

Copyright is owned by the Author of the thesis. Permission is given for a copy to be downloaded by an individual for the purpose of research and private study only. The thesis may not be reproduced elsewhere without the permission of the Author.

**Formalist features determining the  
tempo and mode of evolution in  
*Pseudomonas fluorescens* SBW25**

A thesis submitted in partial fulfilment of the requirements for the degree  
of  
Ph.D.  
In  
Evolutionary Genetics

At Massey University, Auckland, New Zealand

Andrew David Farr

2015



## Abstract

In order to explain the adaptive process, it is necessary to understand the generation of heritable phenotypic variation. For much of the history of evolutionary biology, the production of phenotypic variation was believed to be unbiased, and adaptation the primary outcome of selection acting on randomly generated variation (mutation). While true, 'internal' features of organisms may also play a role by increasing the rate of mutation at specific loci, or rendering certain genes better able to translate mutation into phenotypic variation. This thesis, using a bacterial model system, demonstrates how these internal features – localised mutation rates and genetic architectures – can influence the production of phenotypic variation.

Previous work involving the bacterium *Pseudomonas fluorescens* SBW25 has shown that mutations at three loci, *wsp*, *aws* and *mws*, can cause the adaptive wrinkly spreader (WS) phenotype. For each locus, the causal mutations are primarily in negative regulators of di-guanylate cyclase (DGC) activity, which readily convert mutation into the WS phenotype. Mutations causing WS at other loci were predicted to arise, but to do so with less frequent types of mutation. The data presented in this thesis confirms this prediction. My work began with the identification and characterisation of a single rare WS-causing mutation: an in-frame deletion that generates a translational fusion of genes *fadA* and *fwsR*. The fusion couples a DGC (encoded by *fwsR*) to a membrane-spanning domain (encoded by *fadA*) causing relocalisation of the DGC to the cell membrane and the WS phenotype. This is one of the few examples of adaptation caused by gene fusion and protein relocation in a real-time evolution experiment. I next took an experimental evolution approach to isolate further rare WS types and characterized these, revealing a range of rarely taken mutational pathways to WS. Lastly, I describe an example of extreme molecular parallelism, in which a cell chaining phenotype is caused – without exception – by a single nucleotide substitution within the gene *nlpD*, despite multiple mutational pathways to this phenotype. Characterisation of different *nlpD* mutants suggests this molecular parallelism is caused by a high local mutation rate, possibly related to the initiation of transcription within this gene.

## Acknowledgements

Foremost, I would like to thank my supervisor, Professor Paul Rainey for discussion and ideas, and for providing the opportunity to pursue my own research directions in such a quality environment. I am very grateful for the years of support. I am thankful for the opportunity to discuss and present research at the Gordon Research Conference in Microbial Population Biology (New Hampshire, USA, 2011). Funding for these projects was obtained by Prof Rainey, and was partially supported by the Marsden Fund of New Zealand. Thanks to the Royal Society of New Zealand and Marsden Fund Council for making fundamental research such as this possible. I am also very grateful to the New Zealand Institute for Advanced Study for financial support during my studies.

Thanks to the post-docs who helped school me on the theory and study of microbial evolution. Particular acknowledgement is due to Dr Peter Lind who provided guidance on the *nlpD* project and assisted with both the flow cytometry fitness assays, and performed the *nlpD* mutant reversion-frequency assay. Thanks to Dr Heather Hendrickson, who put an enormous effort in setting up the fluorescent microscope, keeping it maintained, and offering guidance in operation. Thanks to Dr Mike McDonald who made the deletions of the initial PBR716 type, and who – as main author on a review we wrote – pointed the way for the useful histories of the field. Thanks to Dr Jenna Gallie for help with my first transposon assay, provision of primers, advice whenever it was needed, and the pUIC3  $\Delta$ wss construct. Thanks to Dr Xue-Xian Zhang for his expansive protocols, Dr Monica Gerth for sage cloning advice, and to Dr Christian Kost for his encouragement and advice and gift of PBR923 and primers.

Thanks also to all those I've worked with over the years who made and make the Rainey Lab such a rewarding experience!

For providing suggestions for the improvement of this thesis gratefully thank Professor Paul Rainey, Dr Peter Lind, Katrina Rainey and Simon Coverdale.

Finally, sincere gratitude to my friends for making life fun, and the deepest appreciation to my family for tireless personal support – Linda, Gavin and Vera, I'm forever in your debt.

