

Genetics of germination-arrest factor (GAF) production by *Pseudomonas fluorescens* WH6: identification of a gene cluster essential for GAF biosynthesis

Anne Halgren,¹ Maciej Maselko,¹ Mark Azevedo,¹ Dallice Mills,² Donald Armstrong² and Gary Banowetz¹

¹USDA-ARS National Forage Seed Production Research Center, Corvallis, OR 97331, USA

²Department of Botany and Plant Pathology, Oregon State University, Corvallis, OR 97331, USA

Correspondence

Gary Banowetz
banowetzg@onid.orst.edu

The genetic basis of the biosynthesis of the germination-arrest factor (GAF) produced by *Pseudomonas fluorescens* WH6, and previously identified as 4-formylaminoxyvinylglycine, has been investigated here. In addition to inhibiting the germination of a wide range of grassy weeds, GAF exhibits a selective antimicrobial activity against the bacterial plant pathogen *Erwinia amylovora*. We utilized the *in vitro* response of *E. amylovora* to GAF as a rapid screen for loss-of-function GAF phenotypes generated by transposon mutagenesis. A Tn5 mutant library consisting of 6364 WH6 transformants was screened in this *Erwinia* assay, resulting in the identification of 18 non-redundant transposon insertion sites that led to loss of GAF production in WH6, as confirmed by TLC analysis. These insertions mapped to five different genes and four intergenic regions. Three of these genes, including two putative regulatory genes (*gntR* and *iopB* homologues), were clustered in a 13 kb chromosomal region containing 13 putative ORFs. A GAF mutation identified previously as affecting an aminotransferase also maps to this region. We suggest that three of the genes in this region (a carbamoyltransferase, an aminotransferase and a formyltransferase) encode the enzymes necessary to synthesize dihydroGAF, the putative immediate precursor of GAF in a proposed GAF biosynthetic pathway. RT-qPCR analyses demonstrated that mutations in the *gntR* and *iopB* regulatory genes, as well as in a *prtR* homologue identified earlier as controlling GAF formation, suppressed transcription of at least two of the putative GAF biosynthetic genes (encoding the aminotransferase and formyltransferase) located in this 13 kb region.

Received 13 July 2012

Revised 27 September 2012

Accepted 26 October 2012

INTRODUCTION

We have previously shown that a naturally occurring herbicide is produced by particular isolates of rhizosphere bacteria obtained from soils of the Willamette Valley, Oregon (Banowetz *et al.*, 2008, 2009; Armstrong *et al.*, 2009; McPhail *et al.*, 2010). This compound, which we have termed a germination-arrest factor (GAF), specifically and irreversibly blocks germination of the seeds of a large number of grassy weed species without significantly affecting the growth of established grass seedlings or mature plants. The biological and herbicidal properties of GAF have been described, and we have developed a quantitative bioassay for

the herbicide based on its ability to arrest germination of the seeds of annual bluegrass (*Poa annua* L.), one of our target weed species (Banowetz *et al.*, 2008). In addition to its herbicidal properties, GAF was also found to exhibit a selective antimicrobial activity against *Erwinia amylovora*, the bacterial plant pathogen that causes fireblight in orchard crops (Halgren *et al.*, 2011).

TLC of extracts of culture filtrates from the GAF-producing bacterial isolate *Pseudomonas fluorescens* WH6 revealed that GAF activity is associated with a particular ninhydrin-reactive compound (Armstrong *et al.*, 2009). We were subsequently able to purify and identify this compound as 2-amino-4-formylaminoxy-3-butenoic acid (4-formylaminoxyvinylglycine) (Fig. 1), a previously undescribed member of a small group of naturally occurring compounds known as oxyvinylglycines (McPhail *et al.*, 2010). Oxyvinylglycines are known to block reactions catalysed by enzymes dependent upon pyridoxal phosphate as a co-factor (Berkowitz *et al.*,

Abbreviations: Cq, threshold cycle; GAF, germination-arrest factor; HIT protein, HisTriad protein; iPCR, inverse PCR; RT-qPCR, real-time quantitative PCR.

Four supplementary figures and a supplementary table are available with the online version of this paper.

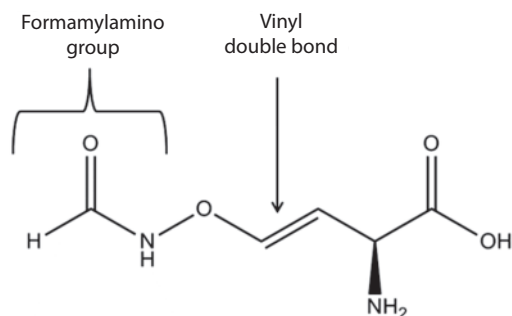


Fig. 1. Structure of GAF, 4-formylaminoxyvinylglycine.

2006). These enzymes include aminotransferases that perform vital functions in nitrogen metabolism, as well as the enzyme aminocyclopropanecarboxylic acid synthase (ACC synthase), which catalyses a critical step in the biosynthesis of the plant hormone ethylene. However, the ability of oxvinynglycines and related compounds to selectively block the germination of grassy weeds and the growth of *E. amylovora* had not previously been recognized.

Transposon mutagenesis of WH6, in an earlier, non-saturating screen based on the *Poa* germination bioassay, produced two mutants of WH6 that had lost the ability to produce GAF (mutants WH6-2 and WH6-3) (Armstrong *et al.*, 2009). The recent sequencing of the WH6 genome (Kimbrel *et al.*, 2010) enabled the sites of the mutations in these strains to be mapped. In the mutant strain WH6-2, the Tn5 insertion disrupted a gene that encodes a putative aminotransferase, suggesting that this gene plays a direct role in GAF biosynthesis. The Tn5 insertion in mutant strain WH6-3 disrupted a gene (*prtR*) that is similar to a putative activator of a sigma factor in *Pseudomonas fluorescens* LS107d2 (Burger *et al.*, 2000), which is consistent with a role for this gene in the regulation of GAF biosynthesis.

The discovery that GAF has selective antimicrobial activity against *E. amylovora* (Halgren *et al.*, 2011) indicated that the response of *E. amylovora* could be used as the basis for a rapid and potentially saturating screen for genes affecting GAF production. The results of a GAF mutational search using an *E. amylovora* screen are reported here. With this approach, two additional regulatory genes were identified that function, together with the previously identified *prtR* gene, to regulate the GAF biosynthetic pathway. Moreover, a number of genes that encode enzymes that appear to be involved, directly or indirectly, in GAF biosynthesis have been identified. Real-time quantitative PCR (RT-qPCR) analysis demonstrated that two genes that encode enzymes likely to be directly involved in the GAF biosynthetic pathway are controlled by the three putative GAF regulatory genes.

