

## ***In vivo* expression technology (IVET) selection of genes of *Rhizobium leguminosarum* biovar *viciae* A34 expressed in the rhizosphere**

Michelle Barr<sup>1</sup>, Alison K. East<sup>2</sup>, Mary Leonard<sup>1</sup>, Tim H. Mauchline<sup>1</sup> & Philip S. Poole<sup>2</sup>

<sup>1</sup>School of Biological Sciences, University of Reading, Reading, UK; and <sup>2</sup>Department of Molecular Microbiology, John Innes Centre, Colney, Norwich, Norfolk, UK

**Correspondence:** Philip Poole, Department of Molecular Microbiology, John Innes Centre, Norwich Research Park, Colney, Norwich, Norfolk NR4 7UH, UK. Tel.: +44 0 1603 450750; fax: +44 0 1603 450045; e-mail: philip.poole@bbsrc.ac.uk

Received 19 December 2007; accepted 8 February 2008.

First published online April 2008.

DOI:10.1111/j.1574-6968.2008.01131.x

Editor: J. Colin Murrell

### **Keywords**

IVET; *Rhizobium leguminosarum*; rhizosphere.

### **Introduction**

The rhizobia are Gram-negative *Alphaproteobacteria* that form a species-specific symbiotic relationship with leguminous plants in which atmospheric N<sub>2</sub> is reduced to NH<sub>3</sub>. To establish this symbiosis, the bacteria must survive for extended periods in soils, often under salt or pH stress, be able to adapt to the nutrient-rich, but highly competitive, plant rhizosphere and finally colonize, infect and nodulate plant roots. Their ability to triumph in this range of environments appears to be reflected in their genomes which encode a wide spectrum of genes showing great metabolic diversity. The rhizobia have large, complex genomes, ranging in size from 6.5-Mb (*Sinorhizobium*) to 9.0-Mb (*Bradyrhizobium*) (Kaneko *et al.*, 2000, 2002; Galibert *et al.*, 2001). The genome of *Rhizobium leguminosarum* strain 3841 (7.75-Mb) contains a circular chromosome and six plasmids, the sequence of which has been determined (Young *et al.*, 2006).

*In vivo* expression technology (IVET), a technique devised to explore the genes expressed under a certain set of conditions (Mahan *et al.*, 1993), has been applied to bacteria

### **Abstract**

IVET was used to identify genes that are specifically expressed in the rhizosphere of the pea-nodulating bacterium *Rhizobium leguminosarum* A34. A library of *R. leguminosarum* A34 cloned in the integration vector pIE1, with inserts upstream of a promoter-less *purN:gfp:gusA*, was conjugated into *purN* host RU2249 and recombined into the genome. After removal of colonies that expressed the reporter genes of the vector under laboratory conditions, the library was inoculated into a nonsterile pea rhizosphere. The key result is that 29 rhizosphere-induced loci were identified. Sequence analysis of these clones showed that a wide variety of *R. leguminosarum* A34 genes are expressed specifically in the rhizosphere including those encoding proteins involved in environmental sensing, control of gene expression, metabolic reactions and membrane transport. These genes are likely to be important for survival and colonization of the pea rhizosphere.

living in many different environments (reviewed by Rediers *et al.*, 2005). A successful IVET strategy has several requirements: Firstly, a modified bacterial strain that cannot grow due to the lack of an essential gene must be available. It must be possible to restore the growth of the bacteria by the introduction and expression of an active copy of this gene or by the addition of supplements to the growth medium. Secondly, a plasmid is required that contains a promoter-less copy of the essential gene, upstream of which DNA fragments can be cloned to drive expression of the gene in the test environment. Such vectors often also contain one or more other promoter-less reporter genes (e.g. *lacZ*, *gusA*) (Jackson & Giddens, 2006). Within plant-associated bacteria in soil and the rhizosphere, IVET has been used to study gene expression in *Pseudomonas fluorescens* in both the sugar beet rhizosphere (Rainey, 1999; Gal *et al.*, 2003) and in soil (Silby & Levy, 2004), *Pseudomonas putida* in the maize rhizosphere (Ramos-Gonzalez *et al.*, 2005), and the colonisation and infection of roots of rice by *Pseudomonas stutzeri* (Rediers *et al.*, 2003). One of the strengths of IVET is that it selects for niche-induced genes, whether or not their expression is essential. Although Oke & Long (1999) used IVET to

study gene expression in bacteroid development in *Sinorhizobium meliloti*, the genes important for survival of rhizobia in soil and the initial interaction between bacteria and their legume-host root has been a neglected area. We therefore developed an IVET system based on complementing a purine auxotroph with *purN* in *R. leguminosarum* A34, and used it to explore the genes specifically expressed in the pea rhizosphere.

## Materials and methods

### Bacterial strains, plasmids and culture conditions

Bacterial strains and plasmids used in this study are detailed in Table 1. *Rhizobium* strains were grown at 28 °C on either tryptone yeast extract (TY) (Beringer, 1974), acid minimal salts medium (AMS) or acid minimal salts agar (AMA) (Poole *et al.*, 1994) with 10 mM D-glucose/10 mM ammonium chloride. Media was solidified either by addition of 1.5% (w/v) Oxoid agar or 1% agarose. *Escherichia coli* strains were grown at 37 °C on Luria–Bertani medium (Sambrook *et al.*, 1989). Antibiotics and additions were

used at the following concentrations ( $\mu\text{g mL}^{-1}$ ): ampicillin (Amp), 50; streptomycin (St), 500; kanamycin (Km), 20; gentamicin (Gm), 20; nyastatin 50; trimethoprim, 10; neomycin (Neo), 40; spectinomycin (Sp), 100, 100; 5-bromo-4-chloro-3-indoyl- $\beta$ -D-glucuronide (Xgluc), 50. All DNA cloning and analysis was performed as described previously (Sambrook & Russell, 2001). Plasmid DNA was sequenced using primers pIE2rev (5'-AACACGACGACCTTCTTCTT-3') and M13uni (5'-GTAAAACGACGGCCAGT-3'). PCR was performed as described previously (Karunakaran *et al.*, 2006).

