

# '*Ca. Liberibacter asiaticus*' Proteins Orthologous with *pSymA*-Encoded Proteins of *Sinorhizobium meliloti*: Hypothetical Roles in Plant Host Interaction

L. David Kuykendall, Jonathan Y. Shao, John S. Hartung\*

United States Department of Agriculture, Agricultural Research Service, Molecular Plant Pathology Laboratory, Beltsville, Maryland, United States of America

## Abstract

*Sinorhizobium meliloti* strain 1021, a nitrogen-fixing, root-nodulating bacterial microsymbiont of alfalfa, has a 3.5 Mbp circular chromosome and two megaplasmids including 1.3 Mbp *pSymA* carrying nonessential 'accessory' genes for nitrogen fixation (*nif*), nodulation and host specificity (*nod*). A related bacterium, psyllid-vectored '*Ca. Liberibacter asiaticus*,' is an obligate phytopathogen with a reduced genome that was previously analyzed for genes orthologous to genes on the *S. meliloti* circular chromosome. In general, proteins encoded by *pSymA* genes are more similar in sequence alignment to those encoded by *S. meliloti* chromosomal orthologs than to orthologous proteins encoded by genes carried on the '*Ca. Liberibacter asiaticus*' genome. Only two '*Ca. Liberibacter asiaticus*' proteins were identified as having orthologous proteins encoded on *pSymA* but not also encoded on the chromosome of *S. meliloti*. These two orthologous gene pairs encode a Na<sup>+</sup>/K<sup>+</sup> antiporter (shared with intracellular pathogens of the family *Bartonellaceae*) and a Co<sup>++</sup>, Zn<sup>++</sup> and Cd<sup>++</sup> cation efflux protein that is shared with the phytopathogen *Agrobacterium*. Another shared protein, a redox-regulated K<sup>+</sup> efflux pump may regulate cytoplasmic pH and homeostasis. The *pSymA* and '*Ca. Liberibacter asiaticus*' orthologs of the latter protein are more highly similar in amino acid alignment compared with the alignment of the *pSymA*-encoded protein with its *S. meliloti* chromosomal homolog. About 182 *pSymA* encoded proteins have sequence similarity ( $\leq E-10$ ) with '*Ca. Liberibacter asiaticus*' proteins, often present as multiple orthologs of single '*Ca. Liberibacter asiaticus*' proteins. These proteins are involved with amino acid uptake, cell surface structure, chaperonins, electron transport, export of bioactive molecules, cellular homeostasis, regulation of gene expression, signal transduction and synthesis of amino acids and metabolic cofactors. The presence of multiple orthologs defies mutational analysis and is consistent with the hypothesis that these proteins may be of particular importance in host/microbe interaction and their duplication likely facilitates their ongoing evolution.

**Citation:** Kuykendall LD, Shao JY, Hartung JS (2012) '*Ca. Liberibacter asiaticus*' Proteins Orthologous with *pSymA*-Encoded Proteins of *Sinorhizobium meliloti*: Hypothetical Roles in Plant Host Interaction. PLoS ONE 7(6): e38725. doi:10.1371/journal.pone.0038725

**Editor:** Paulo Lee Ho, Instituto Butantan, Brazil

**Received:** February 1, 2012; **Accepted:** May 14, 2012; **Published:** June 25, 2012

This is an open-access article, free of all copyright, and may be freely reproduced, distributed, transmitted, modified, built upon, or otherwise used by anyone for any lawful purpose. The work is made available under the Creative Commons CC0 public domain dedication.

**Funding:** No current external funding sources for this study.

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

\* E-mail: John.Hartung@ars.usda.gov

## Introduction

Within the order Rhizobiales of the class alpha-Proteobacteria, *Sinorhizobium meliloti* is a member of a diverse bacterial family, *Rhizobiaceae* [1–3]. Other genera contained within the *Rhizobiaceae* and indeed the *Rhizobium* type genus [3] include *Agrobacterium* and *Rhizobium* [4]. A growing number of species of the family *Rhizobiaceae* and of the order *Rhizobiales* are becoming recognized based on 16 S RNA gene sequence data. Many species interact with plants and other eukaryotes to produce important agricultural, economic and/or environmental consequences. '*Ca. Liberibacter asiaticus*' is recognized as the causal agent of huanglongbing (HLB), also known as citrus greening disease. This phytopathogen was introduced into the New World from Asia [5,6], and is transmitted by the citrus psyllid, *Diaphorina citri* [7]. The full genomic sequence of '*Ca. Liberibacter asiaticus*' was determined by deep sequencing of total DNA obtained from a single citrus psyllid which contained more than 10<sup>8</sup> cells of '*Ca. Liberibacter asiaticus*.' Phylogenetic analysis indicated that '*Ca. Liberibacter asiaticus*' and *S. meliloti* are close bacterial relatives [2]. '*Ca. Liberibacter asiaticus*' has a single small chromosome of

1.23 Mbps with 36.5% GC content compared to the chromosome of *S. meliloti* which is 3.65 Mbps and 62.7% GC [2,8]. '*Ca. Liberibacter asiaticus*' is not available in culture but can colonize intracellularly and move systemically in the phloem vessels of citrus trees [9,10] following introduction by the citrus psyllid, *D. citri*. '*Ca. Liberibacter asiaticus*' also systemically and intracellularly colonizes citrus psyllid salivary glands [11]. The genome of '*Ca. Liberibacter asiaticus*' has many adaptations to facilitate an intracellular lifestyle in both plant and insect cells [2,12].

*S. meliloti* 1021, free living in the soil, also lives in specialized root nodules in symbiosis with alfalfa, *Medicago sativa* L., where *S. meliloti* reduces atmospheric nitrogen and thereby confers plant nitrogen sufficiency. *S. meliloti* 1021 has a composite genome comprised of the chromosome and two megaplasmids, *pSymA* and *pSymB*, of 1.3 Mbps and 1.7 Mbps [8]. Megaplasmids *pSymA* and *pSymB* encode 1293 and 1570 protein-coding genes, respectively [8,13].

Predicted gene products of all open-reading frames (ORFs) in the genome of '*Ca. Liberibacter asiaticus*' were compared by protein BLAST against all of the predicted gene products encoded by chromosomal genes of *S. meliloti* 1021 [2]. However, predicted

gene products of all ORFs encoded on the plant symbiosis-controlling *S. meliloti* megaplasmid *pSymA* had not been previously subjected to protein BLAST compared with the predicted gene products encoded in the ‘*Ca. Liberibacter asiaticus*’ genome. As an “accessory” genome component [14], *pSymA* may have originated from an ancestral plasmid and may have maintained individual and some small blocks of genes originally from host chromosomes. It has long been known that the *pSymA* megaplasmid of *S. meliloti*, in contrast to the circular chromosome or the other *S. meliloti* megaplasmid *pSymB*, specializes in carrying genes essential for establishing and maintaining intimate intracellular plant interactions with alfalfa. The megaplasmid *pSymA* itself is self-transmissible and upon transfer confers nodulation (*nod*), nitrogen fixation (*nif*) and other symbiotic abilities in related bacteria with a similar genetic background or complement of chromosomal genes [15–17]. The chromosome of the tumor-inducing bacterial phytopathogen *Agrobacterium tumefaciens* encodes characteristics practically identical to those of *Rhizobium* microsymbionts, and members of the species also maintain pTi as an accessory genomic component. When bacterial cells recruit or accept transfer of the Ti plasmid or *pSymA*, either tumor-inducing pathogenicity or nitrogen-fixing symbiotic ability, respectively, is conferred on the host bacterium. *Agrobacterium* and *Rhizobium* were combined into a single genus within the Rhizobiaceae largely for this reason and to reject taxonomy based largely on plant interactions determined by extrachromosomal elements, i.e., to avoid plasmid-based nomenclature [4].

As *S. meliloti* is a close phylogenetic relative of ‘*Ca. Liberibacter asiaticus*’, it is likely that the two bacteria deploy a similar repertoire of mechanisms for avoiding defenses elicited in host plant cells by their invasion, or, in the case of beneficial root nodule bacteria, recruitment or “welcomed entry” [18]. The goal of the present study was to compare predicted gene products encoded by the genome of ‘*Ca. Liberibacter asiaticus*’ with those encoded by genes carried on the megaplasmid *pSymA* of *S. meliloti* strain 1021. The rationale for identifying protein products of ‘*Ca. Liberibacter asiaticus*’ genes orthologous to those encoded by genes on *pSymA* is simply that such genes and their protein products may play direct or indirect role(s) in establishing and maintaining an intimate intracellular interaction with eukaryotes. We have identified a number of such interesting protein products.

