Applied Microbiology and Biotechnology

Copper Tolerance in *Frankia* sp. Strain EuI1c involves Surface-Binding and Copper Transport.

Medhat Rehan[1,2], Teal Furnholm[1], Ryan H Finethy[1], Feixia Chu[1], Gomaah El-Fadly[2], and Louis S. Tisa[1,*]

[1]Department of Molecular, Cellular, and Biomedical Sciences, University of New Hampshire, Durham, NH 03824-2617, USA

[2]Department of Genetics, Kafrelsheikh University, 33516, Kafr El-Sheikh, Egypt

* Corresponding Author: **Mailing address:** Department of Genetics, Kafrelsheikh University, 33516, Kafr El-Sheikh, Egypt. **Telephone:** +2-01006976664. **Fax:** +2-0479102930. **E-mail:** medhat.rehan@agr.kfs.edu.eg.

Table S1 Primers used in this study

Locus ID (gene name)	Forward (5' → 3') primer Reverse (5' → 3') primer	Description
FraEuI1c_6235	F-AACACCACCGGTGAGATCAAGA R-ATCACCTTGCAGGAGAACAGGT	rpoB
FraEuI1c_4907	F-AGCTGGACAAGAACCGCAACAA R-AAGTTGACGATGGAGCTGACGA	rpsA
FraEuI1c_0940	F-TCGGCAAGACCGTCATCATCCA R-ATCACCTTGCAGGAGAACAGG	atpD
FraEuI1c_4734	F-TCTACCGCCTGTCCTACCGGAT R-ACGAACCGGACCCTACGTCA	copC1
FraEuI1c_7109	F-TCTACAAGCTGTGGTTCGTGCT R-ATCACCCACCGGTTCGACAGGAA	copCD
FraEuI1c_1869	F-TCCCAGCTGTTCACCACCATGTA R-CGAAGATGAGGATCTTCACGCCGA	copD
FraEuI1c_6307	F-TTCTGGGCGTTCGCCTACAACG R -ATCACCTTGCAGGAGAACAGGT	copA (copper-translocating P-type ATPase)
FraEuI1c_6308	F -GTATGACCTGCGGGCACTGTGT R-TGGTAGCCGGCCTCTTCGACGG	copZ (Heavy metal transport/ detoxification protein)
FraEuI1c_7040	F-TCCTGATGATCACCAACCTGGACT R-AAGCTGACGTAGTACTGAGCGGAA	Periplasmic Binding protein/lacI
FraEuI1c_4628	F-ACACCAAGAAGAAGTCCGTCACCA R-GATCGTGATGCTGTTGTCGGCATA	Hypothetical protein
FraEuI1c_6772	F-TGACCAAGCAGATCAAGTTCACGC R-TTCACCTGGTAGATGTAGAACGACGC	ABC transporter ligand-binding protein
FraEuI1c_1092	F-AAGGTCAACAAGGAGTTCTTCGACGC R-TCATCCGGTAGCGTTCTCGATCTT	Sulfate ABC transporter, periplasmic sulfate-binding protein

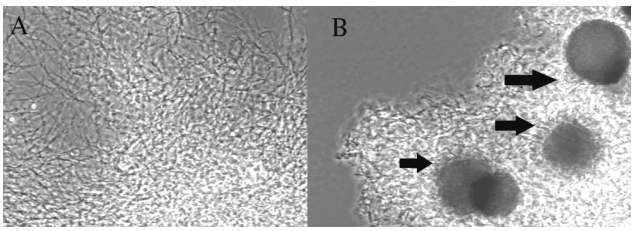

Figure S1. Photomicrograph of *Frankia* strain DC12 grown in growth medium supplemented with (B) and without Cu^{+2} (A). Arrow points to globular structure on surface of hyphae.

