

Figure S1. Alphaproteobacterial HpnP and species tree reconciliation. The 16S rRNA species tree (black) is overlaid with the HpnP tree (blue). The reconciliation was completed in Jane using 1, 2, and 3 as the event costs for loss, duplication, and horizontal transfer respectively. Vertical transfer – open circle, horizontal transfer – closed circle, loss – dashed line, red – best reconciliation, yellow – other equally good reconciliations.

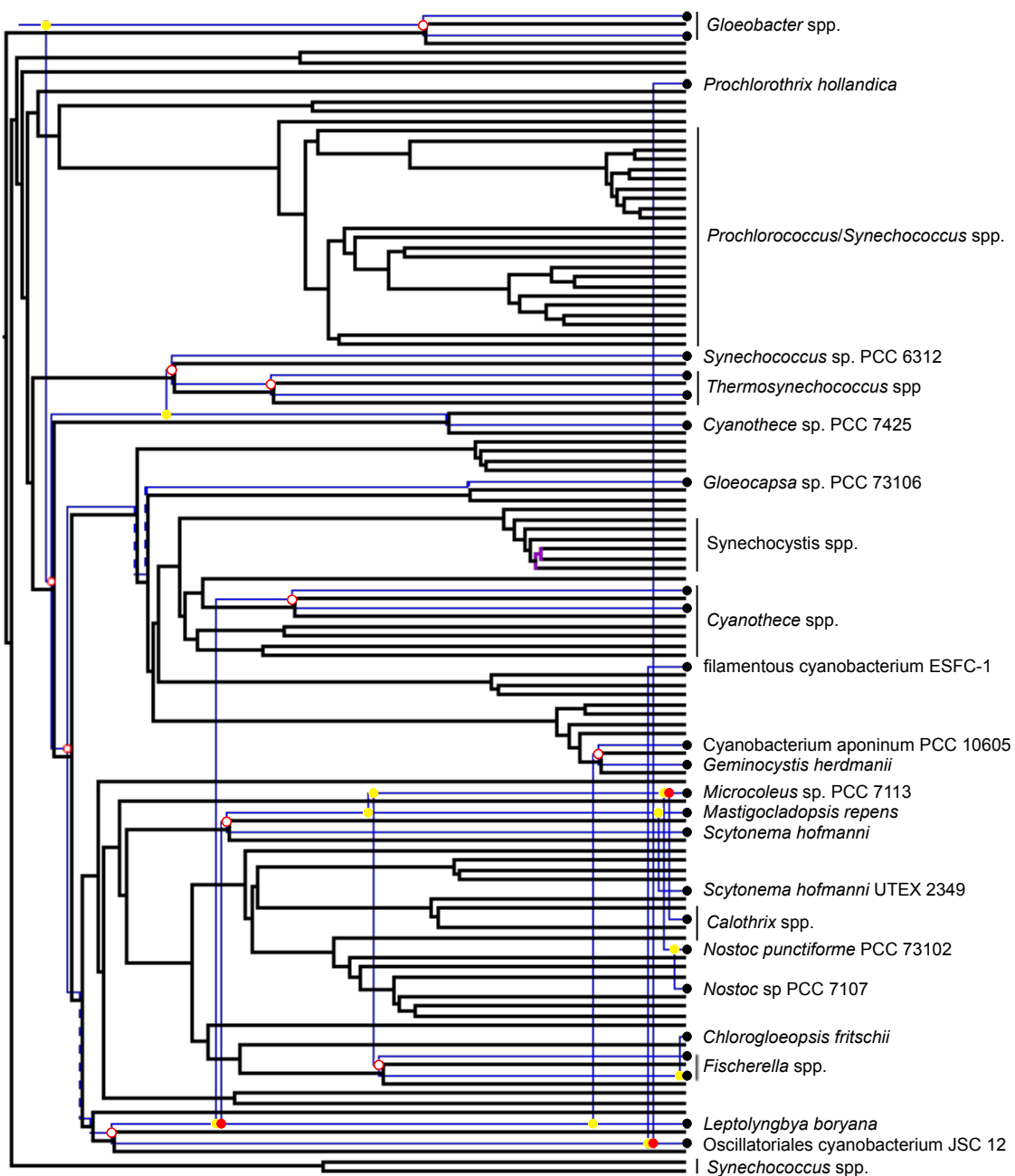


Figure S2. Cyanobacterial HpnP and species tree reconciliation. The 16S rRNA species tree (black) is overlaid with the HpnP tree (blue). The reconciliation was completed in Jane using 1, 2, and 3 as the event costs for loss, duplication, and horizontal transfer respectively. Vertical transfer – open circle, horizontal transfer – closed circle, loss – dashed line, red – best reconciliation, yellow – other equally good reconciliations, purple – polytomy due to identical sequences.

Table S1. Environmental genome walking primers

Primer Name	Target	Method	Primer Sequence
IGU1iPCR1F	U1 IG1	Inverse PCR	AATGATCTGCTGTGGGGGTC
IGU1iPCR1R	U1 IG1	Inverse PCR	GGCAGAAAGAGCGGGGTTAT
IGU1iPCR2F	U1 IG1	Inverse PCR	CGTTCCGTGTGGCTATCCAA
IGU1iPCR2R	U1 IG1	Inverse PCR	CGGTGCAATTTGCCTGTGAA
IGU1SP1F	U1 IG1	SpeedUp Forward	AACAAAGTAGATGGCTCCCG
IGU1SP2F	U1 IG1	SpeedUp Forward	AATGATCTGCTGTGGGGGTC
IGU1SP3F	U1 IG1	SpeedUp Forward	CGTTCCGTGTGGCTATCCAA

Table S2. Improved phylogenetic diversity.

Phylogeny	Number of unique sequences	PD score*	Colless's Imbalance†
Welander et al., 2010	28	8.95	0.387(48)
This study, genomic sequences only	115	14.02	0.111(2)
This study, genomic and environmental sequences	117	14.07	0.132(4)

*Phylogenetic diversity (PD) score is the sum of all branch lengths.

†Colless's Imbalance ranges from 0(even) to 1(lateralized). The percentile of the target tree's imbalance is found in parentheses and is based on an equiprobable distribution of random trees (n=1000).

Table S3. List of sequences used with gi numbers.

