# The role of chance and history in the evolution of a novel trait in *Pseudomonas fluorescens* SBW25

A thesis submitted in partial fulfilment of the requirements for the
degree of
Doctor of Philosophy
in
Evolutionary Genetics

At Massey University, Albany Campus, New Zealand

Sylke Nestmann

2013

# Abstract

The processes that lead to the evolution of novel adaptive traits are poorly understood. It is not clear how the combination of natural selection and random variation determine evolutionary pathways, specifically how evolutionary history affects the outcome of evolution. In a previous experiment in which populations of the bacterium *Pseudomonas fluorescens* were subjected to repeated evolutionary reversals (the *R*everse-*E*volution *E*xperiment (REE)) stochastically switching genotypes evolved *de novo* in two out of 12 replicate lineages. Whole genome sequencing revealed not only causative mutations in *carB* (referred to as *carB**) and *rpoD* but also the entire evolutionary pathways comprised of eight additional mutations. It was hypothesized that evolutionary history played a major role in switcher evolution.

This was tested by ‘reviving’ four REE lineages and the performance of ‘replay evolution’ experiments. Whereas the repeated evolution of switcher genotypes was observed in three of the four lineages, the likelihood of switcher occurrence varied and depended on the genotype. By artificially introducing the original switcher mutation *carB** into genotypes from one lineage, potential fitness benefits of the switcher mutation at each point in evolutionary time were assessed. The introduction of *carB** into the ancestral genotype of SBW25 created a switcher with higher fitness than the ancestor, indicating that evolutionary history is not necessarily required to give rise to an evolutionary successful switcher. This idea was tested further during a real-time evolution experiment using the same genotypes as founder populations. The capacity to evolve a switcher, based on the competitive fitness of *carB**, was only partly reflected in the ability to re-evolve switchers, which is most likely due to the availability of alternative evolutionary pathways, the number of competitors, and the fitness of the founder genotype. In addition, epistasis contributed to an increased capacity to produce a switcher based on *carB**, resulting in an increased likelihood to evolve this novel trait in genotypes from the later evolutionary time points.

Evolutionary history plays an important role in the evolution of switching. Even though switchers evolved repeatedly from different genotypes, the likelihood that this path was taken depended on genetic constraints and ecological factors. These factors and their contribution to the evolutionary outcome were dependent on the genetic composition and changed throughout evolutionary history.

# Acknowledgements

This research project would not have been possible without the support of many people. First and foremost, I wish to thank my supervisor, Professor Paul Rainey for his support, his continuous faith and help throughout my project. I also want to thank my co-supervisor Dr. Austen Ganley for encouragement and discussions during the research. I would like to thank the New Zealand Institute for Advanced Study for funding the project and for the financial support to attend overseas conferences.

I would like to acknowledge former Rainey lab members in particular Dr. Hubertus Beaumont for his extensive work on the Reverse-Evolution Experiment and Dr. Jenna Gallie for her studies on the switcher lineages. Dr. Monica Gerth, thank you for teaching me many methods and tricks in the laboratory. Thanks to Dr. Gayle Ferguson for helping with the material and methods chapter of my thesis. My special thanks are extended to all the Rainey lab members for their help and kindness and for being so weird and entertaining at the same time. Thanks to Katrina Rainey for improving my thesis and for her valuable comments.

Special thanks goes to Dr. Katrin Hammerschmidt for her treasured assistance with statistic problems and for her help with the planning of experiments. You have been a great source of inspiration, hope and motivation.

I am grateful to my parents who supported me throughout the entire time and encouraged me to start this project and to move to New Zealand. Without my parents my PhD would not have been possible. Thanks to my flatmates, Jazz, Stef, Nielvin, Andrew and Sat for being my New Zealand family, for the *flinners* and for cheering me up. Finally, I would like to thank Katrin and Peter for being great friends. Fishing, extensive cooking sessions, delicious food and *Tatort* nights are things that I really enjoy in life and I am glad that I could share this with the two of you. Thanks for looking after me.

# Table of contents

# List of illustrations

# List of tables

# Chapter 1: Introduction

## 11.1 Evolution

Biologists have always been fascinated by the existence of complex life forms and the extraordinary biodiversity that we observe amongst organisms in our environment. Biodiversity is the result of evolution. Evolution can be described as the change in heritable characteristics in populations of organisms over generations in response to the environment. More than 150 years ago Charles Robert Darwin (1809 – 1882) and Alfred Russel Wallace (1823 – 1913) formulated their revolutionary idea about evolution. They were not only able to relate species by providing a theory of a common ancestor but they also delivered an explanation for the process behind it – natural selection (Darwin & Wallace, 1858; Darwin, 1859). During the last century and a half the idea of evolution by natural selection has become more and more concrete. Today, there is no doubt that natural selection is a powerful force driving evolution. Evolution by natural selection can only take place if certain conditions within a population are fulfilled: *replication*, *heredity* and *variation*, with variation affecting reproductive success (Lewontin, 1970; Dennett, 1995). There must be phenotypic variation amongst individuals within a population (morphology, physiology, and behaviour). This phenotypic variation must have fitness consequences causing some individuals to reproduce more often and others to reproduce less. Furthermore this phenotypic variation must be heritable. Whenever these three conditions are fulfilled adaptive evolution by natural selection will occur and transform a population over time.

### 1.1.1 Origin of genetic variation

Evolutionary relevant phenotypic differences are based on genetic variation. Mutation is the ultimate source of genetic variation within populations. Mutations are randomly occurring changes in DNA sequences, which can encompass only a few nucleotides or large sections of the genome such as entire genes. They are the product of random errors during DNA replication and repair. Mutations can be classified according to their fitness effect on its carrier as: lethal, deleterious, neutral or beneficial. Lethal mutations lead to the death of the organism. A variant carrying a deleterious mutation has a reduced fitness, which decreases the probability of reproduction, and thus contribution to the next generation. Mutation accumulation experiments provide evidence that slightly deleterious mutations are the most common type of mutations (Crow & Simmons, 1983; Kibota & Lynch, 1996; Maisnier-Patin *et al.*, 2005). The evolution of novelties depends on the supply of beneficial mutations. If the fitness is elevated by a beneficial mutation then the variant is consequently favoured by natural selection. The variant has an increased likelihood of reproducing and appearing with a high frequency in the next generation. Furthermore, mutations can occur without having an immediate effect on fitness and are therefore neutral. The fitness effect of a mutation is context dependent. For example a beneficial mutation in one environment can be neutral or deleterious in another environment. In addition, the genetic background in which the mutation occurs can have a large impact on the fitness of the variant due to interactions between alleles (epistasis).

Sexual reproduction and recombination (meiosis, chromosomal cross-over) as well as horizontal gene transfer are important processes that redistribute genetic variation between lineages. It is still an open question why sex and recombination evolved in the first place as both mechanisms are considered to be risky and costly for the organism (Otto & Lenormand, 2002). One explanation that has been suggested is that the shuffling of mutations via sex and recombination is a rather quick mechanism that can group beneficial mutations within an organism or remove deleterious mutations from populations without dragging along beneficial mutations (Charlesworth, 2003; Wayne & Miyamoto, 2006).

### 1.1.2 Natural selection

Natural selection changes the phenotypic distribution within a population as a result of differences in reproduction. Genetic variation (mutation) gives rise to variants within a population that show new phenotypic characteristics, which are inheritable. If a new variant performs even slightly better under given environmental conditions, it will produce more offspring in comparison to alternative types and will contribute with higher frequency to the next generation. The new variation thus will spread within the population over time. The ancestor will become less common and perhaps even go extinct. Evolution by natural selection takes place (reviewed in Stearns & Hoekstra, 2005: p.1-4).

### 1.1.3 The impact of stochastic processes on evolution

Many factors that influence evolution are stochastic such as random mutations, genetic drift, environmental change, and migration (Gillespie, 2006). Phenotypic variation within populations is inherently necessary for natural selection to become effective (Lewontin, 1970; Dennett, 1995) and is fuelled by the occurrence of random mutations (see section 1.1.1). When and where a mutation arises in the genome is the result of chance and is therefore unpredictable. Genetic drift is a further element of stochasticity that has an impact on evolution. Drift can be described as a fluctuation of allele frequencies from generation to generation caused by the random distribution of alleles. Here the compositional change of a population is based just on a stochastic process. The effect of random drift will be strongest in small populations. Under these circumstances mutations with a small beneficial effect on fitness are likely to go extinct due to the strong effect of random genetic drift. In small populations, genetic drift can lead in turn to the fixation of a slightly deleterious mutation within a population. Consequently genetic drift prevents populations from adapting to an environment to the degree that we would expect if selection were the only factor (Lynch, 2007; Bell, 2008). Drift becomes for example important after migration events when only view individuals of an original population migrate to distant places and found a new population. In addition, populations have to cope with ever-changing environments. Environmental change happens over time and throughout space and when and in which direction the surrounding environment will change are not predictable.

## 1.2 The role of adaptation, chance, and history in evolution

Since Darwin's time scientists have been fascinated by the question of to what extent random factors (mutations and genetic drift) and (non-random) selection contribute to evolutionary change. Adaptations are novel traits that make new ecological niches accessible and implicate a change in reproductive success (Pigliucci, 2008; Blount *et al.*, 2012). Decades ago the idea that adaptation was solely responsible for evolution of novel traits was popular and is not uncommon today (Adaptationism). Proponents of this view believe that novel structures arise as a consequence of natural selection only, resulting in organisms optimally adapting to their respective environments (Dawkins, 1976; Gould & Lewontin, 1979; Dawkins, 1986; Beatty & Desjardins, 2009). Strong evidence supporting the idea that selection is the only force driving evolution comes from parallel and convergent evolution (Dawkins, 1986; Dennett, 1995; Conway Morris, 2003b; Vermeij, 2006; Conway Morris, 2010). Here closely related (parallel evolution) or unrelated lineages (convergent evolution) evolve similar solutions to the same adaptive problem. Eyes evolved multiple times in distantly related phyla and represent a strong example of convergent evolution (Lamb *et al.*, 2007; Lamb *et al.*, 2009). Parallel evolution has been observed in natural populations (Losos *et al.*, 1998; Allender *et al.*, 2003; Boughman *et al.*, 2005) and during laboratory experiments (Wichman *et al.*, 1999; Woods *et al.*, 2006; Bantinaki *et al.*, 2007; McDonald *et al.*, 2009). These observations of convergent and parallel evolution suggest that evolution might follow very similar evolutionary routes due to a strong impact of natural selection and is therefore rather predictable. Thus the role of chance is considered to be negligible (Vermeij, 2006; Conway Morris, 2009; Conway Morris, 2010).

Darwin recognised that random factors could have an impact on evolutionary pathways. In his work on orchids he emphasised that the starting position of evolution is set by chance and can lead to multiple solutions through different evolutionary pathways in response to the same adaptive challenge (Darwin, 1882). The different outcomes cannot be explained by selection only. Random occurrences of variation in lineages account for different directions of evolutionary pathways and different results (Beatty, 2006). The belief that natural selection is the only force driving evolution was questioned by a new generation of scientists several decades ago (Monod & Wainhouse, 1971; Jacob, 1977; Gould & Lewontin, 1979; Jablonski, 1986; Gould, 1989; Mani & Clarke, 1990; Travisano *et al.*, 1995a). The classic paper from Gould and Lewontin (1979) presents a landmark publication in which they criticise the view of the Adaptationists. They argue that the outcome of evolution is unpredictable because of the impact of stochastic processes. Randomness can justify, for example, innovations that are not logically the best possible adaptation. Instead evolutionary 'accidents' and their consequences for evolutionary pathways could give validation to the existence of novel body structures that are rather unpredictable (Gould & Lewontin, 1979; Gould, 1987). The effects of chance are usually the result of random mutations without any value to the organism and to genetic drift (Crow & Kimura, 1970; Kimura, 1983; Spiess, 1989; Suzuki *et al.*, 1989; Lynch, 2007). Consequently alleles without adaptive significance, that neither improve nor decrease adaptation, are maintained in populations over generations. Travisano *et al.* (1995) point out that chance can be important for phenotypic evolution because beneficial mutations can be lost even in large populations due to stochastic processes (Travisano *et al.*, 1995a).

Every organism that we observe today is the outcome of evolution, which is built upon evolutionary history. Evolutionary history is the entirety of evolutionary events that occurred along the evolutionary path of an organism. Such evolutionary events are influenced by random variation, genetic drift, selection and the surrounding environments. These factors are tightly linked and change over time. Consequently evolution follows certain evolutionary pathways shown in the evolutionary histories of organisms. Many innovations that occurred throughout the history of life are assumed to be solitary events without parallel occurrences anywhere else or at another time. This illustrates the importance of unique evolutionary pathways with a high degree of dependency on initial circumstances (Gould & Lewontin, 1979; Gould, 1989; Gould, 2002). Evolutionary history can become important if certain genetic changes with adaptive significance in the past constrain or promote certain evolutionary outcomes in comparison to other outcomes (Gould & Lewontin, 1979) In such cases evolutionary endpoints are highly dependent on prior (historical) events (Travisano *et al.*, 1995a). Stephen J. Gould was one of the pioneers to evaluate the effect of adaptation, chance and history on evolutionary change.

### 1.2.1 Gould's view of the world's history

Stephen Jay Gould (1941-2002) was a theorist and palaeontologist with a great interest in evolution. He became fascinated by the unique Cambrian fossil discovery in the Burgess Shales in 1909 and in his famous book *Wonderful Life: the Burgess Shales and the Nature of History* (1989) he discusses the possibility of a significant impact of random factors and evolutionary history in evolution. There is no doubt that selection is a powerful driver of evolution (Gould & Lewontin, 1979) but he suggested that randomness can change the foundation for selection to act upon, and thus would lead to different outcomes. Evolution in that case would be unpredictable. He introduced the concept of historical contingency. Here the evolutionary pathway that an organism can take is highly dependent on historical events, which often occur randomly and the evolutionary result is an anatomically imperfect solution. His famous example is described in *The Panda's Thumb* (1987). The panda has evolved a helpful bamboo-stripping 'thumb', which looks like a sixth digit but actually has evolved by the extension of a wrist bone (Gould, 1987). Evolution can follow very different pathways that lead to multiple outcomes. Which path a population takes depends on initial conditions and is rather unpredictable.

Gould points out that constraints play an important role in evolution and have to be taken into account. Constraints can restrict populations in their ability to be best adapted to the environment they live in. Gould describes constraints as *"internal factors"* that *"restrict the freedom of natural selection to establish and control the direction of evolutionary change"* (Gould, 2002: p. 1028). He mentions that constraints, in addition, can offer new opportunities for diversification and can provide the foundation for the evolution in new directions (Gould, 2002). Although the term 'constraint' is still vague in the context of selective processes, many evolutionary biologists have agreed that constraints based on genetic backgrounds, developmental constraints (epigenetic) and environmental constraints exist and have to be taken into account.

#### *1.2.1.1 Replaying the tape of life*

To disentangle the effects of adaptation, chance and history, Gould offered his much-discussed *Gedankenexperiment* in which he tried to envision what our world would be like if we "*replayed life's tape*". He asked: If it was possible to replay the history of life from any point in time in the past would the outcome be very similar or would different paths lead to a different outcome?

> *"I call this experiment 'replaying life's tape'. You press the rewind button and, making sure you thoroughly erase everything that actually happened, go back to any time and place in the past... Then let the tape run again and see if the repetition looks at all like the original."* (Gould, 1989: p.48)

Based on his assumption that random factors and constraints cause the history of life to be contingent on past events, any *"replay of life's tape"* would follow different evolutionary routes leading to a different world from the one we see today (Gould, 1989). It is unlikely that the evolution of *Homo sapiens* would happen again. From his point of view evolution is like a lottery where starting positions are mostly set by chance (Gould, 1989; Pal *et al.*, 2006) and subsequent evolutionary events are determined by the starting position (Gould & Lewontin, 1979). Gould did not dispute the occurrence of convergence but he argued that convergent lineages are usually not independent from each other and share a huge amount of history, for example a common ancestor, therefore according to Gould, even in the face of convergence, evolution is still unlikely to be repeatable. For example there is an ongoing debate as to whether the evolution of eyes in different phyla is a true case of convergence. Former studies claimed that eyes have evolved from a monophyletic origin based on homologous *Pax* genes and are therefore the result of parallel evolution (Quiring *et al.*, 1994; Halder *et al.*, 1995a; Halder *et al.*, 1995b) but more recent studies suggest that eyes are the outcome of true convergent evolution (Kozmik, 2005; Kozmik, 2008).

## 1.3 Experimental evolution with microorganisms

There is no doubt that it is impossible to rewind and replay the history of life. Studying the origin of innovations and the impacts of adaptation, chance and evolutionary history are challenging. Fossil records, which Gould was fascinated with, can be used to compare ancestral phenotypes and reconstruct evolutionary trajectories of different organisms. They can offer some information about convergence, parallelism and predictability. Nevertheless fossils are usually from small, separated areas and in many cases they represent different evolutionary time points and thus can only provide a snapshot of evolution. Little can be inferred on the origin of particular innovations, on the environmental circumstances, and on population composition at that time. In addition, ancient DNA samples are rare and often not well preserved. Mapping the genotype to the phenotype is therefore fragmented and incomplete.

Comparative studies in higher organisms are problematical. Because generation times are in general quite long it is hard to look at multiple generations and to follow evolution in action. The underlying selective mechanism is often unknown and causes of the observed evolutionary patterns can only be indirectly inferred. It is difficult to make assumptions as to why a particular evolutionary outcome occurred and what other evolutionary outcomes might have been possible.

Microorganisms offer the opportunity to follow evolution in real-time, over many generations under controlled environmental conditions in the laboratory. Experiments become replicable and repeatable, which makes the results and observations more powerful. The accessibility of genomic data, due to affordable genome sequencing techniques, provides the opportunity to compare phenotypic and genotypic trajectories. Most importantly, evolutionary stages of microorganisms can be stored in the freezer. This 'frozen fossil record' enables us to go back in evolutionary time and revive earlier stages (rewind) and actually *'replay life's tape'* with microorganisms from different starting positions.

### 1.3.1 The *Escherichia coli* long-term experiment

Richard Lenski and colleagues initiated a key experiment for evolutionary biology in 1988 - the *E. coli* long-term experiment. The aim of the experiment was to understand the roles of adaptation, chance and history in evolutionary change. Twelve *E. coli* lineages were initiated from a common ancestor and maintained in flasks containing glucose-limited and citrate-rich media (Lenski *et al.*, 1991; Lenski & Travisano, 1994). The bacterial cultures were transferred to fresh media daily. Since all 12 lineages were initiated from the same *E. coli* genotype, all adaptations that occurred during the course of selection could only be caused by the appearance of new mutations. Additionally, between-lineage differences could only be the result of novel mutations. The trait affected by evolutionary change was fitness. In accordance with their expectations they found that after 2000 generations all 12 lineages had reached similar elevated fitness endpoints, which was attributable to the effect of natural selection. It was suggested that the high degree of phenotypic parallelism reflects the fixation of similar mutations affecting fitness (Lenski *et al.*, 1991). As well as fitness they monitored bacterial cell size, which was assumed to be only weakly correlated with fitness. They found that phenotypic change in cell size was most likely the result of chance and evolutionary history (Travisano *et al.*, 1995a). Lenski and Travisano (1994) extended the experiment for a further 8000 generations. The fitness of the 12 parallel lineages after this time was still similar between the lines but surprisingly they showed some degree of diversification in their fitness endpoints. They argued that the divergence in fitness reflected underlying genetic differences due to the accumulation of different random mutations. In addition they found that the rate of adaptation declined over time and was likely to be due to both a decreased number of beneficial mutations and a decreased fitness benefit of new mutations (Lenski & Travisano, 1994). The *E. coli* experiment is still running and has now reached over 50,000 generations.

## 1.4 The model organism *Pseudomonas fluorescens* SBW25

The Gram-negative bacterium *Pseudomonas fluorescens* SBW25 was first isolated from the leaf of a sugar beet plant, University Farm, Wytham, Oxford (Rainey & Bailey, 1996). The rod-shaped bacterium can usually be found in soil and water, on plant surfaces like leaves and in the rhizosphere. *P. fluorescens* has multiple flagella and is obligate aerobe. In environments with low iron concentration the cells produce the siderophore pyoverdine. This is a chelating agent for iron and is responsible for the fluorescent colour in *P. fluorescens*.

Like other microorganisms *P. fluorescens* has the ability to adapt to a novel environment very quickly. Observation of real-time evolution under controlled conditions is possible. In addition, the physical and genetic map of *P. fluorescens* (Rainey & Bailey, 1996) and the whole genome sequence (Silby *et al.*, 2009) are available, making *P. fluorescens* an excellent model organism for studying aspects of evolutionary theory.

### 1.4.1 Diversification of *P. fluorescens* SBW25 in a static environment

When wild type *P. fluorescens* SBW25 grows on agar plates containing King's B media (KB) the observed colony morphology is round and smooth. After a colony is propagated in a microcosm containing liquid KB media and incubated undisturbed, novel types occur within 2-3 days which can be differentiated visually on KB agar plates by their different colony morphologies (Fig.1.1). The new variants appear to be niche specialists within a spatially structured environment. Competition between the different niche specialists drives diversification. The smooth looking ancestral type (SM) grows in the broth phase of the media. A distinct novel group of morphotypes are named 'wrinkly spreaders' (WS) because their colonies show wrinkled edges. Another distinct type is the 'fuzzy spreader' (FS), which has larger colonies and blurry edges. WS are well studied and it has been shown that they have the ability to form a biofilm (Fig.1.1) on the surface of the microcosm (Rainey & Travisano, 1998).

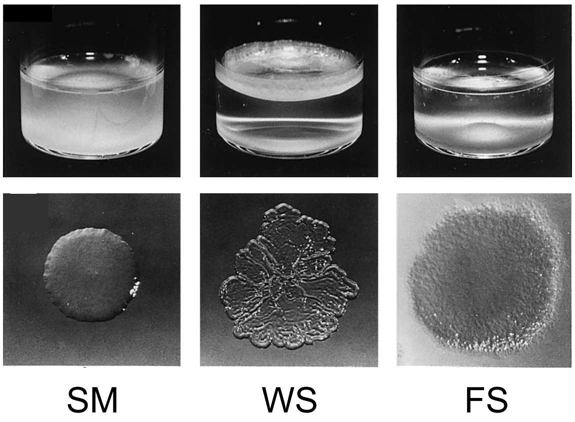

**Figure 1.1: Adaptive radiation in a static environment.** After 2-3 days growth in a static KB microcosm the ancestor *P. fluorescens* SBW25 gives rise to different genotypes with different colony morphologies on KB agar plates: smooth ancestral colonies (SM), wrinkly spreader colonies (WS) and fuzzy spreader colonies (FS) (bottom row from left to right). Each type occupies a different niche: SM grows in the broth (left upper row) and WS occupies the air-liquid interface (middle upper row). FS cells colonise the air-liquid interface but sink to the bottom of the microcosm very quickly (right upper row). Picture was incorporated from (Rainey & Travisano, 1998).

### 1.4.2 The evolutionary race of wrinkly spreaders for the air-liquid interface

Within hours after inoculation of a static KB microcosm with *P. fluorescens* SBW25 the liquid broth turns into a heterogeneous environment due to an oxygen gradient that develops from the surface, which contains the highest level of oxygen, to the bottom, containing the lowest amount of oxygen (Koza *et al.*, 2011). *P. fluorescens* requires oxygen to grow. The lack of oxygen implies increased selection pressure for bacterial cells but also leads to the availability of unoccupied niches. Any newly evolved variant that is able to occupy the air-liquid interface first will gain a benefit because this niche contains oxygen from the air and nutrients from the liquid KB media equally (Fig. 1.2). The advantage will be reflected in an increased cell number of the new variant (Rainey & Travisano, 1998; Koza *et al.*, 2011). WS are well adapted to the air-liquid interface and evolve quickly under static conditions (Fig. 1.2). This is attributed to mutations that spread in the population.

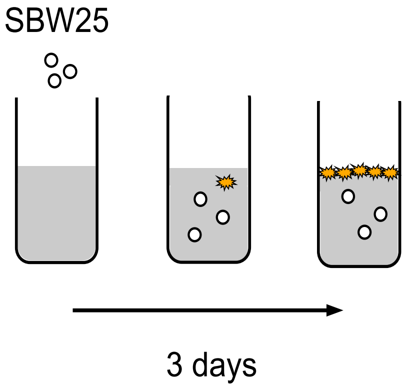

**Figure 1.2: Rise of the WS.** After inoculation of a static environment with *P. fluorescens* SBW25 (white circles) an oxygen gradient within the liquid media develops. After a short time WS appear (orange star) that are able to colonise the air-liquid interface. This novel type gains benefit by obtaining access to oxygen from the air and nutrients from KB media. The subsequent increase in frequency can be seen in biofilm formation.

Detailed analysis revealed mutations in genes that are involved in the production of an extracellular polymer – cellulose. As a consequence of the mutations WS cells start to overproduce cellulose on the cell surface, which adopts the function of glue and prevents cells from separating from each other after cell division and leads to biofilm formation on the surface of the microcosm (Rainey & Rainey, 2003; Spiers *et al.*, 2003).

The overproduction of cellulose is costly for the individual cell of a WS population and only the formation of a biofilm provides an advantage over other types that are present in the static environment. In a shaken environment, with no niche structure, oxygen and nutrients are not limiting resources. Hence, WS types have no selective advantage in such a homogeneous environment. Reduced growth of WS in shaken liquid media (Koza *et al.*, 2011) provides the opportunity for new variants to emerge. Novel types that do not overproduce cellulose can separate after cell division and will increase their cell number because they do not bear the cost of cellulose production (Fig. 1.3). On KB agar plates they show smooth and round colony morphology similar to the wild type *P. fluorescens* SBW25 (Fig. 1.1). The new SM types show mutations that reverse the effect of previous WS mutations. Reversal mutations can occur in the same gene or same operon but have been discovered in other genes as well. As a result cellulose production is shut down or reduced (Ferguson *et al*. 2013, *manuscript in preparation*).

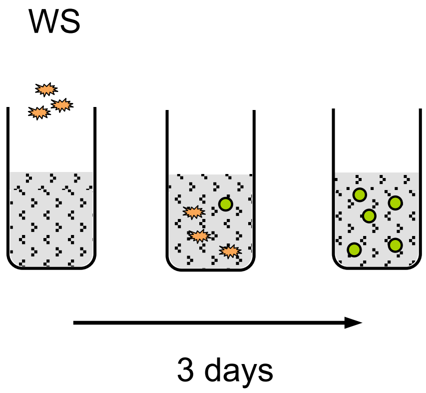

**Figure 1.3: Fall of the WS.** Growth of WS in a shaken environment gives rise to new variants that do not overproduce cellulose (green circles). As a result cells avoid sticking together after cell division. WS still bear the cost of producing cellulose but without any advantage because oxygen and nutrients are available throughout the liquid media and not restricted to a certain niche. The new type can reach higher cell numbers than WS.

### 1.4.3 The genetic origins of the WS phenotype

In the previous section the phenotypic behaviour of WS types on KB agar plates and in liquid KB media was described in detail (Fig. 1.1, Fig. 1.2, Fig. 1.3). In previous studies, suppressor analyses were used to identify the relationship between the genetic changes that gave rise to WS in a static environment and it's phenotypic appearance. Random transposon insertions were performed in WS types, which were then screened for deactivated cellulose production as well as for deficiency in biofilm formation. Here the observed phenotype on agar plates reversed from colonies with wrinkled edges to ancestor-like smooth colony morphology. As a result four key *loci* were identified that contained causal mutations underlying the WS phenotypes. Mutations occurred primarily in the *wss* and *wsp* operon (Kahn, 1998; Bantinaki, 2001; Spiers *et al.*, 2002; Spiers *et al.*, 2003) and in the loci *aws* and *mws* (Gehrig, 2005; McDonald *et al.*, 2009). In the following section each locus will be described in more detail.

#### *1.4.3.1 The* wss *operon*

The *wss* ('*w*rinkly *s*preader' *s*tructural) operon is an assembly of 10 genes (*wssA*-*wssJ*) and encompasses a region of ~15 kb in the genome (Fig. 1.4). It has been shown that the *wss* genes together build a functional unit encoding for the synthesis of an *a*cetylated *c*ellulosic *p*olymer – ACP (Spiers *et al.*, 2002).

**Figure 1.4: The *wss* operon.** It is a cluster of 10 genes (*wssA*-*wssJ*) that encodes for the production of ACP. The genes *wssA* and *wssJ* (orange) are thought to be responsible for the localisation of the operon. Together the subunits *wssBCDE* (green) are likely to build the cellulose synthase complex, whilst *wssGHI* (purple) play a role in acetylation of the cellulose polymer.

The first and the last gene of the operon (*wssA*, *wssJ*) are involved in the localisation of the *wss* operon within the cell. The following genes, *wssBCDE*, together are predicted to build a cellulose synthase complex. The genes *wssGHI* are thought to be involved in the acetylation of the cellulose polymer (Spiers *et al.*, 2002; Rainey & Rainey, 2003; Spiers *et al.*, 2003).

Biochemical analyses of the WS mats and colonies have confirmed that the *wss* operon is the determinant of the WS phenotype. The structure of the WS biofilm is based on ACP. The polymer is causal for mat formation in WS types and the colonisation of the air-liquid interface (Spiers *et al.*, 2002; Rainey & Rainey, 2003; Spiers *et al.*, 2003). The secondary messenger molecule bis-(3'-5')-cyclic dimeric guanosine monophosphate (c-di-GMP) activates ACP production. The whole regulatory mechanism is not fully understood yet but it is known that c-di-GMP is produced by diguanylate cyclases (DGC) (Ross *et al.*, 1987; Goymer *et al.*, 2006; Malone *et al.*, 2007).

### *1.4.3.2 The* wsp *operon*

The *wsp* (*w*rinkly *sp*reader) operon encodes for seven proteins that are part of a chemosensory pathway analogous to the Che chemotaxis pathway in *Escherichia coli* (Bantinaki, 2001). The seven genes *wspABCDEFR* occupy ~8.5 kb of the genome. WspA is likely to be localised in the cell membrane and belongs to the *m*ethyl-accepting *c*hemotaxis *p*roteins (MCP). WspB and WspD are thought to be scaffold proteins. WspC is predicted to be a methyltransferase and WspF a methylesterase. Both proteins regulate the activity of the kinase WspE by an opposing feedback loop, which adds methyl groups to WspA, or removes them (Fig. 1.5).

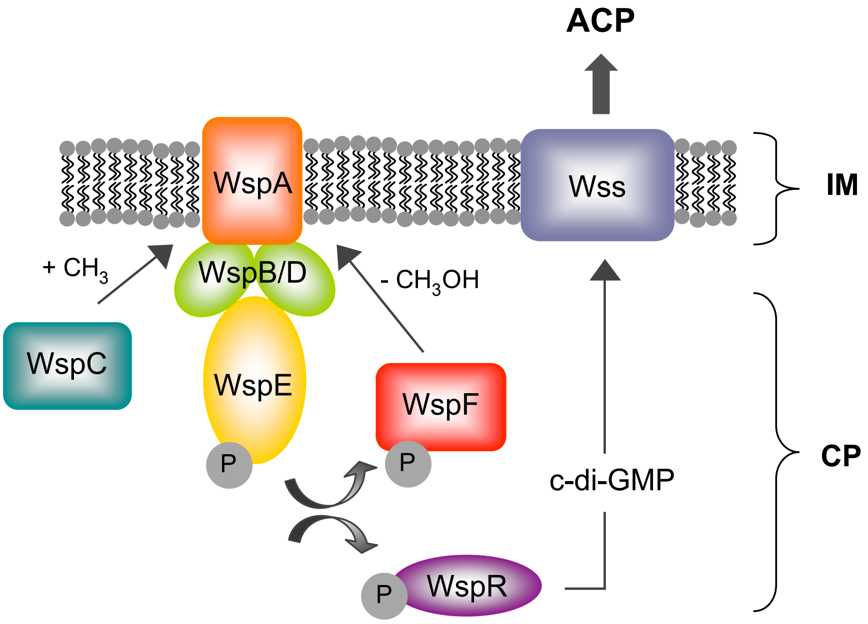

**Figure 1.5: The Wsp signal transduction pathway model.** The *wsp* operon combines seven genes, *wspABCDEFR*, that regulate a chemotaxis pathway similar to the Che pathway in *E. coli*. WspA, a MCP, together with the scaffold proteins WspB and WspD and the kinase WspE form a protein complex that is localised at the inner membrane (IM). Located in the cytoplasm (CP), a methyltransferase (WspC) and a methylesterase (WspF) are components of an opposing feedback loop that adds to and removes methyl groups from WspA, which leads to activation or deactivation of the kinase WspE. The active kinase transfers phosphoryl groups to WspF that removes methyl groups from WspA and phosphoryl groups to WspR, a DGC, which becomes effective and produces c-di-GMP, an activator of ACP biosynthesis by Wss. Figure adapted from Bantinaki *et al.* (2007).

The addition of a methyl group by WspC leads to autophosphorylation of the kinase that transfers phosphoryl groups to WspF and WspR, which become activated. Phosphorylated WspF removes the methyl group from WspA and consequently inactivates the kinase. The addition of a phosphoryl group to WspR, which is predicted to be a DGC, results in the synthesis of c-di-GMP, a secondary messenger molecule and activator of enzymes of the cellulose biosynthesis pathway amongst others (Fig. 1.5) (Malone *et al.*, 2007). Bantinaki (2001) suggests a continuous switch between active and inactive states of Wsp.

The results of suppressor analysis and comparative sequencing suggest that *wspF* is a mutational target for the evolution of WS. A disturbed function of WspF caused by a mutation invalidates the opposing feedback loop with WspC and turns on kinase activity permanently. Consequently phosphoryl groups are transferred to the DGC WspR continuously and cause the overproduction of c-di-GMP, the inducer of ACP production through Wss (Fig. 1.5; Bantinaki *et al.*, 2007).

#### *1.4.3.3 The* **aws** *operon*

The *aws* (*a*lternate *w*rinkly *s*preader) locus was identified using re-evolution experiments (McDonald *et al.*, 2009). The derived WS types from a manufactured *P. fluorescens* SBW25 genotype with a deleted *wsp* operon (Fig. 1.5) were subject to suppressor analysis and comparative sequencing. The results suggested Aws as an alternative mutational pathway that is different from the evolution through the *wsp* operon (McDonald *et al.*, 2009).

Aws encompasses the three genes *awsX*, *awsR* and *awsO*. Together they occupy 2.3 kb in the genome and are part of a putative operon. AwsX is a membrane-associated protein and predicted to be a negative regulator of AwsR, which shows features of a DGC (GGDEF domain) with a c-di-GMP binding site. The third gene, *AwsO*, encodes for a protein with similarity to OmpA and is thought to have porin-like functions in the outer membrane (Fig. 1.6; Sugawara & Nikaido, 1992; Gehrig, 2005; McDonald *et al.*, 2009).

It has been shown that mutations in *awsX* are likely to occur during the evolution of WS. A mutation in this locus is sufficient to induce the phenotype. As a consequence the negative regulation of the DGC AwsR is interrupted and causes the overproduction of c-di-GMP. The activation of *wss* by c-di-GMP leads to overproduction of ACP on the cell surface (Fig. 1.6; McDonald *et al.*, 2009).

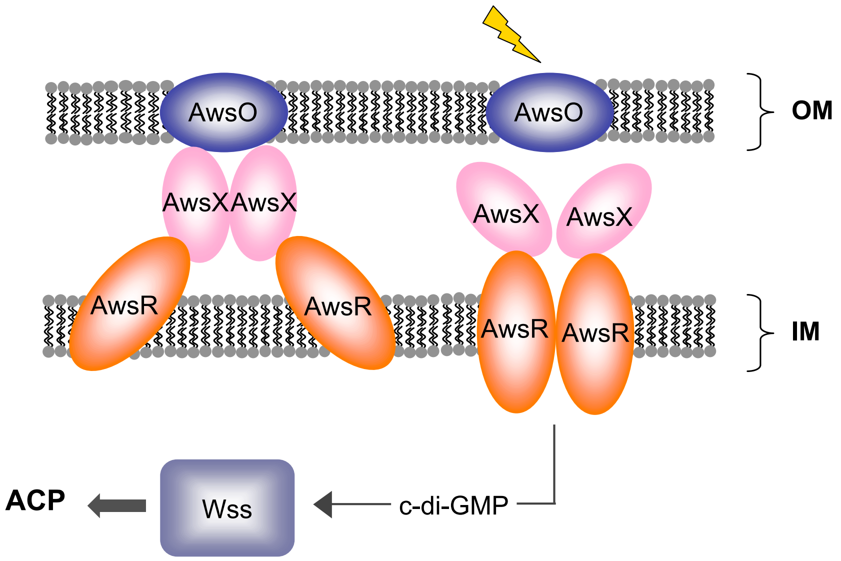

**Figure 1.6: The Aws pathway model.** The *aws* operon encompasses three proteins. AwsO is thought to be an outer membrane (OM) porin. AwsX is a membrane-bound protein and predicted to be a negative regulator of AwsR, which has the function of a DGC. A) When non-activated, AwsX is bound to AwsO and the DGC AwsR stays inactive. B) After an environmental signal AwsX becomes detached from AwsO, which leads to the dimerisation of AwsR (active state). Now AwsR increases the level of c-di-GMP, which in turn activates the cellulose biosynthetic enzymes of Wss, causing the overproduction of ACP (McDonald *et al.*, 2009). (IM=inner membrane)

### *1.4.3.4 The* mwsR *locus*

Further re-evolution experiments and supressor analysis identified an additional alternative mutational pathway leading to a WS phenotype, the *mws* (*m*ike's *w*rinkly *s*preader) locus (Fig. 1.7). Here a constructed genotype (wild type SBW25 with *wsp* and *aws* deleted) was propagated in a static microcosm. After diversification the WS types were subjected to transposon suppressor analysis. Comparative sequencing revealed causal mutations within a single large gene (3.8 kb) – *mwsR*. The gene encodes for a protein with multiple domains including a GGDEF domain, suggesting the function of a DGC. Furthermore an EAL domain was identified, which is predicted to be a phosphodiasterase (PDE) responsible for the degradation of c-di-GMP. Additionally, MwsR has two trans-membrane regions (PAS domains) indicating a membrane-bound localisation within the cell (McDonald *et al.*, 2009).

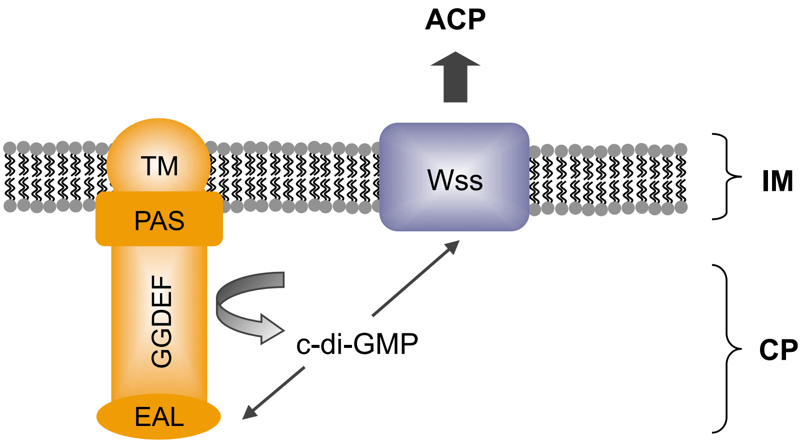

**Figure 1.7: Model of MwsR function.** MwsR is a large membrane-bound protein with multiple domains. The signal is received by the two transmembrane regions of the PAS domain and transmitted to the GGDEF domain, which produces c-di-GMP. The EAL domain has the proposed antagonistic function of degrading c-di-GMP. It has been suggested that the balance between c-di-GMP production and degradation is the determinator for cellulose biosynthesis by Wss. If the balance is shifted and c-di-GMP is overproduced, then Wss becomes active and synthesisis ACP (McDonald *et al.*, 2009). (TM = transmembrane region, IM = inner membrane, CP = cytoplasm)

Because MwsR has the predicted function of a DGC as well as the function of a PDE it was thought that the balance between synthesis and breakdown of c-di-GMP is crucial for the evolution of WS because it can lead to activation or deactivation of the *wss* operon. The MwsR model predicts that the activity of the EAL domain (PDA) negatively regulates the activity of the GGDEF domain (DGC). It has been shown that the effect of a mutation in the EAL domain is the inhibition of the c-di-GMP degradation. Consequently an overactive DGC causes overproduction of the cellulosic polymer (Fig. 1.7; McDonald *et al.*, 2009).

### 1.4.4 The Reverse-Evolution Experiment (REE)

The experiment presented here took advantage of the ability of *P. fluorescens* to quickly diversify in a static microcosm. New variants (WS) arise by mutations in genes that are involved in the production of cellulose (see section 1.4.3). The polymer enables the cells to colonise the air-liquid interface. Within this niche oxygen and nutrients are equally available and provide a benefit enabling WS to increase in frequency (see section 1.4.2). Under shaking conditions, however, this novel trait is lost due to the costs of cellulose production without it providing any benefits. One of the fundamental questions that were addressed by the REE was how repeatable and predictable is evolution. Can a novel trait (WS), once it is lost under shaking conditions, repeatedly evolve when the bacteria are grown again in a static environment? What is the genetic basis of repeated WS evolution?

Dr. Hubertus Beaumont and Professor Paul Rainey initiated the REE in 2006 with 12 replicate lineages of the bacterium *P. fluorescens* SBW25. The lineages were kept in glass microcosms containing 6 ml liquid KB media. All lines were treated equally and were exposed to two alternating environmental conditions – static and shaken. Under static conditions the microcosms were left undisturbed, whereas under shaken conditions the microcosms grew under constant shaking at 160 rpm.

The experiment started by growing 12 replicate cultures of *P. fluorescens* SBW25 in the static environment. After the first selection round (three days), the static cultures were diluted and plated on KB agar plates, which were incubated for two days. The plates were then examined for colonies with morphologies different to that of the ancestral colony phenotype. One colony of the most common new type was picked from the agar plate, transferred into a fresh microcosm and subsequently exposed to shaken conditions. After three days (second selection round) the procedure described above was repeated, and the most common new colony type now growing under shaken conditions was selected and transferred into a new microcosm. In essence, every time a new type was found, it was transferred into new media, and kept under the opposite environmental condition to that in which it had evolved in (Fig. 1.8).

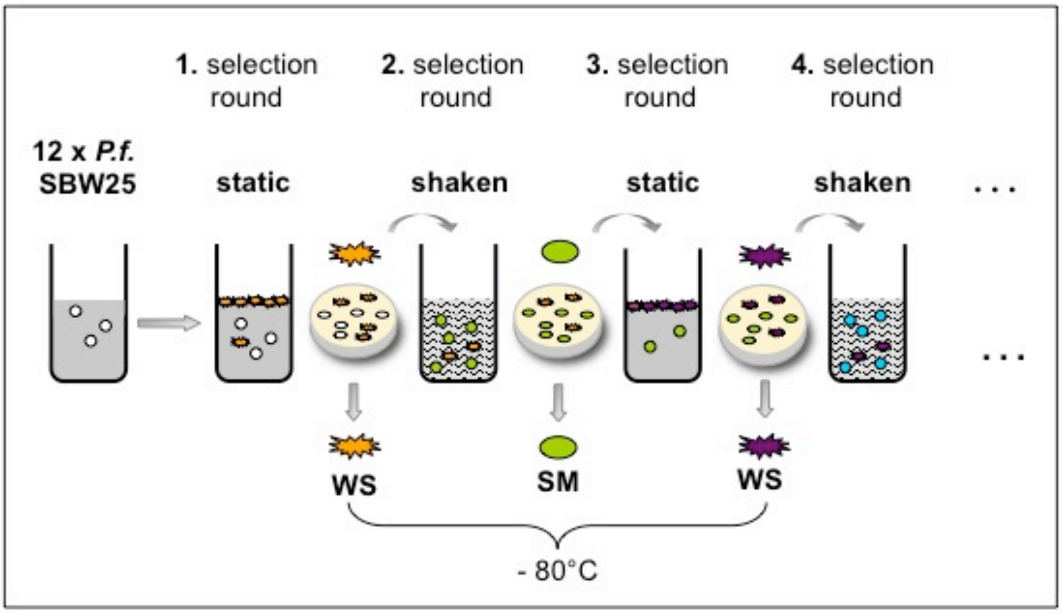

**Figure 1.8: Experimental design of the REE of one of the 12 replicated lines in *P. fluorescens* SBW25.** The initial inoculation was carried out with 12 replicates of SBW25. Cultures were kept in microcosms under static conditions. After one selection round (three days), the cultures were plated on agar plates and screened for new colony types. The most common new type was used to inoculate new KB microcosms that were kept under shaken conditions. After three days the cultures were plated and screened for the most common new colony type. The new type was then used to inoculate the new static microcosms starting a new cycle of evolution. All 12 replicate lines went through three to eight cycles of evolution during the REE (Beaumont *et al.*, 2009).

Where no new type could be detected after three days, 6 μl of the bacterial culture were transferred into fresh microcosms. These were kept under the original conditions parallel to the plating step (see above), either static or shaken, for additional three days (same selection round). Cultures were diluted, plated and screened for new types. Where no new colony morphology type was found on the plates, the procedure described above was repeated until a new colony type was observed. If no new type was observed over a long period of time (10 transfers within the same selection round) the lineage was terminated. Every newly evolved type was stored in small tubes containing a saline-glycerine solution at -80°C. The ‘frozen fossil collection’ was later revived and used for further analysis (Beaumont *et al.*, 2009).

#### *1.4.4.1 Whole-genome sequencing of evolutionary endpoints*

Sequencing the entire genome of an organism has become an indispensable and powerful tool for evolutionary biologists. Primarily it has been used to discover changes in the DNA sequence of an evolutionarily derived genome by comparing it with the known reference ancestral genome sequence. Solexa sequencing is an efficient technology that yields high quality sequences and was used in this experiment. Here the genotypes of the evolutionary endpoint in the 12 lineages that did not produce any more new colony types was revived from the freezer stock and used for Solexa sequencing. This way it was possible to reconstruct the evolutionary pathway that was taken by each lineage depending of the environmental conditions (static or shaken). The phenotype was mapped to the genotype by identifying all causative mutations and the order in which they appeared.

## 1.5 The rise of a stochastically switching phenotype

The experimental design of the REE made it possible to follow real-time evolution of 12 parallel lineages of *P. fluorescens* that evolved alternately in a static and shaken environment, starting from the same clonal founder population (Fig. 1.8; see section 1.4.4). Increased selection pressure under static conditions, due to the lack of oxygen in the media, favours diversification into multiple new types, usually different WS phenotypes. Propagation, typically of a WS type, in a shaken environment often reverses the phenotype to one that resembles the ancestral SM type (Fig. 1.1; Rainey & Travisano, 1998). Each phenotypic change was caused by a mutational modification. In two of the 12 lineages, Line 1 and Line 6, a novel phenotype evolved independently after nine rounds of selection. This new type showed distinctive colony morphology on agar plates and was not detected in any of the 12 lineages during earlier selection rounds. Phenotypic and genetic analysis revealed that the new type had properties of a stochastically switching genotype (Beaumont *et al.*, 2009; Gallie, 2009).

The switcher types of Line 1 and Line 6 were genetically and phenotypically characterised in previous studies (Beaumont *et al.*, 2009; Gallie, 2009). On agar plates, switcher populations appear as a combination of translucent and opaque colonies (Fig. 1.9A) and/or colonies with random opaque sectors (Fig. 1.9B). Comparative sequencing revealed that the phenotypic variation between translucency and opaqueness was not caused by genetic variation. The different colony types were genetically identical. Furthermore, a transfer of a translucent or opaque colony to a separate KB agar plate resulted again in a mixture of these different colony types, indicating that one type can rapidly give rise to the other type and vice versa (Beaumont *et al.*, 2009).

