DOI: 10.1534/genetics.109.107110

Adaptive Divergence in Experimental Populations of *Pseudomonas fluorescens*. IV. Genetic Constraints Guide Evolutionary Trajectories in a Parallel Adaptive Radiation

Michael J. McDonald,* Stefanie M. Gehrig,† Peter L. Meintjes,* Xue-Xian Zhang* and Paul B. Rainey*,[1]

**New Zealand Institute for Advanced Study and Allan Wilson Centre for Molecular Ecology and Evolution, Massey University Albany, North Shore City 0745, New Zealand and* †*Department of Plant Sciences, University of Oxford, Oxford OX1 3RB, United Kingdom*

Manuscript received July 7, 2009
Accepted for publication August 5, 2009

ABSTRACT

The capacity for phenotypic evolution is dependent upon complex webs of functional interactions that connect genotype and phenotype. Wrinkly spreader (WS) genotypes arise repeatedly during the course of a model Pseudomonas adaptive radiation. Previous work showed that the evolution of WS variation was explained in part by spontaneous mutations in *wspF*, a component of the Wsp-signaling module, but also drew attention to the existence of unknown mutational causes. Here, we identify two new mutational pathways (Aws and Mws) that allow realization of the WS phenotype: in common with the Wsp module these pathways contain a di-guanylate cyclase-encoding gene subject to negative regulation. Together, mutations in the Wsp, Aws, and Mws regulatory modules account for the spectrum of WS phenotype-generating mutations found among a collection of 26 spontaneously arising WS genotypes obtained from independent adaptive radiations. Despite a large number of potential mutational pathways, the repeated discovery of mutations in a small number of loci (parallel evolution) prompted the construction of an ancestral genotype devoid of known (Wsp, Aws, and Mws) regulatory modules to see whether the types derived from this genotype could converge upon the WS phenotype via a novel route. Such types—with equivalent fitness effects—did emerge, although they took significantly longer to do so. Together our data provide an explanation for why WS evolution follows a limited number of mutational pathways and show how genetic architecture can bias the molecular variation presented to selection.

UNDERSTANDING—and importantly, predicting—phenotypic evolution requires knowledge of the factors that affect the translation of mutation into phenotypic variation—the raw material of adaptive evolution. While much is known about mutation rate (*e.g.*, Drake *et al.* 1998; Hudson *et al.* 2002), knowledge of the processes affecting the translation of DNA sequence variation into phenotypic variation is minimal.

Advances in knowledge on at least two fronts suggest that progress in understanding the rules governing the generation of phenotypic variation is possible (Stern and Orgogozo 2009). The first stems from increased awareness of the genetic architecture underlying specific adaptive phenotypes and recognition of the fact that the capacity for evolutionary change is likely to be constrained by this architecture (Schlichting and Murren 2004; Hansen 2006). The second is the growing number of reports of parallel evolution (*e.g.*, Pigeon *et al.* 1997; ffrench-Constant *et al.* 1998; Allender *et al.* 2003; Colosimo *et al.* 2004; Zhong *et al.* 2004; Boughman *et al.* 2005; Shindo *et al.* 2005; Kronforst *et al.* 2006; Woods *et al.* 2006; Zhang 2006; Bantinaki *et al.* 2007; McGregor *et al.* 2007; Ostrowski *et al.* 2008)—that is, the independent evolution of similar or identical features in two or more lineages—which suggests the possibility that evolution may follow a limited number of pathways (Schluter 1996). Indeed, giving substance to this idea are studies that show that mutations underlying parallel phenotypic evolution are nonrandomly distributed and typically clustered in homologous genes (Stern and Orgogozo 2008).

While the nonrandom distribution of mutations during parallel genetic evolution may reflect constraints due to genetic architecture, some have argued that the primary cause is strong selection (*e.g.*, Wichman *et al.* 1999; Woods *et al.* 2006). A means of disentangling the roles of population processes (selection) from genetic architecture is necessary for progress (Maynard Smith *et al.* 1985; Brakefield 2006); also necessary is insight into precisely how genetic architecture might bias the production of mutations presented to selection.

Despite their relative simplicity, microbial populations offer opportunities to advance knowledge. The

Supporting information is available online at http://www.genetics.org/cgi/content/full/genetics.109.107110/DC1.

[1]*Corresponding author:* NZIAS, Massey University Albany, Private Bag 102904, North Shore City 0745, New Zealand.
E-mail: p.b.rainey@massey.ac.nz

wrinkly spreader (WS) morphotype is one of many different niche specialist genotypes that emerge when experimental populations of *Pseudomonas fluorescens* are propagated in spatially structured microcosms (Rainey and Travisano 1998). Previous studies defined, via gene inactivation, the essential phenotypic and genetic traits that define a single WS genotype known as LSWS (Spiers *et al.* 2002, 2003) (Figure 1). LSWS differs from the ancestral SM genotype by a single nonsynonymous nucleotide change in *wspF*. Functionally (see Figure 2), WspF is a methyl esterase and negative regulator of the WspR di-guanylate cyclase (DGC) (Goymer *et al.* 2006) that is responsible for the biosynthesis of c-di-GMP (Malone *et al.* 2007), the allosteric activator of cellulose synthesis enzymes (Ross *et al.* 1987). The net effect of the *wspF* mutation is to promote physiological changes that lead to the formation of a microbial mat at the air–liquid interface of static broth microcosms (Rainey and Rainey 2003).

To determine whether spontaneous mutations in *wspF* are a common cause of the WS phenotype, the nucleotide sequence of this gene was obtained from a collection of 26 spontaneously arising WS genotypes ($WS_{A\text{-}Z}$) taken from 26 independent adaptive radiations, each founded by the same ancestral SM genotype (Figure 1): 13 contained mutations in *wspF* (Table 1; Bantinaki *et al.* 2007). The existence of additional mutational pathways to WS provided the initial motivation for this study.

Here we define and characterize two new mutational routes (Aws and Mws) that together with the Wsp pathway account for the evolution of 26 spontaneously arising WS genotypes. Each pathway offers approximately equal opportunity for WS evolution; nonetheless, additional, less readily realized genetic routes producing WS genotypes with equivalent fitness effects exist. Together our data show that regulatory pathways with specific functionalities and interactions bias the molecular variation presented to selection.

MATERIALS AND METHODS

Bacterial strains, growth conditions, and manipulation: The ancestral (wild-type) "smooth" (SM) strain is *P. fluorescens* SBW25 and was isolated from the leaf of a sugar beet plant grown at the University Farm, Wytham, Oxford, United Kingdom, in 1989 (Rainey and Bailey 1996). The ancestral strain was immediately stored at −80° to minimize adaptation to the laboratory environment. The niche-specialist wrinkly spreader (LSWS) genotype [PBR663 (previously PR1200)] was derived from the ancestral genotype after 3 days of selection in a spatially structured microcosm (Spiers *et al.* 2002). Twenty-six independent WS (WS_A–WS_Z) were derived from the ancestral SM genotype after 5 days of selection (a single WS type was isolated from each of 26 independent microcosms) (Bantinaki *et al.* 2007). LSWS *lacZ* (PBR741) was generated by mobilizing mini-Tn *7-lacZ* from *Escherichia coli* (Choi *et al.* 2005).

P. fluorescens strains were grown in lysogeny broth (LB) (Bertani 1951) or King's medium B (KB; King *et al.* 1954) at 28°. *E. coli* strains were grown in LB at 37°. Antibiotics and supplements were used at the following concentrations: ampicillin, 100 $\mu g \cdot ml^{-1}$; cycloserine, 8 $\mu g \cdot ml^{-1}$; gentamycin, 10 $\mu g \cdot ml^{-1}$; kanamycin, 100 $\mu g \cdot ml^{-1}$; spectinomycin, 100 $\mu g \cdot ml^{-1}$; and tetracycline, 12 $\mu g \cdot ml^{-1}$. X-gal (5-bromo-4-chloro-3-indolyl-β-d-galactopyranoside) was added at a concentration of 40 $\mu g \cdot ml^{-1}$. *N*-[5-nitro-2-furfurylidene]-1-arminutesohydantoin was used to counterselect *E. coli* (100 $\mu g \cdot ml^{-1}$ in agar plates). Plasmid DNA was introduced to *E. coli* by transformation and *P. fluorescens* by conjugation, following standard procedures. The helper plasmid pRK2013 was used to facilitate transfer between *E. coli* and *P. fluorescens* (Figurski and Helinski 1979).

Molecular biology techniques: Standard molecular biology techniques were used throughout (Sambrook *et al.* 1989). Gateway technology (vectors pCR8 and pCR2.1; Invitrogen) was used to clone PCR fragments. The fidelity of all cloned fragments was checked by DNA sequencing. Oligonucleotide primers for allelic replacements and reconstructions were designed on the basis of the SBW25 genome sequence (Silby *et al.* 2009).

Construction of deletion mutants and allelic replacements: A two-step allelic exchange strategy was used as described previously (Rainey 1999; Bantinaki *et al.* 2007). For the generation of deletion mutants, oligonucleotide primers were used to amplify ~1000 nucleotide regions flanking the gene(s) of interest. A third PCR reaction was used to SOE flanking sequences together (Horton *et al.* 1989). The product was cloned into pCR8 (and sequenced to check for errors); the resulting fragment was excised using *Bgl*II and introduced into the suicide vector pUIC3 (Rainey 1999). Deletion constructs were mobilized into *P. fluorescens* by conjugation and allowed to recombine with the chromosome by homologous recombination. Recombinants from which pUIC3 had been lost were identified from LB agar plates after a brief period of nonselective growth in KB broth. To generate allelic replacements, PCR fragments (~3 kb) centered upon the allele of interest were cloned as above and introduced into pUIC3, and the replacement was made by recombination. Deletion and allelic exchange mutants were confirmed by PCR; allelic exchange mutants were further checked by DNA sequencing.

