Document S1. Associated data pertaining to phylogeny estimations of P-T4SSs.

Proteins utilized for phylogeny estimation of P-T4SSs

For 75 P-T4SSs encoded within diverse proteobacterial genomes, five proteins (VirB4, VirB8-VirB11) were selected for analysis (**Table 1**). Additionally, second copies of RvhB4, RvhB8 and RvhB9 encoded within Rickettsiales genomes were treated as a minimal T4SS (*rvh-II*), resulting in 90 P-T4SSs in total. NCBI GenBank accession numbers for all proteins are provided.

Phylogeny estimations of P-type type IV secretion systems (P-T4SSs)

Phylogenies were estimated from concatenated alignments of five components (VirB4, VirB8-VirB11), except for *rvhB-II*, which contains homologs to VirB4, VirB8 and VirB9 only (see text for further details on alignment and dataset construction) (**fig. 1**). (*A*) ML-based phylogeny estimated on the unmasked alignment (LG + gamma + *I*). This tree is further described in **fig. 2** of the manuscript. (*B*) ML-based phylogeny estimated on the unmasked alignment (WAG + gamma + *I*). (*C*) ML-based phylogeny estimated on the masked alignment (LG + gamma + *I*). (*D*) ML-based phylogeny estimated on the masked alignment (WAG + gamma + *I*). (*E*) Bayesian inference of phylogeny based on the masked alignment (CAT-GTR). All ML-based estimations were generated using RAxML v.7.2.8 (Stamatakis 2014), with branch support assessed with 1000 bootstrap pseudoreplications. Bayesian analysis employed PhyloBayes v3.3 (Lartillot et al. 2009), with branch support measured by frequency of clades within the posterior distribution of trees. See text for further details pertaining to dataset construction, alignment and ML and Bayesian search strategies. All trees were rooted with four P-T4SSs from species of *Epsilonproteobacteria*. *rvh-I* and *rvh-II* T4SSs are shaded light and dark gray, respectively. Orange circles denote the clade unifying the *rvh-I* and *rvh-II* paralogs. The *Bartonella vir*, *vbh* and *trw* T4SSs are each colored distinct shades of green.

REFERENCES

Lartillot N, Lepage T, Blanquart S. 2009. PhyloBayes 3: a Bayesian software package for phylogenetic reconstruction and molecular dating. Bioinformatics [Internet] 25:2286–2288. Available from: http://www.ncbi.nlm.nih.gov/pubmed/19535536

Stamatakis A. 2014. RAxML version 8: A tool for phylogenetic analysis and post-analysis of large phylogenies. Bioinformatics 30:1312–1313.

Table 1 (Document S1)

Table 1 (Document S1). Proteins utilized for phylogeny estimation of P-T4SSs.

Taxon [@]	**Locale** [~]	**VirB4** [^]	**VirB8** [$]	**VirB9**	**VirB10** [#]	**VirB11** [#]
Wolinella succinogenes str. DSM 1740	Chromosome	NP_907292	NP_907290	NP_907289	NP_907288	NP_907286
Campylobacter jejuni subsp. *jejuni* 81-176	pVir	NP_863348	NP_863298	NP_863299	NP_863300	NP_863302
Helicobacter pylori str. J99	Chromosome	NP_222737	NP_222756	NP_222757	NP_222758	NP_224034
Helicobacter pylori str. PeCan18B	Chromosome	AAM03033	AAM03034	AAM03035	AAM03036	AAM03040
Legionella pneumophila str. Corby	Chromosome	YP_001249518	YP_001249519	YP_001249520	YP_001249522	YP_001249515
Photobacterium profundum str. SS9	Chromosome	YP_133260	YP_133264	YP_133265	YP_133267	YP_133262
Mesorhizobium loti str. MAFF303099	Chromosome	NP_106918	NP_106921	NP_106922	NP_106923	NP_106916
Mesorhizobium loti str. MAFF303099	pMLb	NP_109464	NP_109461	NP_109460	NP_109459	NP_109467
"*Candidatus* Midichloria mitochondrii" str. IricVA	I; Chromosome	---[*]	YP_004679440	YP_004679865	YP_004679438	YP_004679437
Rickettsia typhi str. Wilmington	I; Chromosome	YP_067003	YP_067242	YP_067239	YP_067244	YP_067245
Rickettsia bellii str. RML369-C	I; Chromosome	YP_538436	YP_538184	YP_538182	YP_538186	YP_538187
Rickettsiaceae bacterium str. Os18	I; Chromosome	WP_019230792	WP_019230975	WP_019231305	WP_019230977	WP_019230978
Orientia tsutsugamushi str. Boryong	I; Chromosome	YP_001249140	YP_001248090	YP_001248270	YP_001248092	YP_001248093
Orientia tsutsugamushi str. Ikeda	I; Chromosome	YP_001937462	YP_001938263	YP_001937770	YP_001938261	YP_001938260
"*Candidatus* Xenolissoclinum pacificiensis" str. L6	I; Chromosome	ETO91781	ETO91952	ETO91160	ETO91954	ETO91955
Neorickettsia sennetsu str. Miyayama	I; Chromosome	YP_506730	YP_506619	YP_506105	YP_506617	YP_506616
Neorickettsia risticii str. Illinois	I; Chromosome	YP_003082035	YP_003081929	YP_003081432	YP_003081926	YP_003081925
Wolbachia endosymbiont (*Drosophila melanogaster*)	I; Chromosome	NP_966608	NP_965838	NP_966824	NP_965840	NP_965841
Wolbachia endosymbiont str. TRS (*Brugia malayi*)	I; Chromosome	YP_198627	YP_198110	YP_198421	YP_198112	YP_198113
Anaplasma marginale str. St. Maries	I; Chromosome	YP_153996	YP_154363	YP_153506	YP_154361	YP_154360
Anaplasma phagocytophilum str. HZ	I; Chromosome	YP_504979	YP_505898	YP_504712	YP_505896	YP_505895
Ehrlichia chaffeensis str. Arkansas	I; Chromosome	YP_507310	YP_506876	YP_507042	YP_506874	YP_506873
Ehrlichia canis str. Jake	I; Chromosome	YP_303167	YP_302671	YP_303416	YP_302669	YP_302668
"*Candidatus* Midichloria mitochondrii" str. IricVA	II; Chromosome	YP_004679293	YP_004679866	YP_004679439	---[*]	---[*]
Rickettsia bellii str. RML369-C	II; Chromosome	YP_537250	YP_538183	YP_538185	---[*]	---[*]
Rickettsia typhi str. Wilmington	II; Chromosome	YP_067709	YP_067240	YP_067243	---[*]	---[*]
Rickettsiaceae bacterium str. Os18	II; Chromosome	WP_019231477	WP_019231304	WP_019230976	---[*]	---[*]
Orientia tsutsugamushi str. Ikeda	II; Chromosome	YP_001936694	YP_001937769	YP_001938262	---[*]	---[*]