## Results

### Maintenance of Homeostasis

In general, proteins encoded by *pSymA* genes are more similar in sequence alignment to those encoded by *S. meliloti* chromosomal orthologs than to orthologous proteins encoded by genes carried on the ‘*Ca. Liberibacter asiaticus*’ genome. There are two very notable exceptions where orthologs are shared only by *pSymA* and ‘*Ca. Liberibacter asiaticus*’ and not with the chromosome of *S. meliloti*. YP\_003064907, the predicted protein product of a ‘*Ca. Liberibacter asiaticus*’ gene, is orthologous to NP\_435608.2 ( $E = 5e-71$ ) encoded by *pSymA* yet lacks an orthologous protein encoded by the chromosome of *S. meliloti* (Table 1). SMART analysis predicts that YP\_003064907, has a cation efflux domain ( $E = 1.3e-31$ ) from residues 15 to 313 of the 313 amino acid protein. This predicted protein also has BLAST matches with cobalt, zinc and cadmium efflux proteins (COG3965) [19]. *pSymA*-encoded NP\_436297.2 is orthologous to YP\_003064775 ( $E = 3e-30$ ) encoded on the chromosome of ‘*Ca. Liberibacter asiaticus*’, but neither has an ortholog encoded by the *S. meliloti* chromosome (Table 1). SMART analysis of ‘*Ca. Liberibacter asiaticus*’ YP\_003064775, annotated as a *nhaA* gene, predicts a sodium/

proton antiporter domain ( $E = 2.1e-113$ ) at residues 4 to 379 of the 381 amino acid protein [20,21]. These two protein pairs are the only two protein pairs uniquely shared by *pSymA* and ‘*Ca. Liberibacter asiaticus*’. A third orthologous protein pair comprised of NP\_436118.1 (*pSymA*) and YP\_003065337 (‘*Ca. Liberibacter asiaticus*’) is much more similar in amino acid sequence ( $E = 1e-134$ ) than is NP\_436118.1 and its ortholog encoded by the *S. meliloti* chromosome, NP\_384912 ( $E = 1e-22$ ). Bioinformatic analyses revealed that YP\_003065337 is a potassium efflux protein. Control of intracellular pH is ascribed in its domain architecture. SMART analysis of YP\_003065337 confidently predicts two domains as follows: an amino-terminal sodium/proton exchanger domain ( $E = 2.7e-72$ ) from residues 4 to 394 and a non-overlapping TrkA domain ( $E = 8.3e-25$ ) for redox regulation, from amino acids 451–566 of the 609 amino acid protein. ‘*Ca. Liberibacter asiaticus*’ YP\_003065245, a potassium uptake protein, has  $E = 0$  alignments with both *pSymA* (NP\_436237.1) and *S. meliloti* chromosomal orthologs (Table 1). These proteins likely contribute to the maintenance of intracellular homeostasis through the regulation of pH and the control of mono- and divalent cation concentrations.

### Signal Transduction

There are several orthologous proteins encoded by genes shared between *pSymA* and the ‘*Ca. Liberibacter asiaticus*’ and *S. meliloti* chromosomes that are predicted to function in sensing and signal transduction. For example, YP\_003064826, a PAS/PAC sensor signal transduction kinase has a *pSymA* ortholog and a *S. meliloti* chromosomal ortholog. Several two-component response regulators or sensor histidine kinases are encoded both by genes on *pSymA* and the genome ‘*Ca. Liberibacter asiaticus*’, and include YP\_003065306, which has a characteristic N-terminal CheY-like sensor and Trans\_reg\_C which are DNA-binding transcriptional activator/effector domains. YP\_003064863, a two component sensor histidine kinase/response regulator hybrid protein, is a FixL-like oxygen-controlled histidine kinase homologous to several *pSymA*-encoded proteins ( $E = 3e-49$  for the most similar ortholog) (Table 1).

### ABC-cassette Type Amino Acid Transporters

ATP Binding Cassette type amino acid transporters comprise the largest class of genes shared by ‘*Ca. Liberibacter asiaticus*’, *pSymA* and the *S. meliloti* chromosome (Table 2). For most ABC-type transporter proteins encoded by ‘*Ca. Liberibacter asiaticus*’ genes there are multiple homologs on *pSymA*, dispersed over the megaplasmid (Table S1). YP\_003064586 is an example of an ABC-type transporter protein from ‘*Ca. Liberibacter asiaticus*’. In addition to NP\_435288.1 there are 10 more *pSymA*-encoded proteins homologous to YP\_003064586 ( $E = 4e-67$  up to  $1e-14$ ), and all are annotated as general amino acid transporters with an ATP binding site and other characteristics of an ABC type amino acid transporter. The corresponding genes are dispersed across the *pSymA* (Fig. 1A; Table 2).

YP\_003064753 is a proline/glycine betaine ABC transporter protein with eight *pSymA*-encoded homologs (Table S1). YP\_003064754, another proline/glycine betaine permease/ABC transporter permease protein, has two *pSymA*-encoded homologs. ‘*Ca. Liberibacter asiaticus*’ YP\_003065117 and YP\_003064552, both ATP-binding ABC-type transporter proteins, each have four homologs encoded by genes on *pSymA*. YP\_003065525, a putative amino acid-binding periplasmic protein, is homologous to two *pSymA*-encoded predicted proteins. YP\_003065526, an ABC-type amino acid transporter permease, has 7 homologs encoded by *pSymA* genes. These are examples of multiple, redundant shared

**Table 1.** Proteins shared between the *Sinorhizobium meliloti* plasmid *pSymA*, and the ‘*Ca. Liberibacter asiaticus*’ and *Sinorhizobium meliloti* chromosomes and with roles in the maintenance of homeostasis or signal transduction.

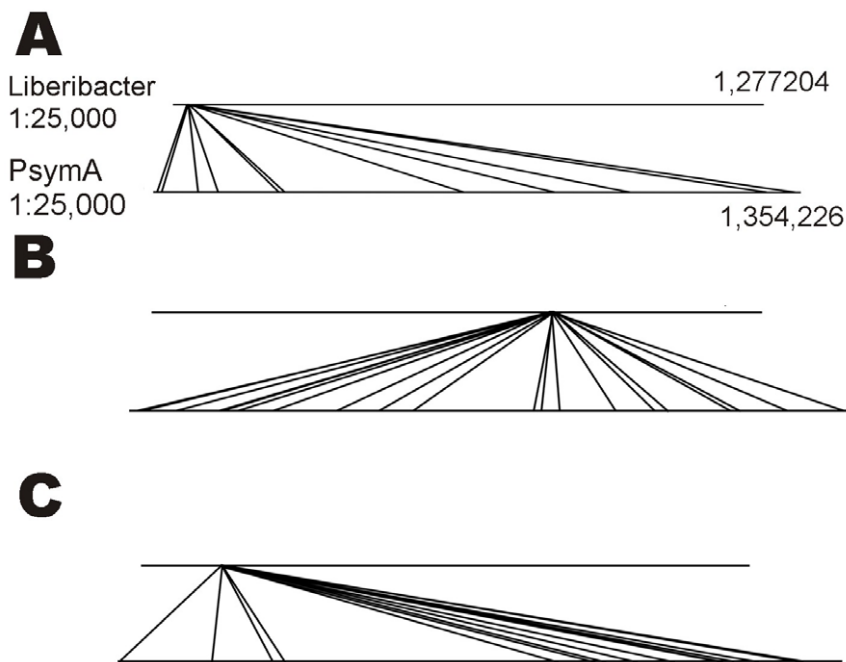
<i>pSymA</i>	<i>L. asiaticus</i>	E	AA	%	<i>S. meliloti</i>	E	Annotation
<b><i>pSymA</i> vs <i>L. asiaticus</i></b>				<b><i>pSymA</i> vs <i>S. meliloti</i></b>			
<b>Maintenance of Homeostasis</b>							
NP_435608.2	YP_003064907	5E-71	306	65	*****	NA	cation diffusion facilitator transporter
NP_436297.2	YP_003064775	3E-30	333	44	*****	NA	Na <sup>+</sup> /H <sup>+</sup> antiporter
NP_436118.1	YP_003065337	1E-134	605	62	NP_384912	1E-22	Potassium efflux protein
NP_436237.1	YP_003065245	0E+00	624	67	NP_384938	0E+00	put potassium uptake protein
<b>Signal Transduction</b>							
NP_436498.1	YP_003064826	3E-16	239	49	NP_384498	6E-22	signal transduction histidine kinase
NP_436497.1	YP_003065306	1E-22	220	47	NP_385120	6E-35	2-component response regulator
NP_435549.2	YP_003065316	1E-12	225	46	NP_386067	2E-22	2-component sensor histidine kinase
NP_435864.1	YP_003064863	3E-49	394	53	NP_385908	1E-53	2-component sensor histidine kinase
NP_435916.1	YP_003064863	2E-16	268	43	NP_385564	4E-19	2-component sensor histidine kinase

Genbank numbers, E-values, number of amino acids (AA) in the homologous region of the protein and the percentage of similar amino acids (%) in pair wise comparisons are provided. The annotations are for the proteins in ‘*Ca. Liberibacter asiaticus*’.  
doi:10.1371/journal.pone.0038725.t001

homologs of ABC-type amino acid transport systems encoded by *pSymA*. ‘*Ca. Liberibacter asiaticus*’ YP\_003064636 is an exception since this putative transmembrane ABC-type transporter permease has only one *pSymA*-encoded homolog, NP\_435480.1 (Table 2). The multiplicity of ABC-type transporter homologs on *pSymA* presumably provides a repertoire of ABC transporters for *S. meliloti*, and is consistent with a high magnitude of importance for

ABC transporter systems in *S. meliloti*. In contrast, ‘*Ca. Liberibacter asiaticus*’ lacks the general overall redundancy of ABC-type transporters present in *S. meliloti*.

