A Key Regulator of the Glycolytic and Gluconeogenic Central Metabolic Pathways in *Sinorhizobium meliloti*

George C. diCenzo, Zahed Muhammed, Magne Østerås, Shelley A. P. O'Brien, and Turlough M. Finan¹

Department of Biology, McMaster University, Hamilton, Ontario L8S 4K1, Canada

ORCID IDs: 0000-0003-3889-6570 (G.C.d.); 0000-0003-4363-6001 (T.M.F.)

ABSTRACT The order Rhizobiales contains numerous agriculturally, biotechnologically, and medically important bacteria, including the rhizobia, and the genera *Agrobacterium*, *Brucella*, and *Methylobacterium*, among others. These organisms tend to be metabolically versatile, but there has been relatively little investigation into the regulation of their central carbon metabolic pathways. Here, RNA-sequencing and promoter fusion data are presented to show that the PckR protein is a key regulator of central carbon metabolism in *Sinorhizobium meliloti*; during growth with gluconeogenic substrates, PckR represses expression of the complete Entner–Doudoroff glycolytic pathway and induces expression of the *pckA* and *fbaB* gluconeogenic genes. Electrophoretic mobility shift assays indicate that PckR binds an imperfect palindromic sequence that overlaps the promoter or transcriptional start site in the negatively regulated promoters, or is present in tandem upstream the promoter motifs in the positively regulated promoters. Genetic and *in vitro* electrophoretic mobility shift assay experiments suggest that elevated concentrations of a PckR effector ligand results in the dissociation of PckR from its target binding site, and evidence is presented that suggests phosphoenolpyruvate may function as the effector. Characterization of missense *pckR* alleles identified three conserved residues important for increasing the affinity of PckR for its cognate effector molecule. Bioinformatics analyses illustrates that PckR is limited to a narrow phylogenetic range consisting of the *Rhizobiaceae*, *Phyllobacteriaceae*, *Brucellaceae*, and *Bartonellaceae* families. These data provide novel insights into the regulation of the core carbon metabolic pathways of this pertinent group of α -proteobacteria.

KEYWORDS LacI/GalR transcriptional regulators; glycolysis; gluconeogenesis; phosphoenolpyruvate; rhizobia

BACTERIA are commonly present in nutritionally complex environments and may move between niches. The fitness of a species is therefore influenced by the ability to utilize diverse nutrient sources, as well as the capability to induce and repress central and peripheral metabolic pathways. *Escherichia coli* and *Bacillus subtilis* have historically served as the primary model species for the study of these topics. However, many species do not follow the rules established in these organisms, including many economically important species such as the photosynthetic and biohydrogen-producing *Rhodospirillum rubrum* and the nitrogen-fixing legume symbiont *Sinorhizobium meliloti*. It is therefore important to study regulation and metabolism in diverse

model species in order to develop a broad understanding and to optimize success in the biotechnological manipulation of these organisms.

S. meliloti is a metabolically versatile organism (Biondi *et al.* 2009; Geddes and Oresnik 2014) of the Rhizobiales order of the α -proteobacteria, which includes N_2 -fixing plant symbionts (rhizobia), plant pathogens (*Agrobacterium*, *Liberobacter*), livestock and human pathogens (*Brucella*, *Ochrobactrum*), and methylotrophs (*Methylobacterium*), among others. The preferred carbon source for *S. meliloti* appears to be C_4 -dicarboxylic acids (succinate, malate, fumarate), which also serve as the primary carbon source available to rhizobia during symbiotic nitrogen fixation (Ronson *et al.* 1981; Finan *et al.* 1983; Bolton *et al.* 1986). There are several differences between the metabolism of *S. meliloti* and its regulation compared to *E. coli* and *B. subtilis* (Geddes and Oresnik 2014). Notable differences include the use of the Entner–Doudoroff (ED) pathway for glycolysis (Stowers 1985), and the presence of succinate-mediated catabolite repression resulting in diauxic growth and the preferential

Copyright © 2017 by the Genetics Society of America

doi: <https://doi.org/10.1534/genetics.117.300212>

Manuscript received July 4, 2017; accepted for publication August 24, 2017; published Early Online August 29, 2017.

Supplemental material is available online at www.genetics.org/lookup/suppl/doi:10.1534/genetics.117.300212/-/DC1.

¹Corresponding author: Department of Biology, McMaster University, 1280 Main St. W., Hamilton, ON L8S 4K1, Canada. E-mail: finan@mcmaster.ca

use of succinate before glucose when both are present (Ucker and Signer 1978; Bringham and Gage 2002; Zhang *et al.* 2016).

Glucose is predominately metabolized through the ED pathway in *S. meliloti*, with a small portion flowing through the pentose phosphate (PP) pathway (Figure 1) (Fuhrer *et al.* 2005). The ED pathway produces one pyruvate and one glyceraldehyde-3-phosphate (G3P). The pyruvate can be directly converted into acetyl-CoA for entry into the tricarboxylic acid (TCA) cycle. The G3P can be further metabolized via the lower half of the Embden–Meyerhoff–Parnas (EMP) pathway to a second pyruvate (Figure 1) (Geddes and Oresnik 2014). When growing with substrates that enter central carbon metabolism directly through the TCA cycle, some carbon is diverted from the TCA cycle to pyruvate for synthesis of acetyl-CoA and continued production of citrate and other TCA cycle intermediates. Additional carbon is converted into phosphoenolpyruvate (PEP) that is metabolized to glucose and other sugars required for biosynthesis of essential polysaccharides, nucleotides, and amino acids. This occurs through reversal of the lower and upper halves of the EMP pathway in a process known as gluconeogenesis (Figure 1) (Finan *et al.* 1988; Geddes and Oresnik 2014). Synthesis of pyruvate from malate is catalyzed by the DME and TME malic enzymes (Driscoll and Finan 1993, 1996, 1997), and synthesis of PEP from oxaloacetate is catalyzed by phosphoenolpyruvate carboxykinase (PCK) (Østerås *et al.* 1995).

The interconnectedness and cyclic nature of the ED, EMP, PP, and TCA pathways (Figure 1) suggest a necessity for strict regulation of the key metabolic crossroads. In *S. meliloti*, this is accomplished in part with the LacI-type transcriptional regulator PckR, which regulates expression of the *pckA* gene encoding for the PCK enzyme (Østerås *et al.* 1997). Expression of *pckA* is low during growth with glucose or lactose and high during growth with succinate or arabinose (Østerås *et al.* 1995, 1997). The *pckR* gene was identified as the locus to which three independent spontaneous mutations that influenced *pckA* expression mapped (Østerås *et al.* 1997). These *pckR* alleles resulted in elevated *pckA* expression and were recessive to the wild-type allele, but the actual nucleotide changes that resulted in the mutant alleles were not determined (Østerås *et al.* 1997). *In silico* studies predicted that PckR may regulate as many as 81 genes in *S. meliloti* (Galardini *et al.* 2015), suggesting that PckR may be a global regulator of carbon metabolism. In *R. sphaeroides*, the CceR regulator fulfills the role of a global regulator of carbon and energy metabolism; RNA-sequencing (RNA-seq) showed that CceR was required for the correct regulation of 225 genes, 31 of which are involved in central carbon and energy metabolism (Imam *et al.* 2015). 6-Phosphogluconate appeared to be the effector molecule of CceR, with elevated levels of 6-phosphogluconate resulting in the dissociation of CceR from its target promoters *in vitro* (Imam *et al.* 2015). However, despite potentially serving similar biological roles, PckR and CceR are nonorthologous proteins and the similarity in their properties is unclear.

Here, we provide RNA-seq data that shows PckR is a key regulator of central carbon metabolism in *S. meliloti*. The biological role and binding activity of PckR is examined with a combination of genetic and *in vitro* biochemical characterizations, and its phylogenetic distribution is presented.