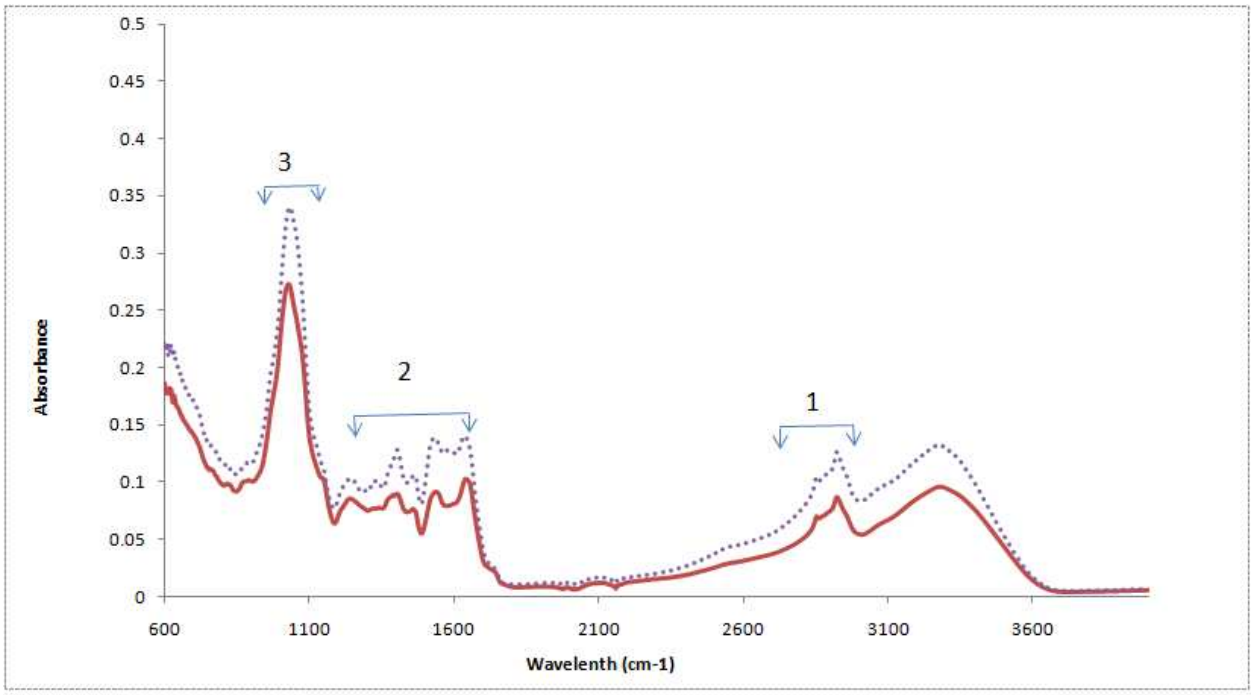

Figure S2. FTIR spectra of Cu^{+2}–stressed *Frankia* cells. Cultures were incubated for 5 days in low phosphate growth medium with (dashed line) or without (solid line) 1 mM $CuSO_4$. After incubation, the cultures were collected and treated as described in the Methods. FTIR scans of lyophilized cells were taken and the averaged scans are presented (n=3). Numbers in figure represents areas of change and correspond to specific chemical signatures: (1) fatty acids, (2) fatty acids - proteins and phosphate-carrying compounds, and (3) cell wall carbohydrates – polysaccharides – PO_4 groups.