Negative staining and examination under the microscope revealed that these isogenic switcher populations are composed of two different cell types. One fraction of the population consists of regular looking cells (non-capsulated cells, Fig 1.9C) and the other fraction of cells displays a coat-like layer on the surface (capsulated cells; Fig 1.9C). Previous studies found a correlation between capsulation and colony opacity, and identified colanic acid as the capsule material. Cells can reversely switch between the two cell stages in a stochastic manner and cause this variation in switcher colony morphologies on agar plates (Beaumont *et al.*, 2009; Gallie, 2009).

Capsulated cells were observed in the ancestral SBW25 population but they occurred at a very low frequency. They were only detectable under the microscope at the cellular level but not at the colony level on agar plates. It has been suggested that the switching mechanism was operating in the ancestral population, but at a very low rate. As evolutionary time went on, an additional mutation occurred in two of the 12 lineages that increased the rate of the transition between capsulated and non-capsulated cells and led to the typical switcher phenotype that was detected after nine rounds of selection (Beaumont *et al.*, 2009; Gallie, 2009). Rapid stochastic switching between the two different cell stages is considered to be a bet-hedging strategy that provides a benefit in a fluctuating environment. The ability to produce phenotypically diverse offspring reduces the risk of being maladaptive in an unpredictable environment and increases the chance of survival (Beaumont *et al.*, 2009; Libby & Rainey, 2011; Rainey *et al.*, 2011).

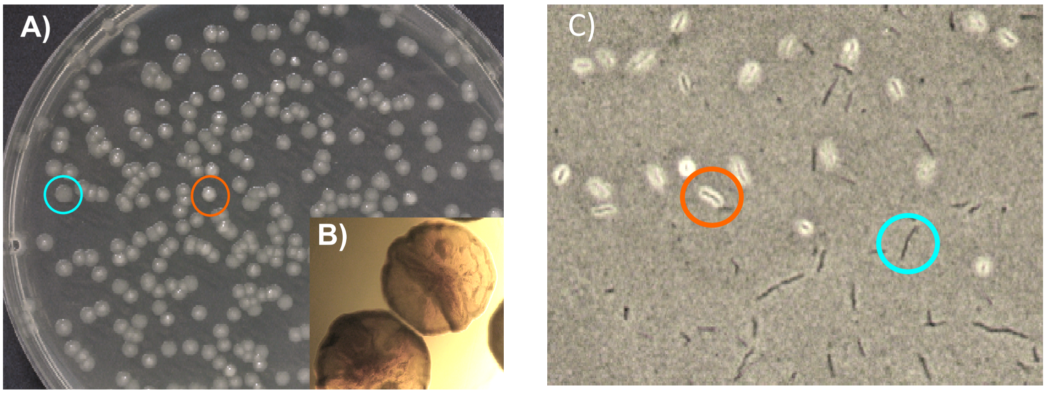

**Figure 1.9: Phenotype of a switcher genotype at colony and cellular level.** Illustrated in A) are colonies of the switcher genotype that evolved in Line 1 during the REE. After 48 hours growth on KB agar plates there were translucent (blue) and opaque colonies (orange). Some colonies showed opaque sectors (B). After negative staining colonies appeared as a mixed population of capsulated (orange) and non-capsulated (blue) cells (C).

In Line 1 and Line 6 the switcher mutations in *carB* and *rpoD* have been identified as the sole cause of switching. The introduction of the *carB* mutation and the *rpoD* mutation into the ancestral SBW25 genome was sufficient to induce the typical switcher phenotype at the colony and cellular level (Fig. 1.9; Beaumont *et al.*, 2009; Gallie, 2009). Previous work found a fitness benefit of the *carB* mutation under static conditions in comparison to the immediate ancestor and showed that stochastic switching is a beneficial phenotypic innovation in *P. fluorescens*. This fitness benefit was not observed when the *carB* mutation was introduced into the ancestral SBW25 background. In addition Line 1 and Line 6 show some remarkable overlap their mutational histories. It was suggested that the evolutionary history, the series of previous evolutionary events might have an effect on switcher evolution, perhaps due to certain mutations or a particular order of mutations (Beaumont *et al.*, 2009).

## 1.6 Aims of the current study

It is widely accepted that natural selection is the main driver in the evolution of key innovations. To what degree and under which circumstances random mutations, genetic drift, history and natural selection contribute to the outcome of evolution is a long-standing controversy (Bock, 1959; King & Jukes, 1969; Gould & Lewontin, 1979; Mayr, 1983; Maynard Smith *et al.*, 1985; Travisano *et al.*, 1995a; Losos *et al.*, 1998; Wagenaar & Adami, 2004; Vermeij, 2006; Pal *et al.*, 2006; Beatty & Desjardins, 2009; Conway Morris, 2010). Some studies have shown that organisms can find similar solutions for the same adaptive problem, which can be seen in similar phenotypes (Travisano *et al.*, 1995a; Losos *et al.*, 1998; Wichman *et al.*, 1999; Allender *et al.*, 2003; Boughman *et al.*, 2005; Woods *et al.*, 2006; Bantinaki *et al.*, 2007; McDonald *et al.*, 2009). The general occurrence of convergent evolution might be attributable to highly selective identical environmental conditions, strong genetic constraints, and epistatic interactions that drive populations along similar evolutionary pathways and lead to similar evolutionary outcomes. Nonetheless similar phenotypes do not necessarily share the same genotype (Korona, 1996; Lenski & Travisano, 1994; Hoekstra *et al.*, 2006; McDonald *et al.*, 2009), which can have consequences for subsequent evolution (Gould, 2002; Blount *et al.*, 2008; Khan, 2011).

I used the well-investigated *P. fluorescens* model system to examine the impact of random mutations and evolutionary history in the evolution of a stochastically switching phenotype. This novel phenotype was observed during the REE in two out of 12 replicate populations of *P. fluorescens* SBW25 after eight preceding mutations. Genetic drift, as mentioned earlier provides an additional source of random variation (see section 1.1.3) and could have caused the occurrence of switcher or other new types during the REE. Real-time evolution experiments with microorganisms however have shown that there is a strong repeatability between replicate lineages that evolve under identical environmental conditions (Treves *et al.*, 1998; Cooper *et al.*, 2003; Herring *et al.*, 2006; Pelosi *et al.*, 2006; Woods *et al.*, 2006; Ostrowski *et al.*, 2008; Conrad *et al.*, 2009; Beaumont *et al.*, 2009). In addition, under strong selection mutational change has been shown to be adaptive; hence genetic drift has a negligible effect in large evolving populations (Brockhurst et al., 2011). Indeed previous work found evidence that stochastically switching is adaptive under static conditions (Beaumont *et al*., 2009).

Line 1 and Line 6 show a high degree of overlap in their mutational history. Both lineages have not only mutations in common but they also show a similar chronological order in which the mutations occurred. It was concluded that evolutionary history (number and order of evolutionary events over time) might have an impact on switcher evolution (Beaumont *et al.*, 2009). The extended knowledge of the *P. fluorescens* model system and the switcher genotype, together with the 'frozen fossil record', provides an excellent opportunity to go back in evolutionary time and to gain insights into the role of chance and history in the evolution of a switcher genotype. In particular this thesis comprises:

### Chapter 3: *Multiple evolutionary pathways to a switcher genotype in P. fluorescens*

In this chapter four independent *P. fluorescens* lineages (Line 1, Line 3, Line 6 and Line 12) that evolved during the course of the REE were examined for their ability to evolve a switcher genotype. Switcher Re-Evolution Experiments (SREE) from genotypes carrying the same number of mutations but that differed in the order and kind of mutation (1s4, 3s4, 6s4 and 12s4) were carried out to investigate whether the four parallel lineages differ in their likelihood to evolve phenotypic switching under static conditions and whether history can explain the variation in switcher occurrence between the four lineages. In addition, parallel occurring switchers in different lineages were examined for parallel genotypic evolution. Furthermore, a non-switcher lineage was tested for its principal capacity to evolve a switcher, and whether genetic constraints based on the mutational history caused some limitations.

### **Chapter 4:** ***Impact of history on the evolution of a switcher genotype in P. fluorescens***

The focus of this chapter was the investigation of genotypes that represent intermediate stages that evolved along the evolutionary pathway of the original switcher in Line 1 during the REE (SBW25 (ancestor), 1s1, 1s2, 1s3 and 1s4 (immediate ancestor of the original switcher)). The original switcher mutation, *carB**, was introduced into the intermediate genetic backgrounds to test whether this mutation can induce the switching phenotype at the colony level (opaque and translucent colonies, colonies with opaque sectors) and the cellular level (mixed population of capsulated and non-capsulated cells). Furthermore fitness experiments clarified whether *carB** has the potential to arise in these earlier genotypes or whether the mutational history is necessary for a *carB** switcher to occur.

### **Chapter 5:** ***Impact of evolutionary history on real time evolution of a switching genotype***

This chapter examined whether the ability to evolve a switcher that was determined in Chapter 4 could be rediscovered during real-time evolution. Different genotypes from Line 1, representing intermediate evolutionary time points, were subject to experimental evolution. SBW25, 1s1, 1s2, 1s3 and 1s4 acted as founder populations for the re-evolution of a switcher genotype (SREE). Here the genotypes that had a significant fitness increase after *carB** was introduced into the genome in the previous experiment were expected to produce the highest switcher number. In addition, the number of ancestral types, the number of competitors, and phenotypic diversity were monitored for each genotype after evolution took place in order to look at the impact of population composition on switcher evolution.

### Chapter 6: *Epistasis and the evolution of stochastic switching in P. fluorescens*

Results from Chapter 4 indicated that one genotype from Line 1, 1s3, has an increased likelihood of switcher occurrence, possibly due to the effect of one or more ealier evolutionary events. One of four different genetic loci (*wspF*, *wssA*, *wssA-J* and *wspR*) or a combination of some loci were assumed to have epistatic interactions with the switcher mutation *carB**, which is likely to be responsible for an increased switcher occurrence during the SREE from 1s3 and 1s4. The deletion of the candidate genes followed by fitness assays revealed the precise epistatic interaction and partly revealed the genetic mechanism that can explain the *carB** benefit when introduced into the genome of 1s3. The results show, how a mutation that occurred early in the history of Line 1 interacts with the switcher mutation (*carB**) later on. The combined effect of both mutations increases the likelihood of switcher occurrence within the population. This epistatic interaction demonstrates the direct impact of evolutionary history on the evolution of a switching genotype.

# Chapter 2: Material and Methods

## 2.1 Bacterial strains, culture conditions and reagents

Bacterial strains listed below (Tab. 2.1) were all stored at -80°C in 45% (v/v) glycerol saline.

**Table 2.1: Designation and genetic properties of the bacterial strains used in this study.**

***Pseudomonas fluorescens***

| **Strain** | **Genetic features** | **Reference** |
|---|---|---|
| SBW25 | Original ancestral strain isolated from the leaf of a sugar beet plant | Rainey & Bailey, 1996 |
| 1s1<br>1s2<br>1s3<br>1s4 | The first four smooth occurring genotypes of Line 1 that evolved during the REE | H.J.E. Beaumont |
| 1w4 | The ninth genotype of Line 1 containing a mutation in the *carB* gene (*carB**) that causes stochastic phenotypic switching | H.J.E. Beaumont |
| SBW25-att::Tn7-*lacZ*<br>1s1-att::Tn7-*lacZ*<br>1s2-att::Tn7-*lacZ*<br>1s3-att::Tn7-*lacZ*<br>1s4-att::Tn7-*lacZ*<br>1w4-att::Tn7-*lacZ* | SBW25, 1s1, 1s2, 1s3, 1s4, 1w4 (Line 1) containing a *lacZ* transcriptional fusion, using a mini-Tn7 transposon system | this study |
| SBW25-*carB**<br>1s1-*carB**<br>1s2-*carB**<br>1s3-*carB**<br>1s4-*carB** | Reconstruction of *carB** into SBW25, 1s1, 1s2, 1s3 and 1s4 (Line 1) | this study |

***Pseudomonas fluorescens***

| Strain | Genetic features | Reference |
|---|---|---|
| SBW25-*carB**-att::Tn7-*lacZ*<br>1s1-*carB**-att::Tn7-*lacZ*<br>1s2-*carB**-att::Tn7-*lacZ*<br>1s3-*carB**-att::Tn7-*lacZ*<br>1s4-*carB**-att::Tn7-*lacZ* | Strains containing *carB** with a *lacZ* transcriptional fusion (mini-Tn7 transposon system) | this study |
| 3s4<br>6s4<br>12s4 | Three genotypes of Line 3, Line 6 and Line 12 that evolved during the REE | H.J.E. Beaumont |
| 12s4-*carB** | Reconstruction of *carB** into the 12s4 genome | this study |
| 12s4-att::Tn7-*lacZ* | 12s4 with a *lacZ* transcriptional fusion | |
| 12s4-*carB**-att::Tn7-*lacZ* | 12s4-*carB** with a *lacZ* transcriptional fusion | |
| 1s3-Δ*wspF* | Deletion of *wspF* from the 1s3 genome | this study |
| 1s3-Δ*wspF*-att::Tn7-*lacZ* | 1s3-Δ*wspF* with a *lacZ* transcriptional fusion | |
| 1s3-*carB**-Δ*wspF* | Deletion of *wspF* from the 1s3-*carB** genome | |
| 1s3-*carB**-Δ*wspF*-att::Tn7-*lacZ* | 1s3-*carB**-Δ*wspF* with a *lacZ* transcriptional fusion | |
| 1s3-Δ*wssA* | Deletion of *wssA* from the 1s3 genome | this study |
| 1s3-Δ*wssA*-att::Tn7-*lacZ* | 1s3-Δ*wssA* with a *lacZ* transcriptional fusion | |
| 1s3-*carB**-Δ*wssA* | Deletion of *wssA* from the 1s3-*carB** genome | |
| 1s3-*carB**-Δ*wssA*-att::Tn7-*lacZ* | 1s3-*carB**-ΔwssA with a *lacZ* transcriptional fusion | |
| 1s3-Δ*wssA-J* | Deletion of *wssA-J* from the 1s3 genome | this study |
| 1s3-Δ*wssA-J*-att::Tn7-*lacZ* | 1s3-*ΔwssA-J* with a *lacZ* transcriptional fusion | |
| 1s3-*carB**-Δ*wssA-J* | Deletion of *wssA-J* from the 1s3-*carB** genome | |
| 1s3-*carB**-Δ*wssA-J*-att::Tn7-*lacZ* | 1s3-*carB**-Δ*wssA-J* with a *lacZ* transcriptional fusion | |

***Pseudomonas fluorescens***

| **Strain** | **Genetic features** | **Reference** |
|---|---|---|
| 1s3-Δ*wspR* | Deletion of *wspR* from the 1s3 genome | this study |
| 1s3-Δ*wspR*-att::Tn7-*lacZ* | 1s3-Δ*wspR* with a *lacZ* transcriptional fusion | |
| 1s3-*carB**-Δ*wspR* | Deletion of *wspR* from the 1s3-*carB** genome | |
| 1s3-*carB**-Δ*wspR*-att::Tn7-*lacZ* | 1s3-*carB**-ΔwspR with a *lacZ* transcriptional fusion | |

***Escherichia coli***

| **Strain** | **Genetic features** | **Reference** |
|---|---|---|
| DH5α-*λpir* | *supE*44, Δ*lacU*169, *hsdR*17, *recA*1, *endA*1, *gyrA*96, *thi*-1, *relA*1, *λpi*r | Invitrogen |
| TOP10 | F', *mcrA*, Δ(*mrr-hsd*RMS-*mcr*BC), Φ80*lacZ*ΔM15, Δ*lac*X74, *deo*R, *recA*1, *ara*D139, Δ(*ara-leu*)7697, *gal*U, *gal*K, *rps*L, StrR, *endA*1, *nup*G | Invitrogen |

## 2.1.1 Culture media

Bacteria were cultured in King's B liquid media (KB) (King *et al.*, 1954) or on KB agar plates. KB broth contained: 10 g.L$^{-1}$ glycerol (Merck), 20 g.L$^{-1}$ Protease Peptone No.3 (Difco Laboratories), 1.5 g.L$^{-1}$ $K_2PO_4 \cdot 3H_2O$ and 1.5 g.L$^{-1}$ $MgSO_4 \cdot 7H_2O$. Solid media included tryptone (Difco Laboraties) instead of Protease Peptone No.3 and 1.5 % bacteriological agar (AppliChem). Lysogeny Broth (LB) (Bertani, 1951) contained: 10 g.L$^{-1}$ $NaCl_2$ (Ajaxs Finechem), 10 g.L$^{-1}$ casein peptone (HiMedia Laboratories) and 5 g.L$^{-1}$ yeast extract (HiMedia Laboratories). Solid media was supplemented with 1.5 % agar.

### 2.1.2 Culture conditions for bacterial strains

#### *2.1.2.1 Pseudomonas fluorescens (P. fluorescens)*

All strains of *P. fluorescens* were revived from frozen stock cultures by transfer to KB agar plates. After two days of incubation, a single colony was chosen from the plate and transferred into a 25 ml glass tube microcosm containing 6 ml of liquid KB media. Unless stated otherwise, the microcosms were incubated at an angle on a shaker (Innova 2300; New Brunswick Science) with access to oxygen under constant shaking at 160 rpm overnight (approximately 16 hours). The growth temperature for *P. fluorescens* was 28°C.

#### *2.1.2.2 Escherichia coli (E. coli)*

*E. coli* strains were used for cloning and transformations. Revival from freezer stock cultures was achieved as described above, except that LB was used as the growth medium. A single colony was picked after one day for inoculation of 5 ml LB media and incubated overnight (ca. 16 hours) at 180 rpm (Innova 2300; New Brunswick Science). The growth temperature for *E. coli* was 37°C.

### 1.1.3 Plasmids

Plasmids that were used in this study are listed below (Tab. 2.2).

**Table 2.2: Plasmids used in this study.**

| Plasmid name | Characteristics | Reference |
|---|---|---|
| pCR8/GW/TOPO | Spe$^R$, pUC *ori*; 2.8 kb sequencing plasmid | Invitrogen |

| | | |
|---|---|---|
| pUIC3 | Tc$^{R}$, *mob, ori*R6K, *bla, ‘lacZY* | Rainey (1999) |
| pRK2013 | Km$^{R}$, IncP4, *tra*, *mob*; mobilization plasmid used as a helper for tri-parental mating | Figurski & Helinski, (1979) |
| pUC18R6KT-mini-Tn7-*lacZ*-Gm | Gm$^{R}$, *bla*, *ori*R6K, *oriT*, $T_0T_1$, Tn7L, Tn7R, *FRT, aac*C1 | Choi et al., (2005) |
| pUX-BF13 | *tnsABCDE, ori*R6K, *mob;* suicide helper plasmid | Choi et al., (2005) |

## 2.1.3 Antibiotics

Antibiotics were used at the following concentrations: tetracycline (Tc; DICHEFA) 10 or 25 μg ml$^{-1}$ as specified (in 1 : 1 ethanol and water), kanamycin (Km; Melford) 100 μg ml$^{-1}$, gentamicin (Gm, Melford) 10 μg ml$^{-1}$ (liquid cultures) or 20 μg ml$^{-1}$ (plates), spectinomycin (Melford) 100 μg ml$^{-1}$, ampicillin (Melford) 100 μg ml$^{-1}$ and cycloserine (Melford) 800 μg ml$^{-1}$. N-[5-Nitro-2-furfurylidene]-1-aminutesohydantoin (NF, Sigma) was dissolved in dimethyl sulfoxide (DMSO, Sigma) to a final concentration of 100 μg ml$^{-1}$ and used in agar plates to inhibit growth of *E. coli*. For blue/white screening 5-Bromo-4-chloto-3-indolyl-β-Dgalactopyranoside (X-gal; Melford Laboratories) was used at 60 μg ml$^{-1}$ as a indicator of β-galactosidase activity. Dimethylsulfoxide (6.7 % (w/v) DMSO; Thermo scientific) was used in PCR to reduce secondary structure and increase PCR yield of GC-rich templates and 55 μg ml$^{-1}$ bovine serum albumin (BSA; New England Biolabs) (Ralser *et al.*, 2006) was used in restriction digest reactions.

## 2.1.4 Enzymes

Taq polymerase (Invitrogen) was used for standard PCR and Phusion™ DNA polymerase (Finnzymes OY) with proofreading activity was used in PCRs for cloning. In this study double restriction digests were performed with *Bgl* II and

*Spe* I (New England Biolabs). The enzymes Exonuclease I (ExoI; New England Biolabs) and Calf-intestinal phosphatase (CIP; New England Biolabs) were used for PCR product purification and T4 DNA Ligase (Invitrogen) for ligation reactions.

### 2.1.5 Primers

Primers were synthesised by either Invitrogen or IDT. After delivery they were resuspended in deionised water to a concentration of 100 pmol μl$^{-1}$ and stored at -20°C. Primers were used at a concentration of 10 pmol μl$^{-1}$.

**Table 2.3: List of primers used in this study.** Restriction sites are underlined and overlap regions for overlap-extension-PCR are in italic.

| Name | Sequence (5' → 3') | Target |
|---|---|---|
| SN003 PF | AGATCTCGACCTGATGTTCCGCCTCAAG | *carB** (*Bgl* II restriction site) |
| SN004 PR | AGATCTAGCGCCAACTGTACGTTCATC | *carB** (*Bgl* II restriction site) |
| SN005 PF | TATCGATCGTGCCGAAGACC | *carB* (no mutation) |
| SN006 PF | TATCGATCGTGCCGAAGACT | *carB** (with C-T mutation) |
| SN007 PR | CAGCATAACTGGACTGATTTCAG | Tn7-*lacZ* |
| SN008 PF | CACCAAAGCTTTCACCACCCAA | Tn7-*lacZ* |
| SN009 PF | CATGAGCGTGCTCAAGCGTGAG | downstream *carB** |
| SN010 PR | GGTGGACTTCATCTCTGGGCCCAG | upstream *carB** |
| SN013 PF | TTCCGTCACCTTGTTCACGCGAC | *carB* gene |
| SN014 PF | CGATTTCCTTCAGGGTCTTACC | *carB* gene |
| SN015 PF | GCGAGCCAGTTTCAACGGAG | *carB* gene |
| SN016 PF | CGCAGGTTCTTCTCGGTCAC | *carB* gene |
| SN017 PF | CGACTGACTTCATCTGAGTGG | *carB* gene |
| SN018 PF | CCATCGGGTCAAAGTTTTCG | *carB* gene |
| SN019 PF | TCTTTGATGAGCCTGCGGGG | *carB* gene |
| SN020 PF | AGATCTGAAGCCACTCGCCAGATCATGG | *wspR* (*Bgl* II restriction site) |
| SN021 PR | *AAGCACTCAGCC*GTGCAGTCACCTGAATTACTGC | *wspR* |
| SN022 PF | *ATTCAGGTGACTGCAC*GGCTGAGTGCTTCTGTGTAC | *wspR* |
| SN023 PR | ACTAGTGATTGACTTCGGTCAGACGCTC | *wspR* (*Spe* I restriction site) |

| | | |
|---|---|---|
| SN024 PF | CAGTCGCCTGCAATCGTTGC | downstream *wspR* |
| SN025 PR | GACGGAATTCAAGCTTGGCGAG | upstream *wspR* |
| SN026 PF | AGATCTTAAATGCCAGTTCCTGTCGTGGAC | *wssA* (*Bgl* II restriction site) |
| SN027 PR | *CGTAGTGTCGGTCAT*TCACTGCCCCTCCCCACCTGG | *wssA* |
| SN028 PF | *GGAGGGGCAGTGA*ATGACCGACACTACGTCCTCCAC | *wssA* |
| SN029 PR | ACTAGTGCGCGGAACGTGTCGAGGTTC | *wssA* (*Spe* I restriction site) |
| SN030 PF | GACCAGCAAGGGTAGGATTCTTC | downstream *wssA* |
| SN031 PR | GAAGTGCAGCATGGCGTTGGCA | upstream *wssA* |
| SN032 PF | CAGCATGTGGACCAAGTGTTC | *wspR* |
| SN033 PF | ATGGCGGTGTGTCAGTTGATGC | *wspR* |
| SN034 PF | GTGTCTGCCCTGTGTCATTCG | *wssA* |
| SN035 PR | GCATTGGCTGTAGAGGTCGAGC | *wssA* |
| SN037 PF | CAAAGCTCGGCGGTGTATGGC | *wspR* |
| SN039 PF | GCTTCGACCATACCCTGCTGCTC | downstream *carB* |
| SN040 PR | TGTTCGACCGTTTCATCAATGAGA | upstream *carB* |
| SN063 PF | GAAGATCTGAGTGAGTCGAGCAGATGAC | *wssA-J* (Gallie, 2009) |
| SN064 PR | GAAGATCTGCAAAGCTCGGTGATATCGTC | *wssA-J* (Gallie, 2009) |
| SN065 PF | TGTTCTTACAGCGAATGCATTCTCT | downstream *wssA-J* (Gallie, 2009) |
| SN066 PR | ACAGAGCCCGTAGCCATCTAC | upstream *wssA-J* (Gallie, 2009) |
| SN067 PF | *CAGCATGCGGATCCGTTGACGGA*CTGATCGAGCGTGCTGAAGG | *wssA-J* (Gallie, 2009) |
| SN068 PR | *TCCGTCAACGGATCCGCATGCTG*GCAATCCCAGTCCGTGGATAAG | *wssA-J* (Gallie, 2009) |
| SN070 PR | TGTCCCTGGCCATGGTCGCTGC | *carB* |
| SN071 PF | CTGAACATGCGACGGAATTCAAGC | upstream *wspR* |

## 2.2 Molecular methods

### 2.2.1 Polymerase Chain Reaction (PCR)

#### *2.2.1.1 Standard Taq-PCR*

A CG1-96 Thermal Cycler (Corbett Life Sciences) or a Mastercycler gradient (Eppendorf) was used to carry out polymerase chain reactions. A standard 25 μl reaction contained: 2.5 μl of 10x PCR Thermo buffer (New England Biolabs), 0.5 μl of 10mM dNTP mix (Bioline), 1 μl of 10 pmol μl forward primer and reverse primer, 0.3 μl Taq polymerase (5 U $\mu l^{-1}$) and 30-100 ng template DNA, made up to 25 μl with deionised water. If the DNA template was rich in GC, 1.5 μl of DMSO was added to the reaction. The initialisation of the PCR was achieved by heating the reaction at 95°C for 5 minutes. Amplification was realised using 30 cycles of denaturation for 20 seconds at 95°C, annealing for 20 sec at 55-65°C, and standard extension at 72°C for 1 minute per kilobase (kb) for the target DNA. Afterwards a final extension step was performed for 5 minutes at 68°C before the sample was cooled to 4°C indefinitely.

#### *2.2.1.2 Overlap-extension-PCR*

Overlap-extension-PCR was used to create gene deletions in *P. fluorescens* strains. Two fragments (~800 bp each) flanking the target gene or operon were amplified separately and then fused together by PCR. For each fragment a pair of inner and outer primers were designed. The inner primers were extended by a nucleotide sequence that overlapped with a region of the other fragment so that they could be later joined together. The outside primers formed the flanking region of the deletion fragment. *Spe* I (5'-ACTAGT) and *Bgl* II (5'-AGATCT) restriction sites were added to the 5'-end of the outer primers (Fig. 2.1) to enable cloning of the final product (Pogulis *et al.*, 1996). An overnight culture of *P. fluorescens* SBW25 was used as the template for PCR generation of the sub-fragments. In a total volume of 50 µl each, 31.7 µl $H_2O$ was mixed with 10 µl 5x Phusion GC PCR buffer, 1 µl 10 mM dNTPs, 2.5 µl of 10 pmol µl$^{-1}$ outer primer and 2.5 µl of 10 pmol µl$^{-1}$ inner primer, 1.5 µl DMSO, 30-100 ng DNA template and 0.5 µl of Phusion enzyme (2.0 U µl$^{-1}$). PCR reactions were carried out according to the instructions of the manufacturer (Finnzymes OY).

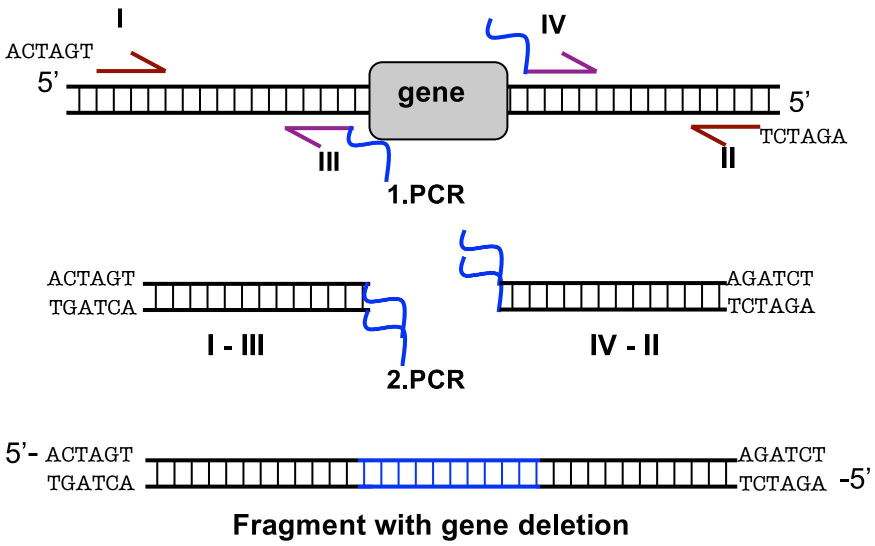

**Figure 2.1: Gene deletion with an overlap-extension-PCR. 1. PCR**: Amplification of flanking regions with two primer pairs (I, III; IV, II), where I and II are the outer primers with restriction sites *Spe* I and *Bgl* II, respectively, and III and IV the inner primers with overlapping nucleotide extensions (blue). **2. PCR**: Fusion of the sub-fragments (I-III and IV-II), and amplification creating a fragment with a shared overlapping region (blue) with the gene deletion.

Initialisation of the PCR was achieved by heating the reaction for 30 seconds at 98°C. Then, denaturation at 98°C for 10 seconds was followed by the annealing reaction at 61-68°C for 25 seconds. The extension took place at 72°C for 50 seconds. The last three steps were repeated 34 times in total. The final elongation was carried out at 72°C for 5 minutes. The products were purified with a clean-up kit and mixed together in the assembly PCR reaction. The protocol for the assembly reaction was the same as that stated above. Approximately 10 µg of each sub-template was used in a 50 μl final reaction volume. Gel electrophoresis and gel purification was used to clean up the PCR product. The DNA fragment containing the gene deletion was stored at -20°C until ligation with pCR8 (see section 2.4.2).

### *2.2.1.3 Agarose gel electrophoresis*

Gel electrophoresis was used to check for size and yield of PCR products. Gels were prepared with UltraPure™ TBE buffer (Invitrogen), 1% UltraPure™ agarose (Invitrogen) and 12 % SYBR® Safe DNA gel stain (Invitrogen). Electrophoresis was carried out at 130 V and DNA bands were visualised using a UV Transilluminator. For gel extraction, electrophoresis was carried out at 100 V. DNA bands were extracted under UV illumination with a sterilised razor blade and separated from the agarose using a QIAquick® Gel Extraction Kit (Qiagen).

## 2.3 Comparative sequencing of genes or gene fragments

A standard Taq PCR was used to amplify a DNA fragment flanking the target region using fragment-specific forward and reverse primers. Sequencing was performed by Macrogen Inc (Soeul, Korea) using forward and reverse amplification primers; where fragments exceeded 800 base pairs (bp) sequencing reactions with additional internal primers were carried out. The obtained sequence was compared to the ancestral DNA sequence of *P. fluorescens* SBW25 using the Sequencher 4.7 Demo software (Gene Codes Corporation, Ann Arbor, MI, USA).

## 2.4 Reconstruction of the *carB** mutation into different strains of *P. fluorescens*

### 2.4.1 Isolation of the DNA fragment carrying *carB**

The *carB* mutation (*carB**) that evolved during the course of a Reverse-Evolution Experiment (REE) in Line 1 (Beaumont *et al.*, 2009; Gallie, 2009) was isolated and introduced into different genotypes of *P. fluorescens*. In the first step a DNA-fragment containing the target point mutation *carB** was amplified using DNA from 1w4 as template. A primer pair flanking the requested DNA fragment was designed; a restriction site was added to the 5'-end of each outer primer. Here, primer SN003 PF and SN004 PR were extended with a *Bgl* II restriction site with the nucleotide sequence AGATCT (Tab. 2.3). Initially, the primers were used to amplify the DNA-fragment by PCR with the proofreading polymerase Phusion. The PCR contained 10 µl of 5x Phusion GC PCR buffer, 1 µl 10 mM dNTP mix, 2.5 µl of 10 pmol $\mu l^{-1}$ forward primer SN003 PF and 2.5 µl of 10 pmol $\mu l^{-1}$ reverse primer SN004 PR, 1 µl DMSO, 0.5 µl Phusion polymerase (2.0 U $\mu l^{-1)}$ and 30-100ng DNA of 1w4 bacterial strain. Deionised water was added to a total volume of 50 µl. The PCR was carried out according to the Phusion protocol (see previous paragraph) with an annealing temperature of 61°C. To confirm the success of the PCR the product was checked for correct size and concentration by electrophoresis in a 1% TBE agarose gel and stored at 4°C until further usage.

### 2.4.2 TOPO®TA cloning

PCR amplified DNA fragments were ligated with vector pCR8 using a pCR8®/GW/TOPO® TA cloning kit (Invitrogen). The vector is linearised and has 3'-T overhangs. As PCRs were performed with Phusion polymerase, a proofreading enzyme that removes all 3'-A overhangs from the ends of the fragment, a successful ligation of the target fragment into the pCR8 plasmid required readdition of 3’-A overhangs (Cheng & Shuman, 2000). Therefore the PCR product was purified (EZNA Cycle-Pure kit, Omega Bio-tek) and mixed with 0.2 μl 10x PCR Thermo buffer, 0.2 μl Taq polymerase (5.0 U $\mu l^{-1}$), 0.4 μl 10 mM dATPs for a final concentration of 0.2 mM, to a final volume of 20 μl. The reaction was incubated for 15 minutes at 72°C and the product subsequently ligated with pCR8 following the standard protocol as follows (Invitrogen): 4 μl of the PCR product was incubated with 1 μl salt solution and 1 μl TOPO vector for 5 minutes at room temperature and then placed on ice. Two μl of the TOPO ligation mixture was added to a vial of thawed One-shot™ chemically competent cells and incubated on ice for 25 minutes. After exposing the cells to a heat shock for 30 sec at 42°C, 250 μl of SOC medium was added. After one hour incubation at 37°C, 25 μl, 50 μl and 75 μl aliquots of bacterial culture were spread on LB-agar plates containing spectinomycin, and incubated overnight at 37°C. Colonies were selected from the plate the following day and sent for sequencing to confirm that the *carB** insert was introduced into the plasmid without any additional mutations. The strains were stored at -80°C.

### 2.4.3 Cloning into the pUIC3 plasmid

*E.coli* carrying pCR8-*carB** was grown in 5 ml LB overnight in the presence of spectinomycin. After purification (EZNA Plasmid mini kit, Omega Bio-tek) the plasmid was digested with *Bgl* II. In a total volume of 50 μl, 1 μg of vector DNA was mixed with 5 μl 10x NEB 3 buffer, 1 μl *Bgl* II enzyme (10 U $\mu l^{-1}$) and $H_2O$ and incubated for 1-2 hours at 37°C. The fragment was separated from the vector by gel electrophoresis in a 1% TAE agarose gel at a voltage of 100, cut out of the gel and purified. The product was stored at -20°C until ligation with plasmid pUIC3.

The suicide reporter plasmid pUIC3 contains a tetracycline resistance gene, a promoter less *lacZY* operon and a multiple cloning site (Rainey, 1999). The vector is relatively large and as a consequence a higher volume of cell culture must be mixed with the plasmid to obtain a sufficient yield of transformants. Therefore, two 5 ml LB overnight cultures were inoculated with *E.coli*pUIC3 and incubated at 37°C overnight. After plasmid purification 2 μg of plasmid DNA was linearised by digestion with 2 μl *Bgl* II (10 U $\mu l^{-1}$), 10 μl 10x NEB 3 buffer, and $H_20$ to a final volume of 100 μl. The linearised vector was isolated by gel electrophoresis in a 1% TAE gel and purified. To prevent self-ligation the digested vector was treated with Antarctic phosphatase to remove 5'-phosphate groups from the ends: 1 μg of linearised vector DNA was mixed with 0.5 units of CIP and incubated at 37°C for one hour. The product was cleaned with a PCR clean up kit (EZNA Cycle-Pure kit, Omega Bio-tek) and stored at -20°C.

Before ligating the *carB* fragment with pUIC3 the concentration of both insert and vector were estimated by spectrometry (BioPhotometer, Eppendorf). The ligase reaction contained a molar ratio of 3:1 of insert to vector: 1.45 μl of linearised pUIC3 was added to 7.05 μl of the *carB* fragment, 1.0 μl 10x NEB buffer, and 0.5 μl ligase, and incubated at 16°C overnight. DNA ligase was inactivated by heating the reaction to 65°C for 20 minutes, and stored at 4°C until transformation into *E. coli*.

#### *2.4.3.1 Manufacture of electrocompetent cells*

Electrocompetent *E.coli*DH5αλpir were prepared by the method of Sambrook and Russell (Sambrook & Russell, 2001). The starter culture, 5 ml LB, was inoculated from the freezer stock and grown overnight at 37°C. 3 ml of the starter culture was used to inoculate a flask with 350 ml LB that was pre-heated to 37°C. The $OD_{600}$ of the culture was monitored every 15-20 minutes. When $OD_{600}$ reached between ~0.35-0.4, the flask was transferred to ice for 20 minutes to stop further growth. Then, the culture was decanted into six x 50 ml chilled centrifuge tubes. The cells were pelleted at 4°C, at 900 x g for 15 minutes. The supernatant was discarded and the cells were washed with 50 ml chilled $H_2O$. The cells were pelleted at 4°C, at 900 x g for 15 minutes and the supernatant was discarded. The second wash followed the same procedure. In the final wash the pellets were resuspended in 50 ml of chilled 10% glycerol, pooled in one tube that was centrifuged at 4°C, at 900 x g for

15 minutes, and the supernatant was discarded. The pellet was resuspended in the residual liquid and aliquots of 50 μL were transferred into pre-chilled microcentrifuge tubes and stored at -80°C.

#### 2.4.3.2 Transformation of pUIC3 into E. coli

An aliquot of *E.coli*DH5αλpir electrocompetent cells was thawed on ice and subsequently mixed with 2.5 μl of the pUIC3*carB**-construct. The mixture was pipetted into a pre-chilled 2 mm cuvette that was inserted into an Elecroporator (Electroporator 2510; Eppendorf). The voltage was set at 2.5 KV and after electroporation 500 μl SOC were immediately added to the cell suspension. The culture recovered for one hour at 37°C and was then plated on LB agar plates containing Ampicillin and X-gal. The next day several blue colonies were selected from the plate and a PCR was performed using vector specific primers. Colonies that showed a correctly-sized insert were inoculated into LB broth, grown overnight at 37°C, and stored at -80°C.

### 2.4.4 Bacterial conjugation

The *carB** mutation was introduced into *P. fluorescens* through tri-parental mating. Conjugative plasmid transmission from the donor strain *E.coli*pUIC3*carB** to a *P. fluorescens* recipient strain was assisted by the Km resistant helper plasmid pRK2013, which provided *mob* and *tra* functions *in trans* (Ditta *et al.*, 1980). The donor and the helper were grown overnight in 5 ml LB with the appropriate antibiotics, and the recipient without any antibiotics. The next day the $OD_{600}$ was measured to ensure equal amounts of each culture to optimise pUIC3*carB** transmission frequency. Approximately 700 μl of the recipient was subjected to heat shock conditions at 45°C for 20 minutes. 300 μl of the donor and helper strains were pelleted by centrifugation for two minutes at 16250 x g. The pellets were washed in 500 μl LB and donor and helper were resuspended together in 250 μl LB. After the heat shock the recipient was pelleted by centrifugation for two minutes at 16250 x g and the 250 μl donor-helper mix was transferred onto the recipient pellet. The mixed culture was pelleted by centrifugation and resuspended in 30 μl water. The conjugation was spread on pre-warmed LB agar plates in a smudge of two cm diameter and dried without lid under sterile condition for approximately 10 minutes. The plates were incubated for six hours at 28°C. The smudge was scraped off and dissolved in one ml LB and vortexed for one minute. The culture was diluted stepwise down to $10^{-3}$ and a volume of 25 μl and 50 μl of each dilution was plated on LB agar plates containing NF, Tc and Xgal and incubated for two days at 28°C. Six blue colonies were selected and purified by streaking on new LB plates. From each sub-streak a single colony was transferred into 5 ml LB with Tc (25 μg $ml^{-1}$) and incubated overnight at 28°C. The cultures were stored at -80°C.

A second step selected from spontaneous excision of the vector and the wild type *carB* from the genome by recombination. Either the resulting transconjugants from the triparental mating or overnight cultures from the frozen stocks were mixed to create a pool of the six randomly-selected genotypes. A flask with 200 ml LB and no antibiotics was inoculated with 10 μl of the mixed culture and incubated overnight at 28°C. From this culture 400 μl was transferred into a new flask containing 20 ml LB. After 30 minutes at 28°C, 10 μg $ml^{-1}$ Tc was added to the culture and incubated at 28°C for a further two hours. The antibiotic D-cycloserine (1 mg $ml^{-1}$) was added to kill all dividing cells over the course of five hours. The culture was harvested by centrifugation of one ml for two minutes at 16250 x g. The pellet was washed once in one ml LB and resuspended in 500 μl LB. The cell suspension was diluted serially from $10^{0}$-$10^{-3}$ and 25 μL and 50 μl of each dilution was spread on LB agar plates containing X-gal. After two days at 28°C several white colonies were selected at random and checked for deletion of the pUIC3 vector backbone and presence of the *carB** mutation. PCR was performed with primers SN004 PR and SN005 PF or SN006 PF. The PCR product generated with SN004 PR and SN006 PF (Tab. 2.3) confirmed the successful introduction of *carB** into the genome. The strains were stored at -80°, and sequenced to ensure that no additional mutations were introduced.

### 2.4.5 Introducing a *lacZ* transcriptional fusion into *P. fluorescens* strains

Neutral phenotypic marking of bacterial strains by a *lacZ* transposition system was used for blue/white screening of mixed bacterial cultures on agar plates, for example in competition experiments. Plasmid DNA was isolated from overnight cultures of *E.coli*HPS1 that carried the delivery plasmid pUC18-Tn7-*lacZ*-Gm (Gm resistance) and *E. coli* that carried the helper plasmid pUX-BF13 (Amp resistance). After estimating concentrations, plasmids were stored at -20°C until further usage.

In the next step the cells of the target strain were made electrocompetent. 6 ml KB was inoculated with *P. fluorescens* target strain and grown overnight. Aliquots of the culture were divided into four pre-chilled microcentrifuge tubes and pelleted at 12700 x g for one minute. Each pellet was washed twice with 1 ml of room temperature 300 mM Sucrose with 30 sec centrifugations at 12700 x g between washes. All four pellets were combined in a final volume of 100 μl 300 mM Sucrose and stored on ice until ready to transform.

For the transformation 500 ng of plasmid DNA – 250 ng of pUX-BF13 and 250 ng of pUC18-Tn7 – were mixed with the 100 μl aliquot of electrocompetent cells and transferred into a chilled 2 mm cuvette. The cuvette was placed into the Electroporator that was set at 2.5 KV. After electroporation the cells recovered in 1 ml SOC for two hours at 28°C. The cells were pelleted by centrifugation, washed, and resuspended in 150 μl SOC. 25 μl, 50 μl and 75 μl aliquots were spread on LB agar plates containing Gm and X-gal.

After two days the plates were screened for single blue colonies. Eight to ten colonies were screened for successful Tn7-Gm transposition by PCR using primers

SN007 PR and SN008 PF (Tab. 2.3). Bacteria that had the insert successfully integrated into the genome were stored at -80°C.

## 2.5 Biological assays

### 2.5.1 Capsule count assay

Purpose of this experiment is to determine the proportion of capsulated cells in a *P. fluorescens* population using a staining technique and microscopy. Capsules consist of colanic acid, which is an extracellular polymeric substance (EPS) appearing as a thick layer around the cell when negative stained with India ink. It stains the background dark against which the capsules around the cells became visible.

Five replicate KB microcosms containing 6 ml liquid KB were inoculated with one colony of the *P. fluorescens* strain and incubated overnight. For acclimatisation to the media 6 μl of the cultures were transferred into fresh media and incubated for an additional night. A tenfold dilution was made of the India ink and 2 μl from the bacterial culture was added to 10 μl of the water – ink solution. The mixture was incubated for approximately 1 minute and 8 μl of it was transferred onto a glass slide, covered with a cover slide, and left at room temperature for a maximum of 10 minutes. After this time, motility of the cells was reduced and pictures could be taken with a camera (Olympus XM10) that was attached to a microscope (Olympus BX61).

For each strain, one slide was prepared per replicate from which ten pictures were taken at magnification 600-fold using Cell* Imaging software. With a program dedicated to this purpose (jennaclick.jar) written by Frederic Bertels (Gallie, 2009), 100 cells were counted in each picture and classified as being either capsulated or non-capsulated.

## 2.5.2 Determination of relative fitness of 12s4*carB**

### *2.5.2.1 Fitness assay*

This experimental approach, namely a measure of the change in density of a strain over time relative to a competitor, was used to reveal fitness differences between two distinct *P. fluorescens* strains when in direct competition within the same environment (Lenski *et al.*, 1991). If colony phenotypes weren't clearly distinguishable, one of the competing strains was tagged with a neutral *lacZ* marker (see section 2.4.5.).

Competitors were revived on KB agar plates or on KB X-gal plates for marked strains. One colony per strain and per replicate was picked from the plate and transferred separately into 6 ml liquid KB media and incubated overnight. The next day the microcosms were vortexed for 30 seconds. After measuring the $OD_{600}$ of the overnight cultures, new microcosms were inoculated with both competitors at an estimated cell ratio of 1:1. To confirm the ratio of the two competitors at the start of the fitness assay ($T_0$), 25 µl and 50 µl of a $10^{-3}$ dilution were plated on KB agar plates containing X-gal. During the fitness assay, the microcosms were incubated statically for 48 hours at 28°C.

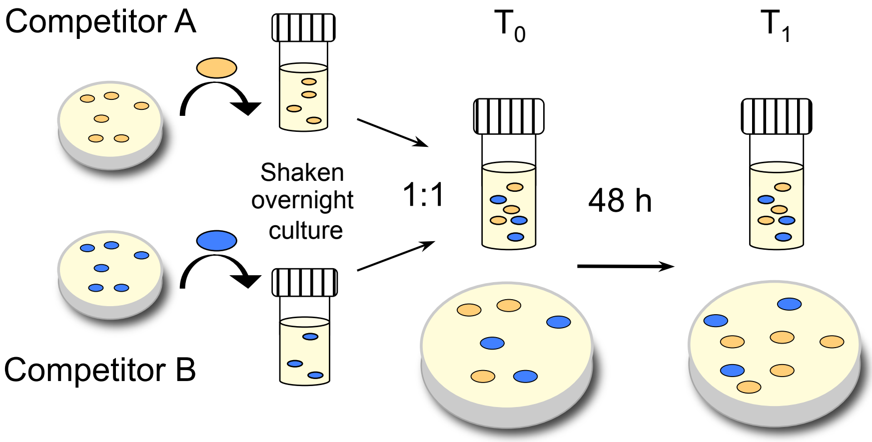

**Figure 2.2: Experimental design of a competition assay.** Competitors A and B were mixed together in a 1:1 ratio in a 6 ml KB microcosm using an overnight culture. A small

amount of the mixed culture was plated and colonies were counted at the beginning of the competition ($T_0$). After 48 hours growth under static conditions the bacterial culture was plated and colony numbers were determined again ($T_1$).