Metabolic cofactors are also likely to be imported by ‘*Ca. Liberibacter asiaticus*’ and *S. meliloti* from the plant host. For example, essential vitamins are up taken by ABC-type transporters encoded by *pSymA* and the ‘*Ca. Liberibacter asiaticus*’ chromo-



**Figure 1.** Distribution of genes on the *S. meliloti* *pSymA* megaplasmid that encode proteins orthologous to YP\_003064586, annotated as an ABC-Type amino acid transporter in ‘*Ca. Liberibacter asiaticus*’ (A), that encode proteins orthologous to YP\_003064948, annotated as a 3-ketoacyl-(acyl carrier protein) reductase in ‘*Ca. Liberibacter asiaticus*’ (B) and that encode proteins orthologous to YP\_003064695, annotated as a LysR type regulator of transcription in ‘*Ca. Liberibacter asiaticus*’ (C).  
doi:10.1371/journal.pone.0038725.g001

**Table 2.** Proteins shared between the *Sinorhizobium meliloti* plasmid *pSymA*, and the '*Ca. Liberibacter asiaticus*' and *Sinorhizobium meliloti* chromosomes with roles in ABC-Type transport and the synthesis of amino acids and metabolic cofactors.

<i>pSymA</i>	<i>L. asiaticus</i>	E	AA	%	<i>S. meliloti</i>	E	Annotation
<b><i>pSymA</i> vs <i>L. asiaticus</i></b>		<b><i>pSymA</i> vs <i>S. meliloti</i></b>					
<b>ABC-Type Transport Systems</b>							
NP_435288.1	YP_003064586	6E-87	257	76	NP_385581	1E-100	ABC transport
NP_436046.1	YP_003064753	9E-41	230	59	NP_386072	1E-50	proline/glycinebetaine ABC transporter
NP_436047.1	YP_003064754	3E-12	138	57	NP_385141	3E-19	proline/glycinebetaine ABC transporter
NP_435405.1	YP_003065117	2E-24	243	51	NP_387074	7E-65	ABC transporter, nucleotide binding
NP_435492.2	YP_003064552	2E-24	238	53	NP_387225	1E-92	put ATP binding ABC transporter
NP_436433.1	YP_003065525	3E-12	243	42	NP_387380	8E-21	put amino acid-binding periplasmic ABC
NP_435285.1	YP_003065526	7E-24	202	59	NP_387382	2E-27	amino acid ABC transporter permease
NP_435480.1	YP_003064636	1E-12	275	41	NP_384592	3E-17	ABC transporter permease
NP_435301.1	YP_003064972	5E-22	204	53	NP_386055	1E-66	thiamine transporter ATP-binding subunit
<b>Synthesis of Amino Acids and Metabolic Cofactors</b>							
NP_436267.1	YP_003064901	5E-35	418	46	NP_386510	2E-70	adenosylmethionine-8-amino-7- transam
NP_436409.1	YP_003064762	1E-150	417	74	NP_385314	1E-161	Serine hydroxymethyl transferase
NP_436412.1	YP_003065324	8E-72	278	66	NP_384204	2E-81	Formyl tetrahydrofolate deformylase
NP_435712.1	YP_003064946	4E-63	422	53	NP_385250	6E-67	3-oxoacyl-(acylcarrierprotein) synthase II
NP_435713.1	YP_003064948	8E-78	247	70	NP_385248	1E-110	3-ketoacyl-(acyl-carrier-protein) reductase
NP_435500.1	YP_003064817	6E-13	232	43	NP_386931	2E-46	inositol monophosphatase family
NP_435501.1	YP_003064878	3E-75	436	56	NP_385027	8E-76	threonine synthase
NP_436504.1	YP_003064845	3E-33	427	43	NP_386695	4E-41	Acetyl ornithine transaminase
NP_435614.1	YP_003064846	5E-47	310	53	NP_384624	1E-47	ornithine carbamoyl transferase
NP_436258.1	YP_003065195	9E-13	218	43	NP_384531	4E-11	succinyl-diaminopimelate desuccinylase

Genbank numbers, E-values, number of amino acids (AA) in the homologous region of the protein and the percentage of similar amino acids (%) in pair wise comparisons are provided. The annotations are for the proteins in '*Ca. Liberibacter asiaticus*'.  
doi:10.1371/journal.pone.0038725.t002

some. YP\_003064972, an ATP-binding component of an ABC transport complex for thiamine, has two homologs on both *pSymA* and the *S. meliloti* chromosome.

### Synthesis of Metabolic Cofactors and Amino Acids

In addition to transporter proteins, there are a large number of genes shared between '*Ca. Liberibacter asiaticus*', *S. meliloti* and *pSymA* that encode proteins involved in the synthesis of metabolic cofactors (Table 2). '*Ca. Liberibacter asiaticus*' YP\_003064901 encodes adenosylmethionine-8-amino-7-oxononanoate transaminase (E = 5e-35). The enzyme 8-amino-7-oxononanoate synthase (AONS), which is pyridoxal 5'-phosphate-dependent, catalyzes 8-amino-7-oxononanoate synthesis, needed for biosynthesis of biotin. '*Ca. Liberibacter asiaticus*' YP\_003064762 is a serine hydroxymethyltransferase (SHMT) that catalyzes production of L-serine and tetrahydrofolate with an exceptionally low E-value when compared with its *pSymA* ortholog NP\_436409. YP\_003065324 encodes formyltetrahydrofolate deformylase, active in glyoxylate and dicarboxylate metabolism. YP\_003064946 is a 3-oxoacyl-(acyl carrier protein) synthase, presumably involved in fatty acid synthesis, orthologous to NodE, a *pSymA*-encoded protein NP\_435712.1 (E = 4e-63). YP\_003064948, annotated as a 3-ketoacyl-(acyl carrier protein) reductase, is a dehydrogenase involved in fatty acid biosynthesis (COG1028). Megaplasmid *pSymA* encodes 20 proteins similar to YP\_003064948 (E = 8e-78 to E = 2e-11) (Fig. 1B; Table S1).

Genes on *pSymA* with orthologs on the '*Ca. Liberibacter asiaticus*' and *S. meliloti* chromosomes also encode enzymes involved in the synthesis of amino acids (Table 2). The predicted protein products of '*Ca. Liberibacter asiaticus*' YP\_003064878 and *pSymA* NP\_435501.1 (E = 3e-75) have domains expected of threonine synthase, ThrC, a member of the TrpB tryptophan synthase superfamily. YP\_003064845 is an acetylornithine transaminase protein, part of the arginine metabolism and lysine biosynthesis pathways. '*Ca. Liberibacter asiaticus*' YP\_003064846 and its orthologs in *S. meliloti*, encoded by genes on *pSymA* and its chromosome, are predicted to be an ornithine carbamoyl transferase, involved in arginine metabolism. YP\_003065195 and its *S. meliloti* homolog are succinyl diaminopimelate desuccinylase involved in lysine biosynthesis. However, NP\_436258.1, a similar protein encoded by *pSymA*, is annotated as an acetylornithine deacetylase, an enzyme which catalyzes a step in arginine metabolism. Plant hosts of '*Ca. Liberibacter asiaticus*' are presumed to supply some amino acids to the obligate phytopathogen since certain genes for the biosynthesis of several amino acids were not found in the comparison of the chromosomes of '*Ca. Liberibacter asiaticus*' and *S. meliloti* [2].