Materials and Methods

Bacterial growth conditions, media, and genetic manipulations

All media [LB (Luria broth), LBmc (LB supplemented with magnesium and calcium), M9 minimal medium, and MM9 (M9 with the phosphate buffer replaced with a 3-[N-morpholino]propanesulfonic acid [MOPS] - potassium hydroxide [KOH] buffer with 2 mM phosphate) minimal medium), antibiotic concentrations, and growth conditions for *S. meliloti* and *E. coli* were as previously described (diCenzo *et al.* 2014). Cultures for RNA isolation were grown in M9 minimal media. Subsequent growth and expression experiments were performed using MM9 minimal media, as we found growth was more rapid and there was somewhat less background GFP fluorescence. Unless stated otherwise, all carbon sources in the minimal media were added to a final concentration of 15 mM. DNA manipulations and recombinant techniques, bacterial matings, and Φ M12 transductions were performed as described before (Finan *et al.* 1984; Sambrook *et al.* 1989; Cowie *et al.* 2006; Milunovic *et al.* 2014). Recombinant plasmids were produced through either ligations with T4 DNA ligase or through sequence and ligation independent cloning (Sambrook *et al.* 1989; Jeong *et al.* 2012). All strains used in this study are listed in Supplemental Material, Table S1 in [File S1](#), and oligonucleotide primer sequences are provided in Table S2 in [File S1](#).

Construction of the *pckR::ΩSp^R* allele

The Ω Sp^R cassette from pHP45 Ω was ligated into HindIII-digested pTH296, introducing the Ω Sp^R cassette 23-bp downstream of the start of the *pckR* open reading frame. (Blondelet-Rouault *et al.* 1997; Østerås *et al.* 1997). The 4.5-kb region from the resulting plasmid was then subcloned into pRK7813 as an EcoRI fragment, producing pTH476. The *pckR::ΩSp^R* allele was recombined into the *S. meliloti* Rm5000 (Finan *et al.* 1984) genome and then transduced into *S. meliloti* RmG212 to produce RmK124, the *pckR* null mutant used in this study.

Construction of the *gfp-lacZ* fusion strains

Plasmids for construction of the *pckA*⁺::*gfp-lacZ* and *zwf*⁺::*gfp-lacZ* alleles were identified in the *S. meliloti* fusion library (Cowie *et al.* 2006). For construction of the *fbaB*⁺::*gfp-lacZ*, *eda2*⁺::*gfp-lacZ*, and *mgsA*⁺::*gfp-lacZ* strains, each promoter region was PCR amplified (oligos: *fbaB* – DF106/DF107; *eda2* – DF108/DF109; *mgsA* – DF110/DF111), cloned into XhoI-digested pTH1522 via SLIC, and transferred via conjugation into the desired *S. meliloti* strain (Cowie *et al.* 2006).

Figure 1 Schematic of central carbon metabolism in *Sinorhizobium meliloti*. The pathways and associated genes for glycolysis, gluconeogenesis, and the TCA cycle are shown. Incomplete representations of the pentose phosphate pathway (PPP), the methylglyoxal pathway (MGP), as well as glycerol and galactose metabolism are also shown in gray. Blue indicates genes repressed by PckR during growth with succinate, and red highlights genes that are induced by PckR during growth with succinate. Light blue is used for *glk* as the level of induction (~ 2.7 -fold) was below the threefold cutoff used in this study. The primary carbon sources used in this study are indicated in magenta. The schematic was produced with the assistance of iPath v2 (Yamada *et al.* 2011), and the gene associations are based on the annotations of an *S. meliloti* metabolic model (diCenzo *et al.* 2016a).

Construction of *pckR* expression vectors

The coding sequence of the *S. meliloti* Rm1021 *pckR* gene (*smc02975*) and the putative *pckR* homolog (*azc_4468*) of *Azorhizobium caulinodans* ORS 571 were PCR amplified (oligos: *smc02975* – DF112/DF113; *azc_4468* – DF114/DF115). In each case, the forward primer contained a ribosome binding site active in *S. meliloti* (Poysti and Oresnik 2007) placed 7-nt upstream of the start codon, as well as a stop codon overlapping the 5' end of the ribosome binding site by 1 nt. Each PCR product was cloned into *Hind*III-digested pRK7813 such that the stop codon was in frame with the N-terminal fragment of the *lacZ α* gene. The constructs were conjugated into *S. meliloti* RmK124 (*pckR*:: Ω Sp^R), and the gene of interest expressed from the *lac* promoter without IPTG induction.

Whole-genome sequencing and analysis

S. meliloti total genomic DNA was isolated as described previously (Cowie *et al.* 2006) from *S. meliloti* RmG212 (*pckR*⁺), RmH166 (*rpk-9*), RmH167 (*rpk-10*), and RmH172 (*rpk-15*). Genome sequences were obtained using the Illumina MiSeq V2 technology, with 250 bp paired-end reads, at the Farncombe Family Digestive Health Research Institute located at McMaster University (Ontario, Canada). The location of mutations were determined by mapping the sequencing reads to the published *S. meliloti* Rm1021 reference genome (Galibert *et al.* 2001) using Geneious R8, followed by the identification of polymorphisms as described elsewhere (diCenzo *et al.* 2016b).

Growth experiments and GFP measurements

Strains were pregrown overnight in LBmc with the appropriate antibiotics. Growth experiments were set up in 96-well microtiter plates as described previously (diCenzo *et al.* 2017), and as described further in the Supplemental Meth-

ods in File S1. The microtiter plate was incubated in a Biotek Cytation 3 plate reader with shaking at 30°, and OD₆₀₀ and GFP fluorescence (excitation/emission: 485/514 nm) were measured every 15 min.

β -Galactosidase assays

Reactions were performed in 96-well microtiter plates largely as described previously (Cowie *et al.* 2006), and as described further in the Supplemental Methods in File S1.

RNA isolation

Duplicate cultures of *S. meliloti* RmG212 and RmK124 were grown overnight in LBmc. Cultures were pelleted, washed once, and resuspended with carbon-free M9 medium. Each cell suspension was used to inoculate 50 ml cultures of M9 with 15 mM glucose and M9 with 15 mM succinate to a starting OD₆₀₀ ~ 0.025 . Cultures were grown with shaking at 30° and cells harvested when cultures reached an OD₆₀₀ between 0.35 and 0.45. Cultures were harvested by immediately mixing 45 ml culture with 5 ml ice-cold cell stop solution (5% unbuffered phenol, 95% ethanol), centrifuging at 3700 $\times g$ for 10 min, and flash-freezing the pellet with liquid nitrogen. Frozen pellets were stored at -80° until use. Total RNA was isolated largely as described previously (MacLellan *et al.* 2005), and is described in detail in the Supplemental Methods in File S1.

RNA-seq and analysis

Depletion of ribosomal RNA was performed using Ribo-Zero rRNA Removal kits (Illumina), and sequencing was performed using an Illumina HiSeq 1500 with 60 nt single reads at the Farncombe Family Digestive Health Research Institute located at McMaster University. Sequencing data were analyzed using the Rockhopper v2.03 software (Tjaden 2015), with default settings and the following modifications: reverse complement reads and verbose output were selected. Reads

were mapped to the reference *S. meliloti* Rm1021 genome sequence, using a file in which the three DNA replicons (NC_003037, NC_003047, and NC_003078) were combined as one. Differentially regulated genes were identified as those with an adjusted *P*-value <0.01, as determined by the Rockhopper v2.03 software, a fold change of at least three based on the average of the normalized read counts, and less than a twofold difference between replicates of the same sample. A *de novo* prediction of noncoding RNAs (ncRNAs) was performed by Rockhopper, and the same thresholds were applied as for the identification of differentially regulated coding regions. Additionally, the ncRNAs were only considered to be true ncRNAs if they overlapped those predicted previously (Sallet *et al.* 2013). Heatmaps were generated with the heatmap.2 function in the gplots package of R (Warnes *et al.* 2016), and the clustering performed using average linkage with a Pearson correlation distance. The output of the Rockhopper analysis is provided as [File S2](#), and the raw and processed RNA-seq data are available through the Gene Expression Omnibus (accession number GSE100765).

Purification of PckR

The *S. meliloti* pckR (smc02975) open reading frame of wild-type *S. meliloti* RmP110 was PCR amplified (oligos: DF116/DF117) and cloned into *Nde*I-digested pTH2976, producing pTH3150, for expression of an N-terminal, H₆-tagged PckR protein. Plasmid pTH3150 was transformed into *E. coli* BL21 (DE3) pLysS for overexpression of the PckR protein. The PckR protein was overexpressed and purified using standard protocols, and these are described in detail in the Supplemental Methods in [File S1](#).

