

RESEARCH ARTICLE

# Phylogenetic Co-Occurrence of ExoR, ExoS, and ChvI, Components of the RSI Bacterial Invasion Switch, Suggests a Key Adaptive Mechanism Regulating the Transition between Free-Living and Host-Invading Phases in Rhizobiales

Mary Ellen Heavner<sup>1</sup>, Wei-Gang Qiu<sup>2</sup>, Hai-Ping Cheng<sup>1,3\*</sup>

**1** Biochemistry Program, The Graduate Center, City University of New York, New York, New York, United States of America, **2** Biological Sciences Department, Hunter College, City University of New York, New York, New York, United States of America, **3** Biological Sciences Department, Lehman College, City University of New York, Bronx, New York, United States of America

\* [haiping.cheng@lehman.cuny.edu](mailto:haiping.cheng@lehman.cuny.edu)



OPEN ACCESS

**Citation:** Heavner ME, Qiu W-G, Cheng H-P (2015) Phylogenetic Co-Occurrence of ExoR, ExoS, and ChvI, Components of the RSI Bacterial Invasion Switch, Suggests a Key Adaptive Mechanism Regulating the Transition between Free-Living and Host-Invading Phases in Rhizobiales. PLoS ONE 10 (8): e0135655. doi:10.1371/journal.pone.0135655

**Editor:** Peter Mergaert, Centre National de la Recherche, FRANCE

**Received:** April 15, 2015

**Accepted:** July 23, 2015

**Published:** August 26, 2015

**Copyright:** © 2015 Heavner et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Data Availability Statement:** Relevant data is contained within the paper, its Supporting Information files, and NCBI's GenBank database. The relevant accession numbers are provided in the Supporting Information files.

**Funding:** This work was supported by the National Institute of Health (SGM081147 to HPC) and the National Science Foundation (CNS-0958379 and CNS-0855217 to CUNY CSI HPCC and the XSEDE CIPRES project). The funders had no role in study

## Abstract

Both bacterial symbionts and pathogens rely on their host-sensing mechanisms to activate the biosynthetic pathways necessary for their invasion into host cells. The Gram-negative bacterium *Sinorhizobium meliloti* relies on its RSI (ExoR-ExoS-ChvI) Invasion Switch to turn on the production of succinoglycan, an exopolysaccharide required for its host invasion. Recent whole-genome sequencing efforts have uncovered putative components of RSI-like invasion switches in many other symbiotic and pathogenic bacteria. To explore the possibility of the existence of a common invasion switch, we have conducted a phylogenomic survey of orthologous ExoR, ExoS, and ChvI tripartite sets in more than ninety proteobacterial genomes. Our analyses suggest that functional orthologs of the RSI invasion switch co-exist in Rhizobiales, an order characterized by numerous invasive species, but not in the order's close relatives. Phylogenomic analyses and reconstruction of orthologous sets of the three proteins in Alphaproteobacteria confirm Rhizobiales-specific gene synteny and congruent RSI evolutionary histories. Evolutionary analyses further revealed site-specific substitutions correlated specifically to either animal-bacteria or plant-bacteria associations. Lineage restricted conservation of any one specialized gene is in itself an indication of species adaptation. However, the orthologous phylogenetic co-occurrence of all interacting partners within this single signaling pathway strongly suggests that the development of the RSI switch was a key adaptive mechanism. The RSI invasion switch, originally found in *S. meliloti*, is a characteristic of the Rhizobiales, and potentially a conserved crucial activation step that may be targeted to control host invasion by pathogenic bacterial species.

design, data collection and analysis, decision to publish, or preparation of the manuscript.

**Competing Interests:** The authors have declared that no competing interests exist.

## Introduction

The Gram-negative soil bacterium *Sinorhizobium meliloti* Rm1021 fixes nitrogen inside the root nodules produced by its plant host alfalfa, *Medicago sativa* [1,2,3,4], and is one of the best characterized symbiotic models of bacterium-plant interactions [1]. *S. meliloti* shares extensive genomic congruence with animal pathogens such as *Brucella suis*; more than 90% of their genes show at least 98% identity [5]. No known work to date has directly compared the full genomes of *S. meliloti* and the prototypical *Brucella* type, *B. abortus*. However, the high genomic similarity of *B. suis* and *B. abortus* has been conclusively demonstrated [6,7], implying a high similarity between *S. meliloti* and *B. abortus*. In addition, *S. meliloti* also shares a high degree of synteny with plant pathogens such as *Agrobacterium tumefaciens* [8]. These similarities suggest that our understanding of *S. meliloti* might also be a prime source of information about pathogenic bacterial invasion of both mammalian and plant hosts.

Effective *S. meliloti* invasion of alfalfa depends on a series of signal exchanges and reciprocal structural developments [1,2,9,10], and is required for the initiation of *S. meliloti*-alfalfa symbiosis [11]. At the site of the invasion, the infection chamber inside curled alfalfa root hairs, *S. meliloti* cells must produce a bacterial exopolysaccharide, succinoglycan, in order to invade the plant root hair cells [11]. The production of succinoglycan is inversely related to the production of flagella by the same *S. meliloti* cells. During this process, *S. meliloti* cells switch from flagella producing free-living cells to succinoglycan producing host-invading cells [12]. Following *S. meliloti* invasion into root hairs, continuous structural and metabolic modifications occur within both the bacterial and the plant cells, resulting in the formation of root nodules filled with *S. meliloti* bacteroids, which convert atmospheric dinitrogen to ammonia for alfalfa's use as a nitrogen source [2,9].

The critical switch of *S. meliloti* from free-living to host-invading cells is linked to up or down regulation of the expression of hundreds of genes, which are represented by succinoglycan and flagellum genes [13,14]. This switch is controlled by the ExoR, ExoS, and ChvI signaling pathway [15,16], which we refer to herein for the first time as the "RSI invasion switch." ExoR is expressed as a cytoplasmic precursor (ExoR<sub>p</sub>), the structure of which has been resolved through computational modeling [17]. ExoR<sub>p</sub> is secreted to the periplasm as a mature and functional protein (ExoR<sub>m</sub>) that interacts with the periplasmic domain of ExoS (ExoS<sub>p</sub>), a sensor kinase [16,18]. It has been experimentally demonstrated that ExoR<sub>m</sub> can be digested by a yet unidentified periplasmic protease to yield a nonfunctional product ExoR<sub>c20</sub> [16]. Our current model suggests that the proteolytic processing of ExoR<sub>m</sub>, which could be a target of environmental changes or host signals, relieves ExoR inhibition of ExoS, and thus activates the RSI switch [16].

ExoS and ChvI comprise a typical two-component regulatory system (TCS) that belongs to the EnvZ/OmpR family [11,19,20,21]. ExoS is a 595-residue protein with two transmembrane domains. The first transmembrane domain is in close proximity to its N-terminus (residues 48 to 67). A periplasmic domain follows between residues 68 and 278, where interaction with ExoR<sub>m</sub> occurs, and is followed by the second transmembrane domain and a cytoplasmic histidine kinase [11,13]. ExoS activation, triggered by the loss of ExoR<sub>m</sub> inhibition, leads to ExoS auto-phosphorylation in its highly conserved cytoplasmic kinase domain, downstream phosphorylation of the 240 residue transcriptional factor ChvI [11], and the activation and/or suppression of multiple lifestyle-associated genes [12,13,14,15,18,22,23,24]. This control mechanism not only prepares cells by switching from flagellum to succinoglycan production, but also regulates metabolism and cell envelope changes necessary for cellular differentiation into endosymbiotic nitrogen fixing bacteroids [14,20,25].

Individual orthologous components of the RSI invasion switch have been found in other bacteria, including mammalian and plant pathogens. The ExoR ortholog from the plant symbiont *Rhizobium leguminosarum* regulates exopolysaccharide production [26], which is similar to the role of *S. meliloti* ExoR [12,27,28,29]. The ExoR ortholog from the plant pathogen *Agrobacterium tumefaciens* regulates the production of succinoglycan and biofilm [30], and the acid induced type VI secretion system [31,32], both of which are essential for host invasion. Mutant screens have identified ExoS and ChvI orthologs in *Agrobacterium tumefaciens*, *Rhizobium leguminosarum*, *Bartonella henselae*, and *Brucella abortus* [33,34,35,36,37,38]. The *A. tumefaciens* ChvG(ExoS)/ChvI system regulates virulence against its plant hosts by modulating succinoglycan and other virulence factor production based on the low pH at the site of infection [31,33,39,40]. The *R. leguminosarum* ChvG(ExoS) sensor kinase contributes to the control of outer membrane protein expression [35,41]. In *B. henselae*, the orthologous BatR/BatS TCS system responds to host-dependent environmental conditions and induces a virulence regulon necessary for intra-erythrocytic mammalian infection [34]. The BvrS(ExoS)/BvrR(ChvI) system in *B. abortus* regulates its invasion of host macrophages [36,37,42,43,44,45]. More recently, whole genome sequencing has identified a large number of putative ExoR, ExoS, and ChvI homologs in numerous other bacteria, many of which lack functional analysis. Most importantly, the discovery of individual components of the pathway in various mammalian and plant pathogens suggests that ExoR, ExoS, and ChvI may regulate host invasions that have otherwise been considered unique, and invasions that eventually bring harm rather than benefit to hosts.

In this report, we characterize the phylogenetic distribution of orthologs of the RSI pathway proteins to test if the presence of an “RSI Switch” is associated with an intracellular bacterial lifestyle. We take the approach of phylogenetic profiling, a bioinformatics method for identifying protein networks based on phylogenetic co-occurrence of network components [46,47,48]. Our findings demonstrate that functionally orthologous RSI invasion switches are only a pervasive, taxon-specific genomic characteristic among Rhizobiales. The ecological characteristics of the bacteria within this “RSI group” suggest that this signaling pathway is a key adaptive mechanism responsible for the success of both symbiotic and pathogenic dimorphic species. Furthermore, we have identified variable regions within ExoR and ExoS that are the best candidates for site-specific mutational analysis to further investigate the function of the RSI switch and its orthologous pathways in the Rhizobiales.

## Materials and Methods

### Phylogenomic analysis with BLAST

Best-best searches in the Kyoto Encyclopedia of Genes and Genomes (KEGG) Sequence Similarity Database (SSDB) [49,50] were conducted to define the phylogenomic limits of our queries. Sequences identified in phmmer Hidden Markov model (HMM) searches (Janelia Farms [51], GenBank nr database [52],  $E \leq 1e^{-5}$  criterion, Alphaproteobacteria) were selected for further analysis. Context Specific (CS) protein blasts [53,54] were used to confirm the phmmer results. CS searches were implemented in three rounds with  $E = 10^{-6}$ ,  $10^{-5}$ , and  $10^{-4}$ . Reciprocal best NCBI blastp [55] searches were conducted against the GenBank [52] nr database (BLOSUM62, alignments of  $\geq 60\%$  of query length and  $E \leq 1e^{-5}$ ). GenBank accession numbers and Blast results are provided in S1 Table. Many sequences demonstrated similarities to ExoS primarily within its highly conserved BaeS histidine kinase domain [CDD: COG0642]. Since the physical interaction of ExoR and ExoS within Rm1021 periplasmic space is crucial to the function of the RSI switch, we required significant similarities and alignments against ExoS periplasmic sensing domain (60% of ExoS length,  $E \leq 1e^{-5}$ ). Genomic loci architectures and

their associated COGs (Cluster of Orthologous Groups) were obtained from the Integrated Microbial Genomes and Metagenomes (IMG) database (<http://img.jgi.doe.gov/>) [56]. COG assignments were based primarily on those of Rm1021 and on consensus Rhizobiales COGs when differences were found.

To confirm that our criterion for putative homologs (i.e.,  $E \leq 1e^{-5}$ ) was sufficiently generous, we demonstrated that the predicted folds of excluded hits differed from the fold of ExoR. To guard against algorithm-intrinsic biases in this analysis, we utilized consensus predictions from four fold-threading servers (SPARKS-X [57], LOOPP [58,59,60,61,62], HHpred [63], and pGenThreader [64]).

When selecting species/strains from multiple sequences within a family or taxon for further investigation, we prioritized candidates based on following hierarchy: (1) type strains, (2) community selected reference organisms, and (3) the organism with the most publically available data.

### Filtering by structural characterization

Structural protein predictions were used to narrow our search for orthologs (results presented in [S2 Table](#)). For the identification of ExoR orthologs we required positive predictions of (1) Sel1 domains and (2) extracellular localization. Both of these sequence characteristics would be necessary for any ortholog to function in a manner similar to ExoR. For ExoS orthologs, we required predictions of structural elements based on the *S. meliloti* prototype: a Sensor Domain [Pfam: PF13756] sandwiched between two trans-membrane helices (TMHs). ChvI putative orthologs were selected using the same methodology, requiring consensus predictions of (1) Response\_reg [Pfam: Pf00072] and (2) Trans\_reg\_C [Pfam: PF00486] domains.