# Table of Contents

<b>Abstract</b> .....	<b>I</b>
<b>Acknowledgements</b> .....	<b>II</b>
<b>Table of Contents</b> .....	<b>III</b>
<b>Table of Abbreviations</b> .....	<b>IX</b>
<b>Chapter 1: Introduction</b> .....	<b>1</b>
1.1 A very brief history of the theory of variation .....	1
1.1.1 Variation within Darwin's theory of natural selection.....	2
1.1.2 The first formalist alternative to Darwinian variation.....	3
1.1.3 The second evolutionary alternative to Darwinian variation: saltation .....	5
1.1.4 Rejection of orthogenesis and saltation in the modern synthesis .....	6
1.2 Genetic architecture and the production of variation .....	8
1.2.1 Contemporary theories of the importance of variation .....	8
1.2.2 Contemporary theories of the role of genetic architecture on evolution.....	9
1.2.3 Insight on the constraint of genetic architecture from evo-devo .....	10
1.2.3.1 Parallel mutations to <i>shavenbaby</i> in <i>Drosophila</i> .....	10
1.2.3.2 Parallel mutations to <i>FRIGIDA</i> in <i>Arabidopsis</i> .....	11
1.2.3.3 Parallel mutations to <i>R2R3MYB</i> across Astrids .....	12
1.2.4 Implications of evo-devo examples on rate of variation production .....	12
1.2.5 Localised mutation rates may alter the production of phenotypic variation .....	14
1.3 Microbes as model systems to explore the relative bias of variation caused by genetic architecture .....	14
1.3.1 Microbes and the ability to identify the locus of mutation.....	15
1.3.2 Microbes and the ability to compare architectures via redundancy .....	16
1.3.3 The <i>Pseudomonas fluorescens</i> SBW25 model system .....	17
1.3.3.1 Adaptive radiation of SBW25 in experimental conditions .....	17
1.3.3.2 The structural basis of the WS phenotype .....	18
1.3.3.3 Mutations causing WS involve changes that affect the regulation of c-di-GMP. 19	
1.3.3.4 The wild type regulation and effects of individual DGCs remain unclear.....	19
1.3.4 DGCs provide a model to assess the effect of architecture on variation.....	20
1.3.4.1 The architectures of known WS-causing DGCs.....	21
1.3.4.1.1 The <i>wsp</i> operon .....	21
1.3.4.1.2 The <i>aws</i> operon .....	23

1.3.4.1.3 The <i>mws</i> composite gene .....	25
1.3.4.2 Summary of <i>wsp</i> , <i>aws</i> and <i>mws</i> architecture and relative rate of evolution of the three pathways .....	27
1.3.5 Negative regulation in <i>wsp</i> , <i>aws</i> and <i>mws</i> may increase the production of variation .....	28
1.4 Research objectives.....	30

**Chapter 2: Materials and methods ..... 31**

<b>2.1 Materials .....</b>	<b>31</b>
2.1.1 Media and growth conditions .....	31
2.1.2 Antibiotics, reagents and enzymes .....	31
2.1.3 DNA Electrophoresis.....	32
2.1.4 Competent cell production .....	32
2.1.5 Photography and microscopy materials.....	32
2.1.6 Strain list .....	33
2.1.7 Transposon and plasmids.....	36
2.1.8 Primers.....	37
<b>2.2 Methods .....</b>	<b>41</b>
2.2.1 Plasmid purifications for electroporation and ligation.....	41
2.2.2 Purification of PCR products.....	41
2.2.3 Triparental conjugation .....	41
2.2.4 Transposon mutagenesis.....	42
2.2.5 Polymerase chain reaction (PCR).....	42
2.2.5.1 Strand overlap extension PCR .....	43
2.2.6 Sanger sequencing .....	44
2.2.7 Evolution of ISWS types .....	45
2.2.8 Allelic exchange .....	45
2.2.9 Production and transformation of competent cells.....	46
2.2.10 Fitness assays .....	47
2.2.10.1 Fitness assays via colony morphology distinction.....	47
2.2.10.2 Competition fitness assays using flow cytometry .....	48
2.2.11 Transcriptional assay .....	49
2.2.12 Quantitative RT-qPCR .....	49
2.2.13 Fluorescence microscopy (ACP and GFP) .....	50
2.2.14 Whole-genome sequencing .....	51
2.2.15 SREE protocol .....	51
2.2.16 Ordering mutations of the SREE.....	52
2.2.17 Measuring <i>rifR</i> mutation rates.....	52