Electrophoretic mobility shift assays

Promoter regions were PCR amplified from *S. meliloti* genomic DNA for use as probes in electrophoretic mobility shift assays (EMSAs; oligos: *zwf* – DF118/DF119, *pckA* – DF120/DF121, *fbaB* – DF122/DF123, *eda2* – DF124/DF125, *mgsA* – DF126/DF127). Each PCR product was between 135 and 200 bp in length, and were centered around the predicted PckR binding motifs. Purified PCR products were 5' end-labeled with T4 polynucleotide kinase (New England Biolabs) using 0.5 pmol of probe and 1 pmol of [γ -³²P] ATP (Perkin Elmer), and purified. Binding reactions and electrophoresis were performed largely as described before (MacLean *et al.* 2008), and described in detail in the Supplemental Methods in [File S1](#). Where applicable, PEP or pyruvate was added to the binding reactions to a final concentration as indicated, or double-stranded oligonucleotides were added to a molar ratio (relative to probe) of 1000:1.

Identification of the PckR binding motif

The 200-nt upstream of the translational start codon of the first gene of each PckR regulated operon were collected, and motifs were identified using the MEME motif discovery webserver (Bailey and Elkan 1994; Bailey *et al.* 2009) with default settings and the following selections: normal mode,

standard DNA alphabet, any number of motif repetitions, and the identification of five motifs. The top hit was 15 nt in length, and manual inspection of the surrounding region led to the enlargement of this motif to 16 nt. The logo of this motif was created with skylign (Wheeler *et al.* 2014) based on the sense strand sequence of the seven copies of the motif, using the “observed counts,” “full length alignment,” and “information content – above background” options.

Modeling the structure of PckR

The tertiary structure of the PckR protein was modeled using the PckR amino acid sequence from *S. meliloti* Rm1021, and the Phyre2 webserver (Kelley *et al.* 2015) with the intensive mode. Phyre2 chose the following six crystal structures (identified according to their Protein Data Base accession) as templates for the modeling, all of which have 20–30% amino acid identity with the *S. meliloti* PckR protein: 1BDH, 3KJX, 1EFA, 3H5T, 2LBG, and 1ZVV. The output from Phyre2 was submitted to the 3DLigandSite webserver (Wass *et al.* 2010) for prediction of ligand binding domains. The tertiary structure of PckR was visualized with Chimera (Pettersen *et al.* 2004).

Blast bidirectional best hit analysis

The Blast bidirectional best hit (Blast-BBH) analysis was performed using an automated pipeline based upon the workflow described previously (Zamani *et al.* 2017), and this process is detailed further in the Supplemental Methods in [File S1](#). All genomes of the order Rhizobiales available through the NCBI Genome database (accessed March 21, 2016) that were annotated as complete were included in this analysis. The PckR Blast-BBHs were aligned with MAFFT-linsi (Katoh and Standley 2013), trimmed with trimAl (Capella-Gutiérrez *et al.* 2009), and a maximum likelihood phylogeny built with the RAXML BlackBox mirror (Stamatakis 2014) on the CIPRES Science Gateway webserver (Miller *et al.* 2010), with a CAT rate heterogeneity and LG amino acid substitution model.

Multilocus sequence analysis

The multilocus sequence analysis was performed using an automated pipeline centered on the use of AMPHORA2 (Wu and Scott 2012), based upon the workflow described previously (Zamani *et al.* 2017). One representative genome for each Rhizobiales species, as well as several outgroup taxa from the α -proteobacteria, were used in this analysis. The phylogeny was built based on 20 proteins present in all proteomes (Frr, RplA, RplB, RplE, RplF, RplK, RplM, RplN, RplP, RplS, RplT, RpoB, RpsB, RpsC, RpsE, RpsI, RpsJ, RpsK, RpsM, and RpsS). Each set of proteins were individually aligned with MAFFT-linsi (Katoh and Standley 2013) and trimmed with trimAl (Capella-Gutiérrez *et al.* 2009). The trimmed alignments were concatenated, and a phylogeny built with the RAXML BlackBox mirror (Stamatakis 2014) on the CIPRES Science Gateway webserver (Miller *et al.* 2010), with a CAT rate heterogeneity and LG amino acid substitution model.

LacI phylogenetic analysis

All proteins of the ENOG4105ETE and ENOG410XPSF orthologous protein groups of the eggNOG database (Huerta-Cepas *et al.* 2016) were downloaded, yielding 4607 proteins. The proteins were aligned with MUSCLE (Edgar 2004) on the CIPRES Science Gateway webserver (Miller *et al.* 2010), with MUSCLE limited to two iterations. The alignment was trimmed with trimAl (Capella-Gutiérrez *et al.* 2009), and an approximate maximum likelihood tree build with FastTree v2 (Price *et al.* 2010). All phylogenetic trees produced in this study were visualized with FigTree v1.4.2.

Data availability

File S1 contains the Supplemental Methods, Figures S1–S7, Tables S1–S3, and the Supplemental References. File S2 contains the output of the Rockhopper analysis of the RNA-seq data. Raw and processed RNA-seq data are available through the Gene Expression Omnibus (accession number GSE100765).

Results

PckR regulates expression of five central carbon metabolism operons

To investigate the role of *pckR* in gene regulation, we examined the transcriptional profile of *pckR*⁺ and *pckR*[−] cells grown with a glycolytic substrate, glucose, or with a gluconeogenic substrate (succinate). Based on RNA-seq analysis, 40 genes are differentially expressed greater than threefold in the wild-type *pckR*⁺ strain RmG212; 11 genes display higher expression with succinate whereas 29 genes display lower expression with succinate (Figure 2A). Of these 40 genes, the differential expression of five operons is not observed in the *pckR* null mutant strain, RmK124. Specifically, the increased transcription of *pckA* and *fbaB*, and the decreased transcription of *zwf-pgl-edd*, *eda2*, and *mgsA*, during growth with succinate appears to be *pckR* dependent (Figure 2A). The phenotypes of the *pckR*::ΩSp^R mutant are not due to polar effects of the insertion, as expression of the downstream gene *smc04662* appeared unaffected (RNA-seq data in File S2), and a *smc04662* mutant has none of the phenotypes of the *pckR* mutant (data not shown).

To independently confirm the role of PckR in the regulation of these five transcripts, the transcription of each was monitored in both *pckR*⁺ and *pckR*[−] strains grown with succinate or glucose using chromosomal *gfp-lacZ* reporter constructs. The β-galactosidase measurements confirm the transcriptional profiles seen in the RNA-seq data (Figure 2B). Consistent with the transcriptional data, the *pckR* null mutant displays impaired growth with gluconeogenic substrates (succinate, arabinose, and pyruvate), but not with glycolytic substrates (glucose, fructose, glycerol, and galactose) or those entering the PP pathway (xylose and ribose) (Table S3 in File S1). Together, these data suggest that PckR functions as a positive regulator of the *pckA* and *fbaB* genes involved in gluconeogen-

Figure 2 Identification of the PckR regulon. Wild-type *S. meliloti* (*pckR*⁺) and a *pckR* insertion mutant (*pckR*[−]) were grown with either glucose (Glu.) or succinate (Suc.) as the sole carbon source. (A) A heatmap showing the expression of all genes that displayed a statistically significant (adjusted *P*-value <0.01, fold change >3) difference in expression level in wild-type cells grown with succinate vs. glucose, as determined through RNA-seq. Expression values are shown as the log₂ of the relative change compared to the average of the *pckR*⁺ strain grown with glucose, and the results for each biological duplicate are presented. Hierarchical clustering analysis, shown on the left, identified four primary groupings, as indicated on the right. (B) The expression of the five genes identified as putatively regulated by PckR through RNA-seq were independently validated using *gfp-lacZ* reporter gene constructs by measuring the β-galactosidase (LacZ) activity of stationary phase cultures. Data points represent the averages of triplicate samples, with error bars indicating the standard deviation.

esis, and as a negative regulator of the *zwf-pgl-edd*, *eda2*, and *mgsA* genes involved in glycolysis (Figure 1).

