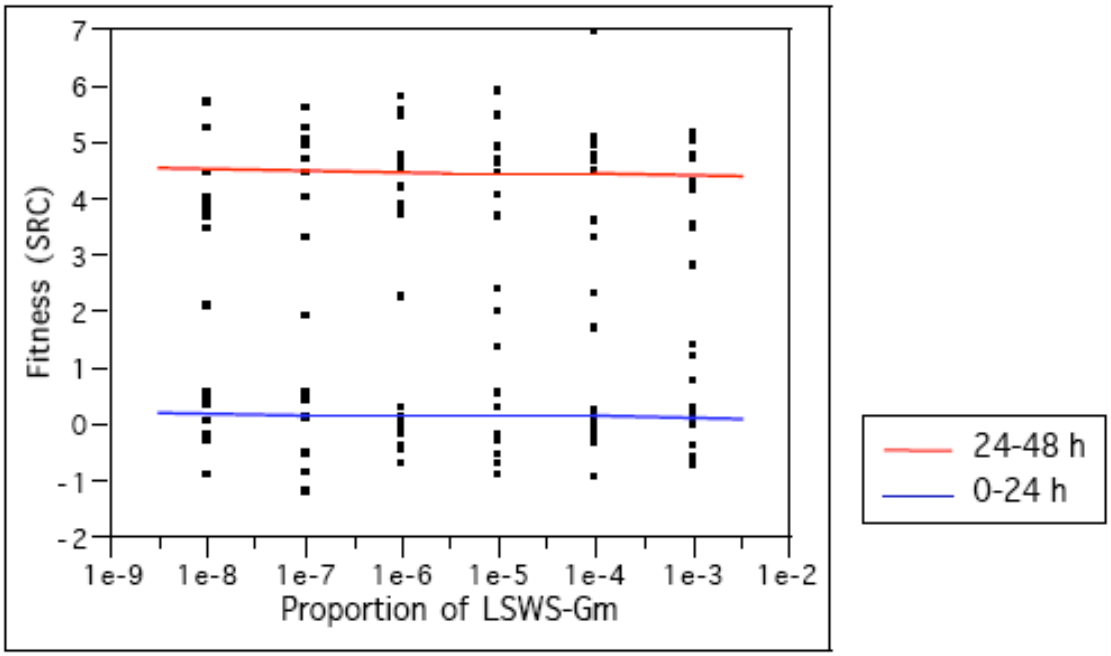

Figure 4-6: SRC for LSWS-Gm at multiple densities for the time intervals 0-24 and 24-48 h.
The blue line was fit to the fitness during the time interval 0-24 h. The red line was fit to the fitness during the time interval 24-48 h. The X-axis is on a log scale for the proportion of LSWS-Gm.

Despite the lack of density dependence, there was a significant difference in fitness between the first and second time intervals suggesting that the benefit of mat formation at the air-liquid interface is *time dependent*. The emergence of positively synergistic interactions over *time* suggested a role for a development-like process in which associations among cells growing together were formed as the WS cells divided. One hypothesis that extends from this suggestion is that disruption of the associations among the WS cells should lead to the loss the emergent fitness benefits associated with the emergent phenotype. Therefore, a second experimental treatment was used to test the hypothesis that the development-like process that leads to the fitness benefit for LSWS-Gm in a broth-saturated microcosm could be destroyed by vortexing. In the LSWS-Gm[interrupted] treatment, LSWS-Gm was inoculated into broth-saturated microcosms at multiple starting densities as previously performed for the first experimental treatment and the control. The difference between the first experimental treatment and the interrupted treatment was that the microcosms due to be incubated for 48 h were vortexed for 30 s at 24 h to destroy the association among the WS cells in the mat before being left to incubate statically for the remaining 24 h. Microcosms were sampled destructively at 24 h and 48 h by vortexing, and fitness (SRC) of SBW25-Gm relative to SBW25-Δ*wss* was calculated (Figure 4-7).

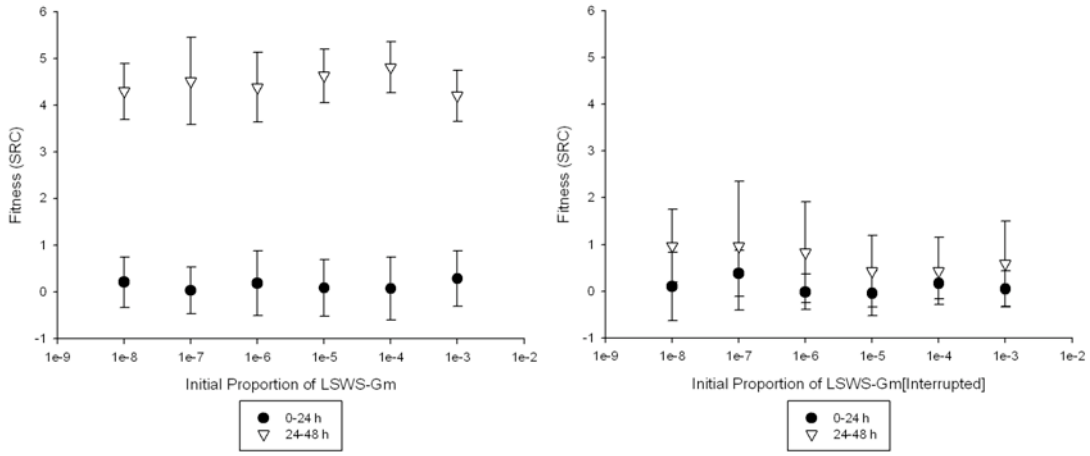

Figure 4-7: Fitness of LSWS-Gm treatment and LSWS-Gm[interrupted] treatment in broth-saturated microcosms.
The fitnesses (SRC) of LSWS-Gm (left) and LSWS-Gm[interrupted] (right) during the first 24 h growth (black circles) and the second 24 h growth (clear triangles). Error bars are 95% confidence intervals from ten replicates.

The comparison between the results for LSWS-Gm and LSWS-Gm[interrupted] show that the observed fitness increase with respect to time for LSWS-Gm can be destroyed by vortexing of the sample at 24 h. Although there is still a significant difference in fitness between 0-24 h and 24-48 h ($F_{1,119}$ = 11.054, p = 0.001), the reduction in fitness due to vortexing is larger and also highly significant ($F_{1,119}$ = 365.2, p < 0.0001).

4.3 DISCUSSION

4.3.1 THE EMERGENT PHENOTYPE FROM WS COOPERATION IS THE PRODUCT OF A DEVELOPMENT-LIKE PROCESS

The defining characteristic of all WS types is the formation of a mat at the air-liquid interface, an unambiguously emergent group-level phenotype under the emergent character criterion (Vrba, 1983). The molecular basis for mat formation is the costly production of cellulose that causes WS cells to adhere to the sides of the microcosm and to one another after cell division, ultimately offsetting the cost of cellulose production by the benefit of exposure to oxygen at the air-liquid interface (Rainey and Travisano, 1998; Rainey and Rainey, 2003). These results have shown that only WS types can obtain the benefits of growing at the air-liquid interface, but also that the fitness benefit can be lost by disruption of the mat. Counter to the expectations for cooperation, this fitness benefit to WS cells at the air-liquid interface was not dependent on density. This challenges the paradigm that suggests that the fitness of cooperators is density dependent. In fact, the coverage of six orders of magnitude for density in this experimental design illustrated that microcosms with fewer than 100 WS cells were capable of enjoying the benefits of occupying the air-liquid interface. However, in all cases the benefit was only observed after 24 h of growth (*i.e.* during the second 24 h of growth). The observation that the fitness benefit is independent of density suggests that the growth advantage is available even to *individual* cells, and the observation that 24 h of prior growth is required suggests that the growth advantage is dependent on time. The importance of this result is two-fold. First, if the trait is individually advantageous when it first arises, the investigation shifts from understanding the *origin* of the cooperative trait to understanding its *maintenance*, because once established the cooperative trait is still vulnerable to exploitation by the evolution of types that do not pay the cost. Secondly, new theory is required to explain how a group-level phenotype might emerge as a function of time.

One suggestion about how time can lead to the emergent phenotype evident in WS mat formation can be made by analogy with development. Development as seen in

many eukaryotes refers to the coordinated cellular growth, differentiation and morphogenesis of a multicellular organism (Gilbert, 2000). Two particular parallels are important when viewing WS mat formation as a development-like process. First, daughter cells at each cell division remain attached to one another, and second, as a consequence of cells remaining attached to one another, the group becomes a three-dimensional structure (morphogenesis)[1]. This group phenotype and its properties are all emergent properties with respect to the cells undergoing development. The critical point is that disruption of the association among the cells of, for example, a developing or developed metazoan, would lead to a catastrophic loss of its emergent phenotype. This is exactly what was observed for WS cells in the interrupted treatment. If the growth of the mat is disrupted during its formation (or later on by the evolution of cheats) there is a total loss of the emergent phenotype. By contrast, in other classical examples of microbial cooperation, *e.g.* the extracellular secretion of invertase (Greig and Travisano, 2004) or siderophores (Griffin *et al.*, 2004), 'disruption' of cells does not disrupt the emergent density-dependent fitness effect (in fact, microcosms are usually continuously shaken to maintain homogeneity). This enhances the distinction between WS mat formation and other classical examples of cooperation. Finally, these differences highlight the fact that WS mat formation is a suitable system for studying the evolution of multicellularity (Chapter 5), while classical examples of extracellular secretion are not.

4.3.2 Inadequacies about Current Cooperation Theory for Microbes

The natural life histories of microorganisms consist of an abundance of complex interactions within and among species that are either antagonistic or synergistic (Thompson, 1999). Recently cooperative interactions among microorganisms have piqued the interests of evolutionists and microbiologist alike, because they appear to

[1] The other characteristics of development, *i.e.* that the growth is coordinated and that cells differentiate, are not necessary for this analogy.

be costly social behaviours (Velicer, 2003; Sachs *et al.*, 2004; West *et al.*, 2006), and because of the hypothesised role in the evolutionary transition from unicellular to multicellular organisms (Buss, 1987; Maynard Smith and Szathmáry, 1995; Michod, 1999; Michod and Roze, 2001; Michod *et al.*, 2005). The key feature of cooperative interactions that generates this interest is that cooperative interactions lead to emergent properties that are 'not otherwise attainable' (Corning, 2002), because the effects are quantitatively or qualitatively 'greater than the sum of the parts' (Gould, 2002). This principle raises two matters of concern with respect to the current cooperation theory for microbes: firstly, the introduction of the term minimal cooperation, and secondly, its (mis)application to biological examples. A third general concern about the limitations of identifying cooperation based on density dependence is also discussed.

The first concern is aimed at the introduction to the literature of a logically inconsistent term, minimal cooperation, in which there is no emergence from a cooperative interaction. As Velicer (2003 p. 322) states, it "is maintained solely by individual level selection", therefore, "it should favour isolated living rather than social life". This is logically inconsistent for cooperative interactions for two reasons that can be demonstrated with reference to the criteria for cooperation established in Section 1.2.2, and by considering the equations at the beginning of the chapter. Cooperation is an interaction, and therefore a component of the fitness for both the actor, X, and the recipient, Y, is determined by that interaction. For minimal cooperation to be valid, and for X to evolve solely by individual selection the interaction, ω_{X*Y}, must equal zero. However, if ω_{X*Y} does equal zero, there is either no interaction, which violates the first criteria for cooperation, *i.e.* that there is an interaction (Section 1.2.2) or the interaction is commensal, amensal or neutral. Thus, minimal cooperation is inappropriate for describing cooperative interactions among microorganisms. However, the appeal of the idea that bacteria evolved to cooperate has led to an uncritical over-acceptance of a cooperative paradigm. As a consequence, the absence of a cooperative interaction (minimal cooperation) has been forced to fit

into a cooperative paradigm in which it does not belong.

The second concern with the current cooperation theory is the misclassification of biological phenomena as minimal cooperation, because given the reasoning above, only one of two things can be true: (1) that the *action actually is* a type of cooperation, and an emergent fitness or phenotype from the cooperative interaction can be determined, or (2) that an *action is not* an example of cooperation. Therefore, the classification of biological phenomena as minimal cooperation is uninformative, and the burden of evidence remains with the researcher to determine whether or not an emergent property due to cooperation exists. The examples of minimal cooperation suggested by Velicer (2003 p. 322) include "the secretion of any diffusible substance (such as a siderophore, cellulolytic enzyme or anticompetitor toxin) that acquires, releases, or protects previously inaccessible or threatened resources." These examples equate to public goods (Section 1.2.2.2), and therefore, must be empirically determined to meet the requirements of a cooperative trait before the results can be interpreted in a cooperative paradigm. However, the cooperative paradigm is commonly assumed rather than determined experimentally (Shapiro and Dworkin, 1997; Shapiro, 1998; Crespi, 2001; Buckling *et al.*, 2003; Velicer, 2003; West and Buckling, 2003; Greig and Travisano, 2004; Griffin *et al.*, 2004; Harrison *et al.*, 2006; West *et al.*, 2006; Buckling *et al.*, 2007; Foster *et al.*, 2007; Ross-Gillespie *et al.*, 2007; Sandoz *et al.*, 2007; Harrison *et al.*, 2008). The evidence for siderophore secretion as an example of cooperation has been examined and found to be lacking the theoretical and empirical support to be considered a cooperative trait, because the correct experiments have not been done with the appropriate bacterial strains (Appendix 9.4.1.1). However, even if experiments show the expected density dependent relationship using appropriate strains and experimental design, this may not be sufficient to conclude cooperation, because density-dependence is also predicted by diffusion sensing (Redfield, 2002) and it may be a response to a cue or chemical manipulation (Keller and Surette, 2006).

The third concern addresses the above-mentioned short-comings that are inherent in

simply using the density-dependent effects of the third and fourth criterion (Section 1.2.2) to identify a trait as cooperative. The objections are clarified by trying to answer two questions. First, does density dependence for an extracellularly secreted compound justify the conclusion that the behaviour is cooperative? Second, given that WS cooperation leads to an emergent phenotype that is not density dependent, what other non-density-dependent cooperative interactions could give rise to emergent group-level properties?

Among bacteria, density dependent behaviours and quorum sensing dynamics have been used as the principal features that verify the extracellular secretion of compounds as cooperative interactions (Henke and Bassler, 2004; Parsek and Greenberg, 2005; Diggle *et al.*, 2007b; Sandoz *et al.*, 2007; Dunny *et al.*, 2008; Nadell *et al.*, 2008). However, this rests on weak foundations, "as neither the need for group action, nor the selective conditions required for its evolution have been demonstrated" (Redfield, 2002 p.365). In a typical example, a small amount of a compound is secreted outside of the cell, and simultaneously the concentration of the compound is sensed at the surface of the cell. If the concentration exceeds a certain threshold, expression is upregulated or, as is common in quorum sensing, the expression of new products is induced. Many superficially plausible, but untested hypothesis about the benefits of density dependent activity and quorum sensing have been proposed (Fuqua *et al.*, 2001; Miller and Bassler, 2001; Foster *et al.*, 2007; Dunny *et al.*, 2008; Nadell *et al.*, 2008); however, a more direct function of extracellularly secreted compounds can be satisfied by diffusion sensing in which cells regulate these compounds by sensing the environment directly and not the density of other cells. Benefits can be realised for a single cell provided that diffusion and mixing are limited in the cell's microenvironment, so that secreted molecules remain close enough to the cell (Redfield, 2002). In addition, it has been claimed that the nature of the some quorum sensing interactions do not involve cooperative signals, but rather are responses to cues and chemical manipulation (Keller and Surette, 2006; Diggle *et al.*, 2007a). Given that predictions from alternative hypotheses like diffusion sensing,

response to cues and chemical manipulation are indistinguishable from the predictions of cooperation, it is concluded that density dependence in examples of public goods is insufficient to claim cooperation.

The current framework assumes an *a priori* role for density in all types of cooperation; however, work in this chapter confirms that group-level phenotypes can emerge as a result of cooperation without an effect of density. This differs from the experimentally verified density-dependent fitness for the production of extracellularly secreted invertase (Greig and Travisano, 2004), all quorum sensing examples (Fuqua *et al.*, 2001; Miller and Bassler, 2001; Dunny *et al.*, 2008; Nadell *et al.*, 2008) and density dependent fitness among cellular aggregates of *M. xanthus* (Rosenberg *et al.*, 1977) or *Dictyostelium discoideum* (Foster *et al.*, 2002) and the theoretical expectations outlined at the beginning of this chapter (Velicer, 2003). However, the criterion of emergence is still satisfied for cooperation among WS cells, because the qualitatively different emergent group-level phenotype can be observed visually and detected by changes in fitness over time. Therefore, it is a concern that the consideration of bacterial cooperation strictly in terms of population density-dependence may overly restrict the research in this area and impede the discovery of genuine examples of cooperation in the microbial world.

4.3.3 Concluding Remarks

In this chapter, the dynamics of an emergent group-level phenotype has been investigated, and it has been determined that the fitness of individuals contributing to the emergent group-level phenotype is not dependent on density, but is dependent on a development-like process. This may initially seem like a paradox, because emergence is predicted to be density dependent; however, this is only true if fitness is measured in an MLS 1 framework in which group fitness is an aggregate fitness of the individuals. Alternatively, if group-fitness is measured in an MLS 2 sense, the density of individuals in the group is arbitrary with respect to fitness, because group fitness is defined independently of particle fitness (Okasha, 2006a). For example, it is clear that small groups and large groups have different characteristics that may affect

the viability of the group that are a function of the number of individuals (*e.g.* mat strength, buoyancy or 'time to collapse'); however, such group-characteristics are arbitrary with respect to actual group fitness in an MLS 2 sense, because they say nothing about the ability of groups to reproduce new group-offspring (group-fecundity). Therefore, the lack of density-dependent emergence in MLS 1 is not actually a paradox when considered in MLS 2. This reinforces the view that the relevant measure of group-fitness is the reproduction of group-offspring (group-fecundity). The importance of group-fecundity rather than group-viability is one of the foundations of the novel theory developed in Chapter 5.

5 SELECTION FOR GROUP REPRODUCTION VIA A DEVELOPMENT-LIKE PROCESS

> *An adaptation at the population level is any mechanism which enhances the tendency of a population to send out propagules and establish additional populations in new or vacant habitats. Under this definition the tendency for a population to persist in a habitat for a long period of time is not considered an adaptation unless the number of new populations founded by the population in question is positively correlated with its persistence. Similarly, an increase in the rate of production of propagules is not considered an adaptation unless this increase results in an increase in the rate of establishment of new populations. A population which sends out fewer but more successful propagules is considered to be better adapted, in terms of this definition, than a population which sends out ten times as many propagules but establishes fewer new populations. This definition, therefore, defines a populational adaptation to be a trait which is the focus of or the result of a process of group selection (Wade, 1977).*

--Wade, M.J. 1977

5.1 INTRODUCTION

5.1.1 PREAMBLE

The transition from free-living single cells to multicellular individuals involves two distinct phases (Bonner, 2001), the cooperation among individual cells to form multicellular groups (MLS 1 framework), followed by the emergence of a higher-order multicellular individual (MLS 2 framework). The distinction between multicellular groups and a multicellular individual dovetails with MLS 1/MLS 2 distinction, because the hereditary particle shifts from being a single cell to a group of cells (Okasha, 2006a). In other words, in multicellular groups, the groups do not reproduce group-offspring, but for multicellular individuals, reproduction of group-offspring (group-fecundity) is required so that the group can meaningfully participate in the process of evolution by natural selection (Lewontin, 1970). Therefore, a multicellular individual is only defined as such if it is capable of reproducing group-offspring. This presents a two-fold paradox for evolutionary theory. First, as with other examples of cooperation, cooperation among lower-level entities will be susceptible to the evolution of selfish types that undermine the integrity of the individual (Buss, 1987; Maynard Smith and Szathmáry, 1995; Michod, 1999). Second, it conflates product with process to require group-offspring to be both a criterion for the identification of a multicellular individual and the process by which group-offspring must arise (Griesemer, 2001; Okasha, 2006a). Various mechanisms have been proposed to limit conflict with selfish types at the lower level (Michod, 1999; Michod and Roze, 2001; Travisano and Velicer, 2004; Michod *et al.*, 2005), but in most cases, these mechanisms assume an MLS 2 framework and ignore the anteceding need to account for the evolution of individuality at the higher level.

Firstly, to address these issues, I will introduce a novel perspective that resolves the paradox presented by the reproduction of group-offspring in the transition to individuality, by including a role for a development-like process (Section 5.1.2). As a consequence of incorporating a development-like process into the transition to individuality, the fates of cells at the lower level become aligned with the fate of the multicellular individual, *i.e.* the individuality of multicellular organisms, becomes a property of the individual cells as well. Thus, the conflict among the cells at the lower

level is ameliorated as a by-product of selecting for a development-like process, by simultaneously selecting at the higher and lower levels. Secondly, I will demonstrate that *P. fluorescens* is a suitable experimental system for determining a response to selection at multiple levels (Section 5.1.3). Thirdly, I will introduce the main aims and results of experiments in this chapter (Section 5.1.4).

5.1.2 A Novel Perspective – Development of Higher Levels of the Biological Hierarchy

The expansion of the biological hierarchy is characterised by emergent properties at the higher-level that result from interactions among lower-level units (Maynard Smith and Szathmáry, 1995; Michod, 1999; Okasha, 2006a). The observation in the previous chapter that a development-like process can lead to an emergent phenotype has fuelled the desire to advance a novel theoretical framework for the major evolutionary transition to multicellular individuals. In this novel perspective, we adopt the starting scenario of Bonner (1998) where a group of cells forms as a consequence of the failure of cells to separate from one another after cell division. Continued division among cells that do not separate leads to the formation of a sessile group of cells that has dimensionality (shape or morphology) and structural organisation not present in populations of free-living single cells – such sessile cells will be called 'group formers' (GF). Interactions among GF cells on surfaces lead to the emergence of non-random structures that are simply dependent on the cellular adhesion and the environment, a minimalistic development-like process. However, in terms of the evolution of individuality, it would appear that structured groups formed by GF cells are an evolutionary *cul de sac*, because, although individual cells within the group continue to divide (group-viability), groups as a whole do not (group-fecundity).

One suggestion is that if types arise that can separate from the group and retain the potential to revert to GF types then group reproduction can exist (Rainey, 2007; Rainey and Kerr, 2009) – such motile types will be called 'swimmers' (SWM). From this perspective, SWM types arising within the group of GF types would lower the integrity of the group, but the evolutionary *cul de sac* is avoided, because the SWM types can swim away from the group to a new location. At the new location, SWM types revert to GF types, adhere to one another and begin to form new group-

offspring to complete one group-generation[1]. The SWM types that swim away from the parent group act as propagules of new groups (*i.e.* a primordial germ-line capable of dispersal), with GF types relegated to a role analogous to the soma.

To simultaneously select at the higher and lower levels, selection must act on the rate of change between the GF and SWM types and *vice versa* (GF ⇔ SWM). Consider an expanded version of the above example. The starting situation is the same as above with a few GF cells that have failed to separate in an environment that favours group formation. Multiple groups begin to form and all the GF cells in all the groups benefit from their group membership. Within each group SWM types arise that can potentially swim away from the group, but since they arise by chance, they make up a different proportion of the group depending on when they arise. However, there is a trade-off for each group, because if too many SWM types arise, the group loses its integrity and goes extinct, but if too few SWM types arise then the group will have too few propagules and be out-competed in the reproduction of group-offspring by groups with more SWM types (Rainey, 2007; Rainey and Kerr, manuscript in preparation). Therefore, because the number of SWM types within a group depends on the rate at which they arise, selection acting at the group level will optimise the rate at which SWM types arise from GF types and ultimately stabilise the multicellular individual. The same is true for the rate at which GF types arise from SWM types, because the most successful groups at the new location will be determined by the rate at which the competing individual SWM types revert to being GF types.

5.1.3 *P. FLUORESCENS* IS A SUITABLE SYSTEM

The *P. fluorescens* system presents a unique opportunity to empirically test predictions derived from this novel perspective. WS types are sessile group formers (GF), because they adhere to one another after cell division through the production of

[1] A reminder that in this section groups are being treated in the MLS 2 sense, so group-generation has the more intuitive meaning of the period of time between the reproduction of successive group offspring.

bacterial cellulose and a proteinaceous adhesin (Spiers *et al.*, 2003). The mat at the air-liquid interface is the product of a development-like process that leads to an emergent group-level phenotype (Chapter 4). Ancestral-like SM types that do not produce cellulose arise within the mat and enjoy a selective advantage by avoiding the cost of cellulose production, but weaken the mat (Rainey and Rainey, 2003). In *Vibrio cholerae* biofilms, types that do not produce exopolysaccharides (EPS) are known to actively increase dispersal (Zhu and Mekalanos, 2003; Nielsen *et al.*, 2006; Liu *et al.*, 2007), which is further evidence to suggest that the SM types that do not produce cellulose that arise within the mat fit the profile for the motile swimmers (SWM) in the broth phase. In addition, the capacity for repeated switching between GF and SWM by compensatory mutation is known for the *P. fluorescens* system (Beaumont *et al.*, 2008). Ultimately, a multicellular individual would have to replace a genetic switch with a phenotypic switch, but this is by no means beyond the capacity of evolution (Beaumont *et al.*, 2008). Although the SM/WS distinction and its correlation with phenotype becomes less clear with increasing numbers of phenotypic reversals, the possibility to select repeatedly for alternation between the two phenotypes (i.e. SW: motile in the broth phase, and GF: group forming at the air-liquid interface) lays the foundation for investigating this novel perspective.