For structural predictions, we utilized a majority rule method to avoid intrinsic algorithm bias. Predictions for highly helical proteins such as ExoR can return inconsistent results as multiple sequence motifs that can fold into similar local structures (e.g., Sel1 and TPR domains). Accordingly, we used the consensus from (1) SMART [65,66], (2) Pfam [67], (3) phmmer [51], and (4) NCBI's Conserved Domain Database (CDD) [68] to predict domains. Consensus trans-membrane predictions were compiled from (1) SMART [65,66], (2) HMMTOP [69], (3) Octopus [70], (4) Phobius [71], and (5) TMHMMv2.0 (<http://www.cbs.dtu.dk/services/TMHMM/>). Consensus secretion signals were obtained using (1) SMART [65,66], (2) Octopus [70], (3) SignalP4.0 [72], and (4) Phobius [71]. Protein localization and predictions of non-canonical secretion were based on PSORTb3.0.2 [73], the SOSUI prokaryota database [74], and the SecretomeP 2.0 algorithm [75].

### Multiple sequence alignments (MSAs)

MSAs were created using Toffee Espresso with standard parameters [76]. MSA quality was verified by visual inspection using Jalview [77] and GenDoc [78] editors. The ExoR ortholog MSA was modified on the N-terminus to align the first Sel1 repeats and signal peptides. The ExoS MSA was trimmed into two regions, periplasmic and cytoplasmic, after the domains and 2D structure alignments were confirmed. No modifications were made to the ChvI alignment after manual verification.

The ExoR, ExoS, and ChvI alignments were blocked, removing all gaps greater than ten residues and increasing the percent continuous conserved sequence (TrimAl v1.3 [79] with 'Gap-pyOut' parameters). Additional manual blocking of the ExoR ortholog alignment was made to (1) remove inter-domain non-conserved sites and (2) to assure the inclusion of predicted protein-protein interaction sites. 0.7% undetermined and gapped positions was achieved for ExoR.

No modifications were made to the ExoS or ChvI TrimAl results. The final blocked alignments for ExoS and ChvI had 0.02% and 0.2% undetermined and gapped sites, respectively.

## Phylogenetic analyses

The ExoR, ExoS, and ChvI ortholog trees were reconstructed with maximal likelihood (RAxML) and Bayesian (MrBayes) algorithms. Since the histidine kinase domain of ExoS is highly conserved, we utilized only the periplasmic region for our phylogenetic analyses [80,81]. ProtTest2.4 [82] was used to rank models based on the blocked alignments using site-built BioNJ trees and AICc criteria. LG+G models were used for all of the data sets and invariant sites were not used as per XSEDE recommendations for RAxML-HPC. WAG+I+G, WAG+G, and JTT+G, were used for ExoR, ExoS, and ChvI ortholog sets, respectively, in MrBayes 3.2.1. Maximum likelihood trees (RAxML-HPC on XSEDE, 7.6.3) [83,84] were built with 1200 initial bootstraps and then to convergence on the CIPRES Science Gateway V.3.3 [85] (<http://www.phylo.org/index.php/portal/#>). Final bootstrapping was performed to auto majority rule (auto-MRE) criterion and SumTrees Python script from the DendroPy3.12.0 package [86] was used to build the consensus trees. Bayesian trees [87,88,89] were built for all sequence sets on the City University of New York High Performance Computer Center (CUNY HPCC) cluster. Four rate categories for all analyses and invariant sites for ExoR orthologs were used. Final Bayesian phylogenies were run for 5 million generations with burn-ins of 750,000, producing consistent convergence for all ortholog sets.

For parsimony inference of gene gains and losses, we used a customized Perl script based on BioPerl [90] to extract the phylogeny of the selected Proteobacteria genomes from a bacterial species tree inferred using 16S RNA sequences [91]. FigTree (<http://tree.bio.ed.ac.uk/software/figtree/>) was used to produce tree images. We have presented pseudo-rooted protein phylogenies as this was more amenable for branching comparisons. Since the RSI protein branching patterns follow accepted Rhizobiales phylogenies [92,93], we have rooted according to the earliest predicted speciation within our dataset. We estimated levels of sequence conservation at each amino-acid position using the computer program Rate4Site [94]. This program outputs standardized scores of amino-acid substitution rates at each residue position based on a multiple protein sequence alignment. The lower the rate score, the more conserved a residue position is evolutionarily. We used *S. meliloti* orthologs as the reference molecule in this analysis.

## Results

### RSI-like pathways are unique to Alphaproteobacteria

In an initial survey, 394 and 398 bacterial genomes that putatively encode Rm1021 ExoS and ChvI orthologs (KEGG best-best analysis, orthology groups K14980 and K14981), respectively, were identified. The lack of KEGG one-to-one ExoS to ChvI putative ortholog matches, as would be expected, is likely due to incomplete annotations and/or the presence of orphan sensor kinases [95]. 39% and 51% of the ExoS and ChvI best-best hits, respectively, were from Alphaproteobacteria. Overall, slightly less than half (41%) of the KEGG hits were from the Rhizobiales, an order of Alphaproteobacteria. In contrast, Rm1021 ExoR initial ortholog searches (KEGG best-best orthology group K07126) returned only 83 genomes within 23 bacterial genera (100% Alphaproteobacteria and 92% Rhizobiales). These results strongly suggested that (1) genomes that encode ExoR orthologs are a subset of ExoS(ChvG)/ChvI genomes and (2) complete, orthologous RSI pathways could only be found in the Alphaproteobacteria. It is likely that outside of the Rhizobiales, TCSs similar to ExoS/ChvI function either without a third partner or with proteins other than ExoR orthologs.

**Table 1. The identification of ExoR orthologs.**

| Selection methods                | Criteria   | Number of candidates |
|----------------------------------|--|----------------------|
| HMM & CS BLASTs                  | Query with Rm1021 ExoR   | 95                   |
| HMM & CS BLASTs                  | Genomes also encode candidate ExoS & ChvI orthologs  | 92                   |
| Reciprocal BLASTp                | Manual best-best hits: ExoR to candidate sequences   | 57                   |
| Consensus structural predictions | Putative ExoR orthologs (1) with signal peptides, (2) without TMHs, & (3) without non-Sel1 domains | 52                   |
| Phylogenetics: ML & Bayesian     | Resolved with majority support   | 47, final set        |

doi:10.1371/journal.pone.0135655.t001

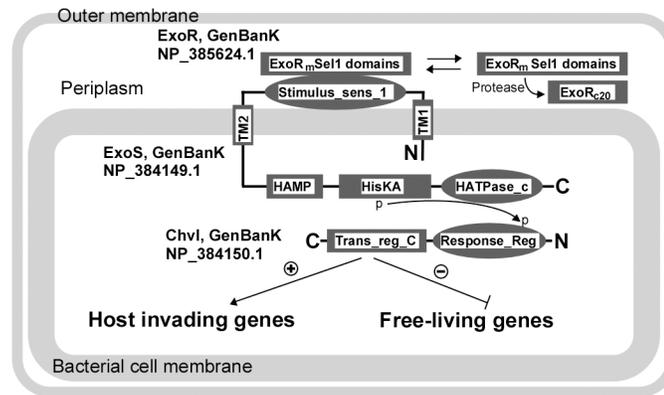
Our curated ortholog searches first focused on Rm1021 ExoR, since this protein, in comparison to ExoS and ChvI, demonstrated the most limited similarity across Alphaproteobacteria genomes. The initial candidate orthologs (non-paralogous sequences from 95 genomes) were taken from HMM/CS BLAST results (Table 1 and S2 Table). However, this group was narrowed from 95 to the 92 genomes that were found to encode candidate orthologs of all three RSI proteins. Manual best-best BLASTs querying the putative ExoR orthologs against the *S. meliloti* Rm1021 genome, were positive for 57 of the 92 genomes (62%) (Table 1). Among this set of 57 candidate ExoR orthologs, 86% originate from Rhizobiales genomes, while the remaining 14% predominantly originate in the Rhodobacterales, a closely related, but potentially non-monophyletic order [93] (Table 1 and S2 Table). We concluded that although systems with proteins like ExoS and ChvI are widespread in Alphaproteobacteria, complete orthologous RSI-like pathways are only consistently found in a subset of this bacterial class.

BLAST queries were used to select putative homologs and the genomes with BLAST hits for all three RSI switch proteins were subsequently identified. Using reciprocal best-best searches, we then differentiated putative orthologs from homologs. The final set of RSI orthologs was obtained by (1) imposing strong predictions for functionality using localization signatures and domain architectures, and (2) requiring resolved ortholog phylogenies. (ExoS ortholog alignment blocking for reconstruction resulted in one set of redundant sequences among the Brucellaceae and reduced the non-redundant number of sequences to 46.)

### Structural characteristics of the RSI proteins and candidate orthologs

*S. meliloti* Rm1021 ExoR (GenBank Accession NP\_385624.1) is a periplasmic protein of 268 residues with a signal peptide and a set of six Sel1 repeats [17] (Fig 1). These structural characteristics determine the function of ExoR within the invasion switch, and we identified the sequences that could potentially fulfill similar functions. Thus, we further narrowed our putative ExoR orthologs by requiring consensus predictions for both secretion and Sel1 domain repeats.

The majority of the 57 candidate ExoR orthologs (91%) were less than 275 residues, while five encoded more than 350 residues (from *Azorhizobium caulinodans* ORS 571, *Hyphomicrobium denitrificans* ATCC 51888, *Novosphingobium aromaticivorans* DSM12444, *Pelagibacterium halotolerans* B2, *Rhodomicrobium vannielii* ATCC17100) (Table 1 and S2 Table). Outlying sequence-encoded structural features among ortholog candidates, as compared to Rm1021 ExoR, included localization signatures and non-Sel1 domains. Four candidate ExoR orthologs unexpectedly returned strong predictions of non-periplasmic localizations, i.e., lack of signal peptide, potential transmembrane helices (TMH), and/or a lack of other secretion mechanisms (from *A. caulinodans* ORS 571, *Beijerinckia indica* ATCC 9039, *Parvibaculum lavamentivorans* DS-1, and *P. halotolerans* B2) (Table 1 and S2 Table). Based on current



**Fig 1. ExoR, ExoS, and ChvI proteins interact to create the RSI Invasion Switch in *S. meliloti*.** ExoR is composed of Sel1 repeats, is secreted to the periplasm as ExoR<sub>m</sub> (lacking the signal peptide), and is cleaved to release its suppression of ExoS. The domain architectures of ExoS and ChvI, along with phosphorylation sites, are illustrated here. Signal transmission is achieved via a phosphorylation cascade and ChvI completes the “switch” by binding to DNA to alter gene expression [96]. Our current model suggests that the levels of ExoR<sub>m</sub> in the periplasm are maintained through the combination of biosynthesis and proteolysis. Proteolysis is possibly sensitive to host signals or to changes in environmental conditions. These mechanisms allow for (1) ExoR<sub>m</sub> levels to be dramatically reduced in the presence of host signals, (2) the turning “ON” of the RSI invasion switch, (3) the activation of host-invading genes, and (4) the suppression of free-living genes.