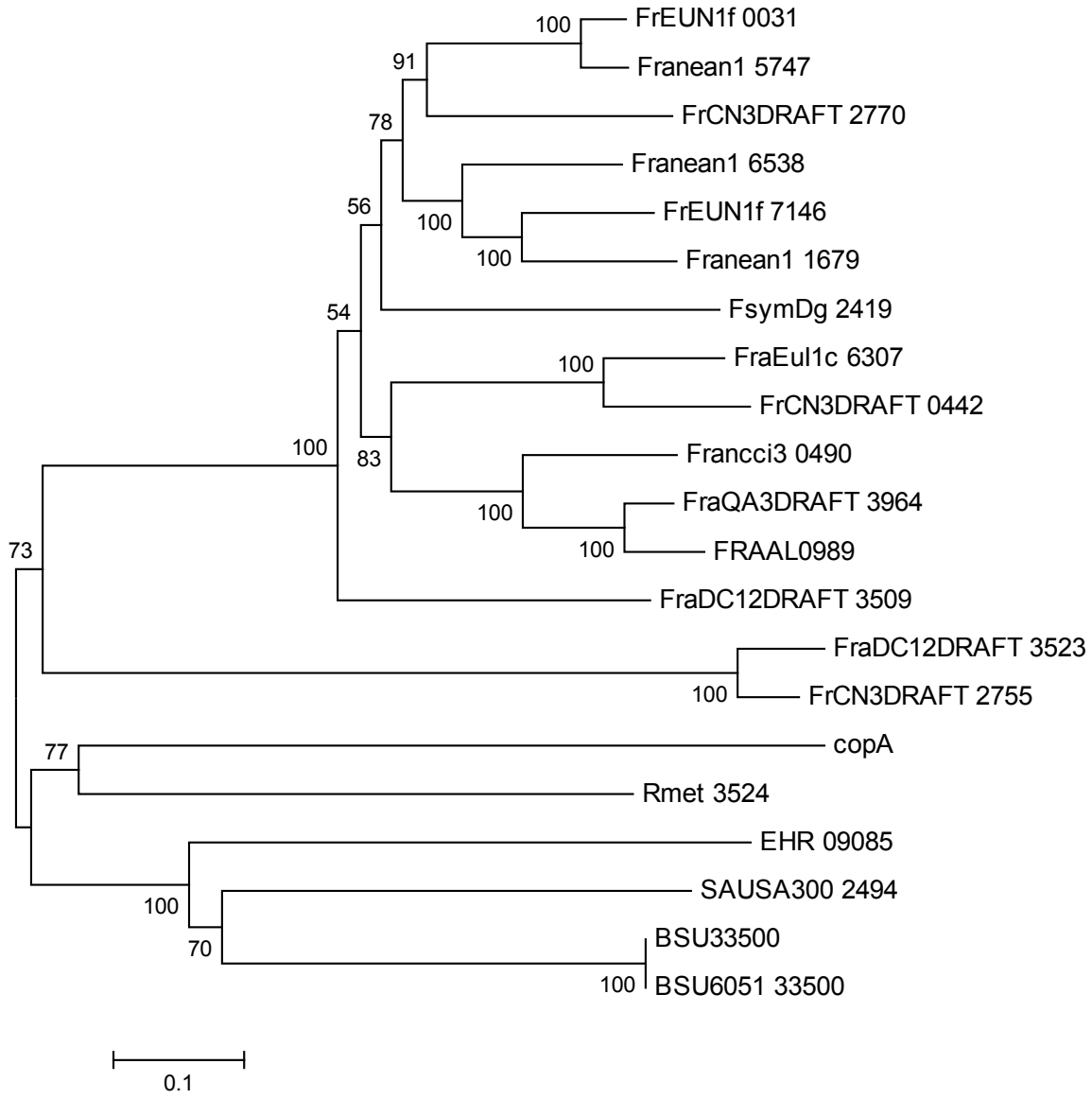

Figure S3. Phylogenetic tree of CopA protein sequences. The evolutionary history was inferred using the Neighbor-Joining method (Saitou and Nei 1987). The optimal tree with the sum of branch length = 5.35347782 is shown. The percentage of replicate trees in which the associated taxa clustered together in the bootstrap test (1000 replicates) are shown next to the branches (Felsenstein 1985). The tree is drawn to scale, with branch lengths in the same units as those of the evolutionary distances used to infer the phylogenetic tree. The evolutionary distances were computed using the JTT matrix-based method (Jones et al. 1992) and are in the units of the number of amino acid substitutions per site. The analysis involved 21 amino acid sequences. All ambiguous positions were removed for each sequence pair. There were a total of 1106 positions in the final dataset. Evolutionary analyses were conducted in MEGA5 (Tamura et al. 2011). Besides the *Frankia* CopA protein sequences, the following protein sequences were used: *Escherichia coli* K12 W3310 (copA)**,** *Enterococcus hire* (EHR_09085), *Bacillus subtilis* strain 168 (BSU33500), *Bacillus subtilis* 6051HGW (BSU6051_33500), *Staphylococcus aureus* USA300 SAUSA300_2494, and *Cupriavidus metallidurans* Rmet_3524).