The cultures were then diluted to $10^{-6}$ and 50 µl and 75 µl volumes were plated on KB agar plates that contained X-gal. The numbers of blue and white colonies gave an estimate of the densities for each competing strain at the beginning of the competition experiment ($T_0$) and after 48 hours ($T_1$; Fig. 2.2). The same procedure was carried out for all replicates.

#### 2.5.2.2 Fitness calculation

Relative fitness was specified as the change of the ratio between the two competing strains, which grew together in the same environment. Initial and final colony counts of each competitor were included in the calculation of the Malthusian parameter for each type that gave an estimate for the change in frequency over time (Lenski *et al.*, 1991).

$$mA = dA/dt \qquad \text{(Equation 1)}$$
$$A_t = A_0 \cdot e^{m_A t} \qquad \text{(Equation 2)}$$
$$m_A = lnA_t - lnA_0/t \qquad \text{(Equation 3)}$$

$$mB = dB/dt \qquad \text{(Equation 4)}$$
$$B_t = B_0 \cdot e^{m_B t} \qquad \text{(Equation 5)}$$
$$m_B = lnB_t - lnB_0/t \qquad \text{(Equation 6)}$$

$A$, $B$ = competitors

$m$ = growth rate of competitor $A$ or competitor $B$

$t$ = time

$m_A$, $m_B$ = the associated Malthusian parameters for each competitor

The relative fitness ($F_{AB}$) of each strain was calculated as the ratio of their Malthusian parameters. The ratio was then log2 normalised. Here, if both competitors showed the same relative fitness, $F_{AB}$ would equal zero. A fitness increase would give values greater than zero and a fitness decrease values below zero. Negative and positive fitness values had the same adicity because a limit of the lower boundary was avoided by log2 normalisation.

$$F_{AB} = log_2(\frac{m_A}{m_B}) \quad \text{(Equation 7)}$$

The $F_{AB}$ values were used for statistical analyses.

## 2.6 Estimation of population growth of *P. fluorescence* lineages

Growth of bacterial cultures was determined by measuring the optical density at a wavelength of 600 nm ($OD_{600}$) over the course of 24 hours using a Synergy 2 plate reader (Biotek®). The *P. fluorescens* strain of interest was revived from the frozen stock culture on KB agar plates. For each replicate one colony was transferred into 6 ml liquid KB media and incubated overnight with constant shaking. On the following day 6 μl of the overnight culture was used to inoculate 6 ml fresh liquid media and incubated overnight. The bacterial cultures were then diluted in liquid KB media to give an $OD_{600}$ between 0.001 and 0.003. From the diluted culture 150 ml aliquots were transferred into separate wells of a 96 well plate (RayLab). Some wells were filled with liquid KB media only and served as blanks. The cultures were incubated at 28C with constant shaking and $OD_{600}$ values were recorded in 10 minutes intervals. Standard analysis was carried out by plate reader software (Gen 5, Biotek) to determine maximal growth rate (Vmax), mean and standard error.

# Chapter 3: Multiple evolutionary pathways to a switcher genotype in *P. fluorescens*

## 3.1 Introduction

Evolution of novel structures is the result of the combined effects of deterministic processes (e.g. selection) and undirected processes (e.g. random mutations, genetic drift). For a long time it was widely accepted that purely deterministic processes drive evolution and selection is the only factor responsible for the establishment of novel phenotypes and key innovations. It was believed that present life forms are the consequences of optimal adaptations to the current environment. In addition scientists argued that natural laws determine how molecules, the fundamental building blocks of life, interact with each other and lead to the structures we observe today (Pross, 2003; Conway Morris, 2010; Boiteau & Pascal, 2011). Based on these assumptions evolution would be to a greater extent the result of deterministic processes and the comparable effect of randomness would be negligible and not relevant for evolutionary outcomes. This perspective is still common today.

In contrast to the deterministic processes in evolution there are random factors, which scientists became increasingly interested in about 40 to 50 years ago. They were inspired by life forms that were not easily understandable by selection alone. They argued that the ultimate sources of variation are mutations, which occur spontaneously throughout the genome. Where and when they occur in the genome happens by chance and is unpredictable. Most of the time mutations are deleterious and sifted out by selection. In other cases they might be neutral, without an immediate effect on the phenotype, and are evolutionarily non-relevant at the time of their occurrence. In a few instances these random mutations are beneficial and likely to become fixed within the population. The order in which mutations accumulate over time is assumed to be random and future evolutionary events may depend on the mutational history (Simpson, 1949; Monod &

Wainhouse, 1971; Gould & Lewontin, 1979; Gould, 1989). Genetic drift provides another source of randomness in evolution. Especially in small populations it is likely that neutral or slightly deleterious mutations reach high frequencies within a population by random drift and lead to suboptimal phenotypes. Furthermore, the surrounding environment of an organism is dynamic. Biotic and abiotic factors change constantly and provide challenges for the organism, which are presumably not predictable. The history of the different environments that a population has experienced over time can be critical for future evolutionary outcomes because it leaves its signatures in the genome and on behaviour (Lewontin, 1966; Beatty & Desjardins, 2009).

The interplay between determinism and randomness raises questions about the predictability of the evolution of key innovations and has fascinated scientists to this day. A heated debate developed decades ago between scientists that believed in the predictability of evolution because of strong selection and those who thought that evolution is more dependent on random factors and rather unpredictable. If selection is the main driver of evolution one would expect the repeated appearance of similar phenotypes under similar environmental conditions, but if chance and randomness have a comparably high impact then similar phenotypes are unlikely to occur repeatedly. With increasing availability of new techniques and genetic information scientists are now able to disentangle phenotypic and genotypic evolution, and this can contribute to a greater understanding of the fundamental processes in adaptive evolution.

### 3.1.1 Convergence, parallelism and predictability of evolution

The occurrence of convergent and parallel evolution provides a powerful argument for the predictability of evolution and for a strong impact of selection on the evolutionary outcomes (Vermeij, 2006; Conway Morris, 2009; Conway Morris, 2010). A classical view of convergent evolution suggests that distantly related lineages evolve a similar solution for a similar adaptive problem. It was thought that the underlying genetic mechanisms in convergent lineages differ from each other because of different genetic starting positions and consequently unrelated lineages follow different evolutionary pathways that lead to the same evolutionary result (Arendt & Reznick, 2008). During parallel evolution closely related lineages evolve the same phenotypic innovation. Here, compared to convergent evolution, lineages share a common ancestor and the genetic changes between lineages are thought to resemble each other because each lineage started from the same genetic starting position (Fig. 3.1; Arendt & Reznick, 2008). In both cases evolution is seemingly predictable, at least on the phenotypic level.

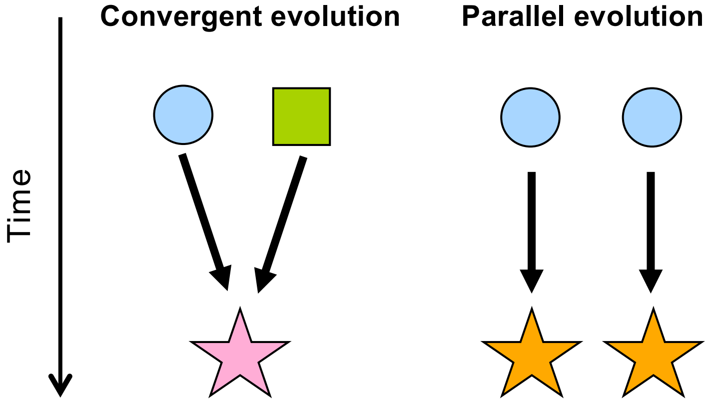

**Figure 3.1: Convergent and parallel evolution.** Convergent evolution is the evolution of a similar phenotype from distantly related lineages. The genetic mechanism is thought to be different. Parallel evolution is the evolution of a similar phenotype from closely related lineages based on similar genetic mechanisms.

Convergent evolution of key innovations has been observed multiple times throughout the history of life, for example the evolution of multicellularity, complex life cycles, complex eyes, mimicry/camouflage mechanisms, as well as the evolution of venom production in predators, occurred multiple times in different taxa. A salient example of convergent evolution is the light and dark colourization in different vertebrate taxa such as lizards (Rosenblum *et al.*, 2004), a variety of birds (Theron *et al.*, 2001; Mundy *et al.*, 2004), pocket mice (Nachman *et al.*, 2003), and the black bear (Ritland *et al.*, 2001). Here many distantly related organisms evolved a similar phenotype that served the same ecological function, for example camouflage in light-coloured environments. Genetic analysis revealed that the phenotype was achieved by the same genetic mechanism, a mutation in the melanocortin-1 receptor (*Mc1r*). Interestingly, a study that focused on parallel evolution of light-coloured hair in 'beach mice' that live in the sand dunes on the Gulf coast of Florida and Florida's Atlantic coast found that subpopulations in both areas evolved the same phenotype independently based on different genetic mechanisms. Florida's gulf coast beach mice showed a mutation in *Mc1r* as in the previous examples but this mutation was not found in the Atlantic coast mice, which led to the conclusion that the underlying genetic changes must have occurred somewhere else in the genome (Hoekstra *et al.*, 2006).

Parallel evolution has been described in *Arabidopsis thaliana*, where 20 different populations evolved early flowering. This is an example of closely related lineages evolving the same phenotype via the same genetic changes. All populations showed mutations in the *Frigida* gene (Shindo *et al.*, 2005). Parallel experimental evolution of two virus populations in different hosts led to the accumulation of many amino acid changes with significance for the adaptive phenotype. Although both populations shared half of the amino acid changes, the order in which they occurred varied between different replicates. Parallel observed changes were not involved in the highest fitness gain and no common trajectory for the adaptation to the new host was identified (Wichman *et al.*, 1999). One of the first observations during the long-time experiment with *E. coli* in the Lenski group was that 12 replicate lineages showed a similar evolutionary trajectory when grown in glucose-limited medium. All 12 lineages increased in cell size and fitness after 2000 generations and converged towards a similar phenotypic endpoint. Nonetheless there were fitness differences between the lineages after a further 8000 generations and it was suggested that this divergence could be attributed to underlying genetic differences. It was thought that time of mutation occurrence and the order of mutations varied between the 12 lineages (Lenski *et al.*, 1991; Lenski & Travisano, 1994; Blount *et al.*, 2008; Barrick *et al.*, 2009). Remarkable parallel evolution has been observed during adaptive radiation in *P. fluorescens* (McDonald *et al.*, 2009). In a static environment *P. fluorescens* diversifies quickly into multiple new types (see Chapter 1, section 1.4.1), including different 'wrinkly spreader' phenotypes (WS). The WS phenotypes have the ability to occupy the air-liquid interface and to form a biofilm, due to the overproduction of a cellulose polymer (Rainey & Travisano, 1998). Previous studies found causative mutations in *wspF*, which is a gene in the *wsp* chemosensory operon (see Chapter 1, Fig. 1.5) and is involved in the regulation of the synthesis of a cellulose polymer (Spiers *et al.*, 2003; Spiers *et al.*, 2002). McDonald and colleagues (2009) investigated 26 independently evolved WS types and revealed two additional mutational pathways, *aws* and *mws* (see Chapter 1, Fig. 1.6 and Fig. 1.7), which were common amongst the different WS genotypes. Out of 26 independent WS types, 25 harboured a mutation in one of the three loci. They concluded that genetic constraints due to specific gene function and regulatory mechanisms within each locus explain this high degree of parallel phenotypic and genotypic evolution (McDonald *et al.*, 2009).

At first glance some of the examples mentioned above support the idea that evolution might be predictable. Similar phenotypes can evolve multiple times in distantly related lineages in response to a similar selective environment. The novel phenotype can occur based on the same genetic mechanism (Ritland *et al.*, 2001; Theron *et al.*, 2001; Nachman *et al.*, 2003; Mundy *et al.*, 2004; Rosenblum *et al.*, 2004; Shindo *et al.*, 2005). This indicates strong genetic constraints that perhaps limit the number of available evolutionary pathways. In other cases closely related lineages use very different mutational pathways to achieve a similar phenotype (Hoekstra *et al.*, 2006). Here it appears that multiple genetic routes can be taken to evolve a similar phenotype. In such cases genetic evolution is less restricted and the path that evolution takes is rather unpredictable.

The examples mentioned above show that evolution at the phenotypic level can be very different from evolution at the genetic level (Manceau *et al.*, 2010). The phenotypic consequences of a mutation might be the same, but the underlying genetic change can be in a similar or a different gene of the same developmental pathway. This raises questions about the impact of deterministic and undirected forces on adaptive processes and the circumstances that define to what extent they contribute to evolutionary outcomes. Is divergence at the genetic level due to the accumulation of random mutations relevant for future evolution? In this study I was interested in the evolution of a novel trait, the stochastic switching between different phenotypic stages, that occurred during the Reverse-Evolution

Experiment (REE) in *P. fluorescens* and whether it can evolve repeatedly (see Chapter, section 1.4.4).

## 3.1.2 Evolution of a stochastically switching genotype in *P. fluorescens*

### *3.1.2.1 The rise of the switcher genotype*

Twelve lineages of *P. fluorescens* evolved parallel during the course of the REE. They started from the same ancestor *P. fluorescens* SBW25 and went through several selection rounds in which they were exposed either to static or shaken conditions (see Chapter1, Fig. 1.8). Environmental change led to the emergence of novel phenotypes (Rainey & Travisano, 1998). Under static conditions usually 'wrinkly spreader' (WS) types arose that can colonise the air-liquid interface and form a biofilm (see Chapter 1, Fig. 1.2). The propagation of a WS in a shaken environment reversed the phenotype often to a smooth, ancestral like colony morphology (SM; see Chapter 1, Fig. 1.3). Every phenotypic change was the result of mutational change. In two out of the 12 lineages (Line 1 and Line 6) extraordinary parallel phenotypic evolution was observed during the ninth selection round. Here a new type occurred independently in Line 1 and Line 6, which has not been observed in any other lineage before. The novel type appeared after eight mutations under static conditions. Subsequent genotypic and phenotypic analyses showed that the new types have characteristics of stochastically switching genotypes (Beaumont *et al.,* 2009; Gallie, 2009). The switcher genotype can randomly switch between two different cell stages (capsulated and non-capsulated, Fig. 1.9C), which lead to a distinctive colony morphology (opaque and translucent colonies and/or colonies with opaque sectors). It has been suggested that this is an adaptive innovation (Beaumont *et al.,* 2009; Gallie, 2009; Libby & Rainey, 2011; Rainey *et al.,* 2011).

### 3.1.3 The mutational history of the switchers in Line 1 and Line 6

The experimental design of the REE allowed the identification of the genetic changes that were responsible for the rise of different phenotypes throughout the evolutionary history of the 12 parallel lineages depending on the environment (static or shaken). Thus it was possible to reproduce the entire mutational history of the switcher genotypes. As already mentioned, the switchers in Line 1 and Line 6 appeared after eight preceding mutations (see Chapter 1, section 1.5). The ninth mutation in both lineages gave rise to the switchers referred to as 1w4 (Line 1) and 6w4 (Line 6). Both switchers are essentially identical. The fact that two out of 12 replicate lineages evolved this novel phenotype independently indicates a high degree of parallel phenotypic evolution. However when comparing the underlying genotypes (Tab. 3.1) it becomes apparent that the causal switcher mutations in both lineages are quite different. The switcher mutation in Line 1 occurred in the *carB* gene whereas in Line 6 it occurred in the *rpoD* gene (Tab. 3.1; Gallie, 2009).

**Table 3.1: Mutational history of the switcher type in Line 1 and Line 6.** After nine rounds of selection (either in a static or a shaken environment) a switcher type (SW) occurred in two out of 12 lineages under static conditions. Listed are the eight preceding mutations and the switcher mutations in Line 1 and Line 6.

| **Selection round** | **Environment** | **Line 1** | | **Line 6** | |
|---|---|---|---|---|---|
| | | Name | Mutational history | Name | Mutational history |
| 1 | static | 1w0 | *mwsR* | 6w0 | *wspF* |
| 2 | shaken | 1s1 | *mwsR* | 6s1 | *wspF* |
| 3 | static | 1w1 | *awsX* | 6w1 | *awsX* |
| 4 | shaken | 1s2 | *awsR* | 6s2 | *awsR* |
| 5 | static | 1w2 | *wspF* | 6w2 | *wspF* |
| 6 | shaken | 1s3 | *wssA* | 6s3 | *wssB* |
| 7 | static | 1w3 | *mwsR* | 6w3 | *nlpD* |
| 8 | shaken | 1s4 | *mwsR* | 6s4 | *nlpD* |
| **9** | **static** | **1w4 (SW)** | ***carB*** | **6w4 (SW)** | ***rpoD*** |

Although the mutational histories of both lineages diverge from each other there are some noticeable parallels between them. Line 1 and Line 6 have genetic loci in common that were affected by a mutation, such as *mwsR*, *aws*, *wspF* and *wss*, but the order of the mutation occurrences shows some variation (Tab. 3.1). It is thought that in most cases the genetic loci have a modular nature in which one mutation activates the WS phenotype (static environment) and a subsequent mutation in the same locus causes the phenotype to reverse to a SM ancestral type (shaken environment; see section 1.4.2). Interestingly, four genes were almost identical between Line 1 and Line 6. Here mutations occurred in *awsX* followed by *awsR*, and afterwards in *wspF* followed by *wssA* in Line 1, or *wssB* in Line 6 during the same selection round (Tab. 3.1).

Of note is the non-modular character of the *wspF* mutation followed by a mutation in *wssA* (Line 1) or *wssB* (Line 6; Tab. 3.1). Here the mutations occurred in different genetic loci, *wspF* and *wssA/wssB*. It is noteworthy that both are linked with each other via a signal transduction pathway (Bantinaki *et al.*, 2007). The *wsp* signal transduction pathway model was described in the introduction (see Chapter 1, Fig. 1.5). A mutation in *wspF* almost certainly abolishes its function as a negative regulator of WspR. As a result WspR is activated and produces c-di-GMP constitutively (Goymer *et al.*, 2006; Malone *et al.*, 2007). Since no deactivating mutation was detected in *wspR*, either in Line 1 or in Line 6 later on (Tab. 3.1), it is likely that all genotypes that occurred after the *wspF* mutation have an increased level of c-di-GMP. It has been shown that increased c-di-GMP synthesis has phenotypic consequences for the genotypes after the *wspF* mutation, such as the production of a cellulosic polymer. It is thought that WssA is responsible for the localization of the cellulose-producing unit within the cell and WssB is part of the cellulose synthase complex (Spiers *et al.*, 2002). Phenotypic analyses have shown that mutations in *wssA* or *wssB* decrease the amount of cellulose production, but the polymer is still produced by the cell (Gallie, 2009). It is likely that increased c-di-GMP concentration might influence the capsulation process with an effect on the switcher phenotypes in Line 1 and Line 6, which is similar to observations in other organisms, such as *Vibrio parahaemolyticus* (Boles & McCarter, 2002; Guvener & McCarter, 2003).

### 3.1.4 Replaying the tape of life in *P. fluorescens*

During the REE the same phenotypic innovation, a stochastically switching phenotype, emerged in two out of 12 replicate lineages of *P. fluorescens*. It has been shown that switching between different phenotypic stages is a winning strategy under the selective regime experienced by the bacteria (Beaumont *et al.*, 2009; Libby & Rainey, 2011; Rainey *et al.*, 2011). The fact that two lineages independently evolved a switcher shows that the evolution of this key innovation can be achieved repeatedly. It is not clear whether this is the result of history or deterministic processes. In both switcher lineages the new phenotype occurred after eight prior mutations and although the causative switcher mutation itself emerged in different genes it is known that the mutational histories overlap partly between the lines, indicating a possible impact of mutational history in the evolution of the switcher (Beaumont *et al.*, 2009).

To investigate these hypotheses I used the opportunity to go back in time, revive frozen bacterial cultures from a 'frozen fossil record', and 'replay' evolution from a particular timepoint. The aim of the experiment was to compare four independently evolved lineages of *P. fluorescens* (Line 1, Line 3, Line 6 and Line 12) in their ability to produce a switcher genotype. The four lineages evolved during the course of the long-term experiment REE (see Chapter 1, section 1.4.4) and each lineage had eight mutations that occurred over time, but the mutations differed in their order and partly in the genes that were affected by a mutation (Tab. 3.2). The immediate ancestors of the switcher in Line 1 (1s4) and Line 6 (6s4) are already known for their ability to evolve this novel type (Beaumont *et al.*, 2009; Gallie, 2009). Replaying evolution in two additional lineages that are not known for switcher evolution from the same starting position as the immediate ancestor of the switcher may clarify to what degree switching can repeatedly evolve. Line 3 and Line 12 were chosen as additional lineages because their mutational histories were except for one mutation in Line 12 completely known at the time when the experiment started. Knowledge of the individual mutational histories of each lineage may shed some light on the significance of the mutations and their order for switcher evolution.

Furthermore the reconstruction of a known switcher mutation (*carB**) into genetic backgrounds that did not produce switcher during the SREE tested for genetic constraints that might prevent a switcher mutation to arise within the population. Here the introduction of *carB** may not lead to the typical switcher phenotype (translucent/opaque colony morphology or colonies with opaque sectors) or *carB** results in a decreased fitness and is therefore unlikely to increase in frequency. In either case a switcher would not be detectable on an agar plate following dilution of a bacteria culture containing $10^9$ cells.

### 3.1.5 Objectives:

1. To test the hypothesis that four parallel lineages of *P. fluorescens* do not differ in their likelihood to evolve phenotypic switching under same environmental conditions. If this is true a Switcher Re-Evolution Experiment (SREE) in four independent linages (Lines 1, 3, 6 and 12) will show no significant variation in switcher occurrence.

2. To test the alternative hypothesis that due to the impact of evolutionary history the four parallel lineages differ in their likelihood to evolve phenotypic switching. In that case different lineages with different kind, number and order of mutations vary significantly in switcher occurrence.

3. To test the hypothesis that parallel phenotypic evolution of switching is the result of parallel genotypic evolution. In that case many of the evolved switcher can be expected to have a mutation in the same gene (e.g. *carB*).

## 3.2 Summary of material and methods

### 3.2.1 The Switcher Re-Evolution Experiment (SREE) in four independent lineages

The experimental conditions were as described for the original REE (see Chapter 1, section 1.4.4). Twelve independent lineages of *P. fluorescens* were propagated in alternately static or shaken environments for three days (one selection round). Selection in contrasting environments promotes diversification, with new types being evident by virtue of distinctive colony morphologies on agar plates. A single colony of the most common new type was chosen at random and used to inoculate a new microcosm and was propagated under the opposing environmental regime for another three days (next selection round). The procedure was repeated for several selection rounds. If no new type occurred within the first three days, a small volume of the bacterial culture was transferred to a new microcosm and exposed to the same conditions for an additional three days (first transfer in the same selection round). This procedure was repeated until a new type was found (several transfers in one selection round; Fig. 1.8). Comparative sequencing of the newly evolved types revealed the underlying mutation for the evolution of the new type in each selection round.

In this study four out of the 12 REE lineages were chosen (Line 1, Line 3, Line 6 and Line 12) to carry out the SREE. Four genotypes that evolved independently in the same selection round and had eight mutations (1s4 in Line 1, 3s4 in Line 3, 6s4 in Line 6 and 12s4 in Line 12) were revived from -80°C freezer stocks on KB agar plates. The genotypes 1s4 and 6s4 are the immediate ancestors of the switcher genotype that evolved during the REE and had the causative mutations in *carB** (1w4) and *rpoD* (6w4; see Chapter 1, section 1.5).

The SREE was carried out in three blocks, with 18 replicates per lineage in Blocks I and II, and 20 replicates in Block III (in total 56 replicate microcosms per lineage). For each replicate an individual colony was chosen and singly cultured overnight under shaken conditions. Six microliter from every overnight culture were used to inoculate new microcosms, which were left undisturbed for three days (one round of selection). Each replicate culture was then diluted and plated on four KB agar plates to ensure the screening of 500–1000 colonies. After two days the plates were screened for colonies with a new morphology. Any replicate that produced a 'new type' was terminated (Fig. 3.2).

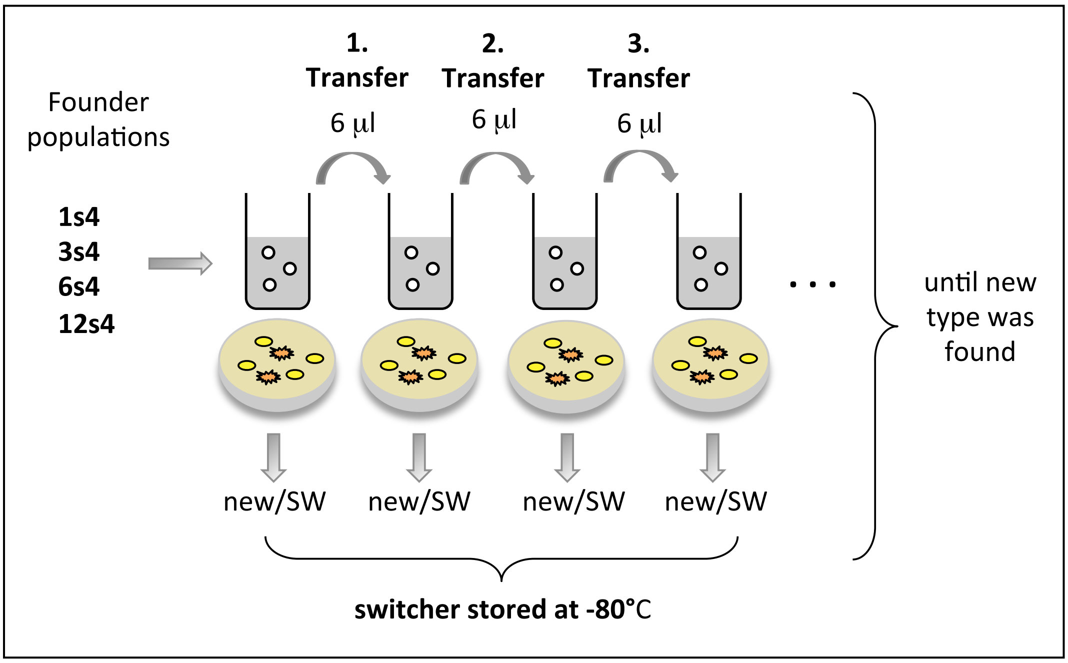

**Figure 3.2: Experimental design of the SREE.** Overnight cultures of 1s4, 3s4, 6s4 and 12s4 acted as founder populations to replay the evolution of phenotypic switching under static conditions (in total 56 replicate microcosms). After three days growth in a static microcosm the cultures were diluted, plated and screened for new colony types. If a new type or a switcher (new/SW) was detected the replicate was terminated. Switchers were stored at –80°C. If no new type was found 6 µl of the culture were transferred to a new microcosm and incubated for another three days under static conditions before being diluted, plated and screened for new types again. This process was repeated until a new type was found.

Any new colony type showing features of a switcher genotype (i.e. opaque and translucent colony formation or opaque sectors within a colony) was stained with Indian ink and examined under the microscope for increased capsule production and stored at -80°C. In addition, to check for heritability of the new trait, a small part of the colony was transferred onto new agar plates several times in a row. In cases where no new type was detected on the four agar plates, 6 μl of the culture were used to inoculate a new microcosm and incubated for another three days (first transfer) under static conditions and then diluted and plated again. The procedure was repeated (several transfers) until a different colony type was detected (Fig. 3.2).

## 3.2.2 Molecular methods

The following molecular methods were used in this study as well as in other chapters of this thesis. They are described in detail in the material and method chapter (Chapter 2).

### *3.2.2.1 Comparative sequencing of the carB gene*

Polymerase chain reaction (PCR) with forward primer SN039 PF and reverse primer SN049 PR (see Chapter 2, Tab. 2.3) was used to amplify a DNA fragment of the re-evolved switcher genotypes (SREE) from 1s4, 3s4, 6s4 and 12s4. The fragment included the *carB* gene. The PCR product was purified and sequenced using *carB* specific primers (SN009 PF, SN010 PF, SN013 PF, SN014 PF, SN015 PF, SN016 PF, SN017 PF, SN018 PF and SN070 PR; Tab. 2.3) and later compared to the ancestral DNA sequence of *P. fluorescens* SBW25 using the Sequencher 4.7 Demo software (Gene Codes Corporation, Ann Arbor, MI, USA).

### *3.2.2.2 Introduction of* carB* *into the genome of 12s4*

The *carB** mutation that occurred *in vitro* in Line 1 after eight mutations (Beaumont *et al.*, 2009; Gallie, 2009) was introduced into the 12s4 genome to gain insights as to whether the genotype 12s4 has the capacity to produce a switcher phenotype. It was tested whether the single point mutation (*carB**) is sufficient to induce a switching phenotype in this genetic background. The first step was to obtain a DNA fragment containing the target point mutation, *carB**, from the 1w4 (which is the name of the REE switcher lineage) for amplification (see Chapter 2, section 2.4.1). A TOPO cloning reaction was carried out to introduce the *carB** fragment into the vector pCR8 (see Chapter 2, section 2.4.2). The vector pCR8 containing the *carB** fragment was digested and the *carB** fragment was then introduced into the vector pUIC3 (see Chapter 2, section 2.4.3). Electrocompetent cells of *E. coli*DH5αλpir (see section *2.4.3.1*) were used to transform pUIC3 containing the *carB** fragment into *E. coli* (see Chapter 2, section *2.4.3.2*). The *E. coli*-puIC3*carB** was used as the donor strain in a bacterial conjugation with the recipient being *P. fluorescens* 12s4 to exchange the ancestral DNA fragment with the *carB** fragment (see Chapter 2, section 2.4.4). A PCR product with the primer pair SN004 PR and SN006 PF (see Chapter 2, Tab. 2.3) confirmed the successful introduction of *carB** into the genome and sequencing ensured no additional mutation.

### *3.2.2.3 Construction of a* lacZ *transcriptional fusion in 12s4*

Neutral marking of bacterial strains by introducing a *lacZ* transcriptional fusion was used for blue/white screening of mixed bacterial cultures on agar plates containing Xgal. This was achieved by using a mini-Tn7 transposon system (see Chapter 2, section 2.4.5). The PCR with primers SN007 PR and SN008 PF (see Chapter 2, Tab. 2.3) tested for the site-specific insert of the right size.

#### *3.2.2.3.1 Capsule count assay for 12s4 and the manufactured 12s4*carB* *strain*

The purpose of this experiment was to determine the proportion of capsulated cells in 12s4*carB** and 12s4 using staining techniques and microscopy. For each strain five replicate KB microcosms containing 6 ml liquid KB were inoculated with a single colony each and incubated overnight. The capsule count assay was carried out following the protocol described in section 2.5.1.

#### *3.2.2.4 Determination of relative fitness of* carB* *in 12s4*

A competition experiment was performed between 12s4 and 12s4*carB** to determine fitness consequences of *carB** in 12s4. Both competitors were grown together for 48 hours in a static microcosm ($N$ = 10; see section *2.5.2.1*). Colony counts at the beginning of the competition and at the end were used to calculate the Malthusian parameters (see Chapter 2, section *2.5.2.2*).

### 3.2.3 Statistical analysis

The results of this study were analysed using the statistical software JMP 9. A general linear model (ANOVA) was used to identify the factor that explained most of the variance in switcher evolution between the lineages. Shapiro-Wilk tests and one-sample $t$-tests were performed in R (R.app GUI 1.23). Shapiro-Wilk was used to test for normal distribution and $t$-test for the significance of the fitness effect of *carB** in 12s4. Two-sample $t$-test was used to reveal differences in capsulation between 12s4 and 12s4*carB**. Values were arcsin transformed. Furthermore the program R was used to illustrate the results of the different $t$-tests. Graphs were partly created in Prism.

## 3.3 Results

The genotypes 1s4, 3s4, 6s4 and 12s4 acted as founder populations for replaying evolution of a switcher genotype under static conditions (SREE; Fig. 3.2). The founder populations evolved during the original REE (see Chapter 1, section 1.4.4) in four independent lineages (Line 1, Line 3, Line 6 and Line 12) and had each experienced eight rounds of the selection regime and thus carried the same number of mutations (eight in total; Tab. 3.2).

**Table 3.2: Mutational history of four lineages from the REE.** The history table shows the name of the strain (e.g. Line 1 in selection round 4: 1s2), selection round, the selective environment and the gene where the mutation occurred. Line 1 and Line 6 evolved a switcher genotype (SW) after eight selection rounds due to two different mutations, *carB* (Line 1) and *rpoD* (Line 6). Line 3 and Line 12 are not known for their ability to evolve a switcher genotype. The mutations that occurred in Line 12 in selection rounds 8 and 9 are still unknown. The “s4” types represent the time point from which the SREE started.

| **Name** | **Selection round** | **Environment** | **Line 1** History | **Line 6** History | **Line 3** History | **Line 12** History |
|---|---|---|---|---|---|---|
| w0 | 1 | static | *mwsR* | *wspF* | *wspC/D* | *awsR* |
| s1 | 2 | shaken | *mwsR* | *wspF* | *wspC* | *awsR* |
| w1 | 3 | static | *awsX* | *awsX* | *mwsR* | *wspF* |
| s2 | 4 | shaken | *awsR* | *awsR* | *mwsR* | *wspC* |
| w2 | 5 | static | *wspF* | *wspF* | *mwsR* | *mwsR* |
| s3 | 6 | shaken | *wssA* | *wssB* | *PFLU0185* | *mwsR* |
| w3 | 7 | static | *mwsR* | *nlpD* | *PFLU0185* | *wspC/D* |
| s4 | 8 | shaken | *mwsR* | *nlpD* | *mwsR* | ? |
| w4 | 9 | static | ***carB*** **(SW)** | ***rpoD*** **(SW)** | *awsR* | ? |

Nonetheless the type and order of the mutations varied between them. The SREE in this study was carried out in three independent blocks, with 56 biological replicates for each lineage. The factor ‘block’ did not have a significant effect on the likelihood of evolving a switcher ($F_{2,9}$ = 0.147, $P$ = 0.86), and was thus dropped from the model.

## 3.3.1 The SREE from genotypically different founder populations

### *3.3.1.1 Average total switcher production in 1s4, 3s4, 6s4 and 12s4*

The plot in Figure 3.3 displays the average number of switchers that were found in a total of 56 replicate microcosms during the SREE of each lineage in three experimental blocks ($N$ = 3). The most switchers were found in 1s4, which is shown in the first bar plot (Fig. 3.3). Here on average seven out of 20 microcosms produced a switcher. Replaying evolution from the genotypes 3s4 and 6s4 led on average to two switchers in both backgrounds. No switcher was detected in 12s4 (Fig. 3.3). Most of the observed variance was explained by the lineage factor ($F_{3,8}$ = 22.3, $P < 0.001$, $R^2 = 0.89$) and a Tukey-Kramer posthoc test over all four strains identified 1s4 as different from 3s4 ($P$ = 0.002), 6s4 ($P$ = 0.002) and 12s4 ($P <$ 0.001) in its average switcher occurrence. The lineages 3s4, 6s4 and 12s4 showed similar behaviour.

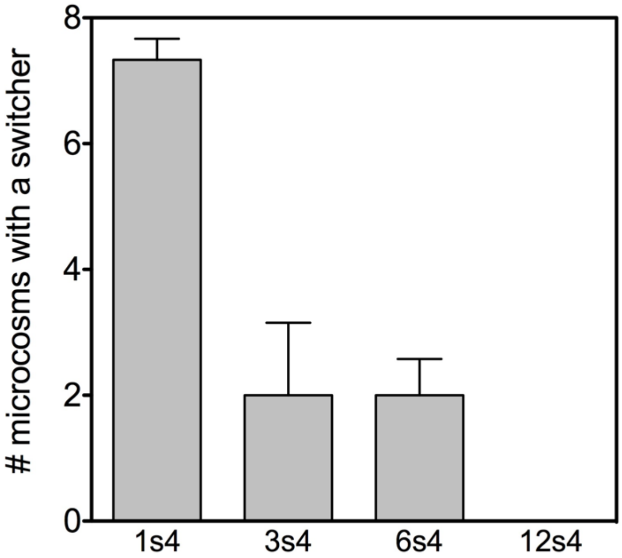

**Figure 3.3: Average number of replicates of 1s4, 3s4, 6s4 and 12s4 that evolved switchers.** The switchers evolved within three days under static conditions. Displayed are SEM based on $N$ = 3.

### *3.3.1.2 Time taken to detect switchers in each strain*

The four lineages 1s4, 3s4, 6s4 and 12s4 differed in the time needed for a switcher genotype to arise. All re-evolved switchers in 6s4 occurred within the first three days of the experiment. In 3s4 one transfer at maximum was necessary to see a switcher on the agar plates. Both strains showed the new phenotype rather quickly in comparison with 1s4 (Fig. 3.4). In 1s4 it took up to five transfers for a phenotypic switcher to be detected. In addition, in 1s4 the number of microcosms that had a switcher increased with the number of transfers. Again re-evolution from 12s4 did not result in a detectable phenotypic switcher (Fig. 3.4). If any kind of new type was found on an agar plate the replicate was terminated.

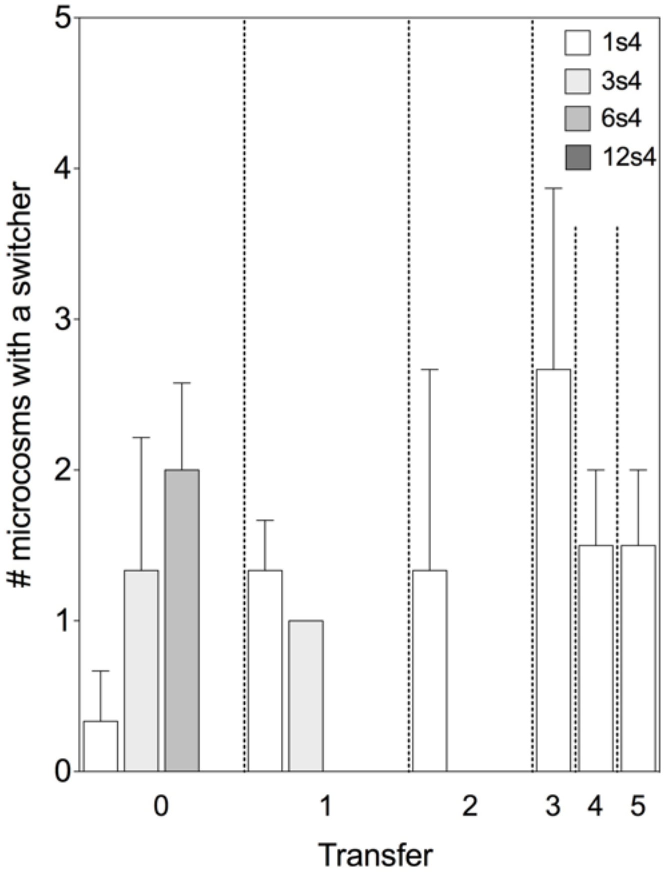

**Figure 3.4: Average number of microcosms of 1s4, 3s4, 6s4 and 12s4 with switchers after each transfer.** The first transfer was performed after the initial three days of bacterial growth in a static environment (zero transfer) if no new type was detected on

agar plates. The new microcosms were then incubated for a further three days under static conditions and plated and checked for new types and switchers. The bacterial cultures were transferred until a new type was found. Displayed are SEM based on $N$ = 3. Vertical dotted lines separate different transfers.

### *3.3.1.3 Time taken to detect new phenotypes in each strain*

The strains 3s4, 6s4 and 12s4 showed the highest average number of microcosms with newly evolved types within the first three days in a static environment (75% of the microcosms). Only a few replicates of 3s4 needed one transfer for a novel colony type to appear. Some replicates of 6s4 and 12s4 were transferred twice before a type with new colony morphology was detected (Fig 3.5).

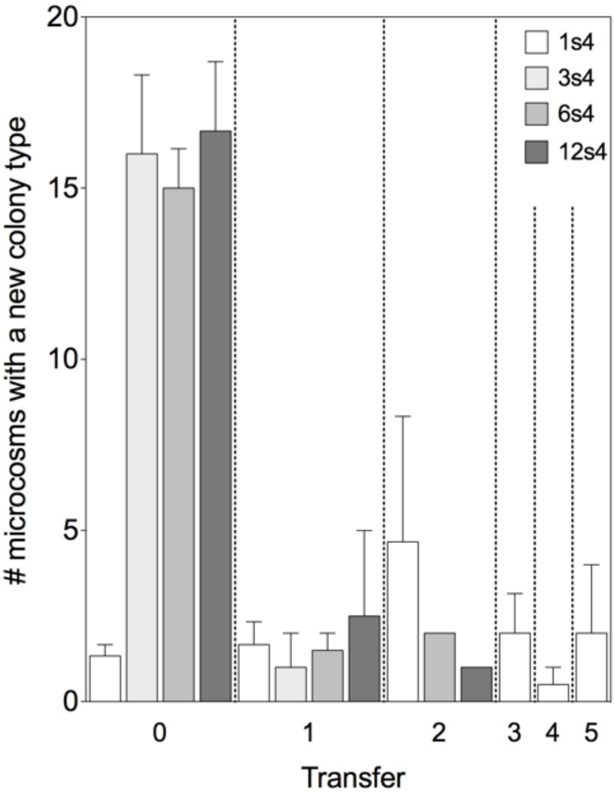

**Figure 3.5: Average number of microcosms of 1s4, 3s4, 6s4 and 12s4 with a novel phenotype after each transfer.** Every three days the cultures were transferred to new microcosms, and screened for new colony morphologies. This process was repeated until a new type was found. Illustrated are the average numbers of microcosms that had a novel

type after each transfer. Displayed are SEM based on *N* = 3. Vertical dotted lines separate different transfers.

Re-evolution of novel types from 1s4 started on average with a small proportion of the microcosms (less then 10%), which stayed on average below five during the entire experiment. All replicates of 1s4 showed a new type after the fifth transfer (Fig. 3.5).

### 3.3.2 Genotypic analysis of re-evolved switcher genotypes

Next it was checked whether the detected switcher colony morphology of the re-evolved switchers were caused by a mutation within the *carB* gene. The *carB* gene was chosen because *carB** was the underlying mutation of the switcher genotype that evolved during the REE (1w4) and previous replay experiments from 1s4 showed a tendency of repeated switcher evolution based on mutations in the *carB* gene (Gallie, 2009). But the results were less powerful due to a small replicate number (N = 36). Therefore, the *carB* gene of all re-evolved switchers was amplified, sequenced, and compared to the ancestral *carB* sequence.

The results of the SREE from 1s4, 3s4, 6s4 and 12s4 showed that from 22 phenotypic switchers that were found in 1s4, only two had a mutation in the *carB* gene. In 6s4, none of the six observed switchers had a mutation in the *carB* gene, and in 3s4 two out of the six switchers had a mutation in the *carB* gene. No phenotypic switcher was found in 12s4.

### 3.3.3 Phenotypic characterization of 12s4*carB**

Genetic analysis and fitness assays were used to gain information as to whether Line 12, where no switcher was observed, had the capacity for a switcher type to arise in this genotype. A mutation in the *carB* gene (*carB**) was causal for the evolution of one of the switchers in the original REE. It is known from previous experiments that *carB**, when introduced into the wild type SBW25 genome, is sufficient to induce a stochastically switching phenotype (Beaumont *et al.*, 2009; Gallie, 2009). The typical colony morphology, such as the co-occurrence of translucent and opaque colonies on an agar plate or colonies with opaque sectors, was found in all lineages except 12s4. To exclude the possibility that a switcher mutation like *carB** occurred in 12s4 during the SREE, but did not lead to the typical switching phenotype (and thus was not detected), the *carB** mutation was introduced into the 12s4 genome.

After the introduction of *carB** into the genome of 12s4, the typical switcher colony morphology was observed on KB agar plates. Colonies showed opaque sectors, indicating switching behaviour. In addition, the percentage of capsulated cells within populations of 12s4*carB** significantly increased compared to 12s4 (two-sample *t*-test: $t = -11.29$, d 5.36, $P < 0.0001$ (arcsin transformed); Fig. 3.6).

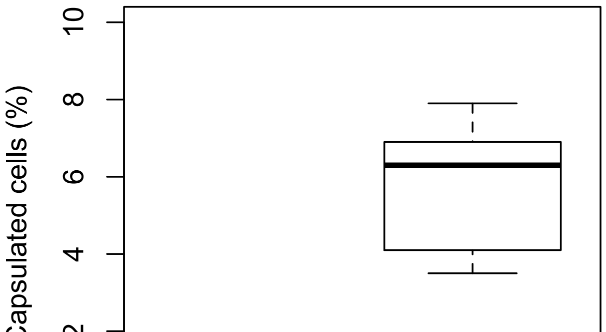

**Figure 3.6: Percentage of capsulated cells in populations of 12s4 and 12s4*carB**.** Illustrated is the average percentage of capsulated cells within populations of 12s4 and 12s4*carB**. Displayed are medians and interquartile ranges based on $N = 5$.

#### 3.3.3.1 *Fitness effect of* carB* *on 12s4*

It was shown in the previous paragraph that the introduction of *carB** induced the switcher phenotype. The fitness assay was performed to test whether the *carB** switcher mutation would have fitness consequences for 12s4. Therefore 12s4 was competed in a head to head competition against 12s4*carB** over 48 hours in a static microcosm. Malthusian parameters were used to calculate the relative fitness of *carB** in 12s4 (see Chapter 2, section *2.5.2.2*). The fitness of 12s4*carB** was significantly greater than zero (one-sample *t*-test: $t$ = 4, df = 9, $P$ < 0.005), indicating a fitness increase caused by *carB** in 12s4 (Fig. 3.7).

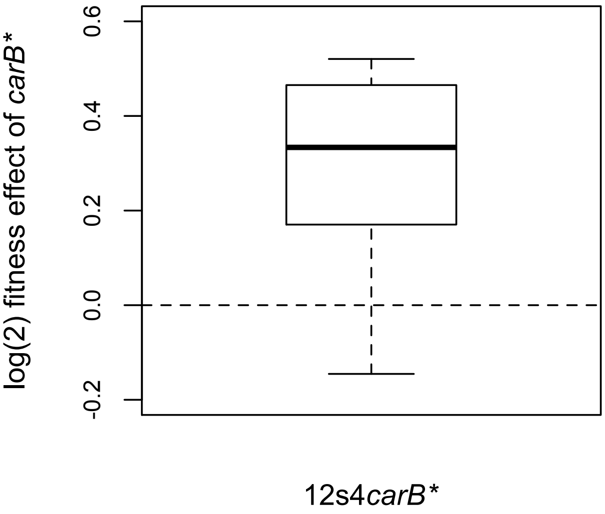

**Figure 3.7: Relative fitness of *carB** in 12s4.** Fitness is specified as log(2) of the Malthusian parameter of 12s4*carB** to the Malthusian parameter of 12s4. The dashed line at zero indicates no fitness difference between 12s4*carB** and 12s4. Displayed is the median and interquartile range based on *N* = 10.

## 3.4 Discussion

There is an ongoing debate about the repeatability and predictability of evolution and the effect of evolutionary history on phenotypic and genotypic evolution.