5.1.4 AIMS AND RESULTS

The primary aim of this chapter was to observe an adaptive response to a selective regime that selects both at the level of the individual and at the level of the group in an MLS 2 framework. Two pre-experiments were required before commencing with the main experiment. Firstly, novel apparatus was developed to meet the major design considerations of the experiment (Section 5.2.1). Secondly, *P. fluorescens* was grown in the novel environment for ten days to determine whether the dynamics of the adaptive radiation exhibited the features required by the novel theoretical perspective (Section 5.2.2). The main experiment was divided into two phases, an evolution phase during which there was an opportunity for adaptation to the experimental regime, followed by a competition phase during which the derived types were competed against the ancestor to determine if the derived types had adapted to the experimental regime (Section 5.2.3). The competition phase is effectively an MLS 2

fitness assay in which the response to the selective regime was determined by counting the number of groups, not individuals.

5.2 RESULTS

5.2.1 THE NOVEL APPARATUS

To investigate this novel framework experimentally, the apparatus needed to meet two environmental requirements. Firstly, multiple independent groups should be able to form within a single environment and secondly, individuals within each of those groups should to simultaneously compete for the same resources. In addition, the apparatus needed to meet three practical requirements. Firstly, the apparatus needed to be easily sterilisable, secondly, it needed to be possible to sample sterilely, and thirdly, it needed to be sufficiently small make repeated experimentation tractable.

After numerous prototypes, the final growing environment was a deep Petri dish, with a 'lid adaptor' resting on the rim of the Petri dish (Figure 5-1). The adaptor was made of a 1 cm thick piece of PTFE (poly-tetrafluoroethene) an autoclavable[1] fluoropolymer. Seven holes were drilled through the PTFE lid adaptor, and the top was designed to be wider than the bottom to create a small 'lip' so that it could be rested on the rim of a Petri dish. The seven holes provided seven independent opportunities for groups to form at the air-liquid interface connected by the media below. Therefore, this modified Petri dish (MPD) afforded the opportunity for the formation of distinct groups, and retained competition among individuals from all groups for shared media resources and oxygen.

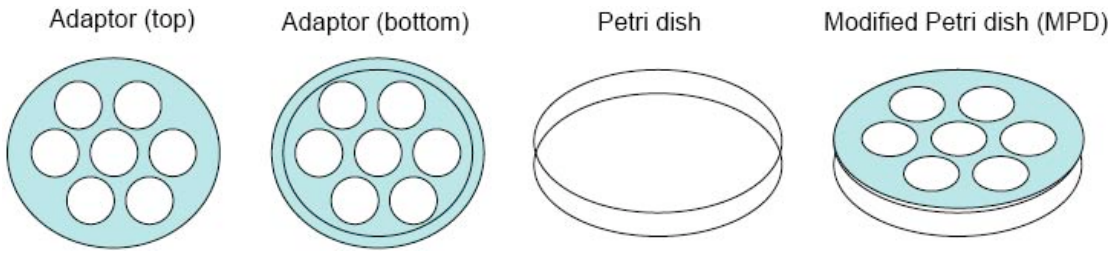

Figure 5-1: Schematic diagram of the Petri dish adaptor and the modified Petri dish (MPD).
From left to right the diagram shows the design of the Petri dish 'lid adaptor' from above, the design from below, an empty Petri dish, and the way the adaptor sits on the empty Petri dish. During incubation, the lid of the Petri dish was placed on top of the MPD.

[1] Autoclave temperatures (121˚C) are well within the maximum operating temperature of 260˚C.

PTFE (commonly used under the trade-name Teflon®) is widely used industrially and commercially, because it has a strongly 'non-stick' surface. Despite this, when the MPDs were filled with KB and inoculated with *P. fluorescens*, WS cells successfully attached to the PTFE surface and formed mats in the holes of the Petri dish lid adaptor (Figure 5-2). A threaded-hole was created in the back of the lid adaptor so that a 6 mm diameter, 10 cm long bolt could be gently screwed in, and that the adaptor could be gently and sterilely placed on or removed from the Petri dish.

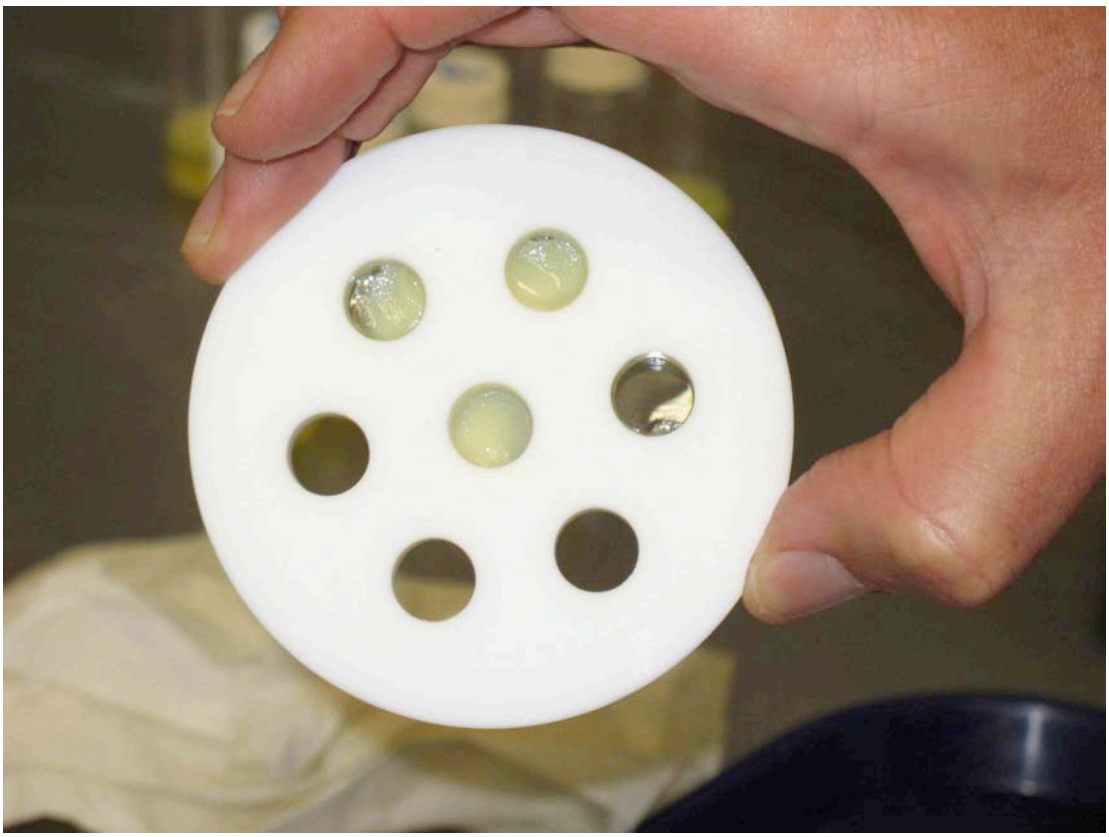

Figure 5-2: A Petri dish lid adaptor showing mat growth.
This picture shows the prototype lid adaptor before a threaded hole was made to allow for insertion of the 10 cm bolt.

Early trials investigated the sampling of mats from the lid adaptor in their entirety by gentle detachment. However, while detachment was possible (Figure 5-3), it was unfeasible, because the process was too destructive to be implemented reliably as part of an experimental protocol. Therefore, in the main experiment (Section 5.2.3) it was determined that individual mats growing at each hole would be sampled with a sterile wire loop rather than removed in their entirety (Methods 7.2.24.2).

Figure 5-3: Mats removed from the holes of the Petri dish lid adaptor.
Six mats were successfully removed from the lid adaptor and gently washed in a Petri dish containing ¼ Ringer's solution.

5.2.2 DIVERSIFICATION IN A MODIFIED PETRI DISH

To determine if the principle features of the known adaptive radiation were conserved for SBW25 growing in the modified Petri dishes (MPDs), wild-type SBW25 was inoculated into 30 MPDs and incubated for between one and ten days (Methods 7.2.24.1). Every 24 h, three MPDs were sampled destructively and the entire contents diluted and plated appropriately to estimate the density of WS and SM types (Figure 5-4).

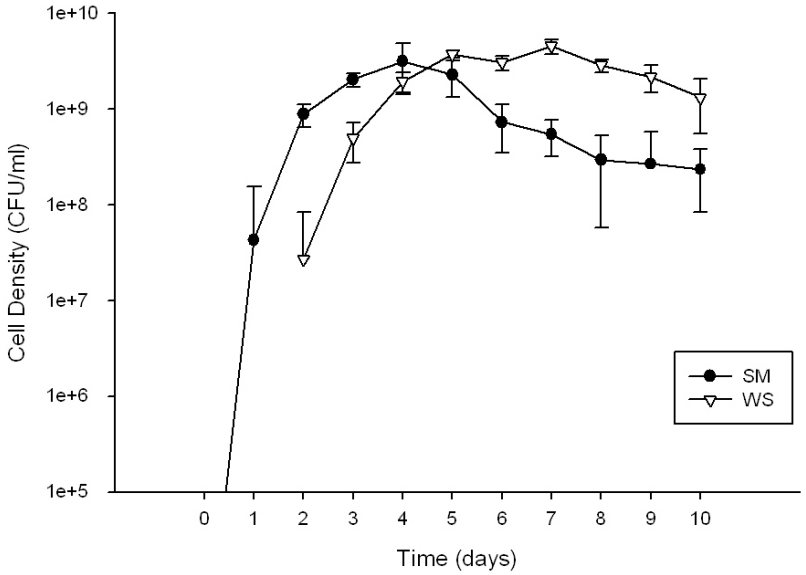

Figure 5-4: Adaptive radiation in a modified Petri dish.
SBW25 was inoculated into an MPD and grown statically for ten days. WS types arise at day two and increase in frequency to become the dominant type by day five. Error bars represent 95% confidence limits.

The adaptive radiation in an MPD showed diversification into a variety of WS genotypes that were first detected at day two. By day five, WS types dominated the MPD, and mats were clearly visible in each of the available niches at the air-liquid interface. Both these observations are comparable to the adaptive radiation known for SBW25 in a standard KB microcosm (Rainey and Travisano, 1998; Section 3.2.1). As a consequence of these data, the incubation period in the MPDs was set at five days, because this was a sufficient incubation period for WS types to arise *de novo* to become the dominant morphological class. Although these data cannot be statistically compared to the adaptive radiation in a standard microcosm for sampling reasons, the major features of the adaptive radiation are at least qualitatively consistent, *i.e.* a variety of the usual derived types were observed[1].

5.2.3 Selection at the Level of the Group and the Level of the Individual

Previous hypotheses about the transition to multicellular individuals had focussed on selection at the level of groups alone. However, with the hypothesised role for development, the primary aim of this chapter was to determine if it were possible to observe an adaptive response to a selective regime that selects both at the level of the individual and the group in an MLS 2 framework. Selection experiments of this nature are typically broken up into two phases, an evolution phase during which there is an opportunity for adaptation to the experimental regime, followed by a competition phase, during which the derived types are competed against the ancestor to determine if the derived types have adapted to the experimental regime. To control for the multi-level effect, ancestral types were evolved under two treatments, the experimental multi-level selected treatment (MLS) and a control treatment that selected at the level of groups alone (group level selected treatment; GLS). The two treatments, each replicated four times, were evolved in a paired-parallel design (Figure 5-5). The experiment was of parallel design, because all four replicates of both treatments were performed simultaneously, and of paired design, because the

[1] FS types occurred occasionally, but were not scored in this experiment, because only the ratio of SM to WS was of interest.

starting genotypes in each MLS replicate were identical to those in each GLS replicate, but the genotypes within each of replicate (replicates 1-4) were different from one another.

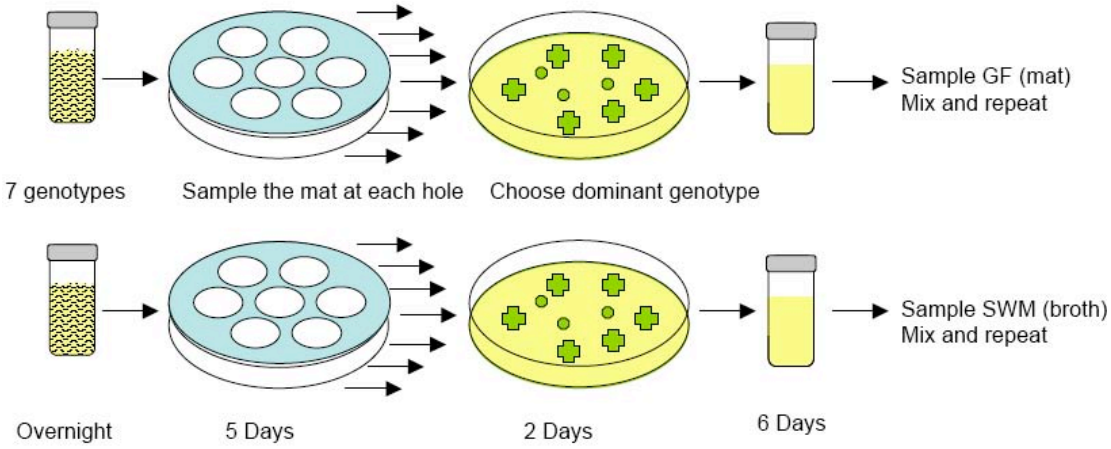

Figure 5-5: Experimental design for the MLS and GLS treatments.
The diagram shows a single replicate for the GLS treatment (top), and the MLS treatment (bottom). For each treatment, seven genotypes were grown up overnight and inoculated in equal volumes into each MPD. The MPDs were incubated for five days then each of the seven mats in the holes at the air-liquid interface were sampled independently with a sterile wire loop, diluted and plated on KB agar plates. After two days of incubation, the dominant genotype was inoculated into a static microcosm and incubated for six days. After this incubation, the treatments diverge. In the GLS treatment the mat was sampled with a sterile wire loop, diluted and inoculated into a fresh MPD to complete one group-generation cycle. In the MLS treatment the broth was sampled, diluted and inoculated into a fresh MPD to complete one group-generation cycle.

To ensure diversity in the paired design, seven independent WS types (one genotype per growth area at the air-liquid interface) were used to initiate each replicate (Methods 7.2.24.2). For each replicate of both treatments, the seven independent WS genotypes were grown overnight and equal volumes of each were added to each MPD. After five days, each of the seven growth areas at the air-liquid interface was sampled independently with a sterile wire loop, diluted and plated. After two days of incubation at 28˚C, the dominant type from each of the seven plates was inoculated into a standard microcosm and incubated for six days. At the end of this incubation, the *mat* was sampled for each microcosm in the GLS treatment to select GF cells, while the *broth* was sampled for each microcosm in the MLS treatment to favour SWM cells. In both treatments, the sample was diluted and inoculated into a fresh MPD to begin the next group-generation. The experimental regime was repeated for six group-generations and the genotypes obtained at each time point stored at -80˚C.

To determine if there was an adaptive response to the selective regime employed during the evolution phase of this experiment, one randomly selected dominant morphotype from the final time point of each replicate of each treatment was selected for further experimentation. Each of these types, called M1-M4 and G1-G4, were competed against the marked strain LSWS-*lacZ* (non-derived – used as the ancestor[1]) under the MLS experimental regime to determine a response to selection. The competition phase is essentially a fitness assay over six group-generations in an MLS 2 framework, where groups are counted instead of individuals as either ancestral groups (blue groups) or derived groups (white groups). A group is counted either 'white' or 'blue' when that colour is the dominant type within the group, because in this selective regime only the dominant type is selected by the experimenter. Therefore the dominant type is the only type that has the opportunity to generate propagules. The number of derived group-offspring was counted for both the GLS treatment and the MLS treatment at each group-generation (expressed as a proportion) and a trendline fitted to the average of the four replicates (Figure 5-6).

[1] Even though the marked LSWS strain is not technically an ancestor, its use in this sense is justified, because it is an independent seven-day mutant genotype, as were all the starting genotypes in the evolution phase of the experiment.

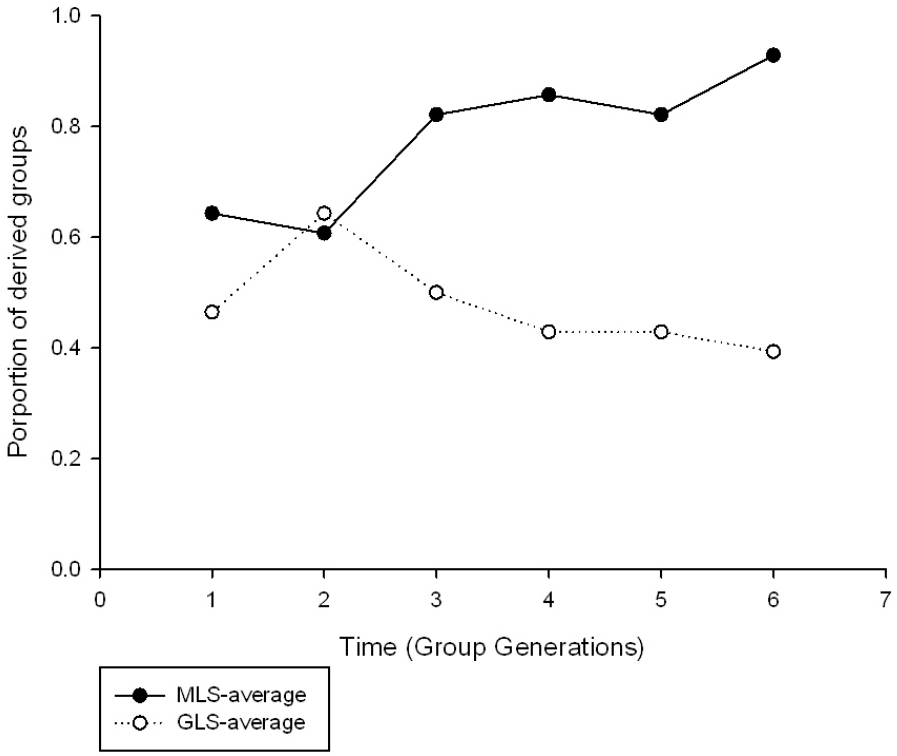

Figure 5-6: Change in proportion of derived groups over time for the GLS and MLS treatments. The trendlines show a general increasing trend for the MLS treatment and a general decreasing trend for the GLS treatment. Error bars were omitted for clarity.

There are three expectations that must be verified to confirm that the response during the experiment shown by the rough trendlines was significant. First, the number of derived groups should increase relative to the number of ancestral groups within the MLS treatment. Second, the number of derived groups should not increase relative to the number of ancestral groups within the GLS treatment. Third, the fitness for the derived groups in the MLS treatment should be significantly higher than the trend for the derived groups in the GLS treatment. Fitness was calculated as the difference between the Malthusian parameters (SRC) for the number of group-offspring between the first and final time points. The SRC was used, because some replicates contained group extinction events, *i.e.* zero groups of either the derived or ancestral type. A *T*-test on the MLS 2 fitness data within the MLS treatment revealed that the number of derived groups was higher[1] than the number of ancestral groups ($t_3 = 2.55$, $p = 0.083$). A *T*-test on the MLS 2 fitness data within the GLS treatment revealed that there was

[1] A *p*-value of < 0.1 is interpreted here as suggestive of statistically significance only.

no significant difference between the number of derived and ancestral groups (t_3 = 1.78, p = 0.17). Finally, the difference in MLS 2 fitness between the MLS treatment and the GLS control treatment showed that the response from the MLS treatment was significantly higher than the GLS control treatment (t_6 = 2.44, p = 0.050) and that the variances could be assumed to be equal (Brown-Forsythe test: $F_{1,6}$ = 4.878, p = 0.069).

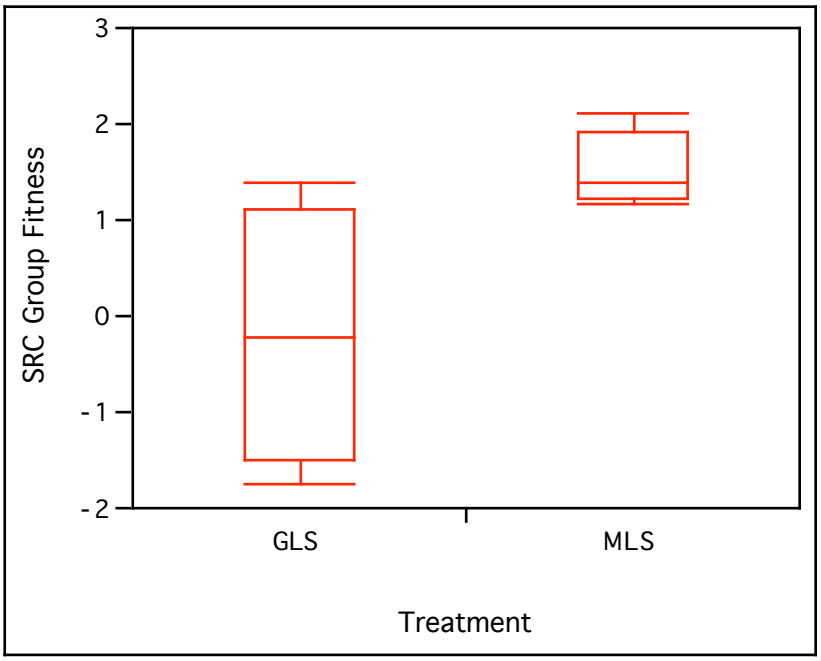

Figure 5-7: Comparison of group fitness between the MLS and the GLS treatments.
The GLS treatment has a relative group-fitness that is not significantly different from zero, while the MLS treatment is both different from zero and from the GLS treatment.

5.3 DISCUSSION

5.3.1 SELECTION FOR A DEVELOPMENT-LIKE PROCESS LEADS TO A HIGHER MLS 2 GROUP FITNESS

The primary aim of this chapter was to see if it were possible to observe a response to a selective regime that selected simultaneously at the level of the individual and the group, and to measure that response in an MLS 2 framework by counting the number of group-offspring. The results of this experiment provide encouraging signs that a response can be observed and measured in an MLS 2 framework. After the evolution phase totalling six group-generations (78 days) and a competition phase totalling a further six group-generations, the trends for the MLS and GLS treatments appeared to be diverging. Of the three expectations outlined for this experiment, two of those three were confirmed with statistical significance, and the third was marginally outside the predefined $\alpha = 0.05$ significance level. First, within the GLS treatment the derived type did not increase in fitness during the course of the experiment compared to the ancestor. Second, when the relative group-fitness (SRC) was compared between the treatments, this trend was confirmed as statistically significant, *i.e.* the MLS treatment reproduced more group-offspring than the GLS treatment. However, the relative increase of the derived type compared to the ancestral type within the MLS treatment did not reach the 95% level of confidence. The lack of statistical significance has identified two areas of improvement for this experimental design – the number of replicate MPDs, and the number of growth areas at the air-liquid interface.

With regards to the first issue, the number of replicate experiments that could be performed simultaneously was restricted to eight in these experiments, due to the availability of lid adaptors. While this may have been sufficient for an MLS treatment alone, by using the paired-parallel design with a GLS control treatment; the power of the statistical test was greatly reduced. This can be rectified for future experiments by manufacturing more lid adaptors, and thereby increasing the number of possible MPD replicates. With regards to the second issue, the number of independent growth areas at the air-liquid interface determines the resolution of the measurements for group fitness in an MLS 2 sense. In other words, because there are seven holes in the lid

adaptor, the only raw values for fitness are discrete values from zero to seven, thus providing only a very coarse scale for measuring group fitness in an MLS 2 sense. This can be improved relatively easily for future experiments by manufacturing lid adaptors with more holes (~20), thus allowing the formation of more independent groups at the air-liquid interface within the same size Petri dish environment.

Determining the fitness of groups of individuals in an MLS 2 framework had never been previously attempted. One challenge concerns the generation time of groups compared to the generation time for individual cells. Two of the primary features of microbes that make them amenable to experimental evolution studies, are their short generation times and their large population sizes. These benefits are largely negated when measuring group fitness in an MLS 2 sense, because instead of reaching population sizes in the billions and reproducing approximately every 2 h, in this experimental design, the total number of groups is only seven and the group-generation time is 13 days. One solution is to increase the strength of selection, which was done by introducing the single-cell bottleneck stage when the cells were plated on KB agar plates. Although the strength of selection does not directly compensate for the orders of magnitude reduction for the generation time and population size, it strongly penalises types that do not respond to the selective regime (*i.e.* they become extinct with respect to the experiment), thus potentially reducing the time required to observe a response to selection. Although this comes with a cost in terms of the biological realism of the experiment, such steps are necessary components of early experimental investigations into completely novel theories. However, for future work, another suggestion that requires less intervention from the experimenter is to remove the single-cell bottleneck, but increase the length of the group-generation and allow the experiment to proceed for more group-generations.