METHODS

Bacterial strains and growth media. *E. amylovora* 153 was obtained from Dr Joyce Loper [United States Department of

Agriculture-Agricultural Research Service (USDA-ARS) Horticultural Crops Research Laboratory, Corvallis, OR]. The origin and characterization of *P. fluorescens* strain WH6 have been described previously (Banowetz *et al.*, 2008; Elliott *et al.*, 1998). All strains used in this study were maintained at -80°C in Luria-Bertani liquid medium (LB medium) (Sambrook & Russell, 2001) with a final concentration of 15% (v/v) glycerol.

The transposon vector pUTmini-Tn5gfp in its host *Escherichia coli* S17-1 λ pir (Matthysse *et al.*, 1996) was obtained as a gift from Ann Matthysse (Department of Biology, University of North Carolina). The mini-Tn5 derivative contains the inner and outer transposase recognition sequences flanking a promoterless GFP gene and a tetracycline resistance cassette.

Tn5 mutagenesis. Both *P. fluorescens* WH6 and *Escherichia coli* S17-1 λ pir containing the transposon delivery vector pUTmini-Tn5gfp were grown in 4 ml LB medium overnight with shaking (200 r.p.m.) at 28 and 37 $^{\circ}\text{C}$, respectively. The *Escherichia coli* medium was supplemented with 10 μg tetracycline ml^{-1} in order to maintain the vector. A 1.5 ml aliquot of each culture was centrifuged at 3330 g for 2 min, and the cell pellet was then washed and resuspended in 500 μl PBS (Sambrook & Russell, 2001). Both cultures were combined into one tube, spread onto an LB plate containing no antibiotic, and left at 28 $^{\circ}\text{C}$ overnight. The following day, the bacterial lawn was collected from the plate using 3 ml PBS, washed, and streaked onto '925 minimal medium' plates (Halgren *et al.*, 2011) containing 10 μg tetracycline ml^{-1} , in order to select against *Escherichia coli* and nontransformants.

Selection of GAF mutants via the *E. amylovora* screen. *E. amylovora* cultures were grown, diluted and plated as previously described (Halgren *et al.*, 2011). Eighteen colonies of transformed *P. fluorescens* WH6 were spotted with a sterile toothpick onto each plate freshly spread with an *E. amylovora* lawn. Plates were incubated at 28 $^{\circ}\text{C}$ and examined for antibiosis 48 h after sample application. Colonies lacking a zone of inhibition were selected and retested against *E. amylovora*, streaked onto plates of LB medium containing 10 μg tetracycline ml^{-1} and cultured for cryopreservation as described above.

Culture filtrate production. Each transformant that had lost the ability to inhibit growth of *E. amylovora* was inoculated into the modified *Pseudomonas* minimal salts medium (PMS medium), cultured and harvested as described previously (Banowetz *et al.*, 2008). To prepare culture filtrates, the *Pseudomonas* culture fluid recovered from 7-day cultures was centrifuged (3000 g, 15 min), and the supernatant was passed through a bacteriological filter (Millipore GP Express Steritop, 0.22 μm pore size). The resulting sterile culture filtrate was stored at 4 $^{\circ}\text{C}$.

TLC analysis and bioassay of GAF activity. TLC analyses of culture filtrates from mutant bacterial cultures were performed on 90% ethanol extracts of the solids from dried culture filtrates as described previously (Armstrong *et al.*, 2009; Halgren *et al.*, 2011). Silica GHL and microcrystalline cellulose TLC plates were purchased from Analtech.

Bioassays for GAF activity in bacterial culture filtrates were performed with annual bluegrass seeds (ABG, *Poa annua* L.) using the standard *Poa* germination bioassay protocol and scoring system described by Banowetz *et al.* (2008). In this scoring system, a score of 4 represents normal germination and plumule development, and a score of 1 represents complete germination arrest immediately after emergence of the radicle and coleoptile. Tests of culture filtrates for antimicrobial activity against *E. amylovora* were performed as described by Halgren *et al.* (2011).

DNA extraction, digestion by restriction enzymes, and self-ligation. DNA was isolated from mutant strains of *P. fluorescens*

WH6 using the ZR Fungal/Bacterial DNA kit (Zymo Research). Purity and concentration were determined using a NanoDrop ND-1000 spectrophotometer (Thermo Scientific).

DNA from putative GAF mutant colonies was digested with *EheI* (*NarI*), which cuts frequently within the pUT-miniTn5gfp delivery vector and within the *P. fluorescens* WH6 genome. *BseSI* was used as a secondary restriction enzyme when *EheI* digests yielded only extremely short inverse PCR (iPCR) products. After the enzyme was inactivated, approximately 12 ng linear digested DNA was circularized by self-ligation using T4 ligase (Fermentas) in a 50 µl reaction mixture at 25 °C for 1 h. The digestions and self-ligation were performed according to the manufacturer's instructions. The self-ligated DNA was purified using the QIAquick PCR Purification kit (Qiagen). The clean ligation products were then used as template for amplification of transposon-flanking DNA sequences in an iPCR reaction.

iPCR and sequencing. All PCR amplifications were conducted in a Robocycler (Stratagene) thermocycler with Platinum *Taq* polymerase using the manufacturer's instructions (Invitrogen). iPCR was performed using outward-facing, transposon-specific primers, and iPCR products were sequenced with a nested transposon-specific primer. Primers used to amplify the products of the *NarI* digest were those developed by Lewenza *et al.* (2005), whereas primers used for amplifying *BseSI* digest products were as noted in Table S1 available with the online version of this paper. Thermocycling consisted of an initial denaturation for 5 min at 94 °C, followed by 40 cycles (30 s, 94 °C; 30 s, 56 °C; 2 min, 72 °C), with a final 10 min extension step at 72 °C. Prior to cycling, samples were heated at 94 °C for 5 min and the extension step was increased to 10 min, 72 °C, as part of the terminal cycle. The samples were subjected to electrophoresis through agarose gels containing 10 µg ethidium bromide ml⁻¹, and amplicons were visualized by exposure to UV light. All samples yielding a single major band were selected for clean-up using either the QIAquick PCR Purification kit or QIAquick Gel Extraction kit (Qiagen). Sequencing was carried out at the Central Services Lab of the Center for Genome Research and Biocomputing (CGRB; Oregon State University, Corvallis, OR). PCR products were identified using BLAST (Altschul *et al.*, 1997), and the region flanking the transposon insertion site was aligned with the annotated genome sequence of *P. fluorescens* WH6 (Kimbrel *et al.*, 2010).