Interestingly, three ncRNAs were identified as differentially expressed between growth with glucose and succinate (*smc06497*, *smc06705* and *sma6570*) in a PckR-independent

Figure 3 Binding of PckR to the promoter regions of PckR regulated genes. (A) Sequences of the DNA regions upstream of each operon belonging to the PckR regulon. Predicted transcriptional start sites (TSS) are indicated by boldface font and “+1.” The location of the putative PckR binding sites, as identified with the help of MEME (Bailey and Elkan 1994; Bailey *et al.* 2009), are in blue font and underlined, and the position of the 3’ end relative to the TSS is indicated. The location of the ZW and NS probes used in D are indicated by the red underlines. The location of the –10 and –35 regions of the *zwf* promoter are indicated by the gray lines above the sequence, and were identified based on the *S. meliloti* consensus promoter sequence (MacLellan *et al.* 2006). (B) A logo and the consensus sequence of the PckR DNA binding sequence, based on the seven predicted motifs, is shown. The logo includes the two positions upstream (–1, –2) and downstream (+1, +2) of the binding site, and the location of the mutations in the oligonucleotides used in (D) are shown. (C) Electrophoretic mobility shift assays (EMSAs) were performed using purified PckR protein and ~135–200-bp ³²P-labeled probes that each included the promoter region of one of the PckR regulated promoters. (D) EMSAs are shown using a probe spanning the *zwf* promoter region and with an additional 25-bp oligonucleotide as indicated. The molar ratio of the oligonucleotide relative to the probe is 1000:1. Probes M1–M4 are ZW derivatives containing T to G, A to C, C to A, or T to C mutations, respectively, at the positions indicated in (B).

manner. In the *pckR*⁺ strain RmG212, the ncRNAs *smc06497*, *smc06705*, and *smc6570* were upregulated ~3.8-, 5.2-, and 4.5-fold during growth with succinate relative to glucose, respectively. These changes were not independently confirmed or pursued further.

PckR directly modulates expression through binding an imperfect palindromic sequence

PckR is homologous to the GalR/LacI family of transcriptional regulators, and we previously noted (Østerås *et al.* 1997) that the *pckA* promoter contains a region from 83 to 70 nt upstream of the transcriptional start site (5’-TTAAATCGATTAAAT-3’) that fits well with the GalR/LacI consensus palindromic binding site (5’-NNNAANCGNTTNNN-3’) (Weickert and Adhya 1992) (Figure 3A). To identify a putative PckR binding site, the five promoter regions whose expression was influenced by PckR were searched for conserved motifs using the MEME software (Bailey and Elkan 1994; Bailey *et al.* 2009). This analysis iden-

tified a conserved 16-nt motif (Figure 3, A and B) that matches the previously identified motif in the *pckA* promoter (Østerås *et al.* 1997), and has a consensus sequence of TTTMAATC GATTWAAA (Figure 3B). One copy of this motif is present in each of the negatively regulated promoters (*zwf*, *eda2*, and *mgsA*), whereas two copies of this motif are present in the positively regulated promoters (*pckA* and *fbaB*) (Figure 3A).

To investigate if PckR could bind the identified motif, EMSAs were used to detect binding of purified PckR protein to labeled DNA fragments. PckR bound all five of the promoter fragments *in vitro* (Figure 3C). In competitive binding assays, the addition of a 25-bp oligonucleotide (termed ZW), which included the predicted PckR binding site of the *zwf* promoter, abolished the shift of the *zwf* promoter probe (Figure 3D). Similar results were obtained if the oligonucleotide included either of the predicted PckR binding sites of the *pckA* promoter (data not shown). In contrast, a control oligonucleotide (termed NS), lacking the predicted PckR binding site,

had no effect on the shift of the *zwf* promoter probe (Figure 3D). Four variants of the ZW oligonucleotide (termed M1–M4) that each had a single residue change were tested. The three oligonucleotides (M2, M3, and M4; Figure 3B) with mutations of conserved residues of the predicted PckR binding site had little effect on the binding of PckR to the *zwf* probe, whereas the M1 oligonucleotide with a mutation outside of the conserved region abolished the shift with the *zwf* promoter probe (Figure 3D). Together, these results strongly suggest that PckR directly regulates expression of the five operons identified above by binding to a palindromic sequence (consensus: TTTMAATCGATTWAAA) in the promoter regions.

Comparison of the location of the PckR binding sites with the predicted location of the promoter motifs, based on transcriptional start site mapping studies (Østerås *et al.* 1995; Sallet *et al.* 2013; Schlüter *et al.* 2013), indicates that the locations of the binding sites differ for positively and negatively regulated promoters. Both of the PckR binding sites in the positively regulated genes are located upstream of the promoter motifs. In contrast, the PckR binding sites of the negatively regulated genes overlap the promoter motif or the transcriptional start site. Together, these data indicate that PckR directly modulates expression of the PckR regulon, with positive vs. negative regulation dependent on the position and number of the PckR binding motifs.

PEP may function as a PckR effector metabolite

As is the case for other LacI/GalR regulatory proteins, we expected an effector molecule to interact with the C-terminal domain of PckR to modulate its DNA binding activity. As we hypothesized the effector molecule was a metabolite of central carbon metabolism, the effect of disrupting central carbon metabolism on the activity of PckR was examined by monitoring *pckA* expression in various carbon metabolic mutants. Mutation of *pckA* results in a strong induction of *pckA* expression in medium containing both glucose and succinate, whereas expression of *pckA* is low in strains carrying wild-type *pckA* when grown in the same medium (Figure 4, A and C). The induction of *pckA* in the *pckA* mutant is dependent on the growth medium, as *pckA* expression in the mutant is relatively low when provided with galactose as a carbon source (Figure 4, B and C). The catabolism of galactose by *S. meliloti* results in the production of G3P (glycolytic intermediate) and pyruvate (gluconeogenic), and is therefore similar to growth with glucose plus succinate, but without the effects of catabolite repression (Figure 1) (Geddes and Oresnik 2012).

One hypothesis for the induction of *pckA* in a *pckA* mutant is that it is a consequence of the accumulation of TCA cycle intermediates during the catabolism of succinate (Figure 1). It was previously shown that TCA cycle intermediates build up in a *S. meliloti dme tme* malic enzyme double mutant when grown in the presence of succinate (Zhang *et al.* 2016). However, *pckA* transcription is not induced in a *dme tme* mutant grown with both glucose and succinate (Figure 4, A and C), inconsistent with *pckA* induction being dependent on a TCA cycle intermediate.

Figure 4 Effect of mutation of central carbon metabolic genes on *pckA* expression. Growth (closed symbols) and *pckA* expression (open symbols) of *S. meliloti* strains containing a *pckA*::*gfp-lacZ* fusion as well as the indicated null mutation(s). The GFP fluorescence units (RFU) shown in the graphs are not standardized by OD₆₀₀, and data points represent the mean of triplicate samples. Strains were grown in minimal medium with either (A) 15 mM glucose and 15 mM succinate or (B) 15 mM galactose as carbon sources. The diauxic growth in glucose and succinate is not well represented due to the plotting of data points only once every 2 hr. (C) A key indicating which genotypes are represented by each symbol. The key also shows the standardized GFP fluorescence (RFU/OD₆₀₀) in the stationary phase, calculated as the mean of the final two time points, when grown in either glucose and succinate (Glu/Suc) or in galactose (Gal).

A second hypothesis is that the observed *pckA* induction is due to a reduced concentration of a metabolite whose synthesis depends on the activity of PckA. We therefore examined *pckA* expression in several mutants with lesions of the EMP pathway (Figure 1). Mutations of *eno* (enolase),

Figure 5 Effect of PEP and pyruvate on the DNA binding activity of PckR. Electrophoretic mobility shift assays were performed using purified PckR protein and a ^{32}P -labeled probe that included the *zwf* promoter region. Phosphoenolpyruvate or pyruvate were added to each reaction mixture at the indicated concentrations (in millimolar).

pgk (phosphoglycerate kinase), and *gap* (glyceraldehyde-3-phosphate dehydrogenase) do not induce strong *pckA* expression when grown with glucose and succinate (Figure 4, A and C). This suggests that the compound of interest is located between PckA and Eno in the gluconeogenic pathway (Figure 1). As PEP is the only metabolite matching this description, we tested whether PEP would inhibit the binding of PckR to the *zwf* promoter using EMSAs. Inclusion of 2 mM PEP in the EMSA binding reactions results in a reduction in the amount of *zwf* probe that is shifted, and an 80% reduction occurs in the presence of 4 mM PEP relative to 0 mM PEP (Figure 5). This indicates that PEP inhibits the DNA binding activity of PckR *in vitro*. Because intracellular PEP concentrations in the millimolar range are detected in *E. coli* (Hogema *et al.* 1998; Hoque *et al.* 2005; Xu *et al.* 2012) and *S. meliloti* (A. Checcucci and A. Mengoni, personal communication), our data are consistent with PEP influencing the ability of PckR to regulate its target promoters at physiologically relevant concentrations. In contrast, pyruvate at concentrations up to 20 mM has no effect on the ability of PckR to bind the *zwf* promoter (Figure 5). Although this suggests that the effect of PEP is specific, consistent with PEP being the PckR effector molecule, it is possible that the effects of PEP occur through nonspecific ionic interactions via the phosphate group. Further *in vitro* experiments are required to conclusively demonstrate direct binding of PEP or another effector to the sensory domain of purified PckR.