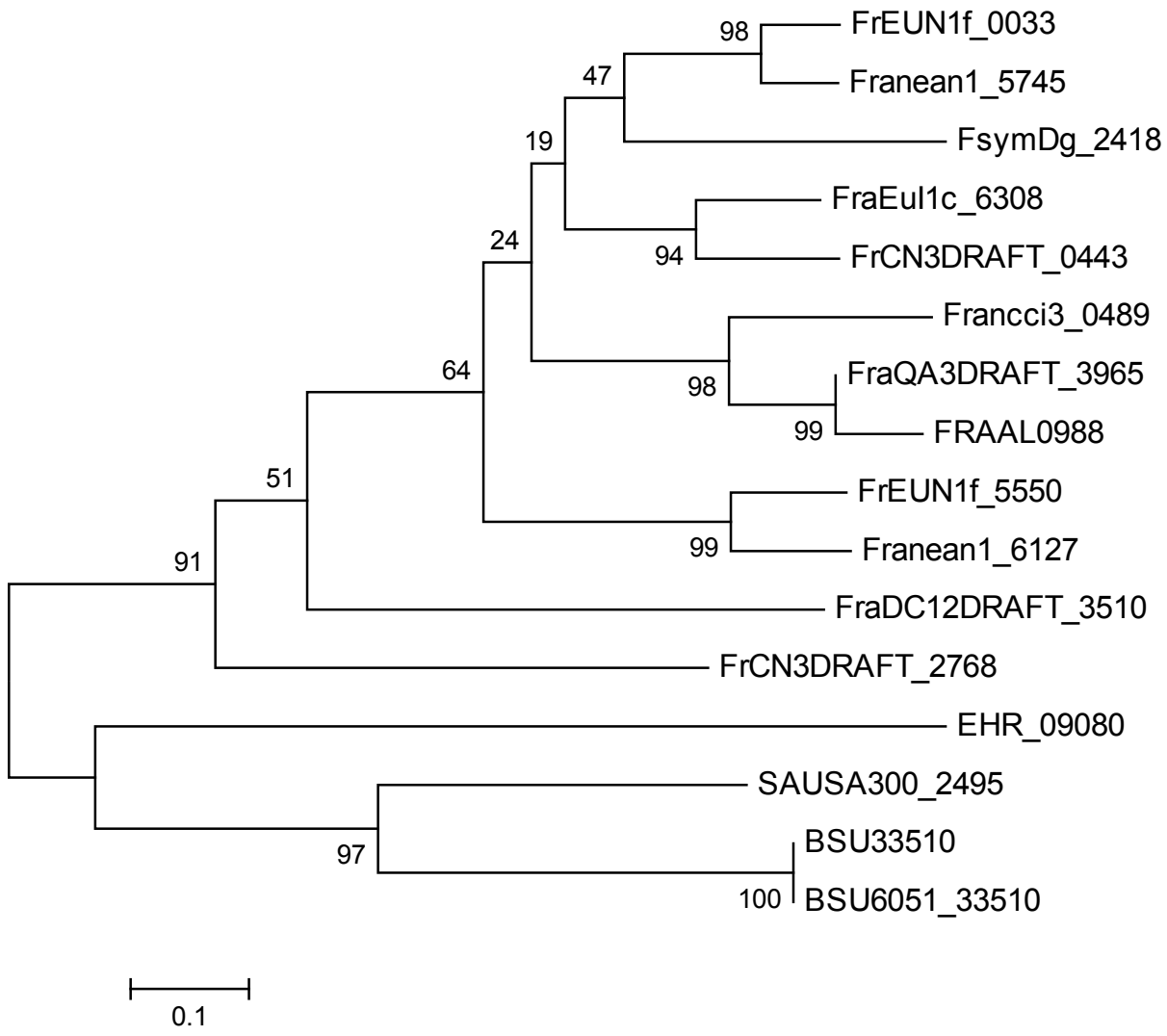

Figure S4. Phylogenetic tree of CopZ protein sequences. The evolutionary history was inferred using the Neighbor-Joining method (Saitou and Nei 1987). The optimal tree with the sum of branch length = 4.84389874 is shown. The percentage of replicate trees in which the associated taxa clustered together in the bootstrap test (1000 replicates) are shown next to the branches (Felsenstein 1985). The tree is drawn to scale, with branch lengths in the same units as those of the evolutionary distances used to infer the phylogenetic tree. The evolutionary distances were computed using the JTT matrix-based method (Jones et al. 1992) and are in the units of the number of amino acid substitutions per site. The analysis involved 16 amino acid sequences. All ambiguous positions were removed for each sequence pair. There were a total of 82 positions in the final dataset. Evolutionary analyses were conducted in MEGA5 (Tamura et al. 2011). Besides the *Frankia* CopZ protein sequences, the following protein sequences were used: *E. hire* (EHR_09080), *B. subtilis* strain 168 (BSU33510), *B. subtilis* 6051HGW (BSU6051_33510), and *S. aureus* USA300 (SAUSA300_2495).