### Rhizosphere competition

*Pisum sativum* (pea cultivar Avola) seeds were grown in washed vermiculite (25 mL) in Falcon tubes (50 mL), to which 12 mL nitrogen-free rooting solution was added. For experiments in a sterile rhizosphere, pea seeds were surface sterilized with 2% hypochlorite, while vermiculite and all solutions were autoclaved. Plants were left open for 7 days so some secondary contamination will have occurred but this was considered insignificant compared with plants grown without prior sterilization. For experiments in nonsterile

**Table 1.** Bacterial strains and plasmids used in this study

Strain or plasmid	Relevant characteristics	Sources or references
<i>E. coli</i>		
DH5 $\alpha$	F <sup>-</sup> $\phi$ 80lacZ $\Delta$ M15 $\Delta$ (lacZYA-argF) U169 recA1 endA1 hsdR17(r <sub>k</sub> <sup>-</sup> , m <sub>k</sub> <sup>+</sup> ) phoA supE44 thi-1gyrA96 relA1	Invitrogen
Library LB10	Library of DH5 $\alpha$ containing <i>R. leguminosarum</i> A34 DNA cloned in pIE1; Km <sup>r</sup> /Neo <sup>r</sup>	This work
<i>S. meliloti</i>		
1021	Wild type; St <sup>r</sup>	Leigh <i>et al.</i> (1985)
<i>R. leguminosarum</i>		
3841	Derivative of <i>R. leguminosarum</i> bv. <i>viciae</i> strain 300; St <sup>r</sup>	Johnston & Beringer (1975)
A34	<i>R. leguminosarum</i> bv. <i>viciae</i> (formerly known as 8401/pRL1J); St <sup>r</sup>	Downie <i>et al.</i> (1983)
J225	Purine auxotroph, <i>purN</i> , Tn5 mutant of strain A34; Km <sup>r</sup> /Neo <sup>r</sup> , St <sup>r</sup>	Stevens <i>et al.</i> (2000)
Library LB12	LB10 conjugated into <i>R. leguminosarum</i> RU2249; Km <sup>r</sup> /Neo <sup>r</sup> , St <sup>r</sup>	This work
Library LB13	Purified white colonies from LB12; Km <sup>r</sup> /Neo <sup>r</sup> , St <sup>r</sup>	This work
RU2249	A34 <i>purN</i> deletion mutant; Sp <sup>r</sup> , St <sup>r</sup>	This work
Plasmids		
pBluescript II SK –	<i>E. coli</i> cloning vector; Amp <sup>r</sup>	Stratagene
pCR4 TOPO	TOPO-adapted <i>E. coli</i> cloning vector; Amp <sup>r</sup> , Km <sup>r</sup> /Neo <sup>r</sup>	Invitrogen
pHP45 $\Omega$	pBR322 derivative carrying $\Omega$ Spc <sup>r</sup> cassette; Amp <sup>r</sup> , Sp <sup>r</sup>	Prentki & Krisch (1984)
pIE1	Integrating IVET vector, promoter-less <i>purN:gfp:gusA</i> ; Single Sall site; Km <sup>r</sup> /Neo <sup>r</sup>	This work
pJQ200SK+	Insertional mutagenesis suicide vector, pACYC derivative containing <i>sacB</i> gene, P15A origin of replication; Gm <sup>r</sup>	Quandt & Hynes (1993)
pOT2	<i>gfp+</i> vector; Gm <sup>r</sup>	Allaway <i>et al.</i> (2001)
pOT3GFP+	pOT2 with Gateway cassette; Gm <sup>r</sup>	Hosie <i>et al.</i> (2002)
pRU877	<i>gusA</i> ; Km <sup>r</sup> /Neo <sup>r</sup>	Lodwig <i>et al.</i> (2004)
pRU969	pOT3- <i>gfp+</i> with <i>purN</i> cloned into SphI/KpnI sites; Gm <sup>r</sup>	This work
pRU978	Replicating IVET vector, promoter-less <i>purN:gfp</i> ; Gm <sup>r</sup>	This work
pRU1478	PCR product of p367/rIpurnm2 cloned in pCR4 TOPO; Amp <sup>r</sup> , Km <sup>r</sup> /Neo <sup>r</sup>	This work
pRU1479	PCR product of p368/rIpurnm3 cloned in pCR4 TOPO;	This work
pRU1479b	BamHI/BstX1 excised <i>phnM</i> gene fragment from pRU1478 ligated into BamHI/BstX1 digested pRU1479; Amp <sup>r</sup> , Km <sup>r</sup> /Neo <sup>r</sup>	This work
pRU1562	<i>phnM</i> - $\Omega$ - <i>purM</i> fragment cloned as a SpeI fragment in pJQS200SK+; Gm <sup>r</sup> , Sp <sup>r</sup>	This work

rhizospheres, all sterilisation was omitted. For all experiments, 7-day-old plants were inoculated at the base of the stem with  $10^3$ – $10^4$  CFU per rhizosphere suspended in sterile water. Bacteria were prepared for inoculation by overnight growth on TY medium followed by three washes in sterile distilled water. After 7 days plants were harvested by removing the shoot, adding 20 mL phosphate-buffered saline to the root and vortex-mixing vigorously for 30 s to release bacteria adhered to the root surfaces. The suspensions were serially diluted and plated onto AMS plates containing nystatin, trimethoprim, streptomycin and neomycin. For library experiments, 24 individual plants were inoculated with  $10^4$  CFU per rhizosphere.