The F230, A289, and A304 residues of PckR influence effector binding

Modeling of PckR against known LacI/GalR family crystal structures using Phyre2 (Kelley *et al.* 2015) reveals prototypical N-terminal DNA binding and C-terminal sensory domains (Figure 6). The C-terminal tetramerization domain present in a subset of LacI/GalR family members, such as LacI (Lewis *et al.* 1996), is not seen. Østerås *et al.* (1997) identified three spontaneous *pckR* mutations, *rpk-9*, *rpk-10*, and *rpk-15*, which altered the regulation of *pckA* expression. We renamed

Figure 6 Predicted structure of the *S. meliloti* PckR protein. The tertiary structure of PckR was predicted using the Phyre2 webserver (Kelley *et al.* 2015) and visualized with Chimera (Pettersen *et al.* 2004). The N-terminal DNA binding domain and the C-terminal sensory domain are indicated, as is the predicted ligand binding domain as determined via the 3DLigand-Site webserver (Wass *et al.* 2010). The three amino acids (F230, A289, and A304) changed in the *pckR* mutants are indicated in red, and are all located in the sensory domain away from the ligand binding site.

these alleles as *pckR9*, *pckR10*, and *pckR15*, respectively, and genome sequencing identified single mutations within *pckR* in each mutant that resulted in F230C, A289T, and A304V amino acid substitutions, respectively. All three of these residues are located within the PckR C-terminal sensory domain (Figure 6), which suggests that these substitutions do not directly impact the DNA binding of PckR, but instead reduce its affinity toward an effector.

The growth and *pckA* and *zwf* expression phenotypes of the *pckR* mutant strains are consistent with the above conclusion. Strains carrying the *pckR9*, *pckR10*, and *pckR15* alleles grow slowly in minimal medium with glucose, whereas the *pckR::ΩSp* null mutant grows like wild type (Figure 7, A and D). The opposite was observed during growth with succinate: strains carrying the *pckR9*, *pckR10*, and *pckR15* alleles grow like wild type, whereas the *pckR::ΩSp* null mutant displays reduced growth (Figure 7, B and E). Similarly, the *pckA* and *zwf* expression profiles of the null mutant are opposite those of the other alleles. In the null *pckR::ΩSp* mutant, *pckA* expression is constitutively low, whereas *zwf* expression is constitutively high (Figure 7). In contrast, *zwf* expression is constitutively low in strains carrying the *pckR9*, *pckR10*, or *pckR15* alleles, whereas *pckA* expression is strongly induced during growth with succinate and during the stationary phase after growth with glucose (Figure 7). These data suggest that the F230C, A289T, and A304V substitutions result in elevated PckR activity, as opposed to the loss of activity of the *pckR::ΩSp* allele. As the activity of PckR appears inhibited by the binding of an effector, the observed phenotypes are likely a consequence of these substitutions reducing the ability of PckR to associate with its effector molecule.

Figure 7 Growth and gene expression phenotypes of various *pckR* alleles. Growth (closed symbols) and gene expression (open symbols) of *S. meliloti* strains with various *pckR* alleles grown in minimal medium with either (A and D) 15 mM glucose or (B and E) 15 mM succinate as the sole carbon source. The GFP fluorescence units (RFU) shown in the graphs are not standardized by OD₆₀₀, and data points represent the mean of triplicate samples. (A and B) All strains carried a *pckA*::*gfp-lacZ* fusion to monitor expression of *pckA* through GFP fluorescence. (D and E) All strains carried a *zwf*::*gfp-lacZ* fusion to monitor expression of *zwf* through GFP fluorescence. (C and F) Keys indicating which genotype is represented by each symbol; (C) corresponds to graphs (A and B), whereas (F) corresponds to graphs (D and E). The keys also show the standardized GFP fluorescence (RFU/OD₆₀₀) in the stationary phase, calculated as the mean of the final two time points, when grown in either glucose (Glu) or succinate (Suc). The extent of the expression differences between samples are partially obscured due to the high background in the GFP fluorescence readings (Figure S6 and Figure S7 in File S1).

PckR is restricted to one clade of the Rhizobiales order

A Blast-BBH approach, starting with the PckR amino acid sequence of *S. meliloti* Rm1021, was used to identify putative PckR orthologs in species of the order Rhizobiales. Note that this approach simply identifies whether two proteins are more similar to each other than to any other protein in the two proteomes, which means that homologous proteins may have been identified as Blast-BBHs even if they do not share functional orthology.

PckR Blast-BBHs are found in all species within the monophyletic group that includes the families Rhizobiaceae, Phyllobacteriaceae, Brucellaceae, and Bartonellaceae, except for the genera *Liberibacter* and *Bartonella* that are obligate intracellular pathogens with reduced genomes (Alsmark *et al.*

2004; Leonard *et al.* 2012) (Figure 8A and Figure S1 in File S1). The three PckR residues (F230, A289, and A305) identified above as important in effector binding affinity are conserved in all of the proteins, and all shared >75% amino acid identity with the *S. meliloti* PckR. These proteins therefore likely represent true PckR orthologs. In contrast, only 7 of the 39 Rhizobiales species outside of the above-mentioned monophyletic group contain a PckR Blast-BBH (Figure 8A). These seven proteins form a monophyletic outgroup in a phylogeny of the PckR Blast-BBHs (Figure S2 in File S1), they share only 30–50% identity with the *S. meliloti* PckR protein, they are quite diverse even from each other, and none contain equivalents of all three of the F230, A289, and A305 residues. Moreover, a plasmid carrying the PckR Blast-BBH from *A. caulinodans* ORS 571 (*azc_4468*) failed to complement the

Figure 8 Phylogenetic analysis of the PckR protein. (A) A RAxML (Stamatakis 2014) maximum likelihood phylogeny based on the concatenated alignments of 20 conserved proteins is shown. The tree was rooted with several α -proteobacteria outgroups, containing one representative genome per species, and was produced as described in the *Materials and Methods*. Taxa are colored based on the presence (blue) or absence (red) of a PckR Blast-BBH. A complete phylogeny containing the outgroups and bootstrap values is provided as Figure S1 in File S1. (B) A subtree of a 4607-member LacI family protein phylogeny produced with FastTree (Price et al. 2010). Proteins in which the F230, A289, and A304 residues are conserved are indicated in blue. The complete LacI family protein phylogeny is provided as Figure S4 in File S1.

pckA and *zwf* expression phenotypes or the succinate growth phenotype of the *S. meliloti* *pckR*:: Ω Sp^R null mutant (Figure S3 in File S1). We therefore conclude that none of these seven proteins represented true orthologs of PckR, and that PckR is specific to, and highly conserved within, one multi-family clade of the Rhizobiales.

In a phylogeny of 4607 LacI family proteins (Figure 8B and Figure S4 in File S1), all of the predicted PckR proteins from the order Rhizobiales form a monophyletic group, which indicates that *pckR* has not been transferred and maintained

outside of this clade. If PckR originated in a distinct bacterial lineage and was then transferred to an ancestral Rhizobiales species, the Rhizobiales PckR group should be nested within a group of proteins from a distinct bacterial taxon; however, this is not observed. Moreover, the two proteins most closely related to the predicted PckR proteins are from the species *Fulvmarina pelagi* and *Aurantimonas manganoxydans*. These species form a deep branching outgroup lineage within the Rhizobiales clade containing the *Rhizobiaceae*, *Phyllobacteriaceae*, *Brucellaceae*, and *Bartonellaceae* families (Cho and

Giovannoni 2003; Denner *et al.* 2003), but only one of the three residues substituted in the *pckR* mutants is conserved in these proteins. These observations are consistent with an early Rhizobiales organism obtaining a gene encoding an LacI family protein, and that this protein subsequently gained the PckR function.