### Export of Bioactive Molecules

In addition to the proteins involved in the import or potential export of amino acids and metabolic cofactors, '*Ca. Liberibacter asiaticus*' shares genes with *pSymA* and the *S. meliloti* chromosomes

**Table 3.** Proteins shared between the *Sinorhizobium meliloti* plasmid pSymA, and the '*Ca. Liberibacter asiaticus*' and *Sinorhizobium meliloti* chromosomes with roles in the export of bioactive molecules, as chaperonins and in the regulation of transcription and translation.

pSymA	<i>L. asiaticus</i>	E	AA	%	<i>S. meliloti</i>	E	Annotation
<b>pSymA vs L. asiaticus</b>				<b>pSymA vs S. meliloti</b>			
<b>Export of Bioactive Molecules</b>							
NP_435263.1	YP_003064686	9E-62	505	51	NP_385265	2E-87	put ABC transporter ATP-binding
NP_435718.1	YP_003065157	4E-30	239	51	NP_385995	1E-35	ABC transporter nucleotide binding
NP_435856.2	YP_003065113	5E-33	241	51	NP_385127	2E-45	Serine protease DO like protease
<b>Chaperonins</b>							
NP_435305.2	YP_003065051	2E-15	175	47	NP_385003	7E-27	chaperone protein DnaJ
NP_435342.1	YP_003065326	8E-13	70	67	NP_386217	1E-18	Cold shock protein
NP_435641.1	YP_003065262	0E+00	526	77	NP_384898	0E+00	Chaperonin GroEL
NP_435311.1	YP_003065261	4E-22	95	71	NP_384899	8E-39	co-chaperonin GroES
<b>Regulation of Transcription and Translation</b>							
NP_436119.2	YP_003064695	2E-20	253	48	NP_384882	1E-43	put Transcriptional regulator
NP_435689.2	YP_003065329	2E-33	245	51	NP_386901	8E-60	put $\sigma$ -54-dependent trans reg
NP_436184.1	YP_003065196	7E-32	132	66	NP_385105	1E-35	transcriptional regulator
NP_436495.1	YP_003065365	7E-49	667	44	NP_386250	3E-58	DNA helicase II
NP_436130.1	YP_003065000	2E-15	311	45	NP_387206	8E-28	site-specific recombinase XerD
NP_435253.1	YP_003064676	7E-19	135	54	NP_385460	1E-16	translation elongation factor Tu
NP_436002.1	YP_003065391	7E-20	311	41	NP_385418	1E-34	asp/glut-tRNA amidotransferase

Genbank numbers, E-values, number of amino acids (AA) in the homologous region of the protein and the percentage of similar amino acids (%) in pair wise comparisons are provided. The annotations are for the proteins in '*Ca. Liberibacter asiaticus*'.  
doi:10.1371/journal.pone.0038725.t003

for several proteins used to export small bioactive molecules (Table 3). These include YP\_003064686 an ABC-type transporter with excellent Blast alignments with chromosomal virulence gene (*chvD*) in *Agrobacterium*, *Bartonella* and *Brucella* [22].

The '*Ca. Liberibacter asiaticus*' genome encodes two *nodI* homologs, likely involved in exporting lipid virulence factors. YP\_003065157, an ATP-binding ABC-type transporter protein with a yhbG domain, is similar to *pSymA nodI*, a multidrug efflux-like protein that secretes lipids and lipopolysaccharides including cell wall and outer membrane or Nod factor components. Interestingly, pSymA encodes six homologs of YP\_003065157, consistent with important roles for these proteins in the *S. meliloti*/alfalfa interaction. When ABC-type transporters of '*Ca. Liberibacter asiaticus*' were blasted against the Transporter Classification Database (<http://www.tcdb.org/>), YP\_003065330, annotated as an ABC-type transporter nucleotide binding/ATPase, was homologous with Q2G2M9 (4.00E-088), a putative multi-drug export ATP-binding/permease protein of *Staphylococcus aureus*-SAOUHSC\_02003. Like YP\_003065157, YP\_003065330 aligns best with *nodI* NP\_435718.1 (2.00E-021) encoded by *pSymA*.

'*Ca. Liberibacter asiaticus*' YP\_003065113, orthologous to pSymA-encoded protein NP\_435856.2 (E = 5e-33), a serine protease DO protease, has a PDZ-metalloprotease domain, a PDZ serine protease domain and a trypsin domain, all fused into a single multi-domain protease product. This protein, a probable virulence factor, is likely to protrude from '*Ca. Liberibacter asiaticus*' into the host cell where it can presumably digest host proteins using zinc as a cofactor.

## Chaperonins

Several chaperonin proteins are encoded by orthologous genes shared by '*Ca. Liberibacter asiaticus*', pSymA and *S. meliloti* (Table 3). YP\_003065051 is a DnaJ molecular chaperone protein with one orthologous protein encoded by pSymA, NP435305.2. Cold shock protein YP\_003065326 has three similar proteins encoded by pSymA. YP\_003065262 (GroEL) is orthologous with both pSymA NP\_435641.1 (E = 0) and with pSymA NP\_435310.1 (E = 1e-166). Co-chaperonin GroES YP\_003065261 is orthologous with pSymA-encoded NP\_535311.1 (E = 4e-22) and NP\_435642.1 (E = 9e-21).

## Regulation of Transcription and Translation

'*Ca. Liberibacter asiaticus*', pSymA and *S. meliloti* share several proteins that regulate transcription (Table 3). YP\_003064695 is a member of the LysR family of transcription regulator proteins. Its importance for plant interaction is indicated by at least 14 LysR-like proteins encoded by and distributed across pSymA (E = 1e-21 to E = 1e-10) (Fig. 1C; Table S1). Other genes encoding transcriptional regulators are not as extensively duplicated on pSymA, consistent with LysR family transcription regulators being of unique importance in plant interaction in *S. meliloti* [23]. The '*Ca. Liberibacter asiaticus*' genome, by contrast, encodes only YP\_003064695, annotated as transcriptional regulator with a LysR region. YP\_003065329, a putative sigma-54-dependent transcription regulator protein is similar to NP\_435689 in pSymA annotated as NifA. NifA is a DNA-binding ATPase and CheY-like two-component sensor-based response regulator that is known to



activate expression of genes involved in nitrogen metabolism. YP\_003065196, encodes a transcription-activating MucR-like protein [24], with two orthologs on *pSymA*. MucR proteins regulate production of extracellular polysaccharides which could contribute to biofilm formation in phloem cells of citrus inhabited by '*Ca. Liberibacter asiaticus*'. YP\_003065365 and *pSymA* NP\_436495 (E = 7e-49) are encoded by orthologous genes for a UvrD2 DNA helicase II, or superfamily I DNA and RNA helicase. '*Ca. Liberibacter asiaticus*' YP\_003065000 and *pSymA* orthologs NP\_436130.1 and NP\_436426.1 encodes a XerD recombinase/integrase. YP\_003064676, a Tu translation elongation factor, is orthologous to *pSymA*-encoded NP\_435253 (E = 7e-19) and to NP\_435715.1 (E = 6e-18), and is involved in the synthesis of proteins. YP\_003065391 encodes aspartyl/glutamyl-tRNA amidotransferase, an enzyme known to allow formation of correctly charged asparagine or glutamine tRNAs in organisms which lack either or both a asparaginyl-tRNA synthetases or a glutaminyl-tRNA synthetase.

### Cell Surface

'*Ca. Liberibacter asiaticus*' and *pSymA* share several genes encoding proteins that contribute to a functional architecture of the cell surface (Table 4). A UDP-glucose-4-epimerase, YP\_003064741 probably has a role interconverting monosaccharides needed for exopolysaccharide synthesis. Cyclic diguanylate controls biofilm formation [25]. YP\_003064881, a putative prokaryotic sensory diguanylate cyclase/phosphodiesterase protein, has two orthologs on *pSymA*, NP\_435318.2 (E = 1e-42) and NP\_436089.2 (E = 7e-42). YP\_003065366 encodes an OmpA/MotB-like protein similar with *S. meliloti* *pSymA* NP\_435574.1 (E = 8e-14) and the *S. meliloti* chromosome NP\_384380 (E = 2e-16). Pilus assembly gene clusters on *pSymA* have orthologs on the '*Ca. Liberibacter asiaticus*' chromosome. Proteins YP\_003065134 (NP\_436096, E = 5e-50), YP\_003065135 (NP\_436097, E = 3e-49), YP\_003065137 (NP\_436098, E = 1e-178; NP\_435955, E = 2e-35), YP\_003065140 (NP\_435332.2, E = 9e-45) and YP\_003065141 (NP\_436102.1, E = 3e-29) are all components of a pilus assembly for '*Ca. Liberibacter asiaticus*'. YP\_003065138, a pilus-associated response regulator receiver protein, likely regulates production of pili based on signals received from the environment.

A transglycosylase, YP\_003064721 in '*Ca. Liberibacter*' *asiaticus*' and *pSymA*-encoded NP\_436359.1 (E = 3e-40), could produce cell wall degradation products. YP\_003064721 signal peptide protein orthologous to *pSymA* NP\_436359.1 and is a putative membrane-bound lytic murein transglycosylase (E = 3e-40).

YP\_003065165, a penicillin-binding peptidoglycan synthase, is orthologous to NP\_436481.1 (E = 0). A transmembrane protein that also binds penicillin, YP\_003065516 is orthologous to *pSymA* NP\_436482.1 (E = 1e-93). '*Ca. Liberibacter asiaticus*' YP\_003065356 (Table 1), a glucosamine-fructose-5-phosphate aminotransferase, needed for N-acetyl D-glucosamine synthesis, aligns with *pSymA*-encoded NP\_435728.1 (E = 0).