2.2.18 Statistical analyses .....53

**Chapter 3: Characterisation of an adaptive pathway created by domain fusion and protein relocalisation..... 55**

**3.1 Introduction .....55**

**3.2 Aims.....56**

**3.3 Results .....57**

3.3.1 Identification of the genetic basis of SWS and measurement of the effects on fitness .....57

3.3.1.1 The SWS mutational locus spans two genes.....57

3.3.1.2 The *fadA-fwsR* fusion is sufficient to form WS .....60

3.3.1.3 The *fadA-fwsR* fusion causes WS of similar fitness to *wspF* mutants .....62

3.3.2 Multiple *fadA-fwsR* fusion mutations cause WS types .....63

3.3.3 The *fadA-fwsR* fusion requires an active DGC to cause WS .....68

3.3.4 Investigation of how *fadA-fwsR* causes WS .....70

3.3.4.1 Testing the hypothesis that promoter capture may increase expression of the DGC coding frame .....70

3.3.4.1.1 Sequence data suggests the *fadA-fwsR* deletion alters the operator of transcription of *fwsR* .....70

3.3.4.1.2 Promoter capture is not sufficient to cause SWS.....71

3.3.4.2 The SWS phenotype requires FadA and FwsR to be translationally fused .....72

3.3.4.3 Testing the hypothesis that relocalisation of the FwsR DGC causes SWS .....74

3.3.4.3.1 Membrane localisation domains necessary for FwsR to cause the SWS phenotype .....74

3.3.4.3.2 Fluorescent tagging shows relocation of the DGC to the membrane associated with the *fadA-fwsR* fusion .....75

**3.4 Discussion .....81**

3.4.1 Summary of results: relocalisation of FwsR to the membrane determines the WS phenotype .....81

3.4.2 The potential contribution to evolution across loci is influenced by genetic architecture.....81

3.4.3 *fadA-fwsR* as a model for the evolution of new genes .....83

3.4.3.1 The *fadA-fwsR* mutation is a model example of the creative power of modularity .....83

3.4.3.2 *FadA-fwsR* provides direct experimental evidence that translational fusion mutations cause adaptive change via protein relocalisation .....84

3.4.3.3 Chimeric relocalisation as a novel mechanism behind new genes in the ‘neofunctionalisation model’ .....85

3.4.4 Implications for secondary messenger research and future directions .....86

<b>Chapter 4: The identification of multiple loci that cause WS.....</b>	<b>89</b>
<b>4.1 Introduction .....</b>	<b>89</b>
<b>4.2 Aims .....</b>	<b>92</b>
<b>4.3 Results .....</b>	<b>93</b>
4.3.1 The SREE methodology and the subsequent identification of mutations ....	93
4.3.2 Identification of genotypes that were not identified via sequencing of the endpoints .....	94
4.3.3 Identity of mutations at GGDEF-containing loci and associated WS phenotypes .....	96
4.3.3.1. Mutations to DGC loci during the SREE are rare and diverse, compared to the original REE .....	96
4.3.3.2 Mutations identified at rare WS loci .....	97
4.3.3.2.1 PFLU0085 .....	97
4.3.3.2.2 <i>fadA-fwsR</i> .....	99
4.3.3.2.3 PFLU0458 and PFLU4744 .....	101
4.3.3.2.4 PFLU0621 .....	103
4.3.3.2.5 PFLU0956 .....	105
4.3.3.2.6 PFLU4781 to PFLU4782 .....	107
4.3.3.2.7 PFLU4858 .....	109
4.3.3.2.8 PFLU2764 .....	111
4.3.3.2.9 PFLU5698 .....	113
4.3.4 The <i>wss</i> operon is the most frequent target of mutation in the SREE .....	115
4.3.5 Parallel evolution at non-DGC regulatory loci of <i>mvaT</i> and <i>nlpD</i> .....	116
4.3.5.1 PFLU4939 .....	117
4.3.5.2 PFLU1301 ( <i>nlpD</i> ) .....	119
<b>4.4 Discussion.....</b>	<b>122</b>
4.4.1 General summary of putative DGC-regulatory mutations .....	122
4.4.2 Mutational set at PFLU0085 suggests post-translational negative regulation encoded at this locus .....	123
4.4.3 The remaining putative c-di-GMP regulators require co-occurring mutations, fusions or promoter mutations to cause WS .....	124
4.4.4 The <i>wss</i> operon is the source of the greatest parallelism in the SREE .....	125
4.4.5 Parallel evolution during the SREE is associated with negative regulators	126
4.4.5.1 Parallelism at locus PFLU4939 is associated with loss-of-function mutations	127
4.4.5.2 Accounting for parallelism in the <i>nlpD</i> gene .....	129
<b>Chapter 5: The causes of parallel molecular evolution of <i>nlpD</i> mutations .....</b>	<b>131</b>