Transposon mutagenesis analysis: Plasmid pCM639 containing IS*phoA*/hah-Tc or IS-Ω-kan/hah (Giddens *et al.* 2007) was introduced from *E. coli* SM10λ_{pir} into *P. fluorescens* by conjugation. The genomic location of the transposon insertions was determined by sequencing the DNA products of arbitrarily primed PCR (Manoil 2000). The precise genomic location of each transposon insertion is recorded in Artemis (Rutherford *et al.* 2000) entry files that can be read into the SBW25 genome (Silby *et al.* 2009) annotation file (http://www.sanger.ac.uk/Projects/P_fluorescens). The Artemis entry files are available as supporting information, File S1, File S2, File S3, and File S4. The SBW25 annotation file is available from the Sanger Web site, above.

As the suppressor study progressed (see results), it became clear that genes encoding DGCs were of primary interest, and yet mutations were commonly (and expectedly) found in the *wss* locus (~35% of insertions). A simple PCR screen was developed to identify these mutants. Each mutant was therefore subjected to a PCR reaction that included a primer specific to the transposon and reading in the outward direction plus a set of 13 *wss*-specific primers (available on request) spaced ~800 bp apart. Mutants from which a PCR fragment was detected were deemed to carry a transposon in the *wss* operon, and these mutants were not further characterized.

Fitness of genotypes: Competitive fitness of WS genotypes was determined by direct competition between each WS genotype and LSWS *lacZ* (PBR947): the ancestral broth colonizing genotype was included to ensure occupancy of

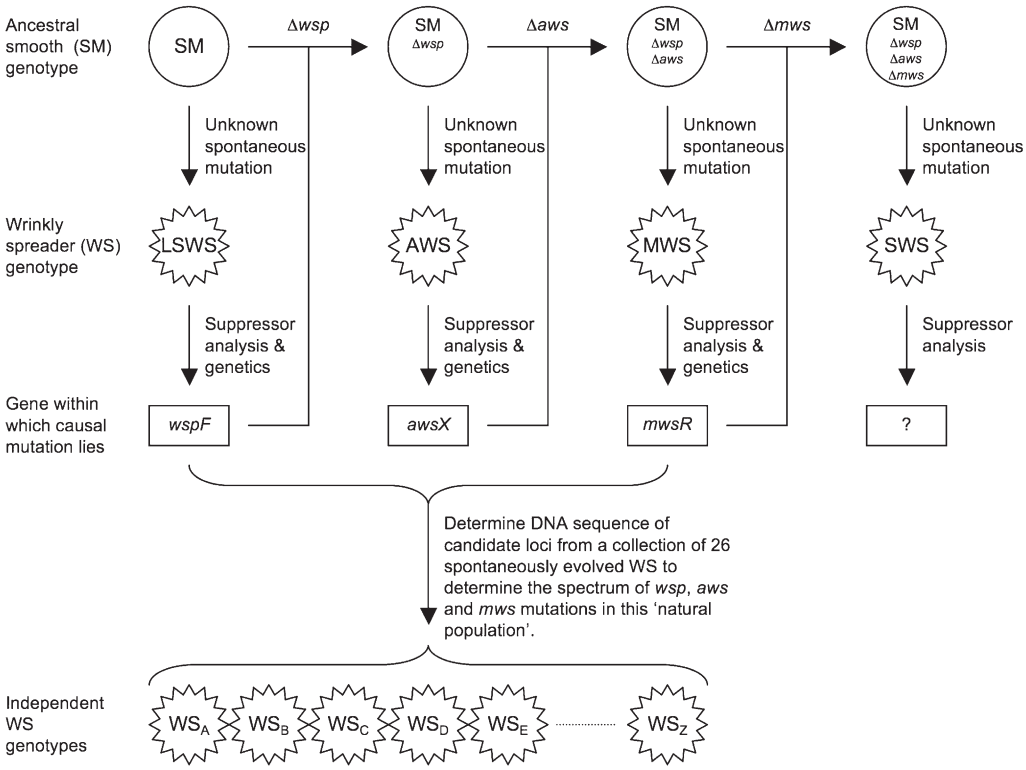


Figure 1.—Outline of experimental strategy for elucidation of WS-generating mutations and their subsequent identity and distribution among a collection of independently evolved, spontaneously arising WS genotypes. The strategy involves, first, the genetic analysis of a specific WS genotype (*e.g.*, LSWS) to identify the causal mutation, and second, a survey of DNA sequence variation at specific loci known to harbor causal mutations among a collection of spontaneously arising WS genotypes. For example, suppressor analysis of LSWS using a transposon to inactivate genes necessary for expression of the wrinkly morphology delivered a large number of candidate genes (top left) (Spiers *et al.* 2002). Genetic and functional analysis of these candidate genes (*e.g.*, Goymer *et al.* 2006) led eventually to the identity of the spontaneous mutation (in *wspF*) responsible for the evolution of LSWS from the ancestral SM genotype (Bantinaki *et al.* 2007). Subsequent analysis of the *wspF* sequence among 26 independent WS genotypes (bottom) showed that 50% harbored spontaneous mutations (of different kinds; see Table 1) in this gene. Failure to account for the mutational origin of all 26 independent WS genotypes by changes in *wspF* indicated the existence of non-*wsp* mutational routes to WS. To identify such routes, the *wsp* operon from the ancestral genotype was deleted (top left), and this genotype was selected in a static broth microcosm. After 5 days, populations were sampled and checked for the presence of genotypes that had converged upon the WS phenotype. Such WS genotypes were detected and one, AWS, was subjected to detailed genetic analysis (this study). Upon identification of the causal mutation in *awsX*, the population of independent WS genotypes (bottom) was again surveyed to see whether any harbored mutations at this locus. A further iteration, which resulted in the identification of *mwsR*, was necessary to account for the mutational origin of all 26 spontaneously arising WS genotypes (see text). Specific details of the mutations are in Table 1.

the broth phase (Bantinaki *et al.* 2007). All strains were grown overnight in shaken KB broth before introduction into spatially structured microcosms ($\sim 10^5$ cells of each competitor) at a ratio of 1:1:1. Relative fitness was calculated as the ratio of Malthusian parameters of the WS strains being compared (Lenski *et al.* 1991).

RESULTS

In previous studies, two complementary experimental approaches were used to identify the spectrum of spontaneous mutations causing the WS phenotype (Figure 1). The first approach involved detailed genetic analysis of a single, defined, spontaneously generated WS genotype to identify the "genes that matter," elucidate their function, and ultimately identify the causal mutation (Spiers *et al.* 2002; Goymer *et al.* 2006; Bantinaki *et al.* 2007). The second approach—possible only after identification of the causal mutation—was to survey a population of independently acquired, spontaneously generated WS genotypes for mutations at the locus of interest (Bantinaki *et al.* 2007). Previous work focused on the WS genotype termed LSWS (PBR663) of which the mutational cause was found to reside in *wspF*. Survey of the nucleotide sequence of *wspF* from 26 independently acquired spontaneous WS genotypes showed that just one-half harbored mutations in *wspF* (Bantinaki *et al.* 2007). Here, to define the unknown mutational routes to WS, the strategy outlined in Figure 1 was repeated, but with each repetition initiated by an ancestral genotype lacking the genes encoding the previously identified mutational path.

Suppressor analysis of LSWS: The strategy devised for identification of *wsp*-independent mutation routes involved (1) deletion of the *wsp* operon from the ancestral genotype; (2) reevolution of WS from the *wsp*-deficient starting genotype (convergence); (3) suppressor analysis of one selected *wsp*-independent WS; and (4) comparison of the results of the *wsp*-deficient WS suppressor study with those from the suppressor analysis of LSWS. To be successful, a comprehensive set of suppressor mutations from both LSWS and the non-

wsp WS must be available so that differences—and thus the likely identity of the non-*wsp* pathway—can be identified. It is important to note that the suppressor analysis carried out here, as in previous studies, is a gene inactivation strategy involving transposon mutagenesis and that, by inactivating genes that determine the WS phenotype, it becomes possible to catalog the set of "candidate genes" that underpin this morphotype. Such knowledge has been an essential first step toward identification of spontaneous mutations causing the WS phenotype (as, for example, outlined in previous articles: Spiers *et al.* 2002; Goymer *et al.* 2006; Bantinaki *et al.* 2007).

The initial suppressor analysis of LSWS reported in Spiers *et al.* (2002) was performed prior to the availability of the SBW25 genome sequence; dependence on cloning, mapping, and cosmid libraries resulted in a limited set of defined WS mutants (8000 mutants were screened; 27 unique WS negative mutants were identified and the nature of genetic defect was determined in 11 instances; Spiers *et al.* 2002). To exploit the comparative suppressor strategy, the suppressor analysis of LSWS was extended using IS*phoA*/hah-Tc (Bailey and Manoil 2002). A total of 113,000 transposon mutants were screened for loss of the wrinkled colony morphology (assuming a Poison distribution and a genome containing 6700 genes, the probability of a gene not being inactivated by a transposon, given the screening of 113,000 mutants, is <0.001). A total of 134 WS suppressor mutants were obtained, and the position of the transposon was determined in each instance. The results are reported in File S1.

All loci previously identified (Spiers *et al.* 2002) were found in this study: 47 (35%) insertions were located in *wss* and 26 (19%) in *wsp*; in addition, insertions were found in genes responsible for the maintenance of rod-shaped cells (*e.g.*, *mreB*, *rodA*, and *pbpA*), indicating that cell shape is an essential determinant of the WS phenotype (Spiers *et al.* 2002). Two previously undiscovered loci were also identified. The first is an eight-gene locus, unique to SBW25, with a predicted role in cell-wall biogenesis (pflu1661–pflu1668). The locus was termed *wsw* (**w**rinkly **s**preader cell-**w**all biogenesis locus). The second locus is composed of five genes (pflu0475–pflu0479) and includes a transcriptional regulator, a membrane-bound de-*N*-acetylase, and two glycosyl transferases—also with a possible role in the modification of cell-wall material. This locus was termed *wsm* (**w**rinkly **s**preader cell-wall **m**odification locus). Eight transposon insertions were found in *wsw* and nine in *wsm*. Mutations in *wsw*, while eliminating the wrinkled morphology of LSWS on an agar plate, did not abolish the ability to form a mat at the air–liquid interface of static broth microcosms (Gehrig 2005).