### Electron Transfer

An important conserved microsytenuous orthologous gene (MOG) cluster, occurs in the '*Ca. Liberibacter asiaticus*' genome and in the genomes of four other members of the Rhizobiales that were analyzed previously [26]. This conserved MOG encodes thirteen subunits that comprise a "mitochondrial-like" Type I respiratory complex (Table 5). Genes encoding this respiratory complex are also found on *pSymA*, and orthologous proteins specified by '*Ca. Liberibacter asiaticus*' chromosomal genes include the following: Chain A, NuoA2 NADH: ubiquinone

oxidoreductase YP\_003065265 and NP\_436080.1 (E = 9e-14); Chain B, or NuoB2 YP\_003065266 and NP\_436079 (E = 6e-48); Chain C, or NuoC2 YP\_003065267 and NP\_436078.1 (E = 6e-33); Chain D, or NuoD2 YP\_003065268 and NP\_436077.1 (E = 1e-110); Chain E or NuoE2, YP\_003065269 and NP\_436076.1 (E = 4e-24) and Chain F or NuoF2 YP\_003065271 and NP\_436075.1 (E = 1e-104). Chains G, H and I/J or NuoG2, NuoH2, a multi-subunit ubiquinone oxidase have the following orthologs: YP\_003065272 and NP\_436074.1 (E = 2e-62); YP\_003065273 and NP\_436071.1 (E = 1e-51), YP\_003065274 and NP\_436072.1 (E = 1e-22), and YP\_003065275 and NP\_436087.2 (E = 8e-11). YP\_003065278 is orthologous with *pSym* proteins NP\_436085.1 and NP\_436083.4 annotated as, respectively, a multidomain K/L oxidoreductase, a NADH ubiquinone oxidase (COG1009), or a multi-domain Na<sup>+</sup>/K<sup>+</sup> antiporter, largely synonymous, involved in energy generation, cation efflux or intracellular pH control. '*Ca. Liberibacter asiaticus*' YP\_003065279 and *pSymA* NP\_436082.1 (E = 2e-92) are chain M of ubiquinone oxidoreductase. '*Ca. Liberibacter asiaticus*' YP\_003065280 and *pSymA* NP\_436081.1 are chain N (E = 4e-57). Nearly all orthologous proteins shared between *pSymA* and *S. meliloti* chromosome have lower E-values aligned with each other than do orthologs shared between *pSymA* and the '*Ca. Liberibacter asiaticus*' chromosome (Tables 1–4). However the amino acid sequence similarities between protein components of the respiratory complex I encoded by *pSymA*, and the '*Ca. Liberibacter asiaticus*' genome and the *S. meliloti* chromosome are highly comparable with each other (Table 5). The organization of the genes in genomic regions encoding the NADH dehydrogenase complex is the same for *S. meliloti* and '*Ca. Liberibacter asiaticus*' (Fig. 2). NuoG and NuoL are encoded by genes in the expected cluster on the *S. meliloti* chromosome. But the e-values obtained by pBlast of *pSymA* NuoG and NuoL vs *S. meliloti* are better for genes found elsewhere on the *S. meliloti* chromosome. The best match for *pSymA* NuoG (NP\_43074) on the *S. meliloti* chromosome is a NAD-dependent formate dehydrogenase alpha subunit (NP\_387115; E = 2e<sup>-95</sup>). The NADH dehydrogenase subunit G (NP\_385378; E = 8e<sup>-77</sup>) is the second best result. This protein is transcribed and translated from a gene with the expected position in the NADH dehydrogenase MOG and aligns well with the '*Ca. Liberibacter asiaticus*' NuoG (e = 0). A similar finding was obtained when *pSymA* NP\_436085 encoding the NuoL subunit was used in a blast search of the chromosome of *S. meliloti*. The best matches were to proteins annotated as monovalent cation/H<sup>+</sup> antiporter subunits located at different positions on the chromosome. NADH dehydrogenase subunit L (NP\_385383; E = 7e<sup>-25</sup>) encoded by the chromosome was significantly less well-aligned. As was seen for NuoG, this protein is transcribed from a gene with the expected position in the NADH dehydrogenase MOG and aligns well with the '*Ca. Liberibacter asiaticus*' NuoL (e = 0). Thus these proteins likely function in the role of NuoL. The NADH dehydrogenase gene cluster on *pSymA* has a different arrangement of genes (Fig. 2). Genes encoding NuoA–NuoG in the center of the Nuo gene cluster are flanked on one side by genes encoding NuoJ – NuoN and by genes, transcribed from the opposite strand, that encode NuoH and NuoI on the other side.

Other proteins are encoded by genes shared by '*Ca. Liberibacter asiaticus*', *pSymA* and the chromosome of *S. meliloti*, and have roles in energy generation. YP\_003065437, is a probable electron transfer flavoprotein-quinone oxidoreductase or FixC. YP\_003064703 is orthologous with both *pSymA*-borne NP\_435265.2 and NP\_435950.1, which are zinc-dependent NADPH quinone oxidoreductases, COG0604, and are important in energy production. YP\_003064781 and YP\_003064782,

**Table 4.** Proteins shared between the *Sinorhizobium meliloti* plasmid pSymA, and the ‘*Ca. Liberibacter asiaticus*’ and *Sinorhizobium meliloti* chromosomes with roles in the modification of cell surface structures.

pSymA	<i>L. asiaticus</i>	E	AA	%	<i>S. meliloti</i>	E	Annotation
<b>pSymA vs <i>L. asiaticus</i></b>				<b>pSymA vs <i>S. meliloti</i></b>			
<b>Cell Surface Structure</b>							
NP_435626.1	YP_003064741	4E-11	188	46	NP_384213	7E-18	UDP-glucose-4-epimerase
NP_436089.2	YP_003064881	7E-42	451	49	NP_385065	2E-95	sensory box/GGDEF family
NP_435318.2	YP_003064881	1E-42	426	49	NP_385065	3E-76	sensory box/GGDEF family
NP_435574.1	YP_003065366	8E-14	96	59	NP_384380	2E-16	OmpA/MotB/outer membrane
NP_436096.2	YP_003065134	5E-50	254	57	NP_384252	1E-62	Pilus component
NP_436097.1	YP_003065135	3E-49	179	72	NP_384251	7E-68	Pilus component
NP_436098.1	YP_003065137	1E-178	408	83	NP_384250	0E+00	component of type IV pilus
NP_435332.2	YP_003065140	9E-45	455	49	NP_384247	2E-43	put pilus assembly
NP_436102.1	YP_003065141	3E-29	200	55	NP_384246	2E-32	put pilus assembly
NP_436101.1	YP_003065138	1E-66	397	55	NP_384249	7E-76	Response regulator receiver (pilus)
NP_436359.1	YP_003064721	3E-40	204	57	NP_387423	1E-64	lytic murein transglycosylase
NP_436481.1	YP_003065165	0E+00	711	70	NP_385436	0E+00	Peptidoglycan synthetase
NP_436482.1	YP_003065516	1E-93	416	60	NP_386288	1E-100	penicillin-binding transmembrane
NP_435728.1	YP_003065356	0E+00	608	76	NP_435728	0E+00	glucosamine-6-Phos transporter

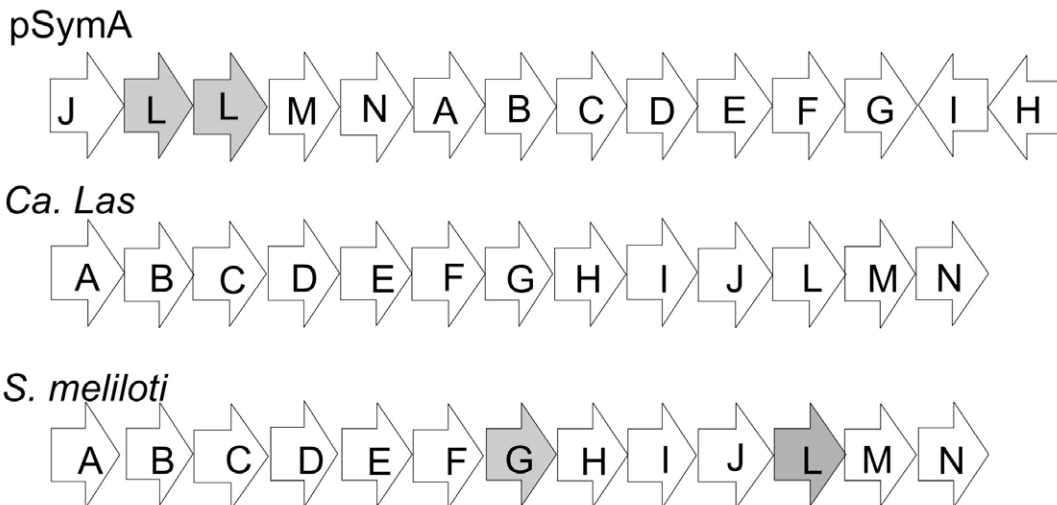
Genbank numbers, E-values, number of amino acids (AA) in the homologous region of the protein and the percentage of similar amino acids (%) in pair wise comparisons are provided. The annotations are for the proteins in ‘*Ca. Liberibacter asiaticus*’.  
doi:10.1371/journal.pone.0038725.t004

encoded by adjacent genes, are orthologs of NP\_436008.1 and NP\_436007.1 (E = 3e-73 and 2e-77 respectively). Both genes have two adjacent ortholog pairs on both pSymA and the *S. meliloti* chromosome. YP\_003065174 and NP\_435375.1 (E = 3e-76) are annotated as energy yielding FAD-dependent dehydrogenases. YP\_003065041 is coproporphyrinogen III oxidase and is orthologous to a pSymA-borne locus NP\_435937.1. The protein has two domains; N-terminal Radical-SAM super family and a C-terminal HemN\_C superfamily, which obtains electrons from a flavodoxin-like system regenerated by a nicotinamide cofactor-dependent flavodoxin [27]. ‘*Ca. Liberibacter asiaticus*’

YP\_003064696 is a NADPH thioredoxin reductase orthologous to NP\_436285.1.