RNA extraction, reverse transcription (RT) and RT-qPCR. Mutant strains WH6-3, WH6-4 and WH6-7 (each containing a Tn5 insertion in a different regulatory gene), and wild-type WH6, were grown as overnight cultures in 5 ml PMS medium at 28 °C with shaking. The following morning, cultures were diluted 1:4 into fresh medium and divided into three replicate cultures with an OD₆₀₀ of approximately 0.13 each. The bacterial RNA was harvested after 4.5 h, when cells had reached mid-exponential phase (OD₆₀₀ ~0.750). RNA was extracted using RNeasy Protect and the RNeasy kit (Qiagen). Samples were treated with DNase using the Turbo DNA-free kit (Ambion), with 1 h incubation and using 2 µl Turbo DNase. RNA concentration was measured using a NanoDrop ND-1000 spectrophotometer. The RNA preparation was determined to be of high quality, based on an analysis of RNA integrity using an Agilent Bioanalyser 2100 (Agilent) to evaluate the 16S and 23S rRNA peaks and their ratio. cDNA was synthesized using Invitrogen's SuperScript First-Strand Synthesis System for RT-PCR kit and random hexamer primers according to the manufacturer's instructions.

The expression of two putative GAF biosynthetic genes, PFWH6_5256 and PFWH6_5257 (putative aminotransferase and formyltransferase genes), was examined and compared by RT-qPCR in each of the three regulatory mutants and in wild-type WH6. For RT-qPCR analysis, the cDNA from each of three replicate cultures of each bacterial strain was analysed using two technical replicates for each cDNA preparation. In addition, each 96-well plate contained no-template controls to screen for possible contamination and dimer formation, as well as

no-reverse-transcriptase controls to screen for DNA contamination. Real-time PCR was performed using a Multicolor Real-Time PCR Detection system (model iQ5; Bio-Rad Laboratories) in 20 µl reaction mixtures with iQ SYBR green Supermix (Bio-Rad Laboratories). Cycling conditions were 95 °C for 3 min, and 40 cycles of 95 °C for 10 s and 57 °C for 30 s. Fluorescence readings were recorded after each cycle. A final melting analysis was obtained by slow heating, with 10 s increments of 0.5 °C from 57 to 95 °C, and fluorescence collection at intervals of 0.5 °C. The threshold cycle (C_q) value of each sample was determined during the exponential phase of amplification. Each dissociation curve trace was checked to ensure that no non-specific products were present in the amplification.

Primer design, optimization and validation. Primers for both RT-qPCR genes of interest and candidate reference genes were designed using Primer3Plus software (Untergasser *et al.*, 2007) (Table S1) from Sigma-Aldrich. Four candidate reference genes were selected for evaluation: *proC*, *rpoD*, *rpsL* and *tufB*. *proC* (encoding pyrroline-5-carboxylate reductase) and *rpoD* (encoding sigma factor RpoD) have been described as stable internal control genes in *Pseudomonas aeruginosa* (Savli *et al.*, 2003). The *rpsL* gene, which encodes the 30S ribosomal protein S12, has also been used as a reference gene in *P. aeruginosa* (Gooderham *et al.*, 2009). *tufB*, which encodes the elongation factor Tu (EF-Tu), has been used in our laboratory as a reference gene for *P. fluorescens* (J. A. Kimbrel, personal communication).

Prior to validating reference genes and performing expression studies, each primer set (designed for either reference genes or genes of interest) was tested for efficiency against a standard curve, consisting of five fivefold serial dilutions, covering a range from 15.4 ng to 4.9 pg using WH6 cDNA. For each primer pair, amplification efficiency was calculated as $E = -1 + 10^{(-1/\text{slope})}$. At least one sample from each primer pair was analysed by agarose gel electrophoresis to verify that the product was a single band of the correct size. Additionally, one sample from each primer pair was sequenced to verify its identity.

To validate the candidate reference genes for use under the conditions of the current study, each candidate primer set was tested against cDNA representing WH6 and each of the mutants used in the analysis. All candidate reference genes were quantified on the same batch of cDNA as used in the experimental analysis. For these validation experiments, triplicate sets of each cDNA were amplified using the same conditions described above. The C_q was averaged for each triplicate and transformed to quantities relative to the sample with the highest expression. These values were imported into geNorm v3.5 software (Vandesompele *et al.*, 2002) for analysis of gene expression stability and subsequent confirmation of optimal reference genes. geNorm identified the genes *tufB* and *rpsL* as the most stably expressed reference genes, and recommended normalization of data to a combination of these two genes as a valid endogenous control. These two reference genes had M values of less than 0.3 (Fig. S1) and gave a V value of 0.146 (Fig. S2).

Data analysis. All post-run relative expression analyses were performed using REST 2009 (Qiagen) (Pfaffl *et al.*, 2002), which included an efficiency correction for each primer set. For data analysis, the C_q values of the genes were converted to relative quantities and normalized using the geometric mean of the two reference genes.

RESULTS

Mutant yield using the *E. amylovora* screen

A total of 6364 *P. fluorescens* WH6 transformants was obtained by Tn5 transposon mutagenesis. From these transformants, 23 putative GAF mutants were identified

using the *E. amylovora* screen, and 18 of these mutations occurred at non-redundant insertion sites (Table 1). Colonies of two of these mutants, WH6-10 and WH6-12, exhibited slight antimicrobial activity in the *E. amylovora* screen, but culture filtrates from all mutants were consistently antibiosis-deficient against *E. amylovora* (Table 2). Loss of GAF production in the mutant culture filtrates was confirmed in all 23 mutants by tests in the *Poa* germination bioassay for GAF activity (Table 2) and by TLC analysis (Figs S3 and S4). iPCR analysis revealed that these 23 mutations involved five unique ORFs and four intergenic regions. The locations of these mutations and the two separate ORF mutations obtained in our previous screening efforts were determined from the draft of the complete WH6 genome obtained by Kimbrel *et al.* (2010) and are illustrated in Fig. 2. The annotations of these mutations, as derived from both this source and from independent iPCR analyses, are summarized in Table 1.

Identification of regulatory genes controlling GAF biosynthesis

Our previous mutant search using the *Poa* germination bioassay as a primary screen resulted in the identification of one regulatory gene (*priR*, PFWH6_3687) controlling GAF

biosynthesis. In the current screen, two genes with homology to known regulatory elements were among the five unique ORFs identified as essential for GAF biosynthesis. One of these genes, PFWH6_5251 in mutant WH6-7, was disrupted by Tn5 insertions at five independent loci (Table 1, Fig. 3). This gene appears to be a homologue of a gene in *Pseudomonas syringae* pv. *maculicola* strain ES4326 that is annotated as producing a hypothetical protein (79 % identity). However, the next most closely related gene (65 % identity) is a gene designated *iopB*, which regulates the production of phenazine 1-carboxamide in *P. chlororaphis* strain PCL1391 (Van Rij, 2006). In strain PCL1391, this gene occurs downstream of another regulatory gene designated *iopA*. The putative *iopB* gene in WH6 also occurs downstream of an *iopA* homologue (PFWH6_5249). In both PCL1391 and WH6, the *iopA* and *iopB* genes are separated by a small putative ORF (PFWH6_5250 in WH6).