### Construction of *purN* host strain RU2249

To make a *purN* deletion strain, DNA fragments from *phnM* and *purM*, the genes either side of *purN* of *R. leguminosarum* A34, were PCR amplified using primer pairs p367/rlpurmn2 and p368/rlpurmn3 (p367, 5'-TTTTACTAGTAGTGAA GGTTCCGTGCGTGTG-3'; rlpurmn2, 5'-TATCGACAG GATCCATCGTTAAGGGTGCCGTCGCTAGTAGA-3'; p368, 5'-TTTTACTAGTTCAAGGCCGCAGCGTTTACCGATCC-3'; rlpurmn3, 5'-TAACGATGGATCCTGTCGATACGCCT TATCGGAGATCACCC-3') and cloned in pCR4 TOPO, forming pRU1478 and pRU1479 respectively. The *phnM* gene fragment was removed from pRU1478 by digestion with BamHI/BstX1 and ligated into BamHI/BstX1 digested pRU1479, forming pRU1479b. A 2-kb BamHI fragment containing the  $\Omega$  Sp<sup>r</sup>-cassette, from pHP45 $\Omega$ , was inserted into the BamHI site between the *phnM* and *purM* fragments in pRU1479b. This *phnM*- $\Omega$ -*purM* fragment was transferred as a SpeI fragment into vector pJQS200SK, forming pRU1562, which was conjugated into *R. leguminosarum* A34. Colonies were plated onto Sp and Neo to select for recombinants by single cross-over and then on AMA sucrose (10%)/Sp to select for double cross-over recombinants. It was confirmed that *purN* had been correctly deleted in a successful recombinant, RU2249, using Southern analysis and probing with an internal fragment of *purN* generated by PCR of *R. leguminosarum* A34 DNA with primers p617/p618 (p617, 5'-TCAGTTGAGCCTTGGGTGCG-3'; p618, 5'-TGCCGAAG GCGATTCCACC-3').

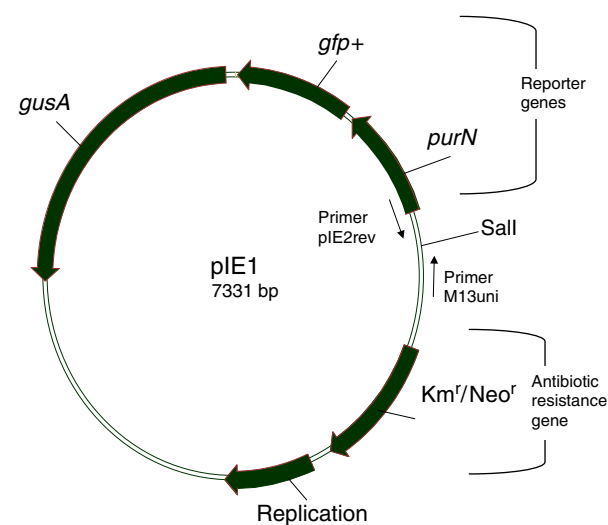
### Construction of vector pIE1

The *purN* gene of *S. meliloti* 1021 was PCR amplified with primers *purNSphI/purNKpnI* (*purNSphI*, 5'-ACCGGT ACCTCAGATGGCGGGGAAATGA-3'; *purNKpnI*, 5'-GAGAGAGCATGCGGAGGAAGAAAAATGAGCATGGTG GCCGCAAG-3'), digested with SphI/KpnI and ligated into pOT3GFP+ also digested with SphI/KpnI, forming pRU969. The *purN* and *gfp+* genes were removed from

pRU969 by digestion with SacII/SphI and ligated into pOT2 digested with the same enzymes, forming pRU978. The *purN* and *gfp+* genes were PCR amplified from pRU978 with primers p397/p398 (p397, 5'-CCTGGGAATTCAT TATTTGTAGAG-3'; p398, 5'-TTTTATGCATTGGCTTTAC TAAGCTGATCCGGTG-3') digested with EcoRI/RsrII and ligated into pRU877 also digested with EcoRI/RsrII, forming pIE1 (Fig. 1). The final vector, pIE1, is a 7.3-kb Km<sup>r</sup>/Neo<sup>r</sup> integration vector with a unique Sall cloning site used for library construction.

### Library construction

Vector pIE1, digested with Sall and 'partially filled' with CTP and TTP, was ligated with fragments greater than 1-kb in size of *R. leguminosarum* A34 Sau3AI-digested genomic DNA, which had been 'partially filled' with ATP and GTP. The 'partial filling' of vector and insert fragments, generates cohesive ends which are complementary to each other, and minimises religation of vector molecules. Transformation of *E. coli* with the ligation mixture gave *c.* 12 000 Km<sup>r</sup> colonies, which were pooled to form library LB10 ( $1 \times 10^{11}$  CFU mL<sup>-1</sup>). Ninety percent of the library plasmids were shown by PCR to contain inserts of 1–5 kb, with the average size being 2 kb. Conjugation of library LB10 into *R. leguminosarum* RU2249 yielded  $10^6$  recombinants mL<sup>-1</sup>, formed by integration via single-crossover homologous recombination between host DNA and the region of A34 genome cloned (library LB12).



**Fig. 1.** Map of vector pIE1. Promoter-less reporter genes *purN*, *gfp+* and *gusA* are found downstream of a unique Sall used for library construction. The position of binding of sequencing primers pIE2rev and M13uni is shown.