6 FINAL DISCUSSION

> *Adaptationist reasoning is not optional; it is the heart and soul of evolutionary biology. Although it may be supplemented, and its flaws repaired, to thin of* displacing *it from central position in biology is to imagine not just the downfall of Darwinism but the collapse of modern biochemistry, and all the life sciences and medicine.*

-- Dennett, D.C. 1995

6.1 OVERVIEW OF MAIN RESULTS

Pseudomonas fluorescens provides a fantastic opportunity to study a real-time adaptive radiation in a controlled experimental environment. The WS cells that arise *de novo* from free-living single cells cooperate to form an emergent group-level phenotype – the mat at the air-liquid interface. This system has been exploited for previous experimental evolution studies, for all the work in this thesis and continues to form the basis of hypothesis-driven experiments within the Rainey Laboratory and in other labs throughout the world. Among the previous studies it has been used to study adaptive radiation (Rainey and Travisano, 1998; Spiers *et al.*, 2002), the causes and maintenance of diversity (Rainey *et al.*, 2000; Spiers *et al.*, 2000; Buckling and Rainey, 2002; Hodgson *et al.*, 2002), the structure-function relationship of the signal-dependent cellulose biosynthesis pathway (Spiers *et al.*, 2002; Gal *et al.*, 2003; Spiers *et al.*, 2003; Goymer *et al.*, 2006; Bantinaki *et al.*, 2007; Malone *et al.*, 2007), the evolution of cooperation (Rainey and Rainey, 2003), the pleiotropic consequences of adaptation (MacLean *et al.*, 2004; Knight *et al.*, 2006), the modularity of bacterial genomes (Rainey and Cooper, 2004), the formation and evolution of biofilms (Spiers *et al.*, 2003; Spiers and Rainey, 2005; Hansen *et al.*, 2007), the effect of immigration history on diversification (Fukami *et al.*, 2007), mutational activation of niche-specific genes (Giddens *et al.*, 2007) and the evolution of multicellularity (Rainey, 2007). Other investigations nearing completion some of which have been referenced in this thesis include studying 'reverse evolution' (Beaumont *et al.*, 2008), the evolution of a bistable, switching phenotype (Beaumont *et al.*, 2008), the molecular mechanism of the bistable, switching phenotype (Gallie and Rainey, personal communication) and genetic constraints on parallel evolution (McDonald *et al.*, 2008).

This thesis was undertaken in a field at the interface of theoretical and empirical research that pushes the horizon of our understanding of both molecular mechanisms and evolutionary principles. In Chapter 2, *P. fluorescens* was used as an experimental tool to elucidate further details about the molecular mechanisms of the signal-dependent cellulose biosynthesis pathway, to delve deeper into the genotypic and phenotypic variation among WS types in an attempt to determine a genotype-phenotype map at a level below the ecological phenotype, and to understand the

fitness consequences of the suite of WS phenotypes. Mutations were identified in 25 of the 26 wrinkly spreaders including a new locus *mws* and three new genes of known loci *wspE*, *awsR* and *awsO*. This new genetic information suggested novel mechanisms for intramolecular phosphotransfer within WspE of the Wsp pathway, provided additional evidence for the working model of Aws pathway and revealed a novel mutation route to the wrinkly spreader phenotype, *mws*. The significance of identifying the *mws* locus has proven to be highly valuable; because it appears to be the first example of a DGC/PDE protein that possesses both DGC and PDE activity. With respect to the phenotypic analyses, multivariate analysis of the phenotypic traits revealed that *wspF* mutants were phenotypically distinct from other genotypes, which suggests that although all wrinkly spreaders share a similar ecology, there are subtle differences between different WS types that can be detected phenotypically. Surprisingly, no correlations with fitness were ever observed, suggesting a large role for non-specific pleiotropic effects.

In Chapter 3, the WS system was tested within the robust MLS 1 theoretical framework of the Haystack model to determine the dynamics of WS and SM with and without group structure. The most significant result was that time spent within a haystack affected the fitness of cooperators, because the 72 h group-generation treatment conformed to the predictions of the model, while the 24 h group-generation treatment did not. In the 24 h group-generation experiment, WS types evolved in the control treatment that were capable of increasing in frequency, however, they were never observed in the group-structured treatment, suggesting that synergistic coevolutionary relationships may have been disrupted by the high rate of migration in this experimental design.

In Chapter 4, the emergent group-level property of WS mats was analysed to determine the nature of the benefit of mat formation. It was determined that there was no density-dependent fitness when fitness was calculated in an MLS 1 framework, but rather that the fitness was dependent on time. This suggested that a development-like process accounted for the benefit of WS mat formation. Crucially, this guided the development of a new theory for the evolution of multicellularity in an MLS 2 framework that incorporates differences in the fecundity of groups, and not simply their viability. This novel theory was empirically tested in Chapter 5 by designing a

novel apparatus and experimental regime. The results showed that adaptation to simultaneous selection at the level of the individual and the group is possible, and that this can be measured in an MLS 2 sense.

In summary, there are three molecular mechanisms responsible for the WS phenotype that only differ subtly in their phenotypes at a level below the ecological niche. WS cooperation leads to the formation of an emergent group-level phenotype at the air-liquid interface and the emergent fitness of WS mats is the product of a development-like process. Finally, a response to selection was observed when selection was simultaneously imposed at the level of the individual and the group in an MLS 2 framework.

6.2 FUTURE DIRECTIONS

The work in this thesis and other work in the Rainey lab suggest that only three pathways are involved in the manifestation of the WS phenotype. Of these, the Wsp pathway is the best described based on experimental studies and homology with the Che chemosensory pathway in *E. coli* (Spiers *et al.*, 2002; Gehrig, 2005; Goymer *et al.*, 2006; Bantinaki *et al.*, 2007; Malone *et al.*, 2007; McDonald *et al.*, 2008); however, the finding of mutations in *wspE* revealed a new interesting mechanism for generating the WS phenotype. As discussed in Chapter 2, important future research on the molecular details of phosphate transfer from $WspE_{CheA}$ to $WspE_{CheY}$ and the competition between the CheY domains of WspF, WspR and $WspE_{CheY}$ is required to develop a complete understanding of the molecular mechanisms of *wspE* wrinkly spreaders. This research would also contribute to a more detailed understanding of the molecular mechanisms in the Wsp pathway as a whole, and determine whether the mechanism of intramolecular phosphotransfer observed in FrzE of *M. xanthus* (Li *et al.*, 2005) is similar to that of *P. fluorescens*.

Another avenue for further molecular research is the Mws pathway. From existing genetic studies, it appears that MwsR is a unique DGC/PDE protein (McDonald *et al.*, 2008), because no other DCG/PDE protein has been determined to have both DGC and PDE function (Jenal and Malone, 2006). This is intriguing, because the coupling of these activities in a single molecule suggests that changing the level of MwsR cannot change the intracellular c-di-GMP levels. Therefore, focussed research in this area may reveal a novel regulation strategy, possibly involving temporal or spatial changes in cellular conditions (*e.g.* pH) to increase the DCG activity relative to the PDE activity and *vice versa* to alter the c-di-GMP levels.

Chapters 3 and 4 provided the impetus and direction for the novel theory and experimental design developed in Chapter 5. The results of this chapter suggest that experiments in an MLS 2 framework can yield valuable insight, and that the experimental system developed here can be used as the basis for much future work. The most obvious next step would be to repeat the experiment with an increased number of group-generation cycles in both the evolution phase and the competition phase of the experiment. As mentioned in the discussion of Chapter 5, with more lid

adaptors the experiment could also be repeated with more replicate MPDs and the robustness of future experiments could be improved by increasing the number of holes in future versions of the lid adaptor. Additionally, the results of experiments using different numbers of holes in the lid adaptor could be compared to one another to determine whether there is an effect of the number of available group-niches.

Another potential future project would be to build on the existing theory by designing experiments in homogeneous chemostat environments. This would minimise the design-specific contrivances of a batch culture experimental regime, allowing a more biologically realistic environment. This design was considered during the development of the experimental design finally adopted in Chapter 5, but no experimentally tractable design solution was found. In this environment it might also be possible to dynamically alter the number of available group-niches by making more niches available during the experiment. Groups with swimmer propagules that take advantage of the newly available niches would be able to reproduce more group offspring than those that do not. It has been suggested by mathematical modelling that different developmental strategies will be favoured depending on the availability of new niches (Rainey and Kerr, manuscript in preparation). If group-niches are always available, mat generation will be short and swimmer production will be maximised (high fecundity), while if group-niches are rarely available, longer mat generation time will be favoured to maximise the absolute number of swimmer cells produced (high viability). Future work exploring these hypotheses empirically would be possible if such a design could be developed.

6.3 FINAL COMMENT

Scientific understanding advances as a slowly evolving iterative process of theorising, testing and revising, occasionally punctuated by a few monumental contributions that transform entire subject areas. Darwin's theory of evolution by natural selection transformed the natural sciences almost 150 years ago, but fine-tuning the details through experimentation and revision remains an important challenge. This thesis has described some novel theoretical and empirical approaches to studying the evolution of cooperative behaviour in the bacterium *P. fluorescens*. These ideas were born out of the contrasting skills of paying intimate attention to the tiny molecular details of microorganisms, and attempting to unify the concepts of an evolutionary theory charged with explaining the existence of everything in our ecosystem that can be observed and detected by the scientific method. The critical review of these novel contributions will ultimately determine their value to the scientific community in the future.

7 MATERIALS AND METHODS

Come on, man. Learn by doing. Learn by doing!

-- Turk, C. 1999

7.1 MATERIALS

7.1.1 MEDIA AND GROWTH CONDITIONS

Unless otherwise specified, all chemicals in this section were obtained from BDH, and all equipment from BioLab. *P. fluorescens* cultures were grown at 28˚C in 30 ml glass vials containing 6 ml of King's Medium B (KB: 10 g/l glycerol; 20 g/l Proteose Peptone No.3; 1.5 g/l Mg_2SO_4; 1.5 g/l K_2HPO_4 (King *et al.*, 1954)) or Lysogeny Broth (LB: 10 g/l NaCl; 10 g/l tryptone; 5 g/l yeast extract (Bertani, 1951)) and supplemented with antibiotics and substrates when necessary (see Section 7.1.6). The standard microcosm is a 30 ml glass vial containing 6 ml of KB and may be incubated statically in a 28˚C room or in a shaking incubator (New Brunswick) at 150 rpm at 28˚C.

E. coli cultures were grown at 37˚C in 5 ml of LB (supplemented with antibiotics and substrates where appropriate) in disposable 30 ml vials. For agar plates, 1.5% w/v agar was added to the media (also supplemented with antibiotics or substrates where appropriate).

7.1.2 BACTERIAL STRAINS

Table 7-1: *Pseudomonas fluorescens* strains used in this study.

Strain	Description	Reference
SBW25	Ancestral, wild-type, environmental isolate, smooth	(Rainey and Bailey, 1996)
SBW25-*lacZ*	SBW25 with a *lacZ* transcriptional fusion	(Zhang and Rainey, 2007)
IWS_A	Independent WS from day 7 of static incubation	(Bantinaki *et al.*, 2007)
IWS_B	Independent WS from day 7 of static incubation	(Bantinaki *et al.*, 2007)
IWS_C	Independent WS from day 7 of static incubation	(Bantinaki *et al.*, 2007)
IWS_D	Independent WS from day 7 of static incubation	(Bantinaki *et al.*, 2007)
IWS_E	Independent WS from day 7 of static incubation	(Bantinaki *et al.*, 2007)
IWS_F	Independent WS from day 7 of static incubation	(Bantinaki *et al.*, 2007)
IWS_G	Independent WS from day 7 of static incubation	(Bantinaki *et al.*, 2007)
IWS_H	Independent WS from day 7 of static incubation	(Bantinaki *et al.*, 2007)
IWS_I	Independent WS from day 7 of static incubation	(Bantinaki *et al.*, 2007)
IWS_J	Independent WS from day 7 of static incubation	(Bantinaki *et al.*, 2007)
IWS_K	Independent WS from day 7 of static incubation	(Bantinaki *et al.*, 2007)
IWS_L	Independent WS from day 7 of static incubation	(Bantinaki *et al.*, 2007)
IWS_M	Independent WS from day 7 of static incubation	(Bantinaki *et al.*, 2007)
IWS_N	Independent WS from day 7 of static incubation	(Bantinaki *et al.*, 2007)
IWS_O	Independent WS from day 7 of static incubation	(Bantinaki *et al.*, 2007)
IWS_P	Independent WS from day 7 of static incubation	(Bantinaki *et al.*, 2007)
IWS_Q	Independent WS from day 7 of static incubation	(Bantinaki *et al.*, 2007)
IWS_R	Independent WS from day 7 of static incubation	(Bantinaki *et al.*, 2007)

IWS_S	Independent WS from day 7 of static incubation	(Bantinaki *et al.*, 2007)
IWS_T	Independent WS from day 7 of static incubation	(Bantinaki *et al.*, 2007)
IWS_U	Independent WS from day 7 of static incubation	(Bantinaki *et al.*, 2007)
IWS_V	Independent WS from day 7 of static incubation	(Bantinaki *et al.*, 2007)
IWS_W	Independent WS from day 7 of static incubation	(Bantinaki *et al.*, 2007)
IWS_X	Independent WS from day 7 of static incubation	(Bantinaki *et al.*, 2007)
IWS_Y	Independent WS from day 7 of static incubation	(Bantinaki *et al.*, 2007)
IWS_Z	Independent WS from day 7 of static incubation	(Bantinaki *et al.*, 2007)
LSWS	Large Spreading Wrinkly Spreader derived from SBW25	(Spiers *et al.*, 2003)
WS-Δ*wspF*	WS type with *wspF* deleted (Δ*wspF* 2A)	(Bantinaki *et al.*, 2007)
SBW25-Δ*panB*	Pantothenate auxotroph of SBW25	(Rainey, 1999)
SBW25-Δ*wss*	Cellulose deficient, carries a Kanamycin resistance marker	(Gehrig, 2005)
SBW25-Gm	SBW25 with a Gm resistance marker	Zhang, X. X. unpublished
SBW25-Gm+GUS	SBW25 with a Gm resistance marker and a GUS transcriptional fusion	Zhang, X. X. unpublished
3GWS	Dominant WS after 36 days in the experimental treatment with a 72-h group-generation	This study
3WS01	Dominant WS after 36 days in control line 1 with a 72-h group-generation	This study
3WS02	Dominant WS after 36 days in control line 2 with a 72-h group-generation	This study
3WS03	Dominant WS after 36 days in control line 3 with a 72-h group-generation	This study
3WS04	Dominant WS after 36 days in control line 4 with a 72-h group-generation	This study
3WS05	Dominant WS after 36 days in control line 5 with a 72-h group-generation	This study
3WS06	Dominant WS after 36 days in control line 6 with a 72-h group-generation	This study
3WS07	Dominant WS after 36 days in control line 7 with a 72-h group-generation	This study
3WS08	Dominant WS after 36 days in control line 8 with a 72-h group-generation	This study
3WS09	Dominant WS after 36 days in control line 9 with a 72-h group-generation	This study
3WS10	Dominant WS after 36 days in control line 10 with a 72-h group-generation	This study
3WS11	Dominant WS after 36 days in control line 11 with a 72-h group-generation	This study
3WS12	Dominant WS after 36 days in control line 12 with a 72-h group-generation	This study
SSM18	SM type at 18 days in the experimental treatment with a 24-h group-generation	This study
SSM27	SM type at 27 days in the experimental treatment with a 24-h group-generation	This study
SSM36	SM type at 36 days in the experimental treatment with a 24-h group-generation	This study

Table 7-2: *Escherichia coli* strains used in this study.

Strain	Description	Reference
DH5α–λ*pir*	*supE*44, Δ*lacU*169, *hsdR*17, *recA*1, endA1, *gyrA*96, *thi*-1, *relA*1, λ*pir*	Gibco-BRL
pRK2013	Contains the helper plasmid pRK2013	(Figurski and Helinski, 1979)

7.1.3 PLASMIDS AND TRANSPOSONS

Table 7-3: Names and characteristics of plasmids and transposons.

Strain	Description	Reference
Plasmids		
pCR8TOPO	Spe^R, pUC *ori*, 2.8kb sequencing plasmid	Invitrogen
pCR8TOPO-Erep	Spe^R, pCR8TOPO containing the WspE mutation construct	This study
pUIC3	Tc^R, *mob*, *oriR6K*, *bla*, Δpromoter-*lacZY*	(Rainey, 1999)
pUIC3-Erep	Tc^R, pUIC3 containing the WspE mutation construct	This study
pRK2013	Km^R, incP4, tra, mob, mobilization plasmid used as a helper in tri-parental mating	(Figurski and Helinski, 1979)
Transposons		
IS-Ωkm/hah	Km^R, ColE1, *ori*, *npt* promoter, *loxP*	(Giddens *et al.*, 2007)

7.1.4 PRIMERS

Table 7-4: Primer names, sequences and targets.

Name	Sequence	Target
18mer	CGA TGA CGC CCT GCT GGA	5' *wspF*
zrwspF2	TAT TTT CTT CAT GGG CCA GG	3' *wspF*
zrwspF3	AGA TCA TGG CCG AGA CGC	Internal *wspF* primer
zrwspF4	AAA TCC TGC TCA AGG GTT	Internal *wspF* primer
zrwspF5	TGT GCT GCT GAC CGG CAT GG	Internal *wspF* primer
Tn*phoA*-II	GTG CAG TAA TAT CGC CCT	IS-Ωkm/*hah*
CEKG2A	GGC ACG CGT CGA CTA GTA CAN NNN NNN NNN AGA G	Non-specific
CEKG2B	GGC CAC GCG TCG ACT AGT CAN NNN NNN NNN ACG CC	Non-specific
CEKG2C	GGC CAC GCG TCG ACT AGT CAN NNN NNN NNN GAT AT	Non-specific
hah-1	ATC CCC CTG GAT GGA AAA CGG	5' end of CEGK2A,B & C
CEKG4	GGC CAC GCG TCG ACT AGT AC	IS-Ωkm/*hah*
TWSF1	ACG TCG CGC GCC AGT CCA	5' *mws*
TWSF2	CTG GTC ACG CCC CTG AGC	Internal *mws* primer
TWSF3	CGA AAG CTT CTG GTC GGA CG	Internal *mws* primer
TWSF4	TGG TGC TGG TGT GGG ACG	Internal *mws* primer
TWSF5	CTG GCC AGC TAT GTG TGC	Internal *mws* primer
TWSR1	CGG CTG AAC ATG GCG GTC	3' *mws*
TWSR2	GCC AGT CCT ACT ACC TGA GC	Reverse internal *mws* primer
Aws11f	GCT GGT TCA GCT TGA TCG AAC CC	3' *aws* primer
Aws12r2	AAC GGA CCG TTA AGC CAA CAC C	5' *aws* primer
Aws15r	CCT TGG CGC CGC TCA TTG CG	Reverse internal *aws* primer
Aws16r	ATG GAT GCT GCC GAT GGT TC	Reverse internal *aws* primer
Aws17r	GTT GGG CAG ACC GGT GAG GC	Reverse internal *aws* primer
Aws18r	CAT TGA GTA AGG TGC CAG GGG TGG	Reverse internal *aws* primer
Aws10	CGA CCG CGC GTT TCA TG	Internal *aws* primer
Aws9	TTC GAA GGC CGT TTG ATC C	Internal *aws* primer
wspA1f	GCT TGA CGT TGC GGA GGT G	5' *wspA*
wspA1r	CAA GTC CTT GGA GGT GGC C	Reverse Internal *wspA*
wspA2f	CAT GAT GAC CGA GCT CAC GG	Internal *wspA*
wspA2r	CGA AGG TCA AGG CGC TGA G	3' *wspA*
wspB1f	GGT CAA CGA AGG CAT GCA GG	5' *wspB*

wspB1r	GTG CTG CCA GTA TTC GCC C	3' *wspB*
wspC1f	GAC CTG CTG ACC GAT GAC G	5' *wspC*
wspC1r	CCG ATG CGG TTC CAG CAG	3' *wspC*
wspD1f	CGC CCA GGT GTT TTA CTG GC	5' *wspD*
wspD1r	GTT GCG TTC CAG GGC GAG	3' *wspD*
wspE1f	CCG CGC ATG TTG ATC ATT GC	5' *wspE*
wspE1r	CCG ATC GAC GAT AGT GCC G	Reverse Internal *wspE*
wspE2f	GAC CAC GGC ATC GAA ATG C	Internal *wspE*
wspE2r	CAT CAG GAT CAG GTC CGG C	3' *wspE*
wspR1f	CGCCCACTGGATAGAATTGC	5' *wspR*
wspR1r	GTA CAC CGA GCC TAG GCG	3' *wspR*
M13F	GTA AAA CGA CGG CCA G	5' PCR8TOPO MCS
M13R	CAG GAA ACA GCT ATG AC	3' PCR8TOPO MCS
Bla	CAG GGT TAT TGT CTC ATG AGC G	5' pUIC3 MCS
P-*lacZ*	TGG GAT TAA CTG CGC GTC GCC	3' pUIC3 MCS
ErepF	AGA TCT GCT GAC GTT CCT GTT CCT GC	~1kb upstream of *wspE* mutation
ErepR	AGA TCT GCA AAG ACC TCC AGC AAC CG	~1kb downstream of *wspE* mutation

Incorporated *Bgl*II sites are underlined.

7.1.5 Strain Storage

All bacterial strains were grown up overnight in either LB or KB to be stored at -80˚C in 70% glycerol saline solution (8.5 g NaCl, 300 ml H_20, glycerol to 1 litre).

7.1.6 Antibiotics and Markers

Unless otherwise specified, all chemicals were obtained from Sigma. Antibiotics were purchased from Melford Laboratories and used in the following concentrations: kanamycin 50µg/ml, gentamicin 40µg/ml, spectinomycin 100µg/ml, tetracycline 10µg/ml, N-[5-Nitro-2-furfurylidene]-1-aminutesohydantoin (NF) 100µg/ml (dissolved in dimethyl formamide – DMF) and 5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside (X-gal; dissolved in DMF) 40µg/ml.

7.2 METHODS

7.2.1 DILUTIONS

All dilutions of bacterial cultures were performed in ¼ Ringer's solution 22.2 µl in 200 µl in microtitre plates unless otherwise specified.

7.2.2 DNA PREPARATION

Plasmid DNA was prepared using the QIAGEN Mini prep kit using 1 ml of an overnight culture of *E. coli* or *P. fluorescens*.

Genomic DNA was prepared by the following method using 1ml of an overnight culture of *E. coli* or *P. fluorescens*. The 1 ml of overnight culture was centrifuged at 13,200 rpm for 2 minutes and resuspended in 567 µl TE buffer, 30 µl of 10% SDS and 3 µl of 20 mg/ml proteinase K. These ingredients were mixed and incubated for 1 h at 37˚C. 100 µl of 5M NaCl was added with thorough mixing. 80 µl CTAB/NaCl solution (4.1 g of NaCl was dissolved in 80 ml of water and 10 g of hexadecyltrimethyl ammonium bromide was added while heating and stirring and adjusted to a final volume of 100 ml) was added, mixed and incubated for 10 minutes at 65˚C. 600 µl of phenol was added mixed and centrifuged for 1 minute. The top layer was pipetted off and placed in a clean, labelled Eppendorf tube. The addition of phenol, centrifugation and pipetting to a clean tube was repeated. This step was repeated again, but 600 µl of chloroform was used instead of phenol. The chloroform step was repeated. 1 ml of 100% ethanol was added and incubated in the freezer at -20˚C for 2 h. The sample was removed from the freezer and centrifuged at 13,200 rpm for 5 minutes. The ethanol was removed and followed by two wash steps with 1 ml of 70% ethanol. The pellet was dried, resuspended in 50 µl of water and DNA concentration was measure using a nanodrop (Nanodrop Technologies, Inc).

7.2.3 POLYMERASE CHAIN REACTION (PCR)

Polymerase chain reactions were performed using a DNA Engine DYAD Peltier Thermal Cycler (Bio-Rad). Reactions were performed in 200 µl tubes with a total volume of 25 µl containing: 2.5 µl 10x PCR Buffer, 1 µl dNTP mix, 1 µl 50 mM $MgCl_2$ and 0.25 µl Taq polymerase, 1 µl of each primer (10 pmol) and the remainder

was deionised water. Template DNA was either pre-prepared genomic DNA (Section 7.2.6) or cells from an overnight culture. For each template-primer combination, the thermal cycler program was optimised. Each cycling program started with an initial denaturation step (96˚C) for 3 minutes (or 10 if the DNA template was cells) followed by 25-30 cycles consisting of a denaturation step (94˚C for 30 s), an annealing step (50-60˚C for 30-45 s specific to each primer pair) and an extension step at 72˚C for an appropriate time depending on the desired product length. The reaction completed with an additional extension step at 72˚C for 10 minutes to finish partially synthesised products. Typically, the PCR was followed by analysis using gel electrophoresis.

7.2.4 ARBITRARY PRIMED PCR (AP-PCR)

This AP-PCR was developed by Manoil (2000) and modified by Jacobs (2003) involing two successive rounds of PCR using a DNA Engine DYAD Peltier Thermal Cycler (Bio-Rad). The first PCR contained: 2.5 µl 10x PCR Buffer, 1 µl dNTP mix, 0.8 µl 50 mM $MgCl_2$ and 0.25 µl Taq polymerase, 2 µl of each primer (10 pmol) and the remainder was deionised water to a reaction volume of 25 µl. Template DNA was either pre-prepared genomic DNA (Section 7.2.6) or cells from an overnight culture. The cycling program started with an initial denaturation step (94˚C) for 10 minutes followed by 6 cycles of 30 s at 94˚C, 30 s at 42˚C, and 3 minutes at 72˚C, where the temperature was decreased by 1˚C for every cycle. Amplification was continued for a further 25 cycles of 30 s at 94˚C, 30 s at 65˚C, and 3 minutes at 72˚C. The products of this PCR were diluted by addition of 80 µl of deionised water, and used as templates in the second round of PCR. In addition, to the template DNA, the second PCR contained 2.5 µl 10x PCR Buffer, 1 µl dNTP mix, 0.8 µl 50 mM $MgCl_2$ and 0.25 µl Taq polymerase, 2 µl of each primer (10 pmol) and the remainder was deionised water to a reaction volume of 20 µl. The cycling program started with a denaturation step of 10 minutes at 94˚C, was amplified for 30 cycles of 30 s at 94˚C, 30 s at 65˚C and 3 minutes at 72˚C. The PCR products were cooled to 4˚C and prepared for sequencing.