<b>5.1 Introduction .....</b>	<b>131</b>
5.1.1 The observed parallelism to <i>nlpD</i> and the associated phenotype change .....	132
5.1.2 The parallel evolution of C565T mutants may have several causes .....	133
<b>5.2 Aim.....</b>	<b>134</b>
<b>5.3 Results .....</b>	<b>135</b>
5.3.1 A guide to the nomenclature of mutations presented in this chapter .....	135
5.3.2 Identification of the possible spectrum of <i>nlpD</i> mutants evolved from SBW25 <i>Δwss</i> .....	135
5.3.3 The chaining phenotype caused by <i>nlpD</i> mutations does not require transcription of <i>rpoS</i> .....	137
5.3.4 Alteration of the <i>nlpD</i> sequence causes a lower frequency of molecular parallelism .....	139
5.3.5 Assessing whether the fitness of <i>nlpD</i> mutants explains parallel evolution	142
5.3.6 The reduced parallelism of <i>nlpD</i> subtle mutants is associated with reduced <i>rpoS</i> transcription .....	144
5.3.7 Increased transcription from $P_{rpoS}$ correlates with the frequency of mutations within <i>nlpD</i> .....	147
<b>5.4. Discussion .....</b>	<b>151</b>
5.4.1 Summary of the causes of parallel mutations to <i>nlpD</i> .....	151
5.4.2 A mutational mechanism may result in the parallelism of C565T .....	152
5.4.3 Implications of the potential mutability of nested promoters at <i>nlpD</i> and other loci .....	154
5.4.4 What is the benefit of a nested <i>rpoS</i> promoter? .....	156
5.4.4.1 The nesting of $P_{rpoS}$ in <i>nlpD</i> may preserve the promoter.....	156
5.4.4.2 The nesting of $P_{rpoS}$ may cause a stochastic change in cell length .....	157
<b>Chapter 6: Concluding remarks.....</b>	<b>159</b>
<b>6.1. Project background.....</b>	<b>159</b>
<b>6.2 Summary of findings.....</b>	<b>160</b>
<b>6.3 Implications of these findings for predicting evolution .....</b>	<b>164</b>
<b>Reference list.....</b>	<b>167</b>
<b>Appendix.....</b>	<b>187</b>
<b>Appendix for Chapter 3.....</b>	<b>187</b>
Section 7.3.1 Transcription assays suggest <i>fadA</i> and <i>fwsR</i> are independently transcribed.....	190
Section 7.3.2 FadA-FwsR-GFP fusions have greater foci intensities than FadA- GFP fusions.....	196

**Appendix for Chapter 4** ..... **199**  
**Appendix for Chapter 5** ..... **206**  
    Section 7.5.1: SBW25 *nlpD* C565T is less fit than WS types ..... 206  
    Section 7.5.2: The degree of change to NlpD is associated with cell-division  
    deficiency ..... 209

## Table of Abbreviations

Abbreviation	Meaning
ACP	Acetylated cellulosic polymer
ALI	Air-liquid interface
BLAST	Basic local alignment search tool
bp	Base pair
c-di-GMP	bis-(3'-5')-cyclic dimeric guanosine monophosphate
CCF	Cross correlation function
CFU	Colony forming units
DGC	Di-guanylate cyclase
DMSO	Dimethyl sulfoxide
DNA	Deoxyribonucleic Acid
dNTP	Dinucleotide triphosphate
EE	Experimental evolution
FAD	Fatty acid desaturase
FS	Fuzzy Spreader
g	Gravity or gram
GFP	Green fluorescent protein
Gm	Gentamicin
h	Hour
HT	Homopolymeric tract
IWS	Independent wrinkly spreader
ISWS	Independent slow wrinkly spreader
Km	Kanamycin
kb	kilobase
KB	King's B medium
LB	Lysogeny broth
LSWS	Large spreading wrinkly spreader
M-W-W	Mann-Whitney-Wilcoxon
MCs	Microcosms
min	Minute
4MU	7-hydroxy-4-methylcoumarin
4MUG	4-methylumbelliferyl-b-d-galactoside
NF	Nitrofurantoin
ORF	Open reading frame
OD	Optical density
PCR	Polymerase chain reaction
PDE	Phosphodiesterase
RBS	Ribosome binding site
REE	Reverse evolution experiment
RNA	Ribonucleic acid
rpm	Revolutions per minute
RT-qPCR	Reverse transcription quantitative real-time PCR

<b>Abbreviation</b>	<b>Meaning</b>
SE	Standard error
SM	Smooth morphotype
SOC	Super optimal broth with catabolic repressor
SREE	Slow reverse evolution experiment
SWS	Slow wrinkly spreader
TAM	Transcription associated mutation
Tc	Tetracycline
TFB	Transformation buffer
TIAM	Transcription initiation associated mutation
TMD	Transmembrane domain
WGS	Whole genome sequencing
WS	Wrinkly spreader
WSS	Wrinkly spreader structural locus
WT	Wild type
X-gal	5-bromo-4-chloro-3-indolyl- $\beta$ -D-galactopyranoside