The evolution of natural competence in

*Streptococcus pneumoniae*

A thesis submitted to the University of Manchester for the degree of Doctor of Philosophy in the Faculty of Life Sciences

2012

Daniel J.P. Engelmoer
<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>List of figures</td>
<td>7</td>
</tr>
<tr>
<td>List of tables</td>
<td>9</td>
</tr>
<tr>
<td>Abstract</td>
<td>11</td>
</tr>
<tr>
<td>Declaration</td>
<td>13</td>
</tr>
<tr>
<td>Copyright</td>
<td>15</td>
</tr>
<tr>
<td>Acknowledgements</td>
<td>17</td>
</tr>
<tr>
<td>Chapter contributions</td>
<td>19</td>
</tr>
<tr>
<td>Preface</td>
<td>21</td>
</tr>
<tr>
<td>1. General introduction</td>
<td>23</td>
</tr>
<tr>
<td>1.1 The evolutionary enigma of recombination</td>
<td>23</td>
</tr>
<tr>
<td>1.2 Recombination in bacteria</td>
<td>25</td>
</tr>
<tr>
<td>1.3 The competent cell state</td>
<td>26</td>
</tr>
<tr>
<td>1.4 Maintenance of competence in bacterial populations</td>
<td>27</td>
</tr>
<tr>
<td>1.4.1 Competence as a DNA repair mechanism</td>
<td>28</td>
</tr>
<tr>
<td>1.4.2 Competence increases the speed of evolution</td>
<td>29</td>
</tr>
<tr>
<td>1.4.3 DNA as a resource</td>
<td>31</td>
</tr>
<tr>
<td>1.4.4 Many hypotheses with only one winner?</td>
<td>32</td>
</tr>
<tr>
<td>1.5 Introducing the model: Streptococcus pneumoniae</td>
<td>32</td>
</tr>
<tr>
<td>1.6 The molecular mechanism of competence</td>
<td>34</td>
</tr>
<tr>
<td>1.7 Project aims and thesis structure</td>
<td>36</td>
</tr>
<tr>
<td>1.8 References</td>
<td>38</td>
</tr>
<tr>
<td>2. Competence increases survival during stress in <em>Streptococcus pneumoniae</em></td>
<td>47</td>
</tr>
<tr>
<td>2.1 Abstract</td>
<td>47</td>
</tr>
<tr>
<td>2.2 Introduction</td>
<td>48</td>
</tr>
<tr>
<td>2.3 Material&amp;Methods</td>
<td>51</td>
</tr>
</tbody>
</table>
2.3.1 Strains and culture conditions 51
2.3.2 Competence induction after or before antibiotic stress 52
2.3.3 Requirements for increased survival 54
2.3.4 The duration of benefits of competence 54

2.4 Results 56
2.4.1 Competence induction following antibiotic stress 56
2.4.2 Competence induction before or during antibiotic stress 56
2.4.3 Benefits of competence versus transformation 58
2.4.4 The duration of benefits of competence 60

2.5 Discussion 62

2.6 Acknowledgements 67

2.7 References 68

3. Fitness associated recombination benefits competence in \textit{Streptococcus pneumoniae} 73

3.1 Abstract 73

3.2 Introduction 75

3.3 Materials & Methods 76
3.3.1 Strains, culture conditions and chemostats 76
3.3.2 Long-term evolution experiment 77
3.3.3 Fitness assays 78
3.3.4 DNA isolation and sequencing 79
3.3.5 Mutation analysis 80
3.3.6 Mutations and mutation rate 80

3.4 Results 82
3.4.1 Costs of competence 82
3.4.2 Transformation is conservative 83
3.4.3 Mutations and the mutation rate 84

3.5 Discussion 86

3.6 References 90

3.7 Supplemental material chapter 3 95
S3.1 - Growth rates of the four ancestors 95
S3.2 - recombination in the chemostat environment 95
S3.3- growth rate of ancestors exposed to different concentrations of kanamycin 96
S3.4 - R-code used for the mixed models 97

4. Phenotypic and genotypic parallel evolution in long-term evolved \textit{Streptococcus pneumoniae} 99
4.1 Abstract

4.2 Introduction

4.3 Material and Methods
    4.3.1 Strains & culture conditions
    4.3.2 Oscillatory population dynamics
    4.3.3 Growth rate
    4.3.4 Biofilm formation
    4.3.5 Parallel genomic changes

4.4 Results
    4.4.1 Oscillatory population dynamics
    4.4.2 Growth rate
    4.4.3 Biofilm formation
    4.4.4 Genotypic changes

4.5 Discussion

4.6 References

5. Testing the Pharaoh’s curse: surface survival and virulence in Streptococcus pneumoniae
    5.1 Abstract
    5.2 Introduction
    5.3 Material & Methods
        5.3.1 Strains and culture conditions
        5.3.2 Quantifying durability
        5.3.3 Durability and pneumococcal life history
    5.4 Results
        5.4.1 Durability depends on genotype and serotype
        5.4.2 Durability, carriage and virulence
    5.5 Discussion
    5.6 References

6. General discussion
    6.1 Short-term benefits
    6.2 Long-term benefits
    6.3 Is competence always conservative?
    6.4 Mutators, parallel evolution
    6.5 The Pharaoh hypothesis
    6.5 Conclusion
    6.6 References
List of figures

1. General introduction
   Figure 1.1. Schematic overview of early com gene regulation. 34

2. Competence increases survival during stress in \textit{Streptococcus pneumoniae}
   Figure 2.1. Schematic picture of the experimental setup used for stress before competence. 53
   Figure 2.2. Growth or death rate of FP5 exposed to various antibiotics before or after competence was induced. 57
   Figure 2.3. Death rate of FP5 exposed to mitomycin C or streptomycin in the presence of independent combinations of CSP and DNA. 59
   Figure 2.4. Mean growth curves of FP5 (black lines) in the presence or absence of CSP and DNA while exposed to sub-MIC concentrations of varying antibiotics. 61
   Figure 2.5. Mean relative fitness of FP5 over FP48 measured during exposure to various antibiotics. 62

3. Fitness associated recombination benefits competence in \textit{Streptococcus pneumoniae}
   Figure 3.1. Experimental environment and schematic setup of the long-term evolution experiment. 78
   Figure 3.2. Mean selection rate constants of evolved populations. 82
   Figure 3.3. Mutations and mutation rates of evolved populations. 85
   Figure S3.1. Mean growth rates of the ancestral strains. 95
   Figure S3.2. Mean growth rates of the ancestral strains while exposed to different concentrations of kanamycin. 96

4. Phenotypic and genotypic parallel evolution in long-term evolved \textit{Streptococcus pneumoniae}
   Figure 4.1. Population density over time for the two ancestral strains. 106
   Figure 4.2. Population density of populations over time after each has evolved for 1000 generations. 107
   Figure 4.3. Long-term density dynamics of four evolved lineages. 107
   Figure 4.4. Relative growth rate of each evolved line after 1000 generations. 108
   Figure 4.5. Relative biofilm formation of evolved populations. 109
5. Testing the Pharaoh’s curse: surface survival and virulence in *Streptococcus pneumoniae*

Figure 5.1. Mean durability on plastic of strains grouped by serotype. 133
Figure 5.2. The relationship between mean durability on a plastic surface and several serotype specific characteristics. 135
List of tables

2. Competence increases survival during stress in *Streptococcus pneumoniae*

   Table 2.1. Details of strains used in this study. 52
   Table 2.2. Bonferroni corrected post-hoc test results for the death rate at both antibiotic stresses of each treatment. 59

4. Phenotypic and genotypic parallel evolution in long-term evolved *Streptococcus pneumoniae*

   Table 4.1. Summary of parallel mutations in loci. 112
   Table 4.2. Summary of parallel mutations in gene groups. 114
   Table 4.2. continued 115

5. Testing the Pharaoh’s curse: surface survival and virulence in *Streptococcus pneumoniae*

   Table 5.1. Strains used in this study with serotype and MLST information. 131
Abstract

The University of Manchester
Daniel Joachim Peter Engelmoer
Doctor of Philosphy
Evolution of natural competence in Streptococcus pneumoniae
29th June 2012

Naturally competent bacterial species are widespread across the bacterial tree-of-life. However, it remains unclear why competence has evolved and how it is maintained in these bacteria. Although it is likely that answers to these questions will be different for each species, a common selective factor cannot be excluded. Currently, three dominant hypotheses, which focus on transformation, try to explain the benefits of competence. Firstly, competence is thought to increase the rate of adaptation by combining beneficial alleles into single genotypes. Secondly, competence can repair DNA-damage by replacing the damaged DNA fragments with undamaged ones. Thirdly, the DNA taken up during competence can be used to recycle environmental DNA fragments for nutrients. One of the best studied naturally competent species is the Gram-positive Streptococcus pneumoniae, which is an opportunistic pathogen generally inhabiting the nasopharyngeal area of young children. Competence in S. pneumoniae is regulated via a density dependent extracellular signaling peptide. Here I use a combination of experiments designed around knockout mutants of the signaling mechanism and next-generation sequencing methods to examine evolutionary mechanisms maintaining competence in this species. I focus, in particular, on tests of the first two hypotheses. First, I extend on the DNA-for-repair hypothesis by showing that competent populations of S. pneumoniae are better protected not only against a DNA-damaging agent, but also against protein synthesis inhibitors. However, the mechanisms underlying this protection differ between types of stress. DNA-damage requires the full process of transformation, while protection against protein synthesis inhibitors only requires the activation of the competent cell state. This shows that benefits of competence cannot be totally explained by the benefits of transformation. Second, I use a long-term evolution experiment, where competent and non-competent strains are evolved in the presence and absence of periodic stress, to determine the importance of competence for adaptation. I find that competence does not increase the rate of adaptation in S. pneumoniae. The fitness of evolved competent populations was significantly lower than those of non-competent populations evolved over the same period of time. However, the intrinsic costs of competence are mitigated by the addition of short periods of stress exposure. These results confirm the prediction of the fitness associated recombination (FAR) hypothesis that competence is favoured in low-fitness situations. Thirdly, whole genome re-sequencing allowed me to explore genomic evolution of the evolved populations. These data revealed that competence reduces the mutational load of deleterious mutations rather than generating combinations of beneficial alleles. In addition I show several case of parallel genomic evolution within each treatment and across treatments. This shows that parallel evolution is not restricted by genotypic background (competence) or environment (periodic stress). In summary, these results show that competence has evolved in populations of S. pneumoniae as a mechanism to enable cells to respond to various forms of stress.
Declaration

No portion of the work referred to in this dissertation has been submitted in support of an application for another degree or qualification of this or any other university or other institute of learning.
Copyright

i. The author of this thesis (including any appendices and/or schedules to this thesis) owns certain copyright or related rights in it (the "Copyright") and s/he has given The University of Manchester certain rights to use such Copyright, including for administrative purposes.

ii. Copies of this thesis, either in full or in extracts and whether in hard or electronic copy, may be made only in accordance with the Copyright, Designs and Patents Act 1988 (as amended) and regulations issued under it or, where appropriate, in accordance with licensing agreements which the University has from time to time. This page must form part of any such copies made.

iii. The ownership of certain Copyright, patents, designs, trade marks and other intellectual property (the "Intellectual Property") and any reproductions of copyright works in the thesis, for example graphs and tables ("Reproductions"), which may be described in this thesis, may not be owned by the author and may be owned by third parties. Such Intellectual Property and Reproductions cannot and must not be made available for use without the prior written permission of the owner(s) of the relevant Intellectual Property and/or Reproductions.

iv. Further information on the conditions under which disclosure, publication and commercialisation of this thesis, the Copyright and any Intellectual Property and/or Reproductions described in it may take place is available in the University IP Policy (see http://www.campus.manchester.ac.uk/medialibrary/policies/intellectualproperty.pdf), in any relevant Thesis restriction declarations deposited in the University Library, The University Library’s regulations (see http://www.manchester.ac.uk/library/aboutus/regulations) and in The University’s policy on presentation of Theses.
Acknowledgements

I have to start thanking my supervisor, Daniel Rozen, for his tireless support during the three plus years I have worked on this thesis. Not only is his incredible enthusiasm for science contagious, but he also was always ready to provide advise, motivation and support at the right time. It was a real pleasure to work with him and I hope that this was not the last project we have worked on together. Being a PhD-student also means that you have to report regularly on the progress of your project and it was the job of Richard Preziosi to read these reports and make sure that I was on track. He did that with genuine interest, which is not the easiest of tasks if you are not a microbiologist. Several parts of this thesis could not have been done without the input of other people. So was it Casey Bergman who provided the opportunity to sequence the evolved populations and it was Ian Donaldson who turned these sequences into data that I could use to produce the results presented in this thesis. Chapter five started out as an undergraduate project that I continued, but it would not have existed without the preliminary work by Sarah Chang. And, off course, I cannot forget the (ex-)members of the Rozen group, who were able to put up with me for the years I worked in Manchester. Thank you Andres, Rob, Dave, Ben, Paul, Cath, Ailsa, Jung-Woo, Hannah and Monica. Working on a project for three years is not possible if you are able to relax every so often. Luckily there were Rob and Alex for nights with games and drinks, Andres, Sam, Dave, Veronica, Rob, Alex, Natalie and Paul for drinks after work, Kevin for the Wednesday evening film and Kim, John and the other Friday night lads for drinks and the occasional Christmas. Last but not least is Jenny, who appeared in the last months I was working on my thesis. She provided me with useful advice, she guarded my sanity, but most importantly she provided, and still provides, timely distraction from the world of science.
Author contributions

Chapter 1
- Daniel Engelmoer: Responsible for all content.

Chapter 2
- Daniel Engelmoer: Contributed to experimental design, experimental procedures, statistical analysis and writing.
- Daniel Rozen: Contributed as supervisor to experimental design and writing.

Chapter 3
- Daniel Engelmoer: Contributed to experimental design, experimental procedures, statistical analysis.
- Ian Donaldson: Contributed to this chapter by transforming the raw SOLiD data into BAM-files, which were used by Daniel Engelmoer for further analysis. He also contributed to the sequencing section in the material & methods section of this chapter.
- Daniel Rozen: Contributed to experimental design and writing.

Chapter 4
- Daniel Engelmoer: Contributed to experimental design, experimental procedures, statistical analysis and writing.
- Daniel Rozen: Contributed to experimental design and writing.

Chapter 5
- Daniel Engelmoer: Contributed to experimental design, experimental procedures, statistical analysis and writing.
- Sarah Chang: No contribution to chapter content. However, her student project initiated further investigation that was reported in this chapter. In addition, her test of different experimental designs was invaluable for this chapter.
- Daniel Rozen: Contributed to experimental design, statistical analysis and writing.

Chapter 6
- Daniel Engelmoer: Fully responsible for all content of this chapter.
This thesis is set in the ‘alternative’ format of the University of Manchester. The alternative format thesis is an assembly of chapters that are written as if they were a published paper with the exception of the general introduction and discussion. I choose to use this thesis format over the traditional format because, in my opinion, the writing of papers is an important skill a young scientist has to master. The thesis is the ideal opportunity for this during a PhD-project. In addition, chapter two is already published and it seems logical to me to insert that into the thesis as a chapter. Because each experimental chapter is written as if it is a paper, I have included all the names of the people that would have been authors if the paper had been published in a peer-reviewed journal. In the acknowledgements I have indicated the contributions of each of these authors. Although no PhD-project is done totally alone, the chapters in this thesis are the result of my own experimental work, interpretation and writing.
The molecular processes underlying evolutionary changes within and across species have been broadly elucidated since Charles Darwin published ‘On the origin of species by means of natural selection’ in 1859. Novel alleles arise through mutations, which result from inaccurate DNA-replication during cell division or from DNA-damage or as a side effect of damage repair. Mutation rates, in absence of DNA-damage, range from $10^{-4}$ mutations per base per generation for RNA viruses to $10^{-11}$ mutations per base per generation for eukaryotes (Drake et al. 1998). In addition to mutation, recombination can generate variation by shuffling existing alleles. In most eukaryotes recombination occurs during meiosis as a component of reproduction. For prokaryotes, by contrast, recombination and reproduction are unlinked.

Often the terms sex and recombination are used interchangeably, likely because of their interdependence in eukaryotes. To avoid confusion in this thesis, which focuses on prokaryotic recombination, I will always refer to recombination as the broad and generic process of ‘gene-shuffling’, which applies to both eukaryotes and prokaryotes. Where necessary, I will specify when I am specifically referring to sex, i.e. recombination and bi-parental reproduction.

1.1 The evolutionary enigma of recombination

It is still unclear why evolutionary processes for recombination (and sex) are maintained in so many different organisms across the entire tree of life (for review e.g. Agrawal 2006a; Otto 2009; Vos 2009). In clonal populations, interference between coexisting beneficial mu-
tations present in different cells can reduce the rate of adaptation by a process called clonal interference (Gerrish and Lenski 1998; de Visser and Rozen 2006; Fogle et al. 2008; Desai et al. 2007). In the 1930s Fisher and Muller first realized that if a population can recombine beneficial mutations into a single genotype, thus reducing clonal interference, then this would accelerate adaptation (Fisher 1930; Muller 1932). There are, however, several limitations to the Fisher-Muller model (for review see e.g. Felsenstein 1974; Barton and Charlesworth 1998; Otto 2009; Vos 2009). First, the Fisher-Muller model does not consider epistatic effects, which can cause beneficial mutations to be dependent on the genetic background they in which they appear (Felsenstein 1974; Kondrashov and Kondrashov 2001). Second, although recombination between individual genomes can combine beneficial alleles, it can also break up existing beneficial gene combinations or replace beneficial alleles with deleterious ones; these effects are often referred to as the recombinational load (Charlesworth and Charlesworth 1975; Feldman et al. 1980; Charlesworth and Barton 1996). In spite of these limitations, several experiments have shown that recombination can increase the rate of adaptation in both eukaryotic and prokaryotic organisms (Colegrave 2002; Goddard et al. 2005; Cooper 2007; Baltrus et al. 2008).

Increasing the rate of adaptation is not the only way recombination can benefit an organism. Recombination can be used to overcome Muller’s ratchet, i.e. the accumulation of deleterious mutations due to genetic drift (Muller 1964; Haigh 1978). Recombination can counter this process by replacing the deleterious allele with the original less deleterious version (known as recreating the “least-loaded class”) or by combining deleterious mutations in a single individual, which is then removed from the population by natural selection (Felsenstein 1974; Keightley and Otto 2006). Muller’s ratchet is especially problematic in small and/or clonal populations where the accumulation of deleterious mutations can lead to extinction through mutational meltdown (Lynch et al. 1993; Andersson and Hughes 1996).

Another possible effect of recombination comes from the benefits it provides in the con-
text of an ecological complex environment, where sexual species interact with pathogens and predators. In these complex networks of interactions antagonistic coevolution between parasites and their hosts is considered to drive the maintenance of sex (Hamilton et al. 1990; Agrawal 2006b). A key factor in this model, known as the Red Queen model, is the strong difference in generation time between host and parasite. The short generation time of the parasites allows them to adapt to the host. Sexual reproduction of the host can offset this high rate of adaptation by creating new allele combinations during reproduction via recombination (Hamilton et al. 1990; King et al. 2009; Morran et al. 2011).

These models on the evolution of recombination describe the current consensus on the benefits of recombination and sex. In the first case, recombination functions as an innovative factor during the process of evolution, while in the second case recombination helps preserve the wild type and is better viewed as a conservative process. The Red Queen hypothesis is a mixture of the previous two pointing to conservation of allele diversity within the host population, which allows the host population to keep up with parasite evolution by continuously recombining new allele combinations. This interpretation, which is mostly based on eukaryotic research, is often applied to prokaryotes. Although recombination in prokaryotes might well be maintained for similar reasons, alternative hypotheses cannot be excluded given that several different mechanisms of recombination exist in prokaryotes. Some of these hypotheses require a biological or organism-specific context. For this reason, I will first briefly introduce the biology of recombination in prokaryotes, and then focus more specifically on recombination in *Streptococcus pneumoniae* before continuing with prokaryote-specific hypotheses.

1.2 Recombination in bacteria

Three different mechanisms of recombination have been identified in bacteria: conjugation, transduction and transformation. The first two mechanisms, conjugation and transduction, occur via either a plasmid or phage vector respectively. For these vectors, recombination of
host DNA is a by-product of their own reproduction and infectious transfer between hosts. This should not be ignored when considering the maintenance of these two mechanisms in bacterial populations, even in cases where plasmids or phage provide direct advantages to bacterial cells (e.g. antibiotic resistance or pathogenicity islands) (Hazen et al. 2010; Ramirez-Diaz et al. 2011; Stokes and Gillings 2011). Recombination through conjugation occurs when plasmids that have integrated into the bacterial chromosome are excised from the chromosome and inadvertently capture bacterial genes, which are passaged during conjugative transfer (Thomas and Nielsen 2005). This can occur both within and between species and has enormous importance for bacterial fitness. However, as the recombination derived from this process is accidental, it is unlikely to have evolved as a mechanism of bacterial sex. Similarly, transduction relies on the bacteriophage mistakenly packaging host genetic material during lysogenic induction prior to viral transmission (Thomas and Nielsen 2005).

In contrast to transduction and conjugation, the genes regulating natural transformation, the process during which a cell takes up, incorporates and expresses ‘naked’ DNA from the environment, are encoded on the bacterial chromosome. This strongly implies that natural transformation is maintained in bacterial populations as a part of the survival strategy of the bacterium, rather than as an aspect of the infectious strategies of mobile elements. In this thesis, I have focused my efforts in understanding the evolutionary function of bacterial recombination on transformation, using the naturally transformable species S. pneumoniae as my model. Transformation is part of a considerably larger cellular physiological state called competence. In this thesis both competence and transformation are often used interchangeably. However, I will narrow my usage where the distinction is important (e.g. Chapter 2).

1.3 The competent cell state

Natural competence has evolved among all major clades of archaea and bacteria (Lorenz and Wackernagel 1994; Jonas et al. 2001). Some well known naturally competent species are Neisseria gonorrhoeae, Haemophilus influenzae, Bacillus subtilis and Streptococcus sp.
In these species, with the exception of *N. gonorrhoeae* which expresses competence continuously, competence is tightly regulated within the cell and is only activated under certain conditions (Goodman and Scocca 1988; Johnsborg et al. 2007). Following the initiation of competence, cells become capable of transformation, during which cells take up, incorporate and express DNA from the environment. DNA uptake can be indiscriminate, as in *S. pneumoniae*, or selective based on a specific uptake sequence, as in *N. gonorrhoeae* and *H. influenzae* (Sisco and Smith 1979; Goodman and Scocca 1988; Lorenz and Wackernagel 1994). After DNA uptake, the DNA can undergo two fates: it can either be degraded or it can be incorporated in the genome via homologous recombination. Homologous recombination requires that DNA contains an overlapping fragment of at least 25-100bp in length (Thomas and Nielsen 2005). This homologous region is generally relatively small compared to the total fragment size. Therefore, it allows for a certain amount of interspecific recombination depending on the sequence similarity between the host genome and donor fragment of DNA (Majewski et al. 2000). For example, *S. pneumoniae* is a notorious user of DNA of closely related species as exemplified by the case of the transformation of penicillin binding protein (PBP), which confers penicillin resistance to recipient cells (Dowson et al. 1989; Havarstein et al. 1997; Hakenbeck et al. 1999; Majewski et al. 2000; Hakenbeck et al. 2001; Chi et al. 2007). Because the genes encoding PBPs are made up of highly divergent sequence blocks, which originated from different species, they are often referred to as mosaic genes. Other examples of mosaic genes in *S. pneumoniae* are tetracycline resistance genes, fluoroquinolone resistance genes and, interestingly, the genes that regulate natural competence itself (Oggioni et al. 1996; Havarstein et al. 1997; Gonzalez et al. 1998). Mosaic structure is a common signature of recombination in all naturally transformable species.

### 1.4 Maintenance of competence in bacterial populations

Three distinct hypotheses have been proposed to explain the maintenance of competence in bacterial populations: 1) Competence is used to repair damaged regions of the chromosome, often referred to as the ‘DNA-for-repair hypothesis’; 2) Competence is used to increase the
rate of adaptation, in a manner analogous to the Fisher-Muller model described for eukaryotes, as outlined above; and 3) DNA is used as a resource, a hypothesis that is commonly designated as ‘DNA for food’. These hypotheses will each be discussed in general terms and with respect to the evolution of natural competence in *S. pneumoniae*.