Aws: a *wsp*-independent mutational route to WS: To identify the first *wsp*-independent mutational route to WS the entire *wsp* operon was deleted from the genome of the ancestral SM genotype (Figure 1). This SM Δ*wspABCDEFR* genotype (PBR693) was phenotypically indistinguishable from the ancestral genotype. When propagated in a spatially structured microcosm, a characteristic WS mat was observed within 5 days at which point the microcosm was destructively harvested and samples were plated on KB agar. Morphological diversity was no different from the previously described 5-day adaptive radiation of the ancestral genotype (Rainey and Travisano 1998). The three major morphological classes were all present, but WS genotypes were the dominant type: a single WS genotype, designated **a**lternate **w**rinkly **s**preader [AWS (PBR210)], was selected for suppressor analysis using IS*phoA*/hah-Tc as described for LSWS.

Approximately 95,000 colonies were screened for loss of the wrinkled morphology: the position of the transposon was determined in 186 mutants, and the results are reported in File S2. Comparison of the results of AWS and LSWS suppressor studies revealed many loci in common; *e.g.*, 64 insertions were identified in *wss* (35%), 34 were found in *wsw* (18%), and 13 in *wsm* (7%), as well as two unique loci. The first was the *muc* locus (Boucher *et al.* 1996) comprising genes pflu1467–pflu1471 (*algU mucABD*) and containing transposon insertions (2 each) in *mucA* and *mucD*. The second locus was defined by 11 insertions in a 2-kb region spanning pflu5211 (2 insertions) and pflu5210 (9 insertions). Comparison with DNA and protein sequence databases showed that these two genes (and a third, pflu5209) are conserved across Pseudomonas and other protobacteria but have no known function. Together the genes form a putative operon termed *aws* (pflu5211 *awsX*, plfu5210 *awsR*, and pflu5209 *awsO*) (Figure 3).

Mutations in the *muc* locus of *P. aeruginosa* have been implicated in expression of the mucoid colony type, and while of some considerable interest (see discussion), our initial attention focused on the *aws* locus because of the presence within this cluster of a di-guanylate cyclase.

Predicted proteins of the Aws locus: AwsX is 171 amino acids in length; with the exception of a proteolytic cleavage site between residues 31 and 32 (suggesting that AwsX is periplasmic), the protein has no defining motifs or database matches. AwsR is 420 amino acids in length and contains two *trans*-membrane domains at the N terminus (residues 19–41 and 151–173), a HAMP domain (residues 159–229, Pfam *E*-value 1.4×10^{-7}), and a GGDEF domain (residues 241–403, Pfam *E*-value 9.2×10^{-58}): the GGDEF domain contains all conserved motifs characteristic of a cyclic-di-GMP (bis-3′-5′-cyclic di-GMP) synthesizing DGC. In addition to the invariable DGC motifs, there is a potential "I-site": the presence of this motif (R**D) indicates that cyclic-di-GMP binds allosterically to AwsR. I-sites have been shown to enable negative feedback control of DGC activity (Christen *et al.* 2006). AwsO is 163 amino acids and contains an OmpA (MotB) domain (residues

TABLE 1

Mutational causes of WS

WS genotype	Gene	Nucleotide change	Amino acid change	Source/reference
LSWS	*wspF*	A901C	S301R	Bantinaki *et al.* (2007)
AWS	*awsX*	Δ100-138	ΔPDPADLADQRAQA	This study
MWS	*mwsR*	G3247A	E1083K	This study
WS_A	*wspF*	T14G	I5S	Bantinaki *et al.* (2007)
WS_B	*wspF*	Δ620-674	P206Δ (8)[a]	Bantinaki *et al.* (2007)
WS_C	*wspF*	G823T	G275C	Bantinaki *et al.* (2007)
WS_D	*wspE*	A1916G	D638G	This study
WS_E	*wspF*	G658T	V220L	Bantinaki *et al.* (2007)
WS_F	*wspF*	C821T	T274I	Bantinaki *et al.* (2007)
WS_G	*wspF*	C556T	H186Y	Bantinaki *et al.* (2007)
WS_H	*wspE*	A2202C	K734N	This study
WS_I	*wspE*	G1915T	D638Y	This study
WS_J	*wspF*	Δ865-868	R288Δ (3)[a]	Bantinaki *et al.* (2007)
WS_K	*awsO*	G125T	G41V	This study
WS_L	*wspF*	G482A	G161D	Bantinaki *et al.* (2007)
WS_M	*awsR*	C164T	S54F	This study
WS_N	*wspF*	A901C	S301R	Bantinaki *et al.* (2007)
WS_O	*wspF*	Δ235-249	V79Δ (6)[a]	Bantinaki *et al.* (2007)
WS_P	*awsR*	222insGCCACCGAA	74insATE	This study
WS_Q	*mwsR*	3270insGACGTG	1089insDV	This study
WS_R	*mwsR*	T2183C	V272A	This study
WS_S	*awsX*	C472T	Q158STOP	This study
WS_T	*awsX*	Δ229-261	ΔYTDDLIKGTTQ	This study
WS_U	*wspF*	Δ823-824	T274Δ (13)[a]	Bantinaki *et al.* (2007)
WS_V	*awsX*	T74G	L24R	This study
WS_W	*wspF*	Δ149	L49Δ (1)[a]	Bantinaki *et al.* (2007)
WS_X[b]	?	?	?	This study
WS_Y	*wspF*	Δ166-180	Δ(L51-I55)	Bantinaki *et al.* (2007)
WS_Z	*mwsR*	G3055A	A1018T	This study

[a] P206Δ(8) indicates a frameshift; the number of new residues before a stop codon is reached is in parentheses.

[b] Suppressor analysis implicates the *wsp* locus (17 transposon insertions were found in this locus). However, repeated sequencing failed to identify a mutation.

52–147, Pfam *E*-value 4.2×10^{-26}), suggesting a porin-like function (Sugawara and Nikaido 2004).

Mutations in awsX are necessary and sufficient for the AWS phenotype: DNA sequence analysis of the *aws* locus from AWS revealed a 39-bp deletion spanning nucleotides 100–138 of *awsX* [*awsX* Δ100-138 (AwsX Δ34-46)]. The deletion is in-frame and flanked by a five-nucleotide repeat (AGGCG); the intervening sequence plus one copy of the repeat is absent in AWS.

Armed with knowledge of a mutation in *awsX*, the DNA sequence of this locus was determined in the 13 independent WS genotypes lacking mutations in *wspF* (Bantinaki *et al.* 2007; see Figure 1). In three of these genotypes, WS_S, WS_T, and WS_V, mutations were found in *awsX*. WS_T contains an in-frame deletion (33 nucleotides) spanning residues 229–261 and flanked by a six-nucleotide repeat (ACCCAG). As in AWS, the deletion in WS_T removed the intervening nucleotides, leaving one copy of the repeated sequence. Both WS_S and WS_V contain single-nucleotide polymorphisms (a transition and transversion, respectively; see Table 1).

To determine whether the *awsX* mutations alone are sufficient to cause the WS phenotype, both the AWS and WS_T mutations were reconstructed in the ancestral genetic background by allelic exchange. Three independent exchanges were made for each allele and in each instance the resulting phenotype was WS.

AwsX is a negative regulator: Previous studies showed that the proximate cause of the LSWS phenotype is overactivation of the WspR DGC (Goymer *et al.* 2006; Bantinaki *et al.* 2007). The fact that the Aws locus also includes a putative DGC (AwsR) suggested the possibility that overactivation of AwsR might cause the WS phenotype. While there are no clues in the amino acid sequence to indicate the means by which activation of the DGC domain might occur, the fact that mutations in *awsX* cause the WS phenotype suggested that AwsX might suppress the activity of AwsR. However, the fact that all *awsX* mutations are in-frame (and in two cases are deletions between repeats) left open the possibility that AwsX might be a positive activator—the activated state being realized upon specific in-frame deletions

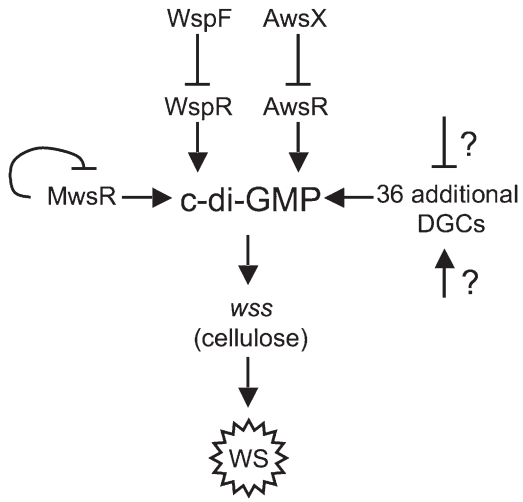


Figure 2.—Network diagram of DGC-encoding pathways underpinning the evolution of the WS phenotype and their regulation. Overproduction of c-di-GMP results in overproduction of cellulose and other adhesive factors that determine the WS phenotype. The ancestral SBW25 genome contains 39 putative DGCs, each in principle capable of synthesizing the production of c-di-GMP, and yet WS genotypes arise most commonly as a consequence of mutations in just three DGC-containing pathways: Wsp, Aws, and Mws. In each instance, the causal mutations are most commonly in the negative regulatory component: *wspF*, *awsX*, and the phosphodiesterase domain of *mwsR* (see text).

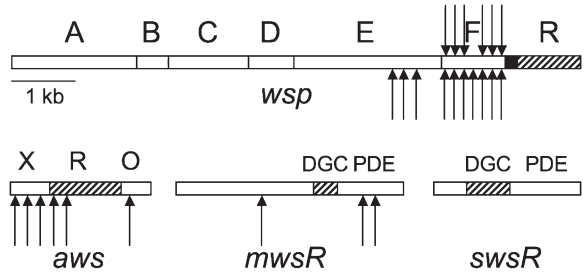


Figure 3.—Genetic organization of the *wsp*, *aws*, *mws*, and *sws* loci and location of mutations causing WS. The *wsp* and *aws* loci are composed of seven and three genes, respectively; *mws* and *sws* are single genes. DGC-encoding genes in *wsp* and *aws* and DGC domains within *mws* and *sws* are indicated by hatching. PDE denotes the phosphodiesterase domain. Deletion of the negative regulatory components (*wspF*, *awsX*, and the PDE domain of *mwsR*) from the ancestral SM genotype results in the expression of the WS morphology. Arrows show the locations of the mutations causing the WS phenotype among the 26 independently obtained (spontaneously arising) WS genotypes used in this study (specific details of the mutations are in Table 1). The mutational basis for the activation of SwsR is unknown.