**Discussion**

The premise for identifying protein products of ‘*Ca. Liberibacter asiaticus*’ orthologous to those encoded by genes on pSymA was an hypothesis that such genes and their protein products in all likelihood play direct or indirect role(s) in establishing and maintaining an intimate intracellular interaction with eukaryotes. ‘*Ca. Liberibacter asiaticus*’ has an obligatory requirement for a



**Figure 2.** Diagram showing the organization of the genes encoding the components of the NADH dehydrogenase complex on pSymA, ‘*Ca. Liberibacter asiaticus*’ and *S. meliloti*.  
doi:10.1371/journal.pone.0038725.g002

**Table 5.** Proteins shared between the *Sinorhizobium meliloti* plasmid *pSymA*, and the '*Ca. Liberibacter asiaticus*' and *Sinorhizobium meliloti* chromosomes with roles in electron transfer.

<i>pSymA</i>	<i>L. asiaticus</i>	E	AA	%	<i>S. meliloti</i>	E	Annotation
<b><i>pSymA</i> vs <i>L. asiaticus</i></b>				<b><i>pSymA</i> vs <i>S. meliloti</i></b>			
NP_436080.1	YP_003065265	9E-14	109	53	NP_385370	3E-14	NADH dehydrogenase subunit A
NP_436079.1	YP_003065266	6E-48	141	79	NP_385371	2E-47	NADH dehydrogenase subunit B
NP_436078.1	YP_003065267	6E-33	142	60	NP_385372	1E-29	NADH dehydrogenase subunit C
NP_436077.1	YP_003065268	1E-110	393	64	NP_385373	1E-106	NADH dehydrogenase subunit D
NP_436076.1	YP_003065269	4E-24	169	53	NP_385375	3E-22	NADH-quinone oxidoreductase subunit E
NP_436075.1	YP_003065271	1E-104	420	64	NP_385376	1E-105	NADH dehydrogenase I subunit F
NP_436074.1	YP_003065272	2E-62	410	54	NP_385378*	2E-65	NADH dehydrogenase subunit G
NP_436071.1	YP_003065273	1E-51	311	55	NP_385379	2E-54	NADH dehydrogenase subunit H
NP_436072.1	YP_003065274	1E-22	136	54	NP_385380	4E-23	NADH dehydrogenase subunit I
NP_436087.2	YP_003065275	8E-11	54	78	NP_385381	6E-10	NADHdehydrogenasesubunit J
NP_436085.1	YP_003065278	2E-23	461	42	NP_385383*	1E-22	NADH dehydrogenase subunit L
NP_436083.4	YP_003065278	3E-20	180	52	NP_385383*	1E-19	NADH dehydrogenase subunit L
NP_436082.1	YP_003065279	2E-92	485	54	NP_385384	1E-100	NADH dehydrogenase subunit M
NP_436081.1	YP_003065280	4E-57	427	50	NP_385385	2E-75	NADH dehydrogenase subunit N
NP_436008.1	YP_003064781	3E-73	249	71	NP_436008	1E-128	electron transfer flavoprotein beta
NP_435693.1	YP_003064781	4E-14	186	52	NP_435693	8E-18	electrontransfer flavoprotein beta
NP_436007.1	YP_003064782	2E-77	316	69	NP_436007	1E-150	Electron transfer flavoprotein alpha
NP_435692.1	YP_003064782	2E-20	245	48	NP_386750	3E-21	Electron transfer flavoprotein alpha
NP_435691.1	YP_003065437	1E-18	326	43	NP_385139	6E-09	flavoprotein ubiquinone oxidoreductase
NP_435265.2	YP_003064703	2E-26	334	43	NP_385059	1E-32	quinone oxidoreductase
NP_435950.1	YP_003064703	4E-12	269	40	NP_386798	5E-42	quinoneoxidoreductase
NP_435375.1	YP_003065174	3E-76	425	56	NP_384973	7E-99	prob FAD-dependent oxidoreductase
NP_435937.1	YP_003065041	9E-20	187	51	NP_384479	3E-19	coproporphyrinogen III oxidase
NP_436285.1	YP_003064696	1E-22	134	61	NP_385673	1E-31	thioredoxin reductase (NADPH)

Genbank numbers, E-values, number of amino acids (AA) in the homologous region of the protein and the percentage of similar amino acids (%) in pair wise comparisons are provided. The annotations are for the proteins in '*Ca. Liberibacter asiaticus*'. \*indicates proteins that were the second or third best match by Blast. doi:10.1371/journal.pone.0038725.t005

eukaryotic host, and there are multiple orthologs of these genes in *Sinorhizobium meliloti*. These factors preclude mutagenesis studies at the present time to test this hypothesis. In particular, we discovered that there are several orthologous genes in the '*Ca. Liberibacter asiaticus*' genome and on *pSymA* that encode proteins with vital roles in maintaining intracellular homeostasis for mono- and divalent cations and pH. Genes encoding a  $\text{Na}^+/\text{H}^+$  antiporter [20,21] and a cation efflux antiporter protein that secretes excess metal ions  $\text{Cd}^{++}$ ,  $\text{Zn}^{++}$  or  $\text{Co}^{++}$  in exchange for  $\text{K}^+$  and  $\text{H}^+$  [19,28] were identified as orthologs shared exclusively by *pSymA* and '*Ca. Liberibacter asiaticus*,' i.e., not found on the *S. meliloti* chromosome. In a previous study we compared the proteins encoded by the circular chromosomes of '*Ca. Liberibacter asiaticus*' and four other members of the Rhizobiales. In that study the cation efflux antiporter (YP\_003064907) was found to be uniquely shared between '*Ca. Liberibacter asiaticus*' and phytopathogenic *Agrobacterium tumefaciens* C58 (NP\_354015; 3E-84). The  $\text{Na}^+/\text{H}^+$  antiporter (YP\_003064775) shared between *pSymA* and '*Ca. Liberibacter asiaticus*' is also shared with 10 species or strains of the *Bartonellaceae* (6E-60–4E-48), but not with the other members of the Rhizobiales studied [26]. This suggests that this protein is used to facilitate an intracellular lifestyle. In addition, another protein encoded by genes shared by *pSymA* and the '*Ca. Liberibacter asiaticus*' and *S. meliloti* chromosomes, is a glutathi-

one-gated potassium efflux protein that exchanges potassium for sodium and protons and thereby modulates cytoplasmic pH [29,30]. This orthologous protein pair encoded by both *pSymA* and the '*Ca. Liberibacter asiaticus*' genome align with an e-value that is 110 orders of magnitude lower than the alignment of the orthologous proteins encoded by *pSymA* and the *S. meliloti* chromosome. Thus, the maintenance of intracellular pH and ion homeostasis are likely critical functions of proteins encoded by genes shared uniquely by '*Ca. Liberibacter asiaticus*' and *pSymA* (but not the *S. meliloti* chromosome). Other '*Ca. Liberibacter asiaticus*' genes with orthologs on *pSymA* likely encode proteins with either direct or indirect nutritional or energy roles in establishing and maintaining an intimate intracellular interaction with eukaryotic hosts.

We identified 86 '*Ca. Liberibacter asiaticus*' genes with 182 probable orthologs on *pSymA* and the *S. meliloti* chromosome. Most of the orthologous proteins encoded by *pSymA* and the *S. meliloti* chromosome produced alignments with substantially more significant E-values than orthologous proteins encoded by '*Ca. Liberibacter asiaticus*' and *pSymA*. Individual genes encoded by the reduced genome of '*Ca. Liberibacter asiaticus*' may have many orthologs or paralogs carried by the *pSymA* megaplasmid besides additional similar genes on the chromosome of *S. meliloti*. An example is YP\_003064948, a 3-ketoacyl-(acyl-carrier-protein)



reductase with 20 similar proteins encoded by *pSymA*. Thus, numerous physiological and niche adaptations conferred upon *S. meliloti* by *pSymA* may not be available to '*Ca. Liberibacter asiaticus*.' However, in a reduced genome such as that of '*Ca. Liberibacter asiaticus*', genes that are retained may function with relaxed substrate specificity [31,32], compensating for the lack of multiple homologs in the '*Ca. Liberibacter asiaticus*' genome. The multiplicity of *pSymA* orthologs complicates any mutagenesis strategy to definitively establish their roles in *S. meliloti*.