### 1.4.1 Competence as a DNA repair mechanism

The DNA-for-repair hypothesis states that absorbed and incorporated DNA can be used to repair damaged DNA, especially double stranded damage, via recombinational repair (Bernstein et al. 1981). Initial experimental evidence in *B. subtilis* supported the hypothesis that competence protected UV exposed cells from DNA damage through recombinational repair (Michod et al. 1988; Wojciechowski et al. 1989). However, it seems unlikely that lesions are repaired via specific replacement from a donor strand. If DNA damage appears randomly on the genome then a damaged cell would need to fortuitously take up the single correct DNA fragment that contains the correct homologous template for repair to be successful. Experiments showed that protection from DNA damage in competent cells of *H. influenzae* did not require the presence of extracellular template covering the entire genome. Instead a small, clearly non-homologous, fragment of the genome sufficed to improve the survival of DNA-damaged cells (Mongold 1992). This result challenged the proposed template-specific mechanism of recombinational repair. Instead Mongold (1992) proposed that the act of transformation itself protected competent cells from DNA damage, rather than that transformation replaced damaged sites. Thus the specific ‘repair’ hypothesis was generalised to include repair that was not template-specific.

The DNA-uptake mechanism in *S. pneumoniae* suggests that template-nonspecific repair may occur during transformation in this species. During DNA uptake the double stranded DNA, which is bound to the outer cell membrane, is actively split into two single strands of DNA (ssDNA) fragments (Clave and Trombe 1989; Mejean and Claverys 1993). One single stranded fragment is internalised into the cell whilst the other fragment is degraded.
(Mejean and Claverys 1988; Berge et al. 2002). The internalised fragment is then protected and processed for homologous recombination by RecA, DprA and SsbB (Martin et al. 1995; Berge et al. 2003; Desai and Morrison 2006; Mortier-Barriere et al. 2007). These same proteins are known to be important for double stranded break repair (Sung et al. 2001; Cirz and Romesberg 2007). Thus the activation of these genes during competence for homologous recombination could simultaneously lead to repair of DNA damage without replacing the damaged fragment, as Mongold (1992) suggested.

At the same time, Redfield (1993) argued against the DNA-for-repair hypothesis because DNA damaging stressors did not activate competence in *B. subtilis* or *H. influenzae*. However, in *Helicobacter pylori*, another naturally competent species, genetic exchange is triggered by DNA damage (Dorer et al. 2010). Similarly, in *S. pneumoniae* low concentrations of antibiotics, including types that do and do not cause DNA damage, can induce competence (Prudhomme et al. 2006). This observation suggests that competence might not only be important for overcoming DNA damage in *S. pneumoniae*, but it might also protect cells from other forms of stress. For example, competent cells could be protected from other forms of stress because competence causes reduced growth, which could give cells extra time to repair damage (Mongold 1992). Experiments in *B. subtilis* showed that the reduced growth during competence indeed protected cells from penicillin exposure (Johnsen et al. 2009). The original DNA-for-repair hypothesis specifically predicts that competence repairs DNA damage. However, Prudhomme et al (2006) and Johnsen (2009) observed that competent cells are, respectively, affected by or protected from antibiotics that do not damage DNA. This strongly suggests that the original DNA-for-repair hypothesis is too narrow. Following these observations Claverys et al (2006) proposed that competence might function as a general stress response mechanism in *S. pneumoniae*.

### 1.4.2 Competence increases the speed of evolution

Benefits of competence are often considered to be intertwined with benefits of transforma-
ation, which naturally led to the hypothesis that competence can benefit prokaryotes in a similar way that recombination can benefit eukaryotes. However, experimental evidence for an increased rate of adaptation in naturally competent bacterial populations is conflicting. Experiments in *H. pylori* have found an increased rate of adaptation attributable to competence, whereas experiments in *Acinetobacter baylyi* showed that competence is selected against during long-term culture (Bacher et al. 2006; Baltrus et al. 2008). These contrasting results show that adaptive benefits arising from recombination in bacteria are not obvious. In earlier models Redfield showed that benefits of competence are restricted by the source of DNA (Redfield 1988; Redfield et al. 1997). Extracellular DNA for transformation generally originates from dead cells, which can be the same or other species (Thomas and Nielsen 2005). Transformation with DNA from dead cells likely means recombining with genes from low-fitness individuals, which can lead to an increased chance of recombination with deleterious mutations and thus reduced fitness as result of an increased recombinational load (Redfield 1988). However, if the fitness of the transforming cell is relatively low than the chance of introducing deleterious mutations in the genome is strongly reduced (Redfield et al. 1997). In fact, if the fitness of the recipient cell is low enough then recombination with DNA from dead cells can be beneficial in both prokaryotes (Redfield et al. 1997; Szollosi et al. 2006) and eukaryotes (Salathe et al. 2009) because it replaces deleterious alleles. A (rapidly) changing environment, which causes low-fitness immediately after environmental change, can maintain eukaryotic sex in the population (Peters and Lively 1999; Gandon and Otto 2007; Becks and Agrawal 2010). These intervals of environmental change have to be unrealistically frequent if competence is continuously active. However, if recombination is tightly regulated to function only in low-fitness situations, e.g. immediately after environmental change, then this may lead to higher fitness than if recombination is continuously active (Hadany and Beker 2003). Following this result, Hadany and Beker (2003) formulated the *Fitness Associated Recombination* (FAR) hypothesis, which predicts that competence should be beneficial in a changing environment if competence is strictly regulated to be active only in low-fitness situations. Experiments in the filamentous fungus *Aspergillus nidu-
lans have provided support for a similar hypothesis (fitness associated sex), but whether this hypothesis also applies to prokaryotic recombination remains to be tested (Hadany and Otto 2007) (Schoustra et al. 2010).

1.4.3 DNA as a resource

The observation of Redfield (1993) that DNA damage could not induce competence in *H. influenzae* and *B. subtilis* led her to conclude that DNA repair is an unlikely explanation for the evolution of natural transformation. Her alternative explanation is that DNA is used as a nutrient. It was determined that *Escherichia coli* can use DNA as a carbon source in stationary phase (Finkel and Kolter 2001); however, *E. coli* is not a naturally competent organism, although its genome retains many of the genes required for transformation. Redfield argues that if DNA is used as a nutrient, competence should be induced by nutritional signals, such as the depletion of the nucleic acids and other precursors of nucleotide biosynthesis (Redfield 2001). This prediction is experimentally supported in *H. influenzae* (MacFadyen et al. 2001), but support is not evident in other species (Redfield 2001).

The DNA for food hypothesis seems unlikely for *S. pneumoniae* for three reasons. Firstly, the DNA for food hypothesis seems quite implausible considering that one of the DNA strands is degraded during DNA uptake whilst the other strand is internalised. In gram-positive bacteria, such as *S. pneumoniae*, the degraded strand will be released in the environment and thus lost or possibly used by other cells. In gram-negative bacteria, such as *H. influenzae*, the degraded strand is released into their periplasmic space making it easier to monopolise nucleic acids if the DNA is used as a resource (Finkel and Kolter 2001; Chen and Dubnau 2004). Secondly, the DNA that is taken up in the cell is directly associated with proteins that protect it against internal nucleases and proteins that are important for recombination (Johnsborg et al. 2007). This suggests that DNA is acquired for the intact template rather than for its parts. Finally, competence in *S. pneumoniae* is induced during exponential growth, when nutrients are abundant (Tomasz 1965), and for only a very short period of time. Thus any
nutrient gains from DNA would be minimal and short-lived. For these reasons, the DNA for food hypothesis is not explored in this thesis.

1.4.4 Many hypotheses with only one winner?
The three previously described hypotheses are not mutually exclusive. Focusing on a single hypothesis might be unnecessarily constraining, as benefits of competence might be organism and/or scenario dependent. However, a returning theme can be identified in these hypotheses: stress, in the form of, for example, DNA damage, a changing environment or resource limitation. These different kinds of stresses affect an organism on different timescales, which allows for different explanations maintaining competence in a population depending on the scenario. DNA damage and resource limitation provide an acute problem to the cell, which needs a short-term solution. Changing environments provide a more long-term challenge because of a directional change in selective pressure making it important to adapt quickly and permanently to the new environment. To date no one has systematically tested these scenarios and the accompanying hypotheses in a single model organism.

1.5 Introducing the model: Streptococcus pneumoniae
To address these different hypotheses I used the model organism *S. pneumoniae*, a Gram-positive bacterium that inhabits the nasopharynx of humans and rarely of other species (Whatmore et al. 1999b). This bacterium is responsible for an enormous disease burden making it one of the leading causes of childhood mortality worldwide because of an extremely high incidence of carriage in children (Gray et al. 1980; Brueggemann et al. 2003). It causes a range of diseases, such as otitis media, sinusitis, pneumonia, bacteremia and meningitis. Despite this high disease burden, *S. pneumoniae* is perhaps better described as a commensal bacterium primarily colonising young children and adults that come into contact with these children rather than as a pathogen (Leiberman et al. 1999; Kaltoft et al. 2008). Often disease is caused as a secondary infection when the mucosal epithelium is compromised by other factors, such as a viral infection (Konradsen and Kaltoft 2002; Charalam-
bous 2007). The propensity to cause invasive disease varies greatly between pneumococcal strains. This can be attributed to variation in the polysaccharide capsule surrounding the cell, which is considered a critical virulence factor (Watson and Musher 1999; Magee and Yother 2001; Weinberger et al. 2010). The capsule is highly immunogenic and therefore is under strong selection to keep changing, for which competence is considered to play an important role (Kelly et al. 1994; Havarstein et al. 1997; Coffey et al. 1998). Similarly, treatment of invasive pneumococcal disease is potentially compromised by the ability to quickly acquire antibiotic resistance via transformation (Dowson et al. 1989; Hakenbeck et al. 2001). For these reasons, the molecular mechanisms of competence have been extensively studied (Johnsborg and Havarstein 2009).

Frederic Griffith discovered transformation in 1928 through a now classic experiment. He injected mice with a small number of viable but avirulent *S. pneumoniae* cells together with a large inoculum of heat-killed virulent *S. pneumoniae*. He found that these mice had an increased mortality compared to mice inoculated with only the avirulent *S. pneumoniae* cells. When he examined the dead mice, he found virulent *S. pneumoniae* cells in the blood of their heart. These observations together with the fact that the heat-killed inoculum did not contain viable cells led to the conclusion that the avirulent *S. pneumoniae* strain must have newly acquired the virulence factors from the dead cells (Griffith 1928). Later, Avery et al. (1944) identified a “desoxyribonucleic acid fraction from pneumococcus type III” as the “transforming principle” of *S. pneumoniae*. They proposed that DNA should be regarded as “...functionally active in determining the biochemical activities and specific characteristics of pneumococcol cells” (Avery et al. 1944). They later confirmed this hypothesis when they found that DNase was “...capable of destroying irreversibly the Type III transforming substance...” (McCarty and Avery 1946). Avery’s work on DNA and transformation in *S. pneumoniae* was the first big step towards identification of DNA as the genetic material, which was later confirmed by Hershey and Chase (1952).
1.6 The molecular mechanism of competence

Since the discoveries by these pioneers the molecular pathways that lead to the competent state in \textit{S. pneumoniae} have been elucidated in great detail. Competence in \textit{S. pneumoniae} is induced through a density dependent signal peptide, Competence Stimulating Peptide (CSP) (Tomasz 1965; Tomasz and Mosser 1966; Yother et al. 1986; Havarstein et al. 1995). CSP is a self-induced signal functioning in a quorum dependent manner (Pestova et al. 1996; Lee and Morrison 1999) (But see Yang et al. (2010) for a more nuanced view). The genes responsible for the induction of competence \textit{comABCDE} are transcribed constitutively at a low level (Pestova et al. 1996). These genes, together with genes such as \textit{comX}, \textit{comW} and \textit{comM}, are often referred to as early com genes because they form the core of the quorum sensing apparatus for the induction of competence (Luo and Morrison 2003; Luo et al. 2004; Peterson et al. 2004; Sung and Morrison 2005; Havarstein et al. 2006).

Briefly, arbitrarily starting at \textit{comC} which encodes the secreted peptide signal CSP, competence is induced in the following way (Fig. 1.1): The 43 amino acid full-length peptide

![Figure 1.1. Schematic overview of early com gene regulation. ComC (pre-CSP) is exported by the ABC-transporter ComAB. Environmental CSP is detected by the histidine kinase ComD, which then phosphorylates ComE. Phosphorylated ComE then will promote transcription of ComAB and ComCDE and sigma factor ComX. ComX will then induce the late com genes. (This figure is based on figure 1 of Karlsson et al. 2007).](image-url)
ComC is transported out of the cell after cleavage of the amino-terminus by its specific ABC-transporter ComAB, resulting in the 17 amino acid peptide called CSP (Havarstein et al. 1995; Hui et al. 1995). ComAB is the bottleneck of the quorum sensing apparatus. Martin et al. (2000) found that overexpression of ComAB resulted in overexpression of ComCDE indicating that the export, rather than production, of CSP is the limiting factor in the induction of competence. Extracellular CSP is detected by the membrane-embedded histidine kinase ComD (Pestova et al. 1996). Detection of CSP by ComD results in the autophosphorylation of ComE. Phosphorylated ComE not only promotes transcription of the comAB and comCDE genes causing a positive feedback to CSP production, but also induces transcription of the sigma-factor ComX, which induces the majority of genes regulated by competence and is a part of the off-switch mechanism (Lee and Morrison 1999; Ween et al. 1999; Martin et al. 2000; Claverys and Havarstein 2002; Luo and Morrison 2003; Peterson et al. 2004; Karlsson et al. 2007; Piotrowski et al. 2009).

After induction of competence by CSP, *S. pneumoniae* enters a competent state for a period of about 30 minutes, with a peak after 20 minutes, during which time more than 100 genes are activated and inactivated (Peterson et al. 2000; Peterson et al. 2004). During this period transformation can occur. However many of the activated genes are not part of the transformation machinery, which is a strong suggestion that benefits of competence are not automatically equivalent to benefits of transformation.

The function of most of the genes activated during competence is unknown. Most known genes are associated with transformation. Other competence associated genes regulate competence-induced cell lysis, which causes cell lysis in a fraction of the non-competent population (Steinmoen et al. 2002; Steinmoen et al. 2003; Guiral et al. 2005; Kausmally et al. 2005; Havarstein et al. 2006). Competence-induced cell lysis has been suggested as the source of DNA for transformation (Steinmoen et al. 2002). Consistent with this possibility, experiments have shown that competence induced cell lysis increases intra- and interspecies gene exchange (Johnsborg et al. 2008), although a study on the kinetics of DNA release showed
that DNA is released long after competence has ended (Moscoso and Claverys 2004).

This extensive body of research offers several advantages to using *S. pneumoniae* as a model to study and understand competence as compared to other naturally competent bacteria, such as *B. subtilis, H. influenzae* or *H. pylori*. Firstly, existing strains containing knockout mutations in early competence genes, together with their isogenic parent strains, can be used in experiments directly comparing differences between the presence and absence of competence (e.g. chapter 3). Secondly, competence stimulating peptide (CSP) can be commercially synthesised because the sequence is known. Synthesised CSP is used in experiments where precise induction of competence is essential (e.g. chapter 2), which is not available in other systems. Thirdly, the wide variety of wild type isolates with differential rates of transformation can be used for experiments on the population dynamics of competence. Although, this specific subject is not addressed in this thesis.

1.7 Project aims and thesis structure

The aim of this thesis is to investigate factors that favour the maintenance of competence in *S. pneumoniae*. In this thesis I will show that the two tested hypotheses, DNA-for-repair and transformation increases the rate of adaptation, are influenced by a single common factor: stress. In addition, I will show that competence and stress have a strong influence on the long-term genomic evolution of *S. pneumoniae*. A final chapter, unrelated to competence evolution, tests the idea that the evolution of *S. pneumoniae* virulence correlates with its ability to survive on foreign surfaces.

In chapter 2, which has been published in Evolution (Engelmoer and Rozen 2011), I investigate short-term benefits of competence using the DNA-for-repair hypothesis as a starting point. I confirm that competence protects against DNA damage by mitomycin C, but also against antibiotics that do not cause DNA damage. The mechanism underlying this protection is different depending on the type of stress encountered by *S. pneumoniae*. DNA damage requires transformation, whereas the effects of stress that does not cause DNA damage...
is reduced by the activation of competence, but not of transformation. These results provide the first evidence that competence is maintained in bacterial population as a mechanism to increase stress responsiveness.

In chapter 3 I investigate the long-term cost and benefits of competence using a long-term evolutionary experiment that tests both the Fisher-Muller model as well as the fitness associated stress hypothesis (Fisher 1930; Muller 1932; Hadany and Beker 2003; Otto 2009). In contrast to predictions of the Fisher-Muller model I show a clear cost of competence in populations evolved in a benign environment. These costs are mitigated by episodes of mild stress, which provides the first experimental evidence for the fitness associated recombination hypothesis in prokaryotes. Whole genome re-sequencing of the evolved populations shows that competent populations accumulated non-synonymous mutations at a lower rate. This indicates that competence functions as a conserving factor of the genome by reducing mutational load, rather than an innovating a factor, which combines beneficial mutations. Finally, I show evidence that mutator phenotypes are associated with the absence of competence (Tenaillon et al. 2000). These results support our short-term experiments that competence is maintained in bacterial populations as a stress response mechanism.

In chapter 4 I explore parallel phenotypic and genotypic changes in our long-term populations, and show several examples of adaptation to the evolutionary environment. I identify several parallel changes, which are found in all evolved lines. I also identify genotypic changes that are environment and treatment specific, showing that selective pressures differed between competent and non-competent populations as well as between the benign and periodic stress environment. I show that parallel genetic changes at the base-pair level are rare, whilst parallel genetic changes per locus and within functional groups are much more common.

In chapter 5, the final experimental chapter, I use the variation in virulence amongst wild type
isolates to test the prediction that \textit{S. pneumoniae} has evolved a ‘sit-and-wait’ transmission strategy. As predicted by the sit-and-wait hypothesis, I show a positive correlation, although weak, between virulence and durability outside of the host. At the same time, I show a negative relationship between durability and the chance that a strain is found in disease. These results are qualitatively consistent with the theoretical observation that mixed transmission strategies can evolve and coexist in pathogen populations (Roche et al. 2011).

In chapter 6 I finish this thesis with a summary addressing the factors that are important for the maintenance of competence in bacterial populations, which concludes the aim of this thesis. In addition, I propose several avenues of further research on the evolution of competence.

1.8 References
Avery, O. T., C. M. MacLeod, and M. McCarty. 1944. Studies on the chemical nature of the substance inducing transformation of pneumococcol types I. Induction of transformation by a deoxyribonucleic acid fraction isolated from pneumococcus type III. J. Exp. Med. 79:137-158.


Cooper, T. F. 2007. Recombination speeds adaptation by reducing competition between ben-
beneficial mutations in *Escherichia coli*. Genetics 172:2093-2100.
Desai, B. V., and D. A. Morrison. 2006. An unstable competence-induced protein, CoiA,
promotes processing of donor DNA after uptake during genetic transformation in
Dorer, M. S., J. Fero, and N. R. Salama. 2010. DNA damage triggers genetic exchange in
protein genes in penicillin-resistant clinical isolates of *Streptococcus pneumoniae*.
*Streptococcus pneumoniae*. Evolution 65:3475-3485.
Finkel, S. E., and R. Kolter. 2001. DNA as a nutrient: novel role for bacterial competence
ford.
and adaptation in large asexual populations. Genetics 180:2163-2173.
abiotic or coevolutionary fluctuations in epistasis. Genetics 175:1835-1853.
asexual population. Genetica 102-103:127-144.


Hershey, A. D., and M. Chase. 1952. Independent functions of viral protein and nucleic acid


munol. 9:358-365.


Prudhomme, M., L. Attaiech, G. Sanchez, B. Martin, and J. P. Claverys. 2006. Antibiotic


Szollosi, G. J., I. Derenyi, and T. Vellai. 2006. The maintenance of sex in bacteria is ensured
by its potential to reload genes. Genetics 174:2173-2180.
Chapter 2

Competence increases survival during stress in *Streptococcus pneumoniae*

Daniel J.P. Engelmoer & Daniel E. Rozen

*Evolution*, 2011, **65**(12):3475-3485

2.1 Abstract

Horizontal gene transfer mediated by transformation is of central importance in bacterial evolution. However, numerous questions remain about the maintenance of processes that underlie transformation. Most hypotheses for the benefits of transformation focus on what bacteria might do with DNA, but ignore the important fact that transformation is subsumed within the broader process of competence. Accordingly, the apparent benefits of transformation might rely less on recombination than on other potential benefits associated with the broader suite of traits regulated by competence. We examined the importance of this distinction in the naturally competent species *Streptococcus pneumoniae*, focussing specifically on predictions of the DNA-for-repair hypothesis. We confirm earlier results in other naturally competent species that transformation protects against DNA-damaging stress. In addition, we show that the stress-protection extends to non-DNA-damaging stress. More important, we find that for some forms of stress transformation is not required for cells to benefit from the induction of competence. This rejects the narrowly defined DNA-for-repair hypotheses and provides the first support for Claverys’ hypothesis that competence, but not necessarily transformation, may act as a general process to relieve stress. Our results highlight the need to distinguish benefits of transformation from broader benefits of competence that do not rely on DNA uptake and recombination.
2.2 Introduction

Horizontal gene transfer is increasingly recognized as a central component of bacterial evolution, leading to the emergence of novel virulence traits and facilitating the spread of antibiotic resistance (Hakenbeck et al. 2001; Fraser et al. 2007; Shapiro et al. 2009; Didelot and Maiden 2010). Despite this realization, numerous questions remain about the evolution and maintenance of the processes that underlie the widespread occurrence of inter- and intraspecific recombination. There are three different mechanisms through which bacteria can recombine DNA: conjugation, the exchange of genetic material by plasmids through direct cell-cell contact; transduction, the transmission of genetic material through a phage intermediate; and transformation, the uptake and recombination of extracellular genetic material from a variety of donor sources (for review e.g. Thomas and Nielsen 2005; Narra and Ochman 2006). Both conjugation and transduction involve the transfer of extra-chromosomal elements, while the genes encoding transformation are located on the bacterial chromosome (Thomas and Nielsen 2005). In contrast to recombination resulting from conjugation or transduction, where chromosomal gene transfer is likely a by-product of processes that evolved to modify the infectious transfer of these agents, recombination mediated by transformation is widely considered to have evolved exclusively for this purpose, although alternative explanations exist (as discussed below). For this reason, transformation is often seen as the bacterial analogue to eukaryotic sex. As with eukaryotic sex, there is currently little consensus regarding the factors that act currently to maintain transformation (e.g. Otto 2009). There are three prevailing hypotheses for the benefits of DNA uptake and recombination in bacteria (for review see e.g. Michod et al. 2008; Vos 2009): 1) DNA is used as a resource; 2) DNA is used to increase the efficiency of natural selection; and 3) DNA is used as a template to repair damaged chromosomal DNA. Here we examine evidence for the third hypothesis in the naturally transformable species Streptococcus pneumoniae.

Bacterial cells in natural environments experience DNA-damaging stress from a variety of sources. Bernstein et al. (1981) proposed that homologous extracellular DNA, mainly from
conspecific individuals, taken up and recombined could be used to repair damaged genomic DNA. Experimental evidence in *Bacillus subtilis* suggested that repair of DNA in competent cells exposed to UV-light occurred through transformation mediated recombination, thereby increasing cellular survival (Michod et al. 1988; Wojciechowski et al. 1989). However, the generality of these results were called into question by experiments in *Haemophilus influenzae* showing that transformation with just a small fraction of conspecific genomic DNA, amounting to less than 1% of the damaged recipient genome and thus unable to replace damaged DNA on a genomic scale, was sufficient to increase survival of cells during UV-exposure (Mongold 1992). Additionally, neither UV-light nor the DNA-damaging antibiotic mitomycin C was able to induce transformation in either *H. influenzae* or *B. subtilis* (Redfield 1993). Both results implied that transformation in these species is unlikely to be maintained if its primary role was a process to repair DNA damage (Redfield 2001).