7.2.5 ELECTROPHORESIS

Electrophoretic gels were prepared with 1% w/v agarose and 0.5 x TBE (90 mM Tris-HCl pH 8.0, 0.55% boric acid, 2 mM EDTA). Ethidium bromide (10 µg/ml) was incorporated into the gel for visualisation of DNA under ultraviolet light. Samples were loaded with a loading buffer containing 0.25% bromophenol blue and 30% glycerol. Gels were submerged in 0.5x TBE in a XX and electrophoressed at a constant voltage of ~110 mV for an appropriate time. A 1 kb DNA ladder (Invitrogen) was used as a size marker suitable for sizing DNA from 500 bp – 10 kb.

7.2.6 DNA EXTRACTION

For single bands of DNA on agarose gels, DNA was extracted directly from the PCR using a QIAquick PCR purification kit (QIAGEN). If multiple fragments were present, the fragment of correct size was isolated by cutting it out of the agarose gel with a sterile razor blade and extracted using the QIAquick gel extraction kit (QIAGEN). DNA samples were eluted in sterilised deionised water and DNA concentration measured using a nanodrop (Nanodrop Technologies, Inc).

7.2.7 DNA SEQUENCING

Sequencing was carried out by the di-deoxy method using the BigDye terminator Cycle Sequencing Ready Reaction (Applied Biosystems, Perkin-Elmer). 10 µl of sample DNA was mixed with 4 µl BigDye reaction mix, 2 µl of 5 x BigDye buffer and 1 µl primer (10 pmol) in a 20 µl volume with deionised water. Sequencing reactions were carried out in a DNA Engine DYAD Peltier Thermal Cycler (Bio-Rad). The thermal cycler program had an initial template denaturation step of 1 minute at 96˚C followed by 25 cycles of amplification of 10 s at 96˚C, 5 s at 50-55˚C and 4 minutes at 60˚C before cooling to 15˚C. Sequencing products were gathered using two different methods. When there were fewer than 16 PCR products to be sequenced, ethanol precipitation was used by adding 2 µl of NaOAc (pH 4.8) and 2 µl of 125 mM EDTA followed by 50 µl of room temperature 100% ethanol. The DNA was recovered by centrifugation at 13,000 rpm for 30 minutes and the supernatant removed. The DNA was washed twice with 250 µl 70% ethanol. When there were greater than 16 PCR products to be sequenced, the CLEANSEQ cleanup method was used. 10 µl of CLEANSEQ magnetic beads and 60 µl of 85% ethanol were added to

the PCR product in a 96 well PCR plate. The mixture was left to incubate at room temperature for 3 minutes in a SPRIPlate 96R. The clear ethanol supernatant was removed with a vacuum hose and a further 40 µl of 85% ethanol was added as a wash step. Again, the 85% ethanol was removed and the plates left to dry upside down in the dark for 5 minutes. In both cases, samples were submitted to the SBS sequencing facility for separation. The sequence traces were analysed using Sequencher (Section 7.2.21).

7.2.8 ALLELIC REPLACEMENT

To test if the mutation detected in IWS_H was both necessary and sufficient for causing the WS phenotype of IWS_H, the mutation in *wspE* was introduced into SBW25. Specially designed primers, ErepF and ErepR (with incorporated *Bgl*II sites) were used in a PCR to amplify the fragment of IWS_H to introduce into SBW25. The PCR product was ligated into pCR8TOPO using 4 µl purified PCR product, 1 µl salt solution and 1 µl pCR8 vector in a total reaction volume of 6 µl. The reaction was mixed gently and left at room temperature for ~30 minutes. Chemically competent *E. coli* were transformed with this ligation mix (Section 7.2.9). Transformed colonies were checked for the presence of an unmutated copy of the insert using colony PCR followed by sequencing. Successfully transformed colonies were grown overnight for storage at -80˚C and used for a plasmid mini-prep. Overnight cultures were centrifuged at 13,200 rpm for 1 minute. The pellet was resuspended in 100 ml Buffer P1 by pipetting up and down, then incubated for 5 minutes at room temperature. 200 µl of Buffer P2 was added, mixed by inverting the tube several times and incubated on ice for 5 minutes. 150 µl Buffer N3 was added, mixed by inverting and incubated on ice for 5 minutes. The mixture was centrifuged for 10 minutes at 13,200 rpm and the supernatant transferred to a new tube. 900 µl (2 volumes) of cold 100% ethanol was added, and left at -20˚C for 30 minutes to precipitate. The DNA pellet was retained by centrifuging the mixture at 13,200 rpm at 4˚C for 30 minutes and pouring off the supernatant. The pellet was washed with cold 70% ethanol and dried on the bench for 10 minutes.

pCR8TOPO+insert and pUIC3 were enzymatically cleaved (Section 7.2.10) and run on a gel. The cleaved products were ligated with 1 µl DNA ligase, 10 x ligation buffer and water to a total reaction volume of 10 µl added and left at 4˚C overnight.

Chemically competent *E. coli* cells were transformed with 10 µl of the ligation mixture and plated on tetracycline plates. 20 colonies were picked and resuspended in 50 µl of water and 2 µl was used in a PCR to check for the presence of the insert with primers bla and p*lacZ*. 20 µl was used to inoculate an overnight culture for use in a conjugation and to make a glycerol stock. pUIC3+insert was conjugated with SBW25 using the tri-parental method (Section 7.2.12).

7.2.9 TRANSFORMATION

Transformations were performed according to Sambrook (1989) with the following modifications.

10 µl of DNA was added to 50 µl of chemically competent cells and incubated on ice for 30 minutes. The cells were heat-shocked at 42˚C for 30 s and SOC medium was added. The suspension was incubated at 37˚C shaking for 1 h and plated out by an appropriate dilution series on selective LB plates.

7.2.10 RESTRICTION ENZYME CLEAVAGE

The restriction enzymes used were from Invitrogen. 20 µl of DNA was digested with 2 µl of restriction enzyme with 8 µl of 5 x Buffer 3 and made up to 40 µl reaction volume. Reactions were left at 37˚C for between 2 h and overnight. The product of the digest was run on a gel. To avoid re-ligation of cleaved plasmids with compatible ends, CIP (calf intestinal phosphatase, NEB) was added at the end of the reaction for 30 minutes at 37˚C before storing at -20˚C. CIP was deactivated by heating to 70˚C for 10 minutes.

7.2.11 BI-PARENTAL CONJUGATION

P. fluorescens (recipient) and *E. coli* containing ISΩkm/hah (donor) were grown overnight. 1 ml of each of the cultures was pelleted at 13,000 rpm and resuspended in 800 µl of pre-warmed LB. 200 µl of *P. fluorescens* was heat-shocked for 20 minutes at 45˚C. 800 µl of *E. coli* was added to the heat-shocked cells and the mixture was centrifuged at 13,200 rpm for 1 minute. The supernatant was removed and the cell mixture was resuspended in 50 µl of ¼ Ringer's Solution. The suspension was gently spread on a sterile, 25 mm-diameter nitrocellulose filter (0.22 µm pores, Millipore) on

an LB plate and incubated for 2-4 h at 28˚C. After incubation the filter was placed in 1 ml of deionised water and vortexed for 15 s. The filter was removed and the suspension appropriately diluted and plated on LB plates supplemented with kanamycin (to select for the transposon) and NF (to counterselect the *E. coli* donor).

7.2.12 TRI-PARENTAL CONJUGATION

Conjugation in this study used *E. coli* DH5α cells as the donor. This strain does not express the *mob* genes required for conjugational transfer of plasmid DNA. A helper strain containing the plasmid pRK2013 expressing the required transfer proteins was used in this tri-parental mating approach. DH5α, pRK2013 and the recipient *P. fluorescens* strain were grown up overnight. 200 µl of the recipient strain was heat-shocked at 42˚C and mixed with 200 µl of DH5α and 200 µl of pRK2013 cells. The cells were centrifuged at 13,200 rpm for 1 minute to form a pellet. The supernatant was removed and the cells were resuspended in 50 µl of LB and applied to an LB plate. The plate was incubated overnight at 28˚C. The cells were scraped off the LB plate and placed into 1 ml of deionised water in a sterile 1.5 ml Eppendorf tube. Dilutions were spread on LB plates containing appropriate antibiotics and NF to counterselect the *E. coli* cells. The plates were incubated for 2 days at 28˚C and then screened for transconjugants. Six transconjugants were re-streaked on plates with NF counterselection to purify them from the *E. coli* lawn and incubated for 2 days at 28˚C. This was followed by enrichment (Methods 7.2.13) for the double cross-over type and excision of pUIC3.

7.2.13 ENRICHMENT

Six transconjugants were grown overnight in 5 ml of LB with tetracycline. 100 µl of each overnight culture was combined in a microcentrifuge tube. 10 µl was added to 200 ml of LB containing no antibiotic in a large flask and left to shake overnight at 28˚C. 400 µl of overnight culture was added to 20 ml of fresh LB without antibiotic and incubated for 30 minutes. 20 ml of 10 µg/ml tetracycline was added to the flask and incubated for a further 2 h. After 2 h, 800 µg/ml cycloserine was added to the flask and incubation continued for a further 4 h. After 4 h, 1 ml of cells was pipetted into a microcentrifuge tube and centrifuged for 1 minute at 13,200 rpm. The cells were washed with 500 µl of sterile water and re-pelleted by centrifugation at 13,200

rpm. The cells were resuspended in 500 µl of ¼ Ringer's solution, diluted appropriately and plated on LB+X-gal plates.

7.2.14 TRANSPOSON MUTAGENESIS

Transposon mutagenesis was performed using the *E. coli* strain ISKm/hah according to a method developed in the Rainey laboratory (Giddens *et al.*, 2007) following a bi-parental conjugation (see Section 7.2.11). The conjugation mix was diluted and plated on LB plates supplemented with kanamycin (to select for the transposon) and NF (to counterselect the *E. coli* donor). After 48 h of incubation at 28˚C the colonies that were picked were those that exhibited a phenotypic change from WS to SM indicating that the phenotype had been altered by the insertion of the transposon. Colonies were grown up overnight with kanamycin and stored at -80˚C. Transformants were sequenced at the site of insertion and were mapped to their position in the *P. fluorescens* genome (see Section 7.2.20) using Artemis software (Rutherford *et al.*, 2000).

7.2.15 FITNESS ASSAYS

Early experiments in the *P. fluorescens* system used an SBW25 strain auxotrophic for pantothenate, SBW25-Δ*panB*, to determine the relative fitness of derived strains compared to the ancestor (Rainey, 1999). However, it was noted that SBW25-Δ*panB* was unsuitable in longer-term experiments (> seven days), because it had a measurable reduction in fitness (MacLean *et al.*, 2005). SBW25-*lacZ* is a modified strain of SBW25 engineered with a Tn*7*:*lacZ* (β-galactosidase) fusion integrated at the Tn*7* attachment site (attTn*7*), a defective prophage locus in the SBW25 genome. When agar plates are supplemented with the substrate X-gal (5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside), the β-galactosidase enzymatic activity provided by the *lacZ* gene cleaves the substrate leaving a blue product. The blue colonies are readily distinguishable from the white colonies and the change in ratio of blue to white colonies provides the raw data for a simple measure of relative fitness over time. In all cases, strains were pre-cultured overnight from a large inoculum of samples stored at -80˚C in glycerol to ensure that the cells were in log phase growth before beginning competition. A sample of overnight culture was diluted appropriately so that the number of cells could be estimated by optical density (OD)

using the calibration curve in Figure 7-1. The actual cell density at t_0 was estimated by plating out appropriate volumes of a dilution series.

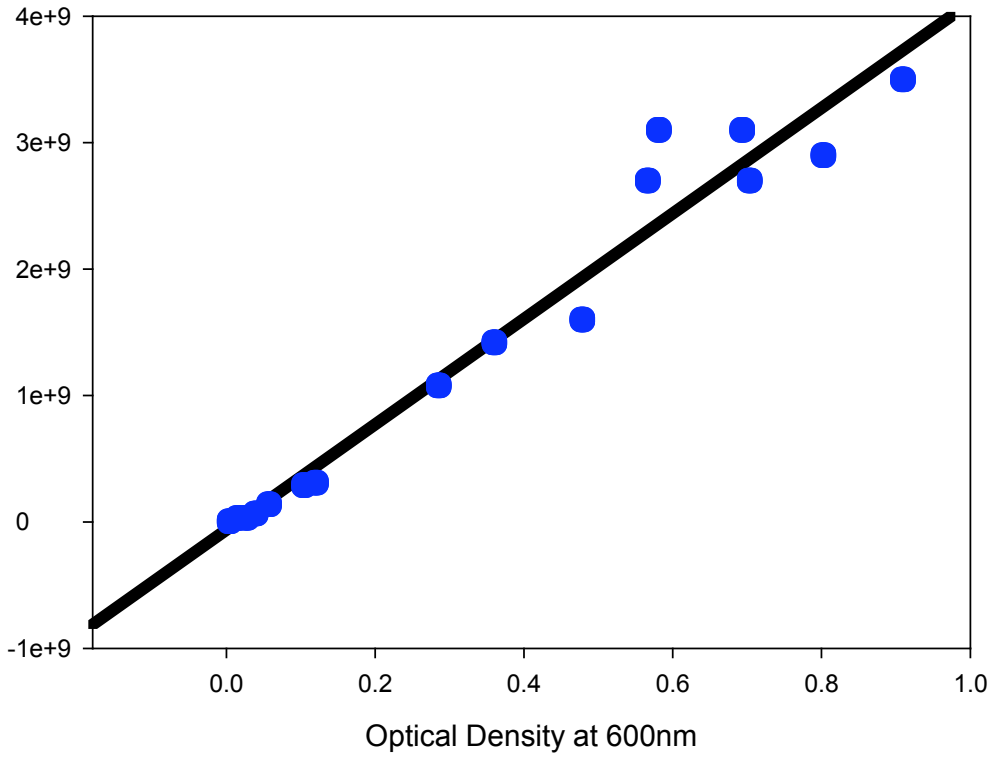

Figure 7-1: Calibration curve for cell density from optical density at 600nm.
Optical density is only reliable for *P. fluorescens* in the range or 0.1 to 1.0, so must be diluted from high cell density before measurements will be accurate.

7.2.15.1 Two Competitor Fitness Assays

Each fitness assay was a competition between a strain of interest, and the marked strain SBW25-*lacZ*. Each strain was inoculated at a ratio of approximately 1:1 with approximately 10^3 cells of each type (Rainey and Rainey, 2003). The number of colony forming units of each strain were counted at t_0 and after 24 h of incubation shaking at 150 rpm at 28˚C. Relative fitness was expressed as the ratio of the two selection rate constants (Lenski *et al.*, 1991).

7.2.15.2 3-way Static Competition Assays

Each 3-way fitness assay was a competition between a WS strain of interest, a marked competitor strain LSWS-*lacZ* and an unmarked ancestral SM type to occupy the broth phase to encourage competition between WS types at the air-liquid

interface. All strains were inoculated in a 1:1:1 ratio with approximately 10^5 cells of each type (Bantinaki *et al.*, 2007). The numbers of colony forming units of the marked and unmarked strains were counted at t_0 and after 24 h of static incubation at 28˚C. Relative fitness was expressed as the ratio of the two selection rate constants (Lenski *et al.*, 1991).

7.2.15.3 Long-Term Fitness Assays

The long-term fitness cost of the presence of the *lacZ* marker in SBW25-*lacZ* was determined by a long-term competition between the SBW25-*lacZ* and the unmarked ancestor SBW25. Each strain was inoculated at 1:1 initial ratio at a density of approximately 10^3 cells at t_0. Cultures were incubated in a shaking environment for 24 h at 150 rpm at 28˚C. At t_{24}, the cultures were diluted by 10^6 and approximately 1000 cells were inoculated into a fresh microcosm. This was repeated for ten days, and after each 24 h period, the numbers of colony forming units of the marked and unmarked strains were counted. Relative fitness was expressed as the ratio of the two selection rate constants (Lenski *et al.*, 1991).

7.2.15.4 Broth-saturated Competitions

Each broth-saturated competition was a competition between the modified strain SBW25-Δ*wss* (unable to produce cellulose or occupy the air-liquid interface) and a gentamicin marked competitor. Each strain was pre-cultured separately for 24 h in a shaken environment at 150 rpm at 28˚C to reach saturation. Each microcosm of SBW25-Δ*wss* was inoculated with varying numbers of gentamicin marked competitor across five orders of magnitude at ratios between $1:10^3$ and $1:10^8$. Microcosms were incubated either statically for 48 h, shaken for 48 h or static for 24 h, interrupted by vortexing for 30 s, and then incubated again for 24 h, all at 28˚C. SBW25-Δ*wss* colonies were counted at 24 and 48 h by plating on KB plates supplemented with kanamycin, while the low frequency of the invading genotype was detected using KB plates supplemented with gentamicin. Relative fitness was expressed as the ratio of the two selection rate constants (Lenski *et al.*, 1991).

7.2.16 MLS 1 SELECTION EXPERIMENTS

A total of four populations (two experimental treatments and two control treatments) each consisting of 12 groups (microcosms), were studied in a long-term experiment totalling 36 days of experimentation. This method is a description for both the experimental and control treatments of the 72-h transfer regime. The 24-h transfer regime changes all 72-h static incubations to 24 h. Samples were taken every 72 h of experimentation to coincide with every transfer in the 72-h transfer regime and every third transfer in the 24-h transfer regime.

Twelve pre-cultures of SBW25 were grown overnight to be used in the experimental treatment. At t_0, 12 microcosms were inoculated with ~ 10^4 cells of ancestral SBW25 and incubated statically for 72 h. After 72 h, each of the 12 microcosms were homogenised for approximately 60 s by vortexing. 100 µl from each microcosm was pipetted into a 1.5 ml Eppendorf tube and vortexed for 10 s to ensure mixing. Twelve samples of the mixture were diluted 100-fold and 6 µl of each diluted sample was used to inoculate fresh microcosms. This transfer regime was repeated until 36 days of experimentation were completed.

Twelve pre-cultures of SBW25 were grown overnight to be used in the control treatment. At t_0, 12 microcosms were inoculated with ~ 10^4 cells of ancestral SBW25 and incubated statically for 72 h. After 72 h, each of the 12 microcosms were homogenised for approximately 60 s by vortexing. Samples from each microcosm were diluted 100-fold and 6 µl of each diluted sample was used to inoculate fresh microcosms. This transfer regime was repeated until 36 days of experimentation were completed.

7.2.17 CONGO RED BINDING ASSAY

This assay was adapted from Malone (2005). Each of the IWS and SBW25-Δ*wspF* were pre-cultured overnight and ten replicates of each were incubated statically for 72 h at 28˚C. Each static culture was homogenized by vortexing for 30 s. 667 µl of 0.005% Congo red was mixed with 333 µl of vortexed sample in Eppendorf tubes. This was incubated for 2 h at 28˚C to allow the Congo red to bind the bacterial cellulose. Cells were pelleted by centrifugation at 13,200 rpm. 200 µl of the

supernatant of each replicate was transferred into a microtitre plate. The microtitre plate was placed in a VERSAmax™ Absorbance Microplate Reader (Molecular Devices) and the samples analysed at 490 nm. The measurements were blanked using water.

7.2.18 PHOTOGRAPHIC ANALYSIS

WS types to be photographed were grown overnight in shaking cultures, diluted appropriately and spread on KB agar plates. The plates were incubated for 48 h. All photographs were taken using equipment provided in the School of Biological Science at the University of Auckland, courtesy of Iain McDonald. Photographs were taken with an eight second exposure time at F11 with 65 mm macro and 4x magnification. Extreme care was taken to generate pictures with conditions as similar as possible to one another for the computer analyses of the pictures. This included using a completely darkened room, using the light from the lamps at identical angles to avoid shadowing or reflection and placement of the colony in the same spot without moving the black background or the lens of the camera.

The pictures were analysed using SigmaScan Pro (Systat Software Inc.) with a modified macro. The source code for the visual basic macro is available in Appendix 9.2.3. The macro creates a threshold based on contrast to differentiate between the colony and the black background. A new object is overlayed consisting only of pixels that constitute the image of the colony. The area and perimeter of the object are calculated using the macro (Appendix 9.2.3). To calculate the circularity of the colony, the measured perimeter is compared to the estimated perimeter, which is calculated from the area assuming the colony was a circle. The ratio of the measured perimeter to the estimated perimeter will be larger than one whenever the colony is wrinkly and the further the deviation from one is an increase in wrinkliness. The estimated perimeter (estimated from the area) of the colony can be calculated by the following equation:

$$estimated\ perimeter = \frac{2\,area}{\sqrt{\frac{area}{\pi}}} \qquad \text{Eq 7-1}$$

While, SM colonies may be expected to have a perimeter that approximately equals the value calculated from the area, WS colonies are not circular, and this can be used to measure their 'wrinkliness' or circularity. The colony circularity is a ratio of the expected perimeter (calculated by Eq 7-1) to the observed perimeter. The observed colony perimeter is calculated by following the edge of the colony against the background. The ratio is expected to equal one for circular colonies, but will be greater than one for colonies that have a wrinkled morphology, because the observed perimeter will be increased relative to the expected perimeter.

7.2.19 MAT STRENGTH ASSAY

Overnight pre-cultures were grown for each WS among the IWS genotypes and for WS-Δ*wspF*. Eight replicates for each of 27 WS were grown statically for 72 h at 28˚C in racks. After 72 h, these racks were gently placed and fastened on an Innova 2300 platform shaker (New Brunswick Scientific). The shaker was set to 40 rpm and run for 10 s. The number of microcosms where the mat was no longer intact was recorded. The speed of the shaker was incremented by 5 rpm and the process repeated. Incrementing continued until the mats of all eight replicates were no longer intact. Mats were deemed destroyed by any partial detachment of their structure from the microcosm walls (most common) or if any ripping or fragmentation of the mats occurred.

7.2.20 ARTEMIS

Artemis (Rutherford *et al.*, 2000) was used to view the entire annotated *P. fluorescens* genome (http://www.sanger.ac.uk/Projects/P_fluorescens/private/). This was used for effective primer design, and the identification of transposon insertion sites based on sequence. Artemis is free and available online and for download at http://www.sanger.ac.uk/Software/Artemis/.

7.2.21 SEQUENCHER

Sequencher 4.7.1 (Gene Codes Corporation) was used to import electropherogram data for visual inspection. Sequential sequences were assembled into contiguous sequences automatically for comparison with known genome sequence to identify mutations.

7.2.22 GENEIOUS

Geneious (Drummond *et al.*, 2007) was used to align sequences and display of alignments. Geneious is available for a free trial at http://www.geneious.com.

7.2.23 STATISTICS

JMP 5.0.1 (S.A.S. Institute) was used for all statistical analyses with the exception of the ANOSIM, which was performed using R v2.7 (Ihaka and Gentleman, 1996).

7.2.23.1 RANDOMISATION TEST FOR A DIFFERENCE BETWEEN CURVES

This randomisation test tests for the difference between curves from types A and B. This approach is necessary for serially sampled non-independent data to avoid problems associated with multiple testing, and incorrect testing of pseudo-independent data. The data for the two curves were pooled and plotted together. A sixth order polynomial was fit to the data to generate a curve for the combined data. The difference between each observed data point and the estimated data point from the line of best fit was calculated to generate the residuals. The residuals were analysed using an ANOVA for a difference between the two types A and B.

7.2.23.2 RANDOMISATION TEST FOR A DIFFERENCE BETWEEN STRAIGHT LINES

This randomisation test is identical to the test between curves except the data are known to fit approximately to a straight line. The data for the two groups was pooled and plotted together. A straight line was fit to the data to generate a line of best fit for the combined data. The difference between each observed data point and the estimated data point from the line of best fit was calculated to generate the residuals. The residuals were analysed using an ANOVA for a difference between the two types A and B. This approach is necessary for non-independent data to avoid problems associated with multiple testing, and incorrect testing of pseudo-independent data.

7.2.24 EXPERIMENTS IN A MODIFIED PETRI DISH

7.2.24.1 DIVERSIFICATION IN A MODIFIED PETRI DISH

Deep Petri dishes were filled with 56 ml of KB media and the Petri dish adaptor was gently placed resting above the edges. This modified Petri dish, with the lid placed over the top will be referred to as a modified Petri dish (MPD). MPDs were inoculated with 5.6 ml of an overnight culture of SBW25 (or LSWS-*LacZ*) and incubated statically at 28˚C. MPDs were destructively sampled every day for a total of ten days. At each sample time point, a 6 mm diameter, 10 cm bolt was gently screwed into the top of the Petri dish adaptor to act as a handle. The mats formed within the holes of the Petri dish adaptor were removed in their entirety by gently manoeuvring the adaptor while holding the bolt. The entire contents of the MPD was pipetted into a 100 ml bottle and vortexed for 60 s to destroy the mats. The contents was diluted and plated appropriately to determine the proportion of SM and WS types.