[some support for a positive regulatory role of AwsX came from an earlier study (Gehrig 2005)].

To test directly the hypothesis that AwsX is a negative regulator, the complete *awsX* reading frame was deleted from the ancestral SM genotype. Constructing such a mutant in which the *awsR* open reading frame was fused to the ATG start codon of *awsX* (replacing *awsX*) proved impossible. However, deletion of *awsX* was achieved when the translational coupling between *awsX* and *awsR* that was evident in the genome sequence of the ancestral genotype was incorporated into the design of the deletion construct. SM Δ*awsX* (PBR717) was phenotypically WS and produced copious calcofluor-staining material (the cellulosic product of the *wss* operon). This demonstrates that AwsX is a negative regulator. That AwsX is the likely negative regulator of AwsR—and not some other DGC—is strongly suggested by both the genetic organization of the locus and the exhaustive suppressor analysis of AWS in which the only DGC involved in expression of the wrinkled phenotype was AwsR. Drawing upon the distinctive parallels with LSWS (where the negative regulator WspF controls the activation of WspR), we suggest that the proximate cause of AWC is overproduction of the acetylated cellulose synthesizing the *wss* operon via the synthesis of c-di-GMP as a consequence of the constitutive activation of AwsR (Figure 2).

MWS: a *wsp*- and *aws*-independent route to WS: While identification of the *aws* mutational route to WS reduced the number of unknown WS-causing mutations among the 26 independent WS genotypes, the existence of WS genotypes with no mutations in either *wspF* or *awsX* raised the possibility of still further pathways. To identify these, the strategy established above (Figure 1) was repeated using the transposon IS-Ω-kan/hah (Giddens *et al.* 2007). This time the starting genotype was the ancestral SM genotype from which both the *wsp* and *aws* loci had been cleanly deleted: SM Δ*wspABCDEFR* Δ*awsXRO* (PBR712).

PBR712 is phenotypically indistinguishable from the ancestral SM genotype and diversified when propagated in a static broth microcosm in a manner fully in accord with the ancestor. When a sample from a 5-day-old microcosm was diluted and plated on KB, WS genotypes were evident. One WS colony was selected for further study and named MWS (**M**ike's **w**rinkly **s**preader: PBR714). Suppressor mutations of MWS that returned the colony phenotype to smooth were sought as above.

A total of 91 WS defective mutants were obtained. Transposon insertions in *wss* (seven insertions) were eliminated by PCR screening (see materials and methods). Given that in both LSWS and AWS the mutational route to WS involved a DGC-containing locus, the remaining suppressor mutants were analyzed only to the point at which several insertions were found in a previously unencountered DGC. After determining the insertion point of the transposon in just 19 suppressor mutants, 7 were found within a large (3.8 kb) gene (Pflu5329), named *mwsR*, which encoded several protein domains including a DGC (Figure 3). The results are reported in File S3. Flanking *mwsR* in SBW25 is *glyA* (a serine hydroxy methyl transferase) and a putative membrane protein. MwsR is present in

most of the sequenced Pseudomonas genomes; where present, it is flanked by *glyA* (always on the downstream side of *mwsR*). A previous study of *P. aeruginosa* and *P. putida* assigned a biofilm phenotype to an ortholog of MwsR named MorA (Choy *et al.* 2004).

The N terminus of MwsR consists of two *trans*-membrane regions: residues 477–569 define a PAS_3 domain (Pfam *E*-value 1.3×10^{-11}), both residues 666–712 and 723–834 define PAS domains (Pfam *E*-values 5.8×10^{-5} and 1.3×10^{-10}, respectively); residues 848–1009 define a GGDEF domain (Pfam *E*-value 7.3×10^{-61}), and residues 1029–1264 define an EAL domain (Pfam *E*-value 3.5×10^{-90}). The GGDEF domain contains all the motifs characteristic of a DGC, although it lacks the I-site. The EAL domain is a predicted phosphodiesterase (PDE), which facilitates hydrolyzation of c-di-GMP to pGpG (Tamayo *et al.* 2005). PDE domains are often found adjacent to DGC domains, leading to the suggestion that DGC/PDE domain proteins may both synthesize and degrade c-di-GMP, although a hybrid protein with both domains functional has not been described (Jenal and Malone 2006).

Predicting the mutational cause of the MWS: Emerging from the analysis of LSWS (where mutations in *wspF* are the cause of the WS phenotype) and AWS (where mutations in *awsX* cause the WS phenotype) is an understanding of the significance of loss-of-function mutations in negative regulators of DGCs. Applying this emerging trend to MwsR led to the prediction that causal mutations might reside in the EAL domain (if both PDE and DGC activities exist in MwsR, then the balance is likely to be crucial: a mutation that reduced or abolished PDE function would cause an oversupply of c-di-GMP with ensuing effects on the production of the cellulosic polymer).

DNA sequence analysis of *mwsR* from MWS revealed a single nonsynonymous nucleotide substitution in the EAL domain (G3247A), resulting in the substitution of lysine for glutamic acid. Sequence analysis of *mwsR* from the collection of 26 independent WS identified three additional *mwsR* mutations: WS_Q contains a six-nucleotide insertion after position 3270 arising from duplication of the preceding six nucleotides; WS_R and WS_Z harbor nonsynonymous mutations (Table 1).

Mutations in mwsR are necessary and sufficient for the MWS phenotype: To test the hypothesis that mutations in *mwsR* are the direct cause of the WS phenotype, two mutant alleles (G3247A and T2183C) were recombined into the ancestral genome replacing the wild-type sequence. Three independent allelic replacements were made for each allele, and in all instances the resulting genotype was WS (each WS genotype tested positive for the production of calcofluor-staining material, the cellulosic product of the *wss* operon).

The EAL domain negatively regulates MwsR activity: Attributing functions to multi-domain proteins presents numerous challenges. We sought to understand the relationship between the EAL and GGDEF domains in MwsR; specifically, we asked whether the EAL domain negatively regulates the GGDEF domain and whether the GGDEF domain is the source of c-di-GMP that causes the WS phenotype. To address the first question, nucleotides 3116–3670 (encoding the EAL domain) were deleted from the ancestral SM genotype: this genotype (PBR935) was morphologically WS. While the proximate cause of WS, in every case examined so far, is a consequence of the oveproduction of c-di-GMP, the source of c-di-GMP caused by deletion of the EAL domain of MwsR need not be the GGDEF domain of MwsR. The fact that no DGC activity has been detected in proteins containing both EAL and GGDEF domains (Jenal and Malone 2006) raised the possibility that the activated GGDEF domain might not reside within MwsR. To explore this further, the remainder of the *mwsR* open reading frame was removed from PBR935, and the resulting genotype (SM Δ*mwsR*; PBR711) was morphologically smooth. If the primary function of MwsR is as a PDE (*i.e.*, the GGDEF domain is nonfunctional) whose activity balanced the DGC activity encoded by some other protein, then the deletion of the GGDEF domain of MwsR in PBR935 would have left the WS phenotype unaltered. Together these data provide evidence that both the EAL and the GGDEF domains of MwsR are functional and that the PDE activity of the EAL domain negatively regulates the DGC activity of the GGDEF domain. Moreover, the fact that the PDE activity, and the allelic exchange mutants (*mwsR* G3247A and T2183C), produce cellulosic material indicates that the net effect of the overactive DGC domain is once again overproduction of acetylated cellulose due to the overproduction of c-di-GMP.

Fitness of AWS and MWS: In all cases, replacement of the wild-type allele by each mutant allele (in the ancestral genetic background) resulted in colony morphology and mat-forming ability that was indistinguishable from that of the original WS mutant. To assess the fitness consequences of these mutations, the fitness of mutants generated by allelic exchange was determined relative to a reference WS (the LSWS genotype) and compared with the fitness of the original mutants [again relative to LSWS containing a *lacZ* marker (PBR741)]. These fitness assays included the ancestral SM genotype whose superior growth in the broth phase ensures that differences between the competing WS strains is due to competitive ability at the air–liquid interface (Bantinaki *et al.* 2007). In all cases, the fitness of the reconstructed mutant was indistinguishable from the fitness of the original mutant, indicating that the mutation identified in each derived genotype accounts fully for both phenotypic and fitness differences between ancestral and derived types. The spectrum of fitness effects is shown in Figure 4, and its legend includes the results of the statistical analyses.

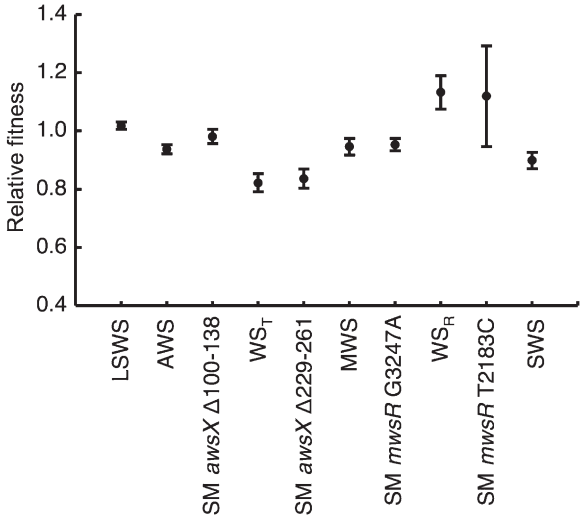


FIGURE 4.—Relative fitness of WS genotypes. Fitness was determined relative to the LSWS genotype carrying a neutral *lacZ* marker. The experiment was designed to test the hypothesis that the mutations identified in the WS genotypes AWS and MWS are causal and to test the hypothesis that the fitness of SWS differs significantly from the known *wspF*-, *awsX*-, and *mwsR*-generated WS genotypes. *t*-Tests revealed no significant differences between derived WS and the reconstructed genotypes for AWS *vs.* SM *awsX* Δ100-138 ($P = 0.081$), WS_T *vs.* SM *awsX* Δ229-261 ($P = 0.381$), MWS *vs.* SM *mwsR* G3260A ($P = 0.424$), and WS_X *vs.* SM *mwsR* T2183C ($P = 0.497$). One-way ANOVA showed a significant difference among the means for the entire data set ($F_{[8,\ 52]} = 10.12$, $P < 0.0001$). Tukey's Honestly Significant Differences test showed that the fitness of SWS was not significantly different from LSWS, MWS, or AWS. Data are means and standard errors of six replicates.