More recently, Claverys et al (2006) proposed a broader hypothesis than Bernstein’s original one, suggesting that competence is used to respond to stress generally rather than just to DNA damage. This idea adds considerably to the debate for two reasons: First, by considering that the benefits of competence may extend beyond a single source of stress this hypothesis is quite general. Of course, stress comes to cells in many forms, and it is unlikely that each stress induces its own dedicated response mechanism. Second, this hypothesis makes the important distinction, lacking in all previous explanations for the benefits of transformation, that “competence” and “transformation” are not equivalent. Competence is the regulated state that bacteria enter and during which transformation occurs. In addition to transformation, however, competence regulates a variety of genes and functions. In *S. pneumoniae*, for example, nearly 150 genes are regulated by competence, yet only a small fraction of these are involved in transformation (Peterson et al. 2004). The current debate is largely framed by the question: “What do bacteria do with DNA?” However, because transformation is subsumed within the broader process of competence, the appropriate question should be more directed towards benefits of competence broadly, and distinguishing which, if any, benefits are as-
Evidence suggesting the importance of this more expansive view comes from observations that transcriptional responses of many classical stress response genes in two naturally transforming bacterial species, *S. pneumoniae* and *B. subtilis*, are altered by competence (Martin et al. 1995; Peterson et al. 2000). Additionally, it was found that competent growth-arrested populations of *B. subtilis* were better protected from penicillin than their growing non-competent counterparts (Johnsen et al. 2009), and that competence is induced by sub-Minimum Inhibitory Concentration (MIC) concentrations of DNA-damaging and non-DNA-damaging antibiotics in both *S. pneumoniae*, *Helicobacter pylori* and *Legionella pneumophila* (Prudhomme et al. 2006; Charpentier et al. 2011; Dorer et al. 2010). Finally, Mongold’s (1992) results showing that protection against UV irradiation in *H. influenzae* could be derived from a small fraction of total genomic DNA implied that the induction of competence itself, rather than transformation *per se*, in DNA damaged cells could be responsible for apparent benefits of recombination. Despite these suggestive results, there are currently no studies specifically directed towards distinguishing fitness benefits of competence versus transformation under various sources of stress. This is the aim of the present work using *S. pneumoniae* as our test organism.

* S. pneumoniae is a naturally competent Gram-positive bacterium long studied as a model for understanding the dynamics and mechanisms of natural transformation (Griffith 1928; Tomasz and Mosser 1966; Havarstein et al. 1995; Johnsborg et al. 2007). As a result of this long history of research the molecular mechanisms of competence induction and transformation are known in great detail (for review e.g. (Dubnau 1999; Claverys and Havarstein 2002; Johnsborg et al. 2007); however, its benefits are surprisingly unclear. Competence arbitrarily ‘begins’ with the secretion of a density dependent extracellular signal, called Competence Stimulating Peptide (CSP) (Tomasz 1965; Havarstein et al. 1995). CSP binds to its dedicated membrane-bound receptor ComD, which in turn phosphorylates its response-regulator, ComE, leading ultimately to the autoinduction of CSP as well as stimulating the production
of the sigma factor ComX (Havarstein et al. 1996; Pestova et al. 1996; Ween et al. 1999; Martin et al. 2000; Luo and Morrison 2003). This latter step, initiating the ‘competent state’, causes the induction of a large collection of genes comprising a diversity of functions. These include genes required for transformation, but also genes required for activities unrelated to transformation, such as bacteriocin production and immunity and several with unknown functions (Peterson et al. 2004). Any of these functions may potentially confer benefits to cells when confronted with acute or chronic stress. The activation of competence-associated functions by CSP induction occurs rapidly, peaking at approximately 30 minutes and then quickly attenuating; reactivation then requires an interval of 1-2 generations (Morrison 1997; Peterson et al. 2004). 

By modifying genes encoding the signal peptide and its receptor it was possible to test predictions of the DNA for repair/stress hypothesis in an environment where the activation of competence was tightly controlled through the addition of synthetic CSP. Using this approach, we investigated predictions of both the strict DNA-for-repair hypothesis by Bernstein et al (1981) and the broader stress response hypothesis of Claverys et al (2006). In short, we find that competence increases the survival of *S. pneumoniae* when exposed to a DNA-damaging antibiotic, confirming earlier findings in *B. subtilis* and *H. influenzae*, as well as to agents that do not damage DNA. Most interestingly, we find that transformation is essential for benefits observed in the presence of DNA damage, while benefits associated with non-DNA-damaging antibiotics only require the activation of competence but do not require DNA uptake. Our results provide support for the hypothesis of Claverys et al (2006) and indicate a critical role for stress responsiveness in the competence regulon of *S. pneumoniae*.

### 2.3 Material & Methods

#### 2.3.1 Strains and culture conditions

*S. pneumoniae* strains were grown in CTM pH 6.8 (Complete Transformation Medium,
which per litre contains: 30g tryptic soy broth, Oxoid Ltd, and 1g yeast extract, Melford Laboratories Ltd, Ipswich, United Kingdom) and starter cultures were stored at -80°C in the same medium supplemented with 30% v/v glycerol.

Strains used in this study are given in table 2.1. Owing to the deletion of the gene encoding the CSP precursor (Iannelli et al. 2005), comC, strain FP5 does not become spontaneously competent, but can be induced to competence with the exogenous addition of 100ng/mL synthetic CSP. By contrast, strain FP48 is totally insensitive to CSP as a result of a deletion of the CSP receptor comD, and is therefore unable to be either spontaneously or artificially induced to become competent.

Table 2.1. Details of strains used in this study.

<table>
<thead>
<tr>
<th>strain</th>
<th>description</th>
<th>reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rx1</td>
<td>Lab strain; unencapsulated derivative of D39</td>
<td></td>
</tr>
<tr>
<td>FP5</td>
<td>ΔcomC of Rx1; CSP production deficient strain</td>
<td>Iannelli et al. 2005</td>
</tr>
<tr>
<td>FP48</td>
<td>ΔcomD of Rx1; CSP sensor deficient strain</td>
<td>Oggioni et al 2004</td>
</tr>
</tbody>
</table>

Before the start of an experiment, an aliquot of the relevant strain was taken from the -80°C freezer and grown in CTM pH 6.8 at 37°C + 5% CO₂ until the cells reached an OD₆₀₀ = 0.30, corresponding to a cell density of roughly 3x10⁸ CFU/mL. The culture was then diluted ten-fold into the experimental environment (CTM pH 7.8) for preconditioning and grown again to an OD₆₀₀ of 0.30. This additional period of growth was used in order to allow cells to fully acclimatise to the experimental environment.

### 2.3.2 Competence induction after or before antibiotic stress

The effects of competence on recovery of *S. pneumoniae* exposed to antibiotic induced stress were assessed by applying stress either before or after the induction of competence. Our experimental design is summarized in figure 2.1 and given in detail below. Our first approach was to determine the effects of competence on cell survival when it was induced immediately following stress. Preconditioned FP5 was diluted ten-fold in pre-warmed CTM pH 7.8
and then divided into 300µL aliquots in 1.5mL eppendorf tubes. Half of these aliquots were exposed to 100ng/mL mitomycin C. After a one hour incubation at 37°C + 5% CO₂, all populations were washed twice in pre-warmed CTM pH 7.8 to remove the mitomycin C. Next, competence was induced in half of the drug exposed and half of the unexposed populations with 100ng/mL synthetic CSP. Simultaneously, 1µg/mL of *S. pneumoniae* genomic DNA, which serves as a template for recombination, was added to all populations. The concentrations of CSP and DNA used result in high levels of transformation. After 30 min. incubation at 37°C + 5% CO₂ 1µL of DNase1 (2,000U/mL) was added to remove DNA. Finally, the aliquots from all treatments were incubated for another 1.5 hours at 37°C + 5% CO₂ to allow recovery after stress exposure. After appropriate dilution of cell aliquots, samples were plated on TS blood agar plates (Tryptic soy agar, LabM Ltd, with 3% horse blood, Oxoid Ltd), and incubated at 37°C + 5% CO₂ for two days before the number of colony forming units (CFU) was determined. The rate of recovery was determined by taking the log of the ratio of CFU at the end of the experiment over CFU directly after drug removal. The experiments were carried out with 5-fold replication.

**Figure 2.1.** Schematic picture of the experimental setup used for stress before competence. (A) and stress after competence (B). In further experiments a variation of experimental setup B is used to determine the effects of CSP and DNA independently.
In our second approach, competence was induced before *S. pneumoniae* was exposed to antibiotic stress. After different periods of time following artificial induction of competence with exogenous CSP, populations of FP5 were exposed to mitomycin C, streptomycin or kanamycin. Briefly, preconditioned, FP5 was diluted ten-fold in pre-warmed CTM pH 7.8 and then divided in 300µL aliquots. To half of the populations 1µg/mL genomic DNA and 100ng/mL synthetic CSP was added. After 0, 20, 40 or 60 min. of incubation at 37°C + 5% CO$_2$, 100ng/ml mitomycin C (n = 6 for each treatment), 100µg/mL streptomycin (n = 6 for each treatment) or 100µg/mL kanamycin (n = 12 for each treatment) was added. Following 30 minutes of drug exposure, populations were washed twice in pre-warmed CTM pH 7.8 and plated after appropriate dilution on TS blood agar plates. Rates of either decline or growth were determined by taking the log of the ratio of CFU after drug exposure over CFU before adding the drug.

### 2.3.3 Requirements for increased survival

To determine whether the benefits of competence were driven by transformation specifically or by competence more generally, FP5 was treated with or without combinations of DNA and CSP and then exposed to mitomycin C or streptomycin. Similar to the previous experiments, preconditioned populations of FP5 were diluted ten-fold in pre-warmed CTM pH 7.8 and divided in 300µL aliquots. These aliquots were exposed to one of four different treatments: (1) +DNA, +CSP; (2) -DNA, -CSP; (3) +DNA, -CSP; or (4) +DNA, +CSP, +1µL DNase1 (2,000U/mL). The concentrations of added genomic DNA and CSP were 1µg/mL and 100ng/mL respectively. At the start and after either 20 minutes for streptomycin or 40 minutes for mitomycin C (the time period corresponding the greatest benefit seen in earlier experiments), either 100µg/mL streptomycin (n=9 for each treatment) or 100ng/mL mitomycin (n=9 for each treatment) was added to each aliquot. After 30 minutes drug exposure populations were washed twice with warm CTM pH 7.8 and plated onto blood agar plates. CFU’s were determined after two days incubation at 37°C + 5% CO$_2$.

### 2.3.4 The duration of benefits of competence
Putative short-term benefits of competence in the presence of antibiotic stress were assessed using both growth curves and direct competitive fitness assays. Growth rates were estimated for FP5 in the presence or absence of 5ng/mL mitomycin C or 10µg/mL streptomycin (both sub-MIC concentrations) as well as 100ng/ml CSP and 1µg/mL DNA (only added to the mitomycin C treatment) (n=4 for each treatment). OD<sub>600</sub> in CTM pH 7.8 was measured every 5 minutes for 24 hours in a 96-well plate using a plate reader at 37°C with continuous shaking. Raw OD values were normalised to a blank well and Ln transformed before analysis.

The relative fitness of FP5 and FP48 in the presence and absence of antibiotic stress was determined following the protocol given in Yang et al (2010). Briefly, preconditioned cells grown to an OD<sub>600</sub> = 0.30 were diluted 100-fold and mixed at a 1:1 ratio in 2ml CTM pH 7.8. Stressed populations were exposed to 5ng/mL mitomycin C or 10µg/mL streptomycin, while control cultures received no antibiotic. Additionally, half of the populations were made competent by adding 100ng/mL CSP at the initiation of the experimental mixture. No exogenous DNA was added in this experiment. However, DNA may have been supplied from cells that either lysed naturally due to drug exposure or from cells that fell victim to competence-induced lysis (in this case FP5 killing FP48) (Steinmoen et al. 2002). Cultures were incubated until the populations reached a density of OD<sub>600</sub> = 0.300, which, in general, took about 14 hours for populations exposed to antibiotics and 5 hours for the populations without antibiotics. Samples of the mixed population were plated to determine the number of CFU of each strain before stress and competence induction and after they reached the final density. Relative fitness was estimated as the ratio of each strain’s Malthusian parameter, calculated from the density of each strain at the start and end of competition, following Lenski et al (1991).
2.4 Results

2.4.1 Competence induction following antibiotic stress

Following the experiments of Michod et al (1988), we initially exposed cells to antibiotic stress, in the form of the DNA--damaging antibiotic mitomycin C, before adding CSP and DNA to induce competence. We predicted that transformed cells would exhibit increased growth/survival during recovery, if competence improved the ability for cells to repair mitomycin C-induced DNA lesions. Consistent with the expectation that mitomycin C damages cells either genetically or physiologically and therefore reduces growth after exposure, exposed cells grew significantly slower during drug recovery than unexposed cells (Fig. 2.2A; one-tailed t-test with unequal variances: $t_{16.799} = -2.097$, $p = 0.026$). However, the growth rate of competent populations of *S. pneumoniae* during the recovery period did not significantly differ from non-competent populations (Fig. 2.2A; two-tailed t-test: $t_8 = 0.185$, $p = 0.857$).

One possible interpretation of this negative result is that the DNA-for-repair hypothesis is not valid for *S. pneumoniae* because in this species competence serves an alternative function. Alternatively, it might be that the temporal order of competence and damage, together with the kinetics of competence induction, influence the degree to which competent cells may benefit versus their non-competent counterparts. As noted by Mongold (1992), the kinetics of competence induction in the model organism used by Michod et al. (1988), *B. subtilis*, ensured that cells had entered a competent state during the entire experiment. That is, cells were competent when they were exposed to stress and this could underlie the beneficial effect of competence, and may also explain our inability to detect benefits to competent pneumococci that were exposed to mitomycin C prior to competence induction.

2.4.2 Competence induction before or during antibiotic stress

In contrast to our previous experiments, we now induced competence in *S. pneumoniae* during or before exposure to mitomycin C. At various times following induction, cells were exposed to drug and their rates of decline estimated. Under these conditions, competent populations exposed to lethal levels of mitomycin C are killed significantly less than non-
Competence increases survival during stress

Figure 2.2. Growth or death rate of FP5 exposed to various antibiotics before or after competence was induced. A) Growth rate of FP5 exposed to mitomycin C (black gradient) or unstressed (light gray bars) before competence was induced; B) Death rate of FP5 exposed to mitomycin C after varying periods in the presence (dark gray bars) or absence (white bars) of CSP and DNA; C) Death rate of FP5 exposed to streptomycin after varying periods in the presence (dark gray bars) or absence (white bars) of CSP and DNA; D) Growth rate of FP5 exposed to sub-MIC concentrations of kanamycin after varying periods in the presence (dark gray bars) or absence (white bars) of CSP and DNA. Error bars indicate 95% CI. * Indicates significant difference (p<0.05) in growth of death rate between presence or absence of CSP and DNA based on Bonferroni post-hoc test.

competent populations (Fig. 2.2B; ANOVA: $F_{1,45} = 4.614$, $p = 0.038$). Post-hoc analysis of each time point shows that the difference in decline between competent and non-competent populations is only significant after both 40 and 60 minutes of competence. After 40 minutes, competent populations decline significantly less than non-competent populations (Fig. 2.2B; two-tailed t-test: $t_{10} = 4.943$, $p= 0.001$). Interestingly, after 60 minutes, the benefit of competence persists, but also now appears in uninduced cells (Fig. 2B). This suggests that this strain can also induce a CSP-independent protective response after 60 minutes of drug exposure. Notably, this response occurs both later and to a lesser degree than in competent cells. These results thus reveal that, on average, competent populations are better protected from DNA-damaging stress than non-competent populations.
To assess if competence induced protection was limited to DNA-damaging agents, we next quantified protection against two other antibiotics, streptomycin and kanamycin, both protein synthesis inhibitors. As for mitomycin C, we found that pneumococcal competence protects cells against both these agents. Competent populations exposed to streptomycin showed a significantly lower death rate than non-competent populations (Fig. 2.2C; ANOVA: $F_{1,46} = 12.141, p = 0.001$). This lower death rate for competent cells was only apparent after 20 minutes of competence before the exposure to the antibiotic (Fig. 2.2C; two-tailed t-test: $t_{9} = 2.402, p = 0.040$), and eroded thereafter. For experiments with kanamycin we used a sub-MIC concentration of drug to determine if protection extended to growth during stress in addition to lethal stress. As with streptomycin, we found a significantly higher growth rate of competent cells after 20 minutes of competence before exposure to kanamycin (Fig. 2.2D; two-tailed t-test: $t_{20} = 2.451, p = 0.024$). Together, the results from these three antibiotics show that the benefits of competence include, but extend beyond DNA-damaging agents. Consistent with Mongold (1992), this suggests that the mere activation of competence, rather than transformation, might be sufficient to cause the positive effect.

### 2.4.3 Benefits of competence versus transformation

In order to distinguish the benefits of competence, broadly, from the more narrow benefits of transformation, we conducted experiments similar to those outlined above, but with one critical difference. Rather than exposing all cells to CSP and DNA simultaneously, which has the effect of inducing competence and providing substrate for recombination by transformation, we exposed cells to either one or both of these products independently. Unsurprisingly, we find a significant difference in death rate between treatments if stress is induced at the time with the largest difference between competent and non-competent survival as determined in the previous experiment (mitomycin C: ANOVA: $F_{3,33} = 7.782, p = 0.001$; streptomycin: ANOVA: $F_{3,34} = 6.364, p = 0.002$). As shown earlier, when both CSP and DNA are supplied together, cells die less than in the case where neither is supplied (Fig. 2.3, table 2.2). More importantly, for mitomycin C we find that increased survival during drug exposure is only
Figure 2.3. Death rate of FP5 exposed to mitomycin C or streptomycin in the presence of independent combinations of CSP and DNA. FP5 exposed to mitomycin C (gray bars) only shows a significant reduced death rate in the presence of CSP and DNA. FP5 exposed to streptomycin (white bars) shows a significant reduced death rate in the presence of CSP while the presence or absence of DNA has no influence on the death rate. The letters a and b indicate significant different groups within an antibiotic treatment when compared to the negative control. Error bars indicate 95% CI.

Table 2.2. Bonferroni post-hoc test results for the death rate at both antibiotic stresses of each treatment. The displayed p-values are the result of tests between the negative control, -CSP –DNA, and each treatment taken from the post-hoc results based on the full model.

<table>
<thead>
<tr>
<th></th>
<th>mitomycin C</th>
<th>streptomycin</th>
</tr>
</thead>
<tbody>
<tr>
<td>treatment</td>
<td>p</td>
<td>treatment</td>
</tr>
<tr>
<td>+CSP +DNA</td>
<td>0.022</td>
<td>+CSP +DNA</td>
</tr>
<tr>
<td>+CSP -DNA</td>
<td>0.355</td>
<td>+CSP -DNA</td>
</tr>
<tr>
<td>-CSP +DNA</td>
<td>1.000</td>
<td>-CSP +DNA</td>
</tr>
</tbody>
</table>
seen when cells are provided with both DNA and CSP (Fig. 2.3, table 2.2), but not when cells are provided with either on their own. This indicates that uptake of DNA, but not necessarily recombination with damaged parts of the genome, is necessary for the observed cellular benefit. By contrast, for cells exposed to streptomycin, only CSP is required to decrease their rate of decline (Fig. 2.3, table 2.2), but DNA is not. In other words, for streptomycin, which does not damage DNA, the activation of the competent state is sufficient to increase survival, but transformation is not required.

### 2.4.4 The duration of benefits of competence

To quantify the longer-term benefits over time of competence during antibiotic stress we first determined the growth of competence induced and uninduced populations of FP5 (competent in the presence of CSP) in the presence and absence of sub-MIC concentrations of antibiotic. In contrast to earlier experiments, CSP/DNA was provided to cells at the same time as drug exposure. As expected, there are no differences between the growth rates of induced and uninduced FP5 in the absence of antibiotics (Fig. 2.4A). This implies that competence is not costly, at least during the period of the experiment. Similarly, there is no added benefit to induced cells when grown in the presence of sub-MIC concentration of mitomycin C (Fig. 2.4B). This same result was observed over a range of sub-MIC concentrations of this drug (data not shown).

In contrast with mitomycin C, competence induced cells grown in sub-MIC concentrations of streptomycin grow significantly faster than uninduced cells (Fig. 2.4C and 2.4D). Specifically, competent cells continue growing, albeit at a slower rate than non-drug-exposed cells, while the growth of their non-competent counterparts is arrested. This difference persists for nearly 8 hours, considerably longer than the peak benefit 20 minutes past induction noted above with much higher concentrations of drug (Fig. 2.2C). Consistent with these growth results, in direct head-to-head competitive assays FP5 and FP48 (a competence-uninducible mutant) have equal fitness in the absence of drug (Fig. 2.5, one-sample t-test vs 1: t₁₀ = -
Figure 2.4. Mean growth curves of FP5 (black lines) in the presence or absence of CSP and DNA while exposed to sub-MIC concentrations of varying antibiotics. Solid lines are populations with CSP and DNA, while dashed lines is where CSP and DNA were absent. A) No antibiotic. B) 5ng/ml mitomycin C, with 95% confidence interval for the populations exposed to DNA and CSP C) 10µg/ml streptomycin D) close up of C with 95% confidence intervals for the populations exposed to DNA and CSP.
Competence increases survival during stress

Chapter 2

0.372; p = 0.718), whether induced or not. By contrast, FP5 has a significantly higher fitness than FP48 in the presence of streptomycin (Fig. 2.5, one-sample t-test vs 1: $t_8 = 5.805; p < 0.0001$). Importantly, we observe a significantly increased benefit that is attributable to CSP induction (Fig. 2.5, two-tailed t-test: $t_r = 3.437, p = 0.011$). While we do observe a direct fitness benefit for FP5 in the presence of mitomycin C (Fig. 2.5, one-sample t-test vs 1: $t_{11} = 5.161; p < 0.0001$), there is no additional benefit to cells as a consequence of competence induction (Fig. 2.5, two-tailed t-test: $t_{10} = -0.184, p = 0.858$). We discuss a possible mechanism underlying the general advantage to FP5 in the presence of both drug stresses in the discussion.

![Figure 2.5. Mean relative fitness of FP5 over FP48 measured during exposure to various antibiotics. Populations were either treated with CSP (white bars) or without CSP (grey bars). Error bars indicate 95% CI. * Indicates significantly different mean relative fitness between CSP treatments within an antibiotic treatment.](image)

2.5 Discussion

The advent of high-throughput sequencing and typing methods has revealed the extent to which horizontal transfer has influenced the structure of bacterial genomes (Fraser et al. 2007; Didelot and Maiden 2010; Hiller et al. 2010; Croucher et al. 2011). Transformation is widely recognized as a critical determinant of microbial evolution and population struc-
tured. However, the factors that select for the maintenance of transformation are still a matter of debate. There are three prevailing hypotheses for the potential benefits of DNA uptake associated with transformation: DNA is used as a resource, DNA is used to increase the efficiency of adaptation, and DNA is used to facilitate repair of genomic damage. Here we have focused on testing predictions of the last hypothesis using the naturally transformable species *S. pneumoniae*. Consistent with predictions and previous tests of the DNA-for-repair hypothesis (Michod et al. 1988; Wojciechowski et al. 1989; Mongold 1992), we find that competent cells of *S. pneumoniae* are better able to withstand DNA-damaging stress than non-competent cells. A natural conclusion from these results is that transformation provides short-term benefits by facilitating the repair of damaged DNA in *S. pneumoniae*. However, following Claverys et al. (2006) we believe that this conclusion would be premature (if still partly correct). First, this conclusion conflates the fact that transformation, the process for uptake and incorporation of DNA, and competence, the broader process regulating transformation, are not equivalent. In order to conclude that transformation *per se* has direct benefits, these must be distinguished from the broader benefits of competence. Second, this conclusion assumes that benefits of transformation do not extent beyond stresses that cause DNA damage. If benefits of transformation do extend to a more broad range of stresses then this too argues against the narrow concept of transformation as solely a mediator of DNA repair. In order to address these limitations of the DNA-for-repair hypothesis, we went on to further consider benefits of transformation/competence against non-DNA-damaging stress using an assay protocol that enabled us to distinguish between putative benefits of transformation and benefits associated with the array of functions regulated by competence in both the short- and longer-term. Our results provide support for the broader stress-response advantage of competence hypothesis by Claverys et al. (2006). When faced with two antibiotics that affect protein biosynthesis, streptomycin and kanamycin, competence provides significant short-term advantages as well as a longer-term fitness advantage for streptomycin. These results, in addition to our evidence of benefits to competent cells exposed to mitomycin C, make clear that competence in *S. pneumoniae* is not merely a process used by cells to repair dam-
aged DNA.