7.2.24.2 A MULTI-LEVEL SELECTION EXPERIMENT

Seven microcosms were inoculated with SBW25 and incubated statically for seven days at 28˚C. The contents of each microcosm were destructively sampled, and the most common WS type from each microcosm was selected to be grown overnight for this experiment. 750 µl of each of the overnight cultures was stored at -80˚C and 8 µl of a ten fold diluted sample (1/7 of 5.6 µl undiluted sample) of each were inoculated into each of two MPDs filled with 56 ml of KB media. One of the MPDs was used as the experimental MLS (multi-level selection) treatment and the other (paired) MPD was used as the control GLS (group level selection) treatment. MPDs were incubated for five days at 28˚C. After this incubation, portions of mats were removed from each of the seven wells in the Petri dish adaptor using a sterile wire loop and place in seven 1.5 ml Eppendorf tube with 500 µl of KB media and vortexed for 30 s. The cells were diluted appropriately, plated on KB plates and incubated for 48 h at 28˚C. A single WS colony from each plate was inoculated into a standard KB microcosm for static incubation and an overnight culture for storing at -80˚C. The inoculated microcosms were incubated for six days statically at 28˚C. For the MLS treatment, each mat was dislodged from the walls of the MPD and 50 µl of supernatant is pipetted into a single

1.5 ml Eppendorf tube with 500 µl KB media. For the GLS treatment, the mat was sampled directly with a sterile wire loop and each sample was placed in a single 1.5 ml Eppendorf tube with 500 µl KB media and vortexed for 30 s. The OD of each is measured to insure equal inocula between the contents of each of the seven microcosms. The contents of each of the seven microcosms were mixed in equal quantities, diluted appropriately and plated. 5.6 µl of the ten fold diluted sample was inoculated into the MPDs to begin a new group-generation. The experiment was continued for a total of six group-generations.

7.2.24.3 MEASURING A RESPONSE TO MULTI-LEVEL SELECTION

To measure a response to the multi-level selection experiment described above, a randomly selected WS from the final time point of the experimental treatment (MLS) was grown in competition with an ancestral WS genotype (LSWS-*lacZ*) according to the protocol of the experimental treatment. After each five-day incubation at 28˚C in the MGPDs, the ratio of blue to white colonies was calculated for mats sampled using a sterile wire loop from each of the seven wells in the Petri dish adaptor to determine the dominant type (blue or white) in each well. The dominant type within each of the wells was determined by Malthusian ratio from the initial global inoculum for that MPD. Given the single cell bottleneck of the dominant type to enforce strong selection, dominance of a group equates to one group-offspring. The experiment was continued as per the above protocol and repeated for six group-generations. The number of group-offspring after six group-generations was used to determine the fitness in an MLS 2 sense.

8 REFERENCES

Acuña, G., W. Y. Shi, K. Trudeau and D. R. Zusman (1995). The CheA and CheY Domains of *Myxococcus xanthus* FrzE Function Independently in-Vitro as an Autokinase and a Phosphate Acceptor, Respectively. *FEBS Letters* **358**(1): 31-33.

Amsler, D. A. and P. Matsumura (1995). Chemotactic Signal Transduction in *Escherichia coli* and *Salmonella Typhimurium*. Two-component Signal Transduction. J. A. Hoch and T. J. Shilhavy. Washington, ASM Press.

Atkinson, Q. D. and R. D. Gray (2005). Curious Parallels and Curious Connections - Phylogenetic Thinking in Biology and Historical Linguistics. *Systematic Biology* **54**(4): 513-26.

Axelrod, R. and W. D. Hamilton (1981). The Evolution of Cooperation. *Science* **211**(4489): 1390-6.

Bantinaki, E., R. Kassen, C. G. Knight, Z. Robinson, A. J. Spiers and P. B. Rainey (2007). Adaptive Divergence in Experimental Populations of *Pseudomonas fluorescens*. III. Mutational Origins of Wrinkly Spreader Diversity. *Genetics* **176**(1): 441-53.

Beaumont, H. J., J. Gallie, C. Kost and P. B. Rainey (2008). The Evolution of a Switching Phenotype in *Pseudomonas fluorescens*. *Manuscript in preparation*.

Beecher, M. D. (1982). Signature Systems and Kin Recognition. *American Zoologist* **22**: 477-90.

Bertani, G. (1951). Studies on Lysogenesis. I. The Mode of Phage Liberation by Lysogenic *Escherichia coli*. *Journal of Bacteriology* **62**(3): 293-300.

Bertani, G. (2004). Lysogeny at Mid-twentieth Century: P1, P2, and other Experimental Systems. *Journal of Bacteriology* **186**(3): 595-600.

Bonner, J. T. (2001). *First Signals: The Evolution of Multicellular Development.* Princeton, NJ, Princeton University Press.

Brockhurst, M. A. (2007). Population Bottlenecks Promote Cooperation in Bacterial Biofilms. *PLoS ONE* **2**: e634.

Brockhurst, M. A., M. E. Hochberg, T. Bell and A. Buckling (2006). Character Displacement Promotes Cooperation in Bacterial Biofilms. *Current Biology* **16**(20): 2030-4.

Buckling, A., F. Harrison, M. Vos, M. A. Brockhurst, A. Gardner, S. A. West and A. Griffin (2007). Siderophore-mediated Cooperation and Virulence in *Pseudomonas aeruginosa*. *Fems Microbiology Ecology* **62**(2): 135-41.

Buckling, A. and P. B. Rainey (2002). The Role of Parasites in Sympatric and Allopatric Host Diversification. *Nature* **420**(6915): 496-9.

Buckling, A., M. A. Wills and N. Colegrave (2003). Adaptation Limits Diversification of Experimental Bacterial Populations. *Science* **302**(5653): 2107-09.

Buss, L. W. (1987). *The Evolution of Individuality*. Princeton, NJ, Princeton University Press.

Campbell, N. and J. Reece (2005). *Biology*. 7th. San Francisco, CA, Benjamin Cummings.

Cavalier-Smith, T. (2006). Rooting the Tree of life by Transition Analyses. *Biology Direct* **1**: 19.

Christen, B., M. Christen, R. Paul, F. Schmid, M. Folcher, P. Jenoe, M. Meuwly and U. Jenal (2006). Allosteric Control of Cyclic di-GMP Signaling. *Journal of Biological Chemistry* **281**(42): 32015-24.

Clarke, K. K. (1993). Non-parametric Multivariate Analysis of Changes in Community Structure. *Austral Ecology* **18**: 117-43.

Connor, R. C. (1986). Pseudo-reciprocity: Investing in Mutualism. *Animal Behaviour* **34**: 1562–84.

Connor, R. C. (1995). Altruism Among Non-relatives: Alternatives to the 'Prisoner's Dilemma'. *Trends in Ecology & Evolution* **10**: 84-86.

Corning, P. A. (2002). The Re-emergence of "Emergence": A Venerable Concept in Search of a Theory. *Complexity* **7**(6): 18-30.

Cosmides, L. M. and J. Tooby (1981). Cytoplasmic Inheritance and Intragenomic Conflict. *Journal of Theoretical Biology* **89**(1): 83-129.

Costerton, J. W., Z. Lewandowski, D. E. Caldwell, D. R. Korber and H. M. Lappin-Scott (1995). Microbial Biofilms. *Annual Review of Microbiology* **49**: 711-45.

Crespi, B. J. (2001). The Evolution of Social Behaviour in Microorganisms. *Trends in Ecology & Evolution* **16**: 178-83.

D'Argenio, D. A. and S. I. Miller (2004). Cyclic di-GMP as a Bacterial Second Messenger. *Microbiology* **150**(Pt 8): 2497-502.

Dale, J., D. B. Lank and H. K. Reeve (2001). Signaling Individual Identity vs. Quality: A Model and Case Studies with Ruffs, Queleas and House Finches *American Naturalist* **158**: 75–86.

Damuth, J. and I. L. Heisler (1988). Alternative Formulations of Multilevel Selection. *Biology & Philosophy* **3**(4): 407-30.

Darwin, C. (1859). *The Origin of Species*. 1st. London, UK, John Murray.

Darwin, C. (1871). *The Descent of Man in Relation to Sex*. New York, NY, Appleton.

Dawkins, R. (1976). *The Selfish Gene*. New York, NY, Oxford University Press.

Dawkins, R. (2006). *The Selfish Gene*. 30th anniversary edition. New York, NY, Oxford University Press.

De Vos, D., M. De Chial, C. Cochez, S. Jansen, B. Tummler, J. M. Meyer and P. Cornelis (2001). Study of Pyoverdine Type and Production by *Peudomonas aeruginosa* Isolated from Cystic Fibrosis Patients: Prevalence of Type II Pyoverdine Isolates and Accumulation of Pyoverdine-negative Mutations. *Archives of Microbiology* **175**(5): 384-8.

Dean, A. M., D. E. Dykhuizen and D. L. Hartl (1986). Fitness as a Function of Beta-Galactosidase Activity in *Escherichia coli*. *Genetical Research* **48**(1): 1-8.

Dennett, D. C. (1995). *Darwin's Dangerous Idea*. New York, NY, Simon and Schuster.

Diggle, S. P., A. Gardner, S. A. West and A. S. Griffin (2007a). Evolutionary Theory of Bacterial Quorum Sensing: When is a Signal not a Signal? *Philosophical Transactions of the Royal Society of London B Biological Sciences* **362**(1483): 1241-9.

Diggle, S. P., A. S. Griffin, G. S. Campbell and S. A. West (2007b). Cooperation and conflict in quorum-sensing bacterial populations. *Nature* **450**(7168): 411-U7.

Djordjevic, S. and A. M. Stock (1997). Crystal Structure of the Chemotaxis Receptor Methyltransferase CheR Suggests a Conserved Structural Motif for Binding S-adenosylmethionine. *Structure* **5**(4): 545-58.

Dobzhansky, T. (1964). Biology, Molecular and Organismic. *American Zoologist* **4**: 443-52.

Drenkard, E. and F. M. Ausubel (2002). *Pseudomonas* Biofilm Formation and Antibiotic Resistance are Linked to Phenotypic Variation. *Nature* **416**(6882): 740-3.

Drummond, A. J., B. Ashton, M. Cheung, J. Heled, M. Kearse, R. Moir, S. Stones-Havas, T. Thierer and A. Wilson (2007). Geneious.

DTU. (2008). Retrieved 11/02/2008, from http://www.cbs.dtu.dk/services/TMHMM-1.0/tmhmm1.0.html.

Dunny, G. M., T. J. Brickman and M. Dworkin (2008). Multicellular Behavior in Bacteria: Communication, Cooperation, Competition and Cheating. *Bioessays* **30**(4): 296-98.

Dykhuizen, D. E. (1990). Experimental Studies of Natural Selection in Bacteria. *Annual Review of Ecology and Systematics* **21**(1): 373-98.

Eldredge, N. (1985). *Unfinished Synthesis: Biological Hierarchies and modern Evolutionary Thought*. New York, NY, Oxford University Press.

Falke, J. J., R. B. Bass, S. L. Butler, S. A. Chervitz and M. A. Danielson (1997). The Two-component Signaling Pathway of Bacterial Chemotaxis: A Molecular View of Signal Transduction by Receptors, Kinases, and Adaptation Enzymes. *Annual Review of Cell and Developmental Biology* **13**: 457-512.

Figurski, D. H. and D. R. Helinski (1979). Replication of an Origin-containing Derivative of Plasmid RK2 Dependent on a Plasmid Function Provided *in trans*. *Proceedings of the National Academy of Sciences of the United States of America* **76**(4): 1648-52.

Fisher, R. A. (1930). *The Genetical Theory of Natural Selection*. London, UK, Oxford University Press.

Fisher, R. A. (1958). *The Genetical Theory of Natural Selection*. 2nd. New York, NY, Dover.

Foster, K. R., A. Fortunato, J. E. Strassmann and D. C. Queller (2002). The costs and benefits of being a chimera. *Proceedings of the Royal Society of London Series B-Biological Sciences* **269**(1507): 2357-62.

Foster, K. R., K. Parkinson and C. R. Thompson (2007). What can Microbial Genetics Teach Sociobiology. *Trends in Genetics* **23**(2): 74-80.

Frank, S. A. (1986). Hierarchical Selection Theory and Sex Ratios. I. General Solutions for Structured Populations. *Theoretical Population Biology* **29**(3): 312-42.

Frank, S. A. (1995). Mutual Policing and Repression of Competition in the Evolution of Cooperative Groups. *Nature* **377**(6549): 520-2.

Frank, S. A. (1998). *Foundations of Social Evolution*. Princeton, NJ, Princeton University Press.

Fukami, T., H. J. Beaumont, X. X. Zhang and P. B. Rainey (2007). Immigration History Controls Diversification in Experimental Adaptive Radiation. *Nature* **446**(7134): 436-9.

Fuqua, C., M. R. Parsek and E. P. Greenberg (2001). Regulation of gene expression by cell-to-cell communication: Acyl-homoserine lactone quorum sensing. *Annual Review of Genetics* **35**: 439-68.

Futuyma, D. J. (1998). *Evolutionary Biology*. 3rd. Sunderland, MA, Sinauer Associates.

Gal, M., G. M. Preston, R. C. Massey, A. J. Spiers and P. B. Rainey (2003). Genes Encoding a Cellulosic Polymer Contribute Toward the Ecological Success of *Pseudomonas fluorescens* SBW25 on Plant Surfaces. *Molecular Ecology* **12**(11): 3109-21.

Gehrig, S. (2005). Adaptation of *Pseudomonas fluorescens* SBW25 to the Air-liquid Interface: A Study in Evolutionary Genetics. Department of Plant Sciences and Wolfson College, University of Oxford.

Giddens, S. R., R. W. Jackson, C. D. Moon, M. A. Jacobs, X. X. Zhang, S. M. Gehrig and P. B. Rainey (2007). Mutational activation of niche-specific genes provides insight into regulatory networks and bacterial function in a complex environment. *Proceedings of the National Academy of Sciences of the United States of America* **104**(46): 18247-52.

Gilbert, S. F. (2000). *Developmental biology*. Sunderland, MA, Sinauer Associates.

Gilpin, M. E. (1975). Group Selection in Predator-prey Communities. *Monographs in Population Biology*(9): 1-108.

Gould, S. J. (1989). *Wonderful Life*. New York, NY, W. W. Norton and Co.

Gould, S. J. (2002). *The Structure of Evolutionary Theory*. Cambridge, MA, Belknap Press.

Gould, S. J. and R. C. Lewontin (1979). Spandrels of San-Marco and the Panglossian Paradigm - a Critique of the Adaptationist Program. *Proceedings of the Royal Society of London Series B-Biological Sciences* **205**(1161): 581-98.

Gould, S. J. and E. S. Vrba (1982). Exaptation - a Missing Term in the Science of Form. *Paleobiology* **8**(1): 4-15.

Goymer, P., S. Kahn, J. Malone, S. Gehrig, A. Spiers and P. B. Rainey (2006). Adaptive Divergence in Experimental Populations of *Pseudomonas fluorescens*. II. The Role of the GGDEF Regulator WspR in Evolution and Development of the Wrinkly Spreader Phenotype. *Genetics* **173**(2): 515-26.

Grafen, A. (1984). Natural Selection, Kin Selection, and Group Selection. Behavioural Ecology: An Evolutionary Approach. J. R. Krebs and N. B. Davies. Oxford, U.K., Blackwell Scientific Puablications**:** 62-84.

Greig, D. and M. Travisano (2004). The Prisoner's Dilemma and Polymorphism in Yeast SUC Genes. *Proceedings of the Biological Society of Washington* **271 Suppl 3**: S25-6.

Griesemer, J. (2001). The Units of Evolutionary Transition. *Selection* **1**(1): 67-80.

Griffin, A. S., S. A. West and A. Buckling (2004). Cooperation and Competition in Pathogenic Bacteria. *Nature* **430**(7003): 1024-7.

Guerinot, M. L. (1994). Microbial Iron Transport. *Annual Review of Microbiology* **48**: 743-72.

Hamilton, W. D. (1963). The Evolution of Altruistic Behavior. *American Naturalist* **97**: 354-56.

Hamilton, W. D. (1964a). The Genetical Evolution of Social Behaviour. I. *Journal of Theoretical Biology* **7**(1): 1-16.

Hamilton, W. D. (1964b). The Genetical Evolution of Social Behaviour. II. *Journal of Theoretical Biology* **7**(1): 17-52.

Hamilton, W. D. (1970). Selfish and Spiteful Behaviour in an Evolutionary Model. *Nature* **228**(5277): 1218-20.

Hamilton, W. D. (1975). Innate Social Aptitudes in Man: An Approach from Evolutionary Genetics. Biosocial Anthropology, Wiley and Sons**:** 133-55.

Hansen, S. K., P. B. Rainey, J. A. Haagensen and S. Molin (2007). Evolution of Species Interactions in a Biofilm Community. *Nature* **445**(7127): 533-6.

Hardin, G. (1968). The Tragedy of the Commons. The Population Problem has no Technical Solution; it Requires a Fundamental Extension in Morality. *Science* **162**(859): 1243-8.

Harrison, F., L. E. Browning, M. Vos and A. Buckling (2006). Cooperation and Virulence in Acute *Pseudomonas aeruginosa* Infections. *BMC Biology* **4**: 21.

Harrison, F. and A. Buckling (2005). Hypermutability Impedes Cooperation in Pathogenic Bacteria. *Current Biology* **15**(21): 1968-71.

Harrison, F., J. Paul, R. C. Massey and A. Buckling (2008). Interspecific Competition and Siderophore-mediated Cooperation in *Pseudomonas aeruginosa*. *Isme Journal* **2**(1): 49-55.

Heisler, I. L. and J. Damuth (1987). A Method for Analyzing Selection in Hierarchically Structured Populations. *American Naturalist* **130**(4): 582-602.

Henke, J. M. and B. L. Bassler (2004). Bacterial Social Engagements. *Trends in Cell Biology* **14**(11): 648-56.

Hodgson, D. J., P. B. Rainey and A. Buckling (2002). Mechanisms Linking Diversity, Productivity and Invasibility in Experimental Bacterial Communities. *Proceedings of the Royal Society of London Series B-Biological Sciences* **269**(1506): 2277-83.

Hull, D. (1980). Individuality and Selection. *Annual Review of Ecology and Systematics* **11**: 311-32.

Huxley, J. (1942). *Evolution: The Modern Synthesis*. London, UK, Allen and Unwin.

Ihaka, R. and R. Gentleman (1996). R: A Language for Data Analysis and Graphics. *Journal of Computational and Graphical Statistics* **5**(3): 299-314.

Jacobs, M. A., A. Alwood, I. Thaipisuttikul, D. Spencer, E. Haugen, S. Ernst, O. Will, R. Kaul, C. Raymond, R. Levy, C. R. Liu, D. Guenthner, D. Bovee, M. V. Olson and C. Manoil (2003). Comprehensive Transposon Mutant Library of *Pseudomonas aeruginosa*. *Proceedings of the National Academy of Sciences of the United States of America* **100**(24): 14339-44.

Jenal, U. (2004). Cyclic Di-guanosine-monophosphate Comes of Age: a Novel Secondary Messenger Involved in Modulating Cell Surface Structures in Bacteria? *Current Opinion in Microbiology* **7**(2): 185-91.

Jenal, U. and J. Malone (2006). Mechanisms of Cyclic-di-GMP Signaling in Bacteria. *Annual Review of Genetics* **40**: 385-407.

Keller, L. and M. G. Surette (2006). Communication in Bacteria: An Ecological and Evolutionary Perspective. *Nature Reviews Microbiology* **4**(4): 249-58.

Kerr, B., P. Godfrey-Smith and M. W. Feldman (2004). What is Altruism? *Trends in Ecology & Evolution* **19**(3): 135-40.

Kerr, B., C. Neuhauser, B. J. M. Bohannan and A. M. Dean (2006). Local Migration Promotes Competitive Restraint in a Host-pathogen 'Tragedy of the Commons'. *Nature* **442**(7098): 75-78.

Kimura, M. (1968). Evolutionary Rate at the Molecular Level. *Nature* **217**(5129): 624-6.

King, E. O., M. K. Ward and D. E. Raney (1954). Two Simple Media for the Demonstration of Pyocyanin and Fluorescin. *Journal of Laboratory and Clinical Medicine* **44**(2): 301-7.

Knight, C. G., N. Zitzmann, S. Prabhakar, R. Antrobus, R. Dwek, H. Hebestreit and P. B. Rainey (2006). Unraveling Adaptive Evolution: How a Single Point Mutation Affects the Protein Coregulation Network. *Nature Genetics* **38**(9): 1015-22.

Lazarsfeld, P. F. and H. Menzel (1961). On the Relation Between Individual and Collective Properties. Complex organizations: A Sociological reader. A. Etzioni. New York, Holt, Rinehart and Winston**:** 422-40.

Lehmann, L. and L. Keller (2006). The Evolution of Cooperation and Altruism - A General Framework and a Classification of Models. *Journal of Evolutionary Biology* **19**: 1365-76.

Lenski, R. E. (1992). Experimental Evolution. Encyclodepia of Microbiology. J. Lederrera. San Diego, Academic Press. **2**.

Lenski, R. E., M. R. Rose, S. C. Simpson and S. C. Tadler (1991). Long-Term Experimental Evolution in Escherichia-Coli .1. Adaptation and Divergence during 2,000 Generations. *American Naturalist* **138**(6): 1315-41.

Levins, R. (1970). Extinction. Some mathematical questions in biology. M. Gerstenhaber. Providence, R. I., American Mathematical Society. **2:** 77-107.

Levit, M. N., T. W. Grebe and J. B. Stock (2002). Organization of the Receptor-kinase Signaling Array that Regulates *Escherichia coli* Chemotaxis. *Journal of Biological Chemistry* **277**(39): 36748-54.

Lewes, G. H. (1875). *Problems of Life and Mind*. London, UK, Trübner.

Lewontin, R. C. (1961). Evolution and the Theory of Games. *Journal of Theoretical Biology* **1**: 382-403.

Lewontin, R. C. (1970). The Units of Selection. *Annual Review of Ecology and Systematics* **1**: 1-18.

Li, Y., V. H. Bustamante, R. Lux, D. Zusman and W. Shi (2005). Divergent Regulatory Pathways Control A and S Motility in *Myxococcus xanthus* through FrzE, a CheA-CheY Fusion Protein. *Journal of Bacteriology* **187**(5): 1716-23.

Liu, Z., F. R. Stirling and J. Zhu (2007). Temporal Quorum-sensing Induction Regulates *Vibrio cholerae* Biofilm Architecture. *Infection and Immunity* **75**(1): 122-26.

Lloyd, E. A. (1988). *The Structure and Confirmation of Evolutionary Theory*. Princeton, NJ, Princeton University Press.

MacLean, R. C. and G. Bell (2003). Divergent Evolution During an Experimental Adaptive Radiation. *Proceedings of the Royal Society of London Series B-Biological Sciences* **270**(1524): 1645-50.

MacLean, R. C., G. Bell and P. B. Rainey (2004). The Evolution of a Pleiotropic Fitness Tradeoff in *Pseudomonas fluorescens*. *Proceedings of the National Academy of Sciences of the United States of America* **101**(21): 8072-7.

MacLean, R. C., A. Dickson and G. Bell (2005). Resource Competition and Adaptive Radiation in a Microbial Microcosm. *Ecology Letters* **8**(1): 38-46.

Malone, J. (2005). Analysis of the Structure-function Relationship of WspR, a GGDEF Response-regulator. Department of Plant Sciences and St. Catherine's College. Oxford, University of Oxford. **D. Phil**.

Malone, J. G., R. Williams, M. Christen, U. Jenal, A. J. Spiers and P. B. Rainey (2007). The Structure-function Relationship of WspR, a *Pseudomonas fluorescens* Response Regulator with a GGDEF Output Domain. *Microbiology* **153**(Pt 4): 980-94.

Manoil, C. (2000). Tagging Exported Proteins using *Escherichia coli* Alkaline Phosphatase Gene Fusions. *Applications of Chimeric Genes and Hybrid Proteins, Pt A* **326**: 35-47.

Marchler-Bauer, A., J. B. Anderson, P. F. Cherukuri, C. DeWeese-Scott, L. Y. Geer, M. Gwadz, S. He, D. I. Hurwitz, J. D. Jackson, Z. Ke, C. J. Lanczycki, C. A. Liebert, C. Liu, F. Lu, G. H. Marchler, M. Mullokandov, B. A. Shoemaker, V. Simonyan, J. S. Song, P. A. Thiessen, R. A. Yamashita, J. J. Yin, D. Zhang and S. H. Bryant (2005). CDD: a Conserved Domain Database for Protein Classification. *Nucleic Acids Research* **33**(Database issue): D192-6.

Maynard Smith, J. (1964). Group Selection and Kin Selection. *Nature* **201**: 1145-47.

Maynard Smith, J. (1976). Group Selection. *Quarterly Review of Biology* **51**(2): 277-83.

Maynard Smith, J. (1982). *Evolution and the Theory of Games*. Cambridge, UK, Cambridge University Press.

Maynard Smith, J. (1988). Evolutionary Progress and Levels of Selection Evolutionary Progress. M. H. Nitecki. Chicago and London, The Univ. of Chicago Press**:** 219–30.

Maynard Smith, J., R. Burian, S. Kauffman, P. Alberch, J. Campbell, B. Goodwin, R. Lande, D. Raup and L. Wolpert (1985). Developmental Constraints and Evolution: A Perspective from the Mountain Lake Conference on Development and Evolution. *Quarterly Review of Biology* **60**(3): 265-87.

Maynard Smith, J. and E. Szathmáry (1995). *The Major Transitions in Evolution*. New York, NY, W.H. Freeman and Company Limited.

Mayr, E. (1980). Prologue: Some Thoughts on the History of the Evolutionary Synthesis. The Evolutionary Synthesis: Perspectives on the Unification of Biology E. Mayr and W. B. Provine. Cambridge, MA, Harvard University Press.

Mayr, E. (1982). *The Growth of Biological Thought: Diversity, Evolution, and Inheritance*. Cambridge, MA, Harvard University Press.

McCleary, W. R., M. J. McBride and D. R. Zusman (1990). Developmental Sensory Transduction in *Myxococcus xanthus* Involves Methylation and Demethylation of FrzCD. *Journal of Bacteriology* **172**(9): 4877-87.

McCleary, W. R. and D. R. Zusman (1990a). FrzE of *Myxococcus xanthus* is Homologous to both CheA and CheY of *Salmonella typhimurium*. *Proceedings of the National Academy of Sciences of the United States of America* **87**(15): 5898-902.