The relative contribution of *wsp*, *aws*, and *mws* to the evolution of WS variation: By the fifth day of propagation, WS genotypes were detected in spatially structured microcosms founded by the ancestral genotype from which the *wsp* operon had been deleted (SM Δ*wsp*); within the same time period, WS genotypes also arose from the ancestral genotype from which both *wsp* and *aws* were deleted (SM Δ*wsp* Δ*aws*). To systematically investigate the contribution of each locus (*wsp*, *aws*, and *mws*) to the production of WS variation, a set of deletion mutants was generated in which all possible single-, double-, and triple-locus mutants were constructed: SM Δ*wsp* (PBR693), SM Δ*aws* (PBR681), SM Δ*mws* (PBR711), SM Δ*wsp* Δ*aws* (PBR712), SM Δ*wsp* Δ*mws* (PBR713), SM Δ*mws* Δ*aws* (PBR718), and SM Δ*wsp* Δ*aws* Δ*mws* (PBR716). Replicated assays in KB broth showed no differences in growth dynamics among genotypes (see Figure S1). Each genotype was inoculated into a series of independent broth-filled microcosms ($\sim 5 \times 10^6$ cells/microcosm), and one microcosm from each series was destructively harvested on a daily basis over the course of 9 days (each assay was replicated threefold). Cells were diluted and plated, and following 48 hr growth on KB agar, the phenotype of 500 colonies from each microcosm was checked for the presence of WS genotypes. With the exception of microcosms founded with the triple-deletion genotype (PBR716), WS genotypes were detected in each and every 2-day-old microcosm. WS genotypes also were detected in microcosms founded by PBR716 at day 5.

SWS: a *wsp*- *aws*-, and *mws*-independent route to WS: A single WS genotype, **s**low **w**rinkly **s**preader (SWS; PBR715), derived from SM Δ*wsp* Δ*aws* Δ*mws* (PBR716), was selected and subjected to suppressor analysis as described above. A total of 41 suppressor mutants were identified (the results are reported in File S4), and transposon insertions were found in previously identified loci, including *wss* (three were identified by PCR screen). As in the analysis of MWS, characterization of suppressor mutants was halted once a previously unencountered GGDEF domain-encoding gene was identified. In this instance, two insertions were found in a single open reading frame (pflu1349) predicted to encode a GGDEF and EAL domain-containing protein (Figure 3). The predicted product of this gene, designated *swsR*, is a protein of 759 amino acids and has two repeat regions and two domains: residues 76–137 and 139–200 define two MHYT N-terminal repeats (GALPERIN *et al.* 2001; Pfam *E*-value 1.2×10^{-15} and 3.5×10^{-13}, respectively) with seven probable membrane-spanning regions between residues 32 and 262; residues 295–452 define a GGDEF domain (Pfam *E*-value 1.6×10^{-56}), and residues 472–708 define an EAL domain (Pfam *E*-value 3.5×10^{-79}).

Deletion of *swsR* from SWS caused the phenotype to revert to the ancestral smooth type, confirming its causal role; however, repeated sequencing of pflu1349 (including the promoter region) revealed no mutations. The fitness of SWS was determined relative to LSWS *lacZ* as described above and found to be statistically indistinguishable from the fitness of LSWS, SWS, MWS, and AWS (Figure 4). Furthermore, with a fitness 0.89 relative to LSWS, the SWS genotype is at the midpoint of the fitness measures determined for WS genotypes (BANTINAKI *et al.* 2007).

The mutational origins of the remaining independent WS genotypes: The data thus far indicate the existence of three—and only three—readily achievable single-step mutational routes to WS: *wsp*, *aws*, and *mws*. Clearly, the discovery of SWS indicates that WS genotypes may arise by additional routes; however, the fact that SWS—with a fitness equivalent to other characterized WS genotypes—took 3 days longer to arise from the "ancestral" SM Δ*wsp* Δ*aws* Δ*mws* (PBR716) genotype than from any of the single- or double-locus ancestral mutants indicates that these other mutational routes, while obtainable, are less likely to be realized. Assuming equivalent per-locus mutation rates, we conclude that the mutational route involving *swsR*, and numerous other potential non-*wsp*, non-*aws*, and non-*mws* genetic routes to WS (see DISCUSSION), is rarely available to selection because of a reduced capacity to translate

mutation into WS variation. If this is true, then the mutational origins of the 26 WS genotypes from independent adaptive radiations should be accounted for by mutations in the three known loci (*wsp*, *aws*, or *mws*). However, sequencing of *wspF*, *awsX*, and *mwsR* from the 26 WS genotypes shows that mutation in these loci accounts for just 19 of the 26 WS genotypes. One possibility is that these genotypes do contain mutations in *wsp*, *aws*, or *mws*, but in components other than *wspF*, *awsX*, and *mwsR*. This led us to return to the independent WS genotypes from which mutations had not been identified.

WS genotypes WS_D, WS_H, WS_I, WS_K, WS_M, WS_P, and WS_Z were each subjected to suppressor analysis using the transposon IS-Ω-kan/hah. The position of the transposon was mapped in each suppressor mutant: in each instance, transposon insertions were found in either *wsp* or *aws*: WS_D, *wsp*; WS_H, *wsp*; WS_I, *wsp*; WS_K, *aws*; WS_M, *aws*; WS_P, *aws*; and WS_z, *wsp* (failure to find insertions in *mws* is consistent with *mwsR* being the sole component of this locus). This strongly implicated the *wsp* and *aws* loci as mutational targets. The DNA sequence of the entire *wsp* or *aws* locus was therefore obtained from each of these independent WS genotypes (the locus sequenced was that indicated by the results of the suppressor analysis) and compared with the ancestral sequence. In six of the seven WS genotypes, a mutation was identified—WS_D, WS_H, WS_I, WS_K, and WS_M—each containing a single nonsynonymous substitution (WS_D,WS_H, and WS_I harbor a mutation in the *wspE* kinase, WS_K contains a mutation in *awsO*, and WS_M contains a mutation in *awsR*). WS_P also carries a defect in *awsR*, but the mutation is caused by a nine-nucleotide (in-frame) insertion (GCCACCGAA) after position 222 (details are given in Table 1; see also Figure 3). Repeated sequencing of *wsp* from WS_X revealed no mutation in any component of this locus, even though this locus is firmly implicated by the results of the suppressor analysis. The sequence of *mwsR* and *awsXRO* was also wild type for WS_X.

DISCUSSION

Systematic change in the genetic structure of populations is brought about by the action of natural selection. But the opportunity for selection to act is dependent upon the presence of adaptive phenotypic variation. The rate at which changes in nucleotide sequence arise (mutation rate) has received (*e.g.*, Drake *et al.* 1998; Hudson *et al.* 2002) and continues to receive considerable attention, but poorly understood are the processes that connect variation in nucleotide sequence with phenotypic variation, the raw material for adaptive evolution. Despite important theoretical advances (reviewed in Rice 2008), knowledge of the rate (per generation, per locus) at which DNA sequence change (mutation) translates into phenotypic variation is scant. Nonetheless, there is no doubt that the value of such knowledge, a measure of an organism's capacity for evolution, would be considerable (Hansen 2006).

In this study, we have defined in some detail two new loci (*aws* and *mws*) within which mutation can generate the WS phenotype; these loci, along with the previously uncovered *wsp* locus, comprise three approximately equally achievable single-step mutational routes to the adaptive WS phenotype. The three pathways are remarkably different from one another, but share two features in common: each contains a gene that encodes a protein with putative—although proven for *wspR*—di-guanylate cyclase activity (*wspR*, *awsR*, and *mwsR*) and in each case the protein is subject to negative regulation (Figure 2). That these pathways do comprise the primary routes by which WS variation arises during the course of the model Pseudomonas radiation is strongly supported by the fact that 25 of 26 WS genotypes from independent radiations harbor mutations in one of the three identified loci (Table 1). WS_X, the one WS genotype with no mutation in *wsp*, *aws*, or *mws* nonetheless requires a functional Wsp pathway for expression of the phenotype, indicating that the causal mutation in WS_X is accessory to, but part of, the Wsp regulatory pathway.

On the basis of previously published genetic and biochemical evidence (in the case of *wsp*, see Goymer *et al.* 2006 and Malone *et al.* 2007) and the results of the analyses presented here for *aws* and *mws*, we infer that in each instance the proximate cause of WS is overproduction of c-di-GMP, which is itself a consequence of constitutive activation of the corresponding DGC (WspR, AwsR, or MwsR). Interestingly, despite the diverse nature of the *wsp*, *aws*, and *mws* regulatory modules, the discovery of transposon insertions in the acetylated cellulose polymer-encoding *wss* operon from suppressor analysis of LSWS, AWS, and MWS (and also SWS) indicates that the primary structural determinant of WS in each instance is the acetylated cellulose polymer (Spiers *et al.* 2002). This raises a number of questions as to the nature of the regulatory connections between the DGC-encoding genes and the product of the Wss locus in the ancestral genotype. At this stage, it is unclear whether the primary role of the numerous DGC-encoding genes in the genome [of which there are 39 (Silby *et al.* 2009)] is, in each instance, to regulate the production of cellulose (presumably in response to different input signals) or whether the many-to-one mapping revealed in our study of WS genotypes is a consequence of the co-option (Gerhart and Kirschner 1997) of DGC-encoding modules whose regulatory targets have been altered as a consequence of constitutive activation. This latter possibility is not improbable, given increasing awareness of the tight spatial and temporal regulation of DGC activity (Jenal and Malone 2006).