That the benefits of competence extend beyond the limited advantages associated with transformation should not be surprising. The competence regulon activated by CSP in *S. pneumoniae* affects more than 100 genes, the precise number depending upon the study (and not considering sRNAs), and only a fraction of these are required for transformation (Peterson et al. 2000; Dagkessamanskaia et al. 2004; Peterson et al. 2004). There are three distinct sets of genes altered by competence induction, early *com*, late *com*, and delayed. Of the early *com* genes, between a third and half are required for transformation (eight out of 17-24) while this is the case for only a quarter of the late *com* genes (14 of 60) and none of the 14-19 delayed genes (Guiral et al. 2007). Regardless of the precise fractions, the point remains that the majority of genes in the competence regulon serve other functions than regulating transformation. Among these functions are genes influencing bacteriocin production, fitness in mice, biofilm production, autolysis and allolysis (Peterson et al. 2004). Rather striking is the apparent excess of genes considered to be involved in stress responses, such as the SOS-response and the bacterial heat shock response (Martin et al. 1995; Peterson et al. 2004; Erill et al. 2007). This abundance of stress response genes regulated by competence, together with data indicating that competence is induced when cells are exposed to a variety of antibiotics from a range of target classes (Prudhomme et al. 2006), led to the suggestion that competence was a general stress response akin to the classical SOS response in Gram-negative bacteria. Our results support this broader role of competence.

Our results have also shown that the benefits of competence can be partitioned into those responses requiring transformation from those which do not, and that this difference appears to depend upon the nature of the stress experienced by cells. To examine this we established conditions that would permit or prevent DNA uptake in cells artificially induced to competence with synthetic CSP. We found that benefits of competence versus mitomycin C required intact DNA. This benefit was eliminated when this DNA was degraded with DNase. By contrast, when competent cells are exposed to streptomycin, they gain benefits of com-
petence induction that are equal whether or not DNA was available for transformation. In this case the benefits from competence appear to be derived from one of the other functions of this regulon. It is possible that GroEL and associated chaperones are able to mitigate the effects of this drug because streptomycin binds to and corrupts ribosome activity. If this were the case, we would predict that DNA would not be required to mitigate the effects of drugs with similar activities to streptomycin. More generally, it would be worthwhile to broadly examine stresses, such as heat, that are known to affect protein folding. In further studies we will aim to address these possibilities directly.

In contrast to the case with streptomycin, intact DNA is required for cells to gain a benefit of competence in the presence of mitomycin C. This suggests that transformation repairs lesions caused by the antibiotic, which is supported by the finding that protection from mitomycin C becomes apparent only in the latter stages of competence. There are, however, alternatives to this explanation that have to be considered, given that our data do not provide direct evidence that template DNA is recombined with the damaged areas on the genome. Mongold (1992) suggested that homologous recombination could possibly improve the survival of UV-exposed competent cells of *H. influenzae* by stalling DNA replication throughout the genome thus providing time for repair of DNA damage before it would fatally interfere with replication. Along these lines, Johnsen et al. (2009) showed that competent cells of *B. subtilis* overcomes penicillin stress by reducing, or even pausing, growth (an outcome not evident here, see Fig. 2.4A). Whether or not this effect would be realized when cells are confronted with antibiotics whose activity is less dependent upon cell growth is unknown, although our growth rate results indicate that *S. pneumoniae* does not stop growing during competence. A related idea to Mongold’s is that internalized DNA could promote the development of new replication forks that would help to bypass mitomycin C induced lesions (Guiral et al. 2007). As yet this possibility remains unexplored.

The nature of stress experienced by *S. pneumoniae* in their natural environment is difficult to
gauge. It will almost certainly be a function of whether cells are residing within the human nasopharynx or moving between hosts. In both locations, however, the stresses will be, by their very nature, unpredictable in kind and intensity. In consideration of this diversity, we utilized two different assay approaches to examine benefits of competence that varied in the lethality of the stress and in the timing of CSP/DNA presentation. In the initial assays carried out in lethal conditions benefits were only observed following induction. In the second assay, carried out with a milder stress over a longer period, delayed benefits were seen in growth curves and fitness assays when CSP and drug were simultaneously applied. These results show that benefits of competence/transformation are present at multiple time scales that can last well beyond the time that the cell is thought to be competent. This raises two related questions about competence and its induction: 1) What conditions induce competence in nature? 2) Why is competence regulated via an extracellular signal, CSP (and, indeed, is it induced solely by this signal)? Natural induction is thought to occur in two ways, either when CSP passively reaches a threshold density (for an alternative view, see e.g. Hense et al. 2007; Yang et al. 2010) or following exposure to sub-lethal concentrations of antibiotic (Prudhomme et al. 2006) which is thought to induce competence via a CSP dependant response. If competence is a response to stress then this implies that the stress precedes CSP production. Unless cells are able to predict future lethal stress, using for example other external proxies, competence induction via CSP secretion would fail to provide benefits. This type of induction could work, however, with sub-lethal stress. It has been suggested that CSP might act as an alarmone that is secreted in response to sub-lethal stress (Havarstein and Morrison 1999; Prudhomme et al. 2006), thereby coordinating a population-wide response. However, this fails to explain why a coordinated response to stress is required in this species as opposed to an autonomous one. Our results may suggest the intriguing possibility that, in fact, an autonomous response is also possible. Results in Figure 2B imply that a CSP-independent protective response arises in mitomycin C exposed populations after 60 minutes of exposure. In addition, the results of competition assays shown in Figure 5 reveal that the ΔcomC strain, FP5, is more fit than the ΔcomD strain, FP48, in the presence of antibiotic
stress, but not in its absence. We speculate that CSP-independent protection appears in FP5 following antibiotic stress, but not in FP48. This is consistent with the idea that stress-induced protection is ComD dependent, an idea not without precedent. Pinas et al (2008), for example, found that the acid stress response in *S. pneumoniae* was CSP-independent, but relied instead on ComE, which is phosphorylated by ComD. If ComD phosphorylation of ComE is also required to induce the antibiotic stress response found here, it would explain the stress-dependent fitness benefits of FP5 over FP48. Future experimental work using a broader range of mutants will be required to test this inference more precisely. However, more generally, these results suggest that *S. pneumoniae* employs a two-pronged response to stress, one that is CSP-dependent and relies, to some degree on population density, and one that is cell autonomous that would afford cells protection even if the density required for CSP mediated induction are not realized. Clearly much remains to be explained about how the regulation of this state maximizes the fitness benefits of competent cells.

Finally, if we return to the original question of ‘what do bacteria do with DNA?’ then our results add to the existing support for the DNA-for-repair hypothesis showing that transformation in *S. pneumoniae* protects cells from DNA damage. More importantly, we show that in the case of competence the question: ‘What do bacteria do with DNA?’ is too narrow by showing that the stress response extends beyond transformation and DNA damage. This is likely not unique for *S. pneumoniae* as in other species, such as *B. subtilis* and *Streptococcus gordonii*, regulatory genes controlling transformation are a small fraction of genes affected by competence (Berka et al. 2002; Vickerman et al. 2007). The competence regulon is not only essential for transformation, but also regulates a series of other functions. In light of our results it is essential now to shift focus from transformation to these broader functions in order to identify long and short-term benefits of this widespread bacterial trait.

### 2.6 Acknowledgements

We thank F. Iannelli for kindly providing the competence mutant strains. We thank T. Coop-
er, O. E. Cornejo, B.A. Evans and two anonymous reviewers for their insightful comment on previous versions of the manuscript. This work was supported through a grant from the BBSRC to DER.

2.7 References


Fraser, C., W. P. Hanage, and B. G. Spratt. 2007. Recombination and the nature of bacterial
Guiral, S., M. Moscoso, A. Dagkessamanskaia, and J. P. Claverys. 2007. Competence in


3.1 Abstract

Natural transformation is a phylogenetically widespread mechanism of horizontal gene transfer in bacteria and a key factor in bacterial genome evolution. Despite this, the evolutionary factors maintaining competence among naturally transformable bacteria remain uncertain. Two potential long-term benefits of transformation are that it increases the rate of adaptation and reduces the accumulation of deleterious mutations. Recent theory, encapsulated in the Fitness Associated Recombination hypothesis (FAR) suggests that benefits of transformation are maximized when bacterial fitness is lowest, and that transformation is costly when bacterial fitness is high. Here we test these novel predictions using long-term evolved populations of the naturally transformable bacterium *Streptococcus pneumoniae*. Using strains that are identical except for their ability to become competent, we show that competence is highly costly under benign conditions, but that these costs are alleviated when cells experience periods of reduced fitness. Next, using whole genome sequencing, we show that competence reduces population diversification and benefits cells by reducing the fixation of new mutations. This benefit is magnified by results showing that non-competent populations evolve increased mutation rates, while evolved competent populations retain the wild-type mutation rate. We conclude that competence in *S. pneumoniae* is a conservative, rather than innovative, process that limits evolutionary change. Our results provide experimental evidence in bacteria for the FAR hypothesis, and show for the first time that the costs and benefits of competence are context-dependent.
3.2 Introduction

Bacterial transformation can have significant effects on the evolution of bacterial genomes and populations by facilitating the horizontal spread of genes conferring novel functions such as antibiotic resistance, new metabolic capabilities, and new modes of virulence (Kelly et al. 1994; Ochman et al. 2000; Boyd et al. 2009; Juhas et al. 2009; Vos 2009; Treangen and Rocha 2011). Because many of the consequences of transformation in bacterial pathogens are advantageous in a clinical context, e.g. the ability to evade antibiotics or immune surveillance, bacterial transformation is widely believed to have evolved to facilitate adaptation (Eshel and Feldman 1970; Levin 1981; Gogarten et al. 2002; Vos 2009). However, this conclusion ignores two potentially important factors. Firstly, pathogenic isolates of most bacterial species may not be representative of the bacterial species as a whole, particularly if these strains experience increased pressure from factors that either increase the rate of transformation or the survival of recombinant strains. In fact, not all naturally transformable bacterial species are pathogens (Didelot and Maiden 2010). Secondly, benefits of transformation inferred from available sequence data represent a highly biased subsample, given that some fraction of the new recombinant strains is eliminated by natural selection. Because of these limitations, any costs and benefits of bacterial transformation remain uncertain.

As for eukaryotic sex, any benefits derived from bacterial transformation must be of sufficient magnitude to offset any associated fitness costs. Several benefits have been proposed: 1) transformed DNA is used to repair DNA damage; 2) transformed DNA is used as a resource; 3) transformation facilitates adaptation (Redfield 2001; Michod et al. 2008). Here we focus on predictions of this third hypothesis. By combining beneficial mutations into a single individual, transformation can reduce clonal interference (Gerrish and Lenski 1998; de Visser and Rozen 2006) and potentially facilitate adaptation in a manner equivalent to the Fisher-Muller model in eukaryotes (Fisher 1930; Muller 1932; Felsenstein 1974; Cohen et al. 2005). Equally, by combining deleterious mutations into a common background, transformation can reduce the mutational load by more efficiently exposing harmful mutations.
to natural selection (Muller 1964; Haigh 1978; Kondrashov 1988). Finally, transformation can directly reduce the mutational load by replacing damaged DNA using free-DNA from the environment (Redfield 1988). Several potential costs of competence, including energetic or metabolic costs as well as risks associated with recombining with free DNA of unknown provenance, oppose these benefits. If transforming DNA comes from dead cells, it may contain the DNA lesions that led to the donor’s demise; when transformed, this DNA may bring along these deleterious effects (Redfield 1988; Redfield et al. 1997). Although metabolic costs of bacterial competence have not been quantified in any system, high costs are likely considering the large number of genes involved in regulating this process. For example, in Streptococcus pneumoniae more than 100 genes are induced by competence, of which only a small fraction are used for the purpose of recombining DNA (Peterson et al. 2000; Peterson et al. 2004). Additionally, in some species the induction of competence coincides with growth arrest (Johnsen et al. 2009).

Considering the competing costs and benefits of competence, it is unsurprising that experiments designed to quantify long-term benefits of bacterial transformation have been equivocal. While Baltrus et al. (2008) showed that transformable strains of Helicobacter pylori evolved more rapidly than non-transformable strains, Bacher et al. (2006) showed the reverse in Acinetobacter baylyi, where transformability was reduced in wild-type strains. While these experimental differences may be caused by particulars of the species investigated, they may have also been caused by experimental approaches that inadvertently exposed either the benefits or costs of competence. One factor that is likely to tip this balance is the initial level of bacterial fitness in the experimental environment (Vos 2009). For example, while transformation is more likely to be costly in a well-adapted cell because transformation may result in the loss of beneficial alleles or combinations of alleles (Redfield 1988; Narra and Ochman 2006), these costs are unlikely to be realised in a poorly adapted cell that has few beneficial alleles to lose.
In this study we use an experimental evolution approach to understand the costs and benefits of bacterial transformation, using the naturally competent gram-positive pathogen *S. pneumoniae* as our model. We test this by allowing wild-type and mutant cells unable to induce competence to evolve in benign and low-fitness conditions. Furthermore, we examine the genetic mechanisms underlying our results using whole genome re-sequencing of evolved isolates. We predict that if the benefits of transformation solely follow the Fisher-Muller model then fitness of competent population is higher than non-competitive populations after 1000 generations of evolution. In addition, we would expect to see a higher rate of fixation of mutations for competent populations. However, if benefits of competence mostly rely on a reduction of the mutational load then we expect to see less fixed mutations for competent populations. Finally, if a combination of both models result in the benefits of competence then it is possible that benefits of competence might depend on the context and we might see a differential result between evolutionary treatments. Our results provide experimental support for competence as a conservative factor mainly reducing mutational load. The benefits of competence, however, are context dependent, with only showing a benefit in the presence of periods of mild stress. Finally, we present serendipitous genotypic and phenotypic data showing a propensity for the evolution of mutator phenotypes in non-competitive bacterial populations.

### 3.3 Materials & Methods

#### 3.3.1 Strains, culture conditions and chemostats

Strains used in this study were derived from Rx1 and its non-competitive derivative FP5 (Iannelli et al. 2005). FP5 is non-competitive due to a deletion in the gene encoding the competence peptide signal *comC* and is otherwise isogenic to the wild-type Rx1. Spontaneous rifampicin or streptomycin resistant mutants were isolated from each strain, and then four independent colonies of each type were further sub-cloned and stored at -80°C. These four ancestors were selected because they had similar growth rates in batch culture (Supplementary material S3.1). These 16 total clones (2 strains * 2 drug resistance types * 4 replicates)
represented the ancestral populations for experimental evolution. Cultures were grown in ¼ CTM pH 7.8 (Complete Transformation Medium), which per litre consists of: 7.5g Tryptic Soy Broth; 0.25g yeast extract; 6g NaCl at pH 7.8. CTM is a medium optimised to permit high rates of natural transformation and at this strength allowed for substantial transformation in the long-term evolutionary environment (supplementary material S3.2). Blood agar plates (Tryptic Soy Agar (TSA) + 3% horse blood), supplemented, where necessary, with either 4µg/mL rifampicin or 100µg/mL streptomycin, were used to enumerate cell density within chemostats and for colony counting during competition assays.

Experimental evolution and competition assays were performed in custom-made chemostats with a 25 mL working volume and a flow rate of 4mL/hr, whilst maintained at 37°C (Fig. 3.1A). Chemostat cultures were inoculated and maintained as described previously (Cornejo et al. 2008), and sampled every 50 generations of growth. Samples were stored at -80°C as freezer stocks in ¼ CTM pH 7.8 + 25% v/v glycerol at an OD600 of 0.20, corresponding to a density of 2x10^8 cells mL⁻¹.

3.3.2 Long-term evolution experiment

Sixteen chemostat populations were inoculated to generate 4 replicates each of a 2 * 2 treatment design pairing competence and stress (Fig. 3.1B). The replicates in each treatment were equally split between the two differently marked versions of Rx1 (competent strain) and Fp5 (non-competent strain). Half of the populations were exposed twice a week to low doses of kanamycin inoculated directly into the chemostat to simulate short periods of low-fitness. Kanamycin concentrations were 5µg/mL upon inoculation, but reduced with the normal outflow rate of the chemostat. This concentration of kanamycin did not reduce the growth rate of the cells (supplementary material S3.3), but is sufficient to cause ribosomal decoding errors during protein production, which promotes the induction of competence (Prudhomme et al. 2006; Stevens et al. 2011). For simplicity, we will hereafter refer to this treatment as “stress” and the basal treatment as “benign”. Every week, after approximately 50 generations, a 1mL sample was taken from each population and tested for the presence of the cor-
rect marker and absence of the opposite marker. Contaminated populations were restarted with 50µL of the previous sampled uncontaminated time point. Populations were maintained for 20 weeks, which corresponds to about 1,000 generations.

3.3.3 Fitness assays

Fitness assays between ancestor and evolved line resulted in an infinitely high fitness of the evolved populations due to termination of the oscillating population density behaviour (Cornejo et al. 2008) in the evolved populations. Therefore, fitness was determined by comparing the change in relative densities of two reciprocally marked evolved populations in a chemostat in mixed culture over a 32-hour span. Competition assays were initiated by inoculating chemostats with equal densities of each competitor. Chemostats were sampled immediately and then again after 32 hours to determine the relative densities of each competitor. The Malthusian parameters per hour were then calculated for each strain based on the density of each strain at the start and end of the competition, as described previously (Lenski et al. 1991). Fitness (selection rate constant) was then calculated as the difference
between Malthusian parameters following Travisano et al (1996) because of often declining population densities of one of the competitors. Two sets of assays were performed. Firstly, competent populations were compared to non-competent populations within either the benign or periodic stress treatment (Fig 1B, vertical comparison). Secondly, populations in the benign environment were compared to populations in the periodic stress treatment within either the competent or non-competent treatment (Fig 1B, horizontal comparison). The selection rate constant values that were the result of these comparisons were subsequently tested for a significant fitness difference by comparing a restricted maximum likelihood mixed model against an intercept of zero, which corresponds to equal fitness. In the mixed model, replicate fitness assays of competitor combinations were nested as a random effect within the fixed effect of treatments (absence/presence periodic stress and absence/presence competence). Secondly, we used the restricted maximum likelihood (REML) mixed model, again with replicate fitness assays as a random factor within treatments, to test for fitness differences between treatments. All analyses were done in R (v2.13.0) with package LME4. P-values were estimated by MCMonte-Carlo simulation with 10,000-fold replication using the languageR package (pvals command; Supplementary material S3.4 for R-code).

3.3.4 DNA isolation and sequencing

Clonal isolates from each of the 16 evolved populations as well as all four ancestor strains were sequenced using the SOLiD4 platform at the University of Manchester genomics facility. Genomic DNA was extracted using phenol-chloroform isolation and ethanol precipitation (Sambrook et al. 1989). Genomic DNA was resuspended in 100µL ddH2O and DNA concentrations were quantified by ND-1000 (Nanodrop products, Delaware, USA). SOLiD data were normalised to an equal number of random reads (8,315,863 per strain) for each sample using a custom perl script (getRandomTags_Index_fastq.pl) developed by I. Donaldson. The normalised reads were then mapped against the fully sequenced reference strain S. pneumoniae R6 (genbank accession: NC_003098) using BFAST (0.6.4e) using default colour space methodology giving an average coverage depth of 151-fold. Mapped reads were then locally
realigned around INDELs using SRMA (0.1.15). The resulting files were subsequently used for downstream analysis in Geneious (Geneious 5.4, Auckland, New Zealand; Drummond et al. 2011), the Galaxy online tool set (Blankenberg et al. 2010; Goecks et al. 2010) and the UCSC genome browser (Schneider et al. 2006).

3.3.5 Mutation analysis

Non-synonymous, synonymous, small INDELs and intergenic mutations were determined using the SNP calling tool in Geneious. SNPs and INDELs were called when the change to the reference was supported in 60% of the reads with at least a coverage depth of 20 reads. This approach minimised false positive and negative SNP calls. The 60% support is slightly less stringent than previous studies (Harris et al. 2010; Croucher et al. 2011), but this is compensated by the high coverage. Variant tables extracted from Geneious were used in Galaxy to identify mutations for each evolved clone compared to its ancestor. Parallel changes were then double checked by hand in the UCSC microbial genome browser to eliminate false positives. Only SNPs that had arisen de novo during the long-term experiment were used for further analysis.

3.3.6 Mutations and mutation rate

To determine the effect of periodic stress and competence the total number of mutations were compared in a GLM model with a Poisson distribution using R. Subsequently, two different methods were used to determine the mutation rate of evolved populations. First, the mutation rate was calculated from genome data using synonymous SNPs (Wielgoss et al. 2011). The mutation rate per base pair per generation was then determined as the number of synonymous SNPs divided by the number of potential synonymous mutations over a 1,000-generations period, which was calculated as the product of the number of synonymous sites in coding regions of the genome multiplied by 1,000. The number of synonymous sites was determined from the known coding region annotation for R6 (Genbank accession: NC_003098) using a previously described Perl-tool (Wielgoss et al. 2011). This analysis
ignores base composition effects or small changes to genome size that might have occurred during the evolution experiment. Mutation rates calculated from sequence data were tested using the GLM-function in R.

Second, a limitation of the preceding analysis is that it cannot account for the possibility that cells may have evolved new mutation rates at different times during the long-term experiment. To overcome this limitation we quantified mutation frequencies of all terminal lineages using a phenotypic assay that quantifies rates of spontaneous mutation to either rifampicin or streptomycin (Bjedov et al. 2003). Each strain and its corresponding ancestor was grown overnight at 37°C+5% CO₂, after which all cells from the confluent lawn were collected in 0.8% NaCl. Cells were washed and concentrated by centrifugation and resuspended in 100µL 0.8% NaCl. 10 µL spots at several different dilutions (five spots per dilution) were placed onto blood agar plates supplemented with either 4µg/mL rifampicin or 100µg/mL streptomycin, whilst total cell densities were determined on unsupplemented plates. Mutation rate was calculated as the ratio of number of mutants to the total population size. Assays were performed in triplicate for each genotype, and the relative mutation rate was determined as the ratio of the mutation rate of each evolved lineage to its the corresponding ancestor. Mean relative mutation rates of an evolved strain were log-transformed before analysis using a two-way ANOVA to determine the effects of periodic stress and competence on the log of the relative mutation rate. One sample t-tests against 0 were done to test if mutation rates of evolved populations had significantly changed compared to the ancestor.

Parallel changes in DNA repair genes were examined at the level of each gene and functional group, as determined from the KEGG classifications. The table of non-synonymous SNPs was used to create a SNP-by-gene table by scoring presence/absence of at least one SNP in a given gene for each strain. Functional group associations were created from the total SNP-by-gene table by summarising presence and absence of SNPs for genes associated with a functional group according to the KEGG-database for S. pneumoniae R6 (http://www.ge-
nome.jp/dbget-bin/www_bget?gn:T00060). Generalised linear models were used to test for differences between treatments for parallel non-synonymous mutations in functional groups.

3.4 Results

3.4.1 Costs of competence

We used pairwise fitness assays between evolved populations to quantify the relative fitness of competent and non-competent populations that have evolved under benign or stress conditions. In contrast to the expectation that competence facilitates adaptation, we found that competent populations of S. pneumoniae are significantly less fit than non-competent populations (Fig. 3.2A, REML mixed model compared to 0: \( t = -2.22; p < 0.001 \)). Thus in a benign environment competence is highly costly to cells, corresponding to a selective rate constant of \( \sim - 0.10/\text{hr} \). By contrast, the fitness of competent and non-competent populations evolved in the presence of periods of stress did not differ from one another (Fig. 3.2A, REML mixed model compared to 0: \( t = -0.18; p = 0.741 \)). These results indicate that costs of competence during long-term evolution are significantly reduced by periodic stress (Fig. 3.2A, REML mixed model: \( F = 2.077; p = 0.017 \)).

![Figure 3.2](image)

**Figure 3.2.** Mean selection rate constants of evolved populations. A) Fitness of competent populations compared to non-competent populations in the two different evolutionary environments. B) Fitness of populations evolved in a periodic stress environment compared to populations evolved in the benign environment for both types of strains. Dashed line indicates the value at which fitness is equal between competitors. Error bars are SE of the mean fitness of eight different evolved strain combinations within a treatment.
There are two, not mutually exclusive, explanations, which explain how costs of competence are mitigated by stress. Firstly, periodic stress could increase the rate of adaptation of competent cells. Secondly, stress could somehow harm non-competent cells. Both possibilities would lead to the same net outcome, namely that costs of competence are eliminated in the presence of stress (Fig. 3.2A). Subsequent tests provide support for both possibilities. We found a significant fitness benefit of approximately 0.05/hr to stressed versus unstressed competent cells (Fig. 3.2B, REML mixed model compared to 0: $t = 1.27$; one-tailed $p = 0.045$). In addition, we identified a significant fitness cost, of approximately -0.05/hr, to non-competent cells evolving in an environment with periodic stress (Fig. 3.2B, REML mixed model compared to 0: $t = -1.67$; $p = 0.023$). The sum of these two fitness effects, i.e. the fitness cost for non-competent populations and the fitness benefit for competent populations evolved in the stress treatment at roughly 0.05/hr each, corresponds to the 0.10/hr cost of competence to cells evolved in an unstressed environment. Thus competence appears to benefit stressed cells in two ways, by increasing their rate of adaptation and by enabling cells to avoid the negative effects of stress during evolution as seen in the non-competent populations.