Dataset	GI	Species Name
Outgroup 1	320160959	<i>Anaerolinea thermophila</i> UNI-1
	39995655	<i>Geobacter sulfurreducens</i> PCA
	57233930	<i>Dehalococcoides ethenogenes</i> 195
	523470333	<i>Desulfovibrio</i> sp. X2
	193212385	<i>Chlorobaculum parvum</i> NCIB 8327
Outgroup 2	645069916	<i>Opitutaceae bacterium</i> TAV5
	501344729	<i>Opitutus terrae</i>
	496472502	<i>Desulfovibrio</i> sp. FW1012B
	501524087	<i>Geobacter bemidjiensis</i>
	506253585	<i>Desulfomicrobium baculatum</i>
	647376358	Dehalococcoidia bacterium DscP2
	504856057	<i>Dehalobacter</i> sp. DCA
	654862234	<i>Desulfatibacillum aliphaticivorans</i>
	655124705	<i>Desulfonatronum lacustre</i>
	501443562	<i>Chlorobium limicola</i>
HpnP ingroup	492877294	<i>Afipia broomeae</i> ATCC 49717
	488798710	<i>Afipia clevelandensis</i> ATCC 49720
	488803967	<i>Afipia felis</i> ATCC 53690
	639244164	<i>Afipia</i> sp. (639244164)
	640480562	<i>Afipia</i> sp. (640480562)
	496697395	<i>Afipia</i> sp. 1NLS2
	571918263	<i>Afipia</i> sp. P52-10
	501352804	<i>Beijerinckia indica</i> subsp. indica ATCC 9039
	497421586	Bradyrhizobiaceae bacterium SG-6C
	499398312	<i>Bradyrhizobium diazoefficiens</i> USDA 110
	517081761	<i>Bradyrhizobium elkanii</i> (517081761)
	654709199	<i>Bradyrhizobium elkanii</i> (654709199)
	654879908	<i>Bradyrhizobium elkanii</i> (654879908)
	654886985	<i>Bradyrhizobium elkanii</i> (654886985)
	654899386	<i>Bradyrhizobium elkanii</i> (54899386)
	504309727	<i>Bradyrhizobium japonicum</i> USDA 6
	636813563	<i>Bradyrhizobium japonicum</i> (636813563)
	648621071	<i>Bradyrhizobium japonicum</i> (648621071)
	654676491	<i>Bradyrhizobium japonicum</i> (654676491)
	654684416	<i>Bradyrhizobium japonicum</i> (654684416)
	654688277	<i>Bradyrhizobium japonicum</i> (654688277)
	654699060	<i>Bradyrhizobium japonicum</i> (654699060)
	654711923	<i>Bradyrhizobium japonicum</i> (654711923)
	654714597	<i>Bradyrhizobium japonicum</i> (654714597)
	654723587	<i>Bradyrhizobium japonicum</i> (654723587)
	505481583	<i>Bradyrhizobium oligotrophicum</i> S58
	653496802	<i>Bradyrhizobium</i> sp. (653496802)
	653552068	<i>Bradyrhizobium</i> sp. Ai1a-2
	639168996	<i>Bradyrhizobium</i> sp. ARR65
	500989536	<i>Bradyrhizobium</i> sp. BTAi1
	404267169	<i>Bradyrhizobium</i> sp. CCGE-LA001