## Results and discussion

### Isolation of mutants suitable for IVET selection in the rhizosphere

To develop an IVET selection it is necessary to have a suitable auxotrophic strain that is attenuated for growth in the rhizosphere. This strain must be capable of complementation *in vivo* by the expression of the auxotrophic marker from an environmentally regulated promoter. An approach based on purine auxotrophy has been used in animal systems; selections based on *purA* were used for studying genes expressed by *Salmonella enterica* serovar Typhimurium (Heithoff *et al.*, 1997) and *Pseudomonas aeruginosa* (Handfield *et al.*, 2000) in the infection of mice. *Rhizobium leguminosarum* Tn5 mutant J225, which is derived from the parent strain A34, was reported as being unable to grow *in vitro* in the absence of purine (Stevens *et al.*, 2000). Strain J225 had been shown to be complemented by a plasmid containing *purMN* but the mutation was not further mapped (Stevens *et al.*, 2000). Therefore, primers p367 and p368, which are specific to the *pur* region, were used in combination with primer IS50 to amplify the DNA flanking the transposon in J225. The transposon position in J225 was mapped to *purN* with a 9-bp Tn5 generated repeat of CCTGCCAGG. When strain J225 was inoculated into the rhizosphere, it was recovered from a sterile rhizosphere but not recovered from a nonsterile rhizosphere, (data not shown). This showed that this purine auxotroph is unable to survive under conditions of suitable competition, and provided the starting point for development of a suitable host strain for IVET.

In order to construct a suitable host strain unable to grow in the environment, a deletion mutant lacking *purN* was made from *R. leguminosarum* strain A34 by replacing *purN* with a Sp<sup>r</sup> cartridge (strain RU2249). It was important to have as host a *purN* deletion mutant (strain RU2249), thus preventing any chance of colonies arising from homologous recombination between *purN* of pIE1 and a genomic *purN*.

When strain RU2249 was inoculated into a nonsterile rhizosphere at 10<sup>3</sup> or 10<sup>4</sup> CFU, no colonies were recovered, while strain A34 was recovered at 10<sup>7</sup>–10<sup>8</sup> CFU plant<sup>-1</sup> rhizosphere.

### Construction of an IVET library

Vector pIE1 was used to make a genomic library of strain A34 using partial infill of the single Sall site. The vector contains the promoter-less *purN* gene from *S. meliloti* and has tandem promoter-less *gusA* and *gfp* reporter genes, expression of which can be monitored on plates by auto-fluorescence or colour change of a chromogenic substrate. This library (LB10) was conjugated into the *purN* strain (RU2249) and single cross-over integrants were selected on

medium containing Neo. Neo<sup>r</sup> colonies were pooled to form library LB12. It was important to have as host a *purN* deletion mutant (strain RU2249), thus preventing any chance of colonies arising from homologous recombination between *purN* of pIE1 and a genomic *purN*.

### Determination of regions of DNA acting as promoters solely in the rhizosphere

Because the insertion in each LB12 library clone lies upstream of the promoter-less *purN:gfp:gusA* the presence of a promoter driving gene expression under different conditions can be determined by measuring expression of one, or more, of these reporter genes. In the laboratory, 87.5% of LB12 clones were blue or pale blue when grown on Neo/X-gluc plates, indicating the presence of a promoter driving gene expression of *gusA in vitro*. One thousand colonies, which were white on Neo/X-gluc plates, were pooled to form library LB13 and a sample tested for growth on agarose plates which, unlike agar plates, do not provide bacteria with an external source of purines (Stevens *et al.*, 2000). Of 900 white colonies tested, 22% failed to grow on agarose, indicating no expression of *purN*. The white colonies that were unable to grow in the absence of purine must either lack a promoter upstream of the reporter genes, or else possess a promoter not expressed under laboratory conditions. Furthermore, this shows the danger of using blue white selection alone because 78% of apparently white colonies still made enough purine to rescue growth. It is therefore essential to screen all potential IVET clones for growth without added purine to confirm the absence of a promoter active in the laboratory.

To isolate promoters specifically expressed in the rhizosphere, 10<sup>4</sup> CFU of LB13 were inoculated into a 7-day-old pea rhizosphere. After a further 7 days' growth 3300 bacteria were recovered from the rhizosphere and plated on TY plus Neo/X-gluc. Of these, 63 (2.4% of the total) were unable to grow in the absence of added purine. These 63 colonies contain putative rhizosphere regulated promoters. Each of the 63 clones was then individually inoculated back into a nonsterile pea rhizosphere; all were able to grow in the rhizosphere and were reisolated on medium containing Neo. All but one of the reisolated strains required addition of purine to laboratory medium, confirming their auxotrophy in the laboratory. The one prototroph was discarded, leaving 62 IVET strains. To ensure that the IVET clones have promoters that really are specifically expressed in the rhizosphere, six random white colonies that do not grow in the absence of added purine in the laboratory were grown in the rhizosphere of pea plants. Unlike the IVET strains, none of these six random white auxotrophic strains could be recovered from the rhizosphere.

From each of these 62 IVET strains, the DNA surrounding the insertion of pIE1 was cloned out of the genome by

digestion of total genomic DNA with SstII, religation, followed by transformation and selection of Km<sup>r</sup> colonies in *E. coli*. Km<sup>r</sup> colonies were obtained for 53 out of the 62 strains. Sequencing of the inserts in Km<sup>r</sup> plasmids allowed the exact position of the insertion of the promoter-less *purN:gfp:gusA* to be identified (Table 2). From restriction enzyme and sequence analysis it was clear that a number of the clones were identical to each other, and, in some cases, represented up to six times (Table 2). Twenty-nine clearly different clone types were identified from the 53 plasmids and these were subjected to sequencing and further analysis.