While these results follow predictions of the FAR-hypothesis and suggest how competence can be maintained in spite of its costs in a benign environment, they do not reveal the mechanisms of these effects. To begin to identify these mechanisms we obtained the complete genomes of evolved clones from each of the 16 experimental populations.

3.4.2 Transformation is conservative

We consider three hypotheses for the role of transformation on the fixation of new mutations in evolved genomes. The null hypothesis is that evolved clones fix mutations at the same rate as one another, irrespective of transformation. Or transformation could increase the fixation rate of mutations. This could occur if, for example, transformation recombines mutations from separate cells into a single genome. Finally, transformation could reduce the fixation
rate of mutations if it either recombines deleterious mutations into a single background, thus facilitating their removal by natural selection, or removes lesions by site-specific recombination using unmutated template DNA. Our results are strongly consistent with the third hypothesis (Fig. 3.3). Specifically, we find that the genetic distance between competent populations is significantly less than populations evolved without competence (Fig. 3.3A; GLM: $X^2 = 9.28$, df = 1 $p = 0.002$). Transformation therefore reduces the fixation rate of non-synonymous mutations. Taken together with the fitness data, we conclude that the conservative nature of transformation comes at the cost of a reduced rate of adaptation in benign conditions. However, the conservative nature of competence allows for better regulation (active or passive) of the rate of fixation of non-synonymous mutations, which increases for competent populations, but not non-competent populations, in the presence of periodic stress (Fig. 3.3A; Welch’s t-test: non-competent: $t_{5.57} = 1.1659$, $p = 0.291$; competent: $t_{7.89} = 7.82$, $p < 0.0001$). Overall, the effect of stress on genetic divergence is only marginally non-significant (Fig. 3.3A; GLM: $X^2 = 3.0835$, df = 1 $p = 0.079$).

### 3.4.3 Mutations and the mutation rate

We find that the total number of mutations in non-competent populations is significantly higher than in competent populations (Fig. 3.3A: $z = -9.344$, df = 1 $p < 0.0001$). Moreover, the total number of mutations increased in populations experiencing periodic stress (Fig. 3.3A: $z = 7.379$, df = 1 $p < 0.0001$). Two interpretations of this result are that: (1) transformation exposes cells with deleterious mutations to selection, thereby facilitating their removal, or (2) that transformation promotes site-specific repair/replacement with un-mutated homologous DNA. A third possibility is that non-competent populations evolved an increased mutation rate relative to competent lineages. To explore this possibility we used complete genome sequences to compare the substitution rate of all terminal populations. We observed a total of 448 synonymous mutations and 1266 non-synonymous mutations across the sixteen evolved lines. However, substitutions were not evenly distributed across populations, with three of eight non-competent populations substituting between two and 10-fold more
mutations than the other populations. The changes in these lineages, which are only present in non-competent populations, are consistent with an increased mutation rate (Fig. 3.3B); accordingly, we find significantly more mutators among non-competent than competent lineages (Binomial Logistic Regression, one-tailed: $X^2 = 3.104, df = 1, p = 0.039$). Consistent with their increased mutation rates, we find that these lineages carry several mutations in DNA repair genes that are diagnostic for genetic mutators (e.g. polC, dnaQ, and mutL; Horst et al. 1999). we then quantified substitutions in DNA repair genes in all 16 populations. By this approach, we detected significantly more non-synonymous mutations in DNA repair
genes in non-competent populations than in competent populations (GLM one-tailed: $X^2 = 162.00$, df = 1, p = 0.0478), an effect that is most pronounced in genes for mismatch repair (Fig. 3.3C; GLM: Mismatch repair: $X^2 = 9.875$, df = 1, p = 0.001). Finally, we used a phenotypic assay to estimate the mutation rate of all evolved lineages relative to their ancestor. Consistent with results based on sequence analyses, we found that the mutation rate of non-competent lineages was significantly higher than competent populations (Fig. 3.3D; Two-way ANOVA: $F_{1,11} = 10.189$; p = 0.009), while there was no overall effect of stress (Fig. 3.3D; Two-way ANOVA: $F_{1,11} = 2.699$; p = 0.129), nor an interaction between stress and competence (Fig. 3.3D; Two-way ANOVA: $F_{1,11} = 0.0003$; p = 0.9876). Thus both at the genetic and phenotypic levels, our data show that non-competent populations have evolved increased mutation rates during the course of this experiment.

### 3.5 Discussion

The enigma of eukaryotic sex revolves around identifying conditions allowing the benefits of sex to overcome its costs. Our work reveals that a similar challenge exists for natural competence in prokaryotes. We find that when populations of *S. pneumoniae* are evolved in benign conditions in chemostats, competence exacts a significant fitness cost (Fig. 3.2A). When compared to non-competent cells, wild-type cells adapted significantly more slowly, consistent with results of Bacher et al. (2006) in *A. baylyi*. However, when evolving cells experience regular pulses of a very mild stress, competent cells adapt as rapidly as their non-competent counterparts (Fig. 3.2). To a first approximation this result supports conclusions of an earlier study in *H. pylori* that competence increases the rate of adaptation (Baltrus et al. 2008). However, our data suggest only that mild stress might permit competence to persist, as wild-type cells are no better than non-competent cells. Therefore, in this instance costs and benefits of competence appear to be fortuitously balanced.

Such a delicate balance is unlikely to be found within the human nasopharynx, where *S. pneumoniae* experiences unpredictable and more profound stress from, e.g. drug exposure,
immune surveillance and physical or chemical stress from coexisting bacterial competitors (Michod et al. 2008). Our results suggest several ways that these forms of stress would favour the maintenance of competence. Firstly, even the relatively minor stress we imposed was sufficient to offset intrinsic costs of competence, as a result of costs to non-competent and benefits to competent populations (Fig. 3.2B). Secondly, the rate of environmental change in our experiment is likely to be considerably lower than that which cells would experience within their human host. Increasing the frequency of periodic stress might increase the advantage of competence (Peters and Lively 1999; Gandon and Otto 2007). Thirdly, in previous short-term experiments we found that competence protects cells from much higher concentrations of antibiotic than we used in the long-term experiment (Engelmoer and Rozen 2011). This, together with our phenotypic results support the major prediction of the FAR hypothesis as applied to bacteria, namely that competence is favoured during periods of fitness-reducing environmental stress (Hadany and Beker 2003; Hadany and Otto 2009).

To identify mechanisms underlying our phenotypic results, we used whole-genome sequencing of all terminal evolved populations. In contrast to previous predictions that competent populations would substitute more mutations than non-competent cells owing to recombination, we observed the reverse. Competent populations were genetically less diverged than non-competent ones because of a reduced number of accumulated mutations. This result clearly argues against the idea that competence benefits cells by recombining beneficial mutations from separate cells into a common genetic background, as expected from the Fisher-Muller model. Instead, these data imply that transformation is beneficial because it reduces the mutational load (Redfield 1988; Redfield et al. 1997). One mechanism by which this can occur is through template directed reversion, owing to the fact that transformed fragments coming from lysed clone-mates are likely to carry the most prevalent un-mutated alleles. This reversion to the wild-type reduces the mutational load but may also be costly for adaptation in a constant environment. Alternatively, transformation can combine deleterious mutations into a single genetic background, thereby facilitating their efficient removal.
by natural selection, in a manner similar to the mutational deterministic hypothesis for eukaryotes (Kondrashov 1988; Barton and Charlesworth 1998; Kouyos et al. 2007). Cells periodically exposed to mild stress appear to suffer from the burden of their mutational load. Accordingly, non-competent cells suffer to a greater degree because they cannot revert to a less loaded state as easily as their recombining counterparts. This may be exacerbated by increased mutation and transformation rates during stress (Bjedov et al. 2003; Henderson-Begg et al. 2006). Redfield (1988) showed that high mutation rates are required for transformation to provide benefits by reducing the mutational load. This condition appears to be satisfied in *S. pneumoniae*. we estimate that the average mutation rate in our experimental isolates of *S. pneumoniae* is $\sim 4 \times 10^{-8}$ per bp per generation, corresponding to $U = 0.08$ mutations per genome per generation, or about 200-fold higher than *Escherichia coli* (Wielgoss et al. 2011), yet similar to other opportunistic pathogens, such as *H. pylori*, *Pseudomonas aeruginosa* and *H. influenzae* (Oliver et al. 2000; Bjorkholm et al. 2001; Watson et al. 2004; Wielgoss et al. 2011). This estimate is consistent with earlier observations in *S. pneumoniae* strains of similar genetic background and some clinical isolates, whilst being about 5 to 10-fold higher than other natural isolates (Morosini et al. 2003; Gould et al. 2007; Cortes et al. 2008; Henderson-Begg et al. 2010). It is thus reasonable to conclude that *S. pneumoniae* experiences a relatively high mutation rate.

Given these high rates of mutation we were surprised to find that some of the non-competent strains, but none of the competent strains, evolved about 5-10-fold higher rates than their ancestor during the course of this long-term experiment (Fig. 3B). As anticipated, these lineages have fixed several mutations in genes responsible for the mutator phenotype in *E. coli*, e.g. *polC*, *dnaQ* and *mutL*). Finding mutators is not surprising by itself, as this has become rather anticipated in experimental evolution studies (Sniegowski et al. 1997; Shaver et al. 2002; Raynes et al. 2011), but rather the fact that mutators were only present in the absence of competence. One could have even predicted that mutators would be more likely to evolve when faced with periodic stress (Travis and Travis 2002; Tanaka et al. 2003). Although we
are uncertain what underlies this difference, we suggest two possibilities. Firstly, transformation can separate mutator alleles from the mutations they cause. Thus while mutators may arise equally in both competent and non-competent cells, they are lost before they become common in competent lineages (Tenaillon et al. 2000). Secondly, Tennailon et al (2000) have proposed that hypermutation and transformation are mutually exclusive adaptive strategies in prokaryotes, whereby conditions favouring transformation necessarily reduce the likelihood that mutators achieve fixation. Distinguishing between these possibilities remains a challenge for future studies in this system.

Little is known about the natural history of most bacteria, which creates important challenges towards understanding the costs and benefits of competence in nature. Patterns of stress are certain to be irregular, as are rates of mutation. Our results suggest that these conditions of stress are favourable to the maintenance of transformation. Similarly, infrequent stress may facilitate the loss of transformation, as it is likely that the conditions favouring transformation may prevail for long enough in some cases while not in others. It is worth noting in this respect that surveys of naturally competent species have revealed significant variation in transformation rates among clones. In *H. influenzae* and *B. subtilis* rates can vary over 6 orders of magnitude (Li et al. 2001; Duitman et al. 2007; Maughan and Redfield 2009). Similar variation has been observed in *S. pneumoniae* (Pozzi et al. 1996; Iannelli et al. 2005; DER unpublished data) suggesting that competence is readily gained and lost in this species.

In summary, we have presented experimental support for the Fitness Associated Recombination hypothesis in bacteria. We show that competence is costly in benign conditions but favoured in low-fitness conditions. In addition, we determine from whole-genome sequences that competence reduces the substitution rate of non-synonymous mutations, and that the absence of competence leads to an increased mutation rate. We conclude that competence in *S. pneumoniae* is conservative, acting to preserve alleles, rather than an innovative process that persists because of benefits it provides by combining beneficial mutations.
3.6 References


Drummond, A. J., B. Ashton, S. Buxton, M. Cheung, A. Cooper, C. Duran, M. Field, J. Heled, M. Kearse, S. Markowitz, R. Moir, S. Stones-Havas, S. Sturrock, T. Thierer,


Stevens, K. E., D. Chang, E. E. Zwack, and M. E. Sebert. 2011. Competence in *Streptococcus pneumoniae* is regulated by the rate of ribosomal decoding errors. MBio 5: e00071


3.7 Supplemental material chapter 3

S3.1 – *Growth rates of the four ancestors*

Growth rates were estimated within a 96-well plate using an automated plate reader (n=5). Cultures were started with 3*10^5 cells in 200µL of CTM pH 7.8. The OD600 was measured every 5 minutes for 24 hours at 37°C with continuous shaking. Raw OD values were normalised to a blank well and Ln transformed before analysis. Then the steepest slope over a 35 minute period was determined. The growth rates were analysed with a one-way ANOVA using R (Fig S3.1).

S3.2 – *recombination in the chemostat environment*

**Material and Methods**

To detect recombination in the chemostat environment, the following experiment was performed. Chemostats were inoculated with either a 1:1 mix of streptomycin and rifampicin resistant FP5 or Rx1 (n=3 for each strain). After 24 hour growth in the chemostats the population density and the number of double ‘mutants’ was determined on blood agar plates supplemented with the appropriate concentration of antibiotics where necessary. The count data was analysed in R using a generalised linear method with a Poisson distribution. A Poisson distribution was used because events over time (mutation/recombination events) were com-

![Figure S3.1 Mean growth rates of ancestral strains. No significant difference was detected between the strains (one-way ANOVA: F_{3,16} = 0.8637; p=0.4801). Error bars are SE of the mean.](image)
pared.

Results

Mixes of two differentially marked strains with either competent or non-competent background were maintained in for 24-h to determine if recombination occurred in the evolutionary environment. Double mutants appeared more often in chemostats with competent populations than in chemostats with non-competent populations (GLM with Poisson distribution: $\chi^2 = 272.37$, df = 1, $p < 0.0001$). This shows that recombination occurred at a higher rate than mutation in the chemostat environment.

S3.3- growth rate of ancestors exposed to different concentrations of kanamycin

Growth rates were estimated at different concentrations of kanamycin for the four ancestors in a 96-well plate using an automated plate reader (n=3). OD600 in CTM pH 7.8 was measured every 5 minutes for 24 hours at 37°C with continuous shaking. Raw OD values were normalised to a blank well and Ln transformed before analysis (Fig S3.2).

![Figure S3.2](image-url) Mean growth rates of the ancestral strains while exposed to different concentrations of kanamycin. Dark greys are non-competent ancestors and light/white are the competent ancestors. Error bars are SE of the mean.
S3.4 - R-code used for the mixed models

# define nested variable, which are the multiple fitness measurements for a combination of evolved lines.
mydata=within(mydata, nest<-factor(comp:treat))

# Mixed effect model (package lme4) with nest as random effect and the treatment (competent vs non-competent or benign vs stress) as a fixed variable
lme=lmer(fitness~treat + (1|nest), data=mydata)

# Mixed effect model difference from 0
lme2=lmer(fitness~0 + treat + (1|nest), data=mydata)

# transform Mixed model output into p-value (package languageR)
pvals.fnc(lme)
Chapter 4

Phenotypic and genotypic parallel evolution in long-term evolved Streptococcus pneumoniae

Daniel J.P. Engelmoer & Daniel E. Rozen

4.1 Abstract

Parallel evolution provides strong evidence for natural selection. Several examples have been described where the same selective pressures on standing genetic variation lead to the similar genotypes and phenotypes. Unclear is to what extent parallelism will arise from de novo mutations. In previous experimental studies parallelism could arise because evolution was occurring in a uniform environment. However, this is hardly a natural environments situation. Therefore, in more diverse environments or in genotypes with differential access to sources of genetic variation (e.g. recombination) genomic parallelisms might not occur. To test this we used long-term evolved strains of Streptococcus pneumoniae, which differed in their ability to recombine and in their exposure to a periodic mild stress, in combination with whole genome re-sequencing. We found that populations had evolved phenotypic and genotypic parallelism. At the genomic level most parallel changes were identified at the level of a locus or gene group, while few parallel changes were found at the nucleotide level. Several parallel genotypic changes were treatment specific, which indicates that these changes were specifically beneficial in these groups (either stress or recombination). Other parallel changes were observed to occur in almost all populations, suggesting that these changes caused general benefits. These results reveal that parallelism can both be bounded by, and also overcome differences cells experience in their evolutionary environments.
4.2 Introduction

Parallel genetic or phenotype changes occurring across diverged populations or species provide some of the most compelling evidence that these changes have arisen due to the action of natural selection. For example, freshwater sticklebacks (*Gasterosteus aculeatus*) within different watersheds have independently diverged into specialist lake and stream morphological types that are more similar across watersheds than within them (Kaeuffer et al. 2012). Moreover, some of the changes underlying this phenotypic parallelism are mirrored at the genomic level, where several QTL’s, which are associated with the phenotypically diverged traits, were also significantly diverged in parallel across multiple watersheds (Kaeuffer et al. 2012). Similar molecular parallelism is found in evolution of flower colour (Streisfeld and Rausher 2009; Smith and Rausher 2011), quinolone resistance in *Pseudomonas aeruginosa* (Wong and Kassen 2011), and pericarp colour in rice (Gross et al. 2010). These examples reveal that natural selection acting on standing genetic variation can produce parallel phenotypes via the same mechanistic route. Less clear is the degree to which parallelism will arise due to novel mutations. This question is challenging to address experimentally in natural populations of animals or plants because it is difficult to control environmental conditions and to ensure a uniform genetic background across starting populations. In addition, the waiting time required for new mutations to arise and become fixed limits the scope of multi-generational studies to organisms with a short life-cycle and large population sizes. Microbes evolved in a controlled laboratory setting combined with next-generation sequencing are therefore ideally suited to studying mechanisms of adaptation and the origin of parallel traits. In this study we examine phenotypic and genetic parallelism in a long-term evolutionary experiment with the bacterium *Streptococcus pneumoniae*.

In a landmark study, Wichman et al. (1999) found that during laboratory evolution bacteriophages evolved parallel adaptive genetic changes at the base pair level. By contrast, extensive parallelism at the base pair level has generally not been observed in similar studies conducted on bacteria (Woods et al. 2006; Lieberman et al. 2011; Tenaillon et al. 2012),
where instead molecular parallelism has been found to occur at the broader level of genes and functional groups. In both groups of organisms, when the effects of these changes have been tested using allelic replacement, they are often found to confer beneficial and largely equivalent phenotypes. For example, the biofilm forming phenotype of *Pseudomonas fluorescens* (Wrinkly Spreader) that readily evolves in static environments can be attributed to many different mutations. Some of these occur in the same genes, but others are found in different pathways but cause similar changes at a global transcriptional level (Knight et al. 2006; McDonald et al. 2009). Similarly, different mutations occurring in the same regulatory locus in laboratory evolved strains of *Escherichia coli* impart equal fitness benefits and cause similar expression changes across the genome (Pelosi et al. 2006; Philippe et al. 2007; Crozat et al. 2010). There are also, however, examples where mutations occurring at different sites within the same locus cause different fitness effects (Cooper et al. 2003; Herring et al. 2006) and are associated with distinct suites of correlated responses (Ostrowski et al. 2008).

In each of these studies, replicate lineages of bacteria were evolved in identical conditions. Thus parallelism in these studies could arise because the initiating strains were identical or because they were subsequently evolved in highly controlled and identical environments. Would parallel changes, either at the phenotypic or genetic levels, prevail in cases where replicate populations experienced contrasting environments? Or would parallel changes occur if the capacity to generate genetic variation differs between strains? The aim of the present work is to answer these questions using an experimental system with *S. pneumoniae*, where replicate populations differed in their exposure to stress and whether or not cells were capable of undergoing recombination due to natural transformation.

In this study long-term evolved populations of *S. pneumoniae* are examined for both parallel phenotypic and genotypic changes. For parallel phenotypic changes we have studied three phenotypes that are thought to be important for fitness in the evolutionary environment: oscillatory changes in population density, biofilm formation and growth rate (Cornejo et al.
2008; Ferenci 2008). We predict to find a reduction in oscillation because the fitness benefits are very high for cells in a population that can maintain a constant density. For biofilm formation we expect a reduction because our low flow rate and thus low nutrient availability would favour cells in where the new nutrients arrive, which is in the broth. We also predict that the growth rates will be reduced because cells will likely adapt to the maximum growth rate in the chemostat (i.e. the flow rate), which is considerably lower than the current maximum growth rate in batch culture. Parallel genotypic changes are determined from whole genome re-sequencing data, which enables the identification of all mutations arising during the period of laboratory evolution in each lineage. Mutations are partitioned into classes of parallelism corresponding to decreasing levels of resolution: base pair, gene, and functional group. Following Dettmen et al (2012), we predict that the level of parallelism will fall in the following order: phenotype > functional group > gene > base pair. Thus we expect that the lowest parallelism will be found at the base pair level because different mutations can cause similar functional changes. Therefore if the functional change is under strong selection then we expect to see the highest level of parallelism at the gene level or functional group level. This approach will also identify the degree to which parallel changes are associated with either the competence or stress treatments. Although, we have no a priori predictions on the amount of parallel changes specific for treatments we can hypothesize on the absence or presence of these. If adaptation to the chemostat environment is the strongest selective pressure then all parallel changes should be across treatments. However, if there is an interaction between the selective pressure of a treatment and chemostat environment then both parallel changes across treatments and within treatments will be observed.

In short, we find parallel changes at both the phenotypic and genotypic level. Both population density oscillations and biofilm formation are reduced across all populations irrespective of treatment, while growth rate of evolved populations did not show change consistent across or between treatments. At the genomic level many cases of parallel evolution were observed at all genetic levels. However, parallel changes were mostly observed at the level
of locus or gene group. These genomic parallel changes were observed within treatments and across treatment.

4.3 Material and Methods

4.3.1 Strains & culture conditions

Strains used in this study were taken from a long-term evolution experiment (Chapter 3). Briefly, competent (Rx1) and non-competent (FP5) strains of *S. pneumoniae* were evolved for 1000 generations in chemostats containing ¼ strength CTM pH 7.8 (Complete Transformation Medium), which per litre consists of: 7.5g Tryptic Soy Broth; 0.25g yeast extract; 6g NaCl and pH set to 7.8) at 37°C. The two focal strains differ in that FP5 carries a lesion in *comC* which regulates the production of the competence stimulating peptide CSP; accordingly, this strain is unable to induce competence or undergo transformation. The strains are otherwise isogenic. Chemostats were propagated in either of two conditions: unperturbed or with periodic stress, where half of the 16 chemostats were exposed to a sub-MIC concentration of kanamycin (5µg/mL) twice a week. Antibiotic was injected into chemostats as a single dose after which it was diluted naturally by the flow of the chemostat. Samples were collected every week (~50 generations) from each chemostat, supplemented with 25% v/v glycerol, and stored at -80°C. Each strain and treatment was replicated four times, resulting in a total of 16 evolved populations.

4.3.2 Oscillatory population dynamics

To quantify the amplitude of bacterial oscillations in chemostats, 50µL of an evolved population in frozen culture was inoculated into chemostats run at the same conditions as during the long-term experiment. Following inoculation, populations were allowed to equilibrate for 24 hours before sampling. Population densities were determined by measuring the OD600 of samples taken at regular intervals for at least 8 hours and over the course of two consecutive days. OD600 values are proportional to CFU; values were natural log transformed before analysis.
4.3.3 Growth rate

We estimated the growth rate of evolved and ancestral populations by inoculating 200µL culture wells of a 96-well plate with $5\times10^5$ cells, and then measuring OD600 every 5 minutes during 24 hours of growth. Populations were maintained at 37°C with continuous shaking in a plate reader (Biotek). Growth rates were determined in $\frac{1}{4}$ strength CTM. Optical density values were log transformed and growth rates were estimated from the growth curves by calculating the steepest slope over a 30 minute period during exponential growth. To facilitate comparisons across populations, we quantified the relative growth rate of each evolved lineage by dividing the growth rate of the evolved population by the average growth rate of its respective ancestor. Statistical analyses were carried out using JMP 9.0.