Protein products not constitutively expressed may be considered as candidates for being involved with niche specialization or pathogenesis; *S. meliloti* has conditionally expressed genes that are especially important for nitrogen fixation and nodulation symbiosis. YP\_003064695 is a putative regulator of transcription by '*Ca. Liberibacter asiaticus*', with 14 related proteins in *pSymA*, many of which are annotated as LysR transcriptional regulatory proteins. Despite considerable conservation both structurally and functionally, LysR-type transcriptional regulators regulate a diverse set of genes, including those involved in virulence, metabolism, quorum sensing and motility [33,34]. In *S. meliloti*, LysR proteins function in exquisite control of a sequence of events leading to root colonization, initiation of nodulation and subsequent differentiation of nodules containing *S. meliloti* [23]. Fatty acid products of the 3-ketoacyl-(acyl-carrier-protein) reductase orthologs are also used to complete the structure of nodulation factors. This may be why the extensive collection of LysR regulatory elements and 3-ketoacyl-(acyl-carrier-protein) reductases encoded on *pSymA* are not found on the '*Ca. Liberibacter asiaticus*' chromosome. Thus, '*Ca. Liberibacter asiaticus*' may have much less capability for finely nuanced control of the expression of genes, consistent with its obligatory intracellular lifestyle. Parasitism of citrus by '*Ca. Liberibacter asiaticus*' follows direct injection of the pathogen into the host and results in systemic invasion of phloem tissues throughout the host [9,10], and differentiated nodules are not produced.

Genes with orthologs on both *pSymA* and the chromosome of *S. meliloti* are largely specialized for symbiosis and nitrogen metabolism in *S. meliloti* and include 48 ABC-type transporter genes. The genome of '*Ca. Liberibacter asiaticus*' encodes only 11 ABC-type transporter proteins that are orthologous to *pSymA*-encoded proteins, compared with a total of about 40 ABC-type transporters encoded by the chromosome [2]. These transporters likely serve to import sugars, amino acids and precursors and purines and pyrimidines into '*Ca. Liberibacter asiaticus*' from the host. The genome of '*Ca. Liberibacter asiaticus*' is known to be deficient in genes used in the biosynthetic pathways of these compounds [2,12].

ABC-type transporters may also be used to export small molecules by '*Ca. Liberibacter asiaticus*'. YP\_003064989 is an ABC-type transporter protein encoded by '*Ca. Liberibacter asiaticus*' that likely functions to export glucans for osmoprotection. A class of proteins that confer multiple drug resistance, are called 'RND' for Resistance, Nodulation and Division [35]. RND-type proteins confer drug resistance based on rapid efflux of drugs from the bacterial cell. YP\_003065157, is an ATP-binding ABC-type transporter protein with a yhbG domain, and is orthologous to *pSymA* NodI, a protein that is required for nodulation of alfalfa by *S. meliloti*. NodI-like proteins secrete lipids and lipopolysaccharides such as cell wall, outer membrane or Nod factor components needed to stimulate the differentiation of nodules by the host. Very interestingly *pSymA* encodes six homologs of YP\_003065157. Orthologs of ABC-type transporters encoded by '*Ca. Liberibacter asiaticus*' were identified with the "Transporter Classification Database" (<http://www.tcdb.org/>). YP\_003065330, annotated as

an ABC-type transporter with nucleotide binding and ATPase domains was homologous to Q2G2M9 (4.00E-88), annotated as a putative multi-drug export ATP-binding/permease protein, *Staphylococcus aureus* SAOUHSC\_02003. YP\_003065330 aligns best with nodI NP\_435718.1 (2.00E-021) encoded by *pSymA*. Thus these proteins at the least share similar structural motifs.

Genes encoding type IV pili may have evolved from structures used for DNA transfer but it is now known that virulence proteins may be secreted by components of a type IV pili [36]. Genes encoding type IV pili are shared between '*Ca. Liberibacter asiaticus*', *S. meliloti*, and *pSymA*. Effectors directly involved with the establishment of '*Ca. Liberibacter asiaticus*' as an intracellular pathogen living within the phloem cells of citrus, may be transferred from the bacterium to its host by type IV secretion involving pilus assembly proteins. If pili are produced by '*Ca. Liberibacter asiaticus*', they may also have roles in twitching motility and biofilm formation as in *Xylella fastidiosa* [37].

Why is the entire *S. meliloti* chromosomal respiratory complex I gene set duplicated on *pSymA*? Plants respond to bacterial pathogens by producing a sustained oxidative burst leading to apoptosis. The respiratory complex protein NuoG can also be used by *Mycobacterium tuberculosis* to pump protons out of the cell to neutralize such superoxide radicals [38]. This may explain the duplication of the genes encoding the respiratory complex on *pSymA* and provide an additional role for NuoG in the '*Ca. Liberibacter asiaticus*'/citrus interaction. Homeostasis with respect to sodium is also of critical importance to both *S. meliloti* and '*Ca. Liberibacter asiaticus*' as discussed above, and in addition to pumping protons out of the cell to create the proton motive force needed for the synthesis of ATP, respiratory complex I of at least some Gram negative bacteria can also transport sodium ions out of the cells [39]. Besides an obvious potential gene dosage effect, redundant copies carried on an extrachromosomal element provide a palette for gene evolution. However in spite of the potential for diversification of these genes afforded to *pSymA*, substantial diversification is not apparent since the E-values for the orthologs encoded by *pSymA*, '*Ca. Liberibacter asiaticus*' and *S. meliloti* are very similar.

Most '*Ca. Liberibacter asiaticus*' protein products that align significantly with proteins encoded by *pSymA* also align significantly with predicted protein products encoded by the *S. meliloti* chromosome. Thus, although present as single copies for the most part in '*Ca. Liberibacter asiaticus*', these genes are present in multiple copies per replicon in the free-living microsymbiont. The presence of multiple orthologs in *pSymA* of '*Ca. Liberibacter asiaticus*' proteins such as the amino acid transporter YP\_003064586, the fatty acid dehydrogenase YP\_003064948 and the LysR transcriptional regulator YP\_003064695, suggests that their products are of particular importance in host/microbe interactions, and are likely to be examples of a single protein acquiring multiple functions in an organism with a reduced genome [32]. '*Ca. Liberibacter asiaticus*' very notably also has two or three genes shared with *pSymA* but not the *S. meliloti* chromosome. These genes are predicted to produce proteins that generate a proton gradient with ATP-producing potential, maintain sodium and potassium homeostasis, prevent over accumulation of divalent metal cations and maintain intracellular pH. Thus the maintenance of intracellular homeostasis is potentially a vital requirement for successful colonization of citrus and psyllids by '*Ca. Liberibacter asiaticus*'. In an obligate pathogen, cellular vitality and pathogenesis are tied together.

## Materials and Methods

The genome of ‘*Ca. Liberibacter asiaticus*’ was compared with both the circular chromosome and megaplasmids, especially pSymA, of *S. meliloti*. To identify orthologous proteins, predicted amino acid sequences (‘*Ca. Liberibacter asiaticus*’ strain psy62, RefSeq NC\_012985; *Sinorhizobium meliloti* strain 1021 RefSeq NC\_003047, and pSymA RefSeq NC\_003037) were downloaded from NCBI. Using default BLAST parameters, each predicted amino acid sequence from the ORFs identified on pSymA of the *S. meliloti* 1021 genome was BLASTed against the predicted amino acid sequences of the ORFs on the chromosomes of ‘*Ca. Liberibacter asiaticus*’ and *Sinorhizobium meliloti* strain 1021. Perl scripts and Excel spreadsheets were created to identify hits between genomes with low, negative e-values and to extract annotations from Genbank. In addition, amino acid similarity and length of the blast hit was extracted from each top hit from the BLAST output. First, a general analysis was done extracting predicted protein products having BLAST alignment values of e - 10 or lower. Similar or homologous proteins were also identified manually where possible to be consistent with the annotations from the authors of each respective genome. That is, annotations of ORFs were extracted from NCBI annotations of the genome of ‘*Ca. Liberibacter asiaticus*’ and megaplasmid pSymA of *Sinorhizobium meliloti* strain 1021, respectively. Thus, the matching homolog might not have the best e-value or match with the top hit as many proteins share the same domain structure or amino acid similarity,