### Analysis of regions driving expression of *purN:gfp:gusA* in the rhizosphere

In order for the *R. leguminosarum* strains to survive in the rhizosphere there must be expression of *purN*, driven by DNA upstream of the reporter cassette. This is wholly

dependent on the position of insertion into the genome of RU2249 and will be different in each of the clonal types. Comparison of the results obtained from sequencing the plasmids with that of the published genome sequence of *R. leguminosarum* 3841 (Young *et al.*, 2006) revealed the genes or noncoding regions into which the insertions had occurred (Table 3). In most cases, the fragments cloned revealed the genome of *R. leguminosarum* strain A34 to be syntenous with that of *R. leguminosarum* strain 3841, however, there were differences in both gene order and gene content. For example, sequencing of pRU1858 showed one end to be in gene RL0068 (an ATP-binding component of an ABC transporter) and the other to be in RL4277 (a methyl-accepting chemotaxis protein) found *c.* 603-kb apart in strain 3841, but obviously within a few kilobases of each other in strain RU2249 (A34 background). In others, e.g. pRU1862, pRU1856, pRU1852, part or all of the plasmid DNA sequence was not highly similar to any gene in 3841

**Table 2.** Plasmids cloned from the genome of strains re-isolated from pea rhizosphere

Plasmid	Original size of insert in library clone (bp)	No. isolated	Mapping of fragment cloned relative to <i>R. leguminosarum</i> 3841*	
			Reporter gene end <sup>†</sup>	Km <sup>r</sup> /Neo <sup>r</sup> gene end <sup>‡</sup>
pRU1858	Unknown	2	RL:85947	RL:4538887
pRU1871	1701	1	RL:395942	RL:394241
pRU1863	Unknown	2	RL:519519	pRL12:418803
pRU1914	2952	1	RL:699307	RL:696355
pRU1877	2170	6	RL:759372	RL:757202
pRU1843	2788	1	RL:1543125	RL:1545913
pRU1849	3321	1	RL:1725024	RL:1728345
pRU1875	1555	1	RL:1903225	RL:1901670
pRU1867	1716	1	RL:2180511	RL:2178795
pRU1853	3145	3	RL:2864197	RL:2861052
pRU1848	2170	1	RL:2924853	RL:2927023
pRU1856	Unknown	2	RL: <i>c.</i> 3161621	NH
pRU1845	1941	1	RL:3279854	RL:3277913
pRU1864	2589	1	RL:3382638	RL:3380049
pRU1865	3580	1	RL:3389803	RL:3386223
pRU1866	3362	1	RL:3855742	RL:3859104
pRU1899	3120	1	RL:4182329	RL:4185449
pRU1868	3896	3	RL:4311081	RL:4307185
pRU1901	2967	1	RL:4637551	RL:4634584
pRU1913	Unknown	2	pRL7: <i>c.</i> 22254	pRL8:120421
pRU1842	4395	6	pRL8:78587	pRL8:82982
pRU1869	2320	2	pRL10:405270	pRL10:402950
pRU1852	Unknown	1	NH	pRL11:19337
pRU1837	Unknown	1	NH	pRL11:620394
pRU1844	2004	6	pRL12:451687	pRL12:449683
pRU1851	<i>c.</i> 2300	1	pRL12:795905	part is NH
pRU1895	Unknown	1	pRL12:866120	RL:3349036
pRU1870	Unknown	1	NH	NH
pRU1847	Unknown	1	NH	NH

\*NH, nonhomologous with *Rhizobium leguminosarum* 3841 sequence.

<sup>†</sup>Determined using primer pE2rev.

<sup>‡</sup>Determined using primer M13uni.

Precise insert sizes in the original library plasmids were deduced by synteny with the 3841 genome.

**Table 3.** Genes up-regulated in the pea rhizosphere

Plasmid	Gene*	Protein function
Class I: Motility and chemotaxis		
Flagellum/pilus biosynthesis		
pRU1877	RL0710	Flagellar basal-body rod protein FlgG
Class II: Nutrient scavenging		
Sulphate acquisition		
pRU1848	RL2767	Putative aliphatic sulfonate transport ATP-binding protein SsuB
Sugar transport		
pRU1858	RL0068	Putative ATP-binding protein of ABC transporter (CUT1 family)
pRU1871	RL0362	Putative MFS family transmembrane (permease) transport protein AraJ
Class III: Central intracellular metabolism		
Amino acid degradation		
pRU1844	pRL120415	Putative alanine catabolic operon regulator DadR
Sugar metabolism		
pRU1914	RL0649	Putative PfkB family sugar kinase
Phospholipid metabolism		
pRU1901	RL4366	Putative ethanolamine ammonia-lyase light chain EutC
Lipid/fatty acid synthesis		
pRU1863	RL0481	Putative 3-oxoacyl-[acyl-carrier-protein] reductase FabG
Aromatic compound metabolism		
pRU1870	Homologue of RHE_CH01085	c. 62% identity over 64 residues with putative esterase protein of <i>Rhizobium etli</i> CFN 42
General metabolism		
PRU1845	RL3121	Putative acetyltransferase
Class IV: Stress response and adaptation		
Detoxification: Antibiotic resistance		
pRU1852	Homologue of SMA0669	c. 71% identity over 140 residues with putative multi-drug efflux system protein of <i>S. meliloti</i> 1021
pRU1847	Homologue of Sala_1765	c. 40% identity over 330 residues with aminoglycoside phosphotransferase of <i>Sphingomonas alaskensis</i> RB2256
Class V: Regulation		
Two-component regulatory systems		
pRU1895	pRL120793	Putative repetitive two-component sensor histidine kinase transcriptional regulatory protein
Transcriptional regulators		
pRU1869	pRL100388	Putative transcriptional regulator (GntR family)
pRU1842	pRL80074	Putative transcriptional regulator (LysR family)
Transcription factors		
PRU1843	RL1473	Conserved hypothetical protein with RNA polymerase sigma factor domain
Environmental sensing		
pRU1853	RL2705	Putative GGDEF sensory box protein
PRU1899	RL3958	Putative sensory box GGDEF/EAL domain protein
Class VII: Virulence and secretion		
Type I secretion		
pRU1866	RL3659–RL3658†	Putative polysaccharidase (PlyA) or could be regulated by the promoter of the next gene - PrsD (Type I secretion system)
Class X: Unknown		
pRU1851	pRL120730	Hypothetical protein (analogous to SMc00958)
pRU1856	RL2995	Conserved hypothetical protein
Class XI: Reverse orientation		
pRU1864	Antisense RL3230	Antisense of putative Rrf2 family transcriptional regulator protein
pRU1849	Antisense RL1649	Antisense of conserved hypothetical protein
pRU1875	Antisense RL1813	Antisense of putative 2-dehydropantoate 2-reductase
pRU1867	Antisense RL2064	Antisense of putative methyltransferase
pRU1868	Antisense RL4081	Antisense of putative UTP-glucose-1-phosphate uridylyltransferase (ExoN)
pRU1865	Antisense RL3242	Antisense of putative transmembrane transporter protein
pRU1837	Antisense acetyl transferase	Antisense of unknown acetyl transferase with 82% id over 166 amino acids with <i>R. etli</i> CFN 42.