4.3.4 Biofilm formation

The ability of evolved populations to form biofilms was estimated in 96-well plates using a crystal violet binding assay (Djordjevic et al. 2002). Populations were inoculated from frozen cultures with $1\times10^6$ cells into 200µL $\frac{1}{4}$ CTM pH 7.8 and maintained in flat-bottomed, plastic 96-well plates at 37°C + 5% CO$_2$. After 24 hours of growth all liquid was removed and the adhered cells growing within surface associated biofilms were washed twice with 0.8% NaCl. Next, 200µL 1% crystal violet solution was added and wells were incubated at room temperature for 15 minutes. Crystal violet was then removed and the wells were washed by rinsing with water to remove any non-adhered crystal violet. The plate was then dried and the crystal violet resuspended in 200µL 95% ethanol and transferred to a new 96-well plate for measurement. To overcome possible differences between experimental blocks, we estimated the relative biofilm formation of each evolved population compared to its respective ancestor, which was run as an internal reference in each experiment. Changes over time in biofilm formation for the different treatments were tested using a repeated-measures ANOVA in SPSS 15.0.
4.3.5 Parallel genomic changes

Genomic sequences for each evolved population at generation 1000 and their ancestors were obtained using the next-generation sequencing platform SOLiD4 at the University of Manchester genomics facility. Raw SOLiD data were normalised to an equal number of reads by randomly selecting a pre-set number of reads (8,315,863) for each population. These reads were then aligned to the reference genome (S. pneumoniae R6) using BFAST (0.6.4e) using default colour space methodology resulting in an average coverage depth of 151-fold. SRMA (0.1.15) was used to locally realign the aligned sequences around INDELs. Subsequently, all SNPs were identified in each population using Geneious (Geneious 5.4, Auckland, New Zealand) (Drummond et al. 2011). A SNP or small INDEL was called if at least 60% of the reads contained the change at sites with at least 20-fold coverage. Mutations were organized by genome position, by gene identity and by functional group. Position level parallelism was assessed by the presence of the same mutation at the identical genome position across more than a single genome. Gene level parallelism was assessed by the presence of mutations in the same gene across multiple genomes. Functional group parallelism was assessed by the presence of a mutation in any gene of a designated functional group based on classifications used in the S. pneumoniae R6 sequencing project (http://www.streppneumoniae.com). For gene level and functional group parallelism only non-synonymous SNPs and INDELs were considered because these mutations are considered to affect gene functioning and therefore are likely to be the target of natural selection. Generalized linear models for functional groups and binomial logistic regression with Firth adjustment for single loci were used to test for treatments effects in JMP 9.0.

4.4 Results

4.4.1 Oscillatory population dynamics

Ancestral populations of S. pneumoniae undergo dramatic oscillations in cell density; populations at their peak reach a density of approximately 10⁹/ml, whereas at their minimum, they are nearly three orders of magnitude lower, leaving completely clear chemostat vessels
Early in this experiment it became apparent that these dramatic chemostat clearings no longer occurred (Fig. 4.2). This was unexpected, because population oscillations remained observable in a previous study even after 300 generations of chemostat evolution (Cornejo et al. 2008). We examined density dynamics in all 16 populations at 1000 generations and found that oscillations were either lost or highly damped in all evolved chemostat populations by this time. Although we were unable to determine the mechanism by which evolved populations have stopped oscillating, our results suggest that evolved populations first evolved resistance to secreted toxin and then potentially lost the ability to produce it. This order is more likely than the reverse, because losing toxin production would confer only a limited direct advantage within a population of toxin producers, whereas resistance would provide an enormous and immediate benefit. Next, we examined the dynamics of trait loss in four focal populations. In three of four focal populations, oscillations were lost by 250 generations, while in the fourth oscillations were lost by 500 generations (Fig. 4.3). The early and parallel loss of this behaviour is consistent with its significant benefit in this evolution environment.

4.4.2 Growth rate

The disappearance of population oscillations makes growth of the populations directly re-
Figure 4.2. Population density of populations over time after each has evolved for 1000 generations. A) non-competent populations evolved in a benign environment; B) non-competent populations evolved in a stress environment; C) competent populations evolved in a benign environment; D) competent populations evolved in a stress treatment. Spotted line is the ancestor measured over the same time period. Grey shades are used to differentiate between replicate lines within a treatment group.

Figure 4.3. Long-term density dynamics of four evolved lineages. Black indicates where the populations oscillates, while white indicates where oscillations have been lost.
lated to the chemostat flow rate. Theory suggests that chemostat competitiveness depends on both growth rate and affinity for the limiting resource, which in a complex medium is unknown. Alternatively, because the maximum growth rate of our *S. pneumoniae* strains in batch culture vastly exceeds the flow rate of our chemostats, selection on growth rate may be minimal or absent. Consistent with this latter possibility, the growth rates of evolved populations in batch culture were highly variable, and increased significantly in only 2 of 16 cases, while declining in 7 others. We observed no differences between strains as a function of either competence or periodic stress (Fig. 4.4; Competence: $F_{1,15} = 1.242, \ p=0.287$; Environment: $F_{1,15} = 0.543, \ p=0.476$), nor did we observe any interactions between these treatments ($F_{1,15} = 2.241, \ p=0.160$). These inconsistent changes across populations support the idea that growth rate in the evolutionary environment was not under strong selection.

### 4.4.3 Biofilm formation

One route by which cells can increase fitness in chemostats is to adhere to the vessel wall and grow as an adhered biofilm (Larsen and Dimmick 1964). To our surprise, all but two populations evolved decreased biofilm formation after 1000 generations (Fig. 4.5A). We observed no differences in biofilm formation as a function of experimental treatments (Competence:
Parallel evolution
Chapter 4

109

Using whole-genome sequencing of evolved populations we partitioned mutations into three classes: intergenic, synonymous and non-synonymous. The total number of mutations across the genome does not differ between treatments, despite the presence of several mutators (Fig. 4.6A; ANOVA: Competence: $F_{1,15} = 1.750, p = 0.211$; Environment: $F_{1,15} = 0.472$, $p = 0.504$).
Figure 4.6. Mutations within evolved populations. A) Number of mutations in each evolved population. B) Number of mutation that occur in more than one population divided by class. C) Number of genes with a mutation in more than one population. Black are non-synonymous SNPs, dark grey are synonymous SNPs, light grey are intergenic SNPs, and white are INDELs.
p=0.505; Interaction: F_{1,15}=0.065, p=0.803), nor are there differences between treatments in the dN/dS ratio (ANOVA: Competence: F_{1,15}=0.556, p=0.470; Environment: F_{1,15}=1.105, p=0.314; Interaction: F_{1,15}=0.807, p=0.387). The mean (+/- SE) dN/dS is 1.01 (0.10), which does not differ significantly from 1 (t-test: t_{15} = 0.078, p = 0.939), despite clear evidence for positive selection from phenotypic data. The lack of an indication of positive selection from the dN/dS ratio suggests relaxed selection, but it could also be a mix of positive and purifying selection. In addition, the dN/dS ratio can be insensitive to changes between closely related individuals, which is the case in our experiment (Kryazhimskiy and Plotkin, 2008). Therefore, it might be better to limit a dN/dS analysis to specific genes. The relative distribution of mutation types is similar between populations (Fig. 4.6A), with most mutations being non-synonymous.

Genomic parallelism across replicate populations can occur across three levels of increasing resolution. First, mutations may occur at the same nucleotide site in multiple lineages. Second, mutations may occur within the same gene across several populations. Finally, functionally related mutations may occur in distinct genes that are part of related pathways or functional groups. In the next sections parallel changes across these three levels are examined. Parallel mutations at the base pair level were relatively rare. Only 3.7% (45 of 1210) of the non-synonymous mutations found across all 16 evolved lines have occurred in two or more populations, which is consistent with earlier findings in experimental populations of *E. coli* (Woods et al. 2006; Tenaillon et al. 2012). For synonymous mutations this fraction is slightly lower (2.2%; 9 of 407), which is similar to intergenic mutations (2.9%; 7 of 237). Interestingly, the fraction of parallel INDELS is somewhat higher (6.8%; 13 of 189). Most parallel mutations were found in only two populations and never in more than three for intergenic and five for synonymous mutations and INDELS (Fig. 4.6B). Given the low number of parallelism at the base pair level it is not surprising that the degree of parallelism between treatments is similar. Much less rare are parallel mutations at the gene level. More than 36% (751 of 2046) of the total number of pneumococcal genes has acquired a
Table 4.1. Summary of parallel mutations in loci. For each treatment the number of populations with at least one mutation in a gene is indicated together with the result of the binomial test. Genes above the dashed line have a mutation in ≥50% of the populations. NS= non-significant.

<table>
<thead>
<tr>
<th>locus</th>
<th>genename</th>
<th>non-competent</th>
<th>competent</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>benign</td>
<td>periodic stress</td>
<td>benign</td>
</tr>
<tr>
<td>spr1310</td>
<td>dcf</td>
<td>4</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>spr1825</td>
<td>gapA</td>
<td>4</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>spr0579</td>
<td>hko9</td>
<td>4</td>
<td>4</td>
<td>3</td>
</tr>
<tr>
<td>spr0146</td>
<td>ABC-SBP</td>
<td>4</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>spr1923</td>
<td>hypothetical</td>
<td>3</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>spr0382</td>
<td>fabF</td>
<td>4</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>spr1300</td>
<td>guaA</td>
<td>3</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>spr1777</td>
<td>rpoB</td>
<td>2</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>spr1444</td>
<td>cobQ</td>
<td>2</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>spr1654</td>
<td>hypothetical</td>
<td>3</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>spr1120</td>
<td>glnP</td>
<td>1</td>
<td>3</td>
<td>1</td>
</tr>
<tr>
<td>spr1443</td>
<td>hypothetical</td>
<td>2</td>
<td>3</td>
<td>1</td>
</tr>
<tr>
<td>spr0215</td>
<td>rpoA</td>
<td>1</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>spr1686</td>
<td>fecE</td>
<td>2</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>spr0250</td>
<td>fusA</td>
<td>0</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>spr0384</td>
<td>fabZ</td>
<td>2</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>spr1403</td>
<td>hypothetical</td>
<td>1</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>spr1584</td>
<td>hypothetical</td>
<td>0</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>spr2045</td>
<td>sphtra</td>
<td>0</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>spr0010</td>
<td>hypothetical</td>
<td>1</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>spr0214</td>
<td>rpsK</td>
<td>3</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>spr0251</td>
<td>polC</td>
<td>3</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>spr0381</td>
<td>fabG</td>
<td>0</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>spr0938</td>
<td>ABC-NBD</td>
<td>3</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>spr1062</td>
<td>ptsI</td>
<td>1</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>spr1093</td>
<td>hypothetical</td>
<td>2</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>spr1430</td>
<td>dpr</td>
<td>1</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>spr1707</td>
<td>amiA</td>
<td>1</td>
<td>3</td>
<td>0</td>
</tr>
<tr>
<td>spr1758</td>
<td>cinA</td>
<td>2</td>
<td>3</td>
<td>0</td>
</tr>
<tr>
<td>spr0006</td>
<td>mfd</td>
<td>2</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>spr0092</td>
<td>capD</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>spr0119</td>
<td>hypothetical</td>
<td>1</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>spr0245</td>
<td>glmS</td>
<td>3</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>spr0247</td>
<td>pulA</td>
<td>1</td>
<td>3</td>
<td>0</td>
</tr>
<tr>
<td>spr0516</td>
<td>pnpA</td>
<td>2</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>spr1040</td>
<td>rexA</td>
<td>1</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>spr1060</td>
<td>phpA</td>
<td>1</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>spr1060</td>
<td>phpA</td>
<td>1</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>spr1101</td>
<td>spnII-interrupt-C</td>
<td>1</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>spr1214</td>
<td>trzA</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>spr1215</td>
<td>ABC-N/P</td>
<td>1</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>spr1315</td>
<td>ABC-MSP</td>
<td>2</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>spr1525</td>
<td>ABC-MSP</td>
<td>1</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>spr1544</td>
<td>secA</td>
<td>3</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>spr1600</td>
<td>hypothetical</td>
<td>1</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>spr1643</td>
<td>ABC-MSP</td>
<td>0</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>spr1837</td>
<td>adhE</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
</tbody>
</table>

a: overrepresented in periodic stress populations
b: overrepresented in competent populations
c: overrepresented in non-competent populations
non-synonymous mutation. Parallel non-synonymous mutations at this level of resolution occurred in 28.1% (199 of 707) of cases (Fig. 4.6C). These levels are similar to previous observations in *E. coli* (Woods et al. 2006; Tenaillon et al. 2012). By contrast, only 14.2% (48 of 338) of synonymous mutations and 16% (19 of 118) INDELs were found in parallel across more than 1 population (Fig. 4.6C). Similarly, while the number of genes with synonymous mutations in more than two populations declines 5-fold (3.0%; Fig. 4.6C), nearly 11.5% of non-synonymous mutations are found in greater than two populations (Fig. 4.6C). Overall, 1.3% of genes with a non-synonymous mutation were found across eight or more of the populations, whereas no synonymous mutations or INDELs were found in parallel at this level (Fig 4.6C). Here, we also did not find that parallel changes were more abundant in a specific treatment.

The 218 cases of non-synonymous gene level parallelism (non-synonymous SNPs and INDELs) were partitioned to identify instances of treatment specific molecular evolution in this long-term experiment. We focused attention on genes that were mutated in at least 4 of the 16 populations, comprising 48 (21.6%) of the 218 cases of parallelism. Of these 48 genes, 12 had a mutation in more than 50% of the populations (table 4.1). This indicates that these genes, with the exception of *glnP* that is more likely to have gained a mutation in lineages

![Figure 4.7](image-url)  
*Figure 4.7. Schematic display of non-synonymous mutations in spr0146 and gapA. Numbers indicate the number of populations containing a mutation at that site.*
Table 4.2. Summary of parallel mutations in gene groups. For each gene group the total number of genes and the fraction of genes with a mutation are given together with the total number of genes with at least one mutation within each treatment and the results of the generalized linear model. Also the number of population in which a mutation occurred within the group is given in the final column. The dashed line indicates groups that have a gene with a mutation in > 50% of the populations. NS = non significant.

<table>
<thead>
<tr>
<th>group/subgroup</th>
<th>genes with mutations</th>
<th>non-competent</th>
<th>competent</th>
<th>test-competence</th>
<th>test-environment</th>
<th>X²</th>
<th>P-value</th>
<th>X²</th>
<th>P-value</th>
<th># populations</th>
</tr>
</thead>
<tbody>
<tr>
<td>Energy metabolism/glycolysis</td>
<td>15</td>
<td>0.53</td>
<td>4</td>
<td>11</td>
<td>6</td>
<td>1.5</td>
<td>NS</td>
<td>4.5</td>
<td>0.0347</td>
<td>16</td>
</tr>
<tr>
<td>Translation/protein modification</td>
<td>9</td>
<td>0.44</td>
<td>6</td>
<td>5</td>
<td>6</td>
<td>5</td>
<td>NS</td>
<td>0.7</td>
<td>NS</td>
<td>16</td>
</tr>
<tr>
<td>Transport proteins/ABC ATP-binding</td>
<td>71</td>
<td>0.44</td>
<td>20</td>
<td>19</td>
<td>6</td>
<td>11</td>
<td>3.1</td>
<td>NS</td>
<td>0.1</td>
<td>15</td>
</tr>
<tr>
<td>Regulatory functions/general</td>
<td>67</td>
<td>0.34</td>
<td>14</td>
<td>12</td>
<td>5</td>
<td>8</td>
<td>3.3</td>
<td>NS</td>
<td>0</td>
<td>15</td>
</tr>
<tr>
<td>Transport proteins/ABC membrane spanning permease</td>
<td>59</td>
<td>0.39</td>
<td>10</td>
<td>16</td>
<td>6</td>
<td>9</td>
<td>1.2</td>
<td>NS</td>
<td>0.8</td>
<td>15</td>
</tr>
<tr>
<td>Transport proteins/ABC substrate binding</td>
<td>32</td>
<td>0.53</td>
<td>9</td>
<td>19</td>
<td>7</td>
<td>9</td>
<td>4.1</td>
<td>0.0434</td>
<td>4.1</td>
<td>0.0434</td>
</tr>
<tr>
<td>Fatty acid metabolism/general fatty acid</td>
<td>22</td>
<td>0.45</td>
<td>7</td>
<td>9</td>
<td>4</td>
<td>6</td>
<td>2.7</td>
<td>NS</td>
<td>1.2</td>
<td>15</td>
</tr>
<tr>
<td>Energy metabolism/other</td>
<td>71</td>
<td>0.45</td>
<td>17</td>
<td>18</td>
<td>8</td>
<td>14</td>
<td>1.8</td>
<td>NS</td>
<td>0.7</td>
<td>14</td>
</tr>
<tr>
<td>Purines and pyrimidines/purine biosynthesis</td>
<td>18</td>
<td>0.50</td>
<td>4</td>
<td>7</td>
<td>3</td>
<td>4</td>
<td>0.8</td>
<td>NS</td>
<td>0.8</td>
<td>12</td>
</tr>
<tr>
<td>Cell envelope/murein sacculus</td>
<td>26</td>
<td>0.58</td>
<td>5</td>
<td>12</td>
<td>3</td>
<td>2</td>
<td>1.9</td>
<td>NS</td>
<td>0.5</td>
<td>10</td>
</tr>
<tr>
<td>Transport proteins/PTS transporter</td>
<td>39</td>
<td>0.36</td>
<td>2</td>
<td>12</td>
<td>4</td>
<td>3</td>
<td>0.8</td>
<td>NS</td>
<td>1.3</td>
<td>10</td>
</tr>
<tr>
<td>Cell envelope/surface polysaccharides</td>
<td>32</td>
<td>0.38</td>
<td>3</td>
<td>12</td>
<td>2</td>
<td>5</td>
<td>1.1</td>
<td>NS</td>
<td>3.4</td>
<td>10</td>
</tr>
<tr>
<td>Cellular processes/competence</td>
<td>20</td>
<td>0.45</td>
<td>4</td>
<td>6</td>
<td>3</td>
<td>2</td>
<td>2.2</td>
<td>NS</td>
<td>0.1</td>
<td>10</td>
</tr>
<tr>
<td>Energy metabolism/sugars</td>
<td>59</td>
<td>0.44</td>
<td>9</td>
<td>17</td>
<td>3</td>
<td>2</td>
<td>2.6</td>
<td>NS</td>
<td>0.3</td>
<td>9</td>
</tr>
<tr>
<td>Other/transposase</td>
<td>101</td>
<td>0.16</td>
<td>4</td>
<td>5</td>
<td>3</td>
<td>6</td>
<td>0</td>
<td>NS</td>
<td>0.6</td>
<td>9</td>
</tr>
<tr>
<td>Translation/ribosomal proteinalse</td>
<td>56</td>
<td>0.23</td>
<td>12</td>
<td>3</td>
<td>2</td>
<td>1</td>
<td>6.3</td>
<td>0.0118</td>
<td>4.7</td>
<td>0.0309</td>
</tr>
<tr>
<td>Replication/DNA replication</td>
<td>25</td>
<td>0.48</td>
<td>11</td>
<td>8</td>
<td>0</td>
<td>2</td>
<td>7.2</td>
<td>0.0073</td>
<td>0</td>
<td>8</td>
</tr>
<tr>
<td>Translation/transcription factor</td>
<td>19</td>
<td>0.47</td>
<td>7</td>
<td>8</td>
<td>0</td>
<td>3</td>
<td>4.7</td>
<td>0.0309</td>
<td>0.2</td>
<td>8</td>
</tr>
<tr>
<td>Central intermediary metabolism/other</td>
<td>21</td>
<td>0.38</td>
<td>9</td>
<td>4</td>
<td>2</td>
<td>1</td>
<td>4.7</td>
<td>0.0309</td>
<td>1.8</td>
<td>8</td>
</tr>
<tr>
<td>Translation/amino acyl tRNA synthetase</td>
<td>27</td>
<td>0.19</td>
<td>2</td>
<td>5</td>
<td>1</td>
<td>1</td>
<td>6.3</td>
<td>0.0122</td>
<td>2.6</td>
<td>8</td>
</tr>
<tr>
<td>Translation/protein modification and translation</td>
<td>10</td>
<td>0.30</td>
<td>2</td>
<td>3</td>
<td>0</td>
<td>5</td>
<td>0</td>
<td>NS</td>
<td>6.5</td>
<td>0.0109</td>
</tr>
<tr>
<td>Energy metabolism/fermentation</td>
<td>16</td>
<td>0.50</td>
<td>1</td>
<td>6</td>
<td>2</td>
<td>2</td>
<td>0.5</td>
<td>NS</td>
<td>1.2</td>
<td>7</td>
</tr>
<tr>
<td>Cellular processes/protein and peptide secretion</td>
<td>11</td>
<td>0.55</td>
<td>4</td>
<td>3</td>
<td>0</td>
<td>2</td>
<td>2.9</td>
<td>NS</td>
<td>0.1</td>
<td>7</td>
</tr>
<tr>
<td>group/subgroup</td>
<td># genes with mutations</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>P-value</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>----------------------------------------------------</td>
<td>------------------------</td>
<td>-------</td>
<td>-------</td>
<td>-------</td>
<td>------</td>
<td>---------</td>
<td>-------</td>
<td>-------</td>
<td>-------</td>
<td>-------</td>
</tr>
<tr>
<td></td>
<td>total fraction</td>
<td>genes</td>
<td>non-competent</td>
<td>competent</td>
<td>test-competence</td>
<td>test-environment</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Purines and pyrimidines/pyrimidine biosynthesis</td>
<td>15</td>
<td>0.40</td>
<td>4</td>
<td>3</td>
<td>1</td>
<td>3.3</td>
<td>0.6</td>
<td>NS</td>
<td>0.1</td>
<td>NS</td>
</tr>
<tr>
<td>Purines and pyrimidines/nucleotide interconversions</td>
<td>13</td>
<td>0.46</td>
<td>1</td>
<td>5</td>
<td>2</td>
<td>1.2</td>
<td>1.2</td>
<td>NS</td>
<td>1.2</td>
<td>NS</td>
</tr>
<tr>
<td>Translation/degradationation proteins, peptides</td>
<td>17</td>
<td>0.53</td>
<td>2</td>
<td>8</td>
<td>1</td>
<td>1.7</td>
<td>1.7</td>
<td>NS</td>
<td>1.7</td>
<td>NS</td>
</tr>
<tr>
<td>Amino acid biosynthesis/aspartate family</td>
<td>17</td>
<td>0.47</td>
<td>4</td>
<td>3</td>
<td>1</td>
<td>1.7</td>
<td>1.7</td>
<td>NS</td>
<td>1.7</td>
<td>NS</td>
</tr>
<tr>
<td>Amino acid biosynthesis/aromatic amino acid</td>
<td>16</td>
<td>0.50</td>
<td>3</td>
<td>4</td>
<td>1</td>
<td>0.6</td>
<td>0.6</td>
<td>NS</td>
<td>0.6</td>
<td>NS</td>
</tr>
<tr>
<td>Replication/restriction, modification</td>
<td>16</td>
<td>0.50</td>
<td>6</td>
<td>6</td>
<td>0</td>
<td>3.3</td>
<td>3.3</td>
<td>NS</td>
<td>3.3</td>
<td>NS</td>
</tr>
<tr>
<td>Replication/recombination</td>
<td>12</td>
<td>0.58</td>
<td>3</td>
<td>5</td>
<td>2</td>
<td>0.1</td>
<td>0.1</td>
<td>NS</td>
<td>0.1</td>
<td>NS</td>
</tr>
<tr>
<td>Cellular processes/cell division</td>
<td>14</td>
<td>0.43</td>
<td>2</td>
<td>3</td>
<td>2</td>
<td>0.1</td>
<td>0.1</td>
<td>NS</td>
<td>0.1</td>
<td>NS</td>
</tr>
<tr>
<td>Translation/translational, tRNA modification</td>
<td>14</td>
<td>0.43</td>
<td>1</td>
<td>3</td>
<td>1</td>
<td>4.1</td>
<td>4.1</td>
<td>NS</td>
<td>4.1</td>
<td>NS</td>
</tr>
<tr>
<td>Energy metabolism/ATP-proton-motive force</td>
<td>11</td>
<td>0.27</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>5.4</td>
<td>5.4</td>
<td>0.0037</td>
<td>0.0037</td>
<td>6</td>
</tr>
<tr>
<td>Replication/repair</td>
<td>17</td>
<td>0.41</td>
<td>2</td>
<td>3</td>
<td>0</td>
<td>0.4</td>
<td>0.4</td>
<td>NS</td>
<td>0.4</td>
<td>NS</td>
</tr>
<tr>
<td>Energy metabolism/pentose phosphate pathway</td>
<td>11</td>
<td>0.55</td>
<td>2</td>
<td>4</td>
<td>1</td>
<td>1.1</td>
<td>1.1</td>
<td>NS</td>
<td>1.1</td>
<td>NS</td>
</tr>
<tr>
<td>Amino acid biosynthesis/glutamate family</td>
<td>7</td>
<td>0.71</td>
<td>2</td>
<td>5</td>
<td>0</td>
<td>1.7</td>
<td>1.7</td>
<td>NS</td>
<td>1.7</td>
<td>NS</td>
</tr>
<tr>
<td>Transport and binding proteins/cation</td>
<td>11</td>
<td>0.45</td>
<td>2</td>
<td>0</td>
<td>1</td>
<td>2.0</td>
<td>2.0</td>
<td>NS</td>
<td>2.0</td>
<td>NS</td>
</tr>
<tr>
<td>Amino acid biosynthesis/serine family</td>
<td>8</td>
<td>0.38</td>
<td>1</td>
<td>3</td>
<td>0</td>
<td>4.0</td>
<td>4.0</td>
<td>0.0449</td>
<td>0.0449</td>
<td>1.1</td>
</tr>
</tbody>
</table>

a: fraction of the total genes in a group that has a mutation in at least one of the populations
b: gene groups with a highly parallel gene
c: gene groups with mutations overrepresented in populations that encountered periodic stress
d: gene groups with mutations overrepresented in non-competent populations
e: gene groups with mutations overrepresented in non-competent populations that encountered periodic stress
f: gene groups with mutations overrepresented in non-competent populations evolved in the benign environment
g: gene groups with mutations overrepresented in competent populations that encountered periodic stress
evolved in the periodic stress environment, have been under general selection for the chemostat environment. In total 10 cases could be identified where mutations were significantly overrepresented in either the stress or competence treatments: 2 genes, *fusA* and *glnP*, were significantly more likely to be mutated in the stressed than unstressed populations, while seven genes, *polC, cinA, glmS, secA, pulA, mfd* and *spr1315*, were more likely to be mutated in parallel in non-competent lineages and one gene, *spr1643*, more likely to be mutated in parallel in competent lineages (table 4.1). In no case did we observe more parallel mutations than expected in strains from unstressed treatments. From these 48 genes, two genes, *gapA* and the putative ABC-transporter binding protein *spr0146*, particularly stand out at the nucleotide level. Whereas *spr 0146* had extensive parallelism at the nucleotide level, *gapA* on the other hand has multiple mutations in 12 of 16 populations that are spread across several sites in the gene. The contrasting cases of both genes are shown in detail in figure 4.7.