doi:10.1371/journal.pone.0135655.g001

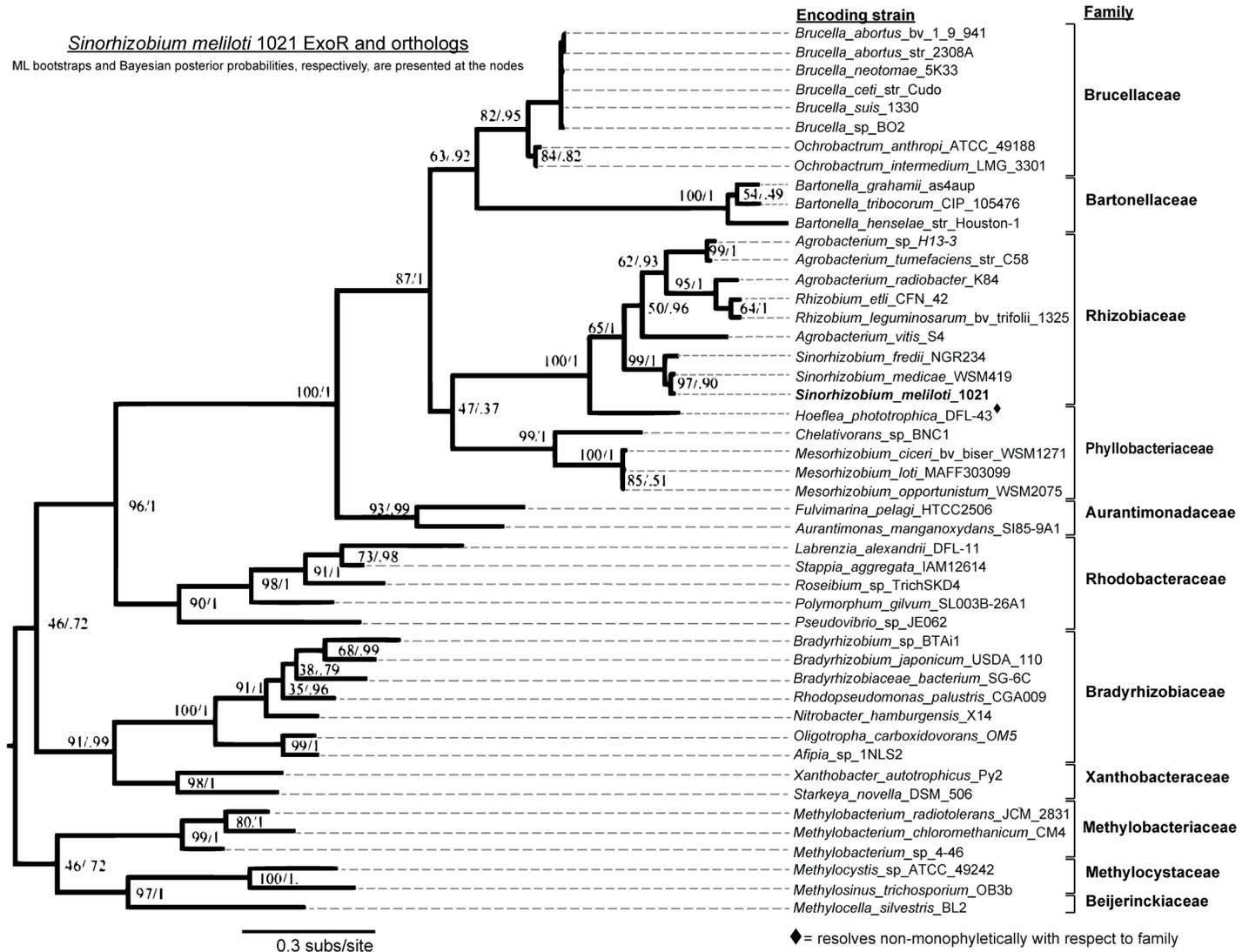
understanding of RSI, proteins that are not secreted to the periplasm cannot function as ExoR-like environmental sensors. One candidate sequence was found to encode a domain other than Sel1 (Sporulation related domain (SPOR), Pfam: PF05036R; *Novosphingobium aromaticivorans*, Table 1 and S2 Table). The final set of 47 candidate ExoR orthologs retained after these analyses were found predominantly in Rhizobiales, and have repeats of a binding domain (Sel1) in addition to consensus predictions for secretion.

All identified ExoS and ChvI orthologs have high levels of sequence similarity (S1 Table) and domain architectures that mirror the *S. meliloti* ExoS and ChvI TCS. These conserved structural characteristics suggest that ExoS and ChvI orthologs may respond to external stimuli and activate transcriptional responses similar to those of *S. meliloti* and *A. tumefaciens*. In our searches, we found ExoS and ChvI orthologs in every genome that encoded an ExoR ortholog, the majority of which were found among the Rhizobiales (Table 1 and S1–S3 Tables).

### Phylogenies of RSI proteins are consistent with each other and with accepted Rhizobiales phylogeny

The Toffee Espresso [76] multiple sequence alignment (MSA) for the ExoR, ExoS, and ChvI ortholog sets required little manual adjustment and supported our domain predictions (S1 Fig). The predicted Sel1 domains of the ExoR orthologs aligned well with few notable insertions. The ExoS and ChvI MSAs also supported domain annotations and predictions.

Well-supported protein phylogenies (Figs 2–4) were produced for the ortholog sets after non-resolvable sequences from five taxa were excluded. These outliers were identified by their comparatively long branch lengths in a preliminary ExoR phylogeny that was poorly supported at numerous nodes (S2 Fig). Three of these five outlying sequences had predictions of divergent localizations (*Ahrensia* sp. R2A130, *Magnetospirillum magnetotacticum* MS-1, *Rhodospirillum rubrum* ATCC17100), while the other two protein sequences are notably longer than ExoR (*Hyphomicrobium denitrificans* ATCC 5188, *Hyphomicrobium* sp. MC1) (Table 1 and S2 Table). Given the exclusion of these sequences, the overall ExoR reconstruction bootstrap support was highly improved (Fig 2 versus S2 Fig), particularly at basal bifurcations.

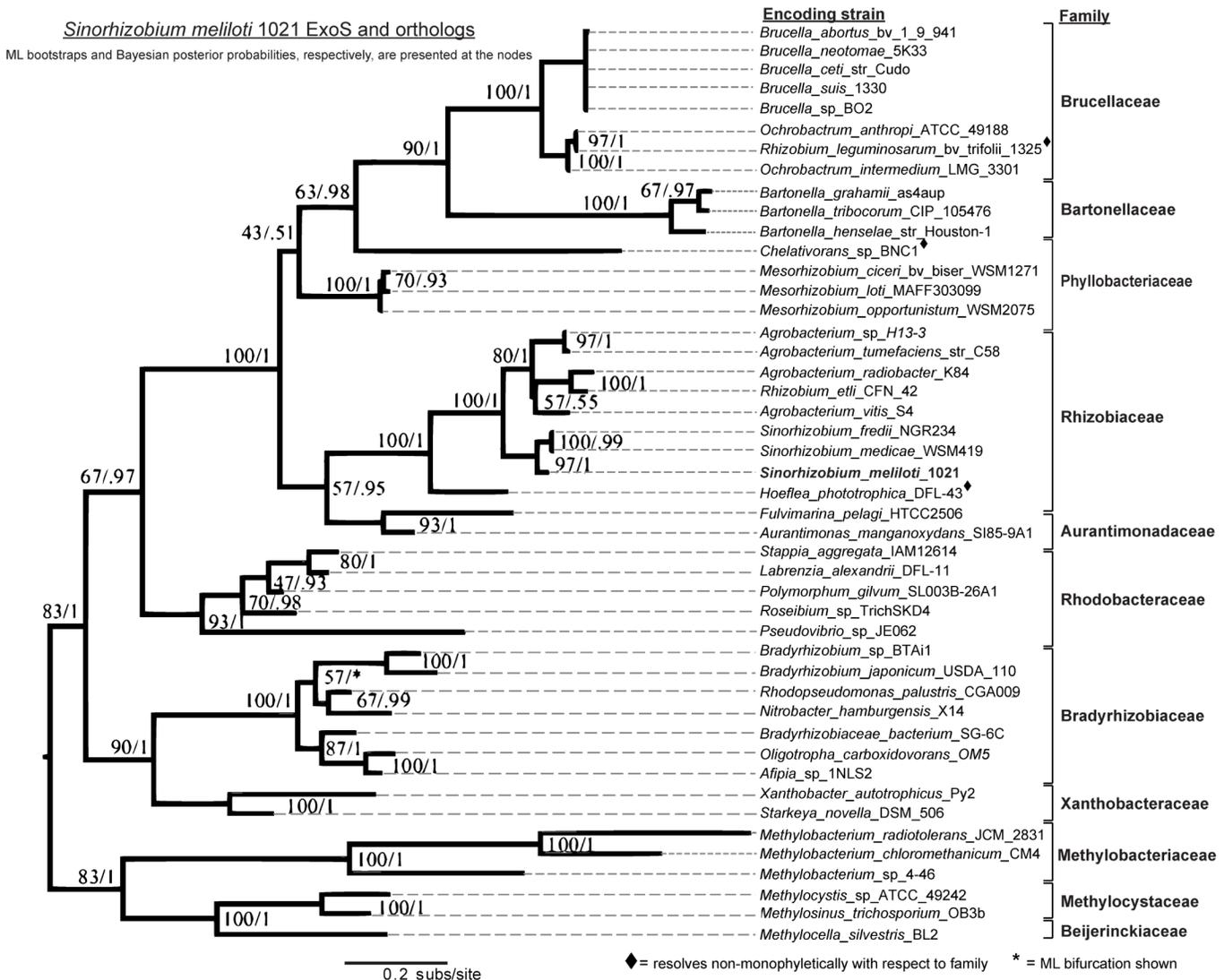


**Fig 2. Evolutionary history of ExoR is congruent with speciation history of Rhizobiales.** ML and Bayesian phylogenetic patterns support the existence of an ExoR ancestor that arose prior to diversification among the *Rhizobiales* and five *Rhodobacterales* species identified in our genomic searches. The ExoR gene was maintained among these species, giving rise to the ortholog set. Phylogenetic reconstruction of ExoR orthologs agrees with currently accepted speciation patterns.

doi:10.1371/journal.pone.0135655.g002

In the final ExoR trees (Fig 2), a node that divides the *Rhizobiaceae* from the *Phyllobacteriaceae* is not well supported by either the likelihood or Bayesian method, but is similar to speciation patterns (PATRIC *Rhizobiales* phylogeny, <http://patricbrc.vbi.vt.edu/> [92,93]). The ExoS reconstructions (i.e., ML and Bayesian) agreed in 39 of the 40 bipartitions (98%) (Fig 3). In the ChvI ortholog reconstructions, the Bayesian tree was better supported than the ML in 68% of the nodes (Fig 4). The most significant variation between the two methods in the ChvI trees was found within the *Sinorhizobium* spp. (Fig 4) and the divergence of the *Phyllobacteriaceae* from the *Rhizobium/Bartonellaceae* clade was unresolved in all of the RSI ortholog trees.

The lower level support in the ChvI reconstruction is likely due to low sequence divergence since ChvI-like proteins are extremely common as bacterial transcription factors. The production of high-quality alignment blocks removed much of the original sequence variation present

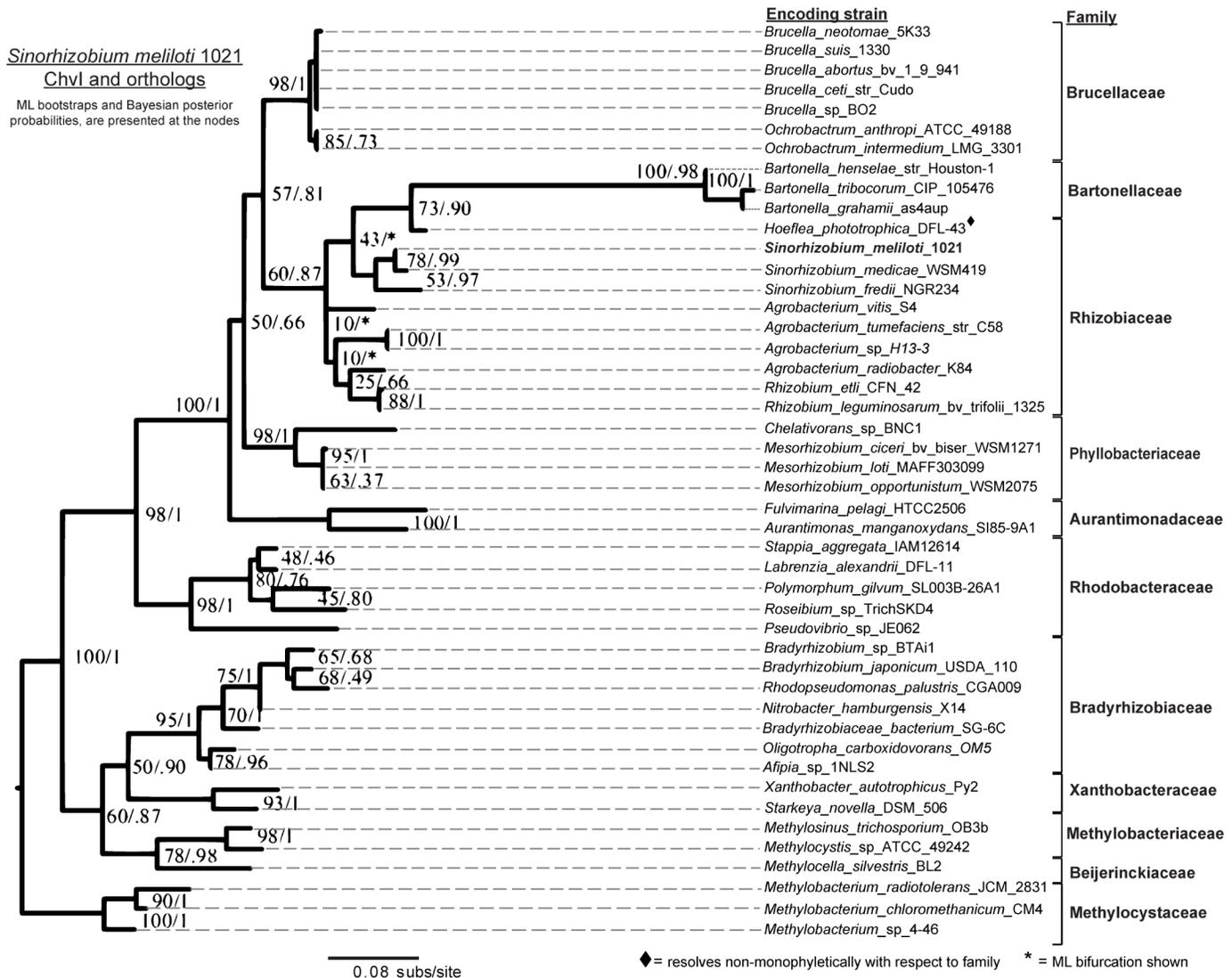


**Fig 3. Concordant evolution of ExoS and ExoR.** The overall evolutionary rates and branching pattern of the ExoS phylogeny agree with those of ExoR (see S3 Fig for mirrored comparisons). Although not all ExoS orthologs resolve strictly with family speciation patterns, most differences are minor. The *Phyllobacteriaceae* orthologs are, in fact, the most significant difference in the predicted histories of the ExoR and ExoS sets. The comparison of the ExoS reconstruction to the ExoR (Fig 2) supports the assertion that the evolution of the two ortholog sets was concordant.

doi:10.1371/journal.pone.0135655.g003

in our ortholog sets. The lack of resolution among *Brucellaceae* RSI orthologs is most likely also due to low genetic, and thus sequence, divergence [97]. For all RSI tree comparisons, we have considered the *Brucella* species proteins as a unified group.