A related issue concerns the degree of overlap and connectivity (pleiotropy and epistasis) of the architec-

ture that defines the WS phenotype in WS genotypes whose phenotype arises from overactivation of different DGCs (Knight *et al.* 2006). A comparison of the results of the comprehensive suppressor studies of LSWS and AWS indicates that the underlying architectures are subtly different: expression of the AWS phenotype requires the *algU mucABD* locus, but this locus appears to have no role in expression of the LSWS phenotype. Just what role the *muc* locus plays in AWS is unclear; however, studies in *P. aeruginosa* show that AlgU is an alternate σ factor (σ^{22}) (Yu *et al.* 1995), MucB an anti-σ factor (of AlgU), and MucC a negative regulator of alginate biosynthesis (Schurr *et al.* 1996). In the ancestral genotype of *P. fluorescens* SBW25, it was previously shown that AlgR is a positive activator and AmrZ (AlgZ) a negative regulator of *wss* transcription (Giddens *et al.* 2007), so the discovery of genes in *P. aeruginosa* that are linked with alginate regulation does not necessarily imply the involvement of alginate in the WS phenotype. Nonetheless, the regulatory nature of this locus and its specific involvement in the expression of the wrinkled morphology of AWS points to further complexity and a network of overlaping regulatory connections, all fertile material for the evolution of phenotypic innovation (Landry *et al.* 2007).

Parallel molecular evolution is influenced by genetic architecture: Discovery that the WS phenotype is repeatedly generated by mutations in a small number of loci provides a further example of parallel molecular (and phenotypic) evolution: for example, 13 of the 26 independent WS genotypes have mutations in *wspF* (WS_N is identical to LSWS), 3 are in *wspE*, 3 are in *awsX*, 4 are in *mwsR*, and 2 have mutations in *awsR* (*awsO* was the only singleton). Some have argued that such striking parallelism is evidence of selection (*e.g.*, Wichman *et al.* 1999; Woods *et al.* 2006); indeed, measurements of the fitness effects of specific mutations following allelic replacements in the ancestral background support this claim. However, our data show that genetic architecture—a consequence of specific interactions and functionalities within the *wsp*, *aws*, and *mws* regulatory modules—biases the spectrum of WS-generating mutations presented to selection.

The clearest evidence of this bias comes from the fact that the ancestral genotype devoid of *wsp*, *aws*, and *mws* [SM Δ*wsp* Δ*aws* Δ*mws* (PBR716)] can generate WS variation via mutations that activate previously unencountered DGCs (in this example, SwsR). This demonstrates that *wsp*, *aws*, and *mws* are not the only mutational routes to WS, and yet the mutational causes of the 26 independent WS genotypes are confined to *wsp*, *aws*, and *mws*. A clue to the cause of this bias comes from the fact that, in experimental adaptive radiations, WS genotypes arising from PBR716 take 3 days longer to reach detectable frequencies compared to WS genotypes arising from the ancestor [or the ancestral genotype devoid of any one or two (but not three) of *wsp*, *aws*, or *mws*]. Given that the fitness of WS generated from PBR716 (SWS) is equivalent to LSWS, AWS, and MWS, and assuming the mutation rate underlying all possible WS-generating mutations is not different from the background spontaneous mutation rate (there is no indication of mutable loci among the spectrum of mutations in Table 1), then the most likely explanation for the bias stems from some feature of the genetic architecture underlying the *wsp*, *aws*, and *mws* pathways that affects the likelihood that spontaneous mutations in these loci are translated into WS variation. If the likelihood is greatest in *wsp*, *aws*, and *mws*, then selection will only rarely see WS variation caused by mutation in other loci. Parallel genetic evolution, therefore, in this instance, is in part a product of the genetic constraints that bias the production of WS variation.

Just why the other mutational routes to WS are less readily achievable, given that the ancestral genome would appear to offer numerous potential routes for WS evolution, is not clear. It is possible that some, such as *swsR* (the proximate cause of the SWS genotype), require very specific mutations to generate WS. Others may require two mutations, while some DGCs may be nonfunctional. Still others, although readily activated, may be unable to interact with the cellulose synthases encoded by Wss and thus be unable to deliver c-di-GMP to the cellulose synthases and related structural components at the appropriate time.

Further and more detailed evidence of bias in the production of WS variation exists within the *wsp*, *aws*, and *mws* pathways, although it is most pronounced within *wsp*. The *wsp* operon is composed of seven genes, all, in principle, capable of acquiring mutations leading to the constitutive activation of the WspR DGC (Figure 3). Goymer *et al.* (2006), for example, showed that it was possible, through *in vitro* manipulation, to generate constitutively active WspR alleles. It is easy to imagine mutations in WspA (the methyl-accepting chemotaxis protein) that might mimic the ligand-bound state, and Bantinaki *et al.* (2007) showed that overactivation of the WspC methyltransferase results in overactivation of WspR. However, despite the existence of seven genetic targets—and thus a large potential mutational target size—mutations causing WS were found in just *wspF* and *wspE*, with the majority (13/26) being in *wspF* (Figure 3). Assuming that the mutation rate is constant across the seven-gene operon, then the conclusion that each gene differs in its capacity to translate mutation into WS variation must be correct. As we have previously argued (Bantinaki *et al.* 2007), this can be explained (and was previously predicted) in terms of the function and regulatory connections among components of the Wsp pathway. WspF is a negative regulator, which, when inactivated by mutation (or otherwise has its activity reduced), results in activation of WspR (Bantinaki *et al.* 2007). This fact, combined with the knowledge that most mutations result in a loss of function, explains

why most *wsp*-generated WS arise from mutations in *wspF*: *wspF*, because of its function and role in the Wsp-signaling pathway, has a high propensity to translate DNA sequence change into WS phenotypic variation.

Arguments with regard to the capacity of *wspF* to translate mutation into WS variation can be equally applied to *awsX* and to the EAL domain of the protein encoded by *mwsR*, which negatively regulate the activity of the AwsR DGC and the DGC domain of MwsR, respectively. In fact, it is highly likely that negative regulation of DGC activity, a feature of all three pathways, explains the prevalence of WS-causing mutations in these pathways.

Mutational target size: Mutational target size is an important factor governing the evolution of phenotypic variation. A trait with a complex functional architecture is likely to be influenced by many loci and thus by definition affords a large mutational target (Houle 1998; Hansen 2006). While in general terms this makes sense, our work shows that the concept can be refined given a more detailed understanding of the specific connections between genotype and phenotype.

On the basis, first, of the results of the comprehensive suppressor studies (for example, of LSWS), we concluded that the WS phenotype is underpinned by a complex architecture of ~100 genes; however, further analysis has allowed a distinction to be drawn between direct structural determinants of the WS phenotype [*e.g.*, the cellulose-encoding *wss* operon and related adhesive factors (Spiers *et al.* 2002, 2003; Spiers and Rainey 2005)], determinants such as cell shape (Spiers *et al.* 2002) and various cell-wall-modifying genes (as shown in this study) that provide a necessary scaffold for correct production of adhesive factors, and genes that define regulatory modules within which mutations generating WS arise (Bantinaki *et al.* 2007). It is this last set of genes that define the mutational target. But further dissection shows that target size as a function of the number of genes within regulatory modules is overly simplistic. Of the 17 genes in the combined *wsp*, *aws*, and *mws* modules, 19 of the 26 causal WS mutations were found in the three negative regulators (*wspF*, *awsX*, and the EAL domain of *mwsR*). What matters, then, is not just the number of genes but their function and regulatory connections. This explains why certain functional components—in the context of their appropriate set of interactions—are better able to translate mutation into phenotypic variation than others.

But genetic target size as refined above is not sufficient to account for the full spectrum of mutations observed here. In particular, it does not explain the relatively large number of mutations in *wspF* relative to *awsX*: both *wspF* and *awsX* are negative regulators, *awsX* is approximately one-half the size of *wspF*, and mutation in *wspF* was the cause of WS on 13 occasions whereas mutation in *awsX* caused WS in just three instances—one-half the number expected. To some extent, this might reflect the small sample size, but it is likely that there are additional constraints operating on *awsX* (and *mwsR*). In terms of transcriptional organization, the negative regulatory components of both *aws* and *mws* precede the DNA sequence that encodes the DGC, and in both instances a functional DGC is critical for the expression of the WS phenotype. Mutations that disrupt the function of the negative regulator, but also impact on the downstream DGC function, will not generate WS, and indeed in the case of *aws*, appear to be lethal. No frameshift mutations were detected in either *awsX* or *mwsR*, which most likely reflects the fact that frameshift mutations in these regulators have deleterious pleiotropic effects. Frameshift mutations, however, are tolerated in *wspF*; the downstream DGC-encoding gene (*wspR*) is separated from *wspF* by 50 nucleotides and contains its own ribosome-binding site (Bantinaki *et al.* 2007). Of the 13 mutations in *wspF*, about one-half (six) are frameshift mutations. When the number of substitutions in *wspF* and *awsX* is compared—and corrected for gene size—the numbers are equivalent.

While it is understood that loss-of-function mutations are more common than gain-of-function mutations (Kimura 1968), it is surprising to find that ~25% of the causal WS mutations (6 of 26) are predicted to be of this kind. The three mutations within the WspE kinase, and the two within the AwsR DGC, cannot, given their phenotypic effects, be loss-of-function mutations. In both cases, these mutations must activate the WspE kinase and AwsR DGC, respectively. The mutation in the putative outer membrane component AwsO most likely mimics the active state of the transporter leading to release of repression of AwsR by AwsX. Strictly speaking, however, these activating mutations might be more appropriately considered a special case of the loss-of-function mutation because the activity of the regulator is no longer inducible. Either way, the prevalence of these mutations is worthy of note and raises questions concerning the relative importance of both kinds of mutations—and the opportunity for their occurrence—in phenotypic evolution.