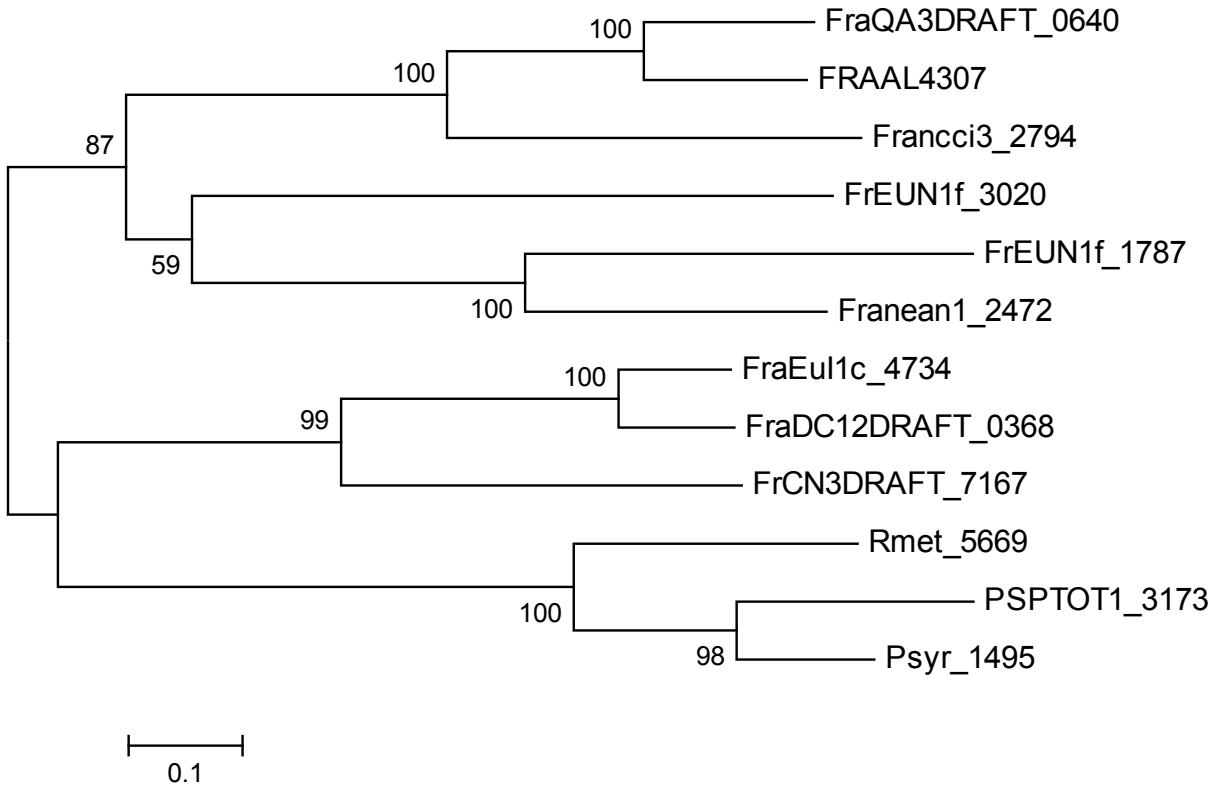

Figure S5. Phylogenetic tree of CopC protein sequences. The evolutionary history was inferred using the Neighbor-Joining method (Saitou and Nei 1987). The optimal tree with the sum of branch length = 5.05269378 is shown. The percentage of replicate trees in which the associated taxa clustered together in the bootstrap test (1000 replicates) are shown next to the branches (Felsenstein 1985). The tree is drawn to scale, with branch lengths in the same units as those of the evolutionary distances used to infer the phylogenetic tree. The evolutionary distances were computed using the JTT matrix-based method (Jones et al. 1992) and are in the units of the number of amino acid substitutions per site. The analysis involved 12 amino acid sequences. All ambiguous positions were removed for each sequence pair. There were a total of 287 positions in the final dataset. Evolutionary analyses were conducted in MEGA5 (Tamura et al. 2011). Besides the *Frankia* CopC protein sequences, the following protein sequences were used: *C. metallidurans* Rmet_3524), *Pseudomonas syringae pv. tomato* (PSPTOT1_3173), *Pseudomonas syringae pv. syringae* B728a (Psyr_1495).