The three sets of ortholog trees were compared pairwise (S3 Fig) to assess the possibility of their synchronized evolution. The similar (73% of family nodes) branching patterns of the ExoR and ExoS ortholog trees (Figs 2 and 3), as well as comparable mutation rates (0.2 and 0.18, respectively) suggest similar evolutionary pressures. Surprisingly, the ExoS-ChvI ortholog tree pair demonstrates fewer congruent (45%, or 5 of 11) family nodes (S3 Fig), even though linked evolutionary pressure in TCS protein pairs is well-established. Since the mutation rate in the ChvI ortholog protein set is low (approximately 25%) in comparison to those of both the ExoR and ExoS orthologs, we propose that purifying pressures lead to high levels of



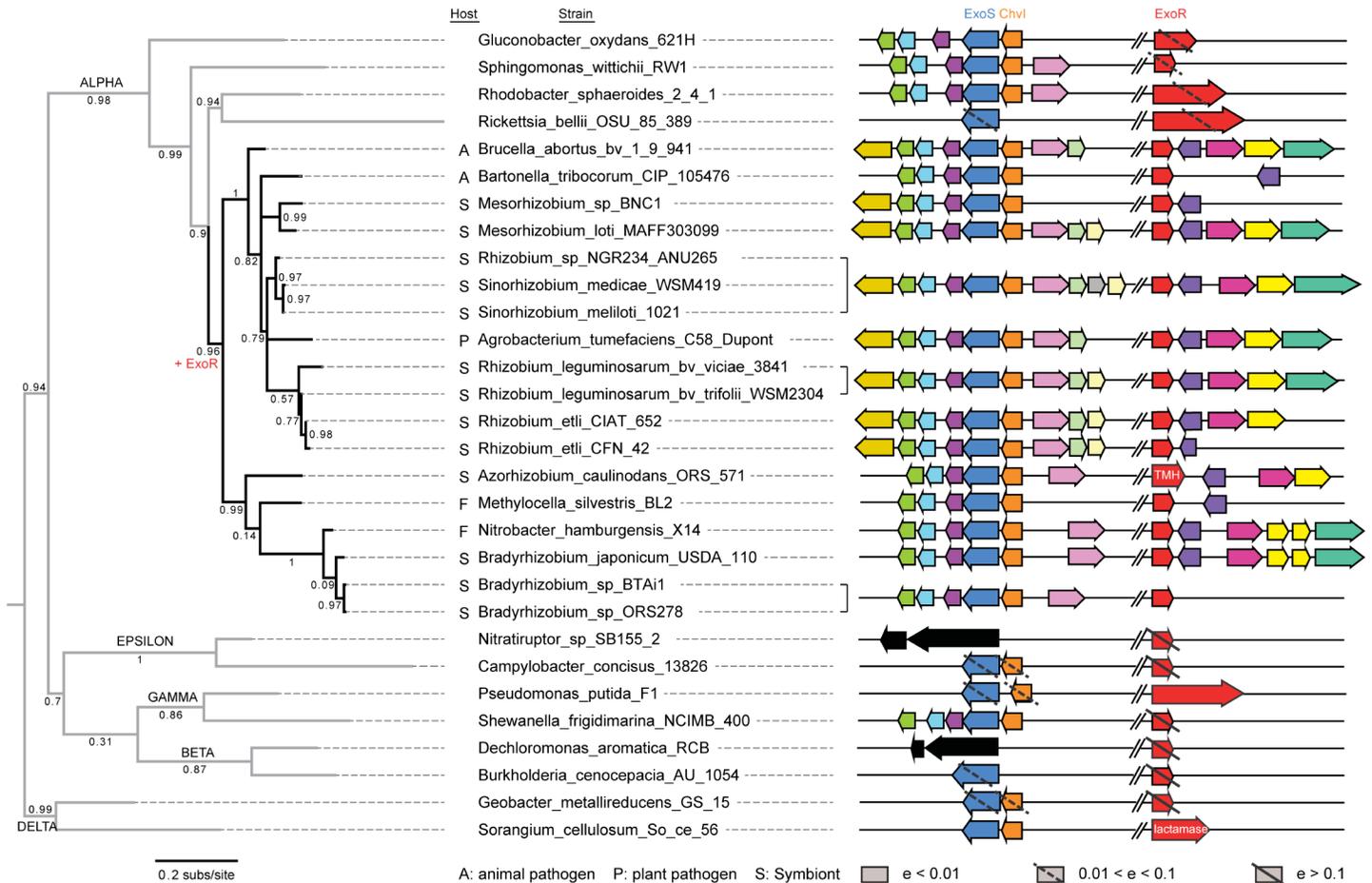
**Fig 4. Conserved sequence evolution of ChvI in Rhizobiales and selected Rhodobacterales.** As a cognate TCS pair, ChvI and ExoS sequences are expected to diversify at similar rates. However, the ChvI phylogeny suggests that it is under strong purifying selection with the lowest amino-acid substitution rate among the three components. It is possible that the functional role of ChvI orthologs (i.e., DNA binding) limits diversification.

doi:10.1371/journal.pone.0135655.g004

conservation and low mutability among the ChvI orthologs. While our comparison of the ExoR-ExoS ortholog trees suggests co-evolution, the initial resolution of the trees, stochasticity, non-monophyletic groups, and divergence in the more ancient lineages must be taken into account. In summary, the phylogenies of the three RSI proteins (Figs 2–4) are highly consistent, not only to each other, but also to the known phylogeny among these Alphaproteobacteria species.

### A single evolutionary origin and the persistence of RSI switch in Rhizobiales

To assess RSI conservation and loci architectures both within and outside of the Rhizobiales, we conducted synteny analysis (Fig 5). Orthologous RSI pathways were mapped onto the



**Fig 5. RSI Synteny analysis within Rhizobiales and ancestral Alphaproteobacteria.** Clusters of orthologous gene (COG) groups for representative Proteobacteria are presented for comparison to the prototype, the loci of Rm1021. Each COG, as defined in the IMG database, is represented by an arrow of different color. For the *exoR-exoS-chvI* genes, a lack of homologs or orthologs is presented as an ORF with a solid or dotted slash, respectively. The sizes and positions of the ORFs are approximate, representing relative expansions/deletions. The ExoR ortholog found in *Azorhizobium caulinodans*, although annotated as an exopolysaccharide regulator, is predicted to encode transmembrane helices and is a structural outlier as compared to ExoR-like molecules found in other Rhizobiales. Lifestyles and host type are given by A (animal pathogen), P (plant pathogen), or S (symbionts). Gene loss/gain is indicated with +/-, respectively. COG# color key (L to R): Dark yellow, 0499; green, 1925; light blue, 2893; dark violet, 1493; blue (*exoS/chvG*), 0642; orange (*chvI*), 0745; light violet, 1866; pale green, 1186; grey, 1652; white, 3145; red (*exoR*), 0790; purple, 0708; magenta, 1502; bright yellow, 0232; turquoise (sporulation-domain encoding). In Rm1021 RSI loci are *exoS* (bp49252–51039), *chvI* (51419–52141), and *exoR* (1637310–1638116) on the SMC chromosome.

doi:10.1371/journal.pone.0135655.g005

Alphaproteobacterial phylogeny with gene neighborhoods organized according to species divergence to visualize (1) gene losses/gains and (2) patterns of co-evolution [95].

ExoR orthologs were not identified in a majority of species within any Alphaproteobacterial order other than Rhizobiales. Conversely, the presence of ExoS/ChvI orthologs is an ancestral state to the order (Fig 5). The evolutionary pathway to this state is not conserved as ExoS/ChvI orthologs were identified, but rarely, in the other Proteobacterial classes (e.g., *Sorangium cellulosum*, Deltaproteobacteria). Sequences with distant relationships to ExoR (BLASTs  $E < 10^{-5}$ , 75% alignment), were identified outside of Rhizobiales, but uniformly lacked best-best character with respect to Rm1021 ExoR and its crucial structural characteristics (S2 Table). Fully orthologous RSI-like protein sets were identified only within Rhizobiales as a taxonomically conserved characteristic, suggesting that this pathway is a lineage-specific adaptive feature of the order.

Among the sampled genomes, the ExoR ortholog neighborhoods (COG0790) show low levels of conservation in comparison to the ExoS/ChvI orthologs. Although the COG 0232 (dGTP triphosphohydrolase) and COG 0018 (arginyl-tRNA synthetase) groups are commonly found with the ExoR COG 0790, there are genomes illustrated in Fig 5 that do not follow this pattern (e.g., *Mesorhizobium etli* CFN42 and *Mesorhizobium sp.* BNC1). COG 0708 (exodeoxyribonuclease III), is the only COG class within which the ExoR orthologs are consistently located. However, since COG 0708 and ExoR orthologs are encoded on opposite strands, it is unlikely that they are under similar transcriptional control.

The ExoS (COG0745)/ChvI (COG0642) ortholog loci show notable conservation which extends into more ancestral orders such as the Rhodobacterales. The phosphotransferase/HPPr related cluster of COG 1493, 2893, and 1925 (incomplete phosphotransferase system among the Alphaproteobacteria), along with *exoS* constitute an operon in Rm1021. This architecture is conserved among the Rhizobiales and suggests that the operon itself is conserved.