Conclusion: The complexities of genetic architecture determine extant phenotypes, but such architectures also underpin the capacity of lineages to generate new genetic and phenotypic variation. Statistical population genetics has traditionally studied the evolution of phenotypes by averaging gene interactions (Fisher 1930) and in so doing has treated the complex network of connectivities that link genotype and phenotype as "noise." While the power of quantitative genetics is undisputed, the complexities arising from the genotype–phenotype map do have evolutionary implications that need to be understood (Schlichting and Murren 2004; Hansen 2006). Indeed, as we have shown here, parallel genetic evolution of the WS phenotype is driven by selection, but the mutations available to selection are a biased subset of all possible WS-generating mutations.

We suggest that this subset is composed of mutations in loci that, because of specific interactions and functionalities, have a heightened propensity to translate mutation into adaptive phenotypic variation. Unequivocal experimental demonstration of the existence of such constraints, combined with insight into their mechanistic bases and consequences, confirms that developmental biases do indeed, along with natural selection, affect the course of phenotypic evolution (Maynard Smith *et al.* 1985; Schluter 1996; Gould 2002; Brakefield 2006).

We thank Christian Kost and Bertus Beaumont and members of the Rainey lab for comments on the manuscript. Bertus Beaumont provided assistance during the early stages of detecting *aws* mutations and Annabel Hurman provided technical assistance. Financial support was provided by The Gatsby Charitable Foundation (United Kingdom), The Foundation for Research Science and Technology (New Zealand), and Massey University.

LITERATURE CITED

Allender, C. J., O. Seehausen, M. E. Knight, G. F. Turner and N. Maclean, 2003 Divergent selection during speciation of Lake Malawi cichlid fishes inferred from parallel radiations in nuptial coloration. Proc. Natl. Acad. Sci. USA **100:** 14074–14079.

Bailey, J., and C. Manoil, 2002 Genome-wide internal tagging of bacterial exported proteins. Nat. Biotechnol. **20:** 839–842.

Bantinaki, E., R. Kassen, C. G. Knight, Z. Robinson, A. J. Spiers *et al.*, 2007 Adaptive divergence in experimental populations of *Pseudomonas fluorescens*. III. Mutational origins of wrinkly spreader diversity. Genetics **176:** 441–453.

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

Boucher, J. C., J. Martinez-Salazar, M. J. Schurr, M. H. Mudd, H. Yu *et al.*, 1996 Two distinct loci affecting conversion to mucoidy in *Pseudomonas aeruginosa* in cystic fibrosis encode homologs of the serine protease HtrA. J. Bacteriol. **178:** 511–523.

Boughman, J. W., H. D. Rundle and D. Schluter, 2005 Parallel evolution of sexual isolation in sticklebacks. Evolution **59:** 361–373.

Brakefield, P. M., 2006 Evo-devo and constraints on selection. Trends Ecol. Evol. **21:** 362–368.

Choi, K.-H., J. B. Gaynor, K. G. White, C. Lopez, C. M. Bosio *et al.*, 2005 A Tn*7*-based broad-range bacterial cloning and expression system. Nat. Methods **2:** 443–448.

Choy, W. K., L. Zhou, C. K. C. Syn, L. H. Zhang and S. Swarup, 2004 MorA defines a new class of regulators affecting flagellar development and biofilm formation in diverse *Pseudomonas* species. J. Bacteriol. **186:** 7221–7228.

Christen, B., M. Christen, R. Paul, F. Schmid, M. Folcher *et al.*, 2006 Allosteric control of cyclic di-GMP signaling. J. Biol. Chem. **281:** 32015–32024.

Colosimo, P. F., C. L. Peichel, K. Nereng, B. K. Blackman, M. D. Shapiro *et al.*, 2004 The genetic architecture of parallel armor plate reduction in threespine sticklebacks. PLoS Biol. **2:** 635–641.

Drake, J. W., B. Charlesworth, D. Charlesworth and J. F. Crow, 1998 Rates of spontaneous mutation. Genetics **148:** 1667–1686.

ffrench-Constant, R. H., B. Pittendrigh, A. Vaughan and N. Anthony, 1998 Why are there so few resistance-associated mutations in insecticide target genes? Philos. Trans. R. Soc. Lond. B Biol. Sci. **353:** 1685–1693.

Figurski, D. H., and D. R. Helinski, 1979 Replication of an origin-containing derivative of plasmid RK2 dependent on a plasmid function provided in *trans*. Proc. Natl. Acad. Sci. USA **76:** 1648–1652.

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

Galperin, M. Y., T. A. Gaidenko, A. Y. Mulkidjanian, M. Nakano and C. W. Price, 2001 MHYT, a new integral membrane sensor domain. FEMS Microbiol. Lett. **205:** 17–23.

Gehrig, S. M., 2005 Adaptation of *Pseudomonas fluorescens* SBW25 to the air-liquid interface: a study in evolutionary genetics. Ph.D. Thesis, University of Oxford, Oxford.

Gerhart, J., and M. Kirschner, 1997 *Cells, Embryos, and Evolution*. Blackwell, Oxford.

Giddens, S. R., R. W. Jackson, C. D. Moon, M. A. Jacobs, X. X. Zhang *et al.*, 2007 Mutational activation of niche-specific genes provides insight into regulatory networks and bacterial function in a complex environment. Proc. Natl. Acad. Sci. USA **104:** 18247–18252.

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

Goymer, P., S. G. Kahn, J. G. Malone, S. M. Gehrig, A. J. Spiers *et al.*, 2006 Adaptive divergence in experimental populations of *Pseudomonas fluorescens*. II. Role of the GGDEF regulator WspR in evolution and development of the wrinkly spreader phenotype. Genetics **173:** 515–526.

Hansen, T. F., 2006 The evolution of genetic architecture. Annu. Rev. Ecol. Syst. **37:** 123–157.

Horton, R. M., H. D. Hunt, S. N. Ho, J. K. Pullen and L. R. Pease, 1989 Engineering hybrid genes without the use of restriction enzymes: gene-splicing by overlap extension. Gene **77:** 61–68.

Houle, D., 1998 How should we explain variation in the genetic variance of traits? Genetica **102–103:** 241–253.

Hudson, E. R., U. Bergthorsson, J. R. Roth and H. Ochman, 2002 Effect of chromosome location on bacterial mutation rates. Mol. Biol. Evol. **19:** 85–92.

Jenal, U., and J. Malone, 2006 Mechanisms of cyclic-di-GMP signaling in bacteria. Annu. Rev. Genet. **40:** 385–407.

Kimura, M., 1968 Evolutionary rate at the molecular level. Nature **217:** 624–626.

King, E. O., M. K. Ward and D. C. Raney, 1954 Two simple media for the demonstration of pyocyanin and fluorescin. J. Lab. Clin. Med. **44:** 301–307.

Knight, C. G., N. Zitzmann, S. Prabhakar, R. Antrobus, R. Dwek *et al.*, 2006 Unravelling adaptive evolution: how a single point mutation affects the protein co-regulation network. Nat. Genet. **38:** 1015–1022.

Kronforst, M. R., D. D. Kapan and L. E. Gilbert, 2006 Parallel genetic architecture of parallel adaptive radiations in mimetic Heliconius butterflies. Genetics **174:** 535–539.

Landry, C. R., B. Lemos, S. A. Rifkin, W. J. Dickinson and D. L. Hartl, 2007 Genetic properties influencing the evolvability of gene expression. Science **317:** 118–121.

Lenski, R. E., M. R. Rose, S. C. Simpson and S. C. Tadler, 1991 Long-term experimental evolution in *Escherichia coli*. I. Adaptation and divergence during 2,000 generations. Am. Nat. **138:** 1315–1341.

Malone, J. G., R. Williams, M. Christen, U. Jenal, A. J. Spiers *et al.*, 2007 The structure-function relationship of WspR, a response regulator with a GGDEF output domain. Microbiology **153:** 980–994.

Manoil, C., 2000 Tagging exported proteins using *Escherichia coli* alkaline phosphatase gene fusions. Methods Enzymol. **326:** 35–47.

Maynard Smith, J., R. Burian, S. Kauffman, P. Alberch, J. Campbell *et al.*, 1985 Developmental constraints and evolution. Q. Rev. Biol. **60:** 265–287.

McGregor, A. P., V. Orgogozo, I. Delon, J. Zanet, D. G. Srinivasan *et al.*, 2007 Morphological evolution through multiple cis-regulatory mutations at a single gene. Nature **448:** 587–590.

Ostrowski, E. A., R. J. Woods and R. E. Lenski, 2008 The genetic basis of parallel and divergent phenotypic responses in evolving populations of *Escherichia coli*. Proc. Biol. Sci. **275:** 277–284.

Pigeon, D., A. Chouinard and L. Bernatchez, 1997 Multiple modes of speciation involved in the parallel evolution of sympatric morphotypes of lake whitefish (*Coregonus clupeaformis*, Salmonidae). Evolution **51:** 196–205.

Rainey, P. B., 1999 Adaptation of *Pseudomonas fluorescens* to the plant rhizosphere. Environ. Microbiol. **1:** 243–257.

Rainey, P. B., and M. J. Bailey, 1996 Physical and genetic map of the *Pseudomonas fluorescens* SBW25 chromosome. Mol. Microbiol. **19:** 521–533.

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

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

Rice, S. H., 2008 Theoretical approaches to the evolution of development and genetic architecture. Ann. NY Acad. Sci. **1133:** 67–86.

Ross, P., H. Weinhouse, Y. Aloni, D. Michaeli, P. Weinbergerohana *et al.*, 1987 Regulation of cellulose synthesis in *Acetobacter xylinum* by cyclic diguanylic acid. Nature **325:** 279–281.

Rutherford, K., J. Parkhill, J. Crook, T. Horsnell, P. Rice *et al.*, 2000 Artemis: sequence visualisation and annotation. Bioinformatics **16:** 944–945.

Sambrook, J., E. F. Fritsch and T. Maniatis, 1989 *Molecular Cloning. A Laboratory Manual.* Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Schlichting, C. D., and C. J. Murren, 2004 Evolvability and the raw material for adaptation, pp. 18–29 in *Plant Adaptation: Molecular Genetics and Ecology*, edited by Q. C. B. Cronk, J. Whitton, R. H. Ree and I. E. P. Taylor. NRC Research Press, Ottawa.