Since most genes operate as a part of pathways and functional groups, it is possible that similar phenotypic changes might result from changes in different genes of the pathway or functional group. Therefore we next identified parallel genotypic changes occurring at the level of functional groups based on annotations from the R6 genome. As observed with parallel changes at the gene level, we find extensive functional parallelism in 39 of 64 functional groups across all treatments. Genes had a mutation in more than 50% of the evolved populations for 17 of 39 (43.6%) gene groups (table 4.2). This type of parallelism could be the result of many mutations in a small number of genes, such as the highly parallel genes from table 4.1, within the functional group. However, on average 44.4% (95% confidence interval: 40.8-48.1%) of the genes in a functional group, which contains more than 3 genes with a mutation, have a mutation in at least one of the evolved lines. This shows that mutations were often dispersed among genes within a functional group. However, for some of these groups, such as glycolysis, general regulatory functions and protein modifications, high levels of parallelism are observed because generally genes containing highly parallel mutations are part of these groups. To overcome this bias from highly parallel genes within a
group, total numbers of mutations per population within a group were considered for further analysis, rather than the absence or presence of mutation in a population.

Of the 39 functional groups with genes mutated in parallel, ten (25.6%) show treatment specific parallelism. Five functional groups were overrepresented among non-competent populations: serine biosynthesis; DNA replication; transcription factors; amino actyl tRNA synthase and modification; and central intermediary metabolism (table 4.2). Parallel changes in two functional groups, glycolosis, and protein modification and translation, were overrepresented in the populations exposed to periodic stress (table 4.2). Mutations for one group, ATP-proton motive force interconversion, were more often found in competent populations evolved in the periodic stress environment (table 4.2). Whereas, mutations in ribosomal proteins were significantly more often found in non-competent populations evolved in a benign environment, mutations in substrate binding ABC transporters are significantly more often found in non-competent populations evolved in the periodic stress environment (table 4.2).

4.5 Discussion
The aim of this work was to combine phenotypic and genotypic analyses in order to quantify and characterize parallel changes occurring in evolving populations of *S. pneumoniae*. As part of this, we also sought to distinguish parallel changes occurring in all or most evolved populations from those that were specific to the two treatment conditions, competence and exposure to periodic stress. We find extensive evidence for phenotypic and genetic parallelism, both generally and more specifically for treatment groups. Of the three phenotypic traits examined in this study we observed parallel reductions in biofilm formation and the loss of oscillating population densities. By contrast, changes in growth rate were idiosyncratic, suggesting that this trait evolved by genetic drift. Examination of genomic evolution of evolved *S. pneumoniae* shows that, like in evolved *E. coli* (Woods et al. 2006; Tenaillon et al. 2012), parallel changes at the base pair level are very rare, while parallel changes at the gene and gene group level are quite common. This follows the general prediction that parallelism will
increase with decreasing levels of genomic resolution.

The oscillation of *S. pneumoniae* population densities in chemostats was first described by Cornejo et al. (2008). These oscillations were the result of toxin production, which killed a large fraction of the growing population before the toxin concentration declined due to chemostat dilution. Interestingly, in pilot studies, Cornejo et al. (2008) observed that oscillations persisted for at least 300 generations of chemostat evolution, which suggested that toxin production or immunity was difficult to lose owing to trade-offs with other aspects of bacterial fitness. This was assumed because of the tremendous advantage that a non-cycling mutant would have in an otherwise wild-type cycling population of cells. However, in our study elimination of the oscillating phenotype occurred in parallel across all 16 chemostats within the first half of the experiment (Fig. 4.2). Although, we were unable to determine the mechanism behind this change it seems most likely that mutations causing toxin immunity would emerge before loss of toxin production. This is because an immune mutant would have a massive and immediate advantage, while a toxin non-producer would only benefit by losing any costs associated with toxin production. Results from Cornejo et al (2008) found evidence supporting the possibility that the toxin was a small secreted peptide. Thus in addition to immunity, changes to toxin production and/or excretion mechanism can potentially lead to reduced oscillation phenotypes.

A second parallel phenotypic change seen in all populations is the reduction of biofilm formation (Fig. 4.5). Wall growth is a well-known response of bacteria maintained in chemostats (Larsen and Dimmick 1964), especially those maintained at high flow rates. Cells in these biofilms form a second population within the chemostat that escapes being washed out with the chemostat flow. This sub-population can reduce the rate of selection if migrants from the biofilm are either wild-type cells lacking new beneficial mutations or if they contain mutations that are harmful during planktonic growth (Dykhuizen and Hartl 1983). Despite possible benefits that biofilm forming cells might have experienced in this experiment, we
instead observed the near uniform loss of this capability, although to different degrees. The fact that we observed consistent losses of biofilms in these populations indicates that this phenotype is adaptive in the chemostat environment, although the mechanism by which any advantages are realized remains unclear.

Besides parallel phenotypic changes across treatments we also observed parallel genetic changes that were limited to certain treatments. Whereas the first class of changes are consistent with general benefits to all evolving populations, some mutations in genes and gene groups are overrepresented only in specific treatments (table 4.1 and 4.2) indicating treatment specific selection on these loci. Both sets of genes are strong candidates for beneficial mutations in this experiment and will be examined in greater detail in future studies. In particular, it will be crucial to generate isogenic lineages carrying only a single mutation in order to distinguish direct benefits from mutations that have become fixed due to hitchhiking or genetic drift. Thus, at present, our assessment of the functional and evolutionary role of any single fixed mutation is conjecture.

At the level of the individual nucleotide, parallelism (mutation in ≥2 populations) was rare and was observed in only 45 of 1210 non-synonymous mutations (Fig. 4.6B). A particularly interesting candidate gene among this set is the ABC-transporter binding protein spr0146. Spr0146 has a non-synonymous point mutation (A -> G) at position 538 in all but one population. ABC-transporters are used to regulate the secretion or import of proteins, peptides or other small molecules. As this gene in pneumococcus is as yet uncharacterized it is not possible to determine its substrate. However, an intriguing possibility is that this mutation is causally involved in either the import or export of the toxin causing population oscillations in chemostats. Mutations preventing import could potentially impart immunity, while changes influencing export could lead to the loss of toxin production. A second possibility is that this mutation increases the transport of nutrients into the bacterial cell, thereby facilitating growth. An important aim of future work is to directly test these possibilities using
Parallel evolution

At the level of a locus or functional group, there are many instances of both treatment specific and non-specific parallelism. Twelve loci, of which eleven are not treatment specific, were identified with a non-synonymous mutation in at least eight of the sixteen evolved populations. Among these candidate genes is \textit{gapA}, which often has multiple mutations per population (Fig. 4.7). In addition, several of these mutations have occurred in parallel at the base pair location across populations. \textit{GapA} encodes GAPDH, which is an important part of the glycolysis pathway (Willemoes et al. 2002). Mutations in \textit{gapA} may influence nutrient accessibility or processing in the chemostat environment. Another interesting locus is the histidine kinase \textit{hk09}. Histidine kinases are transmembrane signalling proteins that are part of two component systems (TCS). Two component systems are key systems for a bacterium to perceive and respond to environmental signals. \textit{Hk09} is part of TCS 09, which in \textit{S. pneumoniae} is an important virulence factor (Blue and Mitchell 2003). The impact of TCSs on virulence depends on strain and infection site (Paterson et al. 2006). It is unlikely, though, that selective pressures on virulence were part of the evolutionary environment. This implies that either TCS09 is very costly to maintain, which suggests that these parallel mutations are knockouts, or, more likely, TCS09 has additional functions with in the pneumococcal cell. Both hypotheses will require further detailed study of the mutations in the evolved populations.

At the level of gene functional group, 17 of 39 functional groups contained mutations across more than 50% of the populations. Several of these groups show strong signals of parallelism owing to parallel mutations in a single gene within the functional group. In others, mutations in the group are more widespread, targeting different genes across populations. Changes in the three different protein groups of ABC-transporters are found in almost all populations (table 4.2). Each of these groups constitutes a specific part of an ABC-transporter. ABC-transporters are highly important for the uptake of nutrients because of their
import/export function. That many mutations have taken place in the different elements of these transporters is a strong indication that all populations have adapted to nutrient availability in the evolutionary environment. This is supported by the observation that mutations in groups of genes influencing energy metabolism are also highly parallel across treatments.

In the light of phenotypic changes in this study, highly parallel changes in genes affecting the cell envelope, i.e. murein sacculus, and surface polysaccharides, lyposaccharides and antigens, are of interest. The murein sacculus, or cell wall, is often the target of lysogenic agents and antibiotics (Lopez et al. 1997; Koch 2003). Changes in this cell structure could have contributed to the reduced oscillation of population densities observed in this study by generating immunity against the toxin. Surface polysaccharides are important for the formation of pneumococcal biofilms (e.g. Lopez 2006). The many changes in this group could be linked to our observed reduction of biofilm formation in the evolved populations. The high mutation frequency in these groups suggests that these have been targets of selection.

Overall ten loci were identified with a mutational bias in a specific treatment. Changes in seven genes were significantly associated with non-competent populations. In contrast, mutations in only one gene were significantly overrepresented in competent populations. This difference is striking and might be explained by the observation that competence reduces the mutational load in \textit{S. pneumoniae} (Chapter 3). If competence mostly repairs lesions then it is not unsurprising that occasionally beneficial mutations are ‘repaired’ too, which according to these results reduces the amount of parallel evolution in competent lineages. It is possible that mutations in these genes could be changes in the regulations of these genes as the result of disabling the competence system. However, only two of the genes, \textit{cinA} and \textit{pulA}, are directly affected by competence induction (Peterson et al. 2004). Two other loci that have mutations predominantly in non-competent populations are \textit{mfd} and \textit{polC}. These genes are part of DNA-repair pathways. Together with the mutation bias in the DNA-replication gene group, which we discuss later, these loci can likely explain the increased mutation rate and mutator phenotypes found in the non-competent populations (Chapter 3).
At the per locus level there is less treatment specific mutational bias as a function of the stress treatment than the competence treatment. Two genes are overrepresented in populations exposed to periodic stress, whilst none are biased in the populations evolved in the benign environment. One of the genes that are overrepresented in populations evolved with periodic stress is *fusA*, which is associated with fusidic acid resistance in *S. pneumoniae* (Belanger et al. 2002). In other species fusidic acid resistance can lead to low levels of cross-resistance to kanamycin (Johanson and Hughes 1994; Price and Gustafson 2001). Although no resistance to kanamycin has been observed in the evolved populations it is possible that mutations in *fusA* lead to better growth in the presence of the sub-MIC concentrations of kanamycin that were used to introduce mild stress in our experiment. This response would be notable because the periods of kanamycin exposure were relatively short and infrequent, and the concentrations of drug very low. Gullberg et al (2011) have recently shown that antibiotic resistance can arise when cells are confronted with very low drug concentrations conferring significant benefits to the resistant cells. Tests of this possibility using mutant genotypes of *fusA* in the presence and absence of kanamycin can be used to examine this possibility.

At the gene functional group level, five groups were found to be overrepresented in populations that lacked competence while two were overrepresented in populations exposed to periodic stress. Three groups were overrepresented in particular treatment groups. Of particular interest are mutations in genes involved in DNA replication, which are overrepresented in non-competent populations. The mutations in this group likely explain the observed mutator phenotypes (Chapter 3) that have arisen in evolving non-competent populations. Many of the genes in this group are associated with mutator phenotypes in *E. coli* (Horst et al. 1999).

Our results reveal the power of combining experimental evolution with whole-genome sequencing to understand bacterial evolution. We observed extensive parallelism in our populations, some which are specific to certain treatments and others that occur generally in all
populations irrespective of treatment or background genotype. The capacity to generate genetic diversity, owing to the presence or absence of competence, does not limit the degree to which parallel changes arise in these populations; the chemostat environment in this experiment is thus sufficiently general in its influence on cells that it selects for parallel phenotypic and genotypic changes. Periodic stress also does not inhibit general change. In spite of this generality, we also observe several parallel genomic changes that are treatment specific. Clarifying the direct effects of these treatment specific mutations may help to understand the effects each treatment has on bacterial populations. Many outstanding questions remain that we will hope to address in future work using this system. Most importantly, we have only obtained sequences from terminal isolates, which prevents us to examine mutational dynamics in these populations. New sequencing approaches that survey all the segregating SNPs in each population through time remain prohibitively expensive, but it is likely that such a detailed approach will be possible in the near future. With such tools, we will be able to ask more detailed questions about the order and effect size of fixed mutations. For example, is the effect on fitness of across treatment parallel changes larger than those within a specific treatment? If so, do global parallel changes become fixed in populations before the treatment specific mutations occur? These are important questions that can be answered with continuation of the experiment described in this study.

In summary, we show that in a long term evolution experiment with *S. pneumoniae* parallel changes occur both phenotypically and genetically. The likelihood of identifying parallel changes increases as we move from the high-resolution analysis of single nucleotide sites to analyses focusing on parallel changes at the same locus or functional group. We observe several parallel changes that are specific to a treatment. Further analysis of these candidate genes using isogenic constructs offers a novel route towards understanding bacterial evolution in the context of recombination and stress.

4.6 References


Chapter 5

Testing the Pharaoh’s curse: surface survival and virulence in *Streptococcus pneumoniae*

*Daniel J.P. Engelmoer, Sarah Chung & Daniel E. Rozen*

5.1 Abstract

Classical theory predicts that pathogen virulence will evolve to an intermediate level because of trade-offs between pathogen reproduction and transmission. However, for pathogens that are transmitted indirectly, e.g. via survival infectious particles in the environment, that do not rely on host mobility for transmission virulence can evolve to higher levels. This “sit-and-wait strategy” called the Pharaoh hypothesis, predicts that survival duration outside of the host should be positively correlated with virulence. Several species are believed to follow this strategy, which includes the gram-positive bacterial pathogen *Streptococcus pneumoniae*. Here we have tested the main prediction of the Pharaoh hypothesis using several genetically distinct wild type strains of *S. pneumoniae* that are diverged in the expression of their surface capsular serotype. As predicted, we find a significant, although weak, correlation between virulence and durability outside of the host. We also find an inverse relationship between durability and the chance that a serotype causes infection. Both results indicate the coexistence of diverse transmission strategies within populations of *S. pneumoniae*. 
5.2 Introduction

Classical evolutionary theory predicts that natural selection will cause pathogen virulence, defined as the effect on host mortality, to evolve towards intermediate levels because of a trade-off between within host pathogen reproduction and pathogen transmission (Anderson and May 1982; Alizon et al. 2009). However, if pathogen transmission is indirect, e.g. via the environment, then both virulence and transmission can evolve to their maximum values because transmission is independent of host survival or mobility (Bonhoeffer et al. 1996; Gandon 1998; Walther and Ewald 2004). One mechanism of indirect pathogen transmission is through fomites, i.e. environmental surfaces that serve as a vector of pathogen transmission, such as clothing, toys, or hospital bedding (Bean et al. 1982; Boone and Gerba 2007; Desai et al. 2011). Fomites are capable of transmitting a pathogen for as long as the infectious agent remains viable outside of a host. Accordingly, pathogens that are transmitted by this route can simply wait patiently until they encounter a susceptible host. This pathogen strategy has been dubbed the “sit-and-wait” strategy, or alternatively the Pharaoh hypothesis in reference to the mysterious death of Lord Carnavon in Egypt, supposedly caused by a pathogen that had survived in the tomb of Tutankhamen (Corelli 1923).

The Pharaoh hypothesis posits that pathogens that can survive outside of their host for extended periods of time become less dependent on host availability (Bonhoeffer et al. 1996; Gandon 1998). As a result they are no longer subject to the direct transmission-virulence trade-off, which allows these pathogens to evolve high virulence without reducing the efficacy of transmission (Ewald 1993; Walther and Ewald 2004). However, these models show that a positive correlation between transmission and virulence is limited to situations where there is at least a short period of survival outside the host before transmission, and where the surviving population of pathogens on fomites is relatively small compared to the size of the host population (Bonhoeffer et al. 1996; Gandon 1998). Additionally, increased competition between pathogens due to multiple infections enhances the positive correlation between survival and virulence (Gandon 1998).
Both the traditional model for the evolution of virulence and the Pharaoh hypothesis assume that pathogen populations evolve towards a single strategy: either one of intermediate virulence in situations of direct transmission, or high virulence in the case of indirect transmission. However, in situations where pathogens exploit both transmission modes multiple strategies can evolve and coexist (Roche et al. 2011). *Vibrio cholerae* and avian influenza are examples of pathogens that have evolved mixed transmission strategies (King et al. 2008; Rohani et al. 2009), and that maintain extensive polymorphism in the degree of virulence. It remains uncertain if in addition, these pathogens are polymorphic for environmental durability, and moreover if durability is positively correlated with virulence. Thus to our knowledge the key prediction of the mixed-model version of the Pharaoh hypothesis has not been tested.

One pathogen that is likely to rely on a mixed transmission strategy is the gram-positive opportunistic pathogen *S. pneumoniae*. Virulence levels in *S. pneumoniae* are highly diverse (Hausdorff et al. 2000; Brueggemann et al. 2003; Kronenberg et al. 2006; Sleeman et al. 2006; Harboe et al. 2009; Weinberger et al. 2010). The extracellular polysaccharide capsule (serotype) is considered the most important of several factors determined to be important for virulence,(Brueggemann et al. 2003). To date more than 90 different serotypes have been described (Hava et al. 2003). While some serotypes are found predominantly associated with asymptomatic carriage within the human nasopharynx others are most often associated with invasive pneumococcal disease (IPD) (Sandgren et al. 2004; Sjostrom et al. 2006). At the same time, *S. pneumoniae* can persist outside of the host on surfaces where it remains infectious (Williams and Kauffman 1978; Mazzola et al. 2003; Walther and Ewald 2004; Kramer et al. 2006; Walsh and Camilli 2011). Although the possibility that pneumococci are transmitted via fomites has been known since the classic study of Hodges and MacLeod (Hodges and MacLeod 1946) (1946), which provided conclusive evidence that pneumococci survive readily in the dust of military barracks, the frequency and effects of indirect trans-
mission are often neglected. Pneumococcal environmental durability and its perceived high virulence have led to the suggestion that *S. pneumoniae* has evolved a sit-and-wait strategy (Walther and Ewald 2004). However, this conclusion fails to consider pneumococcal variation in either virulence or surface survival. Our aim here is to assess the relationship between durability of different bacterial strains from different serotypes with varying virulence. In addition, in light of the fact that this species is a commensal species more often associated with asymptomatic carriage than disease, we examined the relationship between durability and aspects of carriage.

Briefly, we identified significant variation in durability between strains and serotypes of *S. pneumoniae*. We find a weak positive relationship between virulence, the chance of death when ill, and durability, consistent with predictions of the Pharaoh hypothesis. By contrast, we find an inverse relationship between invasiveness, which is the chance of contracting IPD from carriage, and durability. These results are consistent with the suggestion of Roche et al. (2011) that pathogens relying on mixed transmission modes can evolve multiple coexisting strategies of virulence and transmission.

### 5.3 Material & Methods

#### 5.3.1 Strains and culture conditions

Strains in this study have been obtained from a collection of carriage isolates taken from children in day-care centres in the Netherlands except for strain B69, which was collected in Atlanta, US (Table 5.1; Bogaert et al. 2006). All strains were grown in CTM pH = 6.8 (Complete Transformation Medium, which per litre contains: 30g tryptic soy broth (Oxoid Ltd) and 1g yeast extract (Melford Laboratories Ltd)) and stored in the same medium with 25% vol/vol glycerol at -80°C.

#### 5.3.2 Quantifying durability

A dense lawn of cells was created by dispersing a single colony of each strain using a mois-
tened cotton swab on a Tryptone Soy agar plate (LabM ltd) supplemented with 3% vol/vol defibrinated horse blood (Oxoid ltd). After 24 hours growth, all the cells were harvested with a sterile cotton swab and resuspended in CTM. Cell suspensions were then diluted to an OD_{600} = 0.3, which corresponds to a density of approximately 3 \times 10^8 cells/mL. Cells suspensions were then concentrated 10-fold by centrifuging at maximum speed (14,800 rpm/min) for 2 minutes and re-suspending the pellets in fresh medium. Horse blood was added to the cell suspension to a final concentration of 5% vol/vol. For each of seven time points, 10μL drops of each cell suspension (approximately 3 \times 10^7 cells) were spotted in a flat bottom 96-well micro-titer plate with 3-fold replication, avoiding the outer wells. Plates were left to dry at room temperature for an hour after which the initial time point was sampled. Plates were then left at room temperature in the dark and destructively sampled after 1, 2, 5, 10,

<table>
<thead>
<tr>
<th>Strain</th>
<th>Serotype</th>
<th>MLST type</th>
</tr>
</thead>
<tbody>
<tr>
<td>605</td>
<td>1</td>
<td>5319</td>
</tr>
<tr>
<td>986</td>
<td>1</td>
<td>6664</td>
</tr>
<tr>
<td>B69</td>
<td>1</td>
<td>228</td>
</tr>
<tr>
<td>110</td>
<td>6B</td>
<td>6629</td>
</tr>
<tr>
<td>280</td>
<td>6B</td>
<td>176</td>
</tr>
<tr>
<td>805</td>
<td>6B</td>
<td>6624</td>
</tr>
<tr>
<td>25</td>
<td>9v</td>
<td>644</td>
</tr>
<tr>
<td>565</td>
<td>9v</td>
<td>644</td>
</tr>
<tr>
<td>892</td>
<td>9v</td>
<td>5307</td>
</tr>
<tr>
<td>3</td>
<td>14</td>
<td>6642</td>
</tr>
<tr>
<td>345</td>
<td>14</td>
<td>132</td>
</tr>
<tr>
<td>607</td>
<td>14</td>
<td>132</td>
</tr>
<tr>
<td>813</td>
<td>19A</td>
<td>6665</td>
</tr>
<tr>
<td>190</td>
<td>19A</td>
<td>5314</td>
</tr>
<tr>
<td>1533</td>
<td>19A</td>
<td>6211</td>
</tr>
<tr>
<td>310</td>
<td>19F</td>
<td>6662</td>
</tr>
<tr>
<td>458</td>
<td>19F</td>
<td>644</td>
</tr>
<tr>
<td>1017</td>
<td>19F</td>
<td>6682</td>
</tr>
<tr>
<td>7</td>
<td>23F</td>
<td>439</td>
</tr>
<tr>
<td>542</td>
<td>23F</td>
<td>392</td>
</tr>
<tr>
<td>907</td>
<td>23F</td>
<td>6598</td>
</tr>
</tbody>
</table>

Table 5.1. Strains used in this study with serotype and MLST information. The strains were collected from day-care centres in the Netherlands by Bogaert et al. (2006), except for strain B69 which was collected in Atlanta, US.
15 and 21 days. Samples were taken by adding 0.8% saline solution to each well and vigorously pipetting up and down until all the dried blood was resuspended. Samples were diluted and 10µL drops were spotted in 3-fold at each dilution on blood plates to determine Colony Forming Units (CFU) after 48-hours incubation at 37°C +5% CO₂. Durability was calculated as the logarithm of the average decline of CFU per day. Differences between durability for serotypes were determined using a restricted maximum likelihood mixed effects model in R (v2.13.0) with package LME4. Two models were made, using the lmer function in R. One consisting where durability is predicted by both the fixed factor serotype and the random factor strain and one model where durability is only predicted by the random factor strain. The models were then compared in a log-likelihood method to determine the individual effects of serotype and strain on the mixed model using the anova function.