Conway Morris (2010) states that organisms are reproducing, self-organised units made up of molecules that interact with each other according to natural laws and economics. He believes that these circumstances of life under similar environmental conditions can drive the convergence of innovations and make evolution most likely predictable (Vermeij, 2006; Conway Morris, 2010). At the same time he admits that there is margin for exceptions. Morris acknowledges that the evolution of citrate utilisation during Lenski's long-term evolution experiment in *E. coli* is a noticeable case of contingency (Blount *et al.*, 2008; Blount *et al.*, 2012). This key innovation occurred in one out of 12 parallel lineages that started from the same ancestor and grew for more than 30,000 generations in glucose-limited but citrate-rich media. Citrate utilization was dependent upon previous mutations due to epistasis, which increased the likelihood of the occurrence of the novel trait. However, Morris questions the general validation of Lenski's results (Conway Morris, 2010). On the other hand, some scientists believe that evolution of innovations is unpredictable because of random factors (random variations, random drift, unpredictable environmental change) (Lewontin, 1966; Monod & Wainhouse, 1971; Gould & Lewontin, 1979; Gould, 1989; Mani & Clarke, 1990). They maintain that the future states of molecules, organisms, species and environments and how they interact with each other is not predictable even over the short term (Kauffman, 2000).

In this study I was interested in whether four independently evolved lineages of *P. fluorescens* differ in their likelihood of evolving a switcher genotype and whether switching can occur repeatedly in all backgrounds. Furthermore I wanted to know whether the order of mutations in each lineage has an impact in the evolution of this novel phenotype. I used the potential of rewinding and replaying evolution from an earlier time point by reviving bacterial cultures of four lineages (Line 1, Line 3, Line 6 and Line 12) from a 'frozen fossil record' from a particular time point (1s4, 3s4, 6s4 and 12s4; Fig. 3.2).

### 3.4.4 Parallel evolution of phenotypic switching in four independent lineages of *P. fluorescens*

Mutations, the basis of variation and innovation, are random and in theory they have an equal chance of arising at any time point during evolution unless evolutionary history promotes or constrains their likelihood of occurrence. Line 1 and Line 6 produced a switcher in a static environment during the course of the REE due to mutations in *carB* and in *rpoD* (Tab. 3.1). Two out of 12 replicate lineages showed parallel phenotypic evolution but the underlying genetic changes were slightly different (Beaumont *et al.*, 2009; Gallie, 2009). Both lineages acquired mutations in some of the same genes during the selection experiment – even at some of the same time points – although no two mutations were identical (Tab. 3.1). In addition, both switcher mutations occurred well into the evolutionary pathway (after eight mutations) during the same selection round, suggesting that only then could the mutation arise and establish a switcher genotype in the population, perhaps due to dependency upon their mutational history.

When 'replaying evolution' in Lines 1 and 6 (REE switcher lineages) and also in Lines 3 and 12, where no switching type was previously observed during the original REE, switchers were found in Line 1 and Line 6, which was expected, and also in Line 3, but not in Line 12. This shows that the evolution of switching can repeatedly occur and is not restricted to Line 1 or Line 6. Switching can occur in a diverse range of genotypic backgrounds. Three out of four lineages were able to produce a switcher genotype. This extent of phenotypic parallelism in the evolution of switching and the differences in the underlying genotypes suggests that multiple pathways can produce the same phenotype, which coincides with results from other evolution experiments with microorganisms (Lenski *et al.*, 1991; Korona *et al.*, 1994; Lenski & Travisano, 1994; Korona, 1996; Nakatsu *et al.*, 1998; Wichman *et al.*, 1999; Cooper & Lenski, 2010; Meyer *et al.*, 2012). Additionally, these studies emphasise that the evolution of a similar phenotype leads to slightly different evolutionary 'endpoints'.

### *3.4.4.1 Multiple pathways can lead to stochastic switching*

Switcher genotypes can arise as the result of a single mutation in different loci, *carB* or *rpoD* (Beaumont *et al.*, 2009; Gallie, 2009). It is therefore likely that re-evolved switchers from the SREE occurred due to mutations in different loci, such as *carB* or *rpoD*. Indeed, the sequencing results show that only four out of 32 re-evolved switchers in all four lineages had a mutation in the *carB* gene. The remaining 28 re-evolved switchers can be attributed to mutations somewhere else in the genome. Whether a mutation in *rpoD* was a causative switcher mutation like in Line 6 has not yet been investigated. Genotypic evolution seems to be less predictable than the phenotypic evolution of a switcher. It is known that *carB* and *rpoD* are involved in the bistable switch and the synthesis of colanic acid in *P. fluorescens*. The colanic acid pathway is complex and linked to the pyrimidine synthetic pathway. Apart from mutations in *carB* and *rpoD* there are more mutational targets within the pyrimidine pathway, such as the *pyrH* gene (Gallie, 2009), that can lead to a switching phenotype. Even given that the mutations occur in different genes, the functional change due to the mutations is likely to be similar. Here the flux through the pyrimidine pathway is reduced in the cell and as a consequence some cells of the population produce colanic acid (Gallie, 2009). It is likely that there are even more genes within the pathway that can enable parallel phenotypic evolution. Metabolic pathways and networks can be involved in evolution by providing multiple accessible mutational pathways via a variety of genes (Pelosi *et al.*, 2006). The results of this study show that the evolution of a switcher is not a rare event and can be observed in a variety of different genotypic backgrounds.

### *3.4.4.2 Impact of mutational history on switcher evolution*

Even though Lines 1, 3 and 6 all showed the fundamental capacity to evolve a switcher, the lineages varied in their likelihood of switcher occurrence. Line 1 showed the highest chance and was very different from Line 3 and Line 6 (Fig. 3.3).

In Line 1 a third of the microcosms of 1s4 produced a switcher, whereas in Line 3 and Line 6 on average only ten percentages of the replicates had a switcher (Fig. 3.3). This suggests that although the three lineages were able to produce the same novel phenotype, which has been shown to be a beneficial strategy (Beaumont *et al.*, 2009), there are factors that are likely to influence the chance of switcher occurrence. The genetic composition and the order of mutations, which was different between the switcher lineages, may constrain or support the establishment of a switcher mutation within the population. For example, the genotype of Line 1 might favour the evolution of phenotypic switching perhaps due to an earlier evolutionary event (Beaumont *et al.*, 2009). During the long-term evolution experiment with *E. coli* in glucose-limited medium, 12 parallel lineages evolved a similar phenotype after circa 10,000 generations, showing a higher fitness than the ancestor and increased cell size (Lenski & Travisano, 1994). The between-lineage comparison revealed that even with all lineages converging towards a similar phenotype there was still some degree of divergence, which was attributed to differences in the underlying genetic composition due to random mutations (Korona *et al.*, 1994; Korona, 1996). The lineages evolved for an additional 20,000 generations before a key innovation, the ability to utilise citrate, appeared in one of the 12 lineages. It was shown that the evolution of this new phenotype was dependent on a mutation that occurred earlier and led to epistatic interactions, which increased the likelihood of occurrence (Blount *et al.*, 2008; Blount *et al.*, 2012). This example demonstrates how the mutational history, i.e. the order of spontaneous mutations that accumulate over time, can determine the outcome of evolution.

Epistatic effects due to an earlier mutation may promote the evolution of switching in 1s4 because of a large fitness effect. In addition, the rate of beneficial mutations could be decreased in 1s4, which could explain the increased number of transfers that were needed to evolve a new type in general. The lack of beneficial mutations increases the chance of a switcher mutation becoming fixed in the population even if the fitness benefit is smaller. In that case the timing of appearance (Hegreness *et al.*, 2006) determines whether a switcher can reach high frequencies within the population and consequently can be found on agar plates following dilution of cultures containing $10^9$ cells.

### 3.4.5 Differences in appearance time of novel types and switchers

Further differences between the three switcher lineages become apparent when comparing the number of transfers that were necessary for new types to arise. Most of the replicates of 3s4 and 6s4 showed a new type already before the first transfer, whereas some replicates of 1s4 had to be transferred to new media five times before the detection of any new type (Fig. 3.5). Evolution from 1s4 was much slower than in the other lineages. It is unlikely that 1s4 evolved a decreased mutation rate that slows down the supply of variation in general and therefore the occurrence of novel phenotypes. Presumably new variants occurred in 1s4 but could not reach a cell frequency high enough to be observed on an agar plate. Results from previous studies support this hypothesis. Phenotypic analyses have shown that 1s4, the immediate ancestor of a switcher in the REE, produces a cellulosic polymer and can build a biofilm on the surface of the microcosm (Gallie, 2009). Hence, a new variant has to compete with 1s4 that is already growing at the air-liquid interface and occupying the niche. Under these circumstances the new type stays unrecognised on agar plates because it cannot increase in cell numbers. In which case a switcher can have a higher chance to arise in the population, but more time is needed before it would be observed.

Interaction between different genotypes and population dynamics within the microcosm (Stern & Orgogozo, 2009) could also explain the decreased likelihood of switcher occurrence in Line 3 and Line 6. Both lineages showed a high proportion of microcosms with a new type after a short time. This indicates a fast supply of beneficial mutations. This could have limited the ability of a switcher increasing in frequency because other new types arising before a switcher already occupied the niche, i.e. air-liquid interface. Hence, interactions between the switcher and other new types, the appearance time and competition ability determine which novel type will increase in cell numbers (Hegreness *et al.*, 2006).

In the *P. fluorescens* experiment a microcosm becomes heterogeneous due to an oxygen gradient and this provides ecological opportunities for novel types (Rainey & Travisano, 1998). Sometimes multiple new types arise within three days and often the ancestor is still present. Little is known about genotypic interactions and the population dynamics within the microcosm. Nonetheless, the number of new types, the time that they need to arise and how they interact during the SREE are likely to depend on the present genotype in Lines 1, 3 and 6, which also can explain the differences in likelihood of switcher occurrence, although switchers in general can occur in every background.

### 3.4.6 Constraints in the evolution of a switcher genotype in Line 12

Line 12 partly overlaps with the mutational history of the other three lineages but also deviates in order and some genes, but nonetheless switcher evolution was not observed. In addition, one mutation of the evolutionary history has not been identified yet (Tab. 3.2), which might be a crucial evolutionary event that limits the establishment of a switcher genotype within the population. The lack of switcher occurrence in Line 12 might be based upon genetic constraints and raises the question whether the appearance of a switcher mutation would actually lead to a typical switcher phenotype with switcher colony morphologies on agar plates (opaque/translucent colonies or colonies with opaque sectors) and an elevated proportion of capsulated cells. The introduction of *carB** into the 12s4 genome demonstrates that this mutation is sufficient to induce the phenotype. In addition, the results from the fitness assay between 12s4*carB** and 12s4 showed that 12s4*carB** had a higher fitness than 12s4 in a static environment, meaning that a switcher based on a mutation in the *carB* gene would be able to increase in frequency in this background. In theory a switcher based on a *carB** mutation has a realistic chance to arise within the population. Therefore the absence of switchers in Line 12 cannot be attributed just to underlying genotypic constraints. The genotype 12s4 was able to diversify in a static environment rather quickly, suggesting that there were no constraints that reduced overall diversification in this background. It could have happened just by chance that no switcher was found in 12s4. A suitable experiment to test this is an additional SREE with 12s4 with a high number of replicates. If the lack of switcher occurrence was coincidental it is likely that a switcher will be detected in some replicates. Alternatively, interaction between different genotypes within the microcosm could explain the lack of switcher detection. It has been shown that 12s4 quite quickly produces new types (Fig. 3.5) that are likely to occupy the air-liquid interface. If the new switcher is not able to compete, perhaps due to lower competitive ability against other new variants or a late appearance time of the causative mutation (Hegreness *et al.*, 2006), then the switcher cannot increase in frequency and stays undetected on agar plates. The rate in which beneficial mutations occur and the effective selective value of the mutation, i.e. genetic and ecological circumstances, determine whether a novel phenotype becomes established within a population or not.

## 3.5 Summary

Three hypotheses were tested in this study (see section 3.1.5). The first hypothesis 1) suggested that four independent lineages of *P. fluorescens* do not differ in their likelihood of evolving a switcher. Overall the results of this study show that the evolution of a switcher genotype is not unique in *P. fluorescens* and can occur in three out of four independently evolved lineages. Multiple routes shaped by different mutations and mutational order can lead to the switching phenotype. The likelihood of switcher evolution however varied between the lineages and therefore this hypothesis is rejected. The results of this experiment however support the alternative second hypothesis 2), the impact of evolutionary history on the evolution of a switching phenotype. Here even though Lines 1, 6 and 3 showed in general the ability to produce a switcher there were noticeable differences between the lineages such as the time that was needed to produce a new type or a switcher, the average number of switchers and the number of new types. Line 1 showed the highest number of evolved switcher and the highest number of transfers that were needed to produce a switcher indicating that the likelihood of switcher occurrency changes with different mutations and a different order of mutations (evolutionary history) in the different lineages. In addition, no switcher was observed in Line 12. This suggests that there are differences between the lineages in their likelihood of switcher evolution due to the genetic background since the environmental conditions were kept similar (static conditions). The third hypothesis (3) was rejected. Sequencing results revealed that causative mutations in the re-evolved switchers are not restricted to the *carB* gene; there are other mutational targets somewhere else in the genome. Parallel phenotypic evolution was not originated by underlying parallel genotypic evolution and is not genetically constraint.

This study of switcher evolution provides an example of parallel evolution leading to the same phenotype. Switching is somehow predictable at the phenotypic level and it can be expected that many lineages of *P. fluorescens* are able to produce a switcher innovation in a static environment. However, the prediction of the genetic routes that can be taken is more difficult since the mutational target size is presumably large. It is difficult to determine to what extent deterministic factors and random factors influence the evolution of a switcher genotype since both are tightly linked, and factors such as the biotic and abiotic environments have to be taken into account. In addition, the results show that multiple evolutionary outcomes are possible, and switching is only one of the possible solutions.

# Chapter 4: Impact of history on the evolution of a switcher genotype in *P. fluorescens*

## 4.1 Introduction

Heritable variation amongst organisms is the foundation for evolution by both drift and natural selection and is central to the emergence of novel traits. In addition, the appearance of a novel trait depends strongly on given environmental conditions. Genetic variation is based on randomly occurring mutations. Evolutionarily relevant phenotypes can be caused by a single mutation (Hall, 1981; Negoro *et al.*, 1994; Hoekstra *et al.*, 2006; Dobler *et al.*, 2012) or can be based on multiple mutations that accumulate over time. Here it is possible that evolution is a stepwise process and each mutation contributes individually to the success of the novel trait. In that case the successful establishment of the new type depends on the overall mutational history. In some cases the chronological order in which the mutations occur along the evolutionary pathway is crucial (Bridgham *et al.*, 2006; Miller *et al.*, 2006; Ortlund *et al.*, 2007), but other examples exist where the chronology is not important for the evolutionary outcome (Wichman *et al.*, 1999). Both scenarios show that the overall history is necessary and that the novel trait could not evolve at an earlier evolutionary time point from an intermediate genotype. Wichman and colleagues (1999) investigated two virus populations that evolved under strong selection in a novel host. Although both populations shared a large proportion of the mutations, the order in which they occurred differed between them. Here the overall fitness was perhaps the sum of the fitness effects of each individual mutation (Wichman *et al.*, 1999). The evolution of a key innovation may depend on mutations at two or more different loci. On their own, the mutations have no significant effect on the novel phenotype, but in combination they can be deleterious or highly beneficial. Such epistatic interactions are a common phenomenon in evolution (Hall, 2002; Miller *et al.*, 2006; Weinreich *et al.*, 2006; Blount *et al.*, 2008; da Silva *et al.*, 2010; Khan *et al.*, 2011; Salverda *et al.*, 2011; Tenaillon *et al.*, 2012).

One central question in evolutionary theory is how accessible or constrained evolutionary trajectories for novel traits are and whether this accessibility changes throughout progressing evolution? Can novel traits based, for example, on a single mutational step evolve at any time during the course of evolution or does the evolutionary accessibility change with further mutations that accumulate during the course of evolution?

### 4.1.1 Impact of mutational history on the evolution of novel traits

Scientists have traditionally focused on the outcome of evolution to gain deeper understanding about evolutionary mechanisms. Over the last two decades scientists have become more and more interested in the genetic basis of adaptive processes. The increased availability of molecular tools and sequencing data provide the opportunity to investigate not only the end product of evolution but also the intermediate stages that occur along evolutionary pathways. Scientists are able to recreate evolutionary pathways by reconstructing relevant mutations in a particular order in the laboratory and test their likelihood of occurrence in different genetic backgrounds. The generation of so-called empirical fitness landscapes is based on fitness values that mutations have when introduced into an artificially created genetic arrangement of other mutations. This approach is a useful tool to determine, for example, how many routes are accessible for evolution of a certain trait or to identify the most likely adaptive walk of a population (Weinreich *et al.*, 2006; Ortlund *et al.*, 2007; Reetz & Sanchis, 2008; Novais *et al.*, 2010). In addition they can test how evolutionary histories can change evolutionary dynamics over time and what mechanisms constrain or shape the fitness landscape and make pathways accessible. For example empirical fitness landscapes can reveal pleiotropic effects or epistatic interactions between mutations (Elena & Lenski, 1997a; Lunzer *et al.*, 2005; Weinreich *et al.*, 2006; Ortlund *et al.*, 2007; O'Maille *et al.*, 2008; Khan *et al.*, 2011).

The number of studies looking at empirical fitness landscapes is growing. Weinreich and his colleagues (2006), for example, studied the individual contribution of five different point mutations that all together increased cefotaxime resistance in *Escherichia coli* remarkably. In principal, there were 120 possible evolutionary pathways that can be generated by different combinations of the five mutations but only 18 trajectories were accessible by selection. Along the 102 inaccessible routes some mutations decreased resistance in certain genetic backgrounds and the evolutionary pathway ended. Only 18 mutational trajectories would lead to a higher resistance based on the increased resistance by each of the five mutations. The strength of resistance, which was measured as minimum inhibitory concentration (MIC) and used as a proxy for fitness, was highly dependent on the order of the mutations. The results showed that only a few pathways would lead to an elevated antibiotic resistance but more routes would lead to intermediate suboptimal states with only a small increase in resistance. Furthermore, in order to gain all five mutations, the mutations could not occur in a random order because of the lack of resistance in some genetic backgrounds (Weinreich *et al.*, 2006). Using the MIC is an elegant way to generate a fitness landscape but it fails to look at overall fitness consequences for the different genotypes and certain routes may be not as accessible as predicted by the MIC values.

### 4.1.2 Impact of the number and order of mutations on the evolutionary trajectory of a switcher genotype

#### *4.1.2.1 The rise of the switching phenotype in P. fluorescens*

During the course of the Reverse-Evolution Experiment (REE) two out of 12 replicate *P. fluorescens* lineages evolved a genotype that switch rapidly between two phenotypic stages (a switcher genotype). This novel phenotype was not detected in any of the parallel lineages before. The causative mutations occurred after eight preceding mutations in the *carB* gene (Line 1) and in *rpoD* (Line 6) during the ninth selection round in a static environment (see Chapter 1, section 1.5). Both mutations were reconstructed into the genomes of the immediate ancestor in both lines (1s4 and 6s4), and the ancestral SBW25 genome. The change of phenotype was examined each time and the fitness effect was measured. Overall it was shown that a mutation in *carB* or *rpoD* was sufficient to induce the switching phenotype in the immediate ancestor as well as in the ancestral SBW25 genotype. The consequences for fitness of the reconstructed types varied. The introduction of the *carB* mutation or the *rpoD* mutation into the immediate ancestor led to significantly increased fitness, indicating that the novel trait was adaptive under static conditions. After the mutations were introduced into the ancestral SBW25 genome a decreased fitness was observed, indicating some historical effect in switcher evolution (Beaumont *et al.*, 2009; Gallie, 2009).

During the Switcher Re-Evolution Experiment (SREE), which was carried out to re-evolve a switching type in four independently evolved lineages, an additional lineage, Line 3, was identified as a switcher lineage (see Chapter 3). Re-evolution from 3s4, which gained eight preceding mutations during previous selection rounds (similar to the immediate ancestors 1s4 in Line 1 and 6s4 in Line 6), led to evolution of switcher genotypes. In Line 3, only two out of six re-evolved switchers had a mutation in the *carB* gene and it was concluded that multiple pathways could lead to a switcher genotype.

When comparing the mutational histories of all three switcher lineages it became apparent that they overlap in some genes that were affected by a mutation along the mutational pathway, but they differ in the exact position of the mutation and in their order: none of the mutations were identical (Tab. 4.1). Nonetheless, the fitness results from the reconstruction of switcher mutations (*carB*, *rpoD*) into the immediate ancestor and SBW25 suggested that mutational history perhaps influenced the evolution of switcher phenotypes (Beaumont *et al.*, 2009; Gallie, 2009).

**Table 4.1: Mutational histories of three switcher lineages.** Summarised are the results from the original REE (Beaumont *et al.*, 2009; Gallie, 2009) and the SREE in which Line 3 was able to produce a switcher, including via a *carB* mutation (see Chapter 3). Listed are the genes in which a mutation arose in each selection round, the selective environment and the name of the genotype. The switcher was detected in the ninth selection round after eight preceding mutations. Mutations in *carB* and *rpoD* have been identified as possible pathways for a switcher.

| **Selection round** | **Enviro.** | **Line 1** | | **Line 6** | | **Line 3** | |
|---|---|---|---|---|---|---|---|
| | | Name | Mutational history | Name | Mutational history | Name | Mutational history |
| 1 | static | 1w0 | *mwsR* | 6w0 | *wspF* | 3w0 | *wspC/D* |
| 2 | shaken | 1s1 | *mwsR* | 6s1 | *wspF* | 3s1 | *wspC* |
| 3 | static | 1w1 | *awsX* | 6w1 | *awsX* | 3w1 | *mwsR* |
| 4 | shaken | 1s2 | *awsR* | 6s2 | *awsR* | 3s2 | *mwsR* |
| 5 | static | 1w2 | *wspF* | 6w2 | *wspF* | 3w2 | *mwsR* |
| 6 | shaken | 1s3 | *wssA* | 6s3 | *wssB* | 3s3 | *PFLU0185* |
| 7 | static | 1w3 | *mwsR* | 6w3 | *nlpD* | 3w3 | *PFLU0185* |
| 8 | shaken | 1s4 | *mwsR* | 6s4 | *nlpD* | 3s4 | *mwsR* |
| **Switcher** | static | 1w4 | *carB* | 6w4 | *rpoD* | 3w4 | *carB* |

#### *4.1.2.2 The exceptional role of Line 1 in switcher evolution*

The results from the original REE (Beaumont *et al.*, 2009; Gallie, 2009) together with the findings from the SREE from the last chapter (see Chapter 3) indicate some impact of the mutational history on switcher occurrence. So far three independent lineages (Line 1, Line 3 and Line 6) were capable of producing a switcher at the same evolutionary time point (1s4, 3s4 and 6s4) at which all lineages have gained eight mutations (Tab. 4.1) due to the same selection regime (see Chapter 3, Fig. 3.2). However, the likelihood that a switcher occurred differed between lines. Line 1 showed a high probability of generating a switcher, which suggests that after eight mutations this genotype was more accessible for switcher evolution. Line 3 and Line 6 produced significantly fewer switcher genotypes, indicating that the trajectory based on eight different mutations in these lines (Tab. 4.1) was less accessible. The results from the SREE of the previous chapter (see Chapter 3) provided evidence that evolution of a switcher in Line 1 was somehow different from the other switcher lineages (see Chapter 3, Fig. 3.3), which may be attributable to differences in the mutational pathway taken.

Although the results from the SREE suggest that the 1s4 genotype of Line 1 is more likely to generate a switcher than the genotypes 3s4 from Line 3 and 6s4 from Line 6, the questions as to how the switching phenotype primarily occurred and to what extent the mutational history influenced the outcome remain unanswered. It is for example not known yet whether genotypes that occurred previous to the original switcher in Line 1 have the same capacity to give rise to a switcher as the immediate ancestor 1s4. Could a switcher arise at an earlier time point, for example from the genotypes 1s1 or 1s2 (Tab. 4.1), with the same probability? Perhaps the evolution of switching is a stepwise process and all mutations that occurred prior to the original switcher contribute to the emergence of a switching phenotype. In that case the entire history is important (Wichman *et al.*, 1999; Meyer *et al.*, 2012). Alternatively two or more previous mutations might have increased the fitness effect of the *carB** mutation due to epistatic interactions (Weinreich *et al.*, 2006; Ortlund *et al.*, 2007; Blount *et al.*, 2008) and these additional mutations are essential for the higher chance of switcher evolution that was observed in Line 1 (see Chapter 3).

###### *4.1.2.2.1.1 Test for the probability of switcher evolution along the evolutionary pathway in Line 1*

To test the impact of evolutionary history on switcher evolution, the *carB* mutation from Line 1 (*carB**) was artificially introduced into different genotypes of the history in Line 1 and its effect on fitness was measured. The original switcher evolved in a static environment from the 1s4 genotype (Tab. 4.1). Therefore SBW25, 1s1, 1s2, 1s3 and 1s4 (immediate ancestor; Tab. 4.1) from Line 1 were selected to introduce *carB** into their genomes (Tab. 4.2). Previous studies have shown that *carB**, when introduced into the ancestral genome, is sufficient to produce the switching phenotype (translucent and opaque colony formation, colonies with opaque sectors; Beaumont *et al.*, 2009). According to the principles of the REE these "s" genotypes (1s1, 1s2, 1s3 and 1s4) evolved under shaking conditions and would be transferred into a static environment in the next selection round (see Chapter 1, section 1.4.4). By focussing on these genotypes it was possible to see whether *carB** could potentially evolve from these intermediate genotypes under the same conditions that led to the original switcher in Line 1.

In order to be detectable on agar plates following dilution of cultures containing $10^9$ cells, two conditions have to be fulfilled: 1.the *carB** mutation has to induce the typical switching phenotype (opaque and translucent colonies, colonies with opaque sectors, elevated proportion of capsulated cells; see Fig. 1.9) and 2. *carB** must enhance fitness in order to have a chance of increasing in frequency to the level necessary to make detection after plating a reasonable possibility.

**Table 4.2: Genotypes of Line 1 after *carB** introduction.** SBW25, 1s1, 1s2, 1s3 and 1s4 were selected to introduce *carB** into their genomes because they would, according to the REE, grow in a static environment in the next selection round. The original switcher evolved under these conditions from 1s4 and the reconstruction of *carB** into 1s4 (last row) should recover the switcher genotype.

| Name of genotype | Mutations | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
| **SBW25** | ***carB**** | | | | | | | | |
| **1s1** | *mwsR* | *mwsR* | ***carB**** | | | | | | |
| **1s2** | *mwsR* | *mwsR* | *awsX* | *awsR* | ***carB**** | | | | |
| **1s3** | *mwsR* | *mwsR* | *awsX* | *awsR* | *wspF* | *wssA* | ***carB**** | | |
| **1s4** | *mwsR* | *mwsR* | *awsX* | *awsR* | *wspF* | *wssA* | *mwsR* | *mwsR* | ***carB**** |

In order to be evolutionary relevant, *carB** needs to provide a fitness benefit for the genotype in which it occurs. Only then is *carB** able to spread through the population. If fitness decreases upon incorporation of *carB** into the genome, it is unlikely that this genotype could win over the non-*carB** genotype. In that case a *carB**-switcher has no chance to evolve (Khan *et al.*, 2011). The effect on fitness and therefore the capacity for *carB** to evolve might change over the course of evolutionary time.

In principle there are three different ways in which history could affect switcher evolution in Line 1. First, history has no effect and mutations generating *carB** are equally likely at each stage of the evolutionary history. In that case, *carB** will cause a switching phenotype and fitness increase in each genotype independently of the number and kind of mutations. Second, the entire mutational history based on eight mutations is required. In this case, one could imagine that the fitness of *carB** gradually increases from one genotype to the subsequent genotype because every mutation that occurs along the pathway individually contributes to the emergence of a switcher. Third, the ocurrence of *carB** in the immediate ancestor 1s4 may depend on just a few evolutionary events that take place throughout the history. Then the fitness of *carB** would only increase in combination with the relevant mutation in a particular genotype (Blount *et al.*, 2008).

Investigating the fitness effect of *carB** and how the phenotype changes in different genotypes of the evolutionary pathway in Line 1 can identify the genotypes with the highest accessibility for *carB** switcher evolution.

### 4.1.3 Objectives:

1. To test whether the introduction of *carB** can induce the switching phenotype (opaque and translucent colonies or colonies with opaque sectors) in different genetic backgrounds in Line 1.
2. To test the hypothesis that there is no significant difference in the capacity to evolve a *carB** switcher between genotypes that occurred at different evolutionary time points throughout the history of Line 1 (SBW25, 1s1, 1s2, 1s3 and 1s4). In that case the introduction of *carB** into the genomes of the different types has a similar fitness effect.

3. To test whether *carB** has a distinct effect on growth trajectories and growth rates in different genetic backgrounds of the evolutionary history.

## 4.2 Summary of materials and methods

The detailed protocol for each method and the materials used are described in the Materials and methods chapter (see Chapter 2).

### 4.2.1 Reconstruction of *carB** into genotypes of Line 1

The *carB** mutation was artificially introduced into genotypes from Line 1 that evolved during the course of the REE (see Chapter 1, section 1.4.4). The selected genotypes were SBW25, 1s1, 1s2, 1s3 and 1s4 and each acted as recipient in a bacterial conjugation with *E. coli* pUIC3*carB** as donor strain. The manufacture of the donor strain *E. colip*UIC3*carB** and the conjugation with a *P. fluorescens* donor strain are described earlier (see Chapter 2, sections 2.4.1-2.4.4). The successful introduction of the *carB** fragment was confirmed by the appearance of typical switcher colony morphology on KB agar plates and by comparative sequencing of *carB** for all generated genotypes with the ancestral SBW25 sequence of *carB*. The obtained strains SBW25*carB**, 1s1*carB**, 1s2*carB**, 1s3*carB** and 1s4*carB** were stored for further analysis at -80°C.

### 4.2.2 Capsule count assay in *carB** and non-*carB** strains of Line 1

After *carB** was successfully introduced, the effect of *carB** on the phenotype was estimated by counting the proportions of capsulated cells within populations of the manufactured *carB** strains (SBW25*carB**, 1s1*carB**, 1s2*carB**, 1s3*carB** and 1s4*carB**) and comparing them to the proportion of capsulated cells within populations of the original non-*carB** strains (SBW25, 1s1, 1s2, 1s3 and 1s4). Five replicates per strain were used for the capsule count assay, which followed the same protocol as has been described previously (see Chapter 2, section 2.5.1).

### 4.2.3 Measuring relative fitness of *carB** in SBW25, 1s1, 1s2, 1s3 and 1s4

Fitness assays were carried out between the engineered *carB** strains (SBW25*carB**, 1s1*carB**, 1s2*carB**, 1s3*carB** and 1s4*carB**) and the original non-*carB** strains of Line 1 (SBW25, 1s1, 1s2, 1s3 and 1s4) from the REE (Tab 4.3). In order to distinguish the competing strains on agar plates, the non-*carB** strains were phenotypically marked with a *lacZ* transcriptional fusion using a mini Tn7 transposition system (see Chapter 2, section 2.4.5). Parallel competition experiments between the strains SBW25, 1s1, 1s2, 1s3 and 1s4 and the associated *lacZ* strain were performed to reveal fitness effects of the marker system (right column in Tab. 4.3).

**Table 4.3: List of the fitness assays between *carB** strains and non-*carB** strains of Line 1.** Competition was carried out between manufactured *carB** strains and non-*carB** strains of Line 1 (*carB** fitness). In addition fitness consequences of *lacZ* introduction on the non-*carB** strains were tested (*lacZ* fitness) by competition experiments between non-*carB** strains and non-*carB** strains with *lacZ* fusion.

| ***carB** fitness** | | | ***lacZ* fitness** | |
|---|---|---|---|---|
| **Replicates** | **Competitor A vs Competitor B** | | **Competitor B vs Competitor C** | |
| 8 | SBW25*carB** | SBW25 | SBW25 | SBW25*lacZ* |
| 8 | 1s1*carB** | 1s1 | 1s1 | 1s1*lacZ* |
| 8 | 1s2*carB** | 1s2 | 1s2 | 1s2*lacZ* |
| 8 | 1s3*carB** | 1s3 | 1s3 | 1s3*lacZ* |
| 8 | 1s4*carB** | 1s4 | 1s4 | 1s4*lacZ* |

The aim of the experiment was to measure the effect of *carB** on the fitness in five different genotypes of Line 1 to see whether the fitness effect changes with progressing evolution. The competition was carried out in a static environment as described earlier (see Chapter 2, section *2.5.2.1*) and the relative fitness of *carB** was calculated using the equations from Chapter 2 in section *2.5.2.2*. There were eight replicates for each competition pair.

### 4.2.4 Population growth determination

Growth of bacterial cultures was determined by measuring the optical density at a wavelength of 600 nm ($OD_{600}$) over the course of 24 hours using a plate reader (see Chapter 2, section 2.6). The strains SBW25, 1s1, 1s2, 1s3, 1s4 from Line 1 and the equivalent *carB** strains SBW25*carB**, 1s1*carB**, 1s2*carB**, 1s3*carB** and 1s4*carB** were revived on KB agar plates from the stored frozen cultures. Seven colonies from each strain were transferred separately into 6 ml liquid KB media and incubated overnight, constantly shaking. On the following day 6 μl of the overnight culture were used to inoculate 6 ml fresh liquid media and were incubated overnight again. The bacterial cultures were then diluted and transferred into a 96-well plate. The cultures were grown under constant shaking and $OD_{600}$ values in each well were recorded in 10 min intervals. Afterwards maximal growth rate (Vmax), mean and standard error were calculated.

### 4.2.5 Statistical analyses

Statistical analyses were carried out using the software package R (R.app GUI 1.23), applying standard tests like the Shapiro-Wilk test for normal distribution and one-sample *t*-tests for the significance of the fitness effect and for differences in the maximum growth rates ($V_{max}$) and capsulation between *carB** and non-*carB** strains. Percentages of capsulated cells were transformed using arcsin transformation (Quinn & Keough, 2002) before the values were used for further statistical analysis. The software R and Microsoft Excel 11.6.0 (2004 Microsoft Corporation) were used for the graphic representation of the results.

## 4.3 Results

### 4.3.1 Phenotypic effect of *carB** on SBW25, 1s1, 1s2, 1s3 and 1s4

After the introduction of *carB** into the different genetic backgrounds that evolved during the REE in Line 1, colonies on agar plates were characterised and the fraction of capsulated cells was quantified to see whether *carB** caused a change in phenotype at the colony and at the cellular level by producing colanic acid on the cell surface.

The introduction of *carB** induced typical switcher colony morphology in all genotypes. Colonies showed opaque sectors, or a mixed population of translucent and opaque colonies on agar plates were observed. The count of capsulated cells in non-*carB** strains (SBW25, 1s1, 1s2, 1s3 and 1s4) and the *carB** strains (SBW25*carB**, 1s1*carB**, 1s2*carB**, 1s3*carB** and 1s4*carB**) revealed that the introduction of *carB** into different genetic backgrounds changed the extent of capsulation significantly in all strains (Fig. 4.1, Tab. 4.4). After *carB** was introduced into the ancestor SBW25 the proportion of cells producing colanic acid within the population increased from being almost zero to circa 20%. The introduction of *carB** into 1s1 led to approximately 30%, in 1s2 to 21% and in 1s3 to just over 10% of the population being capsulated. In 1s4 the average proportion of cells being capsulated in the presence of *carB** was circa 17%. For strains 1s2 and 1s4 without *carB** being introduced the Shapiro-Wilk test gave a significant result (Tab. 4.4) suggesting a lack of normal distribution. Values here were very small (Fig. 4.1; 0%-2% of the population) and the assay was performed for only five replicates, which could explain the lack of normal distribution.

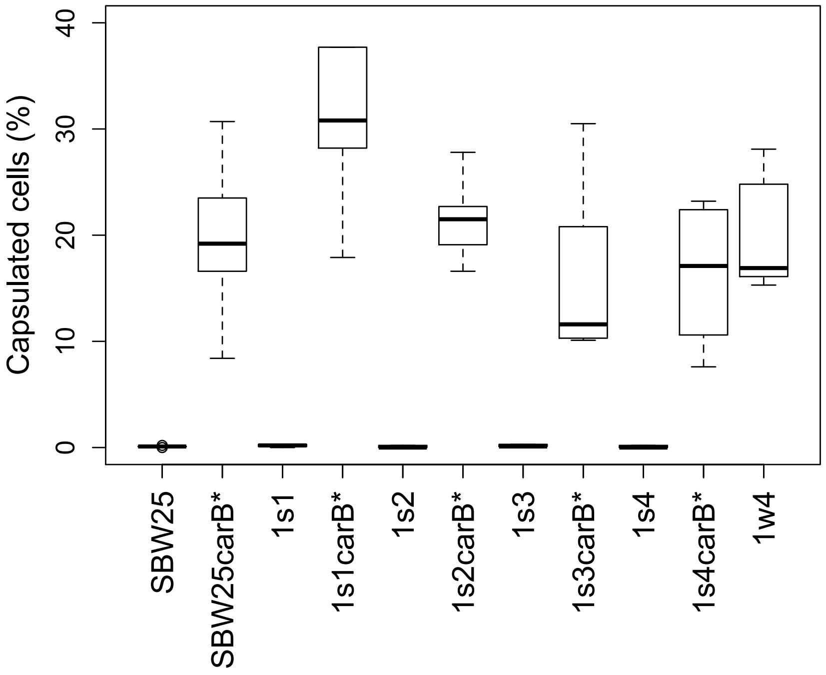

**Figure 4.1: Percentages of capsulated cells in populations of non-*carB** and *carB** strains.** The proportions of capsulated cells in non-*carB** strains (SBW25, 1s1, 1s2, 1s3 and 1s4) and the associated *carB** strains (SBW25*carB**, 1s1*carB**, 1s2*carB**, 1s3*carB**, 1s4*carB** and 1w4) are shown. Capsules were visualised by negative staining under the microscope. Displayed are the medians and interquartile ranges based on $N = 5$.

Nonetheless the values were used in two-sample *t*-tests. The introduction of *carB** into the different genotypes of Line 1 caused a phenotypic change at the cellular level by significantly increasing the frequency of capsulated cells (Tab. 4.4). As a control, the capsule production of the original evolved switcher from the REE, 1w4, was compared with the capsule production of reconstructed 1s4*carB** type, which had supposedly the same genotype. The results show that there was no difference between them (Tab. 4.4).

**Table 4.4: Summary of the statistical analyses of the capsule count assay.** A Shapiro-Wilk test was used to test for normal distribution ($\alpha$=5%). The two-sample *t*-test was applied to test whether the capsule proportion mean of the non-*carB** strains (SBW25, 1s1, 1s2, 1s3, 1s4) and the associated *carB** strains (SBW25*carB**, 1s1*carB**, 1s2*car*B*, 1s3*carB**, 1s4*carB**) were significantly different from each other (95% confidence interval). Analyses were performed with arcsin-transformed data.

| | **Shapiro-Wilk test** | | **Two-sample *t*-test** | | |
|---|---|---|---|---|---|
| **Strain** | **W** | ***P*-value** | ***t*** | **df** | ***P*-value** |
| SBW25 | 0.80 | 0.08 | -8.63 | 4.19 | < 0.001 |
| SBW25*carB** | 0.99 | 0.96 | | | |
| 1s1 | 0.84 | 0.18 | -12.79 | 4.49 | < 0.001 |
| 1s1*carB** | 0.88 | 0.32 | | | |
| 1s2 | 0.68 | 0.01 | -18.34 | 5.76 | < 0.0001 |
| 1s2*carB** | 0.98 | 0.95 | | | |
| 1s3 | 0.87 | 0.26 | -7.18 | 4.30 | < 0.01 |
| 1s3*carB** | 0.83 | 0.13 | | | |
| 1s4 | 0.68 | 0.01 | -8.55 | 4.49 | < 0.001 |
| 1s4*carB** | 0.90 | 0.42 | | | |
| 1s4*carB** | - | - | -1.05 | 7.39 | 0.32 |
| 1w4 | 0.84 | 0.15 | | | |

### 4.3.2 Fitness effect of *carB** in SBW25, 1s1, 1s2, 1s3 and 1s4

It has been shown that the introduction of *carB** into the genotypes SBW25, 1s1, 1s2, 1s3 and 1s4 had a significant effect on the phenotype and increased production of colanic acid capsules in all genotypes (Fig. 4.1). Fitness experiments between *carB** strains (SBW25*carB**, 1s1*carB**, 1s2*carB**, 1s3*carB** and 1s4*carB**) and non-*carB** strains (SBW25, 1s1, 1s2, 1s3, 1s4) were performed to uncover the net fitness effect that can be linked to *carB** after being inserted in the different backgrounds. In a competition experiment each non-*carB** strain was competed against its associated artificially constructed *carB** strain in a static environment. For example, SBW25 was competed against SBW25*carB**, 1s1 against 1s1*carB**, and so on (Tab. 4.3). The relative fitness was calculated from the Malthusian parameters of both competitors. Replicates that showed less than 30 colonies for both competitors on plates were excluded from further analyses.

#### *4.3.2.1 Fitness consequences of the marker system*

No significant marker effect was noticeable in any strain (Tab 4.5; see Appendices, Fig. A1). A *lacZ* transcriptional fusion was introduced into SBW25, 1s1, 1s2, 1s3 and 1s4 and tested for any fitness consequences of the marker system under static conditions.

**Table 4.5: Fitness effect of *lacZ* transcriptional fusion**. Results of Shapiro-Wilk tests (for normal distribution) and the one-sample *t*-tests, which compared the mean of relative fitness of *lacZ* in the genotypes SBW25, 1s1, 1s2, 1s3 and 1s4 with zero, are summarised.

| | **Shapiro-Wilk test** | | **One-sample *t*-test** | | |
|---|---|---|---|---|---|
| **Competing strains** | **W** | ***P*-value** | ***t*** | **df** | ***P*-value** |
| SBW25 vs SBW25*lacZ* | 0.88 | 0.29 | 1.68 | 4 | 0.17 |
| 1s1 vs 1s1*lacZ* | 0.91 | 0.36 | -0.89 | 6 | 0.41 |
| 1s2 vs 1s2*lacZ* | 0.85 | 0.13 | 1.55 | 6 | 0.17 |
| 1s3 vs 1s3*lacZ* | 0.94 | 0.63 | 0.91 | 7 | 0.40 |
| 1s4 vs 1s4*lacZ* | 0.97 | 0.89 | 0.59 | 7 | 0.57 |

#### *4.3.2.2 Fitness consequences of carB**

The results of the competition experiment between SBW25*carB** and SBW25 showed that the introduced mutation caused a significant increase in relative fitness compared to SBW25. The opposite effect was observed when *carB** was introduced into 1s1. Here the relative fitness of *carB** decreased. The competition experiment between 1s2*carB** and 1s2 showed a reduced fitness of *carB**; the mutation led to a negative relative fitness. In the next step the fitness of 1s3*carB** increased when growing together with 1s3 in a static microcosm. The introduction of *carB** into 1s4 resulted in a small but significant increase in relative fitness (Tab. 4.6). One replicate of 1s4 had less than 30 colonies for each competitor on the plate and was therefore excluded from further analyses (df = 6; Tab. 4.6).

**Table 4.6: Fitness effect of *carB** in genotypes of Line 1.** Results of Shapiro-Wilk tests (for normal distribution) and one-sample *t*-tests, which compared the mean of relative fitness of *carB** in the genotypes SBW25, 1s1, 1s2, 1s3 and 1s4 with zero, are summarised.

| | **Shapiro-Wilk test** | | **One-sample *t*-test** | | |
|---|---|---|---|---|---|
| **Competing strains** | **W** | ***P*-value** | ***t*** | **df** | ***P*-value** |
| SBW25*carB** vs. SBW25 | 0.88 | 0.20 | 7.30 | 7 | < 0.0001 |
| 1s1*carB** vs. 1s1 | 0.93 | 0.55 | -6.31 | 7 | < 0.001 |
| 1s2*carB** vs. 1s2 | 0.94 | 0.56 | -3.77 | 7 | 0.003 |
| 1s3*carB** vs. 1s3 | 0.94 | 0.65 | 11.31 | 7 | < 0.0001 |
| 1s4*carB** vs. 1s4 | 0.96 | 0.85 | 3.66 | 6 | 0.005 |

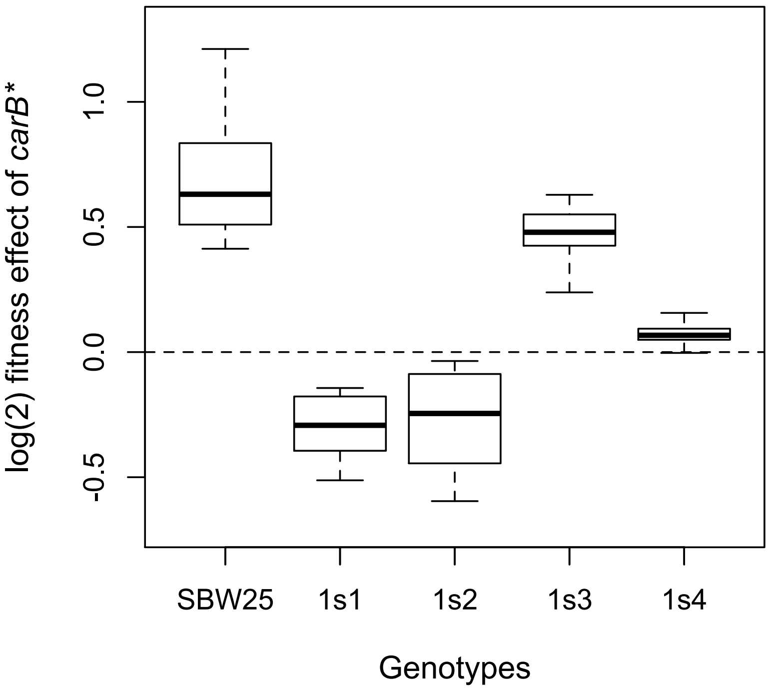

**Figure 4.2: Relative fitness of *carB** in SBW25, 1s1, 1s2, 1s3 and 1s4.** The relative fitness is displayed as log2 of the ratio of the two Malthusian parameters of each competitor. The dashed line at zero indicates no fitness difference after *carB** was

introduced. Values above the dashed line indicate a fitness increase and below the line indicate fitness decrease. Displayed are medians and interquartile ranges based on *N* = 8.

Overall the results of the fitness assays demonstrate that the net fitness effect of *carB** differed between the genotypes of Line 1 (Fig. 4.2).

### 4.3.3 Effects of *carB** on bacterial growth

Growth of the bacterial cultures was determined by measurements of their $OD_{600}$ values over the course of 24 hours in 10 min intervals. The aim of the experiment was to identify differences in growth parameters between non-*carB** and *carB** strains as a phenotypic effect of the *carB** mutation.

#### *4.3.3.1.1 Bacterial growth of carB* and non-carB* strains over 24 hours*

The $OD_{600}$ measurements were taken in a shaken environment every ten minutes. The growth curves in Figure 4.3 illustrate the behaviour of each bacterial culture over time, starting with the lag-phase with almost no noticeable change in growth, followed by the log-phase where bacterial cells grew exponentially, and the beginning of the stationary phase when no change in growth occurred. The *carB** strains (SBW25*carB**, 1s1*carB**, 1s2*carB**, 1s3*carB**, 1s4*carB**) and the non-*carB** strains (SBW25, 1s1, 1s2, 1s3, 1s4) followed similar growth trajectories. It is noticeable that during exponential growth the strains diverge from each other. In general the slopes of the growth curves when *carB** was present in the strains did not vary very much from the strains without *carB** but the *carB** strains took longer to enter the exponential growth phase (Fig. 4.3).