	653477533	<i>Bradyrhizobium</i> sp. Cp5.3

Dataset	GI	Species Name
	544645606	<i>Bradyrhizobium</i> sp. DFCI-1
	608613371	<i>Bradyrhizobium</i> sp. DOA9
	640606658	<i>Bradyrhizobium</i> sp. DOA9
	653423211	<i>Bradyrhizobium</i> sp. Ec3.3
	639269491	<i>Bradyrhizobium</i> sp. OHSU_III
	500305918	<i>Bradyrhizobium</i> sp. ORS 278
	493662555	<i>Bradyrhizobium</i> sp. ORS 285
	496317759	<i>Bradyrhizobium</i> sp. ORS 375
	505725203	<i>Bradyrhizobium</i> sp. S23321
	496249482	<i>Bradyrhizobium</i> sp. STM 3809
	496253242	<i>Bradyrhizobium</i> sp. STM 3843
	656043459	<i>Bradyrhizobium</i> sp. th.b2
	639177735	<i>Bradyrhizobium</i> sp. Tv2a-2
	653447364	<i>Bradyrhizobium</i> sp. URHA0002
	653520650	<i>Bradyrhizobium</i> sp. URHA0013
	657881192	<i>Bradyrhizobium</i> sp. URHD0069
	494867534	<i>Bradyrhizobium</i> sp. WSM1253
	653392484	<i>Bradyrhizobium</i> sp. WSM1417
	653528340	<i>Bradyrhizobium</i> sp. WSM1743
	653462358	<i>Bradyrhizobium</i> sp. WSM2254
	648509787	<i>Bradyrhizobium</i> sp. WSM2793
	653436691	<i>Bradyrhizobium</i> sp. WSM3983
	648541515	<i>Bradyrhizobium</i> sp. WSM4349
	494882022	<i>Bradyrhizobium</i> sp. WSM471
	495419321	<i>Bradyrhizobium</i> sp. YR681
	518325440	<i>Calothrix</i> sp. PCC 7103
	499842390	Candidatus <i>Koribacter versatilis</i> Ellin345
	515386899	<i>Chlorogloeopsis fritschii</i>
	505032897	Cyanobacterium aponinum PCC 10605
	506435442	<i>Cyanothece</i> sp. PCC 7424
	501729632	<i>Cyanothece</i> sp. PCC 7425
	503089159	<i>Cyanothece</i> sp. PCC 7822
	517207725	filamentous cyanobacterium ESFC-1
	515365955	<i>Fischerella muscicola</i>
	652337927	<i>Fischerella</i> sp. PCC 9605
	515865344	<i>Geminocystis herdmanii</i>
	554673747	<i>Gloeobacter kilaueensis</i> JS1
	499454847	<i>Gloeobacter violaceus</i> PCC 7421
	493573935	<i>Gloeocapsa</i> sp. PCC 73106
	648457296	<i>Leptolyngbya boryana</i>
	557879987	<i>Leptolyngbya boryana</i> CCAP 1462/2
	648365781	<i>Mastigocladopsis repens</i>
	489693355	<i>Methylobacterium extorquens</i> PA1
	506304389	<i>Methylobacterium extorquens</i> CM4
	498372741	<i>Methylobacterium mesophilicum</i> SR1.6/6
	506408858	<i>Methylobacterium nodulans</i> ORS 2060
	501431528	<i>Methylobacterium populi</i> BJ001
	501277757	<i>Methylobacterium radiotolerans</i> JCM 2831