McCleary, W. R. and D. R. Zusman (1990b). Purification and Characterization of the *Myxococcus xanthus* FrzE Protein Shows that it has Autophosphorylation Activity. *Journal of Bacteriology* **172**(12): 6661-8.

McDonald, M., S. Gehrig, X. X. Zhang and P. B. Rainey (2008). Genetic Constraints Guide Evolutionary Trajectories in a Parallel Adaptive Radiation. *Manuscript in preparation*.

Mendel, G. (1865). Veruche ueber Pflanzenhybride. *Verhandlungen des Naturforschenden Vereines Brunn* **4**: 3-47.

Michod, R. E. (1982). The Theory of Kin Selection. *Annual Review of Ecology and Systematics* **13**: 23-55.

Michod, R. E. (1999). *Darwinian Dynamics*. Princeton, NJ, Princeton University Press.

Michod, R. E. and D. Roze (1997). Transitions in Individuality. *Proceedings of the Biological Society of Washington* **264**(1383): 853-7.

Michod, R. E. and D. Roze (2001). Cooperation and Conflict in the Evolution of Multicellularity. *Heredity* **86**: 1-7.

Michod, R. E., Y. Viossat, C. A. Solari, M. Hurand and A. M. Nedelcu (2005). Life-history Evolution and the Origin of Multicellularity. *Journal of Theoretical Biology* **239**(2): 257-72.

Miller, M. B. and B. L. Bassler (2001). Quorum Sensing in Bacteria. *Annual Review of Microbiology* **55**: 165-99.

Nadell, C. D., J. B. Xavier, S. A. Levin and K. R. Foster (2008). The Evolution of Quorum Sensing in Bacterial Biofilms. *PLoS Biology* **6**(1): 171-79.

NCBI Conserved Domain Search. (2007). "Conserved Domain Search." Retrieved 11/02/2008, from http://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi.

Nielsen, A. T., N. A. Dolganov, G. Otto, M. C. Miller, C. Y. Wu and G. K. Schoolnik (2006). RpoS controls the *Vibrio cholerae* Mucosal Escape Eesponse. *Plos Pathogens* **2**(10): 933-48.

Nowak, M. (2006). Five Rules for the Evolution of Cooperation. *Science* **314**(5805): 1560.

Nowell, P. C. (1976). Clonal Evolution of Tumor-Cell Populations. *Science* **194**(4260): 23-28.

O'Sullivan, D. J. and F. O'Gara (1992). Traits of Fluorescent *Pseudomonas spp.* Involved in Suppression of Plant Root Pathogens. *Microbiology Reviews* **56**(4): 662-76.

Okasha, S. (2006a). *Evolution and the Levels of Selection*. Oxford, UK, Oxford University Press.

Okasha, S. (2006b). The Levels of Selection Debate. *Philosophy Compass*: 74-85.

Oksanen, J., R. Kindt, P. Legendre, B. O'Hara, G. L. Simpson and M. H. H. Stevens (2008). Vegan: Community Ecology Package.

Palleroni, N. (1984). *Pseudomonadaceae*. Bergey's Manual of Systematic Bacteriology. N. Kreig and J. Holt. Baltimore, Williams & Wilkins Co. **1:** 141-99.

Parkinson, J. S. and E. C. Kofoid (1992). Communication Modules in Bacterial Signaling Proteins. *Annual Review of Genetics* **26**: 71-112.

Parsek, M. R. and E. P. Greenberg (2005). Sociomicrobiology: The Connections Between Quorum Sensing and Biofilms. *Trends in Microbiology* **13**(1): 27-33.

Paul, R., S. Weiser, N. C. Amiot, C. Chan, T. Schirmer, B. Giese and U. Jenal (2004). Cell cycle-dependent dynamic localization of a bacterial response regulator with a novel di-guanylate cyclase output domain. *Genes & Development* **18**(6): 715-27.

Queller, D. C. (1994). Genetic Relatedness in Viscous Populations. *Evolutionary Ecology* **8**(1): 70-73.

Queller, D. C. and J. E. Strassmann (1998). Kin Selection and Social Insects. *Bioscience* **48**(3): 165-75.

Rainey, P. B. (1999). Adaptation of *Pseudomonas fluorescens* to the Plant Rhizosphere. *Environmental Microbiology* **1**(3): 243-57.

Rainey, P. B. (2007). Unity from Conflict. *Nature* **446**(7136): 616.

Rainey, P. B. and M. J. Bailey (1996). Physical and Genetic Map of the *Pseudomonas fluorescens* SBW25 Chromosome. *Molecular Microbiology* **19**(3): 521-33.

Rainey, P. B., A. Buckling, R. Kassen and M. Travisano (2000). The Emergence and Maintenance of Diversity: Insights from Experimental Bacterial Populations. *Trends in Ecology & Evolution* **15**(6): 243-47.

Rainey, P. B. and T. F. Cooper (2004). Evolution of Bacterial Diversity and the Origins of Modularity. *Research in Microbiology* **155**(5): 370-5.

Rainey, P. B. and B. Kerr (2009). The Emergence of Individuality during the Evolutionary Transition from Single-celled Organisms to Multicellularity. *Manuscript in preparation*.

Rainey, P. B. and K. Rainey (2003). Evolution of Cooperation and Conflict in Experimental Bacterial Populations. *Nature* **425**(6953): 72-4.

Rainey, P. B. and M. Travisano (1998). Adaptive Radiation in a Heterogeneous Environment. *Nature* **394**(6688): 69-72.

Rand, D. M., R. A. Haney and A. J. Fry (2004). Cytonuclear Coevolution: The Genomics of Cooperation. *Trends in Ecology & Evolution* **19**(12): 645-53.

Ratledge, C. and L. G. Dover (2000). Iron Metabolism in Pathogenic Bacteria. *Annual Review of Microbiology* **54**: 881-941.

Redfield, R. J. (2002). Is Quorum Sensing a Side Effect of Diffusion Sensing? *Trends in Microbiology* **10**(8): 365-70.

Ridley, M. (2004). *Evolution*. 3rd. Oxford, UK, Blackwell Publishing.

Rosenberg, E., K. H. Keller and M. Dworkin (1977). Cell Density-Dependent Growth of *Myxococcus xanthus* on Casein. *Journal of Bacteriology* **129**(2): 770-77.

Ross, P., R. Mayer, H. Weinhouse, D. Amikam, Y. Huggirat, M. Benziman, E. de Vroom, A. Fidder, P. de Paus, L. A. Sliedregt and et al. (1990). The Cyclic Diguanylic Acid Regulatory System of Cellulose Synthesis in *Acetobacter xylinum*. Chemical Synthesis and Biological Activity of Cyclic Nucleotide Dimer, Trimer, and Phosphothioate Derivatives. *Journal of Biological Chemistry* **265**(31): 18933-43.

Ross-Gillespie, A., A. Gardner, S. A. West and A. S. Griffin (2007). Frequency Dependence and Cooperation: Theory and a Test with Bacteria. *American Naturalist* **170**(3): 331-42.

Rozen, D. E. and R. E. Lenski (2000). Long-term Experimental Evolution in *Escherichia coli* VIII. Dynamics of a Balanced Polymorphism. *American Naturalist* **155**(1): 24-35.

Rutherford, K., J. Parkhill, J. Crook, T. Horsnell, P. Rice, M. A. Rajandream and B. Barrell (2000). Artemis: Sequence Visualization and Annotation. *Bioinformatics* **16**(10): 944-5.

Sachs, J. L., U. G. Mueller, T. P. Wilcox and J. J. Bull (2004). The Evolution of Cooperation. *Quarterly Review of Biology* **79**(2): 135-60.

Sambrook, J., E. F. Fritsch and T. Maniatis (1989). *Molecular Cloning. A Laboratory Manual*. Woodbury, NY, Coldspring Harbour Laboratory Press.

Sandoz, K. M., S. M. Mitzimberg and M. Schuster (2007). Social Cheating in *Pseudomonas aeruginosa* Quorum Sensing. *Proceedings of the National Academy of Sciences of the United States of America* **104**(40): 15876-81.

Sanger Institute. (2007). "*Pseudomonas fluorescens*." Retrieved 22/06/2007, from http://www.sanger.ac.uk/Projects/P_fluorescens/.

Schluter, D. (2000). *The Ecology of Adaptive Radiation*. Oxford, UK, Oxford University Press.

Shapiro, J. A. (1998). Thinking about Bacterial Populations as Multicellular Organisms. *Annual Review of Microbiology* **52**(1): 81-104.

Shapiro, J. A. and M. Dworkin (1997). *Bacteria as Multicellular Organisms*. Oxford, UK, Oxford University Press.

Sharp, S. P., A. McGowan, M. J. Wood and B. J. Hatchwell (2005). Learned Kin Recognition Cues in a Social Bird. *Nature* **434**(7037): 1127-30.

Simms, E. L. and D. L. Taylor (2002). Partner Choice in Nitrogen-fixation Mutualisms of Legumes and Rhizobia. *Integrative and Comparative Biology* **42**(2): 369–80.

Simpson, G. G. (1953). *The Major Features of Evolution*. New York, Columbia University Press.

Smolin, L. (1992). Did the Universe Evolve. *Classical and Quantum Gravity* **9**(1): 173-91.

Sober, E. and D. S. Wilson (1998). *Unto Others*. Cambridge, MA, Harvard University Press.

Spiers, A. J., J. Bohannon, S. M. Gehrig and P. B. Rainey (2003). Biofilm Formation at the Air-liquid Interface by the *Pseudomonas fluorescens* SBW25 Wrinkly Spreader Requires an Acetylated form of Cellulose. *Molecular Microbiology* **50**(1): 15-27.

Spiers, A. J., A. Buckling and P. B. Rainey (2000). The causes of *Pseudomonas* diversity. *Microbiology* **146 (Pt 10)**: 2345-50.

Spiers, A. J., S. G. Kahn, J. Bohannon, M. Travisano and P. B. Rainey (2002). Adaptive Divergence in Experimental Populations of *Pseudomonas fluorescens*. I. Genetic and Phenotypic Bases of Wrinkly Spreader Fitness. *Genetics* **161**(1): 33-46.

Spiers, A. J. and P. B. Rainey (2005). The *Pseudomonas fluorescens* SBW25 Wrinkly Spreader Biofilm Requires Attachment Factor, Cellulose Fibre and

LPS Interactions to Maintain Strength and Integrity. *Microbiology* **151**(Pt 9): 2829-39.

Sterelny, K. (2001). *Dawkins Vs. Gould: Survival of the Fittest*. Cambridge, UK, Icon Books UK.

Stewart, C. B., J. W. Schilling and A. C. Wilson (1987). Adaptive Evolution in the Stomach Lysozymes of Foregut Fermenters. *Nature* **330**(6146): 401-04.

Sundstrom, L., M. Chapuisat and L. Keller (1996). Conditional Manipulation of Sex Ratios by Ant Workers: A Test of Kin Selection Theory. *Science* **274**(5289): 993-5.

Tarpy, D. R., D. C. Gilley and T. D. Seeley (2004). Levels of Selection in a Social Insect: A Review of Conflict and Cooperation During Honey Bee (*Apis mellifera*) Queen Replacement. *Behavioral Ecology and Sociobiology* **55**(6): 513-23.

Taylor, D. R., C. Zeyl and E. Cooke (2002). Conflicting Levels of Selection in the Accumulation of Mitochondrial Defects in *Saccharomyces cerevisiae*. *Proceedings of the National Academy of Sciences of the United States of America* **99**(6): 3690-94.

Thompson, J. N. (1999). The evolution of species interactions. *Science* **284**(5423): 2116-18.

Travisano, M., J. A. Mongold, A. F. Bennett and R. E. Lenski (1995a). Experimental Tests of the Roles of Adaptation, Chance, and History in Evolution. *Science* **267**(5194): 87-90.

Travisano, M., F. Vasi and R. E. Lenski (1995b). Long-Term Experimental Evolution in *Escherichia coli* III. Variation among Replicate Populations in Correlated Responses to Novel Environments. *Evolution* **49**(1): 189-200.

Travisano, M. and G. J. Velicer (2004). Strategies of Microbial Cheater Control. *Trends in Microbiology* **12**(2): 72-78.

Trivers, R. L. (1971). The Evolution of Reciprocal Altruism. *Quarterly Review of Biology* **46**(1): 35-57.

Trivers, R. L. (1972). Parental Investment and Sexual Selection. Sexual Selection and the Descent of Man. B. Campbell. Chicago, Aldine**:** 136-79.

Velicer, G. J. (2003). Social Strife in the Microbial World. *Trends in Microbiology* **11**(7): 330-37.

Vrba, E. S. (1983). Macroevolutionary Trends - New Perspectives on the Roles of Adaptation and Incidental Effect. *Science* **221**(4608): 387-89.

Vrba, E. S. (1984). What is Species Selection? *Systematic Zoology* **33**(3): 318-28.

Wade, M. J. (1977). An Experimental Study of Group Selection. *Evolution* **31**(1): 134-53.

Wade, M. J. (1978). A Critical Review of the Models of Group Selection. *Quarterly Review of Biology* **53**(2): 101-14.

Wandersman, C. and P. Delepelaire (2004). Bacterial iron sources: From siderophores to hemophores. *Annual Review of Microbiology* **58**: 611-47.

Weber, H., C. Pesavento, A. Possling, G. Tischendorf and R. Hengge (2006). Cyclic-di-GMP-mediated Signalling within the Sigma Network of *Escherichia coli*. *Molecular Microbiology* **62**(4): 1014-34.

Weiner, R., E. Seagren, C. Arnosti and E. Quintero (1999). Bacterial Survival in Biofilms: Probes for Exopolysaccharide and its Hydrolysis, and Measurements of Intra- and Interphase Mass Fluxes. *Methods Enzymology* **310**(1): 403-26.

West, S. A. and A. Buckling (2003). Cooperation, Virulence and Siderophore Production in Bacterial Parasites. *Proceedings of the Biological Society of Washington* **270**(1510): 37-44.

West, S. A., A. S. Griffin and A. Gardner (2007a). Evolutionary Explanations for Cooperation. *Current Biology* **17**(16): R661-R72.

West, S. A., A. S. Griffin and A. Gardner (2007b). Social Semantics: Altruism, cooperation, Mutualism, Strong Reciprocity and Group Selection. *Journal of Evolutionary Biology* **20**(2): 415-32.

West, S. A., A. S. Griffin and A. Gardner (2008). Social Semantics: How Useful has Group Selection Been? *Journal of Evolutionary Biology* **21**(1): 374-85.

West, S. A., A. S. Griffin, A. Gardner and S. P. Diggle (2006). Social Evolution Theory for Microorganisms. *Nature Reviews Microbiology* **4**(88): 597-607.

West, S. A., I. Pen and A. S. Griffin (2002). Cooperation and Competition Between Relatives. *Science* **296**(5565): 72-5.

Whitfield, D. P. (1986). Plumage Variability and Territoriality in Breeding Turnstone *Arenaria interpres*: Status Signalling or Individual Recognition? *Animal Behaviour* **34**(5): 1471-82.

Whitfield, D. P. (1987). Plumage Variability, Status Signalling and Individual Recognition in Avian Flocks. *Trends in Ecology & Evolution* **2**(1): 13-18.

Wichman, H. A., M. R. Badgett, L. A. Scott, C. M. Boulianne and J. J. Bull (1999). Different Trajectories of Parallel Evolution during Viral Adaptation. *Science* **285**(5426): 422-24.

Williams, G. C. (1966). *Adaptation and Natural Selection*. Princeton, NJ, Princeton University Press.

Williams, G. C. (1992). *Natural Selection: Domains, Levels and Challenges*. Oxford, UK, Oxford University Press.

Wilson, D. S. (1975). A Theory of Group Selection. *Proceedings of the National Academy of Sciences of the United States of America* **72**(1): 143-46.

Wilson, D. S. (1977). Structured Demes and Evolution of Group-Advantageous Traits. *American Naturalist* **111**(977): 157-85.

Wilson, D. S. (1980). *The Natural Selection of Populations and Communities*. Menlo Park, CA, Benjamin Cummings.

Wilson, D. S. (1983). The Group Selection Controversy: History and Current Status. *Annual Review of Ecology and Systematics* **14**(1): 159-87.

Wilson, D. S. (1987). Altruism in Mendelian Populations Derived from Sibling Groups - the Haystack Model Revisited. *Evolution* **41**(5): 1059-70.

Wilson, D. S. (2008). Social Semantics: Toward a Genuine Pluralism in the Study of Social Behaviour. *Journal of Evolutionary Biology* **21**(1): 368-73.

Wilson, E. O. (1975). *Sociobiology*. Cambridge, MA, Harvard University Press.

Wingreen, N. S. and S. A. Levin (2006). Cooperation among Microorganisms. *PLoS Biology* **4**(9): e299.

Wright, S. (1932). The Roles of Mutation, Inbreeding, Crossbreeding, and Selection in Evolution. *Proceedings of The Sixth International Congress of Genetics* **1**: 356-66.

Wright, S. (1945). Tempo and Mode in Evolution: A Critical Review. *Ecology* **26**(4): 415-19.

Wynne-Edwards, V. C. (1962). *Animal Dispersion in Relation to Social Behavior*. London, UK, Oliver & Boyd.

Zhang, X. X. and P. B. Rainey (2007). Construction and Validation of a Neutrally-marked Strain of *Pseudomonas fluorescens* SBW25. *Journal of Microbiological Methods* **71**(1): 78-81.

Zhu, J. and J. J. Mekalanos (2003). Quorum Sensing-dependent Biofilms Enhance Colonization in *Vibrio cholerae*. *Developmental Cell* **5**(4): 647-56.

9 APPENDICES

This appendix is organised by original chapter. Appendix 9.1 has items referred to in Chapter 1, 9.2 in Chapter 2 *etc*.

9.1 APPENDIX ITEMS FROM CHAPTER 1

9.1.1 CHAPTER 1.3 THE POWER OF MICROBIAL MODEL SYSTEMS

9.1.1.1 SBW25-*LACZ* IS NEUTRALLY MARKED IN LONG-TERM EXPERIMENTS

Integration at the Tn*7* attachment site (attTn*7*) has been shown to have no measurable fitness effect on multiple substrates and *in planta* (Zhang and Rainey, 2007); however, its neutrality in long-term fitness assays was unknown. Therefore, its long-term fitness was determined over a ten-day competition experiment in a shaken environment with daily serial transfer (Section 7.2.15.3). The ratio of Malthusian parameters (Lenski *et al.*, 1991) was measured to determine the fitness of SBW25-*lacZ* relative to the ancestral SBW25 in competition. No significant difference in fitness was found for 10 days of competition between the marked and unmarked strains ($t_9 = 0.447$, $p = 0.665$), and no differences in fitness were observed between any time points within the 10 days of competition (Table 9-1).

Table 9-1: Ten-day competition assay between SBW25 and SBW25-*lacZ*.

Growth Period	Relative fitness of *lacZ* marked strain	*t*-stat for Malthusian Ratio	*p*-value
t_0-t_1	1.003	0.375	0.716
t_1-t_2	1.004	0.565	0.586
t_2-t_3	1.000	0.058	0.955
t_3-t_4	1.003	0.290	0.778
t_4-t_5	0.992	0.910	0.386
t_5-t_6	1.005	0.575	0.580
t_6-t_7	1.006	0.435	0.674
t_7-t_8	1.001	0.059	0.954
t_8-t_9	0.994	0.983	0.351
t_9-t_{10}	1.003	0.389	0.707
t_0-t_{10}	1.001	0.447	0.665

Malthusian growth parameters were measured for SBW25 and SBW25-*lacZ* in competition with one another in shaking cultures for every time point over a ten-day period. *T*-tests on the data revealed that there was no significant different between the *lacZ* marked strain and the wild-type SBW25 (n=10).

9.1.2 Chapter 1.4.1.1 Diversification in a Microcosm

9.1.2.1 Fine-scale Fitness Dynamics of LSWS in Competition with SBW25 in a Static Environment

To determine the fitness dynamics of LSWS more precisely, equal densities of SBW25-*lacZ* and LSWS were inoculated into 54 (3 replicates of 18 time points) microcosms. Microcosms were incubated statically at 28˚C for 48 h. Every 4 h, microcosms were sampled destructively by vortexing and the contents diluted and plated appropriately to determine the relative proportions of SM and WS (Figure 9-1). Relative fitness was expressed as an SRC, the difference between two Malthusian parameters (Lenski *et al.*, 1991). The SRC is more applicable than the ratio of Malthusian parameters in this instance, because it can tolerate negative values for Malthusian parameters under circumstances in which there are a net loss in cell numbers. A decrease in cell numbers was expected, particularly for SBW25, because of the steep oxygen gradient, its inability to form a mat and the short time intervals between sampling. The SRC, illustrated by the dashed line in the figure, is calculated over a time interval of 4 h (*i.e.* 0 – 4, 4 – 8, 8 – 12 etc.), while the cumulative SRC is measured over increasing intervals (*i.e.* 0 – 4, 0 – 8, 0 – 12 etc.). Negative values for the SRC indicate that the proportion of WS has decreased between the two time points and positive values for the SRC indicate the proportion of WS types has increased. The results for the cumulative SRC agree with previous estimates of relative fitness for LSWS during a 24 h incubation period (Malthusian ratio of 0.8 for 24 h), but in addition, the SRC over shorter time intervals shows that LSWS has a greater fitness than SBW25 between 20 – 24 h.

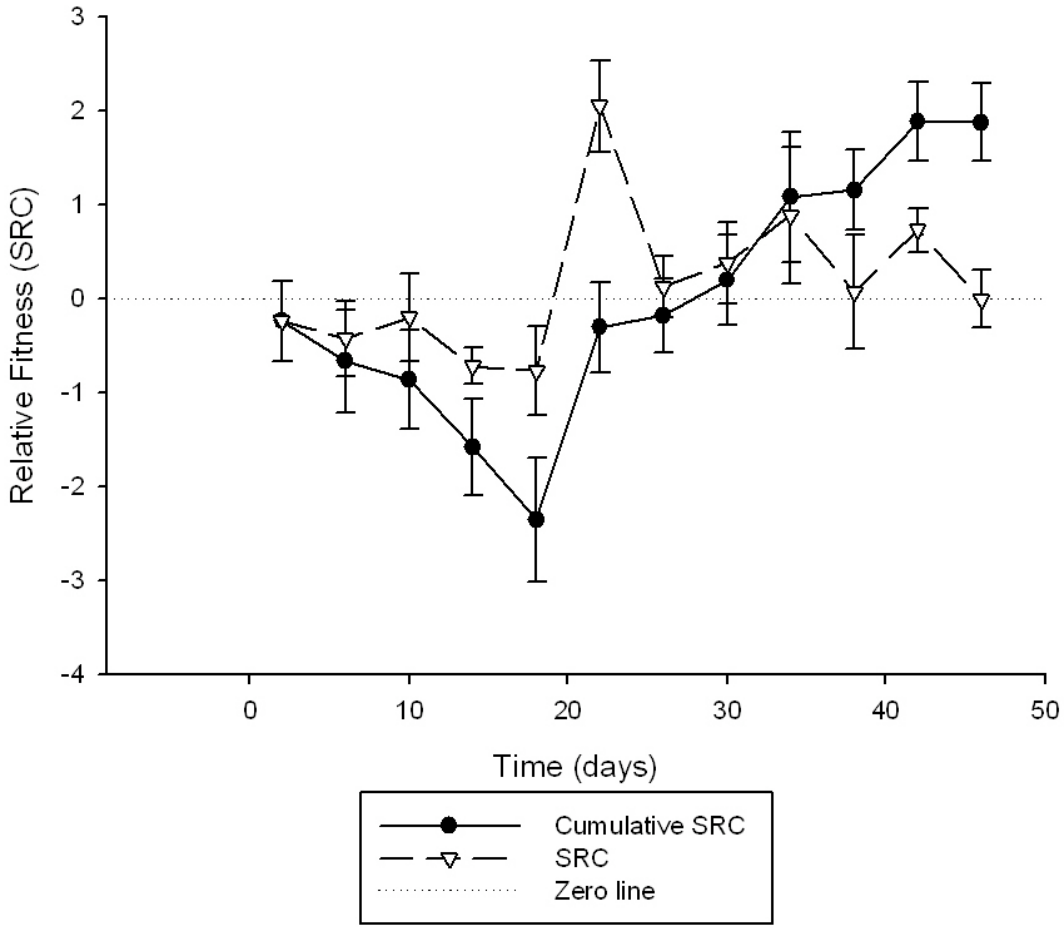

Figure 9-1: Fine-scale fitness (SRC) dynamics for LSWS compared to SBW25 over 48 h.
The dashed line with open triangles shows how the SRC changes over 48 h in a static microcosm for four-h time intervals. The solid line with filled circles shows the cumulative SRC, which is calculated using the same data, but over increasing intervals from t_0 to the time point in question. The dotted 'zero' line for SRC is congruent with a relative fitness of 1 for a Malthusian ratio.

9.1.3 CHAPTER 1.5.3 THE AWS LOCUS

Table 9-2: NCBI BlastP hits for *aws*X on 11.10.2007.