No ORF was found upstream of reporter gene in pRU1913.

\*Orthologue of *Rhizobium leguminosarum* 3841 unless otherwise stated.

†Insertion is in an intergenic region.

and therefore the position of insertion could not be precisely identified relative to the 3841 genome. The Sym plasmid of strain A34 is pRL1JI rather than pRL10JI which is found in the sequenced strain 3841.

### Genes expressed solely in the rhizosphere

The genes identified as driving expression of *purN* in the rhizosphere can be split into groups depending on their function (as described by Rediers *et al.*, 2005) and are shown in Table 3.

Within Class I (motility and chemotaxis), this IVET study has identified RL0710, *flgG*, which lies in the middle of a cluster of genes involved in chemotaxis and motility. Other IVET studies have identified bacterial expression of flagella components in the rhizosphere; flagellar M-ring protein *fliF* in *P. fluorescens* in sugar beet rhizosphere (Gal *et al.*, 2003) and flagellar assembly protein *fliO* in *P. putida* in maize rhizosphere (Ramos-Gonzalez *et al.*, 2005). In *Pseudomonas*, flagella have been shown to be important in root colonization (De Weger *et al.*, 1987) and in seed adhesion (DeFlaun *et al.*, 1994). This study does not distinguish between colonisation and survival in the rhizosphere. A means of getting about is important for survival and growth in this often nutrient-poor environment.

Within the genes of *R. leguminosarum* that are involved in nutrient scavenging (Class II-Table 3) is RL2767 (*ssuB*), encoding an ATP-binding component (ABC) of a putative aliphatic sulphonate ABC transport system. Such a system is likely to be involved in uptake of sulphate-containing compounds. The periplasmic solute binding protein of a sulphate ABC transporter was identified with IVET in *P. putida* in the maize rhizosphere (Ramos-Gonzalez *et al.*, 2005). This illustrates that uptake of sulphate-containing compounds is important in organisms facing survival in the rhizosphere. Two other transporters were identified; RL0068 encodes an ABC of the CUT1 sugar uptake family (Saier, 2000), and RL0362 (*araJ*), encodes a permease of major facilitator system (MFS) family transporter. The latter probably forms an operon with RL0363, encoding a glyoxalase, and may be involved with processing and transport of arabinose polymers. Arabinogalactan is very common in pea mucilage, the complex polysaccharide secreted by plant roots and likely to provide a significant source of carbon for rhizosphere organisms (Knee *et al.*, 2001). Root mucilage has been shown to be able to act as the sole carbon source for growth of *R. leguminosarum* 8401, which is the same strain as A34 (Knee *et al.*, 2001).

According to Jackson & Giddens (2006) the most common class of genes identified by IVET in Class III, central intermediary metabolism, are those concerned with protein synthesis and degradation. One such example of this is pRL120415, encoding DadR, an AsnC family regulator of

the alanine catabolic operon (Allaway *et al.*, 2000). As DadR is a positive regulator, an increase in its concentration enhances the transcription of genes needed for breakdown of alanine. Other genes induced by the rhizosphere include those involved in carbohydrate metabolism (sugar kinase, RL0649), phospholipid metabolism [putative ethanolamine ammonia-lyase light chain (*eutC*), RL4366] and lipid and fatty acid synthesis [putative 3-oxoacyl-[acyl-carrier-protein] reductase (*fabG*), RL0401] (Table 3). RL3121, encoding a putative acyltransferase, was induced in the rhizosphere. Other acyltransferases have been identified in bacteria upon their interaction with plants (Yang *et al.*, 2004; Marco *et al.*, 2005).

We have identified two genes, induced in the rhizosphere, which are concerned with stress response and adaptation (Class IV). In pRU1852 and pRU1847, the DNA upstream of the reporter gene shows homology with genes encoding putative multi-drug resistance proteins. It seems likely that expression of these proteins in soil and rhizosphere is due to the need to remove toxic compounds from the cell. Genes encoding other de-toxication mechanisms have been found to be expressed in other rhizosphere-colonizing bacteria using IVET (Jackson & Giddens, 2006).

Of those IVET-identified genes placed into the group concerned with regulation (Class V), pRL120793 encodes a putative repetitive two-component sensor histidine kinase transcriptional regulator. Two-component systems are widespread in prokaryotes, enabling cells to respond to environmental stimuli through changes in gene expression. Using IVET, the elevated expression of a gene encoding a histidine kinase sensor was shown by *P. putida* in the maize rhizosphere (Ramos-Gonzalez *et al.*, 2005) and also by *P. fluorescens* in soil (Silby & Levy, 2004). Many other examples of expression of two-component systems in bacteria associated with plants are listed by Jackson & Giddens (2006). Two genes (pRL100388 and pRL80074) encoding transcriptional regulators, and a gene encoding a protein with a sigma factor-like domain (RL1473) were identified. Two genes concerned with environmental sensing by the metabolism of the novel secondary messenger cyclic di-GMP were identified. The GGDEF domain (found in RL2705 and RL3958) synthesizes cyclic di-GMP, while the EAL domain (of RL3958) is involved in cyclic di-GMP hydrolysis.