The second putative regulatory element identified in the *Erwinia* screen occurs immediately upstream of the *iopA* homologue in WH6, but in the opposite orientation. This gene, PFWH6_5248 in mutant WH6-4, was disrupted by Tn5 insertion at four independent loci (Table 1, Fig. 3). This gene encodes a putative GntR transcriptional regulator. GntR transcriptional regulators constitute a large family of regulatory proteins that contain both DNA-binding

Table 1. Mutations causing loss of GAF production in *P. fluorescens* WH6

Mutation designation	ORF site of transposon insertion*	Putative ORF product	Accession no.
WH6-2::Tn5†	PFWH6_5256	Aminotransferase	ZP_07777818
WH6-3::Tn5†	PFWH6_3687	PrtR regulator	ZP_07776266
WH6-4A::Tn5			
WH6-4N::Tn5	PFWH6_5248	GntR regulator	ZP_07777810
WH6-4AA::Tn5			
WH6-4BB::Tn5			
WH6-6C::Tn5	PFWH6_5719	HemY protein	ZP_07778272
WH6-7D::Tn5			
WH6-7R::Tn5			
WH6-7U::Tn5	PFWH6_5251	IopB regulator	ZP_07777813
WH6-7DD::Tn5			
WH6-7CC::Tn5			
WH6-8F::Tn5	5' of PFWH6_4690	HIT protein	ZP_07777256
WH6-9H::Tn5	5' of PFWH6_0984	Hypothetical protein	ZP_07773606
WH6-10K::Tn5	PFWH6_1695	dGTPase	ZP_07774306
WH6-11L::Tn5‡			
WH6-11M::Tn5‡			
WH6-11O::Tn5‡			
WH6-11P::Tn5§	PFWH6_5254	Carbamoyltransferase	ZP_07777816
WH6-11T::Tn5§			
WH6-11W::Tn5			
WH6-11X::Tn5			
WH6-12Q::Tn5	5' of PFWH6_5556	ThiG protein	ZP_07778113
WH6-15V::Tn5¶	5' of PFWH6_5255	Unknown	ZP_07777817
WH6-15Y::Tn5¶			

*ORF designations are based on the WH6 draft genome published by Kimbrel *et al.* (2010).

†Mutations identified in earlier screen using the *Poa* bioassay (Armstrong *et al.*, 2009).

‡,§,||,¶ Redundant insertion sites.

Table 2. Biological activity of culture filtrates from mutant strains of *P. fluorescens* WH6

Source of culture filtrate	<i>Poa</i> germination assay (score \pm SEM)*	<i>E. amylovora</i> assay area of zone of inhibition (cm ² \pm SEM)
WH6 (wild-type)	1.0 \pm 0.0	15.1 \pm 0.2
Water control	4.0 \pm 0.0	0.0 \pm 0.0
Culture medium	3.8 \pm 0.2	0.0 \pm 0.0
WH6-2	3.8 \pm 0.1	0.0 \pm 0.0
WH6-3	3.7 \pm 0.1	0.0 \pm 0.0
WH6-4A	3.0 \pm 0.1	0.0 \pm 0.0
WH6-4N	3.6 \pm 0.3	0.0 \pm 0.0
WH6-4AA	3.6 \pm 0.3	0.0 \pm 0.0
WH6-4BB	4.0 \pm 0.0	0.0 \pm 0.0
WH6-6C	3.1 \pm 0.1	0.0 \pm 0.0
WH6-7D	3.2 \pm 0.2	0.0 \pm 0.0
WH6-7R	3.5 \pm 0.1	0.0 \pm 0.0
WH6-7U	4.0 \pm 0.0	0.0 \pm 0.0
WH6-7DD	3.9 \pm 0.1	0.0 \pm 0.0
WH6-7CC	3.9 \pm 0.0	0.0 \pm 0.0
WH6-8F	3.3 \pm 0.2	0.0 \pm 0.0
WH6-9H	3.3 \pm 0.2	0.0 \pm 0.0
WH6-10K	3.2 \pm 0.2	0.0 \pm 0.0
WH6-11L	3.2 \pm 0.2	0.0 \pm 0.0
WH6-11M	3.4 \pm 0.2	0.0 \pm 0.0
WH6-11O	3.6 \pm 0.2	0.0 \pm 0.0
WH6-11P	3.2 \pm 0.2	0.0 \pm 0.0
WH6-11T	3.8 \pm 0.2	0.0 \pm 0.0
WH6-11W	4.0 \pm 0.0	0.0 \pm 0.0
WH6-11X	4.0 \pm 0.0	0.0 \pm 0.0
WH6-12Q	3.1 \pm 0.1	0.0 \pm 0.0
WH6-15V	4.0 \pm 0.0	0.0 \pm 0.0
WH6-15Y	4.0 \pm 0.0	0.0 \pm 0.0

*The scoring system is described in detail by Banowitz *et al.* (2008). A score of 4 indicates normal development, equivalent to the water control, at the end of the 7-day assay period. A score of 1 indicates germination arrest immediately after emergence of the radicle and coleoptile.

helix–turn–helix domains and aminotransferase domains (Haydon & Guest, 1991; Hoskisson & Rigali, 2009). Regulators of this family possess a conserved N-terminal domain that is involved in the DNA binding. Based on homology, the specific GntR regulator disrupted in our screen belongs to the MocR-like subfamily, which contains an aminotransferase domain that requires pyridoxal 5'-phosphate as a co-factor (Sung *et al.*, 1991; Rigali *et al.*, 2002). The closest homologue is a gene from *Marinobacter* sp. ELB1, with which it shares 90% identity. The *gntR* homologue in WH6 and the *iopA* and *iopB* homologues are on opposite strands of the chromosome and separated by an intergenic region with two predicted promoters that also occur on opposite strands of the chromosome (Fig. 3).

Identification of a gene cluster affecting GAF biosynthesis

The regulatory genes *iopB* and *gntR* reside within a 13 kb region of the *P. fluorescens* WH6 genome that includes two other genes (PFWH6_5254 and PFWH6_5256) and an

intergenic region whose disruption by Tn5 insertion resulted in loss of GAF production (Table 1, Fig. 3). The WH6-2::Tn5 mutation identified in our earlier screen using the *Poa* bioassay system also occurs within this 13 kb region. The latter mutation disrupted a putative aminotransferase gene (PFWH6_5256) (Armstrong *et al.*, 2009; Kimbrel *et al.*, 2010). Immediately downstream of this gene is a gene (PFWH6_5527) encoding a putative formyltransferase. The proximity of these two genes suggests that they both may play a role in the synthesis of the formylamino side-chain of GAF (see the structure illustrated in Fig. 1).