NCBI Reference	Description	Organism	Score	E-value
ref|YP_346424.1	Hypothetical protein Pfl_0691	*P. fluorescens* Pf0-1	274	2e-72
ref|YP_257880.1	Hypothetical protein PFL_0743	*P. fluorescens* Pf-5	273	4e-72
gb|AAL79574.1	Unknown	*P. syringae*	228	1e-58
ref|ZP_01638503.1	Conserved hypothetical protein	*P. putida* W619	214	3e-54
ref|YP_606436.1	Hypothetical protein PSEEN0695	*P. entomophila*	213	7e-54
ref|YP_001269839.1	Hypothetical protein Pput_4535	*P. putida F1*	205	1e-51
ref|NP_746780.1	Hypothetical protein PP_4671	*P. putida* KT2440	204	3e-51
ref|ZP_01716930.1	Conserved hypothetical protein	*P. putida* GB-1	202	8e-51
ref|YP_001349602.1	Hypothetical protein PSPA7_4248	*P. aeruginosa* PA7	188	2e-46
ref|ZP_00970461.1	Hypothetical protein PaerC01000097	*P. aeruginosa* C3719	182	8e-45

9.2 APPENDIX ITEMS FROM CHAPTER 2

9.2.1 CHAPTER 2.2.1 IDENTIFYING THE MUTATIONAL ROUTES TO WS

Table 9-3: List of GGDEF/GGEEF domain proteins (PF009900) in *P. fluorescens* SBW25.

CDS	Precise motif	E-value (PF00990)	EAL Domain (PF00563)
pflu0085	GGEEF	9.4e-63	No
pflu0183	GGEEF	6.8e-54	No
pflu0185	GGDEF	8.7e-65	Yes
pflu0458	ASNEF	1.3e-49	Yes
pflu0621	GGEEF	9.3e-50	No
pflu0956	GGEEF	9.0e-71	No
pflu1081	AGSEF	2.6e-20	Yes
pflu1083	IADEF	5.4e-19	Yes
pflu1114	GSDAF	1.1e-44	Yes
wspR	GGEEF	3.2e-16	No
pflu1349	GGDEF	1.7e-57	Yes
pflu2031	GGDDF	3.3e-5	Yes
pflu2361	GGEEF	3.6e-29	No
pflu2606	GGEEF	4.1e-59	No
pflu2753	GGDEF	4.6e-56	Yes
pflu2764	GGEEF	4.9e-12	No
pflu3255	GGDEF	9.6e-52	No
pflu3444	GGDEF	3.7e-61	Yes
pflu3448	GGDEF	2.0e-45	No
pflu3571	GGEEF	9.1e-65	No
pflu3650	GGDEF	5.9e-56	No
pflu3699	AGDRF	6.9e-23	No
pflu3760	GGDEF	2.9e-57	No
pflu4116	GGEEF	1.2e-52	No
pflu4198	GGDEF	8.1e-69	Yes
pflu4306	GGEEF	2.1e-54	No
pflu4308	GGDEF	3.6e-57	Yes
pflu4600	GGEEF	2.5e-53	No
pflu4782	GGDEF	1.1e-56	No
pflu4858	GGDQF	2.2e-39	Yes
pflu5127	GGDEF	9.6e-48	No
pflu5210 (awsR)	GGDEF	7.5e-58	No
pflu5329 (mwsR)	GGDEF	1.1e-62	Yes
pflu5593	GGDEF	1.7e-69	Yes
pflu5608	GGEEF	8.0e-39	No
pflu5698	GGDEF	1.0e-55	Yes
pflu5960	GSDEF	1.1e-34	Yes
pflu6074	SGDEF	2.8e-51	Yes

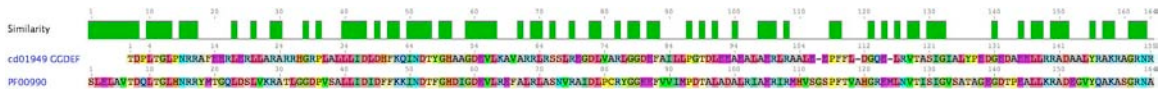

Figure 9-2: Alignment of PF0090 and cd01949.

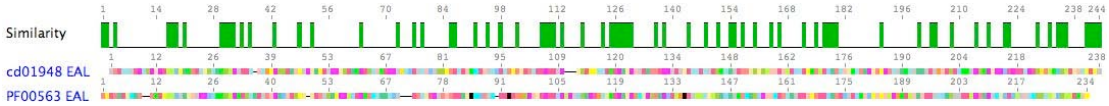

Figure 9-3: Alignment of PF00563 and cd01948.

Table 9-4: IWS_D transposon insertion sequence data.

Transconjugant	Sequence	Insertion site	Unique
1	AGGGGTCGAGTCTGACCAT	*wsp* Operon	Yes
2	AATCGGTACGCCGAGGGAGA	*wsp* Operon	Yes
3	GTTTGCGAGTGCCACGTCCA	*wsp* Operon	Yes
4	GATCGGCCTCGGCTTC	*wsp* Operon	Yes
5	GGCAAACTCCTCCCCCCCA	*wsp* Operon	Yes
6	GCGGCATGTTCGAAGTGCAG	*wsp* Operon	Yes
7	CGCCGGGCGAATACTGGCA	*wsp* Operon	Yes
8	CGGCAGCACTTCGGCCACT	*wsp* Operon	Yes
9	GCGCGAGTGATAGCGGATG	*wsp* Operon	Yes
10	CCTCGACATCCGCCGGCCCG	*wsp* Operon	Yes
11	TGCGCTTTTTCACCGGGGAA	*wsp* Operon	Yes
12	CACCGATCATGGCCTGGTCGT	*pflu5744*	Yes
13	TCGGTCATCATGTCGTTAAAGCCGGTTTC	*wsp* Operon	Yes
14	ATGGGTTGATCGGGATGTCT	*pflu5871*	Yes
15	CAGGGGACGGATTGTCGGCA	*wsp* Operon	Yes
16	CAGGCCAGCTTTGCCATTGA	*wsp* Operon	Yes

Table 9-5: IWS_H transposon insertion sequence data.

Transconjugant	Sequence	Insertion site	Unique
1	TCCTGGCTGACGCTGCCC	*pflu080*	Yes
2	TGGACGGTATTGAAT	*wsp* Operon	Yes
3	TTGGTTGGCAGCGGT	*wsp* Operon	Yes
4	TCGTCGAGCACGTGG	*wsp* Operon	Yes
5	AATGCGCGCATCTGC	*wsp* Operon	Yes
6	CTGACCATCAGTATT	*wsp* Operon	Yes
7	GCTCTTGCTGTTGAC	*wsp* Operon	Yes
8	CGCCTGGCGATCCTCGCCTTG	*wsp* Operon	Yes
9	CCACCACTTGTGCCTCGGT	*wsp* Operon	Yes
10	TTGCAAACGGGCCAGGTAC	*wsp* Operon	Yes
11	CTCGAAGCCTGCATGCG	*wsp* Operon	Yes
12	ACCCTGTTGCAACGGGAT	*wsp* Operon	Yes
13	CTGCGACTGGCCGTTGTGGCA	*wsp* Operon	Yes
14	ACCCTGTTGCCGATCAGCTC	*wsp* Operon	Yes

Table 9-6: IWS_I transposon insertion sequence data.

Transconjugant	Sequence	Insertion site	Unique
1	CCGACCGTTATCCTGCAGGATC	*wsp* Operon	Yes
2	CGCAAAACGGCTTC	*wsp* Operon	Yes
3	AGGGGTCGAGTCTGACCATC	*wsp* Operon	Yes
4	GCTGCCTTTGGCGGTAAAGCC	*wsp* Operon	Yes
5	GCGTTGGCTCTGGA	*pflu5871*	Yes
6	CCTGATGATGCGCATG	*fabG*	Yes
7	TGCTCAAGTACAGCTACTCGCCGG	*wsp* Operon	Yes
8	GACAGTACCGCCTCGTCGA	*wsp* Operon	Yes
9	GCTCCCGACAGCGCCGCGCT	*wsp* Operon	Yes
10	CCTGCGCACCAAGTGGTGT	*wsp* Operon	Yes
11	GACCTGATCCTGATGGACCT	*wsp* Operon	Yes
12	TCAGGTTGATCGCCACC	*wsp* Operon	Yes
13	GCAAACTCCTCCCCCCCATAG	*wsp* Operon	Yes
14	GTTGCTGGAACAGGCGCGTG	*wsp* Operon	Yes
15	GACCGAAGTGTCCACCGCCG	*wsp* Operon	Yes

Table 9-7: IWS_K transposon insertion sequence data.

Transconjugant	Sequence	Insertion site	Unique
1	CAGGCCACCGGCCGT	*aws* Operon	Yes
2	GATCGGCTGGCTGAT	*aws* Operon	Yes
3	CAGGTCGGCCTGAGCTGTG	*pflu2552*	Yes
4	TACACCTCACGGGTGTTG	*mreC*	Yes
5	CTCAATTCCTCCG	*mreC*	Yes
6	CCGCGAGTTGGGGCGCCG	*wssC*	Yes
7	GCTGGCGCAATGGATAC	*aws* Operon	Yes
8	GTCCAGCAGCGTCGGCTTA	*pflu*0892	Yes
9	TGCGCACGCCACTGCGG	*aws* Operon	Yes
10	TTTGCCTTTGACAGG	*pflu5744*	Yes

Table 9-8: IWS_M transposon insertion sequence data.

Transconjugant	Sequence	Insertion site	Unique
1	CGCCGTGCCCGTGGCC	*mrdB*	Yes
2	GGAGACTTTCTGGCCGAGGC	*pflu5423*	Yes
3	ACGCTTGCTTGATCAGCAGG	*mrdB*	Yes
4	GTGGAAAAAATCGCGCAT	*gatC*	Yes
5	GGGCACGGAGATCGGCAGG	*aws* Operon	Yes
6	AGGTCGTCGGTGTACTGGGT	*aws* Operon	Yes
7	GGCGCTGCTGCGGGTGCAAGCG	*pflu0427*	Yes
8	CGAATCCCTCGCCCACCAGGC	*aws* Operon	Yes
9	CTTCCATCTCGCCGAGCAAGGC	*aws* Operon	Yes
10	GCTCAGGCTCAGTCCCGGCAG	*pflu0427*	Yes
11	GCGCCTGGGCGGCGAC	*aws* Operon	Yes
12	CTGATCTATTTGCTGCTGATC	*mrdB*	Yes
13	CTGGCCTGCCAGCGCTCGCCTAA	*pflu0427*	Yes
14	GCTGCCGATGGTTCGCCCTT	*aws* Operon	Yes
15	TTCATCGAGCATCGCGTTGG	*pflu1603*	Yes
16	GTGGTCGACACCAACTTGAA	*fabG*	Yes
17	ACCAGGTCGCCACCGTCGATA	*pflu*0428	Yes
18	GTCAGCTTCGATCTGCGCTT	*icd*	Yes

Table 9-9: IWS_P transposon insertion sequence data.

Transconjugant	Sequence	Insertion site	Unique
1	AACTGGGCACCAAGG	*icd*	Yes
2	GGCAGGTGTATGGCGGTGAACTG	*pflu*2543	Yes
3	GTCATTGATCTCTTTGAAGCGG	*aws* Operon	Yes
4	CTTCCATGACCGCTTCAGGGTCAA	*tolQ*	Yes
5	TGGGTGGCCGATTAATGCGGCAAACAGGT	*aws* Operon	Yes
6	CCGGTGTGCAGGGTTGGCGTG	*pflu5420*	Yes
7	GAGGGAGCCGACCATGACCCGATCATTGCG	*algJ*	Yes
8	ATCAAGGAATTGCAGGCCC	*icd*	Yes
9	CTGAGGAAGCGCGCGAGCGCCTGGTGCA	*pflu4309*	Yes
10	ACTCAATGCTGCCGACGCAGCGATGTACCA	*aws* Operon	Yes
11	ATGCCCTTCTTCGACATGCT	*pflu5217*	Yes
12	GGTTGTGTTCGGCATACACCCGCA	*aws* Operon	Yes
13	GCACCAGCCGCATCATCAACAGGCTCGC	*pflu1603*	Yes
14	CGAACGGCTCGTCGCCGATCAGTGGACGGCC	*pflu2985*	Yes

Table 9-10: IWS_O transposon insertion sequence data.

Transconjugant	Sequence	Insertion site	Unique
1	TCAGCAGCACGATTACCGGGATCAGC	*mwsR*	Yes
2	GCCCCACGAGCGCAGGGTGATGGCGTCG	*pflu4623*	Yes
3	GGCGATGATCGACTGCCAGCCGTTCT	*mwsR*	Yes
4	TGAAATTCTGCGGTGAGGCT	*mwsR*	Yes
5	CGCTCACGCCGCCCACGAGG	*wssE*	Yes
6	TGCCCAGGCCCTGATAGATGCGCGCGG	*wssE*	Yes
7	CAGGTCAGGTTGATCGCCAC	*wssB*	Yes
8	ACTTCCTGCGTGAGCATGGGTGTG	*aws*	Yes
9	CTGGTGGATACCGAA	*wssB*	Yes
10	TCCAACGCCTTGAAAGCCTTGGTCAT	*wssE*	Yes
11	ATCTGGTTTTCCAGCTCGTAGCTGAT	*pflu2985*	Yes
12	CGCTCGGCCTCCGGCTCGCCCAGTTCACGC	*wssE*	Yes
13	GCAGTGGCAACTGTCGCAGAACCTG	*wssC*	Yes
14	ACTATATAATGGGGCGCTTTAA	*pflu1667*	Yes
15	TCAAGTACAGCTACTCGCCG	*wssC*	Yes
16	CAGGATATTTTCGAAGGCACCGG	*mwsR*	Yes
17	GCTGGCGGCCCTGCGCGAAGTCGGCTCCA	*wssB*	Yes
18	CGCCTCACTGTATTTGTTGGCC	*wssE*	Yes
19	CGACAGGTCAGCCAAACGGGTGAACGGG	*wssC*	Yes

Table 9-11: IWS_R transposon insertion sequence data.

Transconjugant	Sequence	Insertion site	Unique
1	CTGGAACAACTGCTGCCGGCGTTC	*mwsR*	Yes
2	ATGGGCTTCTCGGCCTCCAC	wssD	Yes
3	GGCAGGTGGGTCAGGGCGT	*mwsR*	Yes
4	TCAACGCTTTAGTTCGCGCAGTCACAACCGC	*pflu1667*	Yes
5	GCAGCAACAGAGCCTGAATCGCTGGCT	*wssA*	Yes
6	CAACACCGTGTATGTCCATCGCGTCGGC	*wssC*	Yes
7	CATACCCTTCGCTGCTCCGGGT	*pflu3542*	Yes
8	TACATGCGCGCCAAGGCGA	*wssE*	Yes
9	CACCACTTCGGGCGGCTGCAGGTGGTCGGCTT	*wssA*	Yes
10	CGACCTGGTGGCCCAGGCCGAACGCCTGA	*wssE*	Yes
11	CGGATGAGGGCGAGCGAATATGTTCGAGGTC	*wssB*	Yes
12	AGCGCAACTGGCATTGCAATTGGGTGGTG	*mwsR*	Yes
13	ACTCGGCCGCGACCAGGCCGT	*wssB*	Yes
14	GTTCACTCCATCGGTGCGTGCGGTGCCGG	*wssF*	Yes
15	ATATTTTCGAAGGCACCGGCGT	*mwsR*	Yes
16	TGACAGCCAGGACAGCGCCTTGCTC	*mwsR*	Yes
17	CTATAAGCCGGACGCCCGGAGGGCGGCTA	*mwsR*	Yes
18	CCATGCTGCTGCTGGGCATTATC	*wssB*	Yes
19	CTGCACCGTCACGGCCGGTGA	*wssC*	Yes
20	CTCCAACTGGTGCTGTTGGGTCAGCACGGC	*mwsR*	Yes

Table 9-12: IWS_S transposon insertion sequence data.

Transconjugant	Sequence	Insertion site	Unique
1	CGATGACCGAGCGCAGCGTCGGTCGCA	*aws* Operon	Yes
2	GTGCCGGAAAGCGCGGGTGGTGTTCGA	*dapF*	Yes
3	TGCCTGTAGGACGTTGCTCGCAGAA	*pflu5025*	Yes
4	CCGAGGAACAGTGTGTAGGCCGGGTT	*pflu5288*	Yes
5	ACCAAATCGGCCAGCGGCAGCAG	*pflu0629*	Yes
6	GGGCATGATTGACACCGAAGGCC	*pflu2593*	Yes
7	GATGGTGCTCAAATCGACGTCCCTGTG	*ecfE*	Yes
8	TCGTCGCCGCCCAGGCGCGCCACCAGGT	*aws* Operon	Yes
9	TGGACCTGCCGATCTCCGTGCCCGGC	*aws* Operon	Yes
10	CAATTGCGCCTGTGTATCGTCGGCCC	*aws* Operon	Yes
11	CCAATGCGGCACCCAGGCGT	*pflu2766*	Yes
12	TCGGCATACACCCGCAGC	*aws* Operon	Yes
13	TACCTGGGTGTTGCCGGGCACG	*aws* Operon	Yes
14	ACCGTGGTCACGCCGTTGCTGTGGAAAGCCCG	*pflu2544*	Yes
15	CCCAGCGCCAGGGTGTTGCCGGT	*pflu5960*	Yes
16	GACGAACTGGGAGAGGACACT	*envC*	Yes
17	GGTGTCCTTCGAAGTCAACCTGGA	*aws* Operon	Yes

Table 9-13: IWS_Y transposon insertion sequence data.

Transconjugant	Sequence	Insertion site	Unique
1	TTCGGTTTTGTTGACTTGGA	*pflu0679*	Yes
2	CCATTCCGTTGGGGGAGTGG	*pflu4330*	Yes
3	CTGGACTTTTTGCCCGAGAG	*mrdB*	Yes
4	CCTGTTCACCACAGCAAGCCTGTTCACCAC	*pflu4534*	Yes
5	ACCTTTGTTTTCGGTGCCGT	*eno*	Yes
6	TGCCAGCGACTTGATGGGCG	*pflu5574*	Yes
7	AATCCAAAACACTCATATTA	*pflu1662*	Yes
8	GATCCGTTCCAAGGTGATCA	*pflu1102*	Yes
9	GGTGAGTTTTTCTCCGTCACC	*pflu4060*	Yes
10	ACGTGGTCACCACCGGCCTG	*pflu4060*	Yes
11	TTGGGTGAAGAAATACCCGT	*pflu0472*	Yes
12	AGGTCAAGGTGCAACCGCTG	*pflu2749*	Yes
13	AACGCCGTACAGGGCAAGCC	*icd*	Yes
14	ATTCGGTCGTGAGTACATCA	*pflu0405*	Yes
15	GCGCAGGAAGTTTTCTTCCAG	*icd*	Yes
16	GTGCTGATCTGCGTTCC	*mreB*	Yes
17	GAGCGTGTGCGCGCGCAATT	*aws*	Yes

Table 9-14: IWS_X transposon insertion sequence data.

Transconjugant	Sequence	Insertion site	Unique
1	AAATAATCGTTGGCCCCGG	Wsp Operon	Yes
2	ATACGGATCGCCTGGGCAA	Wsp Operon	Yes
3	AATCAGGTGGGGATTGAGTAGG	Wsp Operon	Yes
4	GACATCATCGGCCTGCGCGA	Wsp Operon	Yes
5	CCCCATAGCGCGCCGGCAAA	Wsp Operon	Yes
6	CGACCAGCAACACCATGGCG	Wsp Operon	Yes
7	CTCCGTGTCGACTTCGGCCG	Wsp Operon	Yes
8	CTTCTCAGCCACCAACCGCG	Wsp Operon	Yes
9	CGCTGAAAATTCCGCACATT	Wsp Operon	Yes
10	AGCTGCTCTCGGTGGGTTTC	Wsp Operon	Yes
11	CGCCGCTCACCCATTTGTTA	Wsp Operon	Yes
12	GGCAGCACCAGGGCAAACTC	Wsp Operon	Yes
13	GTATCCGACCCGTCTACTTTGTAGTGCGCT	*insN* and *pflu*5289	No
14	GGAGTTCATCAGCCGTTGCA	Wss Operon	Yes
15	ATCGATGGCATGGATGCCGT	Wsp Operon	Yes
16	TCCAGGTCGACGCCATTGCC	Wsp Operon	Yes
17	AGATGCGCGCATTGCGTATC	Wsp Operon	Yes
18	CGACCTCGGCCGCAGCCTCG	Wsp Operon	Yes

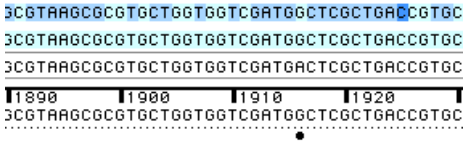

Figure 9-4: Mutation in IWS_D.

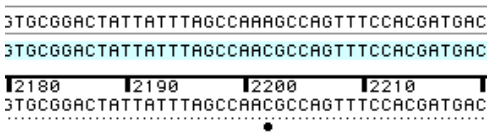

Figure 9-5: Mutation in IWS_H.

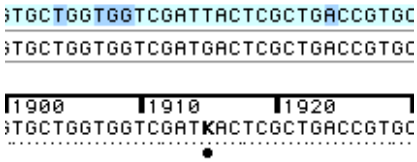

Figure 9-6: Mutation in IWS_I.

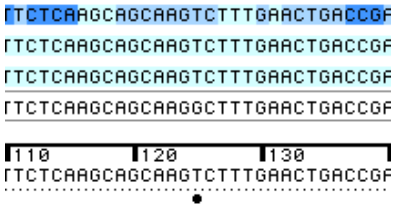

Figure 9-7: Mutation in IWS_K.

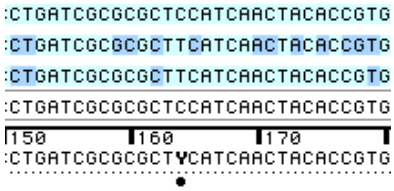

Figure 9-8: Mutation in IWS_M.

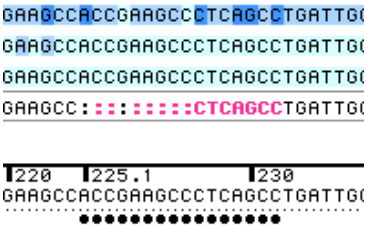

Figure 9-9: Mutation in IWS_P.

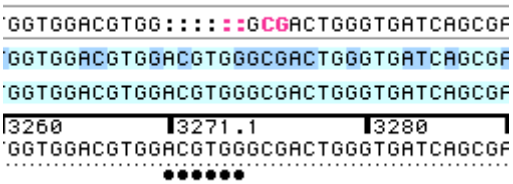

Figure 9-10: Mutation in IWS_Q.

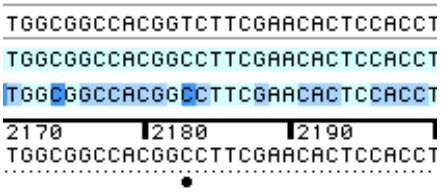

Figure 9-11: Mutation in IWS_R.

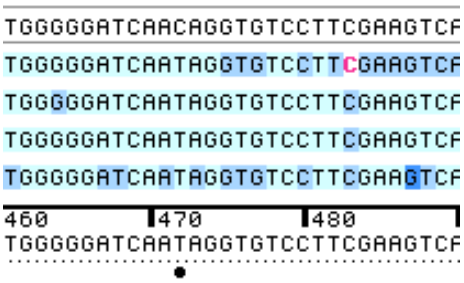

Figure 9-12: Mutation in IWS_S.

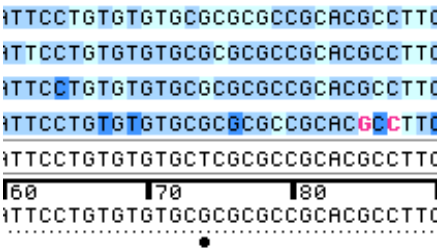

Figure 9-13: Mutation in IWS$_V$.

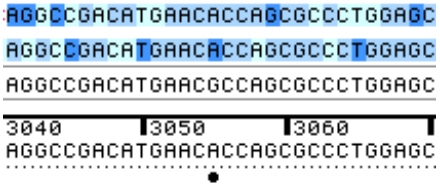

Figure 9-14: Mutation in IWS$_X$.

9.2.2 Chapter 2.2.1.1 Genetic Reconstruction of a *wspE* Mutation

Figure 9-15: Allelic replacement mutation confirmed in SBW25.