Discussion

The PckR transcriptional regulator is found in a subset of the Rhizobiales (Figure 8). PckR regulates expression of central carbon metabolic genes whose products include the entire ED pathway, the committed step of gluconeogenesis (PckA), and FbaB that functions at a crossroads of gluconeogenesis and glycolysis (Figure 1 and Figure 2). We propose a model of PckR-based regulation in which catabolism of glycolytic compounds results in elevated concentrations of an effector molecule that associates with PckR and reduces the affinity of PckR for its DNA binding motif (Figure 9). When concentrations of the effector drop sufficiently, such as during catabolism of gluconeogenic substrates, PckR dissociates from its effector and binds its cognate DNA motif. This results in repression of transcription if the binding site overlaps the promoter motif or transcriptional start site, or induction of transcription if there are two binding sites located upstream of the promoter motifs. The *pckR9*, *pckR10*, and *pckR15* alleles displayed elevated, but not constitutive, expression of *pckA* (Figure 7). This suggests that the PckR proteins in these mutants have reduced affinity for its effector and that the F230, A289, and A305 amino acid residues of PckR are important in increasing the affinity of PckR to its effector. The combined genetic and *in vitro* data presented here (Figure 2, Figure 3, Figure 4, and Figure 5) is consistent with the PckR effector molecule being PEP, although additional studies are required.

We also attempted to identify additional PckR binding sites in promoters that were not identified as part of the PckR regulon. Using both HMMER (Eddy 2009) and PatScan (Dsouza *et al.* 1997) searches, we found only one additional sequence that was almost certainly a PckR binding motif; this motif was upstream of the *feuNPQ* locus, a signal transduction system responsive to osmotic conditions and involved in regulating export of cyclic β -glucans (Griffitts *et al.* 2008; Carlyon *et al.* 2010). However, the location of the motif relative to the transcriptional start site suggests that PckR binding would not impact transcription. Several other sequences that shared resemblance with the PckR consensus binding motif are present in the genome; however, they all contained enough mismatches that it was likely that they were nonfunctional. Nevertheless, it is notable that several of the genes with these sequences are related to carbon metabolism, including *glgP*, *thuE*, *pckR*, and *sdhC*. We hypothesize that at least some of these genes may eventually become integrated into the PckR regulon.

Given the location of the binding sites in the negatively regulated promoters (*zwf-pgl-edd*, *eda2*, and *mgsA* genes),

Figure 9 Proposed model of PckR activity. A schematic illustrating the proposed model of PckR activity is shown using *pckA* and *zwf* as example genes positively and negatively regulated by PckR, respectively. (A) When intracellular concentrations of the effector are high, such as during growth with glucose, PckR is predominately associated with the effector and disassociated from its target DNA binding sites. This results in high transcription of the negatively regulated promoters, and low expression of the positively regulated promoters. (B) When intracellular concentrations of the effector are low, such as during growth with succinate, PckR is predominately not associated with the effector and is bound to its target DNA binding sites. This results in low transcription of the negatively regulated promoters, and high expression of the positively regulated promoters.

PckR binding presumably impairs the binding or progression of the RNA polymerase, thereby decreasing expression. However, it is less clear how PckR induces expression of the positively regulated promoters of *pckA* and *fbaB* (Figure 2). We note that in the absence of PckR, there is sufficient *pckA* expression to allow growth with succinate as the carbon source (Figure 7B). Although this growth is reduced relative to the wild type, a *pckA* null mutant fails to grow with succinate as the sole carbon source (Østerås *et al.* 1995, 1997) (Figure S5 in File S1). One possibility is that PckR directly recruits the RNA polymerase to the target promoters. A second explanation is that PckR indirectly promotes RNA polymerase recruitment by displacing a negative transcriptional regulator. In support of the latter, a *pckA* promoter truncated at an *EcoRI* cut site (Figure 3A) that lacks the PckR binding sites is constitutively active (Østerås *et al.* 1997), even though induction of the full-length *pckA* promoter is dependent on PckR binding (Figure 2). These contradicting observations can be reconciled by theorizing that PckR does not recruit the RNA polymerase to the promoter, but instead outcompetes a negative transcriptional regulator for promoter occupancy, thereby indirectly providing RNA polymerase access to the promoter. Although most LacI/GalR family regulators repress transcription, several are known to act as

repressors and activators and these include the carbon catabolite protein A (CcpA) of low G + C content Gram-positive bacteria (Schumacher *et al.* 2007), and the cytidine repressor protein (CytR) that regulates transcription of genes in *E. coli* and *Vibrio cholerae* (Rasmussen *et al.* 1996; Swint-Kruse and Matthews 2009; Watve *et al.* 2015). Both CcpA and CytR interact with effector ligands, DNA, and other regulatory proteins, and further studies are necessary to determine whether PckR interacts with other regulatory proteins at the *pckA* promoter.

The possibility of PEP modulating PckR activity is interesting for several reasons. PEP is the product of the PCK reaction, which is the committed step of gluconeogenesis. PEP allosterically modulates the activity of several proteins that function early in glycolysis, including phosphofructokinase (Ogawa *et al.* 2007), the committed and rate-limiting step of glycolysis in *E. coli*; however, *S. meliloti* lacks the phosphofructokinase enzyme. Phosphorylation of the proteins of the phosphotransferase system, a global regulatory pathway modulating secondary carbon metabolism in many prokaryotic species, is also influenced by the PEP/pyruvate ratio (Hogema *et al.* 1998). This is noteworthy as genetic evidence has implicated proteins of the *S. meliloti* phosphotransferase system in succinate-mediated catabolite repression (Pinedo *et al.* 2008; Pinedo and Gage 2009). PEP may therefore provide a link between succinate-mediated catabolite repression-based and PckR-based transcriptional regulation that could help coordinate these responses.

Although α -proteobacteria are a metabolically versatile group of biotechnologically important bacteria (Biondi *et al.* 2009; Imam *et al.* 2013), there has been little study of the regulation of central carbon metabolism in these organisms (Imam *et al.* 2015). Recent works suggest that the regulatory pathways in these organisms may be diverse and evolutionarily independent. PckR is specific to a clade of the order Rhizobiales (Figure 8), CceR regulates central carbon and energy metabolism in the order Rhodobacterales (Imam *et al.* 2015), and *in silico* predictions suggest that GluR regulates central carbon metabolism in the order Caulobacterales (Novichkov *et al.* 2013). PckR, CceR, and GluR are LacI-type transcriptional regulators, but all appear to have evolved independently based on a phylogenetic reconstruction of 4607 LacI proteins (Figure S4 in File S1). Experimental analyses indicate that the PckR and the CceR regulons overlap, yet these proteins recognize distinct DNA binding motifs and different metabolites appear to function as their effector (Figure 3 and Figure 7; Imam *et al.* 2015). How and whether the differences between these transcription factors influence their response changes in nutrient availability is currently unclear, but we hypothesize that differences would be most evident during growth in the presence of multiple carbon substrates.

Acknowledgments

This work was supported by the National Science and Engineering Council of Canada (NSERC) through grants to T.M.F. and an NSERC Alexander Graham Bell Canada Graduate Scholarship-Doctoral (CGS-D) award to G.C.d.