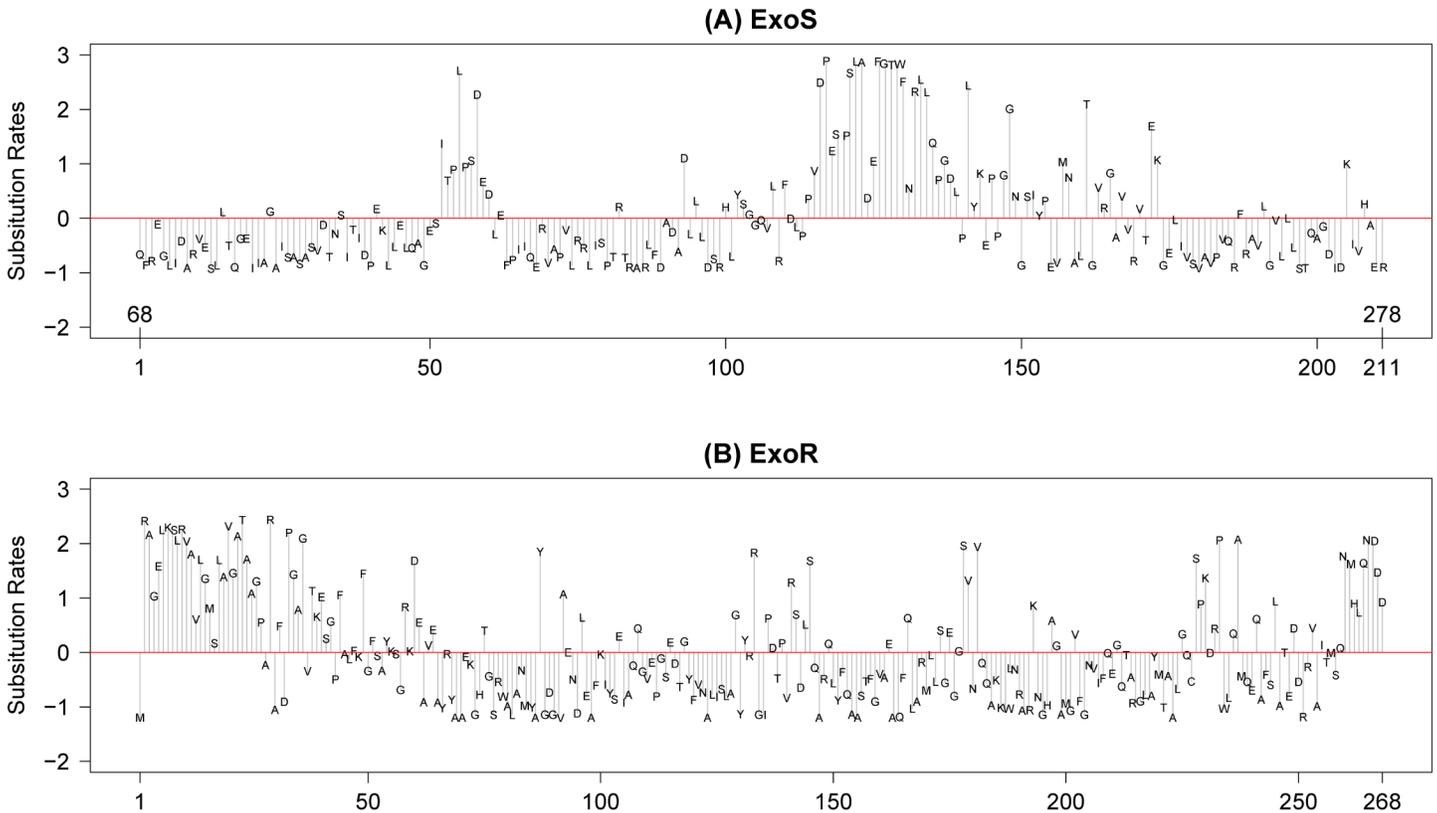
### Non-orthologous ExoR proteins may serve divergent functions in non-Rhizobiales species

Of the 35 sequences identified as potential Rm1021 ExoR homologs, but not as orthologs (Table 1 and S2 Table), approximately one-third were found to have divergent structural and/or functional predictions. Outlying structural predictions included the SPOR (Pfam: PF05036), PG\_binding\_1 (Pfam: PF01471), and Peptidase\_C14 (Pfam: PF00656) domains. Unexpected, predicted functionalities included beta-lactamase activity and organelle localization signaling (e.g., PodJL). These putative ExoR homologs with alternative structural and/or functional predictions were found in more ancestral lineages, such as the Caulobacterales, the Rhodobacterales, the Rhodospirillales, and the Sphingomonadales (S2 Table). SPOR domains and PodJL proteins are associated with peptidoglycan binding. Both the putative beta-lactamase and the peptidase C14 have associated binding predictions. We conclude from gene neighborhood syteny analysis (Fig 5) that there has been a single evolutionary origin for the complete set of RSI switch orthologs in a recent common ancestor of Rhizobiales. The RSI Switch has persisted throughout the subsequent divergence within the order.

### Evolutionary variability within ExoR and ExoS sensing domain

ExoR orthologs evolve faster than those of ExoS, with overall average pairwise sequence identities of 44.6% and 61.4%, respectively, based on comparative alignments among 46 non-redundant orthologs from Rhizobiales species. The N-terminus and C-terminus of ExoR are the most variable; although, conserved and variable residues are dispersed throughout the molecule (Fig 6B). In contrast, sequence variability of ExoS is concentrated in two regions, one short region consisting of the sequence ITPLPSDED (residues #52–60 of periplasmic region/#119–127 of full ExoS) and another longer regions consisting of PVDPEPSLA-DEFGTWFNRLQPGDL (#114–139 of periplasmic/#181–206 full ExoS) (Fig 6A).

Further evolutionary reconstruction based on parsimony revealed differential lineage-specific amino-acid substitutions associated with the emergence of species using animals as host versus plants (Fig 7). In the N-terminal hypervariable region of ExoS, a tryptophan, derived from an otherwise highly conserved leucine, is found among the *Brucellaceae* orthologs, along with the *Rhizobium leguminosarum* sequence. Although the plant symbiont *R. leguminosarum* is an outlier, the other species (*Brucellaceae* and *Bartonellaceae*) that encode tryptophans at this site are intra-erythrocyte, animal-infecting pathogens. The plant-associated *Rhizobiaceae* resolved together according to parsimony and consistently encode an asparagine within the more C-terminal ExoS hypervariable region (Fig 7). *Hoeflea phototrophica*, a species that has



**Fig 6. Site-specific evolutionary rates of ExoR (A) and ExoS (B).** Substitution rates (y-axis) at amino-acid positions (x-axis, numbers based on (A) the periplasmic region of *S. meliloti* ExoS and (B) full-length ExoR) were obtained using Rate4Site [94]. While variable and conserved residues are evenly dispersed in ExoR (B), two variable residues are notable and specific to two regions in ExoS (A). We hypothesize, subject to experimental verification, that these two variable regions form binding domains to ExoR and are the molecular basis of host-specific responses among Rhizobiales species.

doi:10.1371/journal.pone.0135655.g006

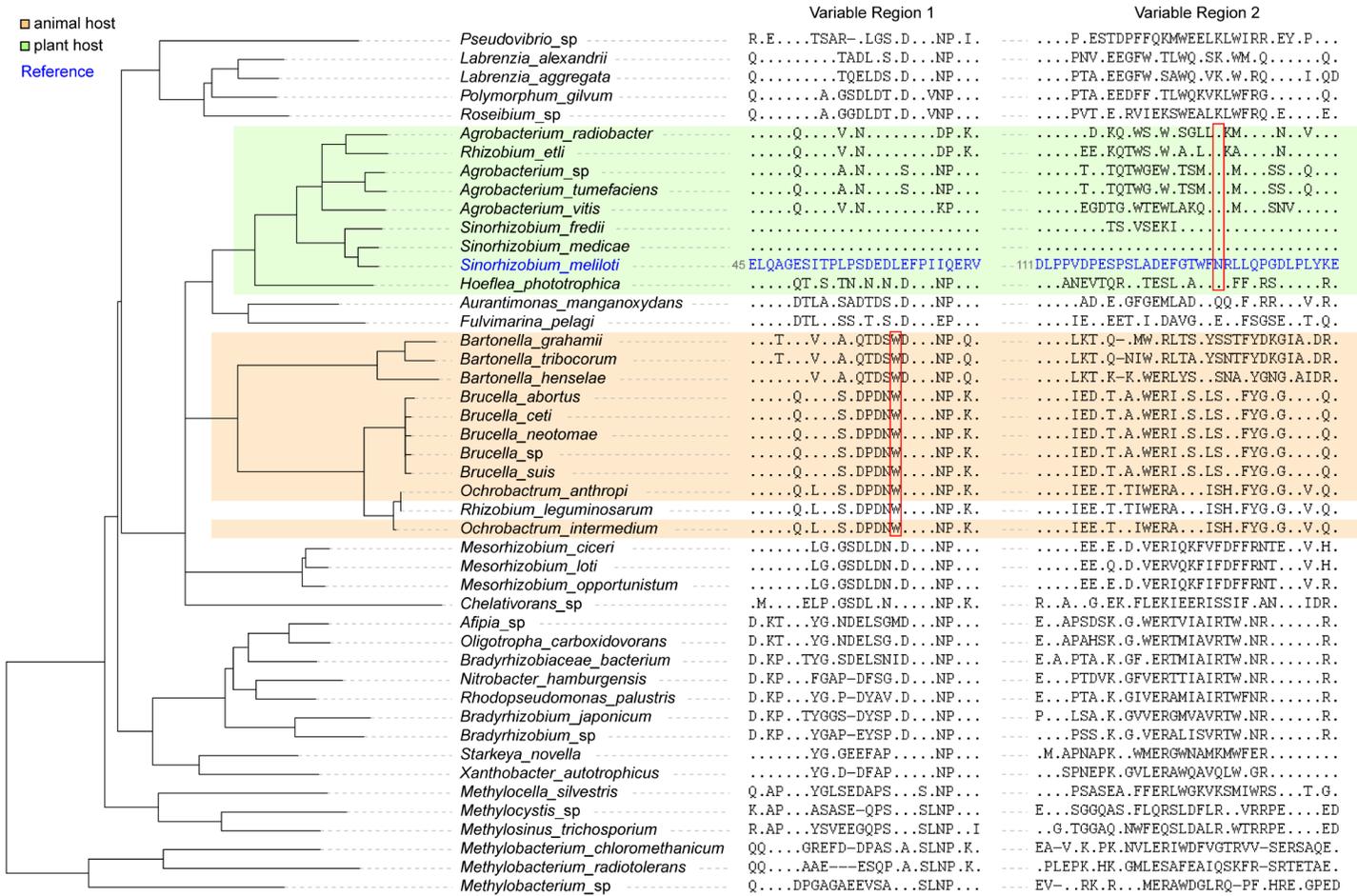
been found to associate with dinoflagellates [98,99], is included in this group, suggesting that a signaling commonality may exist between plant and protist host invasions.

## Discussion

The RSI invasion switch regulates biochemical changes within *Sinorhizobium meliloti* to produce invasion-competent bacteria from free-living forms, a morphogenesis that is necessary for host entry and the establishment of symbiosis [11,12,100]. The genomic encoding of RSI orthologs has been correlated to organisms that significantly impact human health and commerce, including diazotrophs and both animal and plant pathogens [30,34,35,36,37]. As detailed in the result section and following discussion, our analyses suggest that ExoR-ExoS-ChvI-like pathways exist as coevolutionary conserved units of signaling function, and they represent a molecular signature within a specific phylogenomic range. Most importantly, our results show that the RSI invasion switch could function in at least 47 different species which include mammalian and plant pathogens.

## The inferred evolutionary history of the RSI switch

TCS proteins similar to ExoS and ChvI are found in all domains of life [101] and their gains can be adaptive within new environments [95]. The conservation patterns of the RSI orthologs found here suggest that, in this case, a third protein was gained, adding to an existing TCS.



**Fig 7. Two host-associated radical amino-acid substitutions in ExoS.** Sequences within the two hypervariable regions (Fig 6) of ExoS are displayed according to a phylogeny of ExoS sequences. Based on parsimony reasoning, an asparagine residue (N131 of the ExoS periplasmic region) is an evolutionarily derived state that is conserved among plant- or dinoflagellate-associated taxa. A tryptophan residue (W61 of the ExoS periplasmic region) is derived from a leucine ancestor and conserved among animal-infecting species with the exception of *Rhizobium leguminosarum* (a plant pathogen).

doi:10.1371/journal.pone.0135655.g007

Although ExoS- and ChvI-like proteins were found at the base of Alphaproteobacteria, ExoR orthologs were not (e.g., *Novophingomondaceae* and *Rickettsiaceae*).

RSI ortholog loci patterns are consistent across the taxa investigated here. However, *exoS*- and *chvI*-like genes are adjacent while *exoR*-like loci are separate indicating that a transfer event of all three genes as a functional unit likely did not occur. Lineage expansion and domain shuffling [95] is a known diversification process in TCSs and could possibly extend to a third, interacting protein. Five Rhodobacterales (ancestral to Rhizobiales) were found to encode ExoR orthologs, but seven other Rhodobacterales species encode ExoR-like proteins with alternate functional and structural predictions. These sequences were selected from our BLASTs as putative homologs, but failed to satisfy our ortholog criteria.

In the prototypical *S. meliloti* RSI system, ExoR is composed solely of SelI domains, interspersed with linking regions [17]. In contrast, putative ExoR homologs (but not orthologs) in families basal to the Rhizobiales encode alternative and additional domains such as SPOR, PG\_binding\_1, and C14\_peptidase. Notably, SPOR and PG\_binding\_1 domains are associated with the periplasm and binding within this cellular compartment. Such ancestral functional profiles suggest that RSI-like modes of interaction (a negative regulator that binds a kinase

sensing domain) may have specialized from other examples of periplasmic binding, potentially via domain loss, to provide environmental response and transcriptional switching. These findings suggest a diversification event for the gain of ExoR-like orthologs within the orders ancestral to Rhizobiales. Changes within a third protein, such as an externally responsive sensor like ExoR, may have improved the necessary fitness to a Rhizobiales ancestor for adaptation to a new niche, such as nitrogen fixation, host-specific intracellular habitation, or alternative nutritional opportunities.