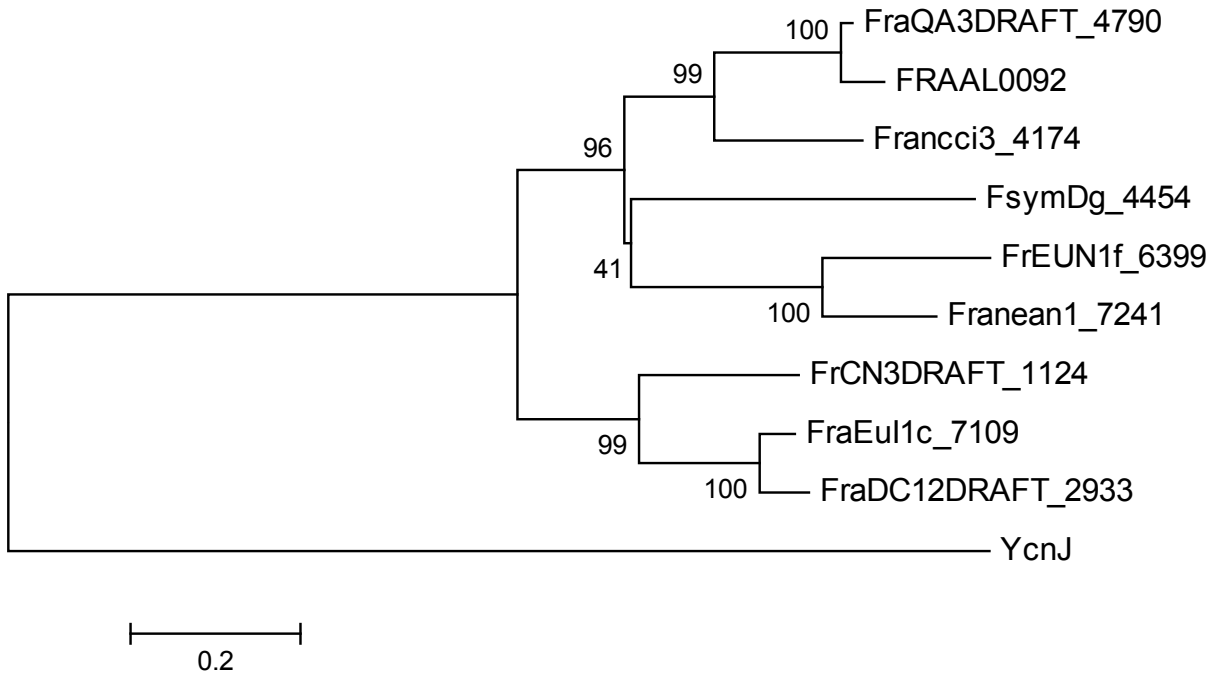

Figure S6. Phylogenetic tree of CopCD protein sequences. The evolutionary history was inferred using the Neighbor-Joining method (Saitou and Nei 1987). The bootstrap consensus tree inferred from 1000 replicates is taken to represent the evolutionary history of the taxa analyzed (Felsenstein 1985). Branches corresponding to partitions reproduced in less than 50% bootstrap replicates are collapsed. The percentage of replicate trees in which the associated taxa clustered together in the bootstrap test (1000 replicates) are shown next to the branches. The evolutionary distances were computed using the JTT matrix-based method (Jones et al. 1992) and are in the units of the number of amino acid substitutions per site. The analysis involved 10 amino acid sequences. All ambiguous positions were removed for each sequence pair. There were a total of 860 positions in the final dataset. Evolutionary analyses were conducted in MEGA5 (Tamura et al. 2011). Besides the *Frankia* CopCD protein sequences, the *B. subtilis* 6051HGW (YcnJ) was used.