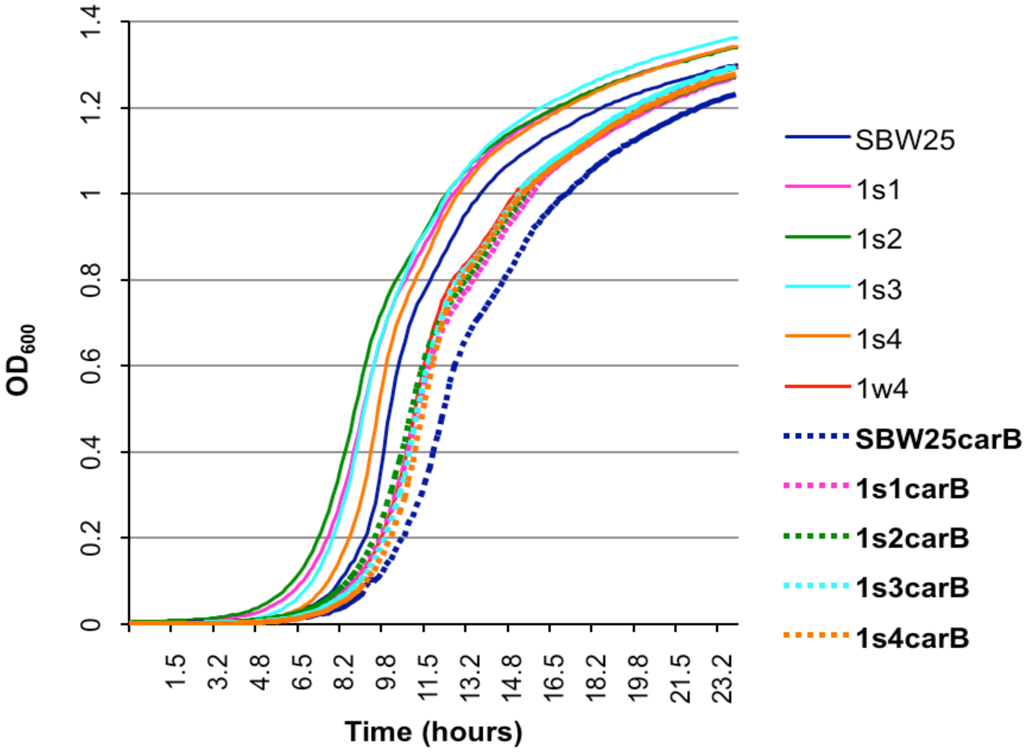

**Figure 4.3: $OD_{600}$ of non-*carB** and *carB** strains over 24 hours under shaking conditions (*N* = 7).** Growth curves are based on $OD_{600}$ values that were measured every 10 min and are illustrated as mean values of seven replicates for each strain.

### 4.3.3.2 Change of growth rates after carB* introduction

Maximum growth rates ($V_{max}$) were calculated for each non-*carB** and *carB** strain to provide additional information about growth properties. The $V_{max}$ represents the maximal increase of cell numbers that occurred during 24 hours under shaking conditions.

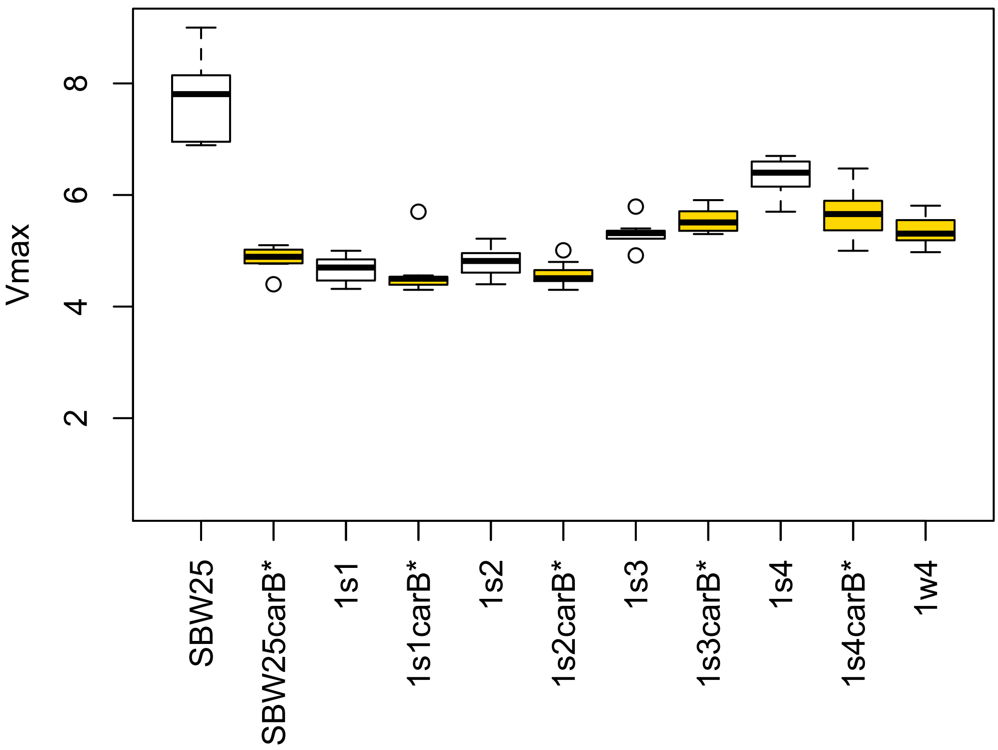

**Figure 4.4: Maximum growth rates ($V_{max}$) for non-*carB** and *carB** strains ($N$ = 7).** The diagram shows $V_{max}$ for the strains of Line 1 (SBW25, 1s1, 1s2, 1s3 and 1s4) without *carB** (white boxplots) and with *carB** introduced into the genome (yellow boxplots). In addition $V_{max}$ for 1w4, the switcher that evolved in the original REE, is illustrated. Thick black lines in the box plot indicate the median of seven replicates. Displayed are medians and interquartile ranges based on $N$ = 7.

**Table 4.7: Effect of *carB** on maximum growth rate ($V_{max}$) in genotypes of Line 1.** Results of Shapiro-Wilk tests (for normal distribution) and two-sample *t*-tests, which compared the mean of $V_{max}$ of *carB** strains (SBW25, 1s1, 1s2, 1s3 and 1s4) with non-*carB** strains (SBW25carB*, 1s1carB*, 1s2carB*, 1s3carB*, 1s4carB* and 1w4 (the original switcher genotype)) in Line 1, are summarised.

| | **Shapiro-Wilk test** | | **Two-sample *t*-test** | | |
|---|---|---|---|---|---|
| **Strain** | **W** | ***P*-value** | ***t*** | **df** | ***P*-value** |
| SBW25<br>SBW25*carB** | 0.89<br>0.90 | 0.28<br>0.34 | 8.98 | 7.04 | < 0.0001* |
| 1s1<br>1s1*carB** | 0.96<br>0.83 | 0.80<br>0.11 | 2.09 | 8.43 | 0.07 |

| | | | | | |
|---|---|---|---|---|---|
| 1s2<br>1s2*carB** | 0.95<br>0.88 | 0.76<br>0.24 | 1.51 | 11.58 | 0.16 |
| 1s3<br>1s3*carB** | 0.92<br>0.82 | 0.44<br>0.07 | -1.72 | 12.00 | 0.11 |
| 1s4<br>1s4*carB** | 0.90<br>0.97 | 0.35<br>0.93 | 2.84 | 10.99 | 0.02* |
| 1s4*carB**<br>1w4 | -<br>0.95 | -<br>0.73 | 1.36 | 9.86 | 0.20 |

It is noticeable that $V_{max}$ is high for the wild type SBW25 in comparison to other strains, irrespective of whether these strains have *carB** or not (Fig. 4.4). After *carB** was integrated into the SBW25 genome $V_{max}$ decreased significantly (Tab. 4.7, $p < 0.0001$). The strains 1s1 and 1s2 already show a lower growth rate and the addition of *carB** does not cause further change. The introduction of *carB** into 1s3 does not result in a significant change of $V_{max}$ (Fig. 4.4, Tab. 4.7). Of all the evolved strains in Line 1, 1s4 has a high maximum growth rate and the artificial introduction of *carB** into its genome significantly decreases $V_{max}$ (Fig. 4.4, Tab. 4.7). In the original REE 1w4 was selected as a novel switching phenotype and the *carB** mutation was obtained from its genotype (see Chapter 1, section 1.5). The strains 1s4*carB** and 1w4 share the same genotype and show the same $V_{max}$ (Fig. 4.4, Tab. 4.7).

## 4.4 Discussion

Over the last two decades more and more studies have addressed the possibility that evolutionary pathways of novel phenotypes can be constrained by the number and order of the mutations that occur along them (Weinreich *et al.*, 2006; Ortlund *et al.*, 2007; Khan *et al.*, 2011). Some pathways are more likely to be taken than others, meaning that some genotypes are evolutionarily more accessible than others (Stern & Orgogozo, 2009). A number of studies have found that the mutational history plays a major role in the accessibility of evolutionary pathways

(Wichman *et al.*, 1999; Blount *et al.*, 2008; Meyer *et al.*, 2012; Wischmann *et al.*, 2012).

In this study I was interested in the effect of evolutionary history on the capacity to evolve a switcher genotype in Line 1 based on a *carB** mutation. It has been suggested that the evolutionary pathway in this lineage is likely to have an impact on switcher evolution (Beaumont *et al.*, 2009) but by what means has not yet been established. To be able to evolve a *carB** switcher, the mutation needs to carry a fitness benefit for the genotype in which it occurs under given environmental conditions. Beaumont and his colleagues (2009) showed that in the genotype 1s4 (the immediate ancestor of the switcher), *carB** improved the performance of the genotype in a static environment by inducing a switching phenotype. The 1s4 genotype gained eight mutations during the course of the REE. Whether *carB** has a similar effect in genotypes that evolved previous to the switcher in Line 1 is not yet known. Testing how *carB** changes the phenotype and the fitness in different genotypes that occurred during the course of evolution in Line 1 can reveal how accessible earlier genotypes are for *carB** switcher evolution and how evolutionary dynamics change over time with proceeding evolution.

### 4.4.1 Phenotypic effect of *carB**

The introduction of *carB** into the genomes of SBW25, 1s1, 1s2, 1s3 and 1s4, which represent different evolutionary time points of the evolutionary trajectory of the switcher in Line 1, caused in each background the appearance of typical switcher colony morphologies on agar plates, such as translucent and opaque colonies and colonies with opaque sectors. The results show that this single point mutation in the *carB* gene is sufficient to induce this phenotype in every genetic background. The findings coincide with the observations from Beaumont et al. (2009) when *carB** was introduced into the ancestral SBW25 genome (Beaumont *et al.*, 2009). This observation that the switcher phenotype was induced by *carB** in every genotype of the evolutionary pathway in Line 1 led to the conclusion that all previous genotypes can potentially lead to the emergence of a switcher when *carB** spontaneously occurs. The lack of switcher emergence during the REE could be the result of not having enough time to evolve and a switcher would have occurred later on, or there are additional factors that changed the likelihood of occurrence.

### 4.4.2 Change of evolutionary accessibility for *carB** along the evolutionary pathway in Line 1

The production of the switcher phenotype (opaque and translucent colonies and/or colonies with opaque sectors) is one of the conditions that have to be fulfilled by a switcher genotype to be able to evolve from the different genetic backgrounds. Moreover, *carB** has to provide a fitness benefit for the genotype in order to increase in cell frequency and to be detectable on an agar plate following dilution of cultures containing $10^9$ cells. The results of this study show that the fitness effect of *carB** in the different genotypes varied along the evolutionary pathway.

In order to determine the effect of history on switcher evolution three possibilities were tested: (1) History has no effect on switcher evolution and *carB** is likely to arise from any given starting position due to an increased fitness in the different genotypes (SBW25, 1s1, 1s2, s3 and 1s4). (2) Every mutation that occurred throughout the history in Line 1 contributes and the entire history is required for the evolution of a switcher. (3) Particular evolutionary events that occurred along the evolutionary trajectory promote the emergence of the switcher later on.

*1. Mutational history has no effect:* The results of the fitness experiments show that in principal the mutational history is not relevant. The introduction of *carB** into the genome of the ancestor SBW25 increased the fitness, indicating that a switcher based on *carB** already has a high probability of occurring at the starting point when none of the eight mutations is present. Nevertheless there was large variation in the fitness effect of *carB** in the subsequent genotypes. For example, due to a negative fitness of *carB**, the mutation has only a limited chance to arise in 1s1 and 1s2. The decreased fitness after *carB** was introduced into 1s1 and 1s2 is likely to cause the extinction of the genotype in the event of a *carB** mutation and the evolutionary path would end. This result is similar to outcomes from other experiments (Weinreich *et al.*, 2006; Khan *et al*, 2011; Meyer *et al.*, 2012). In genotypes of the advanced evolutionary history, such as 1s3 and 1s4, a switcher based on a *carB** mutation has a greater chance to establish within the population because it causes a fitness increase.

*2. The entire mutational history has an effect:* As mentioned earlier in this paragraph there was no stepwise fitness increase of *carB** from the ancestor SBW25 across the subsequent genotypes up to 1s4. The results demonstrate that the evolution of a *carB** based switcher is not dependent upon the combined effect of all mutations that occurred previously to the switcher in 1s4 (Wichman *et al.*, 1999; Ortlund *et al.*, 2007; Meyer, 2012).

*3. Few evolutionary events have an effect:* It was shown that the mutational history was in principal not needed in order to produce a switching genotype. The ancestor SBW25 has already the capacity to generate an evolutionary successful switcher based on *carB**. This capacity was not observed in 1s1 and 1s2 but again in 1s3 and 1s4. This large variation of the fitness results from the different backgrounds indicates that particular evolutionary events that occurred along the evolutionary trajectory promote the emergence of the switcher later on (Blount *et al.*, 2008; Salverda *et al.*, 2011; Blount *et al.*, 2012). In particular the large fitness increase from 1s2 to 1s3 after *carB** was introduced (Fig. 4.2.) suggests that the genotype of 1s3 promotes the evolution of a switcher from this evolutionary position.

As mentioned earlier, the introduction of *carB** into the ancestral SBW25 genome resulted in a large fitness increase. This outcome is not in agreement with the findings of Beaumont et al. (2009). They performed a similar fitness experiment in which SBW25 competed against SBW25*carB** in a static environment and found that no significant fitness increase could be attributed to *carB**. Gallie (2009) on the other hand found that over the course of 48 hours, *carB** can provide a benefit in the SBW25 genotype under static conditions. The performance of the fitness assay over the course of 72 hours, however, made the significant fitness effect of *carB** disappear due to other new types that emerged as a result of progressing evolution (Beaumont *et al.*, 2009). This shows that the results of such fitness experiments are very responsive to time. Furthermore, population dynamics within the microcosms are very sensitive to other factors as well. For example the pH-value of the media or the temperature can change the transition rate between capsulated and non-capsulated cells within a population with a presumable impact on the outcomes of such fitness experiments (personal conversation with Jenna Gallie). However, the increased fitness of SBW25 in the presence of *carB** indicates that in theory a switcher has the potential to evolve from the ancestral genotype without any additional mutation.

The *carB** mutation had a negative fitness effect on 1s1 and 1s2. Even though the phenotype was generated by *carB** in both genotypes, the mutation would not be able to increase in frequency because of a low fitness. This indicates strong genetic constraints, perhaps caused by negative epistasis (Khan *et al.*, 2011). Here two mutations can be beneficial for the organism but when they occur together in the same genotype the fitness decreases because of negative interactions.

A significant increase in fitness was observed after *carB** was introduced into the genome of 1s3. The improved performance of 1s3 when *carB** was present suggests that in this background *carB** has a realistic chance of occurring and to increase its frequency. The genotypes of 1s2 and 1s3 are almost identical. Only two additional mutations separate 1s3 from 1s2 but the fitness effects of *carB** were very different from each other. It is likely that one of the two additional mutations or the combined effect of both contribute to the fitness increase of *carB** in 1s3. Hence the mutations are perhaps crucial evolutionary events that open up an evolutionary pathway for switcher evolution based on *carB**. This will be further investigated in Chapter 6.

As expected, the introduction of *carB** into the genome of the original immediate ancestor 1s4 (REE; see Chapter 1, sections 1.4.4 & 1.5) resulted in increased fitness (Beaumont *et al.*, 2009). However, the advantage given by *carB** was not as high in 1s4 as in 1s3. Previous studies found that 1s4, which evolved in a shaken environment, still produced detectable amounts of a cellulosic polymer (Gallie, 2009). This leads to the ability to produce a biofilm at the surface in a static environment by which 1s4 gains a significant benefit. The slightly decreased advantage of *carB** in 1s4 compared to the huge advantage in 1s3 is therefore perhaps attributable to an already very fit 1s4 type. Nonetheless the *carB** switcher is still fit enough to arise in this genotype.

### 4.4.3 Impact of environmental factors on *carB** fitness

The growth curves and growth rates of the different genotypes of Line 1 were measured without and in the presence of *carB**. Essentially *carB** did not change the trajectory of the growth curves of the different genotypes but all *carB** strains needed more time until they entered the log-phase. Similar results were obtained for the maximal growth rates. In general $V_{max}$ did not change significantly after *carB** was introduced into the genome except for in the ancestral SBW25 strain and 1s4. In both cases $V_{max}$ decreased significantly. These findings do not coincide with the results from the fitness experiments. If fitness was just determined by the cell doubling, one would expect high fitness values to be associated with high maximal growth rates. The variation between the results from the fitness experiments and the maximal growth rates suggest that factors additional to growth rates determine whether *carB** provides a fitness advantage or not. Growth rates were measured under shaking conditions, but the fitness experiments were performed in static microcosms. It is highly likely that ecological interactions not only between the competition partners but also with the surrounding environment (e.g. the heterogeneous microcosm based on an oxygen gradient) influence the population dynamics (Meyer *et al.*, 2012) of the different types, affecting their competitive power. To study ecological dynamics in bacteria in these microenvironments is challenging but essential in order to gain more insight into evolutionary dynamics based on random mutation and deterministic processes in the context of environmental conditions.

## 4.5 Summary

In this study I was interested in the impact of evolutionary history on the evolution of a switcher genotype based on a *carB** mutation in Line 1. I wanted to know whether the capacity to evolve a *carB** switcher (phenotypic change and fitness consequences after *carB** was introduced) changes with progressing evolution, depending on the genotype representing different evolutionary stages with a different number of mutations. The data show that *carB** is sufficient in each genotype to induce the switcher phenotype. Furthermore the hypothesis that there is no difference in the capacity to evolve a *carB** switcher between the different genotypes in Line 1 (see 4.1.1) can be rejected. The variation in fitness effects of *carB** in genotypes from different evolutionary time points in Line 1 shows that the evolutionary accessibility for *carB** changes at each time point and is rather unpredictable. I found no evidence for a continuously increasing fitness of *carB** along the evolutionary pathway. The introduction of *carB** increased the fitness in the ancestor SBW25, decreased fitness in the two subsequent types 1s1 and 1s2, and increased fitness again in 1s3 and 1s4. This indicates that *carB** appearance depends on the genotype in which it occurs. It is likely that the increased fitness in 1s3 is attributable to some additional mutations that are present in this genotype. In principal SBW25 is already capable of giving rise to a *carB** switcher, meaning that the mutational history is not relevant in the ability to produce an evolutionarily successful *carB** switcher. The causes for the lack of switcher occurrence during the course of the REE will be further investigated in the following chapter.

# Chapter 5: Impact of evolutionary history on real-time evolution of a switching genotype

## 5.1 Introduction

Ever since Darwin and Wallace introduced their theory of evolution by natural selection (Darwin & Wallace, 1858; Darwin, 1859), scientists have been fascinated with the questions as to whether evolution is repeatable and predictable. Evolutionary trajectories are shaped by the order in which random mutations occur and are the result of selection acting upon these mutations. To what degree and under what conditions randomly occurring mutations and selection shape evolutionary pathways are still not well understood. In order to predict evolutionary pathways scientists have become interested in evolutionary routes that can be taken by an organism and have investigated their relevance for the evolutionary outcomes. To achieve this, researchers use empirical fitness landscapes to obtain estimates about possible evolutionary pathways and their likelihood of being taken by organisms (Lunzer *et al.*, 2005; Weinreich *et al.*, 2006; de Visser *et al.*, 2009; Franke *et al*, 2011; Khan *et al.*, 2011; Kvitek & Sherlock, 2011). Empirical fitness landscapes are based on fitness measurements, usually between two genotypes. One genotype represents an ancestral state and the other genotype differs by a small set of known mutations. Possible intermediate genotypes are then artificially created and their fitness is measured. If fitness is higher than the ancestor, then this path is accessible by evolution. The highest fitness indicates the most likely evolutionary course. In this way evolutionary trajectories, usually for a single trait, can be reproduced step by step. Investigations of fitness landscapes have shown that the number of mutations and order in which they occur during evolution can have a major impact on the outcome of evolution (Weinreich *et al.*, 2006; Ortlund *et al.*, 2007; Franke *et al.*, 2011; Khan *et al.*, 2011).

Trajectories of real-time evolution can be expected to be more complex than trajectories based on only a few genetic loci because the fitness of an organism depends on many genes. It has been pointed out that under natural conditions selection does not act on a small collection of genetic variations but rather acts on all possible beneficial mutations that occur anywhere in the genome (Franke *et al.*, 2011) that could shape the 'real' fitness landscape (Colegrave & Buckling, 2005). Evolution can lead to multiple possible evolutionary outcomes following different mutational routes. Furthermore the fitness of an organism may depend on the presence of other genotypes within the population. This has been observed during experimental evolution in microorganisms (reviewed by Kassen, 2009; Elena & Lenski, 1997b; Rainey & Travisano, 1998; Friesen *et al.*, 2004; MacLean *et al.*, 2004; Maharjan *et al.*, 2006; Lang *et al.*, 2011). Here, the success of a particular genotype depends on the frequency of other genotypes. For example frequency dependency is the result of the emergence of cross feeding, where one type feeds on the metabolic waste product of another type (Helling *et al.*, 1987; Rosenzweig *et al.*, 1994; Treves *et al.*, 1998). Hence the predictability of evolutionary pathways is complex and difficult because of selection acting on many mutations that occur in the whole genome of an organism and almost incalculable ecological interactions between different genotypes that may occur during diversification. How theoretical fitness landscapes, which is the order in which mutations are likely to occur, and estimates of most probable evolutionary trajectories translate into real-time evolution has hardly been investigated so far.

### 5.1.1 Impact of mutational history on *carB** switcher evolution in Line 1

The *carB** mutation was responsible for the emergence of a switching genotype in one of two switcher lineages (Line 1) that evolved during the course of the Reverse-Evolution Experiment (REE) amongst 12 parallel *P. fluorescens* lineages (see section 1.4.4 and 1.5). After eight mutations, which were gained during eight selection rounds either in a static or shaken environment, the switcher genotype appeared as a result of strong selection acting within a static environment (see section 1.5; Rainey & Travisano, 1998; Beaumont *et al.*, 2009). Previous studies investigated the fitness effect of *carB** on the ancestral SBW25 genotype and the immediate ancestor 1s4 in Line 1. They found that the fitness of the ancestor SBW25 decreased in the presence of *carB**, but in the immediate ancestor 1s4 *carB** had a beneficial effect and was adaptive (Beaumont *et al.*, 2009; Gallie, 2009). In addition, Beaumont and colleagues (2009) performed replay evolution experiments, which started from the common ancestor SBW25 or the immediate ancestor 1s4. After three days in a static environment none of 138 replicate populations of SBW25 produced a switcher genotype, whereas in six of 36 replicates of 1s4 a switcher was observed. In addition, five of the six switcher genotypes showed a causative mutation in *carB* while the other one had a mutation in *rpoD*. They concluded that SBW25 and 1s4 differed fundamentally in their capacity to give rise to a switcher type, suggesting that some of the mutations in the evolutionary history contributed to the evolution of a switcher genotype. (Beaumont *et al.*, 2009; Gallie, 2009). Evidence is still lacking as to which part of the history and to what degree history affects switcher evolution.

As part of the study described in the previous chapter (see Chapter 4), *carB** was artificially introduced into the common ancestor SBW25, the immediate ancestor 1s4 and intermediate genotypes (1s1, 1s2, 1s3) that occurred at different time points in the switcher lineage Line 1 along the evolutionary pathway. The genotypes differed in the number of mutations that had accumulated over the course of the REE. The results revealed that the *carB** introduction had very different fitness consequences (Fig. 5.1). When introduced into the ancestral SBW25 genome, *carB** had a positive effect on fitness. Although this outcome contradicts the findings of Beaumont et al. (2009) due to differences in the experimental design of the fitness assays (duration of the competition experiment), it demonstrates that in principle *carB** has a realistic chance of arising within a population of SBW25 and to increase in frequency. Switcher evolution appears to be already possible in the common ancestor SBW25 (Fig. 5.1). The opposite was observed for *carB** fitness in genotypes that evolved during subsequent selection rounds, 1s1 and 1s2. Here *carB** decreased the fitness of both genotypes and hence a *carB** switcher cannot easily emerge in these genetic backgrounds under static conditions. A large fitness increase was observed after the introduction of *carB** into the genome of 1s3. The benefit of having *carB** was then maintained in 1s4, the immediate ancestor of the original switcher from the REE, although the fitness effect was not as strong as in 1s3. Switcher evolution based on *carB** becomes more likely again in 1s3 and 1s4. These findings suggest that the mutational history is in principal not required in order to give rise to a switcher genotype.

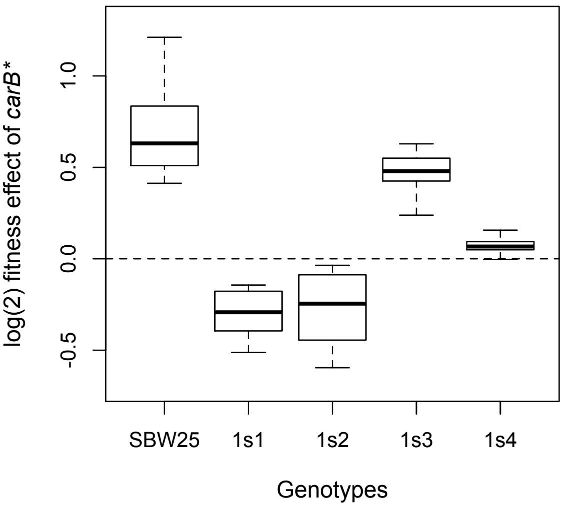

**Figure 5.1: Results of *carB** fitness measurements in static in different genotypes of Line 1.** SBW25, 1s1, 1s2, 1s3 and 1s4 are genotypes that evolved at different time points during the course of evolution in Line 1. The introduction of *carB** in these background and fitness measurements indicate how likely it is for a *carB** switcher to arise in the different backgrounds. Results were obtained in Chapter 4. Displayed are medians and interquartile ranges based on *N* = 8.

## 5.1.2 Lack of switcher occurrence during the REE in Line 1

A switcher based on *carB** was only observed during experimental evolution from 1s4 over the course of the REE in Line 1 (Beaumont *et al.*, 2009). The results from the previous chapter (see Chapter 4, section *4.3.2.2*) indicate the existence of genetic constraints that may explain the lack of switcher occurrence in 1s1 and 1s2. Here the *carB** mutation decreased the fitness of both genotypes and consequently a *carB** switcher was unlikely to increase in frequency and was unlikely to be found during the REE. However, the results from the fitness experiments in other genotypes (Fig. 5.1) revealed that *carB** has a high chance of appearing in the ancestor SBW25 and 1s3 due to a fitness increase. This raises the question as to whether other factors (i.e. ecological interactions between different genotypes; Turner et al., 1996; Rozen & Lenski, 2000) limited switcher occurrence at these earlier time points and could explain why no switcher was detected in SBW25 and 1s3 throughout the REE.

Evolutionary change is driven by ecological opportunity (Bell, 2008). It has been shown that spatial heterogeneity plays an important role in the diversification of bacterial populations, i.e. *P. fluorescens* (Rainey & Travisano, 1998). Structured environments (i.e. static microcosms) offer new unoccupied niches and therefore the opportunity for novel types to arise. The interactions of ecological factors in spatially structured environments are very complex (Rainey *et al*., 2000; Rozen & Lenski, 2000), and the impact of ecological factors on evolutionary outcomes is not fully understood. The outcome of competition between multiple genotypes might not solely depend upon their fitness in a particular environment alone but also on the presence of other genotypes (Chao & Levin, 1981). Complex interactions between different genotypes could potentially explain the lack of switcher evolution in SBW25 and 1s3 during the REE despite the genetic capacity to evolve an evolutionarily successful switcher (see Chapter 4). Here, it is not only critical at which time point the switcher mutation occurs, but also how many different genotypes and competitors are present within a population, and how frequent they are. It is likely that the number of competitors determine whether a switching genotype can increase in frequency within a population. Thus in addition to the genetic capacity at different evolutionary time points, switcher emergence might be influenced by ecological factors. In that case the likelihood of switcher occurrence should be linked to changes in population composition, i.e. the amount of different genotypes, the cell density of new types, and the diversity of genotypes present within the populations at different evolutionary time points.

Empirical fitness landscapes look at the fitness of artificially created genotypes under a particular set of environmental conditions in order to determine the likelihood of their occurrence. Most of the time these fitness landscapes represent very specific evolutionary situations (e.g. strong selection/weak mutation)

(Weinreich *et al.*, 2006), but whether the results would reoccur during real-time evolution has not been explored. In this study a Switcher Re-Evolution Experiment (SREE) was performed to test whether the findings from Chapter 4, the fitness effect of *carB** in different genotypes of Line 1, can be reproduced during real-time evolution. Based on the fitness effect of *carB** in SBW25 (Fig. 5.1) one would expect the ancestor genotype to re-evolve the highest number of switchers. A very low number of switcher genotypes are likely to be found in 1s1 and 1s2, whereas a high likelihood of switcher emergence can be expected in 1s3 and 1s4.

Evolution experiments focus either exclusively on genetic changes or on the change of ecological interactions. Studies that combine genetical analyses and ecology are rare. Evolution is the result of the combined effect of many factors, including randomly occurring genetic changes, natural selection, the impact of the surrounding environment and ecological interactions between different genotypes within a population (Meyers & Bull, 2002). How interactions between different factors change with proceeding evolution and what factors are more dominating under which conditions is unknown. For example it is still obscure as to whether the number and order of mutations, which exist at different time points of evolutionary trajectories, have an impact on the level of effectiveness of different factors and therefore an impact on the outcome of evolution.

### 5.1.3 Objectives

1. To test the hypothesis that the genotypes from different evolutionary time points in Line 1 are similar in their likelihood to evolve switcher phenotypes during real-time evolution (Switcher Re-Evolution Experiment (SREE)). In that case I expect no difference in the numbers of switchers that evolve from SBW25, 1s1, 1s2, 1s3 or 1s4.

2. To test the hypothesis that there is a high degree of parallelism between genotypic and phenotypic switcher evolution in genotypes from different evolutionary time points (SBW25, 1s1, 1s2, 1s3 or 1s4). If this is true then

parallel-evolved switcher should show similar causative mutations, i.e. in the same gene (*carB*).

3. To test whether ecological interactions have an impact on the likelihood of switcher evolution in genotypes from different evolutionary time points. Here, variation in the number of re-evolved switchers should be linked to a change in population composition (i.e. number of new variants, cell density, diversity) over evolutionary time.

## 5.2 Summary of materials and methods

### 5.2.1 Experimental design of the SREE in Line 1

Experimental conditions of the SREE were based on the experimental design of the original REE and followed the same experimental guidelines (see section 1.4.4). In summary, during the REE bacterial cultures were grown for three days in one of the selective environments, static or shaken, which marked one selection round. After three days the bacterial cultures were plated and screened for a new colony morphology type. The most abundant new type was selected for the inoculation of a new microcosm, which was then kept under the opposite environmental condition than that in which it had evolved (the next selection round). This was repeated for several selection rounds, with an alternation between a static and a shaken environment for each selection round. In cases when no new type was detected, a small amount of culture was transferred into a new microcosm and kept for another three days under the same environmental condition (first transfer, same selection round). This was repeated until a new type was detected on agar plates (several transfers in one selection round). The appearance of a type with novel colony morphology marked the end of a selection round.

In this study a replay experiment (SREE) was performed from different starting positions (evolutionary time points) in Line 1. The SREE was carried out to test whether the evolution of a switcher genotype is equally likely from any genotype that evolved at different time points during the course of evolution or whether the eight preceding mutations that were present in 1s4 (immediate ancestor of the original switcher in the REE) are essential. The switcher in the REE occurred in Line 1 and Line 6 after eight rounds of selection (see section 1.5). The focus of this study is on Line 1 as it has been characterised and studied to a great extent in previous studies (Beaumont *et al.*, 2009; Gallie, 2009). The switcher in Line 1 (1w4) evolved in a static environment from 1s4.

**Table 5.1: Mutational history of the switcher (1w4) in Line 1.** Evolutionary history of the switcher in Line 1 is composed of nine mutations, including the switcher mutation *carB** (last column). Mutations that occurred during history occurred in genes that are involved in the regulation of cellulose synthesis due to selection in a static or in a shaken environment. The founder populations were genotypes that would, according to the experimental design of the REE, go into a static environment in the following selection round ("s" genotypes; bold) because the original switcher (1w4) evolved under static conditions.

| **Number of mutations** | **1** | **2** | **3** | **4** | **5** | **6** | **7** | **8** | **9** |
|---|---|---|---|---|---|---|---|---|---|
| **Environment** | static | shaken | static | shaken | static | shaken | static | shaken | **static** |
| **Name** | 1w0 | **1s1** | 1w1 | **1s2** | 1w2 | **1s3** | 1w3 | **1s4** | **1w4** |
| **Mutational History** | *mwsR* | *mwsR* | *awsX* | *awsR* | *wspF* | *wssA* | *mwsR* | *mwsR* | ***carB**** |

For that reason five genotypes that occurred during the REE in Line 1 and would grow according to the rules of the REE in the next selection round in a static environment were selected (Tab. 5.1; SBW25, 1s1, 1s2, 1s3, and the immediate ancestor 1s4). The strains were revived on KB agar plates from a frozen stock culture.

The experiment was divided into three experimental blocks with 20 replicates of each strain in each block (60 replicate microcosms in total for each lineage). For one biological replicate a single colony was picked from the agar plate and transferred into 6 ml liquid KB media. After growing overnight, 6 μl of each culture were used to inoculate fresh microcosms, which were kept under static conditions for three days (one selection round). After three days the cultures were thoroughly vortexed, diluted and spread on KB agar plates (Fig. 5.2).

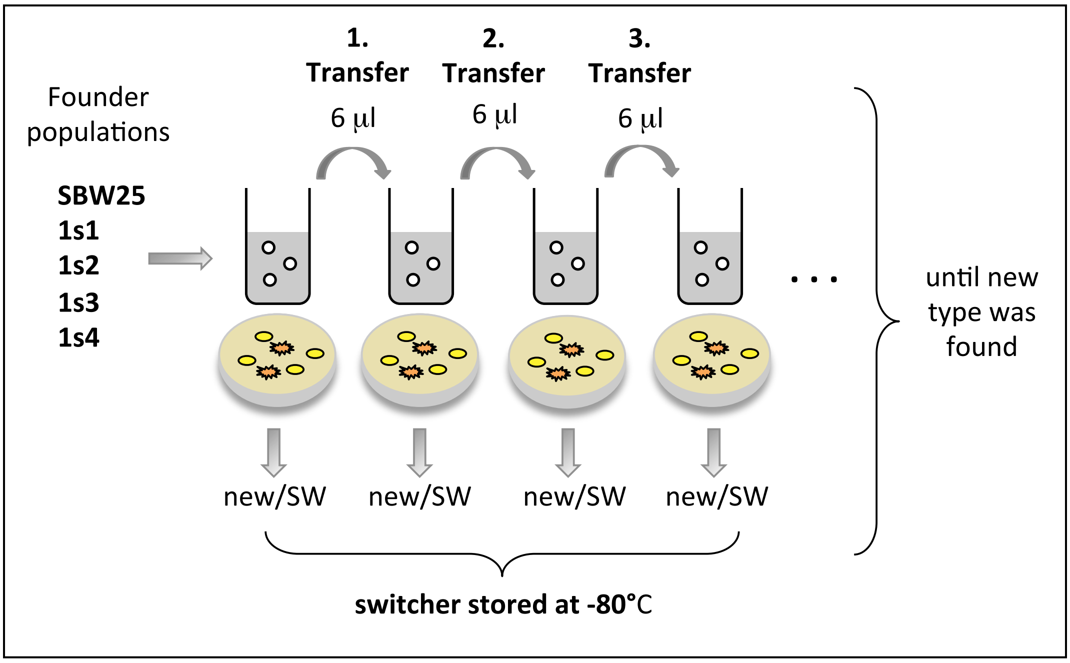

**Figure 5.2: Experimental design of the SREE.** Overnight cultures of SBW25, 1s1, 1s2, 1s3 and 1s4 acted as founder populations to replay the evolution of phenotypic switching under static conditions (60 microcosms per lineage). After three days' growth in a static microcosm the cultures were diluted, plated and screened for new colony types. If a new type or a switcher (new/SW) was detected the replicate was terminated. Switchers were stored at –80°C. If no new type was detected 6 µl of the culture were transferred to a new microcosm and incubated for another three days under static conditions before being diluted, plated and screened for new types again. This process was repeated until a new type was found.

The $OD_{600}$ was randomly measured for two to three microcosms in each strain to estimate the plating volume that was needed for four plates to be able to screen approximately 500-1000 colonies. After two days all plates were screened for new colony types. In addition one plate from each replicate was randomly chosen to count the numbers of ancestral colonies, the numbers of the most common colony type, the total numbers of new type colonies, numbers of switcher colonies and colony diversity (number of different colony morphologies). As soon as a new colony type was found on one of the four plates, the replicate line was terminated (Fig. 5.2).

Colonies that showed switching characteristics were transferred to new KB agar plates from which a single colony was used to inoculate an overnight culture, which was then stored at -80°C the following day. In addition, the new switching type was transferred to new agar plates several times in a row to certify the inheritance of the novel phenotype. Furthermore, negative staining with Indian ink confirmed a switcher phenotype on the cellular level under the microscope (see Chapter 1, Fig. 1.9C). Where no new type was observed on the plates, 6 μl of the three-day culture were transferred to a fresh microcosm, which was left undisturbed for another three days under the same environmental conditions (first transfer). This procedure was repeated parallel to diluting, plating and screening until a new type was found (several transfers; Fig. 5.2).

### 5.2.2 Screening the *carB* gene of re-evolved switchers

The switcher genotypes that evolved during the course of the SREE were further analysed and screened for an underlying mutation in the *carB* gene. A PCR was performed with the primer pair SN039 PF and SN040 PR, using the protocol that was described earlier (see Chapter 2, section *2.2.1.1*). The PCR products were sequenced using *carB* specific primers (SN009 PF, SN010 PF, SN013 PF, SN014 PF, SN015 PF, SN016 PF, SN017 PF, SN018 PF and SN070 PR; see Chapter 2, Tab. 2.3) and screened for changes in the DNA sequence by comparison to the *P. fluorescens* SBW25 ancestral DNA sequence using the alignment software Sequencher.

### 5.2.3 Statistical analyses

The software JMP 9.0 was used to perform generalised linear models using the data obtained during the SREE. The data included number of switcher that were detected on plates after the dilution of a population containing $10^9$ cells as well as number of new colony types and number of ancestors. Illustrations of the results were created with the software Prism 5.

## 5.3 Results

### 5.3.1 The re-evolution of switching genotypes in SBW25, 1s1, 1s2, 1s3 and 1s4 (SREE)

#### *5.3.1.1 Total number of re-evolved switcher genotypes*

The founder lineages SBW25, 1s1, 1s2, 1s3 and 1s4 evolved in Line 1 during the course of the REE at different evolutionary time points (see Chapter 1, section 1.4.4). The evolution of a switcher genotype from 1s4 in a static environment was already observed during the course of the REE (Beaumont *et al.*, 2009, Gallie, 2009). In this study I was interested in the likelihood of switcher occurrence at earlier time points represented by the original common ancestor SBW25 and the genotypes 1s1, 1s2 and 1s3 in addition to the immediate ancestor of the switcher, 1s4. The average number in 60 replicate microcosms in three experimental blocks (N = 3) that produced a switcher genotype was then calculated. The results of a generalised linear model show that the different genotypes vary in their probability to produce switcher genotypes ($X^2_6$ = 17.96, $P < 0.01$). Most of the difference was explained by the lineage itself ($X^2_4$ = 43.16, $P < 0.0001$). The genotypes 1s3 and 1s4 had the highest number of microcosms that produced a switcher (Fig. 5.3) and showed a similar likelihood of switcher occurrence ($X^2_4$ = 2.9 $10^{-31}$, $P$ = 1). Both were distinctive from the other lineages in their ability to produce switchers ($X^2_4$ = 16.47, $P < 0.0001$). No switcher was detected in 1s1 (Fig. 5.3)

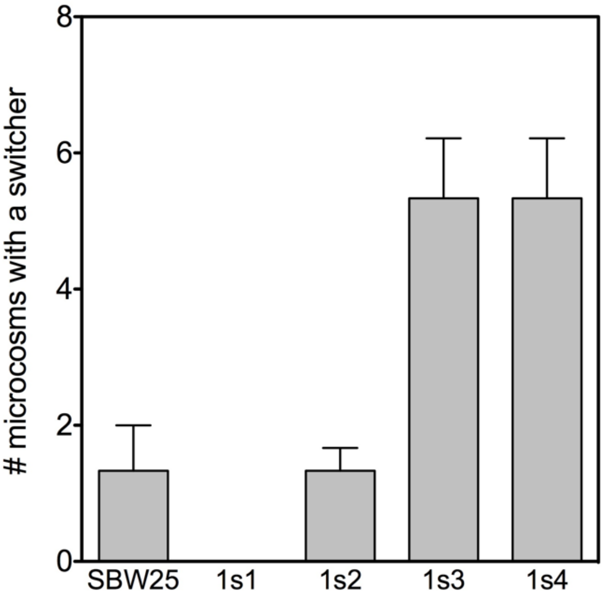

**Figure 5.3: Average number of replicates of SBW25, 1s1, 1s2, 1s3 and 1s4 that evolved switcher genotypes in the SREE.** The lineages acted as founder populations and represented different evolutionary stages of Line 1. Switchers evolved within one round of selection (three days) under static conditions. Displayed are SEM based on *N* = 3.

#### *5.3.1.2 Evolutionary routes of re-evolved switcher genotypes*

All re-evolved switchers that were detected on agar plates during the SREE were checked for a mutation in the *carB* gene using gene specific primers (see section 5.2.2). From a total of 37 re-evolved switchers in all lineages only one was found with a mutation in the *carB* gene. From the 1s4 genotype a *carB* switcher evolved in one replicate. The other switchers must have occurred based on genetic changes somewhere else in the genome.

##### *5.3.1.2.1 Time taken to detect new phenotypes or switchers in each genotype*

The aim was to investigate whether SBW25, 1s1, 1s2, 1s3 and 1s4 showed differences in the number of transfers that were needed for a new type (Fig. 5.4A) and for a switcher (Fig. 5.4B) to occur. The results show that SBW25, 1s1 and 1s2 diversified rather quickly. All microcosms produced a new type within the first three days (switcher not included) and were terminated. Most of the microcosms of 1s3 produced a new type within the first three days as well but in some cases one transfer was needed.

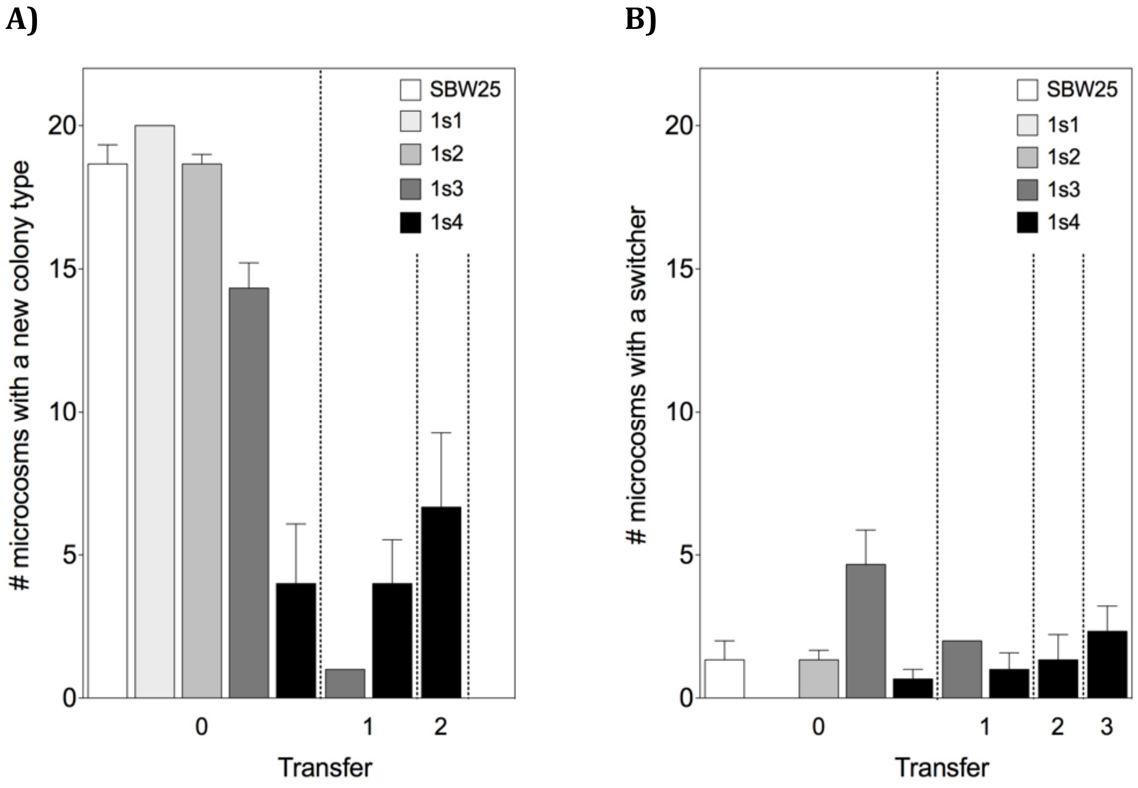

**Figure 5.4: Average number of microcosms of SBW25, 1s1, 1s2, 1s3 and 1s4 with (A) a new type or (B) a switcher genotype after each transfer.** Bacterial cultures were transferred to a new microcosm if no new type with distinctive colony morphology was found on the agar plates after three days and following dilution of cultures containing $10^9$ cells. Microcosms with a novel type were terminated. The switcher was not included in the graph illustrated in (A). Vertical dotted lines separate different transfers.

Evolution from 1s4 was much slower and followed a pattern quite different from the other lineages. Here only a small number of microcosms gave rise to a new type within the first three days. Some replicates were transferred up to two times before a new type was found (Fig. 5.4A), sometimes three times before a switcher was found. For 1s4 the number of microcosms that produced a new type was overall lower after each transfer.

In Figure 5.4B the number of microcosms with a switcher for each genotype and per transfer is illustrated. In SBW25 and 1s2 all switchers occurred within the first three days. No switcher was found in 1s1. Most of the switchers in 1s3 were found within the first three days but in some cases one transfer was needed. Up to three transfers were required for a switcher to occur from 1s4 (Fig. 5.4B). It is noticeable that after the third transfer in 1s4 all new types were switchers.