## Facultative lifestyles are characteristic of organisms that encode RSI Switch orthologs

We have found that tripartite RSI orthologous gene sets are characteristic of Rhizobiales, an order that includes numerous pathogens and symbionts. While lineage-specific presence of any one gene is itself an indication of species adaptation, phylogenetic co-occurrence of all three interacting partners strongly suggests that the RSI switch is a key adaptive pathway in Rhizobiales. Of the multiple RSI switch-encoding bacteria that can reside intracellularly within their hosts, the majority are predominantly facultative rather than obligate. Gross transcriptional switching is consistent with proposed RSI functions, and may regulate opportunistic responses to environmental signals by coordinating facultative changes in metabolism, membrane chemistry, and motility within alternate environments. Of the RSI-encoding species that lack known hosts we suggest that those with unstable genomes may be under selection for host association, such as select *Brucella* spp., perhaps due an advantage in their a free-living state.

Many facultative organisms, both diazotrophs and well-known pathogens, are found in the Rhizobiales. These species switch to intracellular lifestyles when enabled by evasion of, or resistance to, their hosts' immune responses. Accordingly, obligate intracellular pathogens, such as the well-known *Rickettsia* spp., were not found to encode RSI orthologs even though species from their sister clades were identified. [S3 Table](#) presents the lifestyles, hosts, and significant roles of the taxa included in this investigation. More than half of the listed species associate facultatively with other organisms and most frequently in symbiotic or pathogenic relationships. Many of the remaining species utilize or oxidize unique carbon, nitrogen, and metallo-organic molecules. Signaling links between their RSI-like switching pathways and these unusual metabolic pathways are not currently known, but are worthy of investigation.

Our phylogenetic analyses suggest that ExoS is under independent, site-specific positive pressures associated with host-type ([Fig 7](#)). As seen in ML/Bayesian phylogenies ([Figs 2–4](#)), the emerging pathogen *Bartonellaceae* are noticeably under more diversifying pressure than the orthologs from other families of the Rhizobiales order. Other emerging mammalian pathogens were identified with RSI sets (*Brucella* spp., *Methylobacterium radiotolerans*, and *Ochrobactrum* spp.) and outnumber the plant pathogens investigated here (*Agrobacterium* spp.). Organisms with unique facultative metabolisms (e.g., methanotrophism), that may be under switch-like transcriptional control, make up the majority of the non-pathogenic species that encode RSI-like pathways.

Closer investigation of ExoS demonstrates that the most rapidly evolving regions of ExoS are within the periplasmic regions and that characteristic insertions and radical substitutions are present ([Figs 6 and 7](#)). We have uncovered specific sequence motifs that correlate with host type (i.e., plant versus animal) among many of the species investigated here. These results not only provide impetus for functional testing via RSI mutagenesis, but they also suggest that adaptive pressures have acted on switch proteins in plant invading-species, animal pathogens, and other classes of less characterized species that invade basal eukaryotic hosts such as protists.

## The implications of an RSI invasion switch within pathogenic species

*S. meliloti* ExoR mutants show reduced invasion efficiency of their host plant, alfalfa [12,27,28,100]. ExoS/ChvI homolog mutants have also been correlated to reduced invasion abilities [14,34,36,37]. Given the high incidence of facultative intracellular lifestyles within the Rhizobiales, and the genetic similarity of pathogenicity and symbiosis islands, a shared sensing mechanism for both types of invasion is possible. Responding in an opportunistic way to specific external conditions enables both pathogenic and symbiotic lifestyles. RSI switch conservation in both symbionts like *S. meliloti* and pathogens like *Brucella* spp. suggests that the pathway which is mostly understood for its contribution to symbiosis, may also contribute to pathogen success.

In addition to the RSI system, other TCSs have evolved with accessory proteins, or adaptors, that control the functional TCS status based on environmental conditions [102]. The CpxP-CpxA-CpxR system [103] is conserved across a phylogenetically broad group of bacteria which includes Gram negative species such as *E. coli*. In this system, CpxP directly binds the sensor kinase CpxA under normal physiological conditions, but becomes sensitive to proteolysis in the presence of misfolded proteins. CpxP cleavage relieves its suppression of CpxA, a sensor kinase, which allows *E. coli* cells to respond to environmental stresses by increasing levels of chaperones and trafficking factors [103]. In the *Streptomyces* HpbS-SenS-SenR system, HpbS releases its negative regulation SenS through a conformational shift under high oxidative stress [104]. The induction of the *Streptomyces* SenS-SenR TCS leads to an increase in expression of protective anti-oxidative catabolic enzymes, like peroxidases and catalases [104]. Given the restricted and unified nature of the RSI orthology group, in addition to our experimental understanding of RSI, it is possible that the RSI system functions under similar environmental control and activation as these tripartite signaling systems.

Orthologs of the RSI invasion switch show synteny and highly similar histories within Rhizobiales. Our results suggest that these RSI-like orthologs sets have potentially experienced unified evolutionary histories and have responded to similar pressures to provide a key evolutionary innovation (KEI) for these species. Experimental works [12,13,15,16,18,31] support the necessity of RSI within Rhizobiales. Our phylogenomic RSI analyses provide an informed starting point not only to characterize host entry in emerging pathogens, but to study a potential connection between mechanisms of host entry in both symbiosis and pathogenesis. By determining the taxonomic range of this “switch,” we have facilitated the targeting of experimental analyses. By delineating potential conservation in invasion signaling, research aimed at pathogens may be facilitated in non-pathogenic species, creating safer short- and long-term research environments.

## Supporting Information

**S1 Fig. Characteristic domains of the ExoR homolog/ortholog data set.** Alignment has been trimmed at both termini and colored to show putative Sel1 domains. The longest putative homolog (Rhodomiv) included was trimmed as follows: 49 and 101 residues from its N- and C-termini, respectively. 5 sequences were removed from this set of 52 for the final reconstructions and analyses (see S3 Fig). Naming conventions are given in S1 Table. (PDF)

**S2 Fig. Preliminary ExoR ML reconstruction built with 52 species.** Bootstrap values are given. *Ahrensia*, *HypmcrDn*, *HypmcrSp*, *Rhodomiv*, and *MgntspirM* were considered to be non-resolved. Main text Fig 2 is built without these taxa. Color code: *Brucellaceae*, red; *Bartonellaceae*, light green; *Phyllobacteriaceae*, pink; *Rhizobiaceae*, royal blue; *Aurantimonadaceae*,

orange; *Rhodobacteraceae*, dark purple; *Bradyrhizobiaceae*, light blue; *Xanthobacteraceae*, yellow; *Methylobacteriaceae*, coral; *Methylocystaceae*, dark green; *Beijerinckiaceae*, brown. Naming abbreviations given in [S1 Table](#).

(TIF)

**S3 Fig. Mirrored comparisons: ExoR-ExoS and ExoS-ChvI paired ortholog phylogenies.**

(ExoR orthologs on left ExoS on right, first figure; ExoS orthologs on left and ChvI on right, second figure) Similar branching patterns among the RSI switch proteins suggest a unified phylogenetic history. Bootstrap values, followed by posterior probabilities, are given. Starred nodes are non-congruent. Color codes given in [S2 Fig](#).

(PDF)

**S1 Table. RSI ortholog information.** GenBank accession numbers & BLAST results.

(PDF)

**S2 Table. ExoR ortholog analyses.** Results of from reciprocal BLASTs and predictions of signal peptides, domains, and transmembrane helices (TMH).

(PDF)

**S3 Table. Phenotypes for species demonstrating RSI co-occurrence.** Every Alphaproteobacterial genome with an identified ExoR ortholog also encodes ExoS and ChvI orthologs. Many of these species are facultatively intracellular and include well-known examples of both pathogens and symbionts.

(PDF)

## Acknowledgments

We extend special acknowledgment to Dr. E. Wiech for expert assistance with structural analyses. We are thankful for Dr Lia Di's assistance in preparation of figures. Additionally, we would like to thank Drs. A. Litt, J. Rachlin, F. Burbrink, A. Carnaval, A. Berkov, and D. Lohman for their helpful advice and critical reading of the manuscript; to Drs. S. Govind and S. Singh for support and advice; and to the members of the Cheng, Govind, Qiu, and Singh labs.

## Author Contributions

Conceived and designed the experiments: MH WQ HC. Performed the experiments: MH WQ HC. Analyzed the data: MH WQ HC. Contributed reagents/materials/analysis tools: MH WQ HC. Wrote the paper: MH WQ HC.

## References

1. Gibson KE, Kobayashi H, Walker GC (2008) Molecular determinants of a symbiotic chronic infection. *Annu Rev Genet* 42: 413–441. doi: [10.1146/annurev.genet.42.110807.091427](https://doi.org/10.1146/annurev.genet.42.110807.091427) PMID: [18983260](https://pubmed.ncbi.nlm.nih.gov/18983260/)
2. Jones KM, Kobayashi H, Davies BW, Taga ME, Walker GC (2007) How rhizobial symbionts invade plants: the *Sinorhizobium-Medicago* model. *Nat Rev Microbiol* 5: 619–633. PMID: [17632573](https://pubmed.ncbi.nlm.nih.gov/17632573/)
3. Gage DJ (2004) Infection and invasion of roots by symbiotic, nitrogen-fixing rhizobia during nodulation of temperate legumes. *Microbiol Mol Biol Rev* 68: 280–300. PMID: [15187185](https://pubmed.ncbi.nlm.nih.gov/15187185/)
4. Long SR (1989) *Rhizobium*-legume nodulation: life together in the underground. *Cell* 56: 203–214. PMID: [2643474](https://pubmed.ncbi.nlm.nih.gov/2643474/)
5. Paulsen IT, Seshadri R, Nelson KE, Eisen JA, Heidelberg JF, Timothy DR, et al. (2002) The *Brucella suis* genome reveals fundamental similarities between animal and plant pathogens and symbionts. *Proc Natl Acad Sci U S A* 99: 13148–13153. PMID: [12271122](https://pubmed.ncbi.nlm.nih.gov/12271122/)
6. Halling SM, Peterson-Burch BD, Bricker BJ, Zuerner RL, Qing Z, Li L, et al. (2005) Completion of the genome sequence of *Brucella abortus* and comparison to the highly similar genomes of *Brucella melitensis* and *Brucella suis*. *J Bacteriol* 187: 2715–2726. PMID: [15805518](https://pubmed.ncbi.nlm.nih.gov/15805518/)