9.2.3 Chapter 2.2.2.1 Variation in Colony Size and Wrinkliness

Macro source code used in Sigmascan Pro (Systat Software Inc.) to determined the area of a bacterial colony. The majority of this code is provided with Sigmascan Pro. The changed settings are the intensity thresholds. The lower intensity threshold is set to 50 and the higher intensity threshold to 255 (max).

```
'"#Uses "Constant.bas"
 Dim App As Object
Dim Worksheet As Object
Sub Main
        Set App = CreateObject("SigmaScan.Application")
   Set Worksheet = App.GetWorksheet
   Worksheet.Show
   Worksheet.MakePermanent
        '* Turn off all measurements
        For i=0 To NUMMEASURES-1
```

```
                App.DoNotCollectMeasurement(i)
            Next i
            App.DrawMajMinAxes=False
            '* Load the image
    ExeDirectory = App.GetExeFileDirectory()
    Dim bacterIWSA1 As Object
    '* INPUT FILE BELOW
    Set bacterIWSA1 = App.OpenImage("c:\research\phd\experiments\iws photos\Ian's Camera\IWS-A1.jpg")
    ResultCode = bacterIWSA1.Show
    ResultCode = bacterIWSA1.MakePermanent
    ResultCode = bacterIWSA1.AdjustContrast(100)

    '* Threshold the image into overlay 1
    '*Need to convert to 8 bit first?
    Dim Left0(1) As Long
    Left0(0) = 50
    Dim Right1(1) As Long
    Right1(0) = 255
    ResultCode = bacterIWSA1.IntensityThreshold(1, 1, Left0, Right1)
    Dim Left2(1) As Long
    Left2(0) = 50
    Dim Right3(1) As Long
    Right3(0) = 255
    ResultCode = bacterIWSA1.IntensityThreshold(2, 1, Left2, Right3)
    '* Run binary (overlay) filters on the objects
    ResultCode = bacterIWSA1.FilterOverlay(2, 2, 3, 3, 2)          '* Erode, split objects
    ResultCode = bacterIWSA1.FilterOverlay(5, 3, 4, 1, 2)          '* Dilate everything
    ResultCode = bacterIWSA1.FilterOverlay(6, 4, 5, 2, 2)          '* Dilate, don't merge
    ResultCode = bacterIWSA1.FilterOverlay(10, 5, 5, 1, 2)         '* Remove edge objects
    ResultCode = bacterIWSA1.AndOverlays(5, 1, 5)     '* Logical AND overlays 1 and 5
    '* Zoom the image an hid overlays for easier viewing
    ResultCode = bacterIWSA1.HideOverlay(1)
    ResultCode = bacterIWSA1.HideOverlay(2)
    ResultCode = bacterIWSA1.HideOverlay(3)
    ResultCode = bacterIWSA1.HideOverlay(4)
    '* Measure the objects
            '* Count the objects
            App.CollectMeasurement(M_NUMOBJECTS, "I")
    ResultCode = bacterIWSA1.CountObjects(5)
    NumItems = Worksheet.GetCellValue("I",1)
            '* For each object find its area and shape factor
            App.DoNotCollectMeasurement(M_NUMOBJECTS)
            App.CollectMeasurement(11,  "A")
            App.CollectMeasurement(33, "B")
ResultCode = bacterIWSA1.MeasureObjects(5)
'           MsgBox("Operation Performed.  Click OK to continue.")
            '* Eliminate all objects not sufficiently round
    'First File
Ends#########################################################################################################

            '* Turn off all measurements
            For i=0 To NUMMEASURES-1
                App.DoNotCollectMeasurement(i)
            Next i
            App.DrawMajMinAxes=False
            '* Load the image
    ExeDirectory = App.GetExeFileDirectory()
    Dim bacterIWSA2 As Object
    '* INPUT FILE BELOW
    Set bacterIWSA2 = App.OpenImage("c:\research\phd\experiments\iws photos\Ian's Camera\IWS-A2.jpg")
    ResultCode = bacterIWSA2.Show
    ResultCode = bacterIWSA2.MakePermanent
    ResultCode = bacterIWSA2.AdjustContrast(100)
    '* Threshold the image into overlay 1
    '*Need to convert to 8 bit first?
    ResultCode = bacterIWSA2.IntensityThreshold(1, 1, Left0, Right1)
     ResultCode = bacterIWSA2.IntensityThreshold(2, 1, Left2, Right3)
    '* Run binary (overlay) filters on the objects
    ResultCode = bacterIWSA2.FilterOverlay(2, 2, 3, 3, 2)          '* Erode, split objects
    ResultCode = bacterIWSA2.FilterOverlay(5, 3, 4, 1, 2)          '* Dilate everything
    ResultCode = bacterIWSA2.FilterOverlay(6, 4, 5, 2, 2)          '* Dilate, don't merge
    ResultCode = bacterIWSA2.FilterOverlay(10, 5, 5, 1, 2)         '* Remove edge objects
    ResultCode = bacterIWSA2.AndOverlays(5, 1, 5)     '* Logical AND overlays 1 and 5
    '* Zoom the image an hid overlays for easier viewing
     ResultCode = bacterIWSA2.HideOverlay(1)
```

```
    ResultCode = bacterIWSA2.HideOverlay(2)
    ResultCode = bacterIWSA2.HideOverlay(3)
    ResultCode = bacterIWSA2.HideOverlay(4)
    '* Measure the objects
            '* Count the objects
            App.CollectMeasurement(M_NUMOBJECTS, "I")
    ResultCode = bacterIWSA2.CountObjects(5)
    NumItems = Worksheet.GetCellValue("I",1)
            '* For each object find its area and shape factor
            App.DoNotCollectMeasurement(M_NUMOBJECTS)
            App.CollectMeasurement(11,  "A")
            App.CollectMeasurement(33, "B")
ResultCode = bacterIWSA2.MeasureObjects(5)

'FILE ENDS##############################################################################
            '* Turn off all measurements
            For i=0 To NUMMEASURES-1
                    App.DoNotCollectMeasurement(i)
            Next i
            App.DrawMajMinAxes=False
            '* Load the image
    ExeDirectory = App.GetExeFileDirectory()
    Dim bacterbacterIWSA3 As Object
    '* INPUT FILE BELOW
    Set bacterbacterIWSA3 = App.OpenImage("c:\research\phd\experiments\iws photos\Ian's Camera\IWS-A3.jpg")
    ResultCode = bacterbacterIWSA3.Show
    ResultCode = bacterbacterIWSA3.MakePermanent
    ResultCode = bacterbacterIWSA3.AdjustContrast(100)
    '* Threshold the image into overlay 1
    '*Need to convert to 8 bit first?
    ResultCode = bacterbacterIWSA3.IntensityThreshold(1, 1, Left0, Right1)
    ResultCode = bacterbacterIWSA3.IntensityThreshold(2, 1, Left2, Right3)
    '* Run binary (overlay) filters on the objects
    ResultCode = bacterbacterIWSA3.FilterOverlay(2, 2, 3, 3, 2)     '* Erode, split objects
    ResultCode = bacterbacterIWSA3.FilterOverlay(5, 3, 4, 1, 2)     '* Dilate everything
    ResultCode = bacterbacterIWSA3.FilterOverlay(6, 4, 5, 2, 2)     '* Dilate, don't merge
    ResultCode = bacterbacterIWSA3.FilterOverlay(10, 5, 5, 1, 2)    '* Remove edge objects
    ResultCode = bacterbacterIWSA3.AndOverlays(5, 1, 5)             '* Logical AND overlays 1 and 5
    '* Zoom the image an hid overlays for easier viewing
    ResultCode = bacterbacterIWSA3.HideOverlay(1)
    ResultCode = bacterbacterIWSA3.HideOverlay(2)
    ResultCode = bacterbacterIWSA3.HideOverlay(3)
    ResultCode = bacterbacterIWSA3.HideOverlay(4)
    '* Measure the objects
            '* Count the objects
            App.CollectMeasurement(M_NUMOBJECTS, "I")
    ResultCode = bacterbacterIWSA3.CountObjects(5)
    NumItems = Worksheet.GetCellValue("I",1)
            '* For each object find its area and shape factor
            App.DoNotCollectMeasurement(M_NUMOBJECTS)
            App.CollectMeasurement(11,  "A")
            App.CollectMeasurement(33, "B")
    ResultCode = bacterbacterIWSA3.MeasureObjects(5)

   ' File Ends#############################################################################
            '* Turn off all measurements
            For i=0 To NUMMEASURES-1
                    App.DoNotCollectMeasurement(i)
            Next i
            App.DrawMajMinAxes=False
            '* Load the image
    ExeDirectory = App.GetExeFileDirectory()
    Dim bacterIWSA4 As Object
    '* INPUT FILE BELOW
    Set bacterIWSA4 = App.OpenImage("c:\research\phd\experiments\iws photos\Ian's Camera\IWS-A4.jpg")
    ResultCode = bacterIWSA4.Show
    ResultCode = bacterIWSA4.MakePermanent
    ResultCode = bacterIWSA4.AdjustContrast(100)
    '* Threshold the image into overlay 1
    '*Need to convert to 8 bit first?
    ResultCode = bacterIWSA4.IntensityThreshold(1, 1, Left0, Right1)
    ResultCode = bacterIWSA4.IntensityThreshold(2, 1, Left2, Right3)
    '* Run binary (overlay) filters on the objects
    ResultCode = bacterIWSA4.FilterOverlay(2, 2, 3, 3, 2)           '* Erode, split objects
    ResultCode = bacterIWSA4.FilterOverlay(5, 3, 4, 1, 2)           '* Dilate everything
```

```
    ResultCode = bacterIWSA4.FilterOverlay(6, 4, 5, 2, 2)            '* Dilate, don't merge
    ResultCode = bacterIWSA4.FilterOverlay(10, 5, 5, 1, 2)           '* Remove edge objects
    ResultCode = bacterIWSA4.AndOverlays(5, 1, 5)     '* Logical AND overlays 1 and 5
    '* Zoom the image an hid overlays for easier viewing
    ResultCode = bacterIWSA4.HideOverlay(1)
    ResultCode = bacterIWSA4.HideOverlay(2)
    ResultCode = bacterIWSA4.HideOverlay(3)
    ResultCode = bacterIWSA4.HideOverlay(4)
    '* Measure the objects
            '* Count the objects
            App.CollectMeasurement(M_NUMOBJECTS, "I")
    ResultCode = bacterIWSA4.CountObjects(5)
    NumItems = Worksheet.GetCellValue("I",1)
            '* For each object find its area and shape factor
            App.DoNotCollectMeasurement(M_NUMOBJECTS)
            App.CollectMeasurement(11,  "A")
            App.CollectMeasurement(33, "B")
ResultCode = bacterIWSA4.MeasureObjects(5)

  ' File Ends##########################################################################
            '* Turn off all measurements
            For i=0 To NUMMEASURES-1
                    App.DoNotCollectMeasurement(i)
            Next i
            App.DrawMajMinAxes=False
            '* Load the image
    ExeDirectory = App.GetExeFileDirectory()
    Dim bacterIWSA5 As Object
    '* INPUT FILE BELOW
    Set bacterIWSA5 = App.OpenImage("c:\research\phd\experiments\iws photos\Ian's Camera\IWS-A5.jpg")
    ResultCode = bacterIWSA5.Show
    ResultCode = bacterIWSA5.MakePermanent
    ResultCode = bacterIWSA5.AdjustContrast(100)
    '* Threshold the image into overlay 1
    '*Need to convert to 8 bit first?
    ResultCode = bacterIWSA5.IntensityThreshold(1, 1, Left0, Right1)
    ResultCode = bacterIWSA5.IntensityThreshold(2, 1, Left2, Right3)
    '* Run binary (overlay) filters on the objects
    ResultCode = bacterIWSA5.FilterOverlay(2, 2, 3, 3, 2)            '* Erode, split objects
    ResultCode = bacterIWSA5.FilterOverlay(5, 3, 4, 1, 2)            '* Dilate everything
    ResultCode = bacterIWSA5.FilterOverlay(6, 4, 5, 2, 2)            '* Dilate, don't merge
    ResultCode = bacterIWSA5.FilterOverlay(10, 5, 5, 1, 2)           '* Remove edge objects
    ResultCode = bacterIWSA5.AndOverlays(5, 1, 5)     '* Logical AND overlays 1 and 5
    '* Zoom the image an hid overlays for easier viewing
    ResultCode = bacterIWSA5.HideOverlay(1)
    ResultCode = bacterIWSA5.HideOverlay(2)
    ResultCode = bacterIWSA5.HideOverlay(3)
    ResultCode = bacterIWSA5.HideOverlay(4)
    '* Measure the objects
            '* Count the objects
            App.CollectMeasurement(M_NUMOBJECTS, "I")
    ResultCode = bacterIWSA5.CountObjects(5)
    NumItems = Worksheet.GetCellValue("I",1)
            '* For each object find its area and shape factor
            App.DoNotCollectMeasurement(M_NUMOBJECTS)
            App.CollectMeasurement(11,  "A")
            App.CollectMeasurement(33, "B")
            ResultCode = bacterIWSA5.MeasureObjects(5)
            MsgBox("Operation Performed.  Click OK to continue.")
End Sub
```

9.2.3.1 VARIATION IN COLONY SIZE AND WRINKLINESS

Table 9-15: Tukey's pair-wise comparison for Area.

Level											Mean
Y	A										1739658.6
W	A										1637989.4
C		B									1329033.0
E		B	C								1239669.8
Del		B	C								1192181.6
A		B	C	D							1116852.0
L			C	D							1064474.8
U				D	E						920504.0
G					E	F					835671.4
O					E	F	G				810962.2
N					E	F	G	H			741815.4
J					E	F	G	H	I		732857.4
S						F	G	H	I	J	667307.6
R						F	G	H	I	J	645608.2
V						F	G	H	I	J	643281.4
H							G	H	I	J	K 609464.2
X							G	H	I	J	K 606157.6
K								H	I	J	K 556677.2
I								H	I	J	K 552139.0
D								H	I	J	K 534054.4
F								H	I	J	K 528024.4
P									I	J	K 513207.6
Q										J	K 501218.8
B										J	K 493762.6
Z											K 418777.0
M											K 413327.2
T											K 394716.6

Genotypes not connected by same letter are significantly different ($\alpha < 0.05$).

Table 9-16: Tukey's pair-wise comparison for Circularity.

IWS Genotype							Mean
A	A						1.8260093
H	A	B					1.6610697
S	A	B	C				1.6099301
X	A	B	C	D			1.5711359
R		B	C	D	E		1.5584323
D		B	C	D	E	F	1.5169218
G		B	C	D	E	F	1.5027036
Del		B	C	D	E	F	1.5026842
L		B	C	D	E	F	1.4899302
I		B	C	D	E	F	1.4338764
F		B	C	D	E	F	1.4269487
O		B	C	D	E	F	1.4268176
K		B	C	D	E	F	1.4251355
W		B	C	D	E	F	1.4231408
E		B	C	D	E	F	1.4063337
P		B	C	D	E	F	1.4058786
V		B	C	D	E	F	1.4053047
C		B	C	D	E	F	1.4023866
B		B	C	D	E	F	1.3946185
T			C	D	E	F	1.3871968
M			C	D	E	F	1.3769041

IWS Genotype						Mean
U			C	D	E	F	1.3711381
J			C	D	E	F	1.3662559
N				D	E	F	1.3414788
Q				D	E	F	1.3381949
Z					E	F	1.2966669
Y						F	1.2753525

Genotypes not connected by same letter are significantly different ($\alpha < 0.05$).

9.2.3.2 Variation in Mat Strength

Table 9-17: Tukey's pair-wise comparison for Mat Strength.

IWS Genotype											Mean
F	A										110.00000
Del	A	B									105.00000
J	A	B	C								102.50000
O	A	B	C								101.25000
B	A	B	C	D							99.37500
A	A	B	C	D							95.62500
W	A	B	C	D							95.62500
L		B	C	D	E						94.37500
E		B	C	D	E	F					93.12500
C		B	C	D	E	F	G				91.25000
U			C	D	E	F	G				90.00000
N			C	D	E	F	G				90.00000
G			C	D	E	F	G				88.75000
Y				D	E	F	G				86.25000
P					E	F	G	H			80.62500
M					E	F	G	H			80.62500
K					E	F	G	H			80.00000
R						F	G	H	I		78.75000
Q							G	H	I		76.87500
I								H	I	J	71.25000
X								H	I	J	71.25000
T								H	I	J	70.62500
S								H	I	J	K	69.37500
H								H	I	J	K	67.50000
D									I	J	K	64.37500
V										J	K	61.87500
Z											K	55.00000

Genotypes not connected by same letter are significantly different.

9.2.3.3 Variation in Congo Red Binding

Table 9-18: Tukey's pair-wise comparison for Congo Red binding.

IWS Genotype					Mean
N	A				0.36321000
L	A	B			0.35327586
T	A	B			0.35291000
Z	A	B			0.35092000
R	A	B			0.35064000
Q	A	B	C		0.34782000
O	A	B	C	D	0.34243000
S	A	B	C	D	0.34126000
J	A	B	C	D	0.33872414

IWS Genotype								Mean
W	A	B	C	D				0.33775000
Y	A	B	C	D	E			0.33701000
Del	A	B	C	D	E			0.33465000
U		B	C	D	E	F		0.32723000
I		B	C	D	E	F		0.32651724
X		B	C	D	E	F	G	0.32537000
H		B	C	D	E	F	G	0.32382759
P		B	C	D	E	F	G	0.32197000
D			C	D	E	F	G	0.31465517
F				D	E	F	G	0.31265517
K				D	E	F	G	0.31082759
G				D	E	F	G	0.30962069
E				D	E	F	G	0.30886207
A				D	E	F	G	0.30824138
B					E	F	G	0.30296552
C						F	G	0.29444828
V							G	0.29139000
M							G	0.29100000

Genotypes not connected by same letter are significantly different ($\alpha < 0.05$).

9.2.3.4 VARIATION IN FITNESS

Table 9-19: Tukey's pair-wise comparison for Fitness in a shaking environment.

IWS Genotype												Mean
Z	A											0.94992625
B	A	B										0.93801000
P	A	B	C									0.93408875
Q	A	B	C	D								0.90824250
F	A	B	C	D	E							0.89804250
K		B	C	D	E	F						0.88487750
D			C	D	E	F						0.87435250
X				D	E	F						0.87029250
H				D	E	F						0.86687375
N				D	E	F	G					0.84967000
J					E	F	G	H				0.84500750
I					E	F	G	H				0.84473500
V					E	F	G	H				0.84120625
T					E	F	G	H				0.84012500
R						F	G	H				0.83595500
O						F	G	H	I			0.82874625
U						F	G	H	I			0.82813750
S							G	H	I	J		0.80640000
L							G	H	I	J	K	0.79760750
M							G	H	I	J	K	L 0.79011375
G								H	I	J	K	L 0.78672750
Del								H	I	J	K	L 0.78552000
Y									I	J	K	L 0.77270625
A										J	K	L 0.76726750
W											K	L 0.74133125
E											K	L 0.74056250
C												L 0.73161750

Genotypes not connected by same letter are significantly different ($\alpha < 0.05$).

Table 9-20: Tukey's pair-wise comparison for fitness in a static environment.

IWS Genotype									Mean
B	A								1.0927875
F	A								1.0916112
I	A	B							1.0503375
Y	A	B							1.0417025
R		B	C						0.9913050
N		B	C	D					0.9874763
P		B	C	D					0.9868950
T		B	C	D					0.9861263
Q			C	D	E				0.9664738
O			C	D	E				0.9646462
W			C	D	E				0.9625100
L			C	D	E				0.9522363
A			C	D	E				0.9364075
J			C	D	E				0.9360237
S			C	D	E				0.9347075
D			C	D	E				0.9330625
C			C	D	E	F			0.9262487
H			C	D	E	F	G		0.9255388
U				D	E	F	G		0.9200925
M				D	E	F	G		0.9198713
X					E	F	G		0.9119762
E					E	F	G		0.8984575
Del						F	G	H	0.8596450
Z							G	H	0.8560850
K								H	I	0.8026500
V									I	0.7596025
G									I	0.7364887

Genotypes not connected by same letter are significantly different ($\alpha < 0.05$).

9.2.4 CHAPTER 2.2.3.2 MULTIVARIATE ANALYSES OF SIMILARITY

This R code is used to determine a dissimilarity matrix based on Euclidean distance followed by an ANOSIM randomization test with 1000 permutations.

```
#Make sure to load 'Vegan' package using the package manager before you try and run this

data<-read.table(file="/Users/Springbok26/R_work/ANOSIM_data.tab", header=T)

attach(data)

library(vegan)
diss<-dist(data,method = "euclidean", diag = FALSE, upper = FALSE, p = 2)
summary(diss)

data.ano<-anosim(diss,Gene,permutations=1000)
summary(data.ano)
#Need to run this last command from main window. Make sure to have an Quartz window open (press icon third from right in
main R window)
plot(data.ano)
```

9.3 APPENDIX ITEMS FROM CHAPTER 3

No items.

9.4 APPENDIX ITEMS FROM CHAPTER 4

9.4.1 CHAPTER 4.3.2 INADEQUACIES ABOUT CURRENT COOPERATION THEORY FOR MICROBES

9.4.1.1 TRACING THE EVIDENCE FOR SIDEROPHORE EXCRETION AS A COOPERATIVE TRAIT

Secretion of extracellular compounds has become a popular system in which to study social phenomena in microbes (Shapiro and Dworkin, 1997; Crespi, 2001; West and Buckling, 2003; Parsek and Greenberg, 2005; West *et al.*, 2006; Wingreen and Levin, 2006). The canonical example of extracellular secretion as cooperation is the production of iron-scavenging proteins called siderophores (West and Buckling, 2003; Griffin *et al.*, 2004; Harrison and Buckling, 2005; Harrison *et al.*, 2006; West *et al.*, 2006; Buckling *et al.*, 2007; Ross-Gillespie *et al.*, 2007; Harrison *et al.*, 2008)[1]. Siderophores, which have a high affinity for iron, are excreted from the bacterial cell and chelate insoluble iron to form iron-siderophore complexes that can subsequently be taken up by the bacteria (Guerinot, 1994; Ratledge and Dover, 2000). This paradigm has been asserted as cooperative in numerous publications (Velicer, 2003; West and Buckling, 2003; Griffin *et al.*, 2004; Harrison and Buckling, 2005; Harrison *et al.*, 2006; West *et al.*, 2006; Buckling *et al.*, 2007; Ross-Gillespie *et al.*, 2007; Harrison *et al.*, 2008), however, tracing the empirical evidence for such claims has proved challenging.

In the most recent paper, siderophore excretion is said to provide "a useful model system for testing theories of the evolution of cooperation" (Harrison *et al.*, 2008 p.49). Following this, the authors claim that "the production of iron-scavenging

[1] These are all known references for the search terms 'cooperation' and 'siderophore' at Web of Science plus the review on social evolution theory for microorganisms by West *et al.* (2006).

siderophores by bacteria and fungi is a classic example of cooperation via production of a public good" and provide evidence from three citations (Ratledge and Dover, 2000; Griffin *et al.*, 2004; Wandersman and Delepelaire, 2004). Ratledge and Dover (2000) and Wandersman and Delepelaire (2004) are both review articles on iron transport by siderophores, and do not comment about the cooperative nature of siderophore excretion, as evidenced by the omission of the words altruist, altruism, cooperation, cooperate, cooperative or public good in either article. Retracing to Griffin *et al.* (2004 p.1025), these authors claim that "siderophore production is an altruistic cooperative trait that is costly for the individual, but provides a local (group) benefit" and provide evidence from three citations (Guerinot, 1994; Ratledge and Dover, 2000; West and Buckling, 2003), and their own methods. The methods of Griffin *et al.* (2004) shall be dealt with last, after tracing the remaining claims to their origins. Again the Ratledge and Dover (2000) reference appears with an older review on iron transport from Guerinot (2000), but again, the words altruist, altruism, cooperation, cooperate, cooperative or public good are omitted from the entire article. Retracing further, West and Buckling (2003 p.37) make the following claim:

> *A crucial feature of siderophores is that they **potentially** benefit all bacteria within the locality. **Assuming** siderophore production is metabolically costly, there is a clear **potential** for the evolution of selfish bacteria that do not produce siderophores, but are capable of their uptake once iron has been bound.*[1]

Crucially, no is evidence supplied to substantiate the *assumptions* of *potential* benefits; however, they cite another article (De Vos *et al.*, 2001) as the evidence for the evolution of non-producers in the lungs of patients affected by cystic fibrosis (CF). Unfortunately, this observation is not a validation of siderophore secretion as cooperation, because siderophore non-producers could be adaptive mutants that arose in response to self-interested individual selection as either cross-feeders of an equally available resource or because production is not required in the lung environment. In addition, this study does not suggest siderophore secretion is an example of cooperation, because it does not include any statements contain the words altruist, altruism, cooperation, cooperate, cooperative or public good in the entire article. If

[1] Boldface emphasis added.

non-producers are cheats rather than individually adaptive mutants, non-producers are expected to arise in the laboratory under controlled conditions, however, no data to this effect has ever been published and all non-producing mutants in studies of this nature are created using molecular biology techniques. Therefore, in retracing the evidence for siderophore excretion as cooperation all avenues so far have been *cul de sacs* and the only avenue remaining for claims lies in the methods of Griffin *et al.* (2004).

The methods of Griffin *et al.* (2004 p.1027) claim to provide evidence that "Siderophore (pyoverdin) production in these strains has the characteristics that make an appropriate cooperative trait for our study". However, inspection of the methods reveals three major flaws. First, both the cost to cooperation in an iron-replete environment (CAA media) and the benefit of cooperation in an iron-deplete environment (CAA + apo-transferrin) were measured by comparison of *monocultures* of the producing strain and the non-producing strain. For the cost to cooperation it was not stated what was compared between the two strains, while *final density* was used to illustrate the benefit to cooperation. These are not the correct experiments, because it is obligatory to demonstrate these costs and benefits of production in *competition* by comparing *Malthusian parameters* to obtain a measure of relative fitness (Lenski *et al.*, 1991). Secondly, the mutant strain that was suggested to cheat on the siderophore production of the wild-type was derived using UV mutagenesis, and remains uncharacterised. There was no measurable difference in fitness for the mutant strain compared to the wild-type in KB media, which is either iron replete or iron deplete[1]; therefore, there should be either a cost or a benefit to cooperation. If different media with the same limiting conditions cannot produce the same result, then this raises serious concerns that the observation may be a media specific effect. Thirdly, siderophore secretion is a tightly regulated trait that is switched off by the bacteria when sufficient iron is available (Guerinot, 1994; Ratledge and Dover, 2000); therefore, if a cost is demonstrated under 'high-iron conditions', the

[1] Cultures of *P. fluorescens* in KB media supplemented with $FeCl_3$ do not fluoresce (*i.e.* do not produce siderophores), while they fluoresce strongly in standard KB media. While not conclusive, this suggests that KB media is iron deplete.

biologically relevant conditions may, in fact, have been violated. In conclusion, the current empirical evidence for siderophore secretion as a cooperative behaviour is at best circumstantial, and there remains a need for rigorous experimentation.

9.5 Appendix Items from Chapter 5

No Items.

9.6 Appendix Items from Chapter 6

No items.