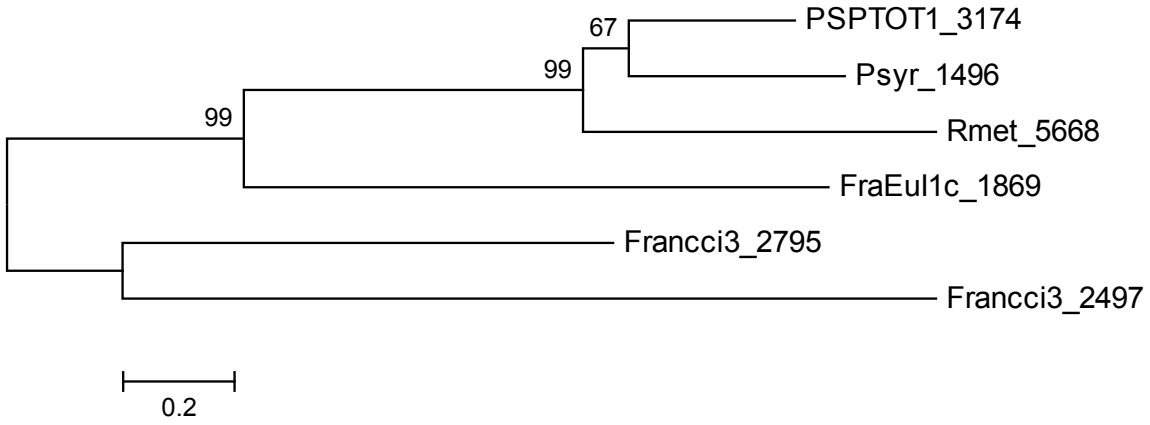

Figure S7. Phylogenetic tree of CopD protein sequences. The evolutionary history was inferred using the Neighbor-Joining method (Saitou and Nei 1987). The optimal tree with the sum of branch length = 6.04157447 is shown. The percentage of replicate trees in which the associated taxa clustered together in the bootstrap test (1000 replicates) are shown next to the branches (Felsenstein 1985). The tree is drawn to scale, with branch lengths in the same units as those of the evolutionary distances used to infer the phylogenetic tree. The evolutionary distances were computed using the JTT matrix-based method (Jones et al. 1992) and are in the units of the number of amino acid substitutions per site. The analysis involved 6 amino acid sequences. All ambiguous positions were removed for each sequence pair. There were a total of 767 positions in the final dataset. Evolutionary analyses were conducted in MEGA5 (Tamura et al. 2011). Besides the *Frankia* CopD protein sequences, the following protein sequences were used: *C. metallidurans* Rmet_5668), *P. syringae pv. tomato* (PSPTOT1_3174), *P. syringae pv. syringae* B728a (Psyr_1496).

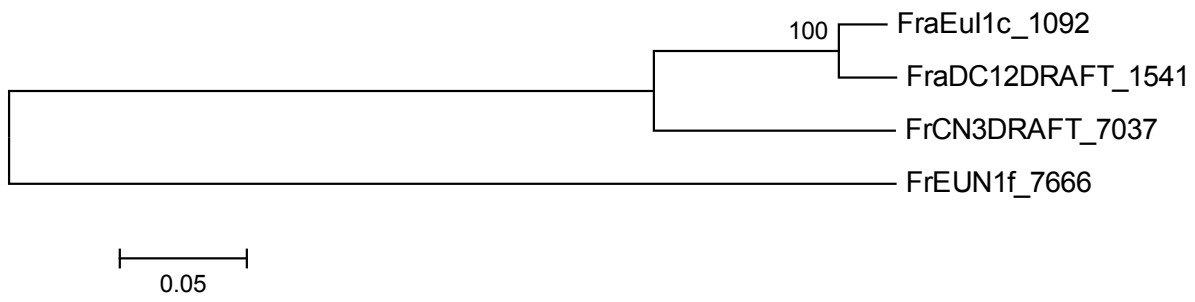