Literature Cited

- Alsmark, C. M., A. C. Frank, E. O. Karlberg, B.-A. Legault, D. H. Ardell *et al.*, 2004 The louse-borne human pathogen *Bartonella quintana* is a genomic derivative of the zoonotic agent *Bartonella henselae*. *Proc. Natl. Acad. Sci. USA* 101: 9716–9721.
- Bailey, T. L., and C. Elkan, 1994 Fitting a mixture model by expectation maximization to discover motifs in biopolymers. *Proc. Int. Conf. Intell. Syst. Mol. Biol.* 2: 28–36.
- Bailey, T. L., M. Boden, F. A. Buske, M. Frith, C. E. Grant *et al.*, 2009 MEME SUITE: tools for motif discovery and searching. *Nucleic Acids Res.* 37: W202–W208.
- Biondi, E. G., E. Tatti, D. Comparini, E. Giuntini, S. Mocali *et al.*, 2009 Metabolic capacity of *Sinorhizobium (Ensifer) meliloti* strains as determined by phenotype MicroArray analysis. *Appl. Environ. Microbiol.* 75: 5396–5404.
- Blondelet-Rouault, M. H., J. Weiser, A. Lebrihi, P. Branny, and J. L. Pernodet, 1997 Antibiotic resistance gene cassettes derived from the omega interposon for use in *E. coli* and *Streptomyces*. *Gene* 190: 315–317.
- Bolton, E., B. Higgison, A. Harrington, and F. O’Gara, 1986 Dicarboxylic acid transport in *Rhizobium meliloti*: isolation of mutants and cloning of dicarboxylic acid transport genes. *Arch. Microbiol.* 144: 142–146.
- Bringhurst, R. M., and D. J. Gage, 2002 Control of inducer accumulation plays a key role in succinate-mediated catabolite repression in *Sinorhizobium meliloti*. *J. Bacteriol.* 184: 5385–5392.
- Capella-Gutiérrez, S., J. M. Silla-Martínez, and T. Gabaldón, 2009 trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. *Bioinformatics* 25: 1972–1973.
- Carlyon, R. E., J. L. Ryther, R. D. VanYperen, and J. S. Griffiths, 2010 FeuN, a novel modulator of two-component signalling identified in *Sinorhizobium meliloti*. *Mol. Microbiol.* 77: 170–182.
- Cho, J.-C., and S. J. Giovannoni, 2003 *Fulvimarina pelagi* gen. nov., sp. nov., a marine bacterium that forms a deep evolutionary lineage of descent in the order “Rhizobiales”. *Int. J. Syst. Evol. Microbiol.* 53: 1853–1859.
- Cowie, A., J. Cheng, C. D. Sibley, Y. Fong, R. Zaheer *et al.*, 2006 An integrated approach to functional genomics: construction of a novel reporter gene fusion library for *Sinorhizobium meliloti*. *Appl. Environ. Microbiol.* 72: 7156–7167.
- Denner, E. B. M., G. W. Smith, H.-J. Busse, P. Schumann, T. Narzt *et al.*, 2003 *Aurantimonas coralicida* gen. nov., sp. nov., the causative agent of white plague type II on Caribbean scleractinian corals. *Int. J. Syst. Evol. Microbiol.* 53: 1115–1122.
- diCenzo, G. C., A. M. MacLean, B. Milunovic, G. B. Golding, and T. M. Finan, 2014 Examination of prokaryotic multipartite genome evolution through experimental genome reduction. *PLoS Genet.* 10: e1004742.
- diCenzo, G. C., A. Checcucci, M. Bazzicalupo, A. Mengoni, C. Viti *et al.*, 2016a Metabolic modelling reveals the specialization of secondary replicons for niche adaptation in *Sinorhizobium meliloti*. *Nat. Commun.* 7: 12219.
- diCenzo, G. C., M. Zamani, B. Milunovic, and T. M. Finan, 2016b Genomic resources for identification of the minimal N₂-fixing symbiotic genome. *Environ. Microbiol.* 18: 2534–2547.
- diCenzo, G. C., H. Sharthiya, A. Nanda, M. Zamani, T. M. Finan, 2017 PhoU allows rapid adaptation to high phosphate concentrations by modulating PstSCAB transport rate in *Sinorhizobium meliloti*. *J. Bacteriol.* 199: e00143–17.
- Driscoll, B., and T. M. Finan, 1993 NAD⁺-dependent malic enzyme of *Rhizobium meliloti* is required for symbiotic nitrogen fixation. *Mol. Microbiol.* 7: 865–873.
- Driscoll, B., and T. M. Finan, 1996 NADP⁺-dependent malic enzyme of *Rhizobium meliloti*. *J. Bacteriol.* 178: 2224–2231.
- Driscoll, B. T., and T. M. Finan, 1997 Properties of NAD⁺- and NADP⁺-dependent malic enzymes of *Rhizobium (Sinorhizobium)*

- meliloti* and differential expression of their genes in nitrogen-fixing bacteroids. *Microbiology* 143: 489–498.
- Dsouza, M., N. Larsen, and R. Overbeek, 1997 Searching for patterns in genomic data. *Trends Genet.* 13: 497–498.
- Eddy, S. R., 2009 A new generation of homology search tools based on probabilistic inference. *Genome Inform* 23: 205–211.
- Edgar, R. C., 2004 MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.* 32: 1792–1797.
- Finan, T. M., J. M. Wood, and D. C. Jordan, 1983 Symbiotic properties of C₄-dicarboxylic acid transport mutants of *Rhizobium leguminosarum*. *J. Bacteriol.* 154: 1403–1413.
- Finan, T. M., E. Hartweig, K. LeMieux, K. Bergman, G. C. Walker *et al.*, 1984 General transduction in *Rhizobium meliloti*. *J. Bacteriol.* 159: 120–124.
- Finan, T. M., I. J. Oresnik, and A. Bottacin, 1988 Mutants of *Rhizobium meliloti* defective in succinate metabolism. *J. Bacteriol.* 170: 3396–3403.
- Fuhrer, T., E. Fischer, and U. Sauer, 2005 Experimental identification and quantification of glucose metabolism in seven bacterial species. *J. Bacteriol.* 187: 1581–1590.
- Galandini, M., M. Brilli, G. Spini, M. Rossi, B. Roncaglia *et al.*, 2015 Evolution of intra-specific regulatory networks in a multipartite bacterial genome. *PLoS Comput. Biol.* 11: e1004478.
- Galibert, F., T. M. Finan, S. R. Long, A. Pühler, A. P. Abola *et al.*, 2001 The composite genome of the legume symbiont *Sinorhizobium meliloti*. *Science* 293: 668–672.
- Geddes, B. A., and I. J. Oresnik, 2012 Inability to catabolize galactose leads to increased ability to compete for nodule occupancy in *Sinorhizobium meliloti*. *J. Bacteriol.* 194: 5044–5053.
- Geddes, B. A., and I. J. Oresnik, 2014 Physiology, genetics, and biochemistry of carbon metabolism in the alphaproteobacterium *Sinorhizobium meliloti*. *Can. J. Microbiol.* 60: 491–507.
- Griffitts, J. S., R. E. Carlyon, J. H. Erickson, J. L. Moulton, M. J. Barnett *et al.*, 2008 A *Sinorhizobium meliloti* osmosensory two-component system required for cyclic glucan export and symbiosis. *Mol. Microbiol.* 69: 479–490.
- Hogema, B. M., J. C. Arents, R. Bader, K. Eijkemans, H. Yoshida *et al.*, 1998 Inducer exclusion in *Escherichia coli* by non-PTS substrates: the role of the PEP to pyruvate ratio in determining the phosphorylation state of enzyme IIA^{Glc}. *Mol. Microbiol.* 30: 487–498.
- Hoque, M. A., H. Ushiyama, M. Tomita, and K. Shimizu, 2005 Dynamic responses of the intracellular metabolite concentrations of the wild type and *pykA* mutant *Escherichia coli* against pulse addition of glucose or NH₃ under those limiting continuous cultures. *Biochem. Eng. J.* 26: 38–49.
- Huerta-Cepas, J., D. Szklarczyk, K. Forslund, H. Cook, D. Heller *et al.*, 2016 eggNOG 4.5: a hierarchical orthology framework with improved functional annotations for eukaryotic, prokaryotic and viral sequences. *Nucleic Acids Res.* 44: D286–D293.
- Imam, S., D. R. Noguera, and T. J. Donohue, 2013 Global insights into energetic and metabolic networks in *Rhodobacter sphaeroides*. *BMC Syst. Biol.* 7: 89.
- Imam, S., D. R. Noguera, and T. J. Donohue, 2015 CceR and AkgR regulate central carbon and energy metabolism in alphaproteobacteria. *MBio* 6: e02461–14.
- Jeong, J.-Y., H.-S. Yim, J.-Y. Ryu, H. S. Lee, J.-H. Lee *et al.*, 2012 One-step sequence- and ligation-independent cloning as a rapid and versatile cloning method for functional genomics studies. *Appl. Environ. Microbiol.* 78: 5440–5443.
- Katoh, K., and D. M. Standley, 2013 MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol. Biol. Evol.* 30: 772–780.
- Kelley, L. A., S. Mezulis, C. M. Yates, M. N. Wass, and M. J. E. Sternberg, 2015 The Phyre2 web portal for protein modeling, prediction and analysis. *Nat. Protoc.* 10: 845–858.
- Leonard, M. T., J. R. Fagen, A. G. Davis-Richardson, M. J. Davis, and E. W. Triplett, 2012 Complete genome sequence of *Liberibacter crescens* BT-1. *Stand. Genomic Sci.* 7: 271–283.
- Lewis, M., G. Chang, N. C. Horton, and M. A. Kercher, 1996 Crystal structure of the lactose operon repressor and its complexes with DNA and inducer. *Science* 271: 1247–1254.
- MacLean, A. M., M. I. Anstey, and T. M. Finan, 2008 Binding site determinants for the LysR-type transcriptional regulator PcaQ in the legume endosymbiont *Sinorhizobium meliloti*. *J. Bacteriol.* 190: 1237–1246.
- MacLellan, S. R., L. A. Smallbone, C. D. Sibley, and T. M. Finan, 2005 The expression of a novel antisense gene mediates incompatibility within the large *repABC* family of α -proteobacterial plasmids. *Mol. Microbiol.* 55: 611–623.
- MacLellan, S. R., A. M. MacLean, and T. M. Finan, 2006 Promoter prediction in the rhizobia. *Microbiology* 152: 1751–1763.
- Miller, M. A., W. Pfeiffer, and T. Schwartz, 2010 Creating the CIPRES science gateway for inference of large phylogenetic trees, pp. 1–8 in *Proceedings of the Gateway Computing Environments Workshop (GCE)*. New Orleans, LA.
- Milunovic, B., G. C. diCenzo, R. A. Morton, and T. M. Finan, 2014 Cell growth inhibition upon deletion of four toxin-antitoxin loci from the megaplasmids of *Sinorhizobium meliloti*. *J. Bacteriol.* 196: 811–824.
- Novichkov, P. S., A. E. Kazakov, D. A. Ravcheev, S. A. Leyn, G. Y. Kovaleva *et al.*, 2013 RegPrecise 3.0—a resource for genome-scale exploration of transcriptional regulation in bacteria. *BMC Genomics* 14: 745.
- Ogawa, T., H. Mori, M. Tomita, and M. Yoshino, 2007 Inhibitory effect of phosphoenolpyruvate on glycolytic enzymes in *Escherichia coli*. *Res. Microbiol.* 158: 159–163.
- Østerås, M., B. T. Driscoll, and T. M. Finan, 1995 Molecular and expression analysis of the *Rhizobium meliloti* phosphoenolpyruvate carboxykinase (*pckA*) gene. *J. Bacteriol.* 177: 1452–1460.
- Østerås, M., S. A. O'Brien, and T. M. Finan, 1997 Genetic analysis of mutations affecting *pckA* regulation in *Rhizobium (Sinorhizobium) meliloti*. *Genetics* 147: 1521–1531.
- Pettersen, E. F., T. D. Goddard, C. C. Huang, G. S. Couch, D. M. Greenblatt *et al.*, 2004 UCSF Chimera—a visualization system for exploratory research and analysis. *J. Comput. Chem.* 25: 1605–1612.
- Pinedo, C. A., and D. J. Gage, 2009 HPrK regulates succinate-mediated catabolite repression in the gram-negative symbiont *Sinorhizobium meliloti*. *J. Bacteriol.* 191: 298–309.
- Pinedo, C. A., R. M. Bringham, and D. J. Gage, 2008 *Sinorhizobium meliloti* mutants lacking phosphotransferase system enzyme HPr or EIIA are altered in diverse processes, including carbon metabolism, cobalt requirements, and succinoglycan production. *J. Bacteriol.* 190: 2947–2956.
- Poysti, N. J., and I. J. Oresnik, 2007 Characterization of *Sinorhizobium meliloti* triose phosphate isomerase genes. *J. Bacteriol.* 189: 3445–3451.
- Price, M. N., P. S. Dehal, and A. P. Arkin, 2010 FastTree 2—approximately maximum-likelihood trees for large alignments. *PLoS One* 5: e9490.
- Rasmussen, P. B., B. Holst, and P. Valentin-Hansen, 1996 Dual-function regulators: the cAMP receptor protein and the CytR regulator can act either to repress or to activate transcription depending on the context. *Proc. Natl. Acad. Sci. USA* 93: 10151–10155.
- Ronson, C. W., P. Lyttleton, and J. G. Robertson, 1981 C₄-dicarboxylate transport mutants of *Rhizobium trifolii* form ineffective nodules on *Trifolium repens*. *Proc. Natl. Acad. Sci. USA* 78: 4284–4288.
- Sallet, E., B. Roux, L. Sauviac, M.-F. Jardinaud, S. Carrère *et al.*, 2013 Next-generation annotation of prokaryotic genomes with EuGene-P: application to *Sinorhizobium meliloti* 2011. *DNA Res.* 20: 339–354.