7. Tsolis RM (2002) Comparative genome analysis of the alpha-proteobacteria: relationships between plant and animal pathogens and host specificity. *Proc Natl Acad Sci U S A* 99: 12503–12505. PMID: [12271145](#)
8. Goodner B, Hinkle G, Gattung S, Miller N, Blanchard M, Qurollo B, et al. (2001) Genome sequence of the plant pathogen and biotechnology agent *Agrobacterium tumefaciens* C58. *Science* 294: 2323–2328. PMID: [11743194](#)
9. Oldroyd GE, Downie JA (2008) Coordinating nodule morphogenesis with rhizobial infection in legumes. *Annu Rev Plant Biol* 59: 519–546. doi: [10.1146/annurev.arplant.59.032607.092839](#) PMID: [18444906](#)
10. Leigh JA, Walker GC (1994) Exopolysaccharides of *Rhizobium*: synthesis, regulation and symbiotic function. *Trends Genet* 10: 63–67. PMID: [8191588](#)
11. Cheng HP, Walker GC (1998) Succinoglycan production by *Rhizobium meliloti* is regulated through the ExoS-ChvI two-component regulatory system. *J Bacteriol* 180: 20–26. PMID: [9422587](#)
12. Yao SY, Luo L, Har KJ, Becker A, Ruberg S, Yu G, et al. (2004) *Sinorhizobium meliloti* ExoR and ExoS proteins regulate both succinoglycan and flagellum production. *J Bacteriol* 186: 6042–6049. PMID: [15342573](#)
13. Wells DH, Chen EJ, Fisher RF, Long SR (2007) ExoR is genetically coupled to the ExoS-ChvI two-component system and located in the periplasm of *Sinorhizobium meliloti*. *Mol Microbiol* 64: 647–664. PMID: [17462014](#)
14. Belanger L, Dimmick KA, Fleming JS, Charles TC (2009) Null mutations in *Sinorhizobium meliloti* *exoS* and *chvI* demonstrate the importance of this two-component regulatory system for symbiosis. *Mol Microbiol* 74: 1223–1237. doi: [10.1111/j.1365-2958.2009.06931.x](#) PMID: [19843226](#)
15. Lu HY, Cheng HP (2010) Autoregulation of *Sinorhizobium meliloti* *exoR* gene expression. *Microbiology* 156: 2092–2101. doi: [10.1099/mic.0.038547-0](#) PMID: [20413557](#)
16. Lu HY, Luo L, Yang MH, Cheng HP (2012) *Sinorhizobium meliloti* ExoR is the target of periplasmic proteolysis. *J Bacteriol* 194: 4029–4040. doi: [10.1128/JB.00313-12](#) PMID: [22636773](#)
17. Wiech EM, Cheng HP, Singh SM (2014) Molecular modeling and computational analyses suggests that the *Sinorhizobium meliloti* periplasmic regulator protein ExoR adopts a superhelical fold and is controlled by a unique mechanism of proteolysis. *Protein Sci* 24: 319–327. doi: [10.1002/pro.2616](#) PMID: [25492513](#)
18. Chen EJ, Sabio EA, Long SR (2008) The periplasmic regulator ExoR inhibits ExoS/ChvI two-component signaling in *Sinorhizobium meliloti*. *Mol Microbiol* 69: 1290–1303. doi: [10.1111/j.1365-2958.2008.06362.x](#) PMID: [18631237](#)
19. Osteras M, Stanley J, Finan TM (1995) Identification of *Rhizobium*-specific intergenic mosaic elements within an essential two-component regulatory system of *Rhizobium* species. *J Bacteriol* 177: 5485–5494. PMID: [7559334](#)
20. Wang C, Kemp J, Da Fonseca IO, Equi RC, Sheng X, Charles TC, et al. (2010) *Sinorhizobium meliloti* 1021 Loss-of-Function Deletion Mutation in *chvI* and Its Phenotypic Characteristics. *Mol Plant Microbe Interact* 23: 153–160. doi: [10.1094/MPMI-23-2-0153](#) PMID: [20064059](#)
21. Wang LC, Morgan LK, Godakumbura P, Kenney LJ, Anand GS (2012) The inner membrane histidine kinase EnvZ senses osmolality via helix-coil transitions in the cytoplasm. *Embo J* 31: 2648–2659. doi: [10.1038/emboj.2012.99](#) PMID: [22543870](#)
22. Lu H-Y, Cheng H-P (2012) *Sinorhizobium meliloti* is the target of periplasmic proteolysis. *J Bacteriol* In press.
23. Bahlawane C, McIntosh M, Krol E, Becker A (2008) *Sinorhizobium meliloti* regulator MucR couples exopolysaccharide synthesis and motility. *Mol Plant Microbe Interact* 21: 1498–1509. doi: [10.1094/MPMI-21-11-1498](#) PMID: [18842098](#)
24. Keating DH (2007) The *Sinorhizobium meliloti* ExoR protein is required for the downregulation of *lpsS* transcription and succinoglycan biosynthesis in response to divalent cations. *FEMS Microbiol Lett* 267: 23–29. PMID: [17233674](#)
25. Chen EJ, Fisher RF, Perovich VM, Sabio EA, Long SR (2009) Identification of direct transcriptional target genes of ExoS/ChvI two-component signaling in *Sinorhizobium meliloti*. *J Bacteriol* 191: 6833–6842. doi: [10.1128/JB.00734-09](#) PMID: [19749054](#)
26. Reeve WG, Dilworth MJ, Tiwari RP, Glenn AR (1997) Regulation of exopolysaccharide production in *Rhizobium leguminosarum biovar viciae* WSM710 involves *exoR*. *Microbiology* 143 (Pt 6): 1951–1958. PMID: [9202471](#)
27. Reed JW, Glazebrook J, Walker GC (1991) The *exoR* gene of *Rhizobium meliloti* affects RNA levels of other *exo* genes but lacks homology to known transcriptional regulators. *J Bacteriol* 173: 3789–3794. PMID: [1711027](#)

28. Doherty D, Leigh JA, Glazebrook J, Walker GC (1988) *Rhizobium meliloti* mutants that overproduce the *Rhizobium meliloti* acidic calcofluor-binding exopolysaccharide. *J Bacteriol* 170: 4249–4256. PMID: [2842307](#)
29. Cheng HP, Yao SY (2004) The key *Sinorhizobium meliloti* succinoglycan biosynthesis gene *exoY* is expressed from two promoters. *FEMS Microbiol Lett* 231: 131–136. PMID: [14769477](#)
30. Tomlinson AD, Ramey-Hartung B, Day TW, Merritt PM, Fuqua C (2010) *Agrobacterium tumefaciens* ExoR represses succinoglycan biosynthesis and is required for biofilm formation and motility. *Microbiology* 156: 2670–2681. doi: [10.1099/mic.0.039032-0](#) PMID: [20576688](#)
31. Wu CF, Lin JS, Shaw GC, Lai EM (2012) Acid-induced type VI secretion system is regulated by ExoR-ChvG/ChvI signaling cascade in *Agrobacterium tumefaciens*. *PLoS Pathog* 8: e1002938. doi: [10.1371/journal.ppat.1002938](#) PMID: [23028331](#)
32. Heckel BC, Tomlinson AD, Morton ER, Choi JH, Fuqua C (2014) *Agrobacterium tumefaciens* *exoR* controls acid response genes and impacts exopolysaccharide synthesis, horizontal gene transfer, and virulence gene expression. *J Bacteriol* 196: 3221–3233. doi: [10.1128/JB.01751-14](#) PMID: [24982308](#)
33. Charles TC, Nester EW (1993) A chromosomally encoded two-component sensory transduction system is required for virulence of *Agrobacterium tumefaciens*. *J Bacteriol* 175: 6614–6625. PMID: [8407839](#)
34. Quebatte M, Dehio M, Tropel D, Basler A, Toller I, Raddatz G, et al. (2010) The BatR/BatS two-component regulatory system controls the adaptive response of *Bartonella henselae* during human endothelial cell infection. *J Bacteriol* 192: 3352–3367. doi: [10.1128/JB.01676-09](#) PMID: [20418395](#)
35. Foreman DL, Vanderlinde EM, Bay DC, Yost CK (2009) Characterization of a gene family of outer membrane proteins (*ropB*) in *Rhizobium leguminosarum* *bv. viciae* VF39SM and the role of the sensor kinase ChvG in their regulation. *J Bacteriol* 192: 975–983. doi: [10.1128/JB.01140-09](#) PMID: [20023026](#)
36. Guzman-Verri C, Manterola L, Sola-Landa A, Parra A, Cloeckeaert A, Garin J, et al. (2002) The two-component system BvrR/BvrS essential for *Brucella abortus* virulence regulates the expression of outer membrane proteins with counterparts in members of the Rhizobiaceae. *Proc Natl Acad Sci U S A* 99: 12375–12380. PMID: [12218183](#)
37. Sola-Landa A, Pizarro-Cerda J, Grillo MJ, Moreno E, Moriyon I, Blasco J, et al. (1998) A two-component regulatory system playing a critical role in plant pathogens and endosymbionts is present in *Brucella abortus* and controls cell invasion and virulence. *Mol Microbiol* 29: 125–138. PMID: [9701808](#)
38. Mantis NJ, Winans SC (1993) The chromosomal response regulatory gene *chvI* of *Agrobacterium tumefaciens* complements and *Escherichia coli* *phoB* mutation and is required for virulence. *J Bacteriol* 175: 6626–6636. PMID: [8407840](#)
39. Li L, Jia Y, Hou Q, Charles TC, Nester EW, Pan SQ, et al. (2002) A global pH sensor: *Agrobacterium* sensor protein ChvG regulates acid-inducible genes on its two chromosomes and Ti plasmid. *Proc Natl Acad Sci U S A* 99: 12369–12374. PMID: [12218184](#)
40. Yuan ZC, Liu P, Saenkham P, Kerr K, Nester EW (2008) Transcriptome profiling and functional analysis of *Agrobacterium tumefaciens* reveals a general conserved response to acidic conditions (pH 5.5) and a complex acid-mediated signaling involved in *Agrobacterium*-plant interactions. *J Bacteriol* 190: 494–507. PMID: [17993523](#)
41. Vanderlinde EM, Yost CK (2011) Mutation of the sensor kinase *chvG* in *Rhizobium leguminosarum* negatively impacts cellular metabolism, outer membrane stability, and symbiosis. *J Bacteriol*.
42. Viadas C, Rodriguez MC, Sangari FJ, Gorvel JP, Garcia-Lobo JM, Lopez-Goni I (2010) Transcriptome analysis of the *Brucella abortus* BvrR/BvrS two-component regulatory system. *PLoS One* 5: e10216. doi: [10.1371/journal.pone.0010216](#) PMID: [20422049](#)
43. Manterola L, Guzman-Verri C, Chaves-Olarte E, Barquero-Calvo E, de Miguel MJ, Moriyon I, et al. (2007) BvrR/BvrS-controlled outer membrane proteins Omp3a and Omp3b are not essential for *Brucella abortus* virulence. *Infect Immun* 75: 4867–4874. PMID: [17664262](#)
44. Manterola L, Moriyon I, Moreno E, Sola-Landa A, Weiss DS, Koch MHJ, et al. (2005) The lipopolysaccharide of *Brucella abortus* BvrS/BvrR mutants contains lipid A modifications and has higher affinity for bactericidal cationic peptides. *J Bacteriol* 187: 5631–5639. PMID: [16077108](#)
45. Lopez-Goni I, Guzman-Verri C, Manterola L, Sola-Landa A, Moriyon I, Moreno E (2002) Regulation of *Brucella virulence* by the two-component system BvrR/BvrS. *Vet Microbiol* 90: 329–339. PMID: [12414153](#)
46. Jothi R, Przytycka TM, Aravind L (2007) Discovering functional linkages and uncharacterized cellular pathways using phylogenetic profile comparisons: a comprehensive assessment. *BMC Bioinformatics* 8: 173. PMID: [17521444](#)

47. Pellegrini M, Marcotte EM, Yeates TO (1999) A fast algorithm for genome-wide analysis of proteins with repeated sequences. *Proteins* 35: 440–446. PMID: [10382671](#)
48. Goh CS, Bogan AA, Joachimiak M, Walther D, Cohen FE (2000) Co-evolution of proteins with their interaction partners. *J Mol Biol* 299: 283–293. PMID: [10860738](#)
49. Kanehisa M, Goto S (2000) KEGG: kyoto encyclopedia of genes and genomes. *Nucleic Acids Res* 28: 27–30. PMID: [10592173](#)
50. Kanehisa M, Goto S, Sato Y, Furumichi M, Tanabe M (2011) KEGG for integration and interpretation of large-scale molecular data sets. *Nucleic Acids Res* 40: D109–114. doi: [10.1093/nar/gkr988](#) PMID: [22080510](#)
51. Finn RD, Clements J, Eddy SR (2011) HMMER web server: interactive sequence similarity searching. *Nucleic Acids Res* 39: W29–37. doi: [10.1093/nar/gkr367](#) PMID: [21593126](#)
52. Benson DA, Karsch-Mizrachi I, Lipman DJ, Ostell J, Sayers EW (2010) GenBank. *Nucleic Acids Res* 39: D32–37. doi: [10.1093/nar/gkq1079](#) PMID: [21071399](#)
53. Angermuller C, Biegert A, Soding J (2012) Discriminative modelling of context-specific amino acid substitution probabilities. *Bioinformatics* 28: 3240–3247. doi: [10.1093/bioinformatics/bts622](#) PMID: [23080114](#)
54. Biegert A, Soding J (2009) Sequence context-specific profiles for homology searching. *Proc Natl Acad Sci U S A* 106: 3770–3775. doi: [10.1073/pnas.0810767106](#) PMID: [19234132](#)
55. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ (1990) Basic local alignment search tool. *J Mol Biol* 215: 403–410. PMID: [2231712](#)
56. Markowitz VM, Korzeniewski F, Palaniappan K, Szeto E, Werner G, Padki A, et al. (2006) The integrated microbial genomes (IMG) system. *Nucleic Acids Res* 34: D344–348. PMID: [16381883](#)
57. Yang Y, Faraggi E, Zhao H, Zhou Y (2011) Improving protein fold recognition and template-based modeling by employing probabilistic-based matching between predicted one-dimensional structural properties of query and corresponding native properties of templates. *Bioinformatics* 27: 2076–2082. doi: [10.1093/bioinformatics/btr350](#) PMID: [21666270](#)
58. Vallat BK, Pillardy J, Majek P, Meller J, Blom T, Cao B, et al. (2009) Building and assessing atomic models of proteins from structural templates: learning and benchmarks. *Proteins* 76: 930–945. doi: [10.1002/prot.22401](#) PMID: [19326457](#)
59. Vallat BK, Pillardy J, Elber R (2008) A template-finding algorithm and a comprehensive benchmark for homology modeling of proteins. *Proteins* 72: 910–928. doi: [10.1002/prot.21976](#) PMID: [18300226](#)
60. Teodorescu O, Galor T, Pillardy J, Elber R (2004) Enriching the sequence substitution matrix by structural information. *Proteins* 54: 41–48. PMID: [14705022](#)
61. Tobi D, Elber R (2000) Distance-dependent, pair potential for protein folding: results from linear optimization. *Proteins* 41: 40–46. PMID: [10944392](#)
62. Meller J, Elber R (2001) Linear programming optimization and a double statistical filter for protein threading protocols. *Proteins* 45: 241–261. PMID: [11599028](#)
63. Soding J, Biegert A, Lupas AN (2005) The HHpred interactive server for protein homology detection and structure prediction. *Nucleic Acids Res* 33: W244–248. PMID: [15980461](#)
64. Lobley A, Sadowski MI, Jones DT (2009) pGenTHREADER and pDomTHREADER: new methods for improved protein fold recognition and superfamily discrimination. *Bioinformatics* 25: 1761–1767. doi: [10.1093/bioinformatics/btp302](#) PMID: [19429599](#)
65. Schultz J, Milpetz F, Bork P, Ponting CP (1998) SMART, a simple modular architecture research tool: identification of signaling domains. *Proc Natl Acad Sci U S A* 95: 5857–5864. PMID: [9600884](#)
66. Letunic I, Doerks T, Bork P (2012) SMART 7: recent updates to the protein domain annotation resource. *Nucleic Acids Res* 40: D302–305. doi: [10.1093/nar/gkr931](#) PMID: [22053084](#)
67. Punta M, Coggill PC, Eberhardt RY, Mistry J, Tate J, Boursnell C, et al. (2012) The Pfam protein families database. *Nucleic Acids Res* 40: D290–301. doi: [10.1093/nar/gkr1065](#) PMID: [22127870](#)
68. Marchler-Bauer A, Zheng C, Chitsaz F, Derbyshire MK, Geer LY, Gonzales NR, et al. (2013) CDD: conserved domains and protein three-dimensional structure. *Nucleic Acids Res* 41: D348–352. doi: [10.1093/nar/gks1243](#) PMID: [23197659](#)
69. Tusnady GE, Simon I (1998) Principles governing amino acid composition of integral membrane proteins: application to topology prediction. *J Mol Biol* 283: 489–506. PMID: [9769220](#)
70. Viklund H, Elofsson A (2008) OCTOPUS: improving topology prediction by two-track ANN-based preference scores and an extended topological grammar. *Bioinformatics* 24: 1662–1668. doi: [10.1093/bioinformatics/btn221](#) PMID: [18474507](#)
71. Kall L, Krogh A, Sonnhammer EL (2004) A combined transmembrane topology and signal peptide prediction method. *J Mol Biol* 338: 1027–1036. PMID: [15111065](#)