In our current screen, two additional types of mutations leading to loss of GAF production were observed in this 13 kb region. Multiple mutations were observed in PFWH6_5254 (mutations labelled WH6-11::Tn5). This gene encodes a putative carbamoyltransferase, which has a closely related homologue (85% identity) in *Pseudomonas chlororaphis* (Van Rij, 2006). Tn5 insertions resulting in loss of GAF production occurred at three independent sites

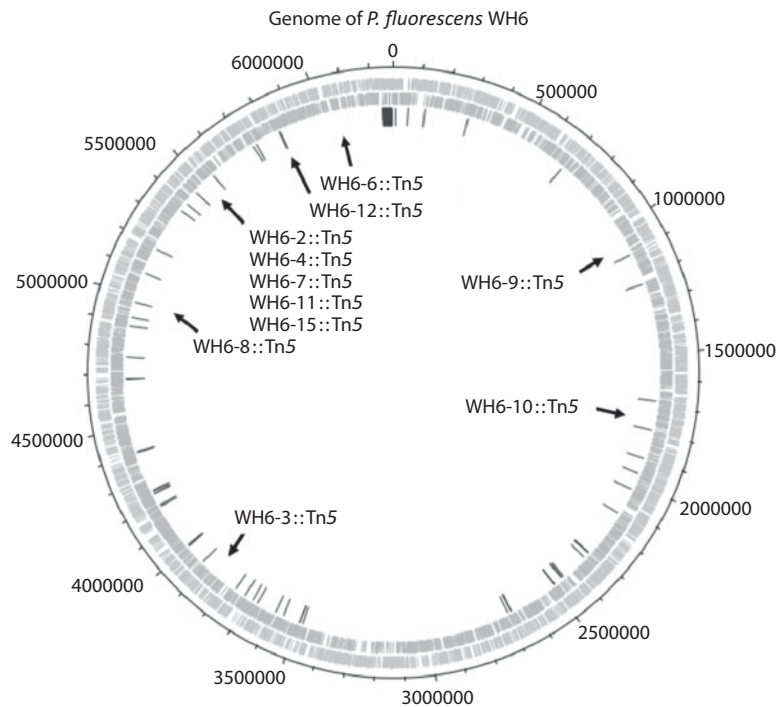


Fig. 2. Genome of *P. fluorescens* WH6 showing the positions of Tn5 insertions that result in loss of GAF production.

in this gene, and insertions at each site were observed more than once. An additional Tn5 insertion (WH6-15::Tn5) that disrupted GAF biosynthesis was within an intergenic region immediately downstream of the putative carbamoyltransferase gene and immediately upstream of a small ORF (PFWH6_5255) of unknown function and significance that precedes the putative aminotransferase gene (PFWH6_5256). It seems likely that mutation in this intergenic region is affecting GAF biosynthesis by disrupting transcription of downstream genes such as the aminotransferase gene.

Characterization of other mutations affecting GAF biosynthesis

The *E. amylovora* screen identified five GAF mutations that are located outside the 13 kb gene cluster described above. (Table 1, Fig. 2). The WH6-3::Tn5 mutation identified in our earlier screen as occurring in a homologue of *prtR*

(PFWH6_3687) (Armstrong *et al.*, 2009; Kimbrel *et al.*, 2010) also lies outside this gene cluster. The genomic context of all of these mutations is illustrated in Fig. 4.

Tn5 insertions affecting GAF production occurred in two ORFs external to the 13 kb gene cluster. WH6-6::Tn5 (located in PFWH6_5719) disrupted a gene encoding a HemY-like protein presumed to function in haem biosynthesis. Immediately downstream of this gene are two ORFs encoding homologues of DsbH, a disulfide-bond reductase, and AlgQ, a global transcriptional regulator. WH6-10::Tn5 (located in PFWH6_1695) disrupted a gene encoding a putative deoxyguanosinetriphosphate triphosphohydrolase-like protein. Downstream of this gene are three additional ORFs, encoding a putative response regulator, a LuxR-family two-component transcriptional regulator and a histidine kinase.

Three GAF mutations external to the 13 kb gene cluster occurred in intergenic regions and presumably affect

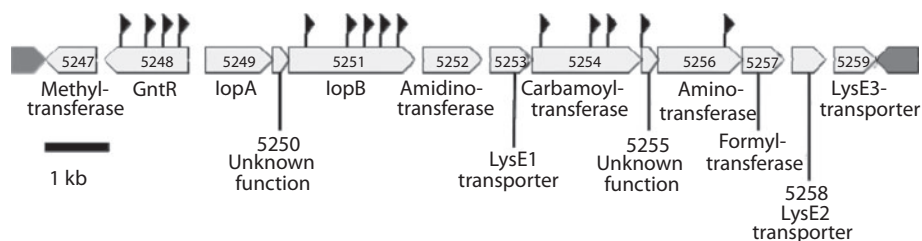


Fig. 3. Structure of the 13 kb gene cluster containing multiple genes affecting GAF biosynthesis. ORF numbers, sites of Tn5 insertions (flags) and putative gene products (based on sequence homology) are indicated.

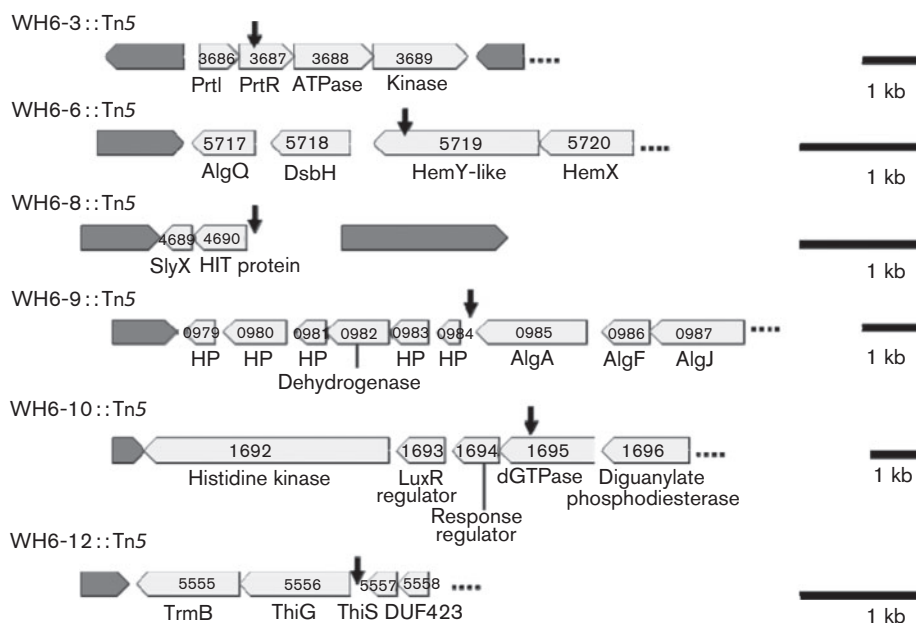


Fig. 4. Position and genetic context of isolated mutations affecting GAF biosynthesis. The positions of Tn5 insertions are indicated by the vertical arrows. Putative gene products, based on sequence homology, are indicated. HP is used to designate hypothetical proteins of unknown function.