## Reference

- Kuykendall LD (2005) Order Rhizobiales, Type Family Rhizobiaceae Genus Rhizobium. In: Brenner, Krieg, Staley, Garrity, editors. *Bergey's Manual of Systematic Bacteriology*, 2nd Edition. New York: Springer. pp 324–361.
- Duan Y, Zhou L, Hall DG, Li W, Doddapaneni H, et al. (2009) Complete genome sequence of citrus Huanglongbing bacterium, ‘*Candidatus Liberibacter asiaticus*’ obtained through metagenomics. *Molecular Plant-Microbe Interactions* 22: 1011–1020.
- Kuykendall LD (2006) List of new names and new combinations previously effectively, but not validly, published. Validation List no. 107. *International Journal of Systematic and Evolutionary Microbiology* 56: 1–6.
- Young JM, Kuykendall LD, Martinez-Romero E, Kerr A, Sawada H (2001) A revision of *Rhizobium* Frank 1889, with an emended description of the genus, and the inclusion of all species of *Agrobacterium* Conn 1942 and *Allorhizobium undicola* de Lajudic, et al. 1998 as new combinations: *Rhizobium radiobacter*, *R. rhizogenes*, *R. rubi*, *R. undicola* and *R. vitis*. *Int J Syst and Evol Microbiol* 51: 89–103.
- Coletta-Filho HD, Targon MLPN, Takita MA, Negri JDD, Pompeu J, et al. (2004) First report of the causal agent of Huanglongbing (‘*Candidatus Liberibacter asiaticus*’) in Brazil. *Plant Disease* 88: 1382.
- Gottwald TR (2010) Current epidemiological understanding of citrus Huanglongbing. *Annual Review of Phytopathology* 48: 119–139.
- Salibe AA, Cortez RE (1966) Studies on the leaf mottling disease of citrus in the Philippines. *FAO Plant Protection Bulletin* 14: 141–144.
- Galibert F, Finan TM, Long SR, Puhler A, Abola P, et al. (2001) The composite genome of the legume symbiont *Sinorhizobium meliloti*. *Science* 293: 668–672.
- Tatineni S, Shankar Sagaram U, Gowda S, Robertson CJ, Dawson WO, et al. (2008) In planta distribution of ‘*Candidatus Liberibacter asiaticus*’ as revealed by polymerase chain reaction (PCR) and real-time PCR. *Phytopathology* 98: 592–599.
- Li W, Levy L, Hartung JS (2009) Quantitative distribution of ‘*Candidatus Liberibacter asiaticus*’ in citrus plants with citrus Huanglongbing. *Phytopathology* 99: 139–144.
- Ammar E-D, Shatters Jr RG, Hall DG (2011) Localization of *Candidatus Liberibacter asiaticus*, Associated with Citrus Huanglongbing Disease, in its Psyllid Vector using Fluorescence in situ hybridization. *J Phytopathology* doi: 10.1111/j.1439-0434.2011.01836.x.
- Hartung JS, Shao J, Kuykendall LD (2011) Comparison of the ‘*Ca. Liberibacter asiaticus*’ genome adapted for an intracellular lifestyle with other members of the Rhizobiales. *PLoS One* 6: e23289.
- Barnett MJ, Fisher RF, Jones T, Komp C, Abola AP, et al. (2001) Nucleotide sequence and predicted functions of the entire *Sinorhizobium meliloti* pSymA megaplasmid. *PNAS* 98: 9883–9888.
- Slater SC, Goldman BS, Goodner B, Setubal JC, Farrand SK, et al. (2009) Genome Sequences of Three *Agrobacterium* Biovars Help Elucidate the Evolution of Multichromosome Genomes in Bacteria. *J Bacteriology* 191: 2501–2511.

but are functionally quite different. Similarly, for proteins of interest that were unique to each genome, proteins were sorted for positive e-values, signifying that an orthologous protein was not encoded in a respective genome. Orthologous genes from the ‘*Ca. Liberibacter asiaticus*’ chromosome, the *S. meliloti* chromosome and megaplasmid pSymA were mapped on linear representations of the respective genophores opened at their origins of replication.

## Supporting Information

### Table S1 Proteins shared between the *Sinorhizobium meliloti* plasmid pSymA, and the ‘*Ca. Liberibacter asiaticus*’ and *Sinorhizobium meliloti* chromosomes.

The protein IDs and annotations used by NCBI and the e-values of pairwise comparisons are provided.

(XLSX)

## Acknowledgments

We acknowledge the technical assistance of Ms. Lauren Conway, Ms. Felicia Davenport and Ms. Larissa Higginbotham.

## Author Contributions

Conceived and designed the experiments: LDK JYS JSH. Performed the experiments: LDK JYS JSH. Analyzed the data: LDK JYS JSH. Contributed reagents/materials/analysis tools: JYS. Wrote the paper: LDK JYS JSH.

- Rogel MA, Hernandez-Lucas I, Kuykendall LD, Balkwill DL, Martinez-Romero E (2001) Nitrogen-fixing nodules with *Ensifer adhaerens* harboring *Rhizobium tropici* symbiotic plasmids. *Appl Environ Microbiol* 67: 3264–3268.
- 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.
- Rosenblueth M, Martinez-Romero E (2006) Bacterial endophytes and their interactions with hosts. *Molecular Plant-Microbe Interactions* 19: 827–837.
- Soto MJ, Sanjuan J, Olivares J (2006) Rhizobia and plant-pathogenic bacteria: common infection weapons. *Microbiology* 152: 3167–3174.
- Goldberg M, Pribylt T, Juhnke S, Nies DH (1999) Energetics and topology of CzcA, a cation/proton antiporter of the resistance-nodulation-cell division protein family. *J Biol Chemistry* 274: 26065–26070.
- Taglicht D, Padan E, Schuldiner S (1991) Overproduction and purification of a functional Na<sup>+</sup>/H<sup>+</sup> antiporter coded by *nhaA* (ant) from *Escherichia coli*. *J Biol Chemistry* 266: 11289–11294.
- Williams KA (2000) Three-dimensional structure of the ion-coupled transport protein NhaA. *Nature* 403: 112–115.
- Goodner B, Hinkle G, Gattung S, Miller N, Blanchard M, et al. (2001) *Agrobacterium tumefaciens* C58: Genome sequence of the plant pathogen and biotechnology agent. *Science* 294: 2323–2328.
- Swanson JA, Mulligan JT, Long SR (1993) Regulation of *SyrM* and *NodD3* in *Rhizobium meliloti*. *Genetics* 134: 435–444.
- Mueller K, Gonzalez JE (2011) Complex regulation of symbiotic functions is coordinated by MucR and quorum sensing in *Sinorhizobium meliloti*. *J Bacteriol* 193: 485–496.
- Tischler AD, Camilli A (2005) Cyclic diguanylate regulates *Vibrio cholerae* virulence. *Infection and Immunity* 73.
- Kuykendall LD, Shao J, Hartung JS (2012) Conservation of Gene Order and Content in the Circular Chromosomes of ‘*Candidatus Liberibacter asiaticus*’ and Other Rhizobiales. *PLoS One* 7: e34673.
- Layer G, Verfurth K, Mahlitz E, Jahn D (2002) Oxygen-independent coproporphyrinogen-III oxidase HemN from *Escherichia coli*. *J Biol Chem* 277: 34136–34142.
- Guffanti AA, Wei Y, Rood SV, Krulwich TA (2002) An antiport mechanism for a member of the cation diffusion facilitator family: divalent cations efflux in exchange for K<sup>+</sup> and H<sup>+</sup>. *Molecular Microbiology* 45: 145–153.
- Ness LS, Booth IR (1999) Different foci for the regulation of the activity of the KefB and KefC glutathione-gated K<sup>+</sup> efflux systems. *J Biol Chemistry* 274: 9524–9530.
- Roosild TP, Castronovo S, Healy J, Miller S, Pliotas C, Rasmussen T, Bartlett W, Conway SJ, Booth IR (2010) Mechanism of ligand-gated potassium efflux in bacterial pathogens. *PNAS USA* 107: 19784–19789.

31. Cordwell SJ, Basseal DJ, Pollack JD, Humphery-Smith I (1997) Malate/lactate dehydrogenase in mollicutes: evidence for a multienzyme protein. *Gene* 197: 113–120.
32. Pollack JD, Myers MA, Dandekar T, Herrmann R (2002) Suspected Utility of Enzymes with Multiple Activities in the Small Genome *Mycoplasma* Species: The Replacement of the Missing “Household” Nucleoside Diphosphate Kinase Gene and Activity by Glycolytic Kinases. *OMICS* 6: 247–258.
33. Kovacicova G, Skorupski K (1999) A *Vibrio cholerae* LysR homologue, AphB, cooperates with AphA at the *topPH* promoter to activate expression of the ToxR virulence cascade. *J Bacteriology* 181: 4250–4256.
34. Harris SJ, Shih Y-L, Bentley SD, Salmond GPC (1998) The *hexA* gene of *Erwinia carotovora* encodes a LysR homologue and regulates motility and the expression of multiple virulence determinants. *Molecular Microbiology* 28: 705–707.
35. Blair JMA, Piddock IJV (2009) Structure, function and inhibition of RND efflux pumps in gram-negative bacteria: an update. *Current Opinion in Microbiology* 12: 512–519.
36. Rego AT, Chandran V, Waksman G (2010) Two-step and one-step secretion mechanisms in Gram-negative bacteria: contrasting the Type IV secretion system and the chaperone usher pathway of pilus biogenesis. *Biochem J* 425: 475–488.
37. Li Y, Hao G, Galvani CD, Meng Y, Fuente Ldl, et al. (2007) Type I and type IV pili of *Xylella fastidiosa* affect twitching motility, biofilm formation and cell-cell aggregation. *Microbiology* 153: 719–726.
38. Velmurugan K, Chen B, Miller JL, Azogue S, Gurses S, et al. (2007) *Mycobacterium tuberculosis nuoG* is a virulence gene that inhibits apoptosis in infected host cells. *PLoS Pathogens* 3: e110.
39. Friedrich T, Scheide D (2000) The respiratory complex I of bacteria, archaea and eukarya and its module common with membrane-bound multisubunit hydrogenases. *FEBS Letters* 479: 1–5.