#### *5.3.1.3 Impact of new types on population composition*

In the previous section it was shown that SBW25, 1s1, 1s2, 1s3 and 1s4 differed in the number of transfers that were needed to give rise to a new type (Fig. 5.4A). I was interested in whether novel types that occurred in the different backgrounds showed differences in their ability to increase in cell frequency within a microcosm. Figure 5.5 demonstrates that the new types that appeared in SBW25, 1s1 and 1s2 reached relatively high cell frequencies within the first transfer. They accounted for 30-50% of the total population.

Novel phenotypes that occurred in 1s3 within the first three days were not able to reach similar cell frequencies. The proportion of the new type stayed below 5% of the total population. It is noticeable that in one replicate of 1s3 after a transfer took place newly occurred types were able to increase their cell frequencies up to 50% of the population. In 1s4 the proportion of the new types remained below 5% of the population regardless of the number of transfers (Fig. 5.5).

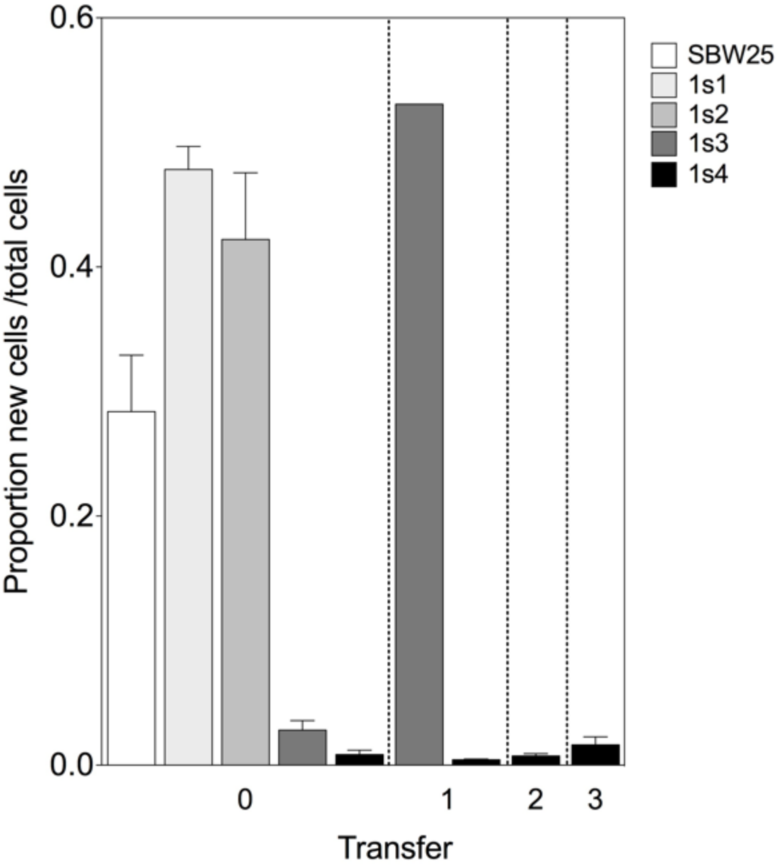

**Figure 5.5: Cell frequencies of novel types in proportion to the total cell number in populations of SBW25, 1s1, 1s2, 1s3 and 1s4 per transfer.** The proportions were calculated from the colony numbers of new types and total colony number on an agar plate that were counted after three days in a static environment after each transfer following dilution of cultures containing $10^9$ cells. Displayed are SEM based on $N$ = 3. Vertical dotted lines separate different transfers.

### *5.3.1.4 Population composition and switcher occurrence*

The most switcher types were observed in 1s3 and 1s4 in comparison to the lineages SBW25, 1s1 and 1s2 (Fig. 5.3). Population composition might have an impact on the likelihood of switcher genotype occurrence. It becomes apparent that the new types that occurred in 1s3 and 1s4 did not reach very high cell frequencies but the numbers of microcosms containing a switcher were relatively high (Fig. 5.6). The opposite scenario may be the case for SBW25 and 1s2. Here not many switchers were found but other new types were found that in addition reached higher cell densities. The lineage 1s1, which did not produce a switcher, showed the highest frequency of novel types in the populations (Fig. 5.6).

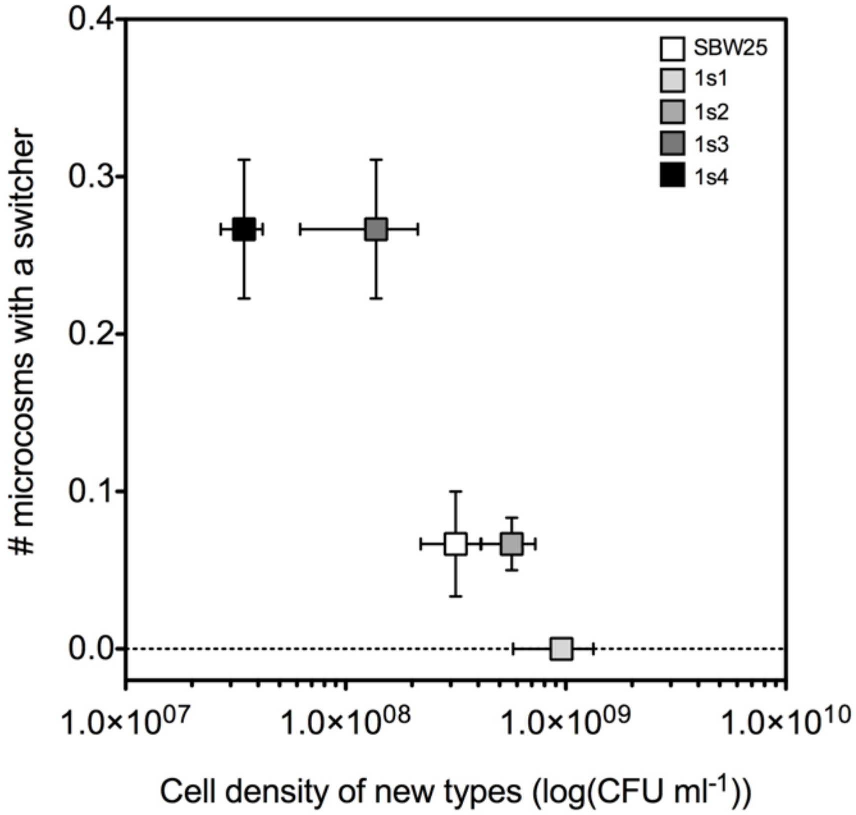

**Figure 5.6: Average proportion of microcosms with a switcher over the average cell densities of new types over all genotypes.** Values for cell densities are based on colony counts on agar plates after three days' growth in a static microcosm. In 1s4 several transfers were needed until a new type arose. Displayed are SEM based on *N* = 3.

### *5.3.1.5 Phenotypic diversity and switcher occurrence*

During the SREE the founder population SBW25, 1s1, 1s2, 1s3 and 1s4 diversified after three days in a static environment into multiple, different novel types, which were observed as new colony morphologies on agar plates. Re-evolution from

SBW25 led on average to four different colony types, which included the ancestor SBW25. The average number of different colony morphologies did not change in replicates in which a switcher genotype was found. In 1s1 the average number of colony types was around four and no switcher was found in this lineage. The founder lineage 1s2 produced on average four different types (including 1s2). In microcosms where a switcher was detected the number increased to an average of five different colony morphologies, likely caused by the additional switcher colony type.

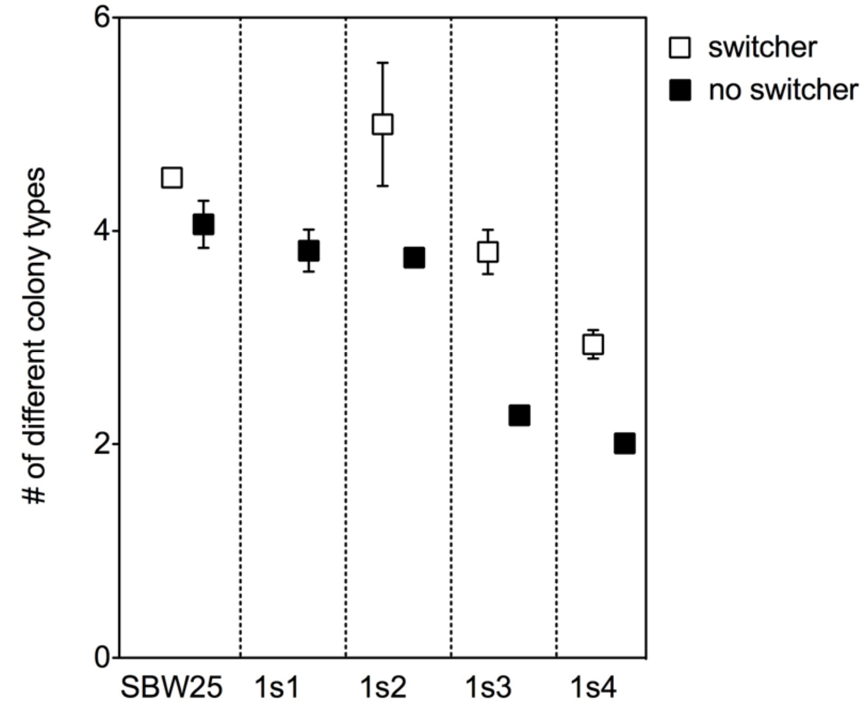

**Figure 5.7: Average number of different colony types in each lineage found on agar plates after diversification.** The different colony types include the colony type of the ancestral strain and the switcher that were found on one agar plate that was randomly chosen from four plates of each replicate in every lineage in the SREE. Displayed are SEM based on *N* = 3.

Overall SBW25, 1s1 and 1s2 produced roughly similar numbers of new types. The lineages 1s3 and 1s4 showed two different colony types, which included the unevolved 1s3 and 1s4. In replicate microcosms with a switcher genotype the number of different types increased up to four (1s3) or three (1s4), which included the switcher as a new type (Fig. 5.7).

Altogether the results of a generalised linear model suggest that there was no significant effect of the number of different colony types on the likelihood of switcher occurrence. The integration of the factors lineage and number of different types into the model however resulted in a significant impact of the two factors together in switcher occurrence ($X^2_4 = 11.7$, $P < 0.02$). Here the emergence of a switcher was dependent on the average number of different colony types within each lineage.

## 5.4 Discussion

Several research groups have studied the genetic basis of the accessibility of evolutionary pathways (Lunzer *et al.*, 2005; Weinreich *et al.*, 2006; Poelwijk *et al.*, 2007; de Visser *et al.*, 2009; Franke *et al.*, 2011; Khan *et al.*, 2011; Kvitek & Sherlock, 2011). They have used genetic tools to artificially construct fitness landscapes and examined possible mutational pathways and their likelihood of being taken by an organism. Scientists have found, for example, that for a particular trait such as resistance against antibiotics there are many mutational pathways that can lead to a resistant phenotype, although only a few trajectories show a high probability (Weinreich *et al.*, 2006). Whether these artificially obtained trajectory estimates are sufficient to predict real-time evolution is still an open question.

In this study experimental evolution was used to investigate real-time evolution of a switching genotype from various founder populations (SBW25, 1s1, 1s2, 1s3 and 1s4) representing different time points of the evolutionary history of the switcher lineage Line 1. During the REE, Line 1 was one of two lineages that evolved a switching genotype amongst 12 replicate *P. fluorescens* lines. The novel switcher type occurred in Line 1 as a result of a mutation in the *carB* gene (*carB**) after eight preceding mutations (Tab. 5.1). Introducing *carB** into the genomes of SBW25, 1s1, 1s2, 1s3 and 1s4 and measuring the fitness effect revealed that *carB**, according to its fitness benefit, has the potential to already occur in the ancestor SBW25 and

1s3. Both genetic backgrounds, in addition to 1s4 (immediate ancestor of the switcher), were highly accessible for *carB** based switcher evolution (Fig. 5.1). In this study I was interested in whether this potential to produce a switching genotype could be observed during real-time evolution or whether other factors limited the detection of a switcher genotype during the REE in preceding genotypes of the switcher. Are there other factors apart from genetic constraints that impact on switcher evolution and how do they affect switcher occurrence with progressing evolutionary history?

### 5.4.1 Impact of evolutionary history on real-time evolution

SBW25, 1s1, 1s2, 1s3 and 1s4 acted as founder populations for the re-evolution of a switching genotype. The results of this experiment clearly show that the likelihood of switcher occurrence varied significantly between SBW25, 1s1, 1s2, 1s3 and 1s4. The average number of switchers was dependent on the genotype in which they occurred (Fig. 5.3). The results are only partly in agreement with the findings of the *carB** reconstruction experiment (Fig. 5.1). The genotypes 1s3 and 1s4 showed a high average switcher number as was to be expected because of a fitness benefit of *carB** in both genotypes. Similarly it was no surprise that no switcher, or a very low number of switchers, was observed in 1s1 and 1s2 since *carB** caused a significant fitness decrease in these genetic backgrounds (Fig. 5.1). Although unexpected, the detection of switching genotypes in 1s2 clearly shows that even with a negative *carB** impact on fitness, switcher evolution is nevertheless achievable. Surprisingly, not many switchers were found in SBW25, which according to the large fitness benefit of *carB** was assumed would produce a high switcher number (Fig. 5.1). All together earlier genotypic stages of the evolutionary history such as the ancestor SBW25, 1s1 and 1s2 had a lower likelihood of switcher occurrence compared to 1s3 and 1s4, which represent genotypes that occurred at a later time point along the evolutionary history of Line 1 (Fig. 5.3). This suggests that history has an impact on the capacity to evolve switcher genotype (Beaumont *et al.*, 2009), but that this is not solely determined by genetic constraints (Lenski *et al.*, 1991; Lenski & Travisano, 1994; Blount *et al.*, 2008; Barrick *et al.*, 2009). Other factors make earlier genotypes (SBW25, 1s1, 1s2) less accessible for switcher evolution than later genotypes (1s3, 1s4; Franke *et al.*, 2011).

In addition to the increased switcher number that was observed in 1s3 and 1s4, they also varied from earlier genotypes in the number of transfers that were needed to produce a new type or a switcher (Fig. 5.4A). In some cases they had to be transferred up to three times before any new type or switcher genotype was observed. Thus, evolution appears to be slower. However, it is likely that new variants arose earlier in both backgrounds but were undetected on agar plates due to a low cell frequency. Phenotypic and genotypic analyses by Gallie (2009) revealed that 1s3 and 1s4 produce, to some degree, a cellulose polymer, which is not the case in SBW25, 1s1 and 1s2 (Gallie, 2009). Consequently the 1s3 and 1s4 types can already survive quite well in a static environment. The cellulose polymer enables them to occupy the air-liquid interface due to biofilm formation (Spiers *et al.*, 2002; Gallie, 2009). As a result other variants that arise have to be able to compete with 1s3 and 1s4 that are already successfully growing in the preferred niche. Under these circumstances new variants perhaps appear but cannot increase in frequency and therefore stay unrecognised. The lack of additional novel beneficial types might offer the opportunity for a switcher to increase in cell numbers.

### 5.4.2 The genetic basis of switcher evolution along the evolutionary pathway in Line 1

Sequencing of the entire *carB* gene of all re-evolved switchers revealed that *carB** is just one out of multiple possible pathways that can lead to a switcher genotype. Parallel phenotypic but divergent genotypic evolution was observed in a previous chapter (see Chapter 3). In that case, three out of four independently evolved *P. fluorescens* lineages produced a switcher based on different genetic mechanisms and confirmed that switcher evolution is not as rare as once thought. In this study re-evolution was performed from various genotypes (different evolutionary time points) within one single lineage, Line 1. Here each subsequent genotype that was used as a founder population for the SREE differed in only two additional mutations from the previous genotype (Tab. 5.1) and one could assume, based on the similarity of the mutations and their order, that the switcher pathway is perhaps predetermined or predictable and a mutation in *carB* is highly likely to be a common route to switcher evolution.

However, the sequencing results suggest otherwise. Only one out of 37 observed switchers had a mutation in *carB*, indicating that the underlying genetic changes in the remaining re-evolved switchers must have appeared elsewhere in the genome. These results are in agreement with the findings of a previous chapter (see Chapter 3). Even when histories overlap, genotypic evolution is not predictable because multiple mutational pathways can give rise to a switcher genotype (Lenski *et al.*, 1991; Korona *et al.*, 1994; Lenski & Travisano, 1994; Korona, 1996; Nakatsu *et al.*, 1998; Wichman *et al.*, 1999; Cooper & Lenski, 2010; Meyer *et al.*, 2012). Within the colanic acid pathway, which is tightly linked to the pyrimidine pathway, other mutational targets can cause a similar functional change. Here the flux through the pyrimidine pathway is reduced and as a consequence some cells of the population start to produce colanic acid (Gallie, 2009).

### 5.4.3 Impact of population dynamics and phenotypic diversity in switcher evolution in Line 1

Studying population dynamics and their impact on evolutionary processes in microorganisms such as bacteria is challenging. Evolution is the result of random mutations and selection under given environmental conditions. Environments are composed of abiotic and biotic factors and how an environment changes over time is presumably unpredictable (Lewontin, 1966; Beatty & Desjardins, 2009). The surrounding environment of an organism is dynamic and constantly changing, with consequences for evolutionary outcomes (Franke *et al.*, 2011). The aim of this study was to obtain information about the impact of population composition on switcher evolution and how population dynamics change during evolution depending on historical genotypes that occurred along the evolutionary path in Line 1.

Earlier genotypes (SBW25, 1s1, 1s2) are less likely to produce a switcher genotype than genotypes that appeared at a later stage along the evolutionary history (1s3, 1s4; Fig. 5.3). In addition to the number of observed switchers, the total number of new types, the proportion of new types in the total population, cell densities and phenotypic diversity (number of different colony morphotypes) within a microcosm were recorded for each replicate to get an estimate of how the population compositions developed within three days of growth and diversification in a static environment. As has been pointed out earlier, SBW25, 1s1 and 1s2 evolved faster than did 1s3 and 1s4. Nearly all replicates of SBW25, 1s1 and 1s2 produced a new phenotype within the first three days in a static microcosm (Fig. 5.4). In addition, new types were able to reach high cell frequencies (Fig. 5.5), indicating that some new variants had a substantial advantage under these conditions. At the same time the likelihood of switcher occurrence decreased (Fig. 5.6). It has been shown that genetic constraints in 1s1 and 1s2 are likely to limit the evolution of a switcher (Fig. 5.1), perhaps due to negative genetic interactions (Khan *et al.*, 2011). In SBW25 however it is possible that the variety of ecological interactions between the many different types that appeared quite quickly decrease the likelihood of switcher occurrence. For example, strong competition for the air-liquid interface may have prevented the establishment of a switcher genotype within the population.

The comparison of the different founder populations shows that diversity was lower in 1s3 and 1s4 than in the genotypes SBW25, 1s1 and 1s2 (Fig. 5.7). Low frequency of new types and lower diversity after diversification of genotypes that occurred later along the evolutionary pathway suggest that the supply of beneficial mutations was higher in the beginning and decreased over time (Lenski & Travisano, 1994; Arjan *et al.*, 1999). Due to the increased supply of beneficial mutations many different novel types evolved simultaneously (McDonald *et al.*, 2009; Sniegowski & Gerrish, 2010) in the beginning and prevented any switcher mutation that might have occurred of becoming established in the population. Later on the rate of beneficial mutations decreased, which may explain the increased number of transfers that were needed to detect a new type in 1s3 and 1s4. The lack of beneficial mutations, on the other hand, might have increased the chance of switcher mutations becoming fixed within populations of 1s3 and 1s4 before another beneficial mutation arose. Here timing of occurrence of the switcher mutation determines the success of the novel phenotype (Hegreness *et al.*, 2006). The lack of beneficial mutations can have different causes. As addressed earlier, the high fitness of 1s3 and 1s4 because of cellulose production might decrease the chances of mutations that imply only a small benefit (e.g. weak biofilm). Alternatively, common evolutionary routes might become depleted over the course of evolution. Other mutational pathways that can improve the performance under present environmental circumstances have to be found, which may take a longer time. This will be discussed in more detail in the next section.

##### 5.4.3.1.1 Depletion of evolutionary pathways with proceeding evolution

The capacity to evolve a switcher based on *carB** has been shown to exist in the ancestor SBW25. Here *carB** had a positive effect on fitness, indicating that a switcher can increase in frequency within a SBW25 population (Fig. 5.1). This raises the question as to why no switcher was found in the first selection round during the REE in any of the 12 parallel lineages, since they started from the same ancestor.

The evolutionary history of Line 1 (Tab. 5.1) is the result of evolution in two alternating environments, static and shaken, starting from the common ancestor *P. fluorescens* SBW25. It is known that *P. fluorescens* SBW25 diversifies quickly when growing in a static KB filled microcosm (Rainey & Travisano, 1998). Within hours usually, WS types arise as a consequence of an oxygen gradient that develops within the media. The WS types are able to colonise the air-liquid interface due to biofilm formation, which is the result of the overproduction of a cellulosic polymer

(Spiers *et al*., 2003; Rainey & Travisano, 2003). The transfer of a WS type into a shaken environment generally causes the reversal to an ancestral-like SM type (Koza *et al.*, 2011). Suppressor analysis of novel genotypes that evolved during the course of the REE in the 12 replicate lineages identified important common genetic pathways that underlie the evolution of WS and SM types. In many cases novel types showed mutations in *wsp*, *wss*, *aws* and *mws* (see Chapter 1, section 1.4.3; Kahn *et al.*, 1998; Bantinaki, 2001; Spiers *et al.*, 2002; Spiers *et al.*, 2003; Gehrig, 2005; McDonald *et al.*, 2009). Noteworthy is the modular nature of mutation occurrences in all 12 lineages, such as that one mutation induces the synthesis of a cellulosic polymer under static conditions followed by a mutation in the same locus that turns off the production under shaken conditions. For example in Line 1 the first mutation appeared in *mwsR* and led to a WS type (Tab. 5.1). The subsequent mutation took place again in *mwsR* and caused a SM-like phenotype. In the next selection round the WS type was induced by a mutation in *aws*. After the change to a shaken environment another mutation in *aws* caused the reversal to the SM type. This modularity was not observed for the fifth and sixth mutations (*wspF* and *wssA*), but after that was observed again for the seventh and eighth mutations, which occurred in *mwsR* (Tab. 5.1). Although the modular organisation was interrupted by mutations in *wspF* followed by *wssA*, they still represent two out of four common mutational pathways involved in WS and SM evolution (Tab. 5.1).

The switcher genotype was detected during the ninth selection round and had a mutation in *carB*, which was not involved in any of the phenotypic changes that occurred previously. Other studies found a correlation between the fitness of a *carB** mutation and the proportion of WS types (Gallie, 2009). For example the fitness benefit of *carB** in SBW25 decreased when competition between SBW25 and SBW25*carB** took place over 72 hours instead of 48 hours, due to the occurrence of novel WS types. This fitness decrease was not observed in 1s4 after 72 hours because evolution was much slower. It was suggested that many of the common mutational routes that can lead to a WS type (*wsp*, *wss*, *aws* and *mws*) are available in genotypes that evolved during earlier selection rounds (e.g. SBW25, 1s1, 1s2). With proceeding evolution these common pathways are gradually removed due to mutations and are therefore not available in 1s4 for WS evolution (Gallie, 2009). The depletion of common mutational pathways over time may explain the increased duration of evolution, the reduced proportion of novel types, the decreased diversity that was observed during the SREE in 1s3 and 1s4 in this study, and an elevated likelihood of switcher emergence (Arjan *et al.*, 1999). Recently, most of the mutational pathways of all 12 replicate *P. fluorescens* lineages have been identified. The data suggest that the overlap in genetic changes is higher in the beginning of the evolutionary pathways and that lineages tend to have mutations in more diverse loci the longer they evolve (data not published). This suggests that after common mutational pathways have been used (McDonald *et al.*, 2009), less likely and more diverse mutational pathways can arise that increase the chance of survival in that particular environment. It has been suggested that common mutational targets, because of genetic architecture and functionalities, have a higher rate of being translated into adaptive phenotypic change (McDonald *et al.*, 2009). This may explain the high number and the fast evolution of new types in SBW25, 1s1 and 1s2, where more of these common mutational pathways are still available, however it remains curious that some pathways such *mwsR* can be used more than once in the evolution of a WS (Tab. 5.1).

### 5.4.4 Comparison of theoretical capacity of switcher evolution and real-time switcher re-evolution

Figure 5.8 summarises the results from the fitness experiment after *carB** was artificially introduced into SBW25, 1s1, 1s2, 1s3 and 1s4 and the results from the SREE in which the same genotypes acted as founder populations to replay the evolution of a switching genotype. The results from the fitness experiments coincide with the results from the SREE for the lineages 1s1, 1s2, 1s3 and 1s4. When considering the fitness effects of *carB** in the different backgrounds it would be expected that 1s1 and 1s2 are unlikely to give rise to a switcher genotype. Here the fitness effect of *carB** was negative and no switcher (1s1) or a low number of switcher genotypes (1s2) were detected during the SREE. Similarly, 1s3 and 1s4 were expected to have a higher likelihood of switcher occurrence based on a positive relative fitness of *carB** in both backgrounds. This is in agreement with the high number of switcher genotypes that were found in both genotypes in the SREE (Fig. 5.8). It is noticeable that for SBW25 the results of the fitness assay and the SREE do not match. Although *carB** improved the performance when introduced into the ancestral genome of SBW25, the re-evolution experiment resulted in a low number of switchers (Fig. 5.8). Evolutionary history has an effect on switcher evolution (Beaumont *et al.*, 2009). Earlier genotypes are less likely to evolve a switcher than are later genotypes. This observation cannot be explained by the impact of genetic constraints alone (Lenski *et al.*, 1991; Lenski & Travisano, 1994; Blount *et al.*, 2008; Barrick *et al.*, 2009). Instead the results of this study suggest that ecological interactions (Franke *et al.*, 2011) within a microcosm between novel types after diversification differ in each genotype and have an affect on switcher occurrence.

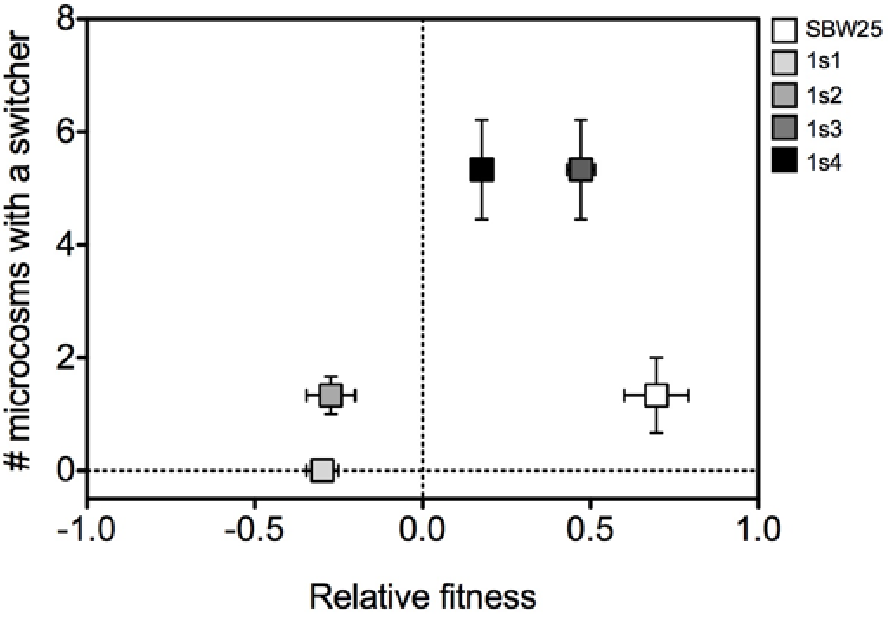

**Figure 5.8: Average number of microcosms with a switcher (SREE) plotted against the relative fitness of *carB** in the different genotypes.** The combined results from the fitness assays after *carB** was introduced into SBW25, 1s1, 1s2, 1s3 and 1s4, and the results from the SREE, where the same lineages acted as founder populations to re-evolve a phenotypic switcher, are illustrated. Displayed are SEM based on $N$ = 3 (y-axis: # microcosms with a switcher) and SEM based on $N$ = 8 (x-axis: relative fitness).

If a genotype gives rise to many new types, usually genotypes that occurred during earlier selection rounds (SBW25, 1s1 and 1s2), it is less likely for a switcher to increase in frequency. Whereas in the presence of fewer new types, a switcher has the opportunity to increase in cell numbers and is more likely to be observed. This was likely the case for genotypes that occurred at a later time point in the evolutionary history (1s3 and 1s4).

The increased likelihood of switcher occurrence in 1s3, together with the large fitness increase after *carB** was introduced into the genome of 1s3, suggest that the underlying genotype promotes the evolution of a switcher. The genotype of 1s3 has only two additional mutations compared to 1s2 (Tab. 5.1). Here *carB** had a negative fitness effect and only a few switchers were found during the course of the SREE. The large variation of *carB** fitness and switcher occurrence between 1s2 and 1s3 indicate that the genetic background in 1s3 contributes to the evolution of a switcher since environmental conditions were kept similar. It is likely that an earlier evolutionary event (*wspF* or *wssA*) increased the likelihood of switcher occurrence as the result of positive epistatic interactions (Segre *et al.*, 2005; Weinreich *et al.*, 2006; Blount *et al.*, 2008; Blount *et al.*, 2012). Epistatic interactions and the impact on switcher evolution will be investigated in the following chapter (see Chapter 6).

## 5.5 Summary

This study represents an example of the impact of evolutionary history on the evolution of a novel trait. Genotypes from different evolutionary time points in Line 1 differ in their likelihood to evolve a switching phenotype during real-time evolution (SREE). The results of the experiment demonstrate that genotypes from late evolutionary time points (1s3 and 1s4) are more likely to produce a switcher than earlier genotypes (SBW25, 1s1 and 1s2). The first hypothesis (see 5.1.3) is therefore rejected. It has been shown that in addition to random mutations that accumulate over time, and selection, there are other factors that can limit or promote the evolution of a novel trait. The results of this study suggest that apart from genetic constraints, ecological interactions, which drive population dynamics, can explain the different likelihoods of switcher occurrence along the evolutionary pathway in Line 1. Therefore the third hypothesis (see 5.1.3) is supported. Genotype interactions such as competition for the air-liquid interface are likely to determine whether a switcher can establish within the population or not. The ability to produce a switcher can change over the course of evolution due to variations in population composition. Evolutionary history can influence the outcome of evolution in several ways. The mutational history can have a direct impact. Genetic constraints in some genotypes might prevent the occurrence of a switcher because of decreased fitness. In other genotypes a preceding mutation perhaps interacts with a switcher mutation (epistasis) and increases the likelihood of switcher evolution due to a fitness increase. In addition, common mutational routes that give rise to beneficial WS types can be depleted over the course of evolution. As evolutionary time goes on the fitness spectrum becomes constrained as adaptation proceeds which is shown in by less ecological diversity. Here switcher evolution becomes more likely because of the lack of fit competitors.

The theoretical capacity to evolve a switcher, which was determined by the performance of fitness experiments between non-*carB** strains and *carB** strains in Chapter 3 (Fig. 5.1), only partly translates into real-time evolution. Furthermore switcher phenotypes that evolved from different founder populations differed in their mutational pathways. The diverse genotypic evolution indicates that multiple pathways to a switcher exist. There was a low degree of parallel genotypic evolution and the second hypothesis (see 5.1.3) is therefore rejected. The results of this study illustrate the importance of ecological interactions for the outcome of switcher evolution, which are also dependent on the present genotype such as the ability to evolve new types. Complex interactions between many different types are difficult to study. Studies that overcome these challenges can provide valuable information about the interplay between genetic variation, selection and ecological dynamics and to what degree and under what circumstances they determine the outcome of evolution.

# Chapter 6: Epistasis and the evolution of stochastic switching in *P. fluorescens*

## 6.1 Introduction

Genetic variation is the foundation for the evolution of key innovations. Understanding how variations at the genetic level can influence future evolutionary change is still incomplete. For example, as organisms evolve, mutations accumulate in genomes. Some of these mutations may have no immediate effect on the phenotype, but interactions with a future mutation may become effective with progressing evolution. Here the interaction can have consequences for the phenotype and the fitness of the organism and propel a population along a particular evolutionary path. It is still unclear as to what degree evolutionary events in the future depend on evolutionary events in the past.

The evolution of new adaptive traits can provide access to new ecological niches and consequently improve the performance of an organism in that particular environment. It is well known that genetic variation such as single mutations can lead to changed gene functions and can provide a selective advantage due to a novel trait. For example, genetic changes can result in the utilization of a new resource (Hall, 1981; Negoro *et al.*, 1994; Blount *et al.*, 2008), or can provide a new function that opens new pathways (Duffy *et al.*, 2007; Meyer *et al.*, 2012). It has been suggested that evolution of a phenotypic innovation can be a stepwise process based on consecutive mutations that have been established over time. In such cases every additional mutation causes a slightly better performance of an already existing function in the ancestor (Lenski *et al.*, 2003; Lamb *et al.*, 2009). Based on phylogenetic comparisons and an increased availability of genetic information it is thought that the evolution of complex organs such as the eye of higher organisms could have evolved by successive aggregations of mutations (Goldsmith, 1990; Yokoyama & Radlwimmer, 2001; Lamb *et al.*, 2009). Furthermore, the evolution of a new trait can depend on single or multiple genetic changes that occurred at earlier time points during an organism's mutational history and may therefore be highly contingent on a particular genotypic background. Here the qualitative performance of a new phenotype may be contingent upon one or more mutations that occurred previously due to epistatic interactions (Weinreich *et al.*, 2006; Blount *et al.*, 2008; Blount *et al.*, 2012; Meyer *et al.*, 2012). I used epistasis as defined by Eckardt (2008): "Epistasis describes a nonadditive genetic interaction that results from the activity (or mutation) of one gene masking the phenotype or effect caused by the activity of another gene" (Eckardt, 2008:p.1).

### 6.1.1 Epistasis and the evolution of novel traits

It is widely accepted that epistasis can give rise to new functions but it is still unknown how commonly it occurs and to what extent it influences evolutionary dynamics. Epistasis plays a central role in processes such as speciation (Gavrilets, 2004), in the evolution of sexual reproduction (Kondrashov, 1988; de Visser & Elena, 2007), and it has been observed during the course of long-term selection (Carlborg *et al.*, 2006; Blount *et al.*, 2008; Blount *et al.*, 2012). Fitness changes of a novel phenotype due to interacting mutations can be bidirectional. If the fitness interaction between mutations results in a lower fitness than each mutation alone, epistasis has a negative effect on fitness, but if the combined fitness is higher than the fitness of the mutations by themselves, epistasis has a positive fitness effect. The fitness consequences due to the combined effect of two or more mutations are crucial for the evolution of novel traits based on such interactions. The new phenotype is likely to increase in frequency and can establish within the population only if the combined fitness of the interacting mutations is higher than the fitness of the ancestor. Hence the evolution of the novel phenotype is highly dependent on the present genotypic background.

Studies of protein evolution demonstrate a strong impact of the chronological order of mutations on the number of available evolutionary pathways. These have shown that epistatic interactions within a single genetic locus can accommodate strong constraints. Out of many possible evolutionary paths (random order of the mutations) only a limited number of paths (specific order of the mutations) are actually accessible by evolution, due to the fitness effect of the epistatic interactions (Bridgham *et al.*, 2006; Miller *et al.*, 2006; Weinreich *et al.*, 2006; Ortlund *et al.*, 2007). In addition to the characterization of the functional nature of the epistatic interactions these studies have given insights into long-term evolutionary consequences of epistasis. For example, Ortlund *et al.* (2007) used phylogenetic analysis to reconstruct the precursor of the vertebrate glucocorticoid and mineralocorticoid receptors (steroid receptors) and tested its function. It appeared that the ancestral protein was not very specific to different steroid ligands. It was concluded that the more specific glucocorticoid receptor evolved from a universal ancestral mineralocorticoid receptor. In order to gain a functional glucocorticoid-specific receptor a set of eight genetic changes was identified that showed a high dependency on the chronological order in which they occurred. Three of the mutations had no direct effect on the function of the receptor but were important for the stabilization of the protein. Only after stabilization were the next three mutations, which were needed to determine specificity of the receptor, able to occur. Two further substitutions were non-functional or even destroyed the receptor on their own but together they changed the conformation of the protein, which was needed to bind to cortisol. These results show that genetic changes that have a stabilizing effect but no immediate functional impact can be crucial for future evolutionary events based on epistasis (Ortlund *et al.*, 2007).

The impact of epistasis in the evolution of a key innovation has been described by Blount *et al.* (2008). In this work a new variant evolved during a long-term evolution experiment with *E. coli* (Lenski *et al.*, 1991; Lenski, 2004) that was able to utilise a new carbon source. Twelve genetically identical *E. coli* populations were propagated independently in a glucose-limited medium that also contained citrate, which cannot be used as a carbon source by *E. coli* in the presence of oxygen. The cultures were daily transferred to fresh media. After 31,500 generations, in one of the 12 lineages, a new type evolved that was able to exploit citrate, and so increased in frequency. ‘Replay-evolution’ experiments from different time points using the ‘frozen fossil record’ provided evidence of a higher likelihood for a citrate utiliser to evolve after 20,000 generations, indicating that a potentiating mutation must have arisen at just about this time point. This suggested a high contingency of the evolution of citrate usage on the specific genotype, which was likely to have been caused by epistatic interactions. In a recent study, Blount *et al*. (2012) described in detail the underlying genetic mechanism that enables *E. coli* to use citrate as a carbon source, but the potentiating mutation and the nature of the epistatic interaction responsible for the establishment of the novel trait has not been identified yet.

### 6.1.2 Investigation of epistatic interactions in the evolution of a switching phenotype

Former studies (Beaumont *et al.*, 2009; Gallie, 2009) have led to a detailed knowledge of the evolutionary history of the stochastic switching phenotype in *P. fluorescens* that evolved during the Reverse-Evolution Experiment (REE) in Line 1 (see Chapter 5, Tab. 5.1). Furthermore, characterisation of mutations that occurred during the history of the REE (see Chapter 1, section 1.4.4) provided insights into the underlying genetic mechanisms that were affected by the mutations and caused phenotypic changes (Goymer *et al.*, 2006; Bantinaki *et al.*, 2007; Malone *et al.*, 2007; McDonald *et al.*, 2009). This information together with the ‘frozen fossil record’ made it possible to go back in evolutionary time as far as the ancestor SBW25 and examine the impact of each mutation that occurred during the REE in the evolution of the switching phenotype. Results from previous chapters and from other studies (Beaumont *et al.*, 2009) suggest that evolutionary history may have an impact on the evolution of phenotypic switching in *P. fluorescens*. It is possible that a potentiating mutation that arose at an earlier evolutionary time point during the REE increased the likelihood of switcher occurrence, perhaps due to an epistatic interaction.

Previous chapters have shown that the evolution of phenotypic switching in *P. fluorescens* SBW25 in Line 1 is not a stepwise additive process where each mutation that occurred throughout the history contributes to the evolution of a switching phenotype. The fact that the introduction of *carB** into the ancestral SBW25 genome had a positive effect on fitness (Fig. 6.1) indicates a lack of dependency on the eight mutations that occurred during history. Nonetheless the introduction of *carB** into the genomes of types that occurred later on caused distinct fitness effects. For example, when *carB** was introduced into 1s1 and 1s2, the beneficial effect of *carB** disappeared (Fig. 6.1). A switcher genotype based on *carB** cannot evolve from 1s1 and 1s2 because it would not be able to increase in frequency. This indicates genetic constraints that limit the evolution of a *carB**-based switcher genotypes in both genotypic backgrounds. It is noticeable that *carB** caused a fitness decrease in 1s2 but a large fitness increase in 1s3 (Fig. 6.1, yellow box plots). In addition, significantly more switchers were found to evolve from 1s3 than from 1s2 (see Chapter 5, Fig. 5.3). The increase in fitness when *carB** was introduced into the genomes of 1s3 and 1s4 showed that the earlier constraints were overcome and the evolution of a switching phenotype based on *carB** became possible (Fig. 6.1).

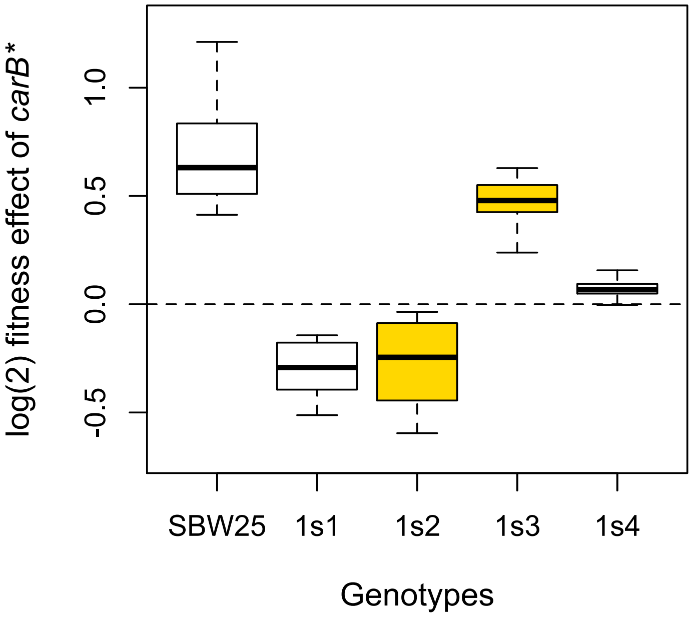

**Figure 6.1: Fitness effect of *carB** on genotypes that occurred during the course of evolution in Line 1.** Yellow box plots shows relative fitness of *carB** in 1s2 (negative fitness) and 1s3 (positive fitness). Displayed are medians and interquartile ranges based on $N = 8$. Results were obtained in Chapter 4.

**Table 6.1: Mutational histories of 1s2 and 1s3.** The evolved strain 1s3 had two additional mutations (in *wspF* and *wssA*) compared to 1s2.

| Mutation | Genotype | |
|---|---|---|
| | **1s2** | **1s3** |
| 1 | *mwsR* | *mwsR* |
| 2 | *mwsR* | *mwsR* |
| 3 | *awsX* | *awsX* |
| 4 | *awsR* | *awsR* |
| 5 | - | ***wspF*** |
| 6 | - | ***wssA*** |

When comparing the underlying genomes of 1s2 and 1s3 (whole genome sequencing; Gallie, 2009) it becomes apparent that 1s3 differs only in two additional mutations from 1s2, a mutation in *wspF* and in *wssA* (both with insertions; Tab. 6.1). One or both mutations can be expected to contribute to the positive fitness effect of *carB** in 1s3 in a static environment due to epistatic interactions and thus increases the likelihood of a switching phenotype evolving.

Unravelling the contribution of successive mutations in *wspF* and *wssA* to fitness effects of *carB** in 1s3 requires recognition of the central role of *wsp.* The *wsp* operon has been described in detail in the introduction (see Chapter 1, section *1.4.3.2*; Fig. 6.4), and although much is known about the role of *wspF* within the operon (Bantinaki *et al.*, 2007), other functions and interactions are possible that have not yet been identified. It is noteworthy that the mutation in *wspF* leads to constitutive activation of *wspR* and the overproduction of c-di-GMP within the cell (Bantinaki *et al.*, 2007). This messenger molecule binds to enzymes of *wssA-J* and initiates the synthesis of cellulose. The *wssA* gene is part of the cellulose-producing unit (*wssA-J*), and is thought to be involved in the localization of the biosynthetic complex within the cell (Spiers *et al.*, 2002). The mutation in *wssA* was discovered via genome sequencing in a genotype that evolved under shaking conditions in Line 1 (1w3). A mutated *wssA* gene causes decreased cellulose production but the production of the polymer however is not abolished completely (Gallie, 2009). This indicates that perhaps another gene of the *wssA-J* operon adopts the function of *wssA* (Spiers *et al.*, 2002) or that *wssA* is still partly active. There are potentially four different loci that could interact with *carB**. Epistasis between *wspF* and *carB** could be true if *wspF* has additional functions that haven't been identified yet. Furthermore *carB** could interact with *wspR* which becomes activated due to the mutation in *wspF* (Fig. 6.4). Another candidate gene is *wssA*, which could still interact with *carB** due to residual function. Alternatively another gene of *wssA-J* might compensate for non-functional *wssA*, in which case the product of second gene of *wssA-J* might interact with *carB** and provides a fitness benefit for 1s3*carB** under static conditions (Spiers *et al.*, 2002).

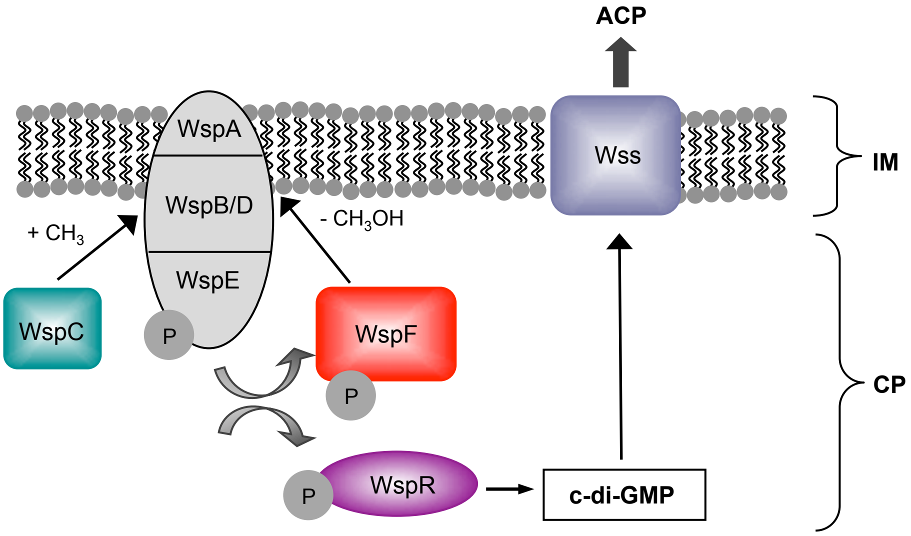

**Figure 6.2: The *wsp* operon**. Detailed description of the operon and the function of the different genes can be found in the overall introduction (see Chapter 1, section 1.4.3.2) and in the preceding paragraph.

It is likely that epistatic interactions between *carB** and one of the loci mentioned above (*wspF*, *wssA*, *wssA-J*, *wspR*), or perhaps a combination of these loci, caused the increase of relative fitness of 1s3*carB** under static conditions and promote the occurrence of switcher genotypes. Identification of the epistatic interaction is crucial in understanding the origin of a switching genotype in *P. fluorescens* and the role of mutational history in the evolutionary process.

In order to identify such epistatic interactions, each potential locus (*wspF*, *wssA*, *wssA-J* and *wspR*) was deleted from the genome of 1s3 and 1s3*carB**. Fitness experiments between non-*carB** and *carB** strains were then performed to test whether the positive fitness effect of *carB** under static conditions holds. Here I was interested in the net fitness effect of *carB** after the deletion of a potentially interacting locus. If the fitness effect of *carB** was still positive I concluded that *carB** still had a chance to arise within the population and no epistatic interaction was causing the fitness increase. A fitness decrease of *carB** in 1s3 after the deletion of *wspF*, *wssA*, *wssA-J* or *wspR* would indicate that the observed previous fitness benefit of *carB** in the 1s3 genotype is the result of the combined effect of *carB** and one or more of the loci. In that case *carB** can only arise within the population as a result of epistasis between *carB** and this particular locus.

### 6.1.3 Objectives

1. To test the hypothesis that epistatic interactions between *carB** and one or more loci/preceding mutations (*wspF, wssA, wssA-J* (the entire *wss* operon, i.e., *wssABCDEFGHIJ*) or *wspR*) are responsible for the positive fitness effect of *carB** in 1s3 under static conditions. In that case the positive fitness effect of *carB** in 1s3 will be abolished after the deletion of one or multiple of these loci. A maintained fitness benefit indicates no interaction between *carB** and *wspF*, *wssA, wssA-J* or *wspR*.