A single gene involved in virulence and secretion (Class VII) was identified in this study. The reporter gene was found to have inserted between genes RL3659 and RL3658 (372 nucleotides downstream from the end of RL3659 and 94 nucleotides before the start of RL3658). The observed expression of the reporter gene may be due to the promoter of RL3659 (encoding PlyA, an exopolysaccharidase-glucanase) or be driven by the region upstream of RL3658, encoding PrsD, part of a Type I secretion system (ABC export system) responsible for the export of PlyA.

Expression and export of PlyA is required for effective biofilm formation (Russo *et al.*, 2006). In the relatively nutrient-rich environment around plant roots, in order to survive bacteria need to interact and compete with other bacteria within the community. This community is likely to take the form of a biofilm, in which the individual cells stick to surfaces, interfaces and/or to each other. Biofilms have been observed on plant root and leaf surfaces (Morris & Monier, 2003) and it is likely to be the way bacteria exist in the natural environment (Davey & O'Toole, 2000).

As in many IVET studies, we have isolated several colonies in which the reporter gene fusion is in the reverse orientation, i.e. on the antisense strand of a 'genuine' annotated coding region (Class XI). Such constructs have been routinely isolated in IVET studies and two hypotheses have been put forward to explain their occurrence; either that they are artifacts of the IVET system or that the promoter activity is real and fulfils a purpose (Jackson & Giddens, 2006). Jackson & Giddens (2006) consider from the routine occurrence of such isolates that they are likely to be genuine. Possible explanations for the expression of the reporter gene include; being driven by antisense RNA of the gene into which the reporter is inserted, from 'read through' of mRNA from a gene further upstream or from signals within the DNA, immediately upstream of the reporter, acting as promoter, possibly promoters of small RNA molecules now being more frequently recognized as regulators of gene expression. The isolates in this class are listed in Table 3 and make up *c.* 24% of the total.

Overall this IVET strategy has led to the successful isolation and identification of genes of *R. leguminosarum* A34 expressed specifically in the pea rhizosphere. They include genes encoding proteins involved in metabolic reactions, membrane transport, environmental sensing and gene regulation. Perhaps the greatest strength of IVET is the ability to probe micro-niches where obtaining sufficient material for microarray analysis is difficult. It is therefore ideally suited to investigation of the pea rhizosphere.

## Acknowledgements

The work in this laboratory is supported by the Biotechnology and Biological Sciences Research Council.

## References

- Allaway D, Lodwig E, Crompton LA, Wood M, Parsons TR, Wheeler T & Poole PS (2000) Identification of alanine dehydrogenase and its role in mixed secretion of ammonium and alanine by pea bacteroids. *Mol Microbiol* **36**: 508–515.
- Allaway D, Schofield NA, Leonard ME, Gilardoni L, Finan TM & Poole PS (2001) Use of differential fluorescence induction and optical trapping to isolate environmentally induced genes. *Environ Microbiol* **3**: 397–406.
- Beringer JE (1974) R factor transfer in *Rhizobium leguminosarum*. *J Gen Microbiol* **84**: 188–198.
- Davey ME & O'Toole GA (2000) Microbial biofilms: from ecology to molecular genetics. *Microbiol Mol Biol Rev* **64**: 847–867.
- DeFlaun MF, Marshall BM, Kulle EP & Levy SB (1994) Tn5 insertion mutants of *Pseudomonas fluorescens* defective in adhesion to soil and seeds. *Appl Environ Microbiol* **60**: 2637–2642.
- De Weger LA, van der Vlugt CI, Wijffes AH, Bakker PA, Schippers B & Lugtenberg B (1987) Flagella of a plant-growth-stimulating *Pseudomonas fluorescens* strain are required for colonization of potato roots. *J Bacteriol* **169**: 2769–2773.
- Downie JA, Ma QS, Knight CD, Hombrecher G & Johnston AWB (1983) Cloning of the symbiotic region of *Rhizobium leguminosarum* - the nodulation genes are between the nitrogenase genes and a *nifA*-like gene. *EMBO J* **2**: 947–952.
- Gal M, Preston GM, Massey RC, Spiers AJ & Rainey PB (2003) Genes encoding a cellulosic polymer contribute toward the ecological success of *Pseudomonas fluorescens* SBW25 on plant surfaces. *Mol Ecol* **12**: 3109–3121.
- Galibert F, Finan TM, Long SR *et al.* (2001) The composite genome of the legume symbiont *Sinorhizobium meliloti*. *Science* **293**: 668–672.
- Handfield M, Lehoux DE, Sanschagrín F, Mahan MJ, Woods DE & Levesque RC (2000) *In vivo*-induced genes in *Pseudomonas aeruginosa*. *Infect Immun* **68**: 2359–2362.
- Heithoff DM, Conner CP, Hanna PC, Julio SM, Hentschel U & Mahan MJ (1997) Bacterial infection as assessed by *in vivo* gene expression. *Proc Natl Acad Sci USA* **94**: 934–939.
- Hosie AHF, Allaway D & Poole PS (2002) A monocarboxylate permease of *Rhizobium leguminosarum* is the first member of a new subfamily of transporters. *J Bacteriol* **184**: 5436–5448.
- Jackson RW & Giddens SR (2006) Development and application of *in vivo* expression technology (IVET) for analysing microbial gene expression in complex environments. *Infect Disord - Drug Targets* **6**: 207–240.
- Johnston AWB & Beringer JE (1975) Identification of the *Rhizobium* strains in pea root nodules using genetic markers. *J Gen Microbiol* **87**: 343–350.
- Kaneko T, Nakamura Y, Sato S *et al.* (2000) Complete genome structure of the nitrogen-fixing symbiotic bacterium *Mesorhizobium loti*. *DNA Res* **7**: 331–338.
- Kaneko T, Nakamura Y, Sato S *et al.* (2002) Complete genomic sequence of nitrogen-fixing symbiotic bacterium *Bradyrhizobium japonicum* USDA110. *DNA Res* **9**: 189–197.
- Karunakaran R, Ebert K, Harvey S, Leonard ME, Ramachandran V & Poole PS (2006) Thiamine is synthesized by a salvage pathway in *Rhizobium leguminosarum* bv. *viciae* strain 3841. *J Bacteriol* **188**: 6661–6668.
- Knee EM, Gong FC, Gao MS, Teplitski M, Jones AR, Foxworthy A, Mort AJ & Bauer WD (2001) Root mucilage from pea and