Schluter, D., 1996 Adaptive radiation along genetic lines of least resistance. Evolution **50:** 1766–1774.

Schurr, M. J., H. Yu, J. M. Martinez-Salazar, J. C. Boucher and V. Deretic, 1996 Control of AlgU, a member of the sigma E-like family of stress sigma factors, by the negative regulators MucA and MucB and *Pseudomonas aeruginosa* conversion to mucoidy in cystic fibrosis. J. Bacteriol. **178:** 4997–5004.

Shindo, C., M. J. Aranzana, C. Lister, C. Baxter, C. Nicholls *et al.*, 2005 Role of FRIGIDA and FLOWERING LOCUS C in determining variation in flowering time of Arabidopsis. Plant Physiol. **138:** 1163–1173.

Silby, M. W., A. M. Cerdeno-Tarranga, G. Vernikos, S. R. Giddens, R. W. Jackson *et al.*, 2009 Genomic and genetic analyses of diversity and plant interactions of *Pseudomonas fluorescens*. Genome Biol. **10:** R51.

Spiers, A. J., and P. B. Rainey, 2005 The *Pseudomonas fluorescens* SBW25 wrinkly spreader biofilm requires attachment factor, cellulose fibre and LPS interactions to maintain strength and integrity. Microbiology **151:** 2829–2839.

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

Spiers, A. J., J. Bohannon, S. M. Gehrig and P. B. Rainey, 2003 Biofilm formation at the air-liquid interface by the *Pseudomonas fluorescens* SBW25 wrinkly spreader requires an acetylated form of cellulose. Mol. Microbiol. **50:** 15–27.

Stern, D. L., and V. Orgogozo, 2008 The loci of evolution: How predictable is genetic evolution? Evolution **62:** 2155–2177.

Stern, D. L., and V. Orgogozo, 2009 Is genetic evolution predictable? Science **323:** 746–751.

Sugawara, E., and H. Nikaido, 2004 OmpA/OprF: slow porins or channels produced by alternative folding of outer membrane proteins, pp. 119–138 in *Bacterial and Eukaryotic Porins: Structure, Function, Mechanism*, edited by R. Benz. Wiley-VCH, Weinheim, Germany.

Tamayo, R., A. D. Tischler and A. Camilli, 2005 The EAL domain protein VieA is a cyclic diguanylate phosphodiesterase. J. Biol. Chem. **280:** 33324–33330.

Wichman, H. A., M. R. Badgett, L. A. Scott, C. M. Boulianne and J. J. Bull, 1999 Different trajectories of parallel evolution during viral adaptation. Science **285:** 422–424.

Woods, R., D. Schneider, C. L. Winkworth, M. A. Riley and R. E. Lenski, 2006 Tests of parallel molecular evolution in a long-term experiment with *Escherichia coli*. Proc. Natl. Acad. Sci. USA **103:** 9107–9112.

Yu, H., M. J. Schurr and V. Deretic, 1995 Functional equivalence of Escherichia coli sigma E and *Pseudomonas aeruginosa* AlgU: E. coli *rpoE* restores mucoidy and reduces sensitivity to reactive oxygen intermediates in *algU* mutants of *P. aeruginosa*. J. Bacteriol. **177:** 3259–3268.

Zhang, J. Z., 2006 Parallel adaptive origins of digestive RNases in Asian and African leaf monkeys. Nat. Genet. **38:** 819–823.

Zhong, S., A. Khodursky, D. E. Dykhuizen and A. M. Dean, 2004 Evolutionary genomics of ecological specialization. Proc. Natl. Acad. Sci. USA **101:** 11719–11724.

Communicating editor: J. Lawrence

GENETICS

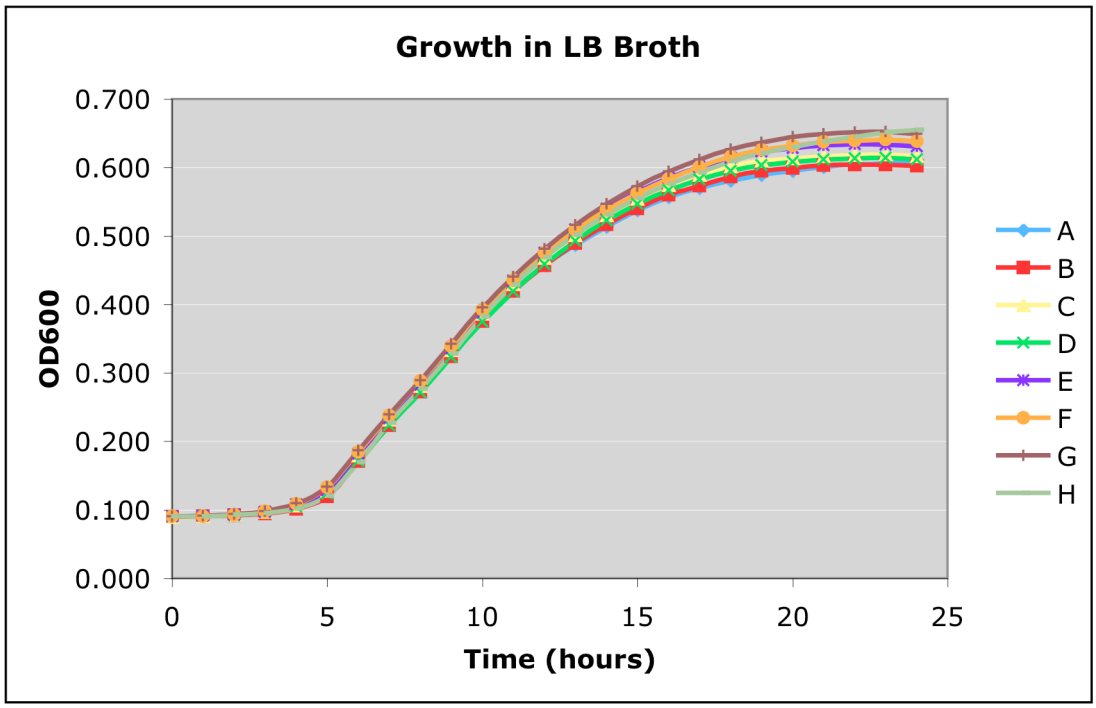
Supporting Information
http://www.genetics.org/cgi/content/full/genetics.109.107110/DC1

Adaptive Divergence in Experimental Populations of *Pseudomonas fluorescens*. IV. Genetic Constraints Guide Evolutionary Trajectories in a Parallel Adaptive Radiation

Michael J. McDonald, Stefanie M. Gehrig, Peter L. Meintjes, Xue-Xian Zhang and Paul B. Rainey

DOI: 10.1534/genetics.109.107110

A



B

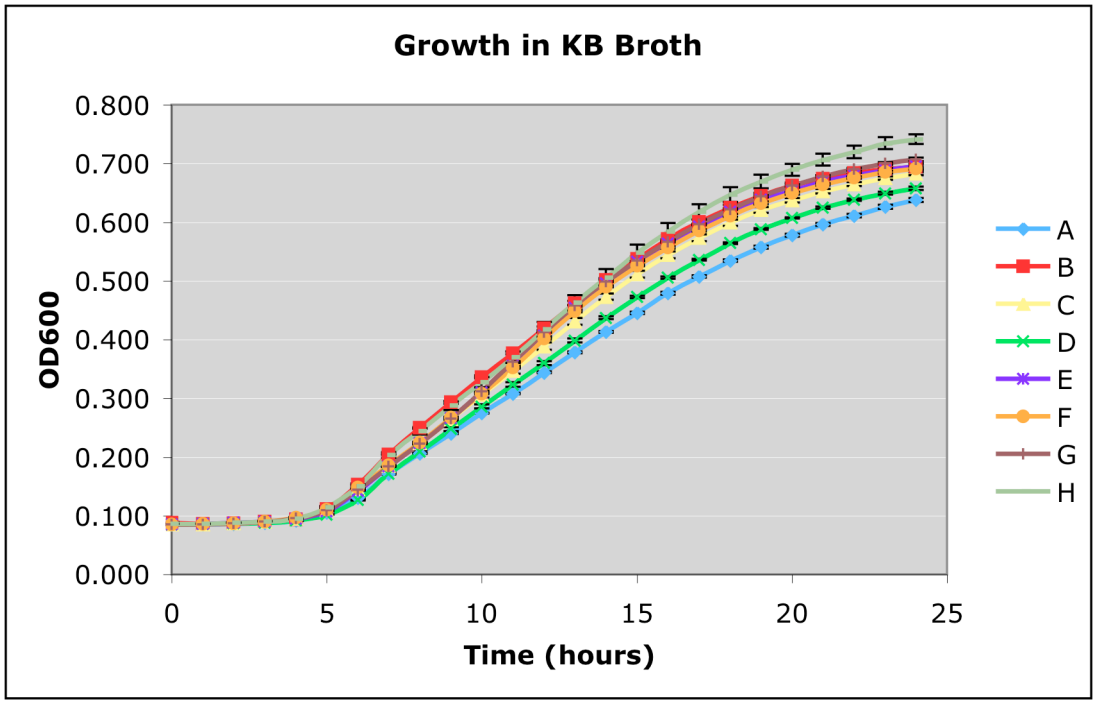


FIGURE S1.—Growth dynamics of *wsp*, *aws*, *mws* deletion mutants in LB (A) and KB (B). A: SM Δ*wsp* (PBR693), B: SM Δ*aws* (PBR681), C: SM Δ*mws* (PBR711), D: SM Δ*wsp* Δ*aws* (PBR712), E: SM Δ*wsp* Δ*mws* (PBR713), F: SM Δ*mws* Δ*aws* (PBR718), G: SM Δ*wsp* Δ*aws* Δ*mws* (PBR716), H: ancestral SBW25. Data are means and standard errors of six replicates (error bars are contained within the symbols). The relevant environment is KB – deletion of DGC-encoding loci did not negatively affect growth in this medium.

FILES S1-S4

Files S1-S4 are available for download as Artemis files at http://www.genetics.org/cgi/content/full/genetics.109.107110/DC1.