expression of downstream genes. These genes encode both putative regulatory and enzymic products. The WH6-8::Tn5 mutation mapped 658 bp upstream of a gene (*oprD*) encoding a protein of unknown function. However, this insertion may have affected a gene on the antisense strand that occurs just 85 bp upstream of the site of insertion. This gene (PFWH6_4690) encodes a putative HisTriad (HIT) protein that likely functions as a nucleotide hydrolase and transferase acting on the alpha-phosphate of ribonucleotides. Immediately 3' of the HIT protein-encoding gene is an ORF encoding a product, SlyX, of unknown function. The WH6-9::Tn5 insertion occurred 230 bp upstream of the start site of another gene of unknown function (PFWH6_0984). Downstream of this gene, and in the same orientation, are two more genes of unknown function and a gene encoding a putative dehydrogenase. WH6-12::Tn5 mapped to an intergenic site 253 bp upstream of a gene with homology to *thiG* (PFWH6_5556), which has a putative role in thiazole biosynthesis. The single co-directional ORF located 3' of *thiG* is *trmB*, which encodes a putative tRNA methyltransferase.

RT-qPCR analysis of putative regulatory gene functions affecting GAF biosynthesis

Tn5 insertions in three putative regulatory genes (*prtR*, *gntR* and *iopB*) affected GAF production. RT-qPCR was performed to determine whether these three genes actually control transcription of two genes presumed to encode enzymes involved in GAF biosynthesis. The two presumptive biosynthetic

genes examined, PFWH6_5256 and PFWH6_5257, encode putative aminotransferase and formyltransferase activities, respectively, and appear likely to be involved in the formation of the formylamino group that characterizes the GAF side-chain located distal to the oxygen bridge in the GAF molecule (Fig. 1). Both of these genes reside within the 13 kb gene cluster discussed above. When expression of these genes in the *prtR*, *iopB* and *gntR* mutants (WH6-3, WH6-4 and WH6-7, respectively) was compared with their transcription in wild-type WH6, a substantial decrease in transcription of both genes was observed in the mutants (Fig. 5), with the *iopB* mutation having the greatest effect in downregulating transcription.

DISCUSSION

The ability to utilize *E. amylovora* as a mutant screen has facilitated our efforts to identify the genes in *P. fluorescens* WH6 that are involved in GAF biosynthesis. Our earlier mutagenic efforts, using the *Poa* germination bioassay as a direct screen for loss of GAF activity, resulted in the recovery of two GAF mutants (WH6-2 and WH6-3) after screening 1214 Tn5-induced mutations (Armstrong *et al.*, 2009). In the current study, we used *E. amylovora* to screen 6364 Tn5-induced mutations and identified 23 mutations affecting GAF biosynthesis. These mutations, combined with the two mutations generated in our earlier study, occurred in seven unique genes and four intergenic regions. Interestingly, none of the mutations observed in the most recent screen occurred in the two GAF-related genes identified in our earlier screen. Thus, neither screen was

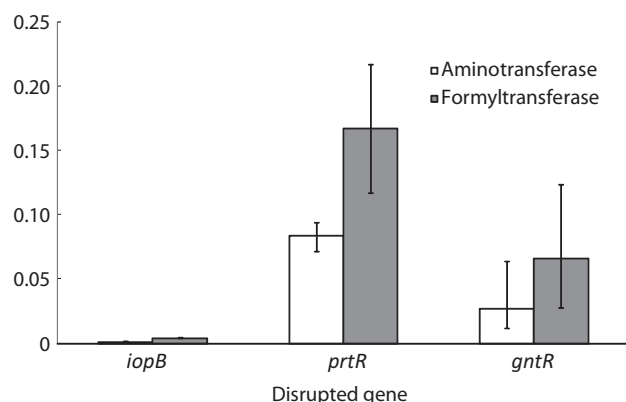


Fig. 5. Effects of mutations in regulatory genes on the transcription of the putative aminotransferase and formyltransferase genes (PFWH6_5256 and PFWH6_5257) suspected to be involved in the synthesis of dihydroGAF. Transcription levels, as evaluated by RT-qPCR, are compared with those observed in wild-type *P. fluorescens* WH6. The mutated regulatory genes are *gntR* (PFWH6_5248), *iopB* (PFWH6-5251) and *prtR* (PFWH6_3687). Error bars, SEM. All of the quantitative PCR results for the mutant strains were significantly different from wild-type with $P < 0.05$.

truly saturating, although multiple mutations were observed in three of the GAF-related genes identified in the *E. amylovora* screen.

We have now identified three genes in *P. fluorescens* WH6 that are involved in the regulation of GAF biosynthesis. Mutations in WH6 homologues of the genes *prtR*, *iopB* and *gntR* suppressed GAF formation. The *prtR* homologue and its immediate upstream neighbour (a *prtI* homologue) have high percentage identity (88%) to the *prtIR* dicistronic operon that occurs in *P. fluorescens* LS107d2 (Burger *et al.*, 2000), which has been shown to regulate the temperature-sensitive production of an extracellular protease. *PrtI* appears to have structural features similar to the sigma factors involved in regulating extracytoplasmic functions (ECF sigma factors), and *prtR* has been suggested to encode a novel transmembrane activator of *PrtI*. The role of the WH6 *iopB* homologue in the regulation of GAF production is unclear, but the gene in *P. chlororaphis* PCL1391 has been reported to regulate phenazine production in that organism through effects on quorum sensing (Van Rij, 2006). The WH6 *gntR* gene identified in our mutant screen belongs to a very large family of transcriptional regulators with over 8500 members described to date (Hoskisson & Rigali, 2009). These regulatory factors control a wide variety of metabolic processes and environmental responses, including antibiotic production (Haydon & Guest, 1991; Hoskisson & Rigali, 2009). Therefore, the role of the WH6 *gntR* gene in regulating GAF biosynthesis is consistent with the function of other members of this family of transcriptional regulators.