Dataset	GI	Species Name
	501287335	<i>Methylobacterium</i> sp. 4-46
	652919088	<i>Methylobacterium</i> sp. 10
	518940350	<i>Methylobacterium</i> sp. 285MFTsu5.1
	651602142	<i>Methylobacterium</i> sp. 77
	516687765	<i>Methylobacterium</i> sp. 88A
	494839951	<i>Methylobacterium</i> sp. GXF4
	516053564	<i>Methylobacterium</i> sp. MB200
	651611548	<i>Methylocapsa acidiphila</i>
	501586616	<i>Methylocella silvestris</i> BL2
	648235176	<i>Methylocystis parvus</i>
	519020310	<i>Methyloferula stellata</i>
	504998593	<i>Microcoleus</i> sp. PCC 7113
	499830400	<i>Nitrobacter hamburgensis</i> X14
	497485859	<i>Nitrobacter</i> sp. Nb-311A
	499634763	<i>Nitrobacter winogradskyi</i> Nb-255
	501377424	<i>Nostoc punctiforme</i> PCC 73102
	504925828	<i>Nostoc</i> sp. PCC 7107
	501559316	<i>Oligotropha carboxidovorans</i> OM5
	497453748	<i>Oscillatoriales cyanobacterium</i> JSC-12
	516314663	<i>Prochlorothrix hollandica</i>
	499472646	<i>Rhodopseudomonas palustris</i> CGA009
	499759880	<i>Rhodopseudomonas palustris</i> HaA2
	499791441	<i>Rhodopseudomonas palustris</i> BisB18
	499823235	<i>Rhodopseudomonas palustris</i> BisB5
	499982521	<i>Rhodopseudomonas palustris</i> BisA53
	501488665	<i>Rhodopseudomonas palustris</i> TIE-1
	503266706	<i>Rhodopseudomonas palustris</i> DX-1
	653026104	<i>Rhodopseudomonas palustris</i>
	653046421	Rhodospirillales bacterium URHD0088
	495664219	<i>Rhodovulum</i> sp. PH10
	648443022	<i>Scytonema hofmanni</i>
	657933118	<i>Scytonema hofmanni</i> UTEX 2349
	504936600	<i>Synechococcus</i> sp. PCC 6312
	499369037	<i>Thermosynechococcus elongatus</i> BP-1
	571030898	<i>Thermosynechococcus</i> sp. NK55a