2. Testing includes the estimation of the fitness effect of *carB** in different genotypes by performing competition experiments between non-*carB** strains (1s3Δ*wspF,* 1s3Δ*wssA*, 1s3Δ*wssA-J* and 1s3Δ*wspR*) and *carB** strains (1s3*carB**Δ*wspF*, 1s3*carB**Δ*wssA*, 1s3*carB*Δ*wssA-J* and 1s3*carB**Δ*wspR*; Tab. 6.2).

3. To examine phenotypic change after *wspF*, *wssA-J* and *wspR* are removed from the genome of 1s3 and 1s3*carB** particularly the change in niche preference of genetically engineered strains using photo-documentation of microcosms over 72 hours.

## 6.2 Summary of materials and methods

The description of all the materials and each method that was used in this study are detailed in Chapters 2.

### 6.2.1 Creating a gene deletion of *wspF*, *wssA*, *wss* or *wspR* in 1s3 and 1s3*carB**

An overnight culture of *P. fluorescens* SBW25 was used as a template for the amplification of two fragments flanking both sides of the target gene or operon (*wspF*, *wssA*, *wssA-J* or *wspR*). Then both fragments were fused together using overlap extension PCR. *Spe* I and *Bgl* II restriction sites were added to both sides of the final fragment to make it suitable for cloning (Pogulis *et al.*, 1996). The amplification reactions followed the Phusion PCR protocol. The products were purified and used together in the assembly PCR in which the overlap regions of both fragments were fused together to create one fragment (see Chapter, section *2.2.1.2*). After gel electrophoreses and gel purification, the fragment was cloned into pCR8 using TOPO® TA cloning (see Chapter, section 2.4.2). Separation of the fragment from the vector was achieved by double digest with two restriction enzymes acting simultaneously (*Spe* I & *Bgl* II). The fragment was then ligated into the suicide reporter plasmid pUIC3 and transformed into *E. coli* (see Chapter 2, section 2.4.3). The manufactured strains (*E. coli*-pUIC3Δ*wspF*, *E. coli*-pUIC3Δ*wssA, E. coli*-pUIC3Δ*wssA-J, E. coli*-pUIC3Δ*wspR*) were stored at -80°C.

### 6.2.2 Bacterial conjugation

The *E. coli*-pUIC3 strains carrying a deletion fragment with Δ*wspF*, Δ*wssA,* Δ*wssA-J* or Δ*wspR* served now as donor. A tri-parental conjugation was performed between the *E. coli* donor strains and the *P. fluorescens* recipient strains 1s3 and 1s3*carB**. The helper strain *E. coli* pRK2013 was used to enable the transfer of the plasmid between the *E. coli* and *P. fluorescens*. Purpose of a tri-parental conjugation was to exchange the present wild type fragment in 1s3 or 1s3*carB** against a new fragment carrying Δ*wspF*, Δ*wssA*, Δ*wssA-J* or Δ*wspR*. Tri-parental-conjugation was carried out to create 1s3Δ*wspF,* 1s3Δ*wssA,* 1s3Δ*wssA-J,* 1s3Δ*wspR* (non-*carB** strains) and 1s3*carB**Δ*wspF,* 1s3*carB**Δ*wssA,* 1s3*carB**Δ*wssA-J* and 1s3*carB**Δ*wspR* (*carB** strains) following the conjugation protocol (see Chapter 2, section 2.4.4).

### 6.2.3 Determination of relative fitness of *carB**

Fitness assays were performed to determine the net fitness effect of *carB** in 1s3 after *wspF*, *wssA*, *wssA-J* or *wspR* were deleted from the genome. Therefore competition experiments were carried out between non-*carB** strains and *carB** strains of 1s3 with one of the genes mentioned above deleted. There were 20 replicates of the fitness assay between 1s3*carB**Δ*wspF* and 1s3Δ*wspF*, 30 replicates of 1s3*carB**Δ*wssA* competing against 1s3Δ*wssA*, 20 replicates of 1s3*carB**Δ*wssA-J* against 1s3Δ*wssA-J* and 20 replicates of 1s3*carB**Δ*wspR* competing against 1s3Δ*wspR* (Tab 6.2).

**Table 6.2: Fitness assay between *carB** (Competitor A) and non-*carB** (Competitor B) strains.**

| Replicates | Competitor A vs | Competitor B |
|---|---|---|
| 20 | 1s3*carB**Δ*wspF* | 1s3Δ*wspF* |
| 30 | 1s3*carB**Δ*wssA* | 1s3Δ*wssA* |
| 20 | 1s3*carB**Δ*wssA-J* | 1s3Δ*wssA-J* |
| 20 | 1s3*carB**Δ*wspR* | 1s3Δ*wspR* |

In preparation of the fitness assays both competitors were marked with a *lacZ* transcriptional fusion using a mini-Tn7 transposon system. That caused blue colouration of colonies on agar plates containing Xgal (see Chapter 2, section

2.4.5). To control for fitness effects due to the *lacZ* marker system, replicates of each competition assay were divided into two replicate groups. In one half of the replicates the non-*carB** strains contained the *lacZ* transcriptional fusion and showed blue colonies whereas the colonies of the *carB** strains stayed white. The other half of the replicates was carried out with unmarked non-*carB** strains but the *carB** strain had a *lacZ* transcriptional fusion and formed blue colonies.

The experimental performance of the competition assays followed the same protocol that was applied in previous experiments (see Chapter 2, section 2.5.2). Both competitors (Tab. 6.2) grew together in a static microcosm for 48 hours and cell numbers were monitored at the beginning and at the end of the competition. Subsequently the Malthusian parameters of both competitors were determined and used to calculate the fitness of the *carB** strain relative to the non-*carB** strain (Lenski *et al.*, 1991).

### 6.2.4 Documentation of niche occupation in static microcosms over time

To follow the ecological behaviour over time in a static environment of 1s3 and 1s3*carB** carrying Δ*wspF*, Δ*wssA-J* or Δ*wspR*, the strains were revived by streaking out a small amount of the frozen stock culture on KB agar plates. Afterwards three colonies of each strain were transferred into separate microcosms containing fresh liquid media. The cultures were grown overnight with shaking and 6 μl was used to inoculate new microcosms for an additional overnight incubation. The next day 6 μl of each overnight culture was used to inoculate a new microcosm that was kept undisturbed for 72 hours. After 48 hours and 72 hours photographs were taken of every microcosms (three replicates per strain) with focus on mat formation and broth colonization.

### 6.2.5 Statistical analysis

The statistical evaluation of the fitness assays results was performed with the software R (R.app GUI 1.23). The Shapiro-Wilk test was used to verify whether the fitness values came from a normal distributed population with a set alpha level of 0.05. If data were normally distributed then a one-sample *t*-test was applied to determine whether fitness was significantly greater or smaller than zero with a 95 % confidence interval. In addition, two-sample *t*-test was used to compare means between replicate groups with *lacZ* transcriptional fusion and replicate groups without *lacZ* to control for effects of phenotypic marking on fitness. The software R was used for the graphic presentation of the data.

## 6.3 Results

The contrasting effects of *carB** on fitness in 1s2 and 1s3 as described above (Fig. 6.1) indicates the existence of epistatic interactions between *carB** and four potential loci in 1s3 (*wspF*, *wssA*, *wssA-J* and *wspR*). In order to disentangle the interactions and relative contribution of the loci, these genes were deleted from the genomes of 1s3*carB** (*carB** strains) and 1s3 (non-*carB** strains). Fitness experiments were then performed between *carB** and non-*carB** strains with *wspF*, *wssA*, *wssA-J* or *wspR* deleted (Tab. 6.2) to test whether the previous observed fitness advantage of *carB** in 1s3 would withstand. Replicates where more than 50 % of the colonies of one competitor showed novel colony morphology after competition were excluded from further analyses.

### 6.3.1 Fitness effect of *lacZ* transcriptional fusion

In order to perform competition experiments to determine the fitness of *carB** in different genetic background it is necessary to be able to distinguish competing strains on agar plates. The *lacZ* gene from *E. coli* serves as a useful marker because of the blue colony colouration when growing on agar plates that contain X-gal. In order to control for any effect due to the *lacZ* marker system each competitor of the fitness assay was marked reciprocally. For example in assays to determine the fitness of 1s3*carB**Δ*wspF* compared to 1s3*Δ*wspF*, 20 replicates were used in total. In one half of the replicates 1s3*carB**Δ*wspF* strains had a *lacZ* transcriptional fusion and formed blue colonies and 1s3*Δ*wspF* strains appeared as white colonies. In the other half of the replicates 1s3*carB**Δ*wspF* showed white colonies and 1s3*Δ*wspF* was marked with a *lacZ* fusion and showed blue colonies (Tab.6.3). The results of Shapiro-Wilk tests show that the values for the relative fitness for each replicate group were normal distributed although 1s3*carB**Δ*wssA-J*-***lacZ*** vs 1s3*wssA-J* was marginal for normality (Tab. 6.3).

**Table 6.3: Fitness effect of *lacZ* fusion in two replicate groups for each competition experiment between *carB** and non-*carB** strain.** Displayed are the results of Shapiro-Wilk test (for normal distribution) and two-sample *t*-test (95% confidence interval) of two replicate groups for each competition experiment between 1s3*carB** and 1s3 strains containing Δ*wspF*, Δ*wssA*, Δ*wssA-J* or Δ*wspR*.

| | **Shapiro-Wilk test** | | | **Two sample *t*-test** | | |
|---|---|---|---|---|---|---|
| **Competing strains** | *N* | W | *P*-value | *t* | df | *P*-value |
| 1s3*carB**Δ*wspF*-***lacZ*** vs. 1s3Δ*wspF* | 10 | 0.94 | 0.58 | -2.04 | 18 | 0.056 |
| 1s3*carB**Δ*wspF* vs. 1s3Δ*wspF*-***lacZ*** | 10 | 0.97 | 0.9 | | | |
| 1s3*carB**Δ*wssA*-***lacZ*** vs. 1s3Δ*wssA* | 15 | 0.93 | 0.28 | 2.56 | 25.87 | 0.02 |
| 1s3*carB**Δ*wssA* vs. 1s3Δ*wssA*-***lacZ*** | 14 | 0.93 | 0.32 | | | |
| 1s3*carB**Δ*wssA-J*-***lacZ*** vs. 1s3*wssA-J* | 10 | 0.85 | 0.051 | -0.36 | 13.51 | 0.73 |
| 1s3*carB**Δ*wssA-J* vs .1s3*wssA-J*-***lacZ*** | 10 | 0.92 | 0.34 | | | |
| 1s3*carB**Δ*wspR*-***lacZ*** vs. *1s3ΔwspR* | 10 | 0.94 | 0.59 | -1.16 | 12.42 | 0.27 |
| 1s3*carB**Δ*wspR* vs. *1s3ΔwspR*-***lacZ*** | 9 | 0.92 | 0.39 | | | |

In principle, no fitness effect of the *lacZ* fusion was observed in the strains apart from the replicate group of 1s3*carB**Δ*wssA*-***lacZ*** vs. 1s3Δ*wssA* and of 1s3*carB**Δ*wssA* vs. 1s3Δ*wssA*-***lacZ***. The two replicate groups showed a significant difference in their mean fitness values from each other (Tab. 6.3, two-sample *t*-

test: $P < 0.05$; see Appendices, Fig. A2). Nonetheless mean fitness of both replicate groups showed the same trend and was significantly greater than zero (1s3*carB**Δ*wssA*-***lacZ*** vs 1s3Δ*wssA* one-sample *t*-test: $t = 11.16$, df = 14, $P < 0.0001$; 1s3*carB**Δ*wssA* vs 1s3Δ*wssA*-***lacZ*** one-sample *t*-test: $t = 6.05$, df = 13, $P < 0.0001$). Both replicate groups were combined to one group for later analyses.

## 6.3.2 Detection of epistatic interactions between *carB** and the genes *wspF*, *wssA*, *wssA-J* or *wspR* in 1s3

In the case of epistatic interactions meaning that the fitness advantage of *carB** in 1s3 is based on the interaction between *carB** and *wspF, wssA, wssA-J* or *wspR,* the deletion of one or more of the genes is expected to alter the fitness effect of *carB** in 1s3. In cases without epistasis the fitness of *carB** is not expected to change after the deletion of one of the loci and *carB** continues to improve the performance of 1s3 under static conditions.

The first gene that could potentially interact with *carB** is *wspF*. Here a yet unknown function of *wspF* in combination with *carB** can provide a benefit to 1s3*carB** in a static environment. The results of the competition experiment between 1s3*carB**Δ*wspF* and 1s3Δ*wspF* showed that the relative fitness of *carB** was still positive (Fig. 6.2, second box plot; Tab. 6.4). This indicates that there is no epistasis between *wspF* and *carB** and it is likely that one of the subsequent mutations is responsible for the epistatic interaction with *carB**. The deletion of the second gene, *wssA*, that could cause the fitness increase of *carB** in 1s3 based on residual function after a mutation occurred in the gene, had no effect on the net fitness of *carB** when 1s3*carB**Δ*wssA* competed against 1s3Δ*wssA* (Fig 6.2, third box plot; Tab.6.4).

It is possible that the loss of *wssA* was compensated by another gene within *wssA-J*, which could explain the maintained positive fitness effect of *carB** after *wssA* was deleted in 1s3. To account for this situation *wssA-J* was removed from the genome and the relative fitness of *carB** was remeasured between 1s3*carB**Δ*wssA-J* and 1s3Δ*wssA-J*. The relative fitness of *carB** in 1s3 was still positive (Fig. 6.2, fourth box plot; Tab. 6.4). The deletion of *wspF*, *wssA* and *wssA-J* didn't suppress the beneficial effect of *carB** in 1s3. On that account *wspR*, which is constitutive activated by the mutation in *wspF* and could interact with *carB**, was deleted. The results of the fitness experiments between 1s3*carB**Δ*wspR* and 1s3Δ*wspR* demonstrate a significant decrease in fitness when *carB** was present in the genotypes (Fig. 6.2, orange box plot; Tab. 6.4;).

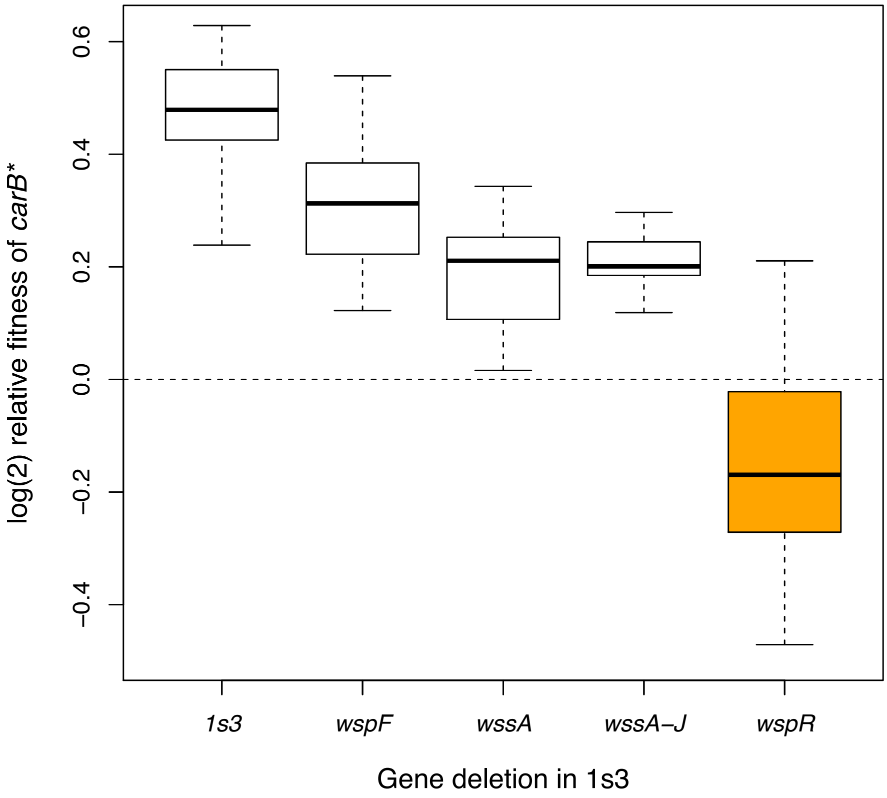

**Figure 6.3: Relative fitness of *carB** in 1s3 after the deletion of *wspF*, *wssA*, *wssA-J* or *wspR*.** The graph illustrates the relative fitness of *carB** in 1s3 resulting from fitness experiments of 1s3*carB** vs. 1s3 (N = 8), 1s3*carB**Δ*wspF* vs. 1s3Δ*wspF* (N = 20), 1s3*carB**Δ*wssA* vs. 1s3Δ*wssA* (N = 29), 1s3*carB**Δ*wssA-J* vs. 1s3Δ*wssA-J* (N = 19) and 1s3*carB**Δ*wspR* vs. 1s3Δ*wspR* (N = 17). Displayed are medians and interquartile ranges.

**Table 6.4: Summarised results from the statistical analysis of fitness effect of *carB**.** Shapiro-Wilk test (for normal distribution) and one-sample *t*-test (95% confidence interval) of the competition experiments between 1s3*carB** and 1s3 strains carrying Δ*wspF*, Δ*wssA*, Δ*wssA-J* or Δ*wspR.*

| | **Shapiro-Wilk test** | | **One-sample *t*-test** | | |
|---|---|---|---|---|---|
| **Competing strains** | W | *P*-value | *t* | df | *P*-value |
| 1s3*carB**Δ*wspF* vs. 1s3Δ*wspF* | 0.98 | 0.9 | 12.8 | 19 | < 0.0001 |
| 1s3*carB**Δ*wssA* vs. 1s3Δ*wssA* | 0.95 | 0.18 | 10.9 | 28 | < 0.0001 |
| 1s3*carB**Δ*wssA-J* vs 1s3Δ*wssA-J* | 0.97 | 0.85 | 20.3 | 18 | < 0.0001 |
| 1s3*carB**Δ*wspR* vs 1s3Δ*wspR* | 0.95 | 0.45 | -2.9 | 16 | 0.005 |

### 6.3.3 Comparison of niche preferences of 1s3 and 1s3*carB** after the deletion of *wspF*, *wssA-J* or *wspR*

Photographs of static microcosms were taken after 48 hours and 72 hours to visualise the performance of 1s3 and 1s3*carB** after *wspF*, *wss*A-J or *wspR* were deleted from both backgrounds. Aim of the study was to identify a change in niche preference in response to the different gene deletions. The pictures indicate that both strains with unimpaired genes have different niche preferences. There was clear biofilm formation after 48 hours in microcosms containing 1s3*carB** whereas 1s3 grew predominantly in the broth. Only after 72 hours, a thin biofilm produced by 1s3 was observed (Fig. 6.3, first row). After *wspF* was deleted the strains didn't show a noticeable change of niche occupation (Fig. 6.3, second row).

The deletion of *wssA-J* operon prevented biofilm formation in either of the two strains. Here after 48 hours the broth was slightly cloudy in microcosms with 1s3 and 1s3*carB**, and after 72 hours there was a very thin white layer at the air-liquid interface in the microcosm containing 1s3*carB** (Fig. 6.3, third row). It is noticeable that the broth of 1s3Δ*wssA-J* and 1s3*carB**Δ*wssA-J* stayed relatively clear. This indicates reduced growth. Interestingly, there was no effect of the *wspR* deletion after 48 hours on niche preference in 1s3 and 1s3*carB**, whereas after 72 hours a distinct biofilm was observed in the non-*carB** strain but no biofilm and a very clear broth in microcosms containing 1s3*carB** (Fig. 6.3, last row). There was reduced growth of 1s3Δ*wspR* and 1s3*carB**Δ*wspR* after 48 hours, which was observed for 1s3*carB**Δ*wspR* after 72 hours too but not for 1s3Δ*wspR.* Here the broth appeared cloudy and a biofilm was seen. The results show that only the deletion of *wspR* resulted in a different ecological behaviour of 1s3*carB** compared to 1s3, which goes hand in hand with the results from the fitness assays (Fig. 6.2).

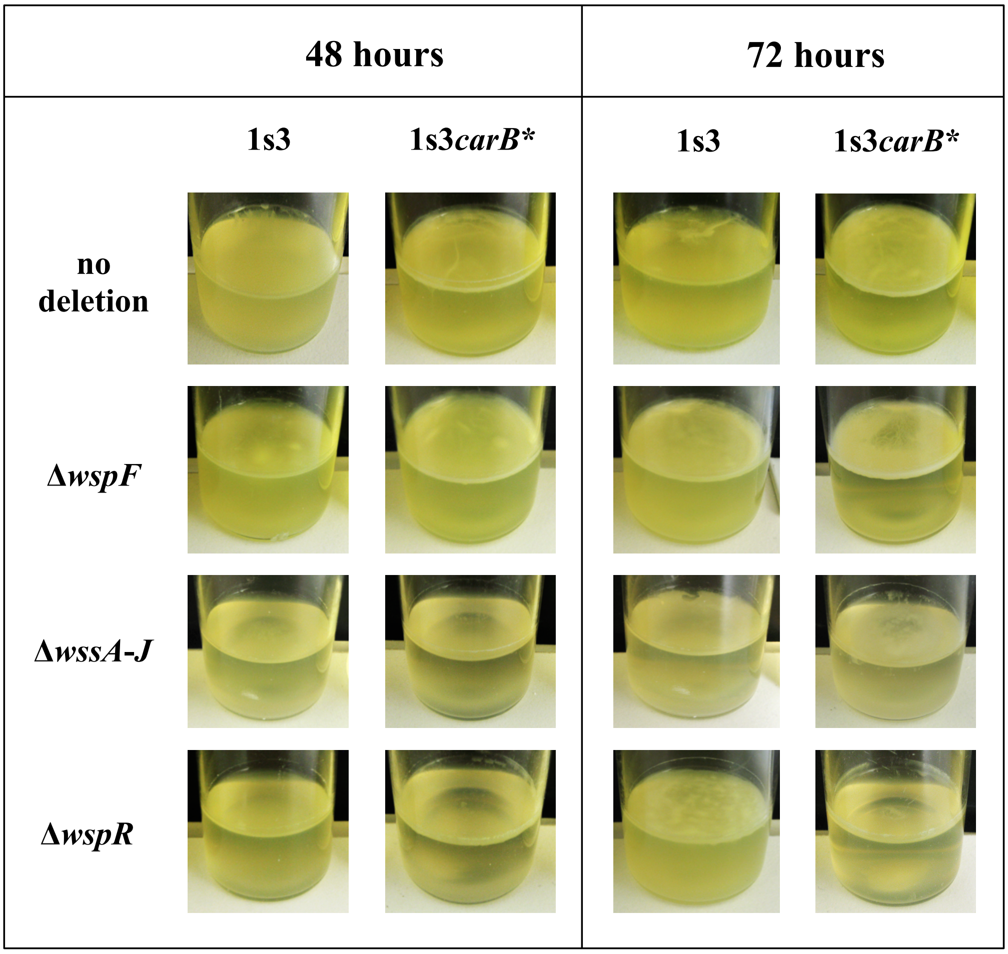

**Figure 6.4: Niche preference of 1s3 and 1s3*carB** in a static environment with *wspF*, *wssA-J* or *wspR* deleted.** Illustrated is one out of three microcosms that shows the ability of 1s3 and 1s3*carB** to occupy the air-liquid interface in a static microcosm when no gene is deleted (first picture row), and when *wspF* (second row), *wssA-J* (third row) or *wspR* (last row) were removed from 1s3 and 1s3*carB**.

## 6.4 Discussion

Epistasis is an often observed phenomenon during evolution (Bonhoeffer *et al.*, 2004; Cooper *et al.*, 2008; da Silva *et al.*, 2010; Khan *et al.*, 2011). Although mutations occur randomly, such epistatic interactions are most deterministic. Interacting loci together can either enhance or decrease the fitness compared to the additive fitness effect of each mutation on its own without the interaction (Maisnier-Patin *et al.*, 2005; Eckardt, 2008). If the combined effect of two or more mutations is higher than the fitness of the ancestor, the novel function can arise and can provide ecological opportunities. New functions that are created by epistatic interactions are highly dependent on the genetic background in which they occur (Weinreich *et al.*, 2006; Ortlund *et al.*, 2007). In addition, it has been shown that they can shift the probabilities of future evolutionary pathways (Weinreich *et al.*, 2005; Blount *et al.*, 2008). It is still unclear to what degree the evolution of new functions depends on evolutionary events in the past and how epistasis acts on evolutionary dynamics.

Evolutionary constraints by epistatic interactions have been observed previously by Blount *et al.* (2008). A variant evolved during a long-term selection experiment in *E. coli* that can utilise citrate under aerobic conditions. The performance of ‘replay’-experiments from different time points revealed that the evolution of Cit$^+$ variants could only occur after one or more potentiating mutations arose within the population. Citrate utilisation was highly dependent on the genetic background in which it occurred. In a recent study (Blount *et al.*, 2012) they characterised the underlying genetic mechanism for the Cit$^+$ phenotype. They found that tandem copies of a gene cluster created multiple *rnk-citT* modules, which put *citT* under the control of the *rnk*-promoter and increased the expression level of *citT*, which was thought to induce citrate utilization under aerobic conditions. The introduction of tandem copies into genotypes from different generations by plasmid transformation resulted in different growth trajectories supporting the hypothesis that a particular genetic background was needed to increases the likelihood of Cit$^+$ evolution after 20,000 generation. A mutation in *arcB* has been suggested to be responsible for the increase in efficiency of citrate usage by interaction with *rnk-citT* modules but the precise mechanism underlying the epistatic interaction is not clear yet (Blount *et al.*, 2012).

In this study four possible genetic interactions (*wspF*, *wssA*, *wss*, *wspR*) with *carB** were analysed in order to determine their impact on the fitness of *carB** in 1s3 and therefore the likelihood of switcher occurrence. In previous experiments it has been shown that the fitness effect of *carB** in two different genotypes (1s2 and 1s3) varied significantly (Fig. 6.1). The genotypes differed only in two mutations (Tab. 6.1). Of particular interest were potential epistatic interactions between *carB** and *wspF*, *wssA*, *wssA-J* and *wspR*, which could explain the fitness difference of *carB** in 1s2 and in 1s3. In addition epistasis may contribute to an increased likelihood of switcher occurrence in 1s3 (see Chapter 5, Fig. 5.3). It was hypothesised that one ore multiple of these loci (*wspF*, *wssA*, *wssA-J* and *wspR*) contribute to the positive fitness effect of *carB** in 1s3 in a static environment. The results of fitness experiments between engineered 1s3*carB** and 1s3 strains with one of the four genes deleted suggest a high degree of epistatic interaction between *carB** and *wspR* which becomes activated by a mutation in *wspF.* Only the deletion of *wspR* from the genome of 1s3*carB** resulted in a negative fitness when 1s3*carB**Δ*wspR* competed against 1s3Δ*wspR* in a static microcosm. Overall there was a general decrease in fitness after removing *wpsF*, *wssA*, *wssA-J* or *wspR* from the genome of 1s3 and 1s3*carB**. This indicates the general importance of these loci for the health of the cells under static conditions. Especially the deletion of *wssA-J* and *wspR* resulted in a decreased growth under static conditions, which can be seen in a clear broth (Fig. 6.3) suggesting that the genes are involved in important processes within the cell. Nonetheless *carB** maintained its positive effect on 1s3 fitness after the deletion of *wspF*, *wssA* and *wssA-J*. The following paragraph will discuss the epistatic interaction in more detail and provides some insights into the underlying epistatic mechanism.

### 6.4.1 The role of epistasis in the evolution of phenotypic switching in *P. fluorescens*

Studying the complex properties and the underlying mechanisms of epistatic interactions is challenging. Epistasis involves key events that occurred in the distant past usually without effect on the phenotype. They are difficult to detect. Only if the interaction partners are known the nature of the particular epistatic interaction on the genetic level and the phenotypic effects can be estimated (Khan *et al.*, 2011). In this study, similar to the studies of Blount *et al.* (2008, 2012), a potentiating mutation occurred earlier during the course of evolution (mutation in *wspF*) and increased the likelihood of a future event a *carB** switcher. In this study the precise epistatic interaction that favours the evolution of a switching phenotype was identified.

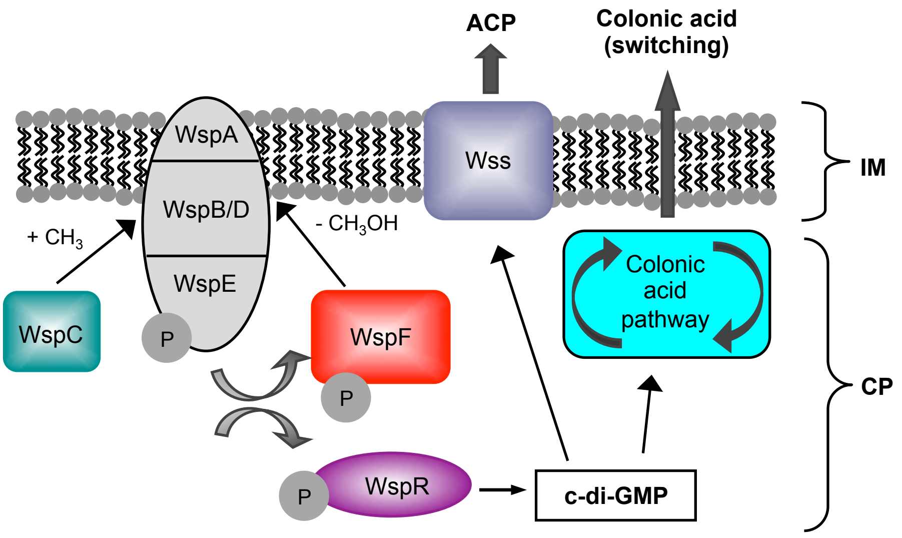

**Figure 6.5: The *wsp* operon associated with the colanic acid pathway.** A mutation in *wspF* activates *wspR* (DGC), which causes the overproduction of c-di-GMP within the cytoplasm (CP). The second messenger initiates ACP production by activating enzymes of *wssA-J*, which is localised at the inner membrane (IM). It is likely that c-di-GMP is partly used for colanic acid production, which can be induced by a mutation in the colanic acid pathway (*carB**) and causes switching behaviour.

I found evidence that a mutation in *wspF* is a crucial event in the past for the evolution of switching behaviour in 1s3. WspF is a negative regulator of WspR and a mutation in the *wspF* gene led to the constitutive activation of WspR. The decreased fitness of *carB** after *wspR* was removed from the 1s3 genome (Fig. 6.2) suggests that without an operating WspR *carB** cannot rise in frequency within the population. The mutation in *wspF* and the deletion of the entire gene led to the loss of negative regulation of *wspR*. As a consequence WspR is active and synthesises c-di-GMP, a second messenger molecule that is involved in the regulation of cellulose production (Fig. 6.5; Goymer *et al.*, 2006; Bantinaki *et al.*, 2007; Malone *et al.*, 2007).

The activation of WspR increases the likelihood of a switching phenotype by *carB** and explains the fitness difference between 1s2*carB** and 1s3*carB**, which was observed in a previous experiment (see Chapter 4, Fig. 4.2). It has been shown that *carB** is involved in the production of colanic acid and the process of capsulation in *P. fluorescens* (Beaumont *et al.*, 2009; Gallie, 2009) but a specific role of the DGC in the colanic acid metabolic pathway in *P. fluorescens* is not known yet. Studies in *Vibrio parahaemolyticus* suggest that the capsulation process is likely to be linked to a protein with presumable DGC function as part of a sensory signaling pathway (Boles & McCarter, 2002; Guvener & McCarter, 2003). The results of this study show that the DGC WspR could be involved in the colanic acid production pathway. Although the precise role of WspR within the pathway is not clear yet, it is possible that the overproduced c-di-GMP is partly used for colanic acid production (Fig. 6.5) and the lack of supply of c-di-GMP due to the *wspR* deletion has detrimental consequences in the presence of *carB** in 1s3 because colanic acid synthesis is not assured anymore.

### 6.4.2 Ecological consequences of epistatic interaction

The study shed light on the principal mechanism underlying the epistatic interaction in 1s3*carB** that causes the fitness increase compared to 1s2*carB**. Furthermore, photo documentation of media-filled microcosms that were inoculated with *carB** or non-*carB** strains of 1s3 with *wspF*, *wssA-J* or *wspR* deleted, were used to obtain information on whether the interaction on the genetic level translates into a change of niche preferences. Pictures of microcosms were taken after 48 hours and 72 hours (Fig. 6.3). The difference in niche occupation between 1s3 and 1s3*carB** (Fig. 6.3, first picture row) is clearly visible. Whereas 1s3 (48/72 hours) primarily occupies the broth, the introduction of *carB** led to the formation of a clear biofilm on the surface of the microcosm. This change in niche preference between 1s3 and 1s3*carB** suggests that *carB**, and therefore colanic acid production can help the bacteria to occupying the air-liquid interface. The access to oxygen and nutrients provides an advantage under static conditions (Rainey & Travisano, 1998). The deletion of *wspF* had no effect on the ecological behaviour in either 1s3 or 1s3*carB**. Without *carB** the cells of 1s3 still grew in the broth and biofilm formation was observed when *carB** was present in the genome of 1s3. This is in line with the results from the fitness test, and suggests that apart from the activation of WspR there is no additional function of *wspF* that could explain the fitness advantage of *carB**. After the deletion of *wssA-J* an overall trend was seen in non-*carB** and *carB* strains. The broth was relatively clear - even after 72 hours. Here, a very thin biofilm developed in microcosms with 1s3*carB**, most probably the result of progressing evolution. Even though, a quantitative measurement was not obtained, the clear visual difference between strains with and without *wssA-J* indicates that the latter showed decreased growth in the broth under static conditions. It is likely that *wssA-J* has additional functions (apart from cellulose production) that have not been identified yet. The loss of these functions could explain the decrease in growth. The unchanged fitness of *carB** when *wssA-J* was deleted suggests that even without the operon, which is detrimental for the cell, *carB** has still a chance to occur within the population. I found interaction between *carB** and an activated WspR that caused a fitness increase. The microcosms containing the 1s3Δ*wspR* and 1s3*carB**Δ*wspR* strains showed, similar to the strains with the *wssA-J* deletion, decreased bacterial growth. The broth in the microcosms appeared less cloudy for the first 48 hours. After 72 hours, the microcosms with 1s3Δ*wspR* were cloudier and had a biofilm. This could be caused by increased growth or progressing evolution that gave rise to a new type able to occupy the air-liquid interface. With *carB** in the genome the same strain did not grow much even after 72 hours and no biofilm was observed. This suggests that the successful biofilm production of 1s3*carB** depends on the presence of an operating WspR.

## 6.5 Summary

I showed that epistatic interactions are responsible for the increased performance of a switching phenotype based on *carB** in 1s3 under static conditions. The *carB** mutation no longer had a positive fitness effect on 1s3 after *wspR* was deleted from the genome. This suggests that an activated *wspR* is necessary for having the capacity to evolve an evolutionarily successful *carB** switcher. An interaction between *wspR*, which became activated by a mutation in *wspF*, and *carB** is the reason for an increased fitness of 1s3*carB** compared to 1s2*carB** and a higher chance of switcher occurrence in 1s3. The ability to form a biofilm is crucial for a switcher to increase in frequency in this background under static conditions. Therefore the first hypothesis (see 6.1.3) is supported. Although the exact molecular mechanism of the epistatic interaction is not fully clear yet, the results suggest that WspR, which is a DGC that syntheses c-di-GMP, is needed for biofilm formation by the switcher based on *carB**. This study shows the importance of the deterministic interaction between two different genetic loci for the evolution of a switching phenotype. The evolution of switching in 1s3 was dependent on a single mutation that occurred earlier (in *wspF*) and therefore on the particular genetic background in which it occurred. The actual cause for the epistatic interaction was the activation of *wspR.* Any mutation that abolishes negative regulation of *wspR* therefore has the potential to lead to a similar epistatic effect.

# Chapter 7: Final discussion

## 7.1 Theoretical background of the project

A central focus of evolutionary biology has been to identify the factors that drive populations along certain evolutionary pathways, and so be able to determine the outcome of evolution. It is still not (fully) understood to what degree chance, natural selection and evolutionary history affect the evolution of adaptive innovations in the face of unpredictable environments. An intense debate has developed over the last decades between two groups of scientists. One group argues that the evolution of innovations is the result of selection only, and that organisms should therefore evolve towards an adaptive optimum under given environmental circumstances. In that case, evolution along different evolutionary pathways should lead to the same result and the impact of random forces and evolutionary history is negligible. Evolution under these circumstances is thought to be predictable (Dawkins, 1986; Dennett, 1995; Schluter, 1996; Conway Morris, 2003; Vermeij, 2006; Conway Morris, 2009; Conway Morris, 2010). The second group claims that the evolution of novel structures is contingent upon their history. The evolutionary path that an organism can take depends on initial circumstances, which are set by random factors such as spontaneous mutations, and environmental conditions. Which mutations arise when, and how the environment changes are unpredictable. Consequently, populations can take different evolutionary routes that do not necessarily lead to the best possible adaptive solution (Lewontin, 1966; Monod & Wainhouse, 1971; Jacob, 1977; Gould & Lewontin, 1979; Jablonski, 1986; Gould, 1989; Mani & Clarke, 1990; Travisano *et al.*, 1995a; Beatty, 2008).

Evolution experiments with microorganisms have been of great value in providing insights into processes that are involved in adaptive evolution and in helping to reconcile the ongoing debate (Lenski *et al.*, 1991; Lenski & Travisano, 1994; Travisano *et al.*, 1995a; Travisano *et al.*, 1995b; Rainey & Travisano, 1998;

Wichman *et al.*, 1999; Beaumont *et al.*, 2009; Saxer *et al.*, 2010; Meyer *et al.*, 2012). The long-term evolution experiment with populations of *Pseudomonas fluorescens* SBW25 (REE; see Chapter 1, section 1.4.4) was initiated in the Rainey Laboratory in 2006 to investigate the extent to which evolution is repeatable, and whether the evolution of novel traits is contingent upon evolutionary history (Beaumont *et al.*, 2009). During the course of the experiment, 12 initially identical *P. fluorescens* lineages evolved alternately in two different environments, static and shaken, which caused the occurrence of new variants. Every time new variants were observed, a single colony of the most dominant type was transferred into the opposite environment from the one in which it had evolved. This procedure was repeated until no new type was detected anymore. Whole genome sequencing of one colony of the evolutionary endpoints was performed to identify the causative mutations that led to the phenotypic change dependent on the two alternate environments. In this way it was possible to follow both phenotypic and genotypic evolution throughout the evolutionary history of the 12 replicate *P. fluorescens* lineages. Under static conditions, 'wrinkly spreader' types (WS) generally arose and under shaking conditions the WS type usually returned to the ancestral-like smooth colony morphology (SM; Rainey & Travisano, 1998; Beaumont *et al.*, 2009; Koza *et al.*, 2011)

This generally observed pattern was 'interrupted' by the occurrence of a novel phenotype, a stochastically switching genotype, in two out of the 12 lineages (Line 1 and Line 6). Both switcher lineages were found to undergo parallel phenotypic, but divergent genotypic, evolution. Although they shared some parts of the mutational history, none of the switcher mutations were identical (Beaumont *et al.*, 2009; Gallie, 2009). The causative mutations occurred in *carB* (Line 1) and *rpoD* (Line 6) after the lineages had gained eight mutations during preceding selection rounds. Both genes are involved in the colanic acid pathway. Mutations in *carB* and *rpoD* caused the overproduction of colanic acid in some cells of the population, leading to the typical switcher phenotype on agar plates (opaque and translucent colony morphologies, and colonies with opaque sectors). The introduction of the mutations into the immediate ancestor and into the ancestral SBW25 genome was sufficient to induce the switcher phenotype. The fitness effect of *carB* and *rpoD* was positive in the immediate ancestor, indicating that the novel trait was an adaptive heritable trait (Beaumont *et al.*, 2009; Rainey *et al.*, 2011; Libby & Rainey, 2011). The mutations introduced into the ancestor SBW25, however, did not provoke this fitness benefit and since both switcher lineages shared some parts of their history it was inferred that the evolutionary history was of importance for the evolution of a switcher genotype. The mechanism underpinning the effect of history on the evolution of switching was not investigated, but it was proposed that epistatic interactions between the switcher mutations and historical events could promote the occurrence of this novel type (Beaumont *et al.*, 2009).

Owing to extensive earlier work, there is a large amount of information available about the underlying genetic and molecular mechanisms causing phenotypic change in the *P. fluorescens* lineages (Rainey & Bailey, 1996; Rainey & Travisano, 1998; Kassen *et al.*, 2000; Spiers *et al.*, 2002; Rainey & Rainey, 2003; Spiers *et al.*, 2003; Gehrig, 2005; Brockhurst *et al.*, 2006; Goymer *et al.*, 2006; Kassen & Bataillon, 2006; Bantinaki *et al.*, 2007; Jasmin & Kassen, 2007; Beaumont *et al.*, 2009; Gallie, 2009; McDonald *et al.*, 2009; Silby *et al.*, 2009; McDonald *et al.*, 2011). The broad knowledge of the *P. fluorescens* model system and the occurrence of the switcher, combined with a 'frozen fossil record', provided the opportunity to study the role of chance and history in the evolution of a switching phenotype.

## 7.2 Summary of the findings

Parallel Switcher Re-Evolution Experiments (SREE) were performed in four independently evolved lineages of the REE, which included Line 3 and Line 12 in addition to the original switcher lineages Line 1 and Line 6 (Chapter 3). All lineages carried eight mutations although they occurred partly in different genes and in a different order. The SREE was performed to test whether the different lineages with different evolutionary histories have the same likelihood of switcher occurrence. It has been suggested that Line 1 and Line 6 are highly likely to produce a switcher than other lineages because of a shared history (Beaumont *et al.*, 2009). The results show that the occurrence of a switcher genotype is not restricted to Line 1 and Line 6 as switching also arose in Line 3. However, no switcher was found in Line 12 (see Chapter 3; Fig. 3.3). This high degree of parallel phenotypic evolution was not observed at the level of the genotype (see Chapter 3, section 3.3.2; Manceau *et al.*, 2010). Hence, multiple evolutionary pathways can lead to a switching phenotype even when the lineages are closely related. Similar observations have been made in other studies (Lenski & Travisano, 1994; Wichman *et al.*, 1999; Hoekstra *et al.*, 2006; McDonald *et al.*, 2009). Although three out of four lineages showed the capacity to produce a switcher, there was variation in the likelihood of switcher occurrence, indicating the impact of additional factors. For example only a few switchers were found in Line 3 and Line 6 but Line 1 gave rise to a significantly greater number of switchers (see Chapter 3, Fig. 3.3). This variation in switcher occurrence was caused by the underlying genotype, which is the composition of different random mutations, and/or the different order of mutations.

In addition, the lines differed in the time needed to generate a new type. Line 3, Line 6 and Line 12 produced new types rather quickly but Line 1 needed much longer (see Chapter 3, Fig. 3.5). This supports the idea that genetic composition plays an important role in the evolution of novel types. The different genotypes are likely to differ in their supply of beneficial mutations, depending on the underlying genotype. Even though Line 12 had in principal the capacity to produce a switcher (see Chapter 3, Fig. 3.7), other new types occurred rapidly under static conditions, and so decreased the chance of a switcher establishing within the population. Altogether the results show that switcher evolution is not a unique evolutionary event, and can repeatedly occur in different lineages. Parallel evolution has been seen as evidence for a negligible impact of chance and history and a strong influence of selection on evolution (Vermeij, 2006; Conway Morris, 2009; Conway Morris, 2010). The results of this study show that although parallel phenotypic switcher evolution can be observed, the likelihood of occurrence depends on the underlying mutational history. If a genotype produces many other new variants that are slightly better in coping with the static environment than the ancestor, these types will increase in frequency rather quickly. Under these circumstances it is difficult for a switcher to increase in cell numbers, and to be detected on an agar plate. There are multiple pathways leading to multiple solutions in a static environment, and whether the switcher path is taken or not depends on the evolutionary history and on the time point of switcher mutation occurrence and is therefore rather unpredictable.

The study in Chapter 4 focused on the impact of evolutionary history on the capacity for switcher evolution within a single lineage, Line 1. The original switcher evolved in Line 1 due to a mutation in the *carB* gene (*carB**) after eight preceding mutations (see Chapter 1, section 1.5). The aim of this study was to investigate whether the capacity to give rise to a switcher genotype changes during the course of evolution. Therefore *carB** was introduced into the genomes of types that evolved at different time points in Line 1 before the original switcher was detected during the REE: the ancestor SBW25, 1s1, 1s2, 1s3 and 1s4. Afterwards the effect of *carB** on the phenotype and on the fitness were examined in the different genotypes. The results demonstrate that *carB** produces the switching phenotype in every genotype tested (see Chapter 4, Fig. 4.1). Unlike in other studies, the evolution of a *carB** switcher is not a stepwise process, or due to the combined effect of all mutations that accumulated over time (Weinreich *et al.*, 2006; Ortlund *et al.*, 2007; Meyer *et al.*, 2012).

The capacity to evolve a *carB** switcher varied between the different genotypes along the evolutionary pathway in a rather unpredictable way. Surprisingly, the ancestor SBW25 already showed the ability to evolve a switcher genotype based on *carB**, indicating that none of the other mutations occurring in the evolutionary history of Line 1 was actually relevant. This result was in contradiction to the findings of Beaumont and colleagues (2009), because of a different experimental design. They performed the fitness assay over a period of 72 hours instead of 48 hours, which led to the evolution of novel types, which had a deleterious effect on the switcher genotype (Beaumont *et al.*, 2009; Gallie, 2009). The genotypes 1s1 and 1s2 did not have the capacity to produce a *carB** switcher due to negative fitness effect of *carB** in these genotypes, indicating some genetic constraints (Weinreich *et al.*, 2006; Khan *et al.*, 2011). Genotypes occurring later in the evolutionary history, 1s3 and 1s4, are again capable of giving rise to a switcher based on *carB** (se Chapter 4, Fig. 4.2). The large difference of the fitness effect of *carB** between 1s2 and 1s3 indicated the existence of epistatic interactions between *carB** and other mutations that occurred during the REE between 1s2 and 1s3. Overall the results show that in principle the evolution of a *carB**-based switcher is not contingent upon the mutational history in Line 1. None of the eight mutations that occurred previous to the original switcher are required to produce an evolutionarily succsessfull switcher. The ancestor SBW25 has already the capacity to evolve a switcher according to the fitness increase after *carB** was introduced into the genomes.

The experiment in Chapter 5 was performed to test whether the findings from Chapter 4 can be reproduced during real-time evolution (Lunzer *et al.*, 2005; Weinreich *et al.*, 2006; de Visser *et al.*, 2009; Franke *et al*, 2011; Khan *et al.*, 2011; Kvitek & Sherlock, 2011). A re-evolution experiment (SREE) was carried out using the same genotypes of Line 1 (SBW25, 1s1, 1s2, 1s3 and 1s4) as founder populations to evolve in a static environment (see Chapter 5, Fig. 5.2). In addition to the number of switchers, the number of all new types, the most common type, and the diversity of different types that was observed on agar plates were monitored. The aim was to get an estimate of how population composition changed during diversification depending on the different founder population.

One of the main findings are that genotypes from early in the evolutionary history (SBW25, 1s1 and 1s2) are less likely to produce a switcher than genotypes that occurred later on (1s3 and 1s4) based on the numbers of detected switcher on agar plates (see Chapter 5, Fig. 5.3). In addition earlier types evolved much quicker and produced more new types, which was reflected in an increased number of new types and diversity on agar plates (see Chapter 5, Fig. 5.5 & Fig. 5.7). The likelihood of beneficial mutations decreased along the evolutionary pathway due to the depletion of common evolutionary pathways. This implies that for improving performance, and ensuring survival under static conditions, alternative mutational pathways will be taken, which is likely to lead to the occurrence of switchers. In addition, the genotypes 1s3 and 1s4 are already quite well adapted to a static environment, probably due to the production of cellulose (Gallie, 2009).