5.3.3 Durability and pneumococcal life history

Although *S. pneumoniae* lives predominantly as a commensal species during asymptomatic carriage within the human nasopharynx, it also can produce a range of acute invasive infections (IPD). Here we determined the relationships between durability and both of these aspects of pneumococcal life history. Serotype specific behavior during carriage was partitioned into carriage duration (i.e. the length of time a given serotype is carried), serotype specific invasiveness (i.e. the chance that a given serotype will cause IPD, the chance of acquisition of each serotype given a naïve or previously colonized host and the serotype-specific odds-ratio of infection (i.e. the likelihood of IPD given carriage prevalence). Relevant data were taken from Sleeman (2006) and Brueggemann (2003), both analyzing population biology of pneumococcal carriage and disease among children in the UK. Serotype-specific virulence data were taken from a meta-analysis (Weinberger et al. 2010) and a more recent study from Inverarity (2011). All analyses were carried out using IBM’s SPSS statistics 19.0 package.
5.4 Results

5.4.1 Durability depends on genotype and serotype

We compared the durability on plastic surfaces between different serotypes using a mixed model with strains as a random nested factor within serotypes. We found a significant difference between the seven serotypes for (Fig. 5.1; log-likelihood ratio: $X^2=13.689$, df =6, $P=0.0166$). This suggests that differences in surface survival are dependent on serotype. In addition we also observed significant variation in durability between strains within a serotype (log-likelihood ratio: $X^2=41.772$, df =1, $P < 0.0001$). These results suggest that serotype and genetic background jointly influence the durability of each strain.

5.4.2 Durability, carriage and virulence

To investigate if *S. pneumoniae* follows a high virulence – high durability strategy predicted by the Pharaoh-hypothesis we examined the relationship between durability on plastic surfaces grouped by serotype.
faces and virulence. Serotype-specific virulence is defined as the risk ratio of death 30 days post infection with a given serotype. Data were taken from a meta-analysis combining the results from several independent studies globally (Weinberger et al. 2010) and a more recent study focusing on invasive pneumococcal disease in Scotland (Inverarity et al. 2011). Although the relationship between virulence and durability is positive in both cases, the correlation is only marginally non-significant in the Scottish study (Fig. 5.2A; Pearson correlation: $r=0.409$, $P=0.066$, $N=21$) and not significant in the Weinberger study (Fig. 5.2A; Pearson correlation: $r=0.243$, $P=0.289$, $N=21$). When combined into a single analysis we found stronger, but still only marginally significant support for a positive relationship between durability and virulence (ANCOVA: $F_{1,39}=4.35$, $P=0.044$). These results thus provide moderate support for the central prediction of the Pharaoh hypothesis that \textit{S. pneumoniae} follows a sit-and-wait strategy.

Having identified evidence of a relationship between surface survival and virulence in \textit{S. pneumoniae}, we next examined whether survival is correlated with pneumococcal acquisition, the duration of carriage once acquired and the probability of causing disease given carriage. If \textit{S. pneumoniae} relies on a sit-and-wait transmission strategy, we predict that more durable strains will be more likely to be acquired and will be carried for longer duration. However, we failed to find support for either prediction (Pearson correlation: Serotype acquisition: $r=-0.095$, $P=0.682$, $N=21$; Carriage Duration: $r=-0.343$, $P=0.164$, $N=18$). Indeed, in both cases the direction of the non-significant relationships are negative, suggesting the opposite conclusion: that more durable strains are less likely to be acquired or maintained during carriage. Similarly, while we observe a marginally non-significant relationship between the infective odds ratio and durability in the study by Brueggemann et al. (2003) (Fig. 5.2B; Pearson correlation: $r=-0.414$, $P=0.062$, $N=21$) this was not the case in a second study (Sleeman et al. 2006) (Fig. 5.2C; Pearson correlation: $r=-0.376$, $P=0.124$, $N=21$). As above, in neither case is the direction of the correlation consistent with predictions of the sit-and-wait model. Considering disease incidence overall, unweighted by carriage prevalence, we
Figure 5.2. The relationship between mean durability on a plastic surface and several serotype specific characteristics. A) Virulence, i.e. the RR of death given disease. Squares are data points from Inverarity et al. (2011) and triangles are data from Weinberger et al. (2010); B) Odds ratio of infection from Brueggemann et al. (2003); C) Attack rate of serotypes causing invasive disease from carriage, taken from Sleeman et al. (2006); D) Incidence of invasive pneumococcal disease (IPD) (give ref for data).
found a significant negative correlation between durability and serotype-specific IPD (Fig. 5.2D; Pearson correlation: $r=-0.515$, $P=0.017$, $N=21$); this result implies that more durable strains are less likely to cause disease. Although the correlation with disease prevalence is fully dependable by serotype 14, which is a serotype strongly associated with disease (three points at the far right of figure 5.2D; Pearson Correlation without these points: $r=0.004$, $P=0.0989$, $N=18$).

### 5.5 Discussion

The central aim of this study was to test the prediction that *S. pneumoniae* has evolved a sit-and-wait strategy, following the Pharaoh-hypothesis as considered theoretically (Bonhoeffer et al. 1996) (Gandon 1998) and empirically (Walther and Ewald 2004). The Pharaoh hypothesis predicts that virulence and durability outside of the host will be positively correlated, because the constraints on virulence will be relaxed in pathogens that do not rely solely on host mobility for transmission. A previous study of this hypothesis using cross-species comparisons has characterized *S. pneumoniae* as a pathogen in the “high virulence high survival group” (Walther and Ewald 2004). However, there are several potential concerns with this interpretation. First, there is significant variation in virulence between serotypes in this species (Hausdorff et al. 2000; Kronenberg et al. 2006; Weinberger et al. 2010). Second, our results provide clear evidence of significant differences in pneumococcal surface survival that is attributable to both serotype and genotype (Figure 5.1). Taking this diversity into account, we examine the predictions of “sit-and-wait” within *S. pneumoniae*. Our results provide significant, although not overwhelming, support for the key prediction of this hypothesis. Specifically we find (Figure 5.2A) that more durable strains of *S. pneumoniae* are associated with more virulent infections, defined as the serotype-specific Risk-Ratio of death given infection.

As originally formulated, the sit-and-wait hypothesis views pathogen transmission as either direct or indirect, rather than a mixture of the two transmission modes. More recent work by
Roche et al (2011) shows that in cases of mixed-transmission diverse transmission/virulence strategies can coexist. This prediction appears well suited to the strains of *S. pneumoniae* examined here, which are all derived from a contemporaneous sample of isolates collected from Dutch daycare centres. While *S. pneumoniae* is generally thought to rely on direct transmission via respiratory droplets, our results and those of another recent study by Walsh and Cammili (2011), indicate that indirect transmission may also be of relevance. Understanding the frequency and effects of this mode of transmission is particularly important if transmission mode is informative about virulence. Making the direct link between durability and virulence requires further manipulative studies that overcome limitations of our correlative analysis. First, it is crucial to obtain direct estimates of survival of strains in the setting in which they were isolated, e.g in daycare centers, hospitals or schools. Durability in this study was assessed on polystyrene in the dark; whether these conditions increase or decrease bacterial survival as compared to conditions they face within natural fomites remains an area for future study. Second, it is necessary to independently assess virulence of collected strains using animal models, particularly as there is a clear difference in durability (Figure 1) due to genotype (beyond serotype), which could also be the case for virulence. In our analysis we had to rely on serotype specific virulence data, which might have introduced considerable inaccuracies.

While certain microbes are obligate pathogens, e.g. *Mycobacterium tuberculosis*, others such as *S. pneumoniae* cause disease opportunistically. Because pneumococcus is predominantly a commensal microbe, we examined the relationship between durability and aspects of carriage. In contrast with virulence, we found no association between durability and the likelihood of acquisition or duration of carriage. Indeed, if anything, these results suggest that more durable strains are less likely to be acquired and carried. Consistent with this, we also found that more durable strains are significantly less likely to cause disease (Fig. 5.2D). This suggests a somewhat complicated interaction between durability and pneumococcal life history. While on the one hand, durable strains appear to be more virulent when they
cause disease; they are simultaneously less likely to cause disease to begin with. This may be because more durable strains are less common overall, although this seems unlikely given that the serotypes we studied are among the most common pneumococcal serotypes. Another possibility is that durability trades-off with fitness during carriage. The human nasopharynx is colonised by several species besides *S. pneumoniae*. Competitive interactions between these species, as well as among pneumococcal clones, within the host can influence the evolution of virulence (Staves and Knell 2010; Lysenko et al. 2010). Disentangling the complex ecological interactions between and among coexisting species and their effects on virulence and durability are areas we hope to explore in future studies.

In summary, we have shown that pneumococcal durability is positively correlated with virulence. This supports the main prediction of the Pharaoh hypothesis. At the same time our results reveal the complex interaction between durability, carriage and disease, and also suggests that several different transmission strategies coexist within populations of this opportunistic pathogen.

5.6 References


Boone, S. A., and C. P. Gerba. 2007. Significance of fomites in the spread of respiratory and
Mazzola, G. J., J. E. Mortensen, L. A. Miller, and J. A. Poupard. 2003. The growth and


Chapter 6

General discussion

Daniel J.P. Engelmoer

In this thesis I have tested hypotheses concerning both short-term and long-term benefits of competence in *Streptococcus pneumoniae*. The aim was to identify factors that explain the maintenance of competence in *in vitro* populations of *S. pneumoniae*. My results have shown that competence increases bacterial fitness in both the short and long-term and also that stress is an important factor that modulates the degree of the benefits from competence. The results of this thesis build upon and expand the mechanisms by which natural competence is maintained among bacteria. In this concluding chapter I will summarize my results and interpret these within the broader context of bacterial evolution.

6.1 Short-term benefits

In Chapter 2, I show that short-term benefits of competence in *S. pneumoniae* are the result of increased resilience when cells are exposed to lethal concentrations of antibiotics. The antibiotics used in this study are known to induce competence in *S. pneumoniae* (Prudhomme et al. 2006; Stevens et al. 2011), which strongly suggests a direct link between the stress I imposed and the evolution of competence as a stress response. Our results can most simply be interpreted in the light of the “DNA-for-repair hypothesis”. This idea posits that natural competence in bacteria is beneficial because it allows bacterial cells to use environmental DNA to repair lesions caused by DNA damaging agents (Bernstein et al. 1981; Michod et al. 2008). However, I believe that my results expand upon this original idea in several important ways. Most importantly, the “Repair” hypothesis narrowly predicts benefits of competence in the form of recombinational repair via transformation, but did not consider the possibility that competence might confer other benefits more generally (Claverys et al. 2006). This
difference is crucial, especially in species where the activation of competence drives the transcription of many genes, only some small fraction of which are required for transformation (Peterson et al. 2000; Peterson et al. 2004). The experiments in this thesis show that it is necessary to distinguish benefits derived from these distinct processes. Specifically, the mode of action of an antibiotic determines if the activation of competence or transformation per se are required for the stress response. In experiments where competent populations were exposed to mitomycin C (DNA-damage) both the activation of competence and the presence of environmental DNA were required for a reduction of cell death. However, in similar experiments where competent populations were exposed to streptomycin (protein synthesis inhibition) only the activation of competence was required to reduce the rate of cell death. This shows that in response to a particular stress different competence associated functions are used to impart benefits. Ultimately it is the sum of benefits of these different responses that promotes the maintenance and evolution of competence in *S. pneumoniae*.

### 6.2 Long-term benefits

Competence is thought to be especially important in the longer term as a mechanism to speed up the rate of evolution through transformation, by recombining beneficial mutations from separate cells into a single genetic background (e.g. (Gogarten et al. 2002; Vos 2009). In Chapter 3, I use an experimental evolution approach to show that competence in *S. pneumoniae* does not necessarily result in a higher fitness after 1000 generations of laboratory adaptation. Surprisingly, we found that in a constant environment the fitness gains of non-competent populations of cells exceeds that of competent populations. That is, during growth under benign conditions, competence is costly. The addition of short periods of mild stress changes this picture drastically. Now instead of costs of competence, competent and non-competent populations evolved equal fitness. This result is consistent with the idea that competence is favoured in populations that encounter periods of reduced fitness (i.e. stress) (Hadany and Beker 2003; Hadany and Otto 2009). Finding support for the fitness associated recombination hypothesis (FAR) is perhaps unsurprising given that most organisms will
encounter stress on a regular basis, as no environment is truly constant. These results help to reconcile earlier studies that indicated uncertain benefits of competence (Bacher et al. 2006; Baltrus et al. 2008), and suggest that future experimental studies of this question need to incorporate environmental variation and stress. These results help to reconcile earlier studies that indicated uncertain benefits of competence (Bacher et al. 2006; Baltrus et al. 2008), and suggest that future experimental studies of this question need to incorporate environmental variation and stress.

Overall, the fitness results together with the surprising observation that competent populations accumulated significantly less mutations overall than non-competent populations argue against the idea that competence increases the rate of adaptation by combining beneficial mutations. Instead, these results argue that competence rather reduces the mutational load (Redfield 1993b; Redfield et al. 1997). Whilst these results are somewhat surprising, a study by Gray and Goddard (2012) in yeast (Saccharomyces cerevisiae) shows a very similar outcome. They show that sex reduces the detrimental effects of stress. They argue that this is because of the increased clearance of detrimental mutations, rather than differences in the accumulation of beneficial mutations as the adaptive rates are similar between sexual lines with different mutation rates, but a lot higher than asexual populations. The results presented in this thesis and those of Gray and Goddard (2012) show that a reduction in mutational load as a function of sex is more widespread than first thought, but more importantly that this function is context dependent.

How does competence reduce mutational load? The current view based on eukaryotic populations is that recombination creates genetic variation through generating new combinations of alleles, which will be subject to selection, without changing allele frequencies (Barton 1995). However, this is not the case for transformation. During bacterial transformation DNA for recombination is taken from the environment, which is necessarily not a part of the reproducing population. Therefore allele frequencies in bacterial populations can change as
the result of natural transformation. The change in allele frequencies depends on the source of environmental DNA, which most likely is derived from dead cells or recently killed cells in the case of *S. pneumoniae* (Johnsborg et al. 2008). This pool of DNA template will likely contain alleles that are at high frequencies among dividing cells. Thus when new alleles arise through mutations, a chance exists that transformation will replace this new allele with the old wild-type allele rather than generating new combinations of alleles. As a result competence appears to mainly reduce mutational load rather than recombining novel alleles in individuals. A similar function has been suggested on in naturally competent *Neisseria* (Treangen et al. 2008). Here specific DNA uptake sequences ensure protection of the core genome leaving the rest of the genome capable of changing via mutation. DNA uptake sequences are unknown for *S. pneumoniae*. Also, *Neisseria* is continuously competent while this is not the case for *S. pneumoniae*. Despite these mechanistic differences, it is an intriguing possibility that these two species may use distinct processes to a similar end—that is, to conserve the genome in the face of mutational damage.

6.3 Is competence always conservative?

The conservative nature of competence observed in my experiments does not mean that competence cannot result in bacterial innovation. In my experiments the pool of DNA template consisted of DNA of closely related strains of the same species, which will have restricted the chances that a novel allele will be transformed. In a natural environment this pool of DNA sources will be much more diverse. This is supported by several cases of lateral gene transfer from other species into *S. pneumoniae* (Kelly et al. 1994; Ochman et al. 2000; Hakkenbeck et al. 2001; Boyd et al. 2009; Juhas et al. 2009; Treangen and Rocha 2011). Among other functions, imported genes can confer novel antibiotic resistance and modified modes of virulence, e.g. the pneumococcal capsule. While environmental DNA can provide beneficial novel alleles, transformation can also entail risks if the incorporated DNA is damaged or contains deleterious mutations. Thus *S. pneumoniae* appears to have evolved a mechanism to ensure a supply of undamaged DNA, called competence induced cell lysis (Steinmoen et
al. 2002; Steinmoen et al. 2003), which may reduce the risk of transformation with damaged DNA. Competence induced cell lysis comes with a new risk, however, as the DNA may come from cells that are incapable of becoming competent. If lesions abolishing competence are transformed, cellular populations may lose the ability to become transformed (Claverys et al. 2006). Bacher et al (2006) observed such an outcome, where in long-term evolved Acinetobacter baylyi transformation was lost. How cells balance the costs and benefits of transformation remain uncertain. Stress, however, may be of importance for its maintenance.

Several questions remain on the subject of competence as a generalized stress response. Firstly, the work in this thesis has been limited to antibiotic stress and it would be interesting to see if these results can be extended to other forms of stress, such as U.V. light, pH, osmotic stress and antibiotics with different modes of action. Not all antibiotics are able to induce competence (Prudhomme et al. 2006), but these non-inducing drugs can still cause stress. Secondly, the rate of transformation varies markedly between S. pneumoniae strains (Pozzi et al. 1996). This variation results from many different factors, among which are genetic variation in the CSP receptor comD and the thickness of the polysaccharide capsule surrounding S. pneumoniae cells. The variation in comD is driven by the fact that there are two dominant forms of the competence signaling peptide, CSP pherotypes (Pozzi et al. 1996), and each receptor is believed to specifically bind its cognate signal. Some versions of the receptor can detect both pherotypes, but at the cost of transformation efficiency (Iannelli et al. 2005). These polymorphisms do not obviously result in substructures of S. pneumoniae populations (Cornejo et al. 2010), as strains carrying either pherotype can recombine; however, currently no explanation exists for the maintenance of this polymorphism. The polysaccharide capsule is highly diverse and is one of the important virulence factors of S. pneumoniae (Brueggemann et al. 2003). Strains with thinner capsules have higher transformation rates and also are generally considered to be invasive strains (Pozzi et al. 1996; Sandgren et al. 2004; Sjostrom et al. 2006). Invasive strains may encounter more stress, as they will be directly exposed to the immune system of the host. Equally, however, carriage isolates will
experience immune surveillance as well as extensive competition from coexisting bacteria within the human nasopharynx. At present, it is unclear if invasive strains have a higher frequency of transformation than those isolated from carriage. Thirdly, the generality of stress as the factor promoting the maintenance of competence remains unclear. In other species, such as *Bacillus subtilis* and *Haemophylus influenzae*, UV-light and mitomycin C, which are two clear DNA-damaging stresses, do not induce competence even though competence provides protection against them (Michod et al. 1988; Mongold 1992; Redfield 1993a). The timing of competence induction may also provide information about whether or not this response arises due to cellular stress. For example, in *B. subtilis* competence is induced at the start of spore formation, which is generally considered a stress response (Johnsborg et al. 2007). Furthermore, in *Neisseria* competence is believed to function as a mechanism to reduce the accumulation of deleterious mutations, which one can also consider a form of stress (Treangen et al. 2008). The factors that maintain competence might not be the same for every bacterial species, but I believe that current research indicates that stress does play an important role.

### 6.4 Mutators, parallel evolution

The long-term evolution study not only showed the importance of stress for the maintenance of competent populations, but it also allowed for two other important observations. Firstly, the presence or absence of competence influences the rate of mutation (Chapter 3). I observed that non-competent populations evolved increased mutation rates and mutator phenotypes during the long-term experiment, while competent cells maintained a much lower mutation rate. Mutator alleles can potentially permit an increased rate of adaptation at the cost of increased numbers of deleterious mutations. Such mutator alleles often are difficult to purge from a non-competent population because they often become linked to a beneficial mutation (Taddei et al. 1997). However, even low frequencies of recombination can break these hitchhiking associations preventing the fixation of mutator alleles. This may make mutator alleles and competence two
mutually exclusive adaptive strategies (Tenaillon et al. 2000). It will be interesting to
determine if competence did prevent the fixation of mutator phenotypes in the evolved
populations. Secondly, I examined phenotypic and genotypic parallel changes in the
evolved populations. I observed that both population density oscillations and biofilm
formation was strongly reduced across all evolved populations. At the genomic level
parallel changes followed earlier observations in bacteria where the extent of parallelism
increased with decreasing resolution of analysis. In order of commonness, parallelism was
seen most often in functional groups, than in genes, and finally at individual nucleotide
sites. (Woods et al. 2006; Lieberman et al. 2011; Dettman et al. 2012; Tenaillon et al.
2012). Parallel changes were observed in specific treatments. In addition, several parallel
changes were found across treatments all populations. Additional experiments are required
to quantify the effects of each mutation and to distinguish between the possibility that
these mutations arose due to the direct benefits they causes versus the possibility that they
became fixed due to genetic drift or hitchhiking. Nevertheless these observations show that
parallel change is not necessarily limited by boundaries such as the presence or absence of
competence, or the presence or absence of periodic stress.

6.5 The Pharaoh hypothesis

In Chapter 5, a brief experimental side-step into virulence evolution of *S. pneumoniae*
revealed that serotypes that have higher rates of survival on plastic surfaces tend also to be
the most virulent ones (i.e. these serotypes have a high risk ratio of resulting in death after
infection). This supports the “Pharaoh hypothesis”, which predicts that pathogens can maxi-
mize their virulence if transmission does not depend on host survival. Thus pathogens that
can ‘sit-and-wait’ for a new host can evolve maximum virulence (Bonhoeffer et al. 1996). *S.
pneumoniae* goes through a carriage stage before it (infrequently) goes on to cause disease.
Therefore one can expect that the sit-and-wait strategy applies to carriage in this species
rather than to virulence per se; however, no relationship between surface survival and car-
riage prevalence were found. Instead an inverse relationship was found for the propensity
to cause disease from carriage and surface survival. Overall these data indicate that diverse survival and virulence strategies have evolved with in the *S. pneumoniae* population. Disentangling these diverse virulence strategies in *S. pneumoniae* will be an aim for the future.

### 6.6 Conclusion

To summarise, this thesis contributes to literature on the evolution of competence, genomic evolution during experimental laboratory evolution, and to the literature on the Pharaoh hypothesis. I have addressed the question of why competence is maintained in *S. pneumoniae*. Starting from existing theory I found that stress is an important factor that promotes the maintenance of competence. I have expanded on the original DNA-for-repair hypothesis showing that competence also protects against stress that does not damage DNA. In the long-term competence appears to help cells to overcome their mutational load rather than increasing the rate of adaptation. The addition of periodic stress mitigates intrinsic costs of competence, thus supporting the fitness associated recombination hypothesis. The unique combination of hypothesis based long-term evolution experiment and next generation sequencing allowed me to explore genomic evolution in addition to fitness changes. I have described several cases of parallel phenotypic and genomic evolution, which will be good starting points for further study. Most interestingly, I found that genomic parallel changes occur in both populations within a treatment and in populations across all treatments showing that parallel evolution is not restricted by differences in genotype (competence) and environment (periodic stress). Finally, my results reveal that competence is an important trait for *S. pneumoniae*, not because it increases the rate of adaptation, but because competence reduces direct effects of stress and prevents the accumulation of deleterious mutations.

### 6.7 References


Prudhomme, M., L. Attaiech, G. Sanchez, B. Martin, and J. P. Claverys. 2006. Antibiotic stress induces genetic transformability in the human pathogen *Streptococcus pneum-