72. Petersen TN, Brunak S, von Heijne G, Nielsen H (2011) SignalP 4.0: discriminating signal peptides from transmembrane regions. *Nat Methods* 8: 785–786. doi: [10.1038/nmeth.1701](https://doi.org/10.1038/nmeth.1701) PMID: [21959131](https://pubmed.ncbi.nlm.nih.gov/21959131/)
73. Yu NY, Wagner JR, Laird MR, Melli G, Rey S, Lo R, et al. (2010) PSORTb 3.0: improved protein sub-cellular localization prediction with refined localization subcategories and predictive capabilities for all prokaryotes. *Bioinformatics* 26: 1608–1615. doi: [10.1093/bioinformatics/btq249](https://doi.org/10.1093/bioinformatics/btq249) PMID: [20472543](https://pubmed.ncbi.nlm.nih.gov/20472543/)
74. Gomi M, Sawada R, Sonoyama M, Mitaku S (2005) Comparative proteomics of the prokaryota using secretory proteins. *Chem-Bio Infomatics J* 5: 56–64.
75. Bendtsen JD, Kiemer L, Fausboll A, Brunak S (2005) Non-classical protein secretion in bacteria. *BMC Microbiol* 5: 58. PMID: [16212653](https://pubmed.ncbi.nlm.nih.gov/16212653/)
76. Di Tommaso P, Moretti S, Xenarios I, Orobítz M, Montanyola A, Chang JM, et al. (2011) T-Coffee: a web server for the multiple sequence alignment of protein and RNA sequences using structural information and homology extension. *Nucleic Acids Res* 39: W13–17. doi: [10.1093/nar/gkr245](https://doi.org/10.1093/nar/gkr245) PMID: [21558174](https://pubmed.ncbi.nlm.nih.gov/21558174/)
77. Waterhouse AM, Procter JB, Martin DM, Clamp M, Barton GJ (2009) Jalview Version 2—a multiple sequence alignment editor and analysis workbench. *Bioinformatics* 25: 1189–1191. doi: [10.1093/bioinformatics/btp033](https://doi.org/10.1093/bioinformatics/btp033) PMID: [19151095](https://pubmed.ncbi.nlm.nih.gov/19151095/)
78. Nicholas KB, Nicholas HB, Deerfield DW (1997) GeneDoc: analysis and visualization of genetic variation. *Embnew news* 4: 14.
79. Capella-Gutierrez S, Silla-Martinez JM, Gabaldon T (2009) trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. *Bioinformatics* 25: 1972–1973. doi: [10.1093/bioinformatics/btp348](https://doi.org/10.1093/bioinformatics/btp348) PMID: [19505945](https://pubmed.ncbi.nlm.nih.gov/19505945/)
80. Mintseris J, Weng Z (2005) Structure, function, and evolution of transient and obligate protein-protein interactions. *Proc Natl Acad Sci U S A* 102: 10930–10935. PMID: [16043700](https://pubmed.ncbi.nlm.nih.gov/16043700/)
81. Mintseris J, Wiehe K, Pierce B, Anderson R, Chen R, Janin J, et al. (2005) Protein-Protein Docking Benchmark 2.0: an update. *Proteins* 60: 214–216. PMID: [15981264](https://pubmed.ncbi.nlm.nih.gov/15981264/)
82. Abascal F, Zardoya R, Posada D (2005) ProtTest: selection of best-fit models of protein evolution. *Bioinformatics* 21: 2104–2105. PMID: [15647292](https://pubmed.ncbi.nlm.nih.gov/15647292/)
83. Stamatakis A, Hoover P, Rougemont J (2008) A rapid bootstrap algorithm for the RAxML Web servers. *Syst Biol* 57: 758–771. doi: [10.1080/10635150802429642](https://doi.org/10.1080/10635150802429642) PMID: [18853362](https://pubmed.ncbi.nlm.nih.gov/18853362/)
84. Stamatakis A, Ott M, Ludwig T (2005) RAxML-OMP: an efficient program for phylogenetic inference on SMPs. *Lecture Notes in Computer Science* 3606: 4.
85. Miller MA, Pfeiffer W, Schwartz T. Creating the CIPRES Science Gateway for inference of large phylogenetic trees; 2010; New Orleans, LA. pp. 1–8.
86. Sukumaran J, Holder MT (2010) DendroPy: a Python library for phylogenetic computing. *Bioinformatics* 26: 1569–1571. doi: [10.1093/bioinformatics/btq228](https://doi.org/10.1093/bioinformatics/btq228) PMID: [20421198](https://pubmed.ncbi.nlm.nih.gov/20421198/)
87. Huelsenbeck JP, Ronquist F, Nielsen R, Bollback JP (2001) Bayesian inference of phylogeny and its impact on evolutionary biology. *Science* 294: 2310–2314. PMID: [11743192](https://pubmed.ncbi.nlm.nih.gov/11743192/)
88. Altekar G, Dwarkadas S, Huelsenbeck JP, Ronquist F (2004) Parallel Metropolis coupled Markov chain Monte Carlo for Bayesian phylogenetic inference. *Bioinformatics* 20: 407–415. PMID: [14960467](https://pubmed.ncbi.nlm.nih.gov/14960467/)
89. Huelsenbeck JP, Larget B, Miller RE, Ronquist F (2002) Potential applications and pitfalls of Bayesian inference of phylogeny. *Syst Biol* 51: 673–688. PMID: [12396583](https://pubmed.ncbi.nlm.nih.gov/12396583/)
90. Stajich JE, Block D, Boulez K, Brenner SE, Chervitz SA, Dagdigian C, et al. (2002) The Bioperl toolkit: Perl modules for the life sciences. *Genome Res* 12: 1611–1618. PMID: [12368254](https://pubmed.ncbi.nlm.nih.gov/12368254/)
91. Wu D, Jospin G, Eisen JA (2013) Systematic identification of gene families for use as "markers" for phylogenetic and phylogeny-driven ecological studies of bacteria and archaea and their major sub-groups. *PLoS One* 8: e77033. doi: [10.1371/journal.pone.0077033](https://doi.org/10.1371/journal.pone.0077033) PMID: [24146954](https://pubmed.ncbi.nlm.nih.gov/24146954/)
92. Williams KP, Sobral BW, Dickerman AW (2007) A robust species tree for the alphaproteobacteria. *J Bacteriol* 189: 4578–4586. PMID: [17483224](https://pubmed.ncbi.nlm.nih.gov/17483224/)
93. Gupta RS, Mok A (2007) Phylogenomics and signature proteins for the alpha proteobacteria and its main groups. *BMC Microbiol* 7: 106. PMID: [18045498](https://pubmed.ncbi.nlm.nih.gov/18045498/)
94. Pupko T, Bell RE, Mayrose I, Glaser F, Ben-Tal N (2002) Rate4Site: an algorithmic tool for the identification of functional regions in proteins by surface mapping of evolutionary determinants within their homologues. *Bioinformatics* 18 Suppl 1: S71–77. PMID: [12169533](https://pubmed.ncbi.nlm.nih.gov/12169533/)
95. Alm E, Huang K, Arkin A (2006) The evolution of two-component systems in bacteria reveals different strategies for niche adaptation. *PLoS Comput Biol* 2: e143. PMID: [17083272](https://pubmed.ncbi.nlm.nih.gov/17083272/)
96. Belanger L, Charles TC (2013) Members of the *Sinorhizobium meliloti* ChvI regulon identified by a DNA binding screen. *BMC Microbiol* 13: 132. doi: [10.1186/1471-2180-13-132](https://doi.org/10.1186/1471-2180-13-132) PMID: [23758731](https://pubmed.ncbi.nlm.nih.gov/23758731/)

97. Foster JT, Beckstrom-Sternberg SM, Pearson T, Beckstrom-Sternberg JS, Chain PS, Roberto FF, et al. (2009) Whole-genome-based phylogeny and divergence of the genus *Brucella*. *J Bacteriol* 191: 2864–2870. doi: [10.1128/JB.01581-08](https://doi.org/10.1128/JB.01581-08) PMID: [19201792](https://pubmed.ncbi.nlm.nih.gov/19201792/)
98. Palacios L, Arahall DR, Reguera B, Marin I (2006) *Hoeflea alexandrii* sp. nov., isolated from the toxic dinoflagellate *Alexandrium minutum* AL1V. *Int J Syst Evol Microbiol* 56: 1991–1995. PMID: [16902042](https://pubmed.ncbi.nlm.nih.gov/16902042/)
99. Fiebig A, Pradella S, Petersen J, Michael V, Pauker O, Rohde M, et al. Genome of the marine alpha-proteobacterium *Hoeflea phototrophica* type strain (DFL-43(T)). *Stand Genomic Sci* 7: 440–448. doi: [10.4056/sigs.3486982](https://doi.org/10.4056/sigs.3486982) PMID: [24019991](https://pubmed.ncbi.nlm.nih.gov/24019991/)
100. Cheng HP, Walker GC (1998) Succinoglycan is required for initiation and elongation of infection threads during nodulation of alfalfa by *Rhizobium meliloti*. *J Bacteriol* 180: 5183–5191. PMID: [9748453](https://pubmed.ncbi.nlm.nih.gov/9748453/)
101. Koretke KK, Lupas AN, Warren PV, Rosenberg M, Brown JR (2000) Evolution of two-component signal transduction. *Mol Biol Evol* 17: 1956–1970. PMID: [11110912](https://pubmed.ncbi.nlm.nih.gov/11110912/)
102. Gross R, Beier D, editors (2012) *Two-Component Systems in Bacteria*. Caister Academic Press.
103. Vogt SL, Nevesinjac AZ, Humphries RM, Donnenberg MS, Armstrong GD, Raivio TL (2010) The Cpx envelope stress response both facilitates and inhibits elaboration of the enteropathogenic *Escherichia coli* bundle-forming pilus. *Mol Microbiol* 76: 1095–1110. doi: [10.1111/j.1365-2958.2010.07145.x](https://doi.org/10.1111/j.1365-2958.2010.07145.x) PMID: [20444097](https://pubmed.ncbi.nlm.nih.gov/20444097/)
104. Ortiz de Orue Lucana D, Groves MR (2009) The three-component signalling system HbpS-SenS-SenR as an example of a redox sensing pathway in bacteria. *Amino Acids* 37: 479–486. doi: [10.1007/s00726-009-0260-9](https://doi.org/10.1007/s00726-009-0260-9) PMID: [19259771](https://pubmed.ncbi.nlm.nih.gov/19259771/)