When comparing the results from this Chapter 5 with the findings from Chapter 4 it becomes apparent that SBW25 has a high capacity to evolve a switcher, but only a few switchers were detected during the SREE (see Chapter 5, Fig. 5.8). These controvertial results emphasise that having the capacity to produce an evolutionarily successful switcher does not always lead to the evolution of a switcher. There are other factors in addition to fitness that determine whether a switcher can become established within a population or not. For example, it is important whether alternative evolutionary pathways are available and can lead to other successful new types. In SBW25 the known common mutational routes to evolve a WS (McDonald *et al.*, 2009) are still open. Under these circumstances a switcher genotype may have occurred but cannot increase in frequency because of the presence of many multiple competitors that successfully occupy the niche. Here ecological interactions between different types within a microcosm influence the outcome of evolution (Elena & Lenski, 1997b; Rainey & Travisano, 1998; Friesen *et al.*, 2004; MacLean *et al.*, 2004; Maharjan *et al.*, 2006; Lang *et al.*, 2011). Similar to findings in Chapter 3, there was only a small degree of parallel genotypic evolution amongst the different re-evolved switcher genotypes, suggesting that multiple mutational pathways can lead to this innovation.

The comparison of the results from Chapter 4 with the results from Chapter 5 revealed some differences between the estimated likelihood of expected evolutionary endings (Weinreich *et al.*, 2006; Ortlund *et al.*, 2007; Reetz & Sanchis, 2008; Novais *et al.*, 2010; Franke *et al.*, 2011; Khan *et al.*, 2011) and the outcomes of real-time evolution. The results suggest that apart from the genetic capacity to evolve a switcher, there are other factors such as the depletion of common mutational pathways, interactions between different competitors, and the fitness of the ancestor that can have an impact on the outcome of evolution. Furthermore, the increased number of switchers in 1s3 and 1s4 support the idea that epistatic interactions between switcher mutations, and mutations that occurred earlier, promote switcher evolution. The interaction of these factors appears to be very complex, and how this interaction changes with proceeding evolution seems to be highly dependent on the genetic composition.

The nature of a potential epistatic interaction between *carB** and mutations that occurred throughout the evolutionary history of Line 1 was studied in Chapter 5. Interestingly, there was a significant fitness effect of *carB** in 1s2 and 1s3 (see Chapter 6, Fig. 6.1), with *carB** causing a fitness decrease in 1s2, and a dramatic fitness increase in 1s3. This, in combination with the fact that both genotypes differed in only two mutations (in *wspF* and in *wssA*) from one another, indicated that epistatic interactions could lead to the increased fitness of *carB** in 1s3. In this study, four potential loci were tested for interactions with *carB**: *wspF* and *wssA* (difference between 1s2 and 1s3) as well as *wssA-J* and *wspR*. The function of *wspF* and *wssA* are not fully understood yet. It was thought that residual functionality after a mutation had occurred might cause the interaction with *carB**. The latter two, *wssA-J* and *wspR*, were chosen because a mutation in *wssA* could in principle be compensated for by another gene in *wssA-J* and then the deletion of *wssA* would have almost no effect on the fitness of *carB** in 1s3. Furthermore it is known that the mutation in *wspF* leads to the removal of the negative regulation of *wspR* (Bantinaki 2007). The activation of *wspR* could lead to epistatis. In order to determine which of the loci interacts with *carB**, each gene or operon was successively deleted from the genome of 1s3. Competition experiments were then performed to test whether the previously observed fitness benefit of *carB** in 1s3 would remain.

The results show that the deletion of *wspF*, *wssA* and *wssA-J* did not significantly alter the fitness effect of *carB** in 1s3. The presense of *carB** in these genomes of 1s3 still resulted in an increased fitness. Only the removal of *wspR* decreased the fitness of *carB** significantly, suggesting that the combined effect of *wspR* and *carB** leads to the observed benefit. The causal mutation however occurs in *wspF*, which is a negative regulator of *wspR*. The mutation in *wspF* almost certainly abolished negative regulation of *wspR*. As a result *wspR*, which is a di-guanylate cyclase (DGC), is constitutively activated and produces c-di-GMP (Bantinaki *et al.*, 2007; Malone *et al.*, 2007). It is likely that c-di-GMP is used in the colanic acid production. The deletion of *wspR* causes a decrease in the c-di-GMP concentration, which has a negative effect on the fitness in the presence of *carB** because colanic acid production is interrupted. Epistasis is a frequently observed phenomenon during evolution (Bonhoeffer *et al.*, 2004; Blount *et al.*, 2008; Cooper & Lenski, 2008; da Silva *et al.*, 2010; Khan *et al.*, 2011; Blount *et al.*, 2012). Epistatic interactions provide a direct example of how evolutionary events in the past, which occur by chance, increase the likelihood of the evolution of adaptive innovations. Here the outcome of evolution is highly dependent on a particular genotype and therefore contingent upon the mutational – and therefore evolutionary – history.

## 7.3 Final comment and future prospects

The predictability of evolutionary outcomes has been the centre of a long and ongoing debate (Gould & Lewontin, 1979; Gould, 2002; Vermeij, 2006; Beatty & Desjardins, 2009; Conway Morris, 2009; Conway Morris, 2010). Empirical evidence showing to what degree and under which circumstances chance, history and natural selection influence evolutionary pathways is rare. 'Rewinding and replaying' the tape of life in *P. fluorescens* can repeatedly lead to the evolution of a switcher genotype in closely related lineages. This case of parallel phenotypic evolution suggests that the evolution of this novel type is likely to occur and is to some extent predictable. Nonetheless, there was some variation in the likelihood of occurrence due to the underlying genetic composition. Here the mutations, which happened by chance, and the order in which they occurred, influenced the outcome of evolution. I was able to show that in Line 1 the theoretical capacity to evolve a switcher changes with proceeding evolution in an erratic manner, and that a switcher is already highly likely to arise from the ancestor SBW25. In principal the evolution of a switcher genotype does not depend on the mutational history. The results of the SREE, however, indicate that this estimated capacity to evolve a switcher is not always realised during real-time evolution. It is difficult to make predictions about switcher occurrence based on the fitness value of *carB** in certain backgrounds because other factors become effective, such as the number and strength of competitors, the availability of alternative evolutionary pathways, and the fitness of the ancestor. These factors depend on the genetic composition, so that history influences switcher evolution. Epistasis has a high impact on switcher evolution in Line 1 at a later time point of the evolutionary history. Here a *carB** switcher has a higher chance of arising due to a mutation that occurred earlier. Epistatic interactions show how randomly occurring mutations in combination with selection determine future evolutionary pathways.

The debate between Simon Conway Morris and Stephen J. Gould became almost a personal dispute about evolution and the history of life. Conway Morris argued for the predictability of evolution. For him the observation of convergent evolution is the result of strong selection, which has the strongest effect on evolution. Gould on the other hand believed in a powerful impact of chance and history. According to his view, evolution is contingent upon history (Conway Morris & Gould, 1998). I have shown that actually 'rewinding and replaying the tape of life' in populations of *P. fluorescens* provides a window into a world that has only previously been explored in theory and can partly disentangle the roles of natural selection, chance and history. The results of this study neither fully support the view of Conway Morris nor the view of Gould, but they show that selection, chance and history are tightly linked and that evolution is the result of their combined effect. The complexity of the interactions identified in this example of 'replaying the tape of live' with a simple organism like *P. fluorescens* suggest that both Gould and Conway Morris underestimated how interwoven these different processes in fact are. Even though parallel (phenotypic) evolution was observed in populations of *P. fluorescens,* ultimately the underlying genotype, carrying the information of the entire evolutionary history of the type, determined how likely the evolution of switching was.

In order to gain more insight into how chance and evolutionary history determine the outcome of evolution further experiments could be performed to test the generality of the findings of this study. For example screening other potential genetic loci in the re-evolved switcher genotypes from the SREE (Chapter 2 & Chapter 4) could provide more information about the number of different evolutionary pathways and the likelihood of them being taken. Furthermore, it has been proposed that the availability of common evolutionary routes that lead to a WS type in *P. fluorescens* may influence the likelihood of switcher evolution. The stepwise deletion of such common evolutionary pathways and the impact on real-time switcher evolution could demonstrate how much the ability to give rise to a switcher genotype depends on the number of available alternative pathways. The investigation of ecological interactions within microcosms is a challenging task but it could lead to a deeper understanding of how the surrounding biotic (and abiotic) environment influences switcher evolution, and how this changes with proceeding evolution and the genetic composition. Epistatic interactions seem to play a central role in switcher evolution. The genetic mechanism underlying the interaction between *wspR* and *carB** in Line 1 is not fully understood. Interestingly, in Line 6 a similar sequence of mutations occurred during the course of the REE. Here a mutation in *wspF* was followed by a mutation in *wssB*. In theory, this would lead to a similar epistatic interaction with a switcher mutation as was observed in Line 1. Nonetheless, not many switching types re-evolved from 6s4. The introduction of *carB** into the genomes of 6s3 and 6s4 could reveal whether an epistatic interaction exists. Overall, the *P. fluorescens* model system, together with the occurrence of a switcher genotype, has much potential for further research that could contribute to a broader understanding of the impact of chance and history in the evolution of a novel trait.

# Reference List

**Allender CJ, Seehausen O, Knight ME, Turner GF & Maclean N (2003)** Divergent selection during speciation of Lake Malawi cichlid fishes inferred from parallel radiations in nuptial coloration. *Proc Natl Acad Sci USA* **100**, 14074-14079.

**Arendt J & Reznick D (2008)** Convergence and parallelism reconsidered: what have we learned about the genetics of adaptation? *Trends Ecol Evol* **23**, 26-32.

**Arjan JA, Visser M, Zeyl CW, Gerrish PJ, Blanchard JL & Lenski RE (1999)** Diminishing returns from mutation supply rate in asexual populations. *Science* **283**, 404-406.

**Bantinaki E (2001)** Characterisation of a novel chemosensory pathway underlying adaptive evolution in experimental populations of *Pseudomonas fluorescens* SBW25. Ph.D. dissertation, University of Oxford, Oxford.

**Bantinaki E, Kassen R, Knight CG, Robinson Z, Spiers AJ & Rainey PB (2007)** Adaptive divergence in experimental populations of *Pseudomonas fluorescens*. III. Mutational origins of wrinkly spreader diversity. *Genetics* **176**, 441-453.

**Barrick JE, Yu DS, Yoon SH, Jeong H, Oh TK, Schneider D, Lenski RE & Kim JF (2009)** Genome evolution and adaptation in a long-term experiment with *Escherichia coli*. *Nature* **461**, 1243-1247.

**Beatty J (2006)** Chance variation: Darwin on Orchids. *Philos Science* **73**, 629-641.

**Beatty J (2008)** Chance variation and evolutionary contingency: Darwin, Simpson (*The Simpsons*) and Gould. In *The Oxford Handbook of Philosophy of Biology*, pp. 189-210. Oxford: Oxford University Press.

**Beatty J & Desjardins EC (2009)** Natural selection and history. *Biol Philos* **24**, 231-246.

**Beaumont HJE, Gallie J, Kost C, Ferguson GC & Rainey PB (2009)** Experimental evolution of bet hedging. *Nature* **462**, 90-97.

**Bell G (2008)** *Selection - The Mechanism of Evolution*, 2nd ed. New York: Oxford University Press.

**Bertani G (1951)** Studies on lysogenesis. I. The mode of phage liberation by lysogenic *Escherichia coli*. *J Bacteriol* **62**, 293-300.

**Blount ZD, Barrick JE, Davidson CJ & Lenski RE (2012)** Genomic analysis of a key innovation in an experimental *Escherichia coli* population. *Nature* **489**, 513-518.

**Blount ZD, Borland CZ & Lenski RE (2008)** Historical contingency and the evolution of a key innovation in an experimental population of *Escherichia coli*. *Proc Natl Acad Sci USA* **105**, 7899-7906.

**Bock WJ (1959)** Preadaptation and multiple evolutionary pathways. *Evolution* **13**, 194-211.

**Boiteau L & Pascal R (2011)** Energy sources, self-organization, and the origin of life. *Orig Life Evol Biosph* **41**, 23-33.

**Boles BR & McCarter LL (2002)** *Vibrio parahaemolyticus scrABC*, a novel operon affecting swarming and capsular polysaccharide regulation. *J Bacteriol* **184**, 5946-5954.

**Bonhoeffer S, Chappey C, Parkin NT, Whitcomb JM & Petropoulos CJ (2004)** Evidence for positive epistasis in HIV-1. *Science* **306**, 1547-1550.

**Boughman JW, Rundle HD & Schluter D (2005)** Parallel evolution of sexual isolation in sticklebacks. *Evolution* **59**, 361-373.

**Bridgham JT, Carroll SM & Thornton JW (2006)** Evolution of hormone-receptor complexity by molecular exploitation. *Science* **312**, 97-101.

**Brockhurst MA, Hochberg ME, Bell T & Buckling A** (2006) Character displacement promotes cooperation in bacterial biofilms. *Curr Biol* **16**, 2030-2034.

**Brockhurst MA, Colegrave N, Rozen DE (2011)** Next-generation sequencing as a tool to study microbial evolution. *Mol Ecol* **20**, 972-980.

**Carlborg O, Jacobsson L, Ahgren P, Siegel P & Andersson L (2006)** Epistasis and the release of genetic variation during long-term selection. *Nat Genet* **38**, 418-420.

**Chao L & Levin BR (1981)** Structured habitats and the evolution of anticompetitor toxins in bacteria. *Proc Natl Acad Sci USA* **78**, 6324-6328.

**Charlesworth B (2003)** The organization and evolution of the human Y chromosome. *Genome Biol* **4**, 226.

**Cheng CH & Shuman S (2000)** DNA strand transfer catalyzed by vaccinia topoisomerase: ligation of DNAs containing a 3' mononucleotide overhang. *Nucleic Acids Res* **28**, 1893-1898.

**Choi KH, Gaynor JB, White KG, Lopez C, Bosio CM, Karkhoff-Schweizer RR & Schweizer HP (2005)** A Tn7-based broad-range bacterial cloning and expression system. *Nat Methods* **2**, 443-448.

**Colegrave N & Buckling A (2005)** Microbial experiments on adaptive landscapes. *Bioessays* **27**, 1167-1173.

**Conrad TM, Joyce AR, Applebee KM, Barrett CL, Xie B, Gao Y, Palsson B (2009)** Whole-genome resequencing of Escherichia coli K-12 MG1655 undergoing short-term laboratory evolution in lactate minimal media reveals flexible selection of adaptive mutations. *Genome Biology* **10**, R118.

**Conway Morris S (2003)** *Life's Solution: Inevitable Humans in a Lonely Universe*. Cambridge: Cambridge University Press.

**Conway Morris S (2009)** The predictability of evolution: glimpses into a post-Darwinian world. *Naturwissenschaften* **96**, 1313-1337.

**Conway Morris S (2010)** Evolution: like any other science is predictable. *Phil Trans R Soc B* **365**, 133-145.

**Cooper TF & Lenski RE (2010)** Experimental evolution with *E. coli* in diverse resource environments. I. Fluctuating environments promote divergence of replicate populations. *BMC Evol Biol* **10**, 11.

**Crow JF & Simmons MJ (1983)** *The Genetics and Biology of Drosophila*. New York: Acadamic.

**Crow JS & Kimura M (1970)** *An Introduction to Population Genetic Theory*. New York: Harper and Row.

**da Silva J, Coetzer M, Nedellec R, Pastore C & Mosier DE (2010)** Fitness epistasis and constraints on adaptation in a human immunodeficiency virus type 1 protein region. *Genetics* **185**, 293-303.

**Darwin CR (1859)** *On the Origin of Species by Means of Natural Selection*. Murray, London.

**Darwin CR (1882)** *On the Various Contrivances by which British and Foreign Orchids are Fertilised by Insects*, 2nd ed. New York: D. Appleton.

**Darwin CR & Wallace A (1858)** On the tendency of species to form varieties; and on the perpetuation of varieties and species by natural means of selection *J Linn Soc Lond* **3**, 45-62.

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

**Dawkins R (1986)** *The Blind Watchmaker*. London: Penguin Books Ltd.

**de Visser JA & Elena SF (2007)** The evolution of sex: empirical insights into the roles of epistasis and drift. *Nat Rev Genet* **8**, 139-149.

**de Visser JA, Park SC & Krug J (2009)** Exploring the effect of sex on empirical fitness landscapes. *Am Nat* **174**, 15-30.

**Dennett DC (1995)** *Darwin's Dangerous Idea: Evolution and the Meanings of Life*. London: Simon & Schuster.

**Ditta G, Stanfield S, Corbin D & Helinski DR (1980)** Broad host range DNA cloning system for gram-negative bacteria - Construction of a gene Bank of *Rhizobium meliloti*. *Proc Natl Acad Sci USA* **77**, 7347-7351.

**Dobler S, Dalla S, Wagschal V & Agrawal AA (2012)** Community-wide convergent evolution in insect adaptation to toxic cardenolides by substitutions in the Na,K-ATPase. *Proc Natl Acad Sci USA* **109**, 13040-13045.

**Duffy S, Burch CL & Turner PE (2007)** Evolution of host specificity drives reproductive isolation among RNA viruses. *Evolution* **61**, 2614-2622.

**Eckardt NA (2008)** Epistasis and genetic regulation of variation in the Arabidopsis Metabolome. *Plant Cell* **20**, 1185-1186.

**Elena SF & Lenski RE (1997a)** Test of synergistic interactions among deleterious mutations in bacteria. *Nature* **390**, 395-398.

**Elena SF & Lenski RE (1997b)** Long-term experimental evolution in *Escherichia coli.* VII. Mechanisms maintaining variability within populations. *Evolution* **51**, 1058-1967.

**Ferguson GC, Bertels F & Rainey PB (2013)** Adaptive divergence in experimental populations of *Pseudomonas fluorescens*. V. Genetic and phenotypic bases of fuzzy spreader fitness. (in press). *Genetics* **195**, 1319-1335.

**Figurski DH & Helinski DR (1979)** Replication of an origin-containing derivative of plasmid RK2 dependent on a plasmid function provided in trans. *Proc Natl Acad Sci USA* **76**, 1648-1652.

**Franke J, Klozer A, de Visser JA & Krug J (2011)** Evolutionary accessibility of mutational pathways. *PLoS Comput Biol* **7**, e1002134.

**Friesen ML, Saxer G, Travisano M & Doebeli M (2004)** Experimental evidence for sympatric ecological diversification due to frequency-dependent competition in *Escherichia coli*. *Evolution* **58**, 245-260.

**Gallie J (2009)** Evolutionary and molecular origins of a phenotypic switch in *Pseudomonas fluorescens* SBW25. Ph.D. dissertation, Massey University, Auckland.

**Gavrilets S (2004)** *Fitness Landscapes and the Origin of Species*. Princeton, N.J. ; Oxford: Princeton University Press.

**Gehrig SM (2005)** Adaptation of *Pseudomonas fluorescens* SBW25 to the air-liquid interface : a study in evolutionary genetics. Ph.D. dissertation, University of Oxford, Oxford.

**Gillespie JH (2006)** Stochastic Processes in Evolution. In *Evolutionary Genetics - Concepts and Case Studies*, pp. 65-79. [CW Fox and JB Wolf] New York: Oxford University Press.

**Goldsmith TH (1990)** Optimization, constraint, and history in the evolution of eyes. *Q Rev Biol* **65**, 281-322.

**Gould SJ (1987)** *The Panda's Thumb*. New York: Penguin Books Ltd.

**Gould SJ (1989)** *Wonderful Life : the Burgess Shale and the Nature of History*, 1st ed. New York: W.W. Norton.

**Gould SJ (2002)** *The Structure of Evolutionary Theory*. Cambridge, Mass.; London: Belknap.

**Gould SJ & Lewontin RC (1979)** The spandrels of San Marco and the Panglossian paradigm: a critique of the adaptationist programme. *Proc R Soc Lond B Biol Sci* **205**, 581-598.

**Goymer P, Kahn SG, Malone JG, Gehrig SM, Spiers AJ & Rainey PB (2006)** Adaptive divergence in experimental populations of *Pseudomonas fluorescens*. II. Role of the GGDEF regulator WspR in evolution and development of the wrinkly spreader phenotype. *Genetics* **173**, 515-526.

**Guvener ZT & McCarter LL (2003)** Multiple regulators control capsular polysaccharide production in *Vibrio parahaemolyticus*. *J Bacteriol* **185**, 5431-5441.

**Halder G, Callaerts P & Gehring WJ (1995a)** New perspectives on eye evolution. *Curr Opin Genet Dev* **5**, 602-609.

**Halder G, Callaerts P & Gehring WJ (1995b)** Induction of ectopic eyes by targeted expression of the eyeless gene in *Drosophila*. *Science* **267**, 1788-1792.

**Hall BG (1981)** Changes in the substrate specificities of an enzyme during directed evolution of new functions. *Biochemistry* **20**, 4042-4049.

**Hall BG (2002)** Predicting evolution by in vitro evolution requires determining evolutionary pathways. *Antimicrob Agents Chemother* **46**, 3035-3038.

**Hegreness M, Shoresh N, Hartl D & Kishony R (2006)** An equivalence principle for the incorporation of favorable mutations in asexual populations. *Science* **311**, 1615-1617.

**Helling RB, Vargas CN & Adams J (1987)** Evolution of *Escherichia coli* during growth in a constant environment. *Genetics* **116**, 349-358.

**Herring CD, Raghunathan A, Honisch C (2006)** Comparative genome sequencing of *Escherichia coli* allows observation of bacterial evolution on a laboratory timescale. *Genetics* **38**, 1406-1412.

**Hoekstra HE, Hirschmann RJ, Bundey RA, Insel PA & Crossland JP (2006)** A single amino acid mutation contributes to adaptive beach mouse color pattern. *Science* **313**, 101-104.

**Jablonski D (1986)** Background and mass extinctions: the alternation of macroevolutionary regimes. *Science* **231**, 129-133.

**Jacob F (1977)** Evolution and tinkering. *Science* **196**, 1161-1166.

**Jasmin JN & Kassen R (2007)** Evolution of a single niche specialist in variable environments. *Proc R Soc Lond B Biol Sci* **274**, 2761-2767.

**Kahn SG (1998)** Molecular characterisation of genes essential for adaptive evolution in *Pseudomonas fluorescens* SBW25. Ph.D. dissertation, University of Oxford, Oxford.

**Kassen R (2009)** Toward a general theory of adaptive radiation: insights from microbial experimental evolution. *Annals of The New York Academy of Sciences* **1168**, 3-22.

**Kassen R & Bataillon T (2006)** Distribution of fitness effects among beneficial mutations before selection in experimental populations of bacteria. *Nat Genet* **38**, 484-488.

**Kassen R, Buckling A, Bell G & Rainey PB (2000)** Diversity peaks at intermediate productivity in a laboratory microcosm. *Nature* **406**, 508-512.

**Kauffman SA (2000)** *Investigations*. New York ; Oxford: Oxford University Press.

**Khan AI, Dinh DM, Schneider D, Lenski RE & Cooper TF (2011)** Negative epistasis between beneficial mutations in an evolving bacterial population. *Science* **332**, 1193-1196.

**Kibota TT & Lynch M (1996)** Estimate of the genomic mutation rate deleterious to overall fitness in *E. coli*. *Nature* **381**, 694-696.

**Kimura M (1983)** *The neutral theory of molecular evolution*. Cambridge: Cambridge University Press.

**King EO, Ward MK & Raney DE (1954)** Two simple media for the demonstration of pyocyanin and fluorescin. *J Lab Clin Med* **44**, 301-307.

**King JL & Jukes TH (1969)** Non-Darwinian evolution. *Science* **164**, 788-798.

**Kondrashov AS (1988)** Deleterious mutations and the evolution of sexual reproduction. *Nature* **336**, 435-440.

**Korona R (1996)** Genetic divergence and fitness convergence under uniform selection in experimental populations of bacteria. *Genetics* **143**, 637-644.

**Korona R, Nakatsu CH, Forney LJ & Lenski RE (1994)** Evidence for multiple adaptive peaks from populations of bacteria evolving in a structured habitat. *Proc Natl Acad Sci USA* **91**, 9037-9041.

**Koza A, Moshynets O, Otten W & Spiers AJ (2011)** Environmental modification and niche construction: developing $O_2$ gradients drive the evolution of the wrinkly spreader. *ISME J* **5**, 665-673.

**Kozmik Z (2005)** Pax genes in eye development and evolution. *Curr Opin Genet Dev* **15**, 430-438.

**Kozmik Z (2008)** The role of Pax genes in eye evolution. *Brain Res Bull* **75**, 335-339.

**Kvitek DJ & Sherlock G (2011)** Reciprocal sign epistasis between frequently experimentally evolved adaptive mutations causes a rugged fitness landscape. *PLoS Genet* **7**, e1002056.

**Lamb TD, Arendt D & Collin SP (2009)** The evolution of phototransduction and eyes. *Philos Trans R Soc Lond B Biol Sci* **364**, 2791-2793.

**Lamb TD, Collin SP & Pugh EN, Jr. (2007)** Evolution of the vertebrate eye: opsins, photoreceptors, retina and eye cup. *Nat Rev Neurosci* **8**, 960-976.

**Lang GI, Botstein D & Desai MM (2011)** Genetic variation and the fate of beneficial mutations in asexual populations. *Genetics* **188**, 647-661.

**Lenski RE (2004)** Phenotypic and genomic evolution during a 20,000-generation experiment with the bacterium *Escherichia coli*. In *Plant Breeding Reviews*, pp. 225-265. John Wiley & Sons, Inc.

**Lenski RE, Ofria C, Pennock RT & Adami C (2003)** The evolutionary origin of complex features. *Nature* **423**, 139-144.

**Lenski RE, Rose MR, Simpson SC & Tadler SC (1991)** Long-term experimental evolution in *Escherichia coli* .1. Adaptation and divergence during 2,000 generations. *Am Nat* **138**, 1315-1341.

**Lenski RE & Travisano M (1994)** Dynamics of adaptation and diversification: a 10,000-generation experiment with bacterial populations. *Proc Natl Acad Sci USA* **91**, 6808-6814.

**Lewontin RC (1966)** Is nature probable or capricious? *BioScience* **16**, 25-27.

**Lewontin RC (1970)** The units of selection. *Annu Rev Ecol and Syst* **1**, 1-18.

**Libby E & Rainey PB (2011)** Exclusion rules, bottlenecks and the evolution of stochastic phenotype switching. *Proc R Soc Lond B Biol Sci* **278**, 3574-3583.

**Losos JB, Jackman TR, Larson A, Queiroz K & Rodriguez-Schettino L (1998)** Contingency and determinism in replicated adaptive radiations of island lizards. *Science* **279**, 2115-2118.

**Lunzer M, Miller SP, Felsheim R & Dean AM (2005)** The biochemical architecture of an ancient adaptive landscape. *Science* **310**, 499-501.

**Lynch M (2007)** The frailty of adaptive hypotheses for the origins of organismal complexity. *Proc Natl Acad Sci USA* **104**, 8597-8604.

**MacLean RC, Bell G & Rainey PB (2004)** The evolution of a pleiotropic fitness tradeoff in *Pseudomonas fluorescens*. *Proc Natl Acad Sci USA* **101**, 8072-8077.

**Maharjan R, Seeto S, Notley-McRobb L & Ferenci T (2006)** Clonal adaptive radiation in a constant environment. *Science* **313**, 514-517.

**Maisnier-Patin S, Roth JR, Fredriksson A, Nystrom T, Berg OG & Andersson DI (2005)** Genomic buffering mitigates the effects of deleterious mutations in bacteria. *Nat Genet* **37**, 1376-1379.

**Malone JG, Williams R, Christen M, Jenal U, Spiers AJ & Rainey PB (2007)** The structure-function relationship of WspR, a *Pseudomonas fluorescens* response regulator with a GGDEF output domain. *Microbiology* **153**, 980-994.

**Manceau M, Domingues VS, Linnen CR, Rosenblum EB & Hoekstra HE (2010)** Convergence in pigmentation at multiple levels: mutations, genes and function. *Philos Trans R Soc Lond B Biol Sci* **365**, 2439-2450.

**Mani GS & Clarke BC (1990)** Mutational order: a major stochastic process in evolution. *Proc R Soc Lond B Biol Sci* **240**, 29-37.

**Maynard Smith J, Burian R, Kauffman S, Alberch P, Campbell J, Goodwin B, Lande R, Raup D & Wolpert L (1985)** Developmental constraints and evolution. *Q Rev Biol* **60**, 265-287.

**Mayr E (1983)** How to carry out the adaptationist program? *Am Nat* **121**, 324-334.

**McDonald MJ, Cooper TF, Beaumont HJE & Rainey PB (2011)** The distribution of fitness effects of new beneficial mutations in *Pseudomonas fluorescens*. *Biology Letters* **7**, 98-100.

**McDonald MJ, Gehrig SM, Meintjes PL, Zhang XX & Rainey PB (2009)** Adaptive divergence in experimental populations of *Pseudomonas fluorescens*. IV. Genetic constraints guide evolutionary trajectories in a parallel adaptive radiation. *Genetics* **183**, 1041-1053.

**Meyer JR, Dobias DT, Weitz JS, Barrick JE, Quick RT & Lenski RE (2012)** Repeatability and contingency in the evolution of a key innovation in phage lambda. *Science* **335**, 428-432.

**Meyers LA & Bull JJ (2002)** Fighting change with change: adaptive variation in an uncertain world. *Trends Ecol Evol* **17**, 551-557.

**Miller SP, Lunzer M & Dean AM (2006)** Direct demonstration of an adaptive constraint. *Science* **314**, 458-461.

**Monod J & Wainhouse A (1971)** *Chance and Necessity : An Essay on the Natural Philosophy of Modern Biology*. New York: Alfred A. Knopf.

**Mundy NI, Badcock NS, Hart T, Scribner K, Janssen K & Nadeau NJ (2004)** Conserved genetic basis of a quantitative plumage trait involved in mate choice. *Science* **303**, 1870-1873.

**Nachman MW, Hoekstra HE & D'Agostino SL (2003)** The genetic basis of adaptive melanism in pocket mice. *Proc Natl Acad Sci USA* **100**, 5268-5273.

**Nakatsu CH, Korona R, Lenski RE, de Bruijn FJ, Marsh TL & Forney LJ (1998)** Parallel and divergent genotypic evolution in experimental populations of *Ralstonia* sp. *J Bacteriol* **180**, 4325-4331.

**Negoro S, Kato K, Fujiyama K & Okada H (1994)** The nylon oligomer biodegradation system of *Flavobacterium* and *Pseudomonas*. *Biodegradation* **5**, 185-194.

**Novais A, Comas I, Baquero F, Canton R, Coque TM, Moya A, Gonzalez-Candelas F & Galan JC (2010)** Evolutionary trajectories of beta-lactamase CTX-M-1 cluster enzymes: predicting antibiotic resistance. *PLoS Pathog* **6**, e1000735.

**O'Maille PE, Malone A, Dellas N, Andes Hess B, Jr., Smentek L, Sheehan I, Greenhagen BT, Chappell J, Manning G & Noel JP (2008)** Quantitative exploration of the catalytic landscape separating divergent plant sesquiterpene synthases. *Nat Chem Biol* **4**, 617-623.

**Ostrowski EA, Woods RJ, Lenski RE (2008)** The genetic basis of parallel and divergent phenotypic response in evolving populations of Escherichia coli. *Proc R Soc Lond B Biol Sci* **275**, 277-284.

**Ortlund EA, Bridgham JT, Redinbo MR & Thornton JW (2007)** Crystal structure of an ancient protein: evolution by conformational epistasis. *Science* **317**, 1544-1548.

**Otto SP & Lenormand T (2002)** Resolving the paradox of sex and recombination. *Nat Rev Genet* **3**, 252-261.

**Pal C, Papp B, Lercher MJ, Csermely P, Oliver SG & Hurst LD (2006)** Chance and necessity in the evolution of minimal metabolic networks. *Nature* **440**, 667-670.

**Pelosi L, Kuhn L, Guetta D, Garin J, Geiselmann J, Lenski RE & Schneider D (2006)** Parallel changes in global protein profiles during long-term experimental evolution in *Escherichia coli*. *Genetics* **173**, 1851-1869.

**Pigliucci M (2008)** What, if anything, is an evolutionary novelty? *Philos Science* **75**.

**Poelwijk FJ, Kiviet DJ, Weinreich DM & Tans SJ (2007)** Empirical fitness landscapes reveal accessible evolutionary paths. *Nature* **445**, 383-386.

**Pogulis RJ, Vallejo AN & Pease LR (1996)** In vitro recombination and mutagenesis by overlap extension PCR. *Methods Mol Biol* **57**, 167-176.

**Pross A (2003)** The driving force for life's emergence: kinetic and thermodynamic considerations. *J Theor Biol* **220**, 393-406.

**Quinn G, Keough M (2002)** *Experimental Design and Data Analysis for Biologists.* Cambridge: Cambridge University Press.

**Quiring R, Walldorf U, Kloter U & Gehring WJ (1994)** Homology of the eyeless gene of *Drosophila* to the *Small eye* gene in mice and *Aniridia* in humans. *Science* **265**, 785-789.

**Rainey PB (1999)** Adaptation of *Pseudomonas fluorescens* to the plant rhizosphere. *Environ Microbiol* **1**, 243-257.

**Rainey PB & Bailey MJ (1996)** Physical and genetic map of the *Pseudomonas fluorescens* SBW25 chromosome. *Mol Microbiol* **19**, 521-533.

**Rainey PB, Beaumont HJ, Ferguson GC, Gallie J, Kost C, Libby E & Zhang XX (2011)** The evolutionary emergence of stochastic phenotype switching in bacteria. *Microb Cell Fact* **10** (Suppl 1), S14.

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

**Rainey PB & Rainey K (2003)** Evolution of cooperation and conflict in experimental bacterial populations. *Nature* **425**, 72-74.

**Rainey PB & Travisano M (1998)** Adaptive radiation in a heterogeneous environment. *Nature* **394**, 69-72.

**Ralser M, Querfurth R, Warnatz HJ, Lehrach H, Yaspo ML & Krobitsch S (2006)** An efficient and economic enhancer mix for PCR. *Biochem Biophys Res Commun* **347**, 747-751.

**Reetz MT & Sanchis J (2008)** Constructing and analyzing the fitness landscape of an experimental evolutionary process. *Chembiochem* **9**, 2260-2267.

**Ritland K, Newton C & Marshall HD (2001)** Inheritance and population structure of the white-phased "Kermode" black bear. *Curr Biol* **11**, 1468-1472.

**Rosenblum EB, Hoekstra HE & Nachman MW (2004)** Adaptive reptile color variation and the evolution of the *Mc1r* gene. *Evolution* **58**, 1794-1808.

**Rosenzweig RF, Sharp RR, Treves DS & Adams J (1994)** Microbial evolution in a simple unstructured environment: genetic differentiation in *Escherichia coli*. *Genetics* **137**, 903-917.

**Ross P, Weinhouse H, Aloni Y, Michaeli D, Weinberger-Ohana P, Mayer R, Braun S, de Vroom E, van der Marel GA, van Boom JH & Benziman M (1987)** Regulation of cellulose synthesis in *Acetobacter xylinum* by cyclic diguanylic acid. *Nature* **325**, 279-281.

**Rozen DE & Lenski RE (2000)** Long-term experimental evolution in *Escherichia coli*. XIII. Dynamics of a balanced polymorphism. *Am Nat.* **155**, 24-35.

**Salverda ML, Dellus E, Gorter FA, Debets AJ, van der Oost J, Hoekstra RF, Tawfik DS & de Visser JA (2011)** Initial mutations direct alternative pathways of protein evolution. *PLoS Genet* **7**, e1001321.

**Sambrook J & Russell DW (2001)** *Molecular Cloning: a Laboratory Manual, Vol. 3*, 3rd ed. New York: Cold Spring Harbor Laboratory Press (CSHL Press).

**Saxer G, Doebeli M & Travisano M (2010)** The repeatability of adaptive radiation during long-term experimental evolution of *Escherichia coli* in a multiple nutrient environment. *PLoS One* **5**, e14184.

**Schluter D (1996)** Adaptive radiation along genetic lines of least resistance. *Evolution* **50**, 1766-1774.

**Segre D, Deluna A, Church GM & Kishony R (2005)** Modular epistasis in yeast metabolism. *Nat Genet* **37**, 77-83.

**Shindo C, Aranzana MJ, Lister C, Baxter C, Nicholls C, Nordborg M & Dean C (2005)** Role of *FRIGIDA* and *FLOWERING LOCUS C* in determining variation in flowering time of *Arabidopsis*. *Plant Physiol* **138**, 1163-1173.

**Silby MW, Cerdeno-Tarraga AM, Vernikos GS, Giddens SR, Jackson RW, Preston GM, Zhang XX, Moon CD, Gehrig SM, Godfrey SA, Knight CG, Malone JG, Robinson Z, Spiers AJ, Harris S, Challis GL, Yaxley AM, Harris D, Seeger K, Murphy L, Rutter S, Squares R, Quail MA, Saunders E, Mavromatis K, Brettin TS, Bentley SD, Hothersall J, Stephens E, Thomas CM, Parkhill J, Levy SB, Rainey PB & Thomson NR (2009)** Genomic and genetic analyses of diversity and plant interactions of *Pseudomonas fluorescens*. *Genome Biol* **10**, R51.

**Simpson GG (1949)** *The Meaning of Evolution: A Study of the History of Life and of its Significance for Man*. New Haven: Yale University Press.

**Sniegowski PD & Gerrish PJ (2010)** Beneficial mutations and the dynamics of adaptation in asexual populations. *Philos Trans R Soc Lond B Biol Sci* **365**, 1255-1263.

**Spiers AJ, Bohannon J, Gehrig SM & Rainey PB (2003)** Biofilm formation at the air-liquid interface by the *Pseudomonas fluorescens* SBW25 wrinkly spreader requires an acetylated form of cellulose. *Mol Microbiol* **50**, 15-27.

**Spiers AJ, Kahn SG, Bohannon J, Travisano M & Rainey PB (2002)** Adaptive divergence in experimental populations of *Pseudomonas fluorescens*. I. Genetic and phenotypic bases of wrinkly spreader fitness. *Genetics* **161**, 33-46.

**Spiess EB (1989)** *Genes in Populations*. New York: Wiley.

**Stearns SC & Hoekstra RF (2005)** *Evolution: an Introduction*, 2nd ed. Oxford: Oxford University Press.

**Stern DL & Orgogozo V (2009)** Is genetic evolution predictable? *Science* **323**, 746-751.

**Sugawara E & Nikaido H (1992)** Pore-forming activity of OmpA protein of *Escherichia coli*. *J Biol Chem* **267**, 2507-2511.

**Suzuki DT, Griffiths AJF, Miller JH & Lewontin RC (1989)** *An Introduction to Genetic Analysis*, 4th ed. New York: Freeman.

**Tenaillon O, Rodriguez-Verdugo A, Gaut RL, McDonald P, Bennett AF, Long AD & Gaut BS (2012)** The molecular diversity of adaptive convergence. *Science* **335**, 457-461.

**Theron E, Hawkins K, Bermingham E, Ricklefs RE & Mundy NI (2001)** The molecular basis of an avian plumage polymorphism in the wild: a melanocortin-1-receptor point mutation is perfectly associated with the melanic plumage morph of the bananaquit, *Coereba flaveola*. *Curr Biol* **11**, 550-557.

**Travisano M, Mongold JA, Bennett AF & Lenski RE (1995a)** Experimental tests of the roles of adaptation, chance, and history in evolution. *Science* **267**, 87-90.

**Travisano M, Vasi F & Lenski RE (1995b)** Long-term experimental evolution in *Escherichia coli*. III. Variation among replicate populations in correlated responses to novel environments. *Evolution* **49**, 189-200.

**Treves DS, Manning S & Adams J (1998)** Repeated evolution of an acetate-crossfeeding polymorphism in long-term populations of *Escherichia coli*. *Mol Biol Evol* **15**, 789-797.

**Turner PE, Souza V & Lenski RE (1996)** Tests of ecological mechanisms promoting the stable coexistence of two bacterial genotypes. *Ecology* **77**, 2119-2129.

**Vermeij GJ (2006)** Historical contingency and the purported uniqueness of evolutionary innovations. *Proc Natl Acad Sci USA* **103**, 1804-1809.

**Wagenaar DA & Adami C (2004)** Influence of chance, history, and adaptation on digital evolution. *Artif Life* **10**, 181-190.

**Wayne ML & Miyamoto MM (2006)** Genetic Variation. In *Evolutionary Genetics - Concepts and Case Studies*, pp. 14-31. [CW Fox and JB Wolf] New York: Oxford University Press.

**Weinreich DM, Delaney NF, Depristo MA & Hartl DL (2006)** Darwinian evolution can follow only very few mutational paths to fitter proteins. *Science* **312**, 111-114.

**Weinreich DM, Watson RA & Chao L (2005)** Perspective: Sign epistasis and genetic constraint on evolutionary trajectories. *Evolution* **59**, 1165-1174.

**Wichman HA, Badgett MR, Scott LA, Boulianne CM & Bull JJ (1999)** Different trajectories of parallel evolution during viral adaptation. *Science* **285**, 422-424.

**Wischmann S, Floreano D & Keller L (2012)** Historical contingency affects signaling strategies and competitive abilities in evolving populations of simulated robots. *Proc Natl Acad Sci USA* **109**, 864-868.

**Woods R, Schneider D, Winkworth CL, Riley MA & Lenski RE (2006)** Tests of parallel molecular evolution in a long-term experiment with *Escherichia coli*. *Proc Natl Acad Sci USA* **103**, 9107-9112.

**Yokoyama S & Radlwimmer FB (2001)** The molecular genetics and evolution of red and green color vision in vertebrates. *Genetics* **158**, 1697-1710.

# Appendices

## A1 Table of Abbreviations

**Table A1:**

| Abbreviation | Meaning |
|---|---|
| ACP | Acetylated cellulosic polymer |
| aws | Alternate wrinkly spreader |
| bp | Base pairs |
| BSA | Bovine serum albumin |
| c-di-GMP | Cyclic-dimeric-guanosine monophosphate |
| CIP | Calf-intestinal phosphatase |
| CP | Cytoplasm |
| DGC | Di-guanylate cyclase |
| dNTP | dinucleotide triphosphate |
| EPS | Extracellular polymeric substance |
| FS | Fuzzy spreader |
| Gm | Gentamycin |
| IM | Inner membrane |
| kb | Kilobase |
| KB | King's medium |
| Km | Kanamycin |
| LB | Lysogeny broth |
| MCP | Methyl-accepting chemotaxis protein |
| MIC | Minimal inhibitory concentration |
| mws | Mike’s wrinkly spreader |
| NF | Nitrofurantonin |
| OD | Optical density |
| OM | Outer membrane |
| PCR | Polymerase-chain-reaction |
| PDE | Phosphodiasterase |
| REE | Reverse-Evolution Experiment |
| SEM | Standard error of mean |
| SM | Smooth type |
| SOC | Super optimal broth with Catabolite repression |
| SREE | Switcher Re-Evolution Experiment |
| SW | Stochastically switching phenotype |
| TAE | Tris-acetate-EDTA buffer |

| | |
|---|---|
| TBE | Tris-borate-EDTA buffer |
| Tc | Tetracyclin |
| TM | Transmembrane region |
| WS | Wrinkly spreader |
| wsp | Wrinkly spreader operon |
| wss | Wrinkly spreader structural operon |
| Xgal | 5-bromo-4-chloro-3-indolyl-beta-D-galactopyranoside |

## A2 Appendix material from Chapter 4

See Chapter 4, section *4.3.2.1* for an explanation of the data.

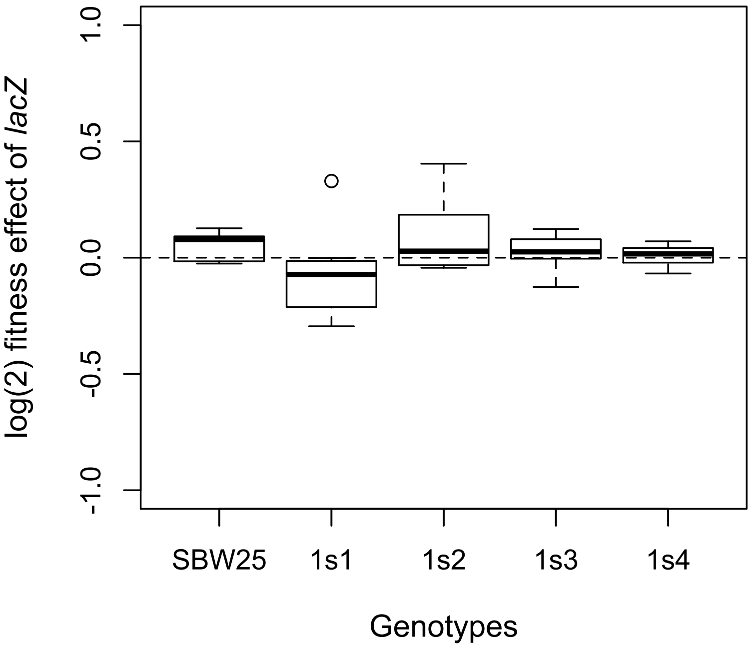

**Figure A1: Relative fitness of *lacZ* transcriptional fusion in SBW25, 1s1, 1s2, 1s3 and 1s4.** The relative fitness is displayed as log2 of the ratio of the two Malthusian parameters of each competitor. The dashed line at zero indicates no fitness difference after *lacZ* was introduced into the genomes. Values above the dashed line indicate a fitness increase and below the line indicate fitness decrease. None of the strains showed a significant change in fitness by the introduction of l*acZ* (see Chapter 4, Tab. 4.5). Displayed are medians and interquartile ranges based on *N* = 5-8.

## A3 Appendix material from Chapter 6

See Chapter 6, section 6.3.1 for an explanation of the data.

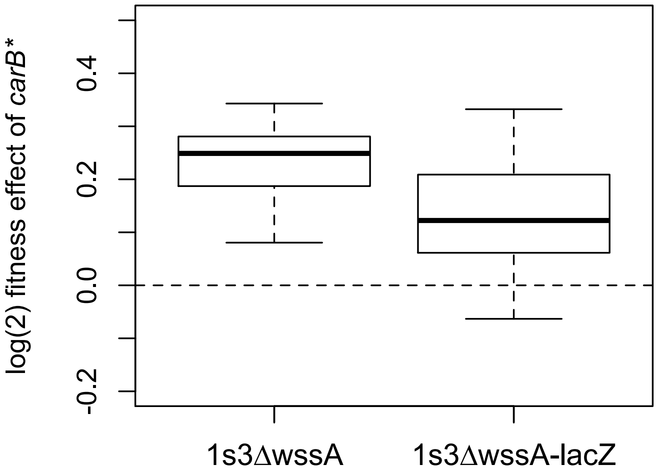

**Figure A2: Relative fitness of *carB** in 1s3Δ*wssA* and 1s3Δ*wssA-lacZ.*** The relative fitness is displayed as log2 of the ratio of the two Malthusian parameters of each competitor. The dashed line at zero indicates no fitness difference between competitors. Values above the dashed line indicate a fitness increase and below the line indicate fitness decrease. Displayed are medians and interquartile ranges based on *N* = 15. Both strains, 1s3Δ*wssA* and 1s3Δ*wssA-lacZ* showed a significant fitness increase after *carB** was introduced into their genomes (see Chapter 6).