The WH6 *iopB* and *gntR* homologues identified in the present study occur in a 13 kb gene cluster that includes at least three other genes that appear likely to be involved in GAF biosynthesis (Fig. 3). This region includes an additional putative regulatory gene (an *iopA* homologue), five genes that appear to encode enzymes involved in biosynthetic processes, two very small ORFs whose functions and significance are unclear, and three genes whose products appear to be involved with the transport of amino acids (or amino acid-related compounds) out of the bacterial cells. At least two of the genes with putative biosynthetic functions in this cluster, genes encoding a carbamoyltransferase and the aminotransferase described above, are essential for GAF production as demonstrated by the loss of GAF production that occurred upon disruption of these genes by Tn5 insertions. A putative formyltransferase gene (PFWH6_5257) suspected to be involved in GAF biosynthesis is located immediately downstream of the aminotransferase gene.

On balance, the genetic elements identified in this study are consistent with the basic outline of a hypothetical GAF biosynthetic pathway proposed in our earlier publication (McPhail *et al.*, 2010). Based on precedents provided by pathways established for other naturally occurring oxyvinylglycines, we are reasonably confident that GAF biosynthesis begins with the amino acid homoserine (2-amino-4-hydroxybutanoic acid). Homoserine has been demonstrated as the starting point for the biosynthesis of the vinylglycine rhizobitoxine in *Burkholderia andropogonis* (Mitchell & Coddington, 1991) and *Bradyrhizobium elkanii* (Yasuta *et al.*, 2001), and for AVG in *Streptomyces* sp. NRRL 5331 (Fernández *et al.*, 2004). Addition of an amino group to the hydroxyl group of homoserine would produce canaline (2-amino-4-aminooxybutanoic acid), a naturally occurring compound known to be produced in certain legumes (Miersch, 1967; Rosenthal, 1997). The transfer of a formyl group to the 4-amino group of canaline would form a compound that can be termed dihydroGAF (2-amino-4-formylaminooxybutanoic acid). Finally, insertion of the vinyl double bond would then result in the formation of GAF itself (2-amino-4-formylaminooxy-3-butenoic acid) (Fig. 1). The GAF-related 13 kb gene cluster identified here appears to contain genes that could encode the enzymes necessary to make the dihydroGAF precursor of GAF, but this gene cluster does not contain a candidate for a gene encoding the desaturase required for insertion of the vinyl double bond. Tentative support for the pathway is provided by the results of our RT-qPCR studies of the effects of mutations in the regulatory genes *prtR*, *iopB* and *gntR*. Mutations in any one of these three genes, in addition to eliminating GAF production, also suppressed transcription of the aminotransferase and formyltransferase genes that we suspect encode the enzymes involved in the formation of the formylamino group found in the GAF side-chain.

To our knowledge, the 13 kb GAF-related gene cluster is absent from other sequenced strains of pseudomonads, with one exception. *P. syringae* pv. *maculicola* strain ES4326 (GenBank accession no. AEAK00000000) contains a gene

cluster that is almost identical to the WH6 gene cluster. Every gene of the WH6 cluster, with the exception of the gene immediately downstream of the *gntR* homologue, is present in high homology in ES4326, although not unanimously as the first BLAST hit. The missing gene (corresponding to PFWH6_5427) encodes a putative *trans*-aconitate 2-methyltransferase, which is unlikely to be involved in GAF biosynthesis. In spite of the similarity of this particular gene cluster in WH6 and ES436, the latter strain does not produce GAF (data not shown). Therefore, we suspect that ES4326 lacks one or more genes essential for the formation of the GAF vinyl double bond.

The functional diversity of the GAF-related genes identified in the present study indicates that the details of the GAF biosynthetic pathway may be somewhat more complicated than the simple model postulated by McPhail *et al.* (2010). Moreover, the biosynthesis of another oxyvinylglycine, 4-methoxyvinylglycine (2-amino-4-methoxy-3-butenonic acid) produced by *P. aeruginosa*, has been shown to involve a five-gene cluster that includes non-ribosomal peptide synthetases (Lee *et al.*, 2010), suggesting that significant diversity may exist in the biosynthetic pathways by which vinylglycines are produced. In the case of GAF, the enzyme encoded by the carbamoyltransferase gene (corresponding to PFWH6_5254), located in the GAF-related 13 kb gene cluster, may function in the synthesis of amino and/or formyl donor compounds required by the pathway, but the roles of other genes shown to affect GAF biosynthesis, particularly those located outside of the GAF-related gene cluster, are uncertain. In some instances, the effects of Tn5 insertions on GAF production could be due to polar effects on genes downstream of the point of insertion. Based on the context in which these mutations occurred, some may have affected additional regulatory functions (e.g. WH6-10::Tn5) or processes related to the desaturase function yet to be identified (e.g. WH6-9::Tn5). To clarify these issues, site-specific mutagenic studies, combined with transcriptional analyses and appropriate gene transformations, are under way.

ACKNOWLEDGEMENTS

Support from the United States Department of Agriculture Cooperative State Research, Education and Extension Service (USDA CSREES) Grass Seed Cropping Systems for Sustainable Agriculture Special Grant Program and from the Oregon State University Agricultural Research Foundation is gratefully acknowledged by D.M and D.A. Technical assistance for parts of this study was provided by Donald D. Chen. The use of trade, firm, or corporation names in this publication is for the information and convenience of the reader. Such use does not constitute an official endorsement or approval by the United States Department of Agriculture or the Agricultural Research Service of any product or service to the exclusion of others that may be suitable.