- its utilization by rhizosphere bacteria as a sole carbon source. *Mol Plant Microbe Interact* **14**: 775–784.
- Leigh JA, Signer ER & Walker GC (1985) Exopolysaccharide-deficient mutants of *Rhizobium meliloti* that form ineffective nodules. *Proc Natl Acad Sci USA* **82**: 6231–6235.
- Lodwig E, Kumar S, Allaway D, Bourdès A, Prell J, Priefer U & Poole P (2004) Regulation of L-alanine dehydrogenase in *Rhizobium leguminosarum* bv. *viciae* and its role in pea nodules. *J Bacteriol* **186**: 842–849.
- Mahan MJ, Schlauch JM & Mekalanos JJ (1993) Selection of bacterial virulence genes that are specifically induced in host tissues. *Science* **259**: 686–688.
- Marco ML, Legac J & Lindow SE (2005) *Pseudomonas syringae* genes induced during colonization of leaf surfaces. *Environ Microbiol* **7**: 1379–1391.
- Morris CE & Monier J-M (2003) The ecological significance of biofilm formation by plant-associated bacteria. *Annu Rev of Phytopathol* **41**: 429–453.
- Oke V & Long SR (1999) Bacterial genes induced within the nodule during the *Rhizobium*-legume symbiosis. *Mol Microbiol* **32**: 837–849.
- Poole PS, Schofield NA, Reid CJ, Drew EM & Walshaw DL (1994) Identification of chromosomal genes located downstream of *dctD* that affect the requirement for calcium and the lipopolysaccharide layer of *Rhizobium leguminosarum*. *Microbiology* **140**: 2797–2809.
- Prentki P & Krisch HM (1984) *In vitro* insertional mutagenesis with a selectable DNA fragment. *Gene* **29**: 303–313.
- Quandt J & Hynes MF (1993) Versatile suicide vectors which allow direct selection for gene replacement in Gram-negative bacteria. *Gene* **127**: 15–21.
- Rainey PB (1999) Adaptation of *Pseudomonas fluorescens* to the plant rhizosphere. *Environ Microbiol* **1**: 243–257.
- Ramos-Gonzalez MI, Campos MJ & Ramos JL (2005) Analysis of *Pseudomonas putida* KT2440 gene expression in the maize rhizosphere: *in vitro* expression technology capture and identification of root-activated promoters. *J Bacteriol* **187**: 4033–4041.
- Rediers H, Bonnecarrere V, Rainey PB, Hamonts K, Vanderleyden J & De Mot R (2003) Development and application of a *dapB*-based *in vivo* expression technology system to study colonization of rice by the endophytic nitrogen-fixing bacterium *Pseudomonas stutzeri* A15. *Appl Environ Microbiol* **69**: 6864–6874.
- Rediers H, Rainey PB, Vanderleyden J & De Mot R (2005) Unraveling the secret lives of bacteria: use of *in vivo* expression technology and differential fluorescence induction promoter traps as tools for exploring niche-specific gene expression. *Microbiol Mol Biol Rev* **69**: 217–261.
- Russo DM, Williams A, Edwards A, Posadas DM, Finnie C, Dankert M, Downie JA & Zorreguieta A (2006) Proteins exported via the PrsD-PrsE type I secretion system and the acidic exopolysaccharide are involved in biofilm formation by *Rhizobium leguminosarum*. *J Bacteriol* **188**: 4474–4486.
- Saier MH (2000) A functional-phylogenetic classification system for transmembrane solute transporters. *Microbiol Mol Biol Rev* **64**: 354–411.
- Sambrook J & Russell DW (2001) *Molecular Cloning: A Laboratory Manual*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York.
- Sambrook J, Fritsch EF & Maniatis T (1989) *Molecular Cloning: A Laboratory Manual*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York.
- Silby MW & Levy SB (2004) Use of *in vivo* expression technology to identify genes important in growth and survival of *Pseudomonas fluorescens* Pf0-1 in soil: discovery of expressed sequences with novel genetic organization. *J Bacteriol* **186**: 7411–7419.
- Stevens JB, deLuca NG, Beringer JE, Ringer JP, Yeoman KH & Johnston AWB (2000) The *purMN* genes of *Rhizobium leguminosarum* and a superficial link with siderophore production. *Mol Plant Microbe Interact* **13**: 228–231.
- Yang S, Perna NT, Cooksey DA, Okinaka Y, Lindow SE, Ibekwe AM, Keen NT & Yang CH (2004) Genome-wide identification of plant-upregulated genes of *Erwinia chrysanthemi* 3937 using a GFP-based IVET leaf array. *Mol Plant Microbe Interact* **17**: 999–1008.
- Young JP, Crossman L, Johnston A *et al.* (2006) The genome of *Rhizobium leguminosarum* has recognizable core and accessory components. *Genome Biol* **7**: R34.