- Sambrook, J., E. F. Fritsch, and T. Maniatis, 1989 *Molecular Cloning: A Laboratory Manual*. Cold Spring Harbor Laboratory Press, New York.
- Schlüter, J.-P., J. Reinkensmeier, M. J. Barnett, C. Lang, E. Krol *et al.*, 2013 Global mapping of transcription start sites and promoter motifs in the symbiotic α -proteobacterium *Sinorhizobium meliloti* 1021. *BMC Genomics* 14: 156.
- Schumacher, M. A., G. Seidel, W. Hillen, and R. G. Brennan, 2007 Structural mechanism for the fine-tuning of CcpA function by the small molecule effectors glucose 6-phosphate and fructose 1,6-bisphosphate. *J. Mol. Biol.* 368: 1042–1050.
- Stamatakis, A., 2014 RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* 30: 1312–1313.
- Stowers, M. D., 1985 Carbon metabolism in *Rhizobium* species. *Annu. Rev. Microbiol.* 39: 89–108.
- Swint-Kruse, L., and K. S. Matthews, 2009 Allostery in the LacI/GalR family: variations on a theme. *Curr. Opin. Microbiol.* 12: 129–137.
- Tjaden, B., 2015 *De novo* assembly of bacterial transcriptomes from RNA-seq data. *Genome Biol.* 16: 1.
- Ucker, D. S., and E. R. Signer, 1978 Catabolite-repression-like phenomenon in *Rhizobium meliloti*. *J. Bacteriol.* 136: 1197–1200.
- Warnes G. R., B. Bolker, L. Bonebakker, R. Gentleman, W. H. A. Liaw, *et al.*, 2016 gplots: various R programming tools for plotting data. R package version 3.0.1.
- Wass, M. N., L. A. Kelley, and M. J. E. Sternberg, 2010 3DLigand-Site: predicting ligand-binding sites using similar structures. *Nucleic Acids Res.* 38: W469–W473.
- Watve, S. S., J. Thomas, and B. K. Hammer, 2015 CytR is a global positive regulator of competence, type VI secretion, and chitinases in *Vibrio cholerae*. *PLoS One* 10: e0138834.
- Weickert, M. J., and S. Adhya, 1992 A family of bacterial regulators homologous to Gal and Lac repressors. *J. Biol. Chem.* 267: 15869–15874.
- Wheeler, T. J., J. Clements, and R. D. Finn, 2014 Skylign: a tool for creating informative, interactive logos representing sequence alignments and profile hidden Markov models. *BMC Bioinformatics* 15: 1.
- Wu, M., and A. J. Scott, 2012 Phylogenomic analysis of bacterial and archaeal sequences with AMPHORA2. *Bioinformatics* 28: 1033–1034.
- Xu, Y.-F., D. Amador-Noguez, M. L. Reaves, X.-J. Feng, and J. D. Rabinowitz, 2012 Ultrasensitive regulation of anapleurosis via allosteric activation of PEP carboxylase. *Nat. Chem. Biol.* 8: 562–568.
- Yamada, T., I. Letunic, S. Okuda, M. Kanehisa, and P. Bork, 2011 iPath2.0: interactive pathway explorer. *Nucleic Acids Res.* 39: W412–W415.
- Zamani, M., G. C. diCenzo, B. Milunovic, and T. M. Finan, 2017 A putative 3-hydroxyisobutyryl-CoA hydrolase is required for efficient symbiotic nitrogen fixation in *Sinorhizobium meliloti* and *Sinorhizobium fredii* NGR234. *Environ. Microbiol.* 19: 218–236.
- Zhang, Y., L. A. Smallbone, G. C. diCenzo, R. Morton, and T. M. Finan, 2016 Loss of malic enzymes leads to metabolic imbalance and altered levels of trehalose and putrescine in the bacterium *Sinorhizobium meliloti*. *BMC Microbiol.* 16: 163.

Communicating editor: M. Johnston