REFERENCES

- Altschul, S. F., Madden, T. L., Schäffer, A. A., Zhang, J., Zhang, Z., Miller, W. & Lipman, D. J. (1997). Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res* **25**, 3389–3402.
- Armstrong, D., Azevedo, M., Mills, D., Bailey, B., Russell, B., Groenig, A., Halgren, A., Banowetz, G. & McPhail, K. (2009). Germination-arrest factor (GAF): 3. Determination that the herbicidal activity of GAF is associated with a ninhydrin-reactive compound and counteracted by selected amino acids. *Biol Control* **51**, 181–190.
- Banowetz, G. M., Azevedo, M. D., Armstrong, D. J., Halgren, A. B. & Mills, D. I. (2008). Germination-arrest factor (GAF): biological properties of a novel, naturally-occurring herbicide produced by selected isolates of rhizosphere bacteria. *Biol Control* **46**, 380–390.
- Banowetz, G. M., Azevedo, M. D., Armstrong, D. J. & Mills, D. I. (2009). Germination arrest factor (GAF): part 2. Physical and chemical properties of a novel, naturally occurring herbicide produced by *Pseudomonas fluorescens* strain WH6. *Biol Control* **50**, 103–110.
- Berkowitz, D. B., Charette, B. D., Karukurichi, K. R. & McFadden, I. M. (2006). α -Vinyl amino acids: occurrence, asymmetric synthesis, and biochemical mechanisms. *Tetrahedron Asymmetry* **17**, 869–882.
- Burger, M., Woods, R. G., McCarthy, C. & Beacham, I. R. (2000). Temperature regulation of protease in *Pseudomonas fluorescens* LS107d2 by an ECF sigma factor and a transmembrane activator. *Microbiology* **146**, 3149–3155.
- Elliott, L. F., Azevedo, M. D., Mueller-Warrant, G. W. & Horwath, W. R. (1998). Weed control with rhizobacteria. *Soil Sci Agrochemi Ecol* **33**, 3–7.
- Fernández, M., Cuadrado, Y., Aparicio, J. F. & Martín, J. F. (2004). Role of homoserine and threonine pathway intermediates as precursors for the biosynthesis of aminoethoxyvinylglycine in *Streptomyces* sp. NRRL 5331. *Microbiology* **150**, 1467–1474.
- Gooderham, W. J., Gellatly, S. L., Sanschagrín, F., McPhee, J. B., Bains, M., Cosseau, C., Levesque, R. C. & Hancock, R. E. W. (2009). The sensor kinase PhoQ mediates virulence in *Pseudomonas aeruginosa*. *Microbiology* **155**, 699–711.
- Halgren, A., Azevedo, M., Mills, D., Armstrong, D., Thimmaiah, M., McPhail, K. & Banowetz, G. (2011). Selective inhibition of *Erwinia amylovora* by the herbicidally active germination-arrest factor (GAF) produced by *Pseudomonas* bacteria. *J Appl Microbiol* **111**, 949–959.
- Haydon, D. J. & Guest, J. R. (1991). A new family of bacterial regulatory proteins. *FEMS Microbiol Lett* **79**, 291–296.
- Hoskisson, P. A. & Rigali, S. (2009). Chapter 1 variation in form and function: the helix-turn-helix regulators of the GntR superfamily. *Adv Appl Microbiol* **69**, 1–22.
- Kimbrel, J. A., Givan, S. A., Halgren, A. B., Creason, A. L., Mills, D. I., Banowetz, G. M., Armstrong, D. J. & Chang, J. H. (2010). An improved, high-quality draft genome sequence of the germination-arrest factor-producing *Pseudomonas fluorescens* WH6. *BMC Genomics* **11**, 522–526.
- Lee, X., Fox, A., Sufrin, J., Henry, H., Majcherczyk, P., Haas, D. & Reimann, C. (2010). Identification of the biosynthetic gene cluster for the *Pseudomonas aeruginosa* antimetabolite L-2-amino-4-methoxy-*trans*-3-butenonic acid. *J Bacteriol* **192**, 4251–4255.
- Lewenza, S., Falsafi, R. K., Winsor, G., Gooderham, W. J., McPhee, J. B., Brinkman, F. S. L. & Hancock, R. E. W. (2005). Construction of a mini-Tn5-*luxCDABE* mutant library in *Pseudomonas aeruginosa* PAO1: a tool for identifying differentially regulated genes. *Genome Res* **15**, 583–589.
- Matthysse, A. G., Stretton, S., Dandie, C., McClure, N. C. & Goodman, A. E. (1996). Construction of GFP vectors for use in Gram-negative bacteria other than *Escherichia coli*. *FEMS Microbiol Lett* **145**, 87–94.

- McPhail, K. L., Armstrong, D. J., Azevedo, M. D., Banowetz, G. M. & Mills, D. I. (2010).** 4-Formylaminoxyvinylglycine, an herbicidal germination-arrest factor from *Pseudomonas* rhizosphere bacteria. *J Nat Prod* **73**, 1853–1857.
- Miersch, J. (1967).** [Detection and isolation of canaline (2-amino-4-amino-hydroxybutyric acid)]. *Naturwissenschaften* **54**, 169–170 (in German).
- Mitchell, R. E. & Coddington, J. M. (1991).** Biosynthetic pathway to rhizobitoxine in *Pseudomonas adropogonis*. *Phytochemistry* **30**, 1809–1814.
- Pfaffl, M. W., Horgan, G. W. & Dempfle, L. (2002).** Relative expression software tool (REST) for group-wise comparison and statistical analysis of relative expression results in real-time PCR. *Nucleic Acids Res* **30**, e36.
- Rigali, S., Derouaux, A., Giannotta, F. & Dusart, J. (2002).** Subdivision of the helix-turn-helix GntR family of bacterial regulators in the FadR, HutC, MocR, and YtrA subfamilies. *J Biol Chem* **277**, 12507–12515.
- Rosenthal, G. A. (1997).** L-Canaline: a potent antimetabolite and anti-cancer agent from leguminous plants. *Life Sci* **60**, 1635–1641.
- Sambrook, J. & Russell, D. W. (2001).** *Molecular Cloning: a Laboratory Manual*, 3rd edn. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory.
- Savli, H., Karadenizli, A., Kolayli, F., Gundes, S., Ozbek, U. & Vahaboglu, H. (2003).** Expression stability of six housekeeping genes: A proposal for resistance gene quantification studies of *Pseudomonas aeruginosa* by real-time quantitative RT-PCR. *J Med Microbiol* **52**, 403–408.
- Sung, M. H., Tanizawa, K., Tanaka, H., Kuramitsu, S., Kagamiyama, H., Hirotsu, K., Okamoto, A., Higuchi, T. & Soda, K. (1991).** Thermostable aspartate aminotransferase from a thermophilic *Bacillus* species. Gene cloning, sequence determination, and preliminary X-ray characterization. *J Biol Chem* **266**, 2567–2572.
- Untergasser, A., Nijveen, H., Rao, X., Bisseling, T., Geurts, R. & Leunissen, J. A. M. (2007).** Primer3Plus, an enhanced web interface to Primer3. *Nucleic Acids Res* **35** (Web Server issue,), W71–W74.
- Van Rij, E. T. (2006).** *Environmental and molecular regulation of phenazine-1-carboxamide biosynthesis in Pseudomonas chlororaphis strain PCL1391*. PhD thesis, Leiden University.
- Vandesompele, J., De Preter, K., Pattyn, F., Poppe, B., Van Roy, N., De Paepe, A. & Speleman, F. (2002).** Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. *Genome Biol* **3**, research0034–research0034.11.
- Yasuta, T., Okazaki, S., Mitsui, H., Yuhashi, K., Ezura, H. & Minamisawa, K. (2001).** DNA sequence and mutational analysis of rhizobitoxine biosynthesis genes in *Bradyrhizobium elkanii*. *Appl Environ Microbiol* **67**, 4999–5009.

Edited by: G. Preston