Figure S8. Phylogenetic tree of the sulfate-binding periplasmic transport protein sequences. The evolutionary history was inferred using the Neighbor-Joining method (Saitou and Nei 1987). The optimal tree with the sum of branch length = 0.81330031 is shown. The percentage of replicate trees in which the associated taxa clustered together in the bootstrap test (1000 replicates) are shown next to the branches (Felsenstein 1985). The tree is drawn to scale, with branch lengths in the same units as those of the evolutionary distances used to infer the phylogenetic tree. The evolutionary distances were computed using the JTT matrix-based method (Jones et al. 1992) and are in the units of the number of amino acid substitutions per site. The analysis involved 4 amino acid sequences. All ambiguous positions were removed for each sequence pair. There were a total of 387 positions in the final dataset. Evolutionary analyses were conducted in MEGA5 (Tamura et al. 2011).

References:

Felsenstein J (1985) Confidence-Limits on Phylogenies - an Approach Using the Bootstrap. Evolution 39(4):783-791 doi:Doi 10.2307/2408678

Jones DT, Taylor WR, Thornton JM (1992) The rapid generation of mutation data matrices from protein sequences. Computer applications in the biosciences : CABIOS 8(3):275-82

Saitou N, Nei M (1987) The Neighbor-Joining method - a new method for reconstructing phylogenetic trees. Mol Biol Evol 4(4):406-425

Tamura K, Peterson D, Peterson N, Stecher G, Nei M, Kumar S (2011) MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. Mol Biol Evol 28(10):2731-2739 doi:DOI 10.1093/molbev/msr121