Table 1 (Document S1)

Taxon [@]	**Locale** [~]	**VirB4** [^]	**VirB8** [$]	**VirB9**	**VirB10** [#]	**VirB11** [#]
Orientia tsutsugamushi str. Boryong	II; Chromosome	YP_001248630	YP_001248271	YP_001248091	---[*]	---[*]
"*Candidatus* Xenolissoclinum pacificiensis" str. L6	II; Chromosome	ETO91710	ETO91161	ETO91953	---[*]	---[*]
Neorickettsia sennetsu str. Miyayama	II; Chromosome	YP_506641	YP_506106	YP_506618	---[*]	---[*]
Neorickettsia risticii str. Illinois	II; Chromosome	YP_003081947	YP_003081433	YP_003081928	---[*]	---[*]
Wolbachia endosymbiont str. TRS (*Brugia malayi*)	II; Chromosome	YP_198580	YP_198471	YP_198111	---[*]	---[*]
Wolbachia endosymbiont (*Drosophila melanogaster*)	II; Chromosome	WP_010963077	WP_010082435	WP_010962300	---[*]	---[*]
Anaplasma marginale str. St. Maries	II; Chromosome	YP_154174	YP_153956	YP_154362	---[*]	---[*]
Anaplasma phagocytophilum str. HZ	II; Chromosome	YP_505671	YP_505033 [$]	YP_505897	---[*]	---[*]
Ehrlichia canis str. Jake	II; Chromosome	YP_303468	YP_303086	YP_302670	---[*]	---[*]
Ehrlichia chaffeensis str. Arkansas	II; Chromosome	YP_507826	YP_507390	YP_506875	---[*]	---[*]
Legionella longbeachae str. D-4968	Chromosome	EEZ95953	EEZ95948	EEZ95947	EEZ95946	EEZ95945
Legionella pneumophila subsp. *pneumophila* str. Philadelphia 1	Chromosome	YP_095285	YP_095280	YP_095279	YP_095278	YP_095277
Legionella pneumophila str. Paris	Chromosome	YP_122517	YP_122522	YP_122523	YP_122524	YP_122525
Legionella pneumophila str. Lens	Chromosome	YP_125529	YP_125534	YP_125535	YP_125536	YP_125537
Burkholderia multivorans str. CGD1	Chromosome	WP_006402578	WP_006402582	WP_006402583	WP_006402584	WP_006402586
Pseudomonas fluorescens str. HK44	pUTK21	EXF91007	EXF91012	EXF91013	EXF91014	EXF91015
Pseudomonas putida (unknown strain)	pWW0	NP_542928	NP_542923	NP_542922	NP_542921	NP_542920
Ralstonia solanacearum str. CMR15	pRSC35	YP_005995023	YP_005995018	YP_005995017	YP_005995016	YP_005995015
Xanthomonas citri subsp. *citri*	pXcB	NP_942610	NP_942615	NP_942616	NP_942617	NP_942618
Cupriavidus basilensis str. OR16	Chromosome	WP_006156658	WP_006156661	WP_006156662	WP_006156663	WP_006156664
Shinella zoogloeoides str. DD12	Chromosome	EYR82770	EYR82763	EYR82762	EYR82761	EYR82760
Nitrobacter hamburgensis str. X14	"plasmid 2"	YP_571861	YP_571856	YP_571855	YP_571854	YP_571853
Mesorhizobium loti str. MAFF303099	pMLa	NP_085793	NP_085797	NP_085798	NP_085799	NP_085800
Sandarakinorhabdus limnophila (unknown strain)	Chromosome	WP_022678817	WP_022678815	WP_022678814	WP_022678813	WP_022678812
Sphingomonas phyllosphaerae (unknown strain)	Chromosome	WP_022685007	WP_022685009	WP_022685010	WP_022685011	WP_022685012
Sphingobium chlorophenolicum str. L-1	Chromosome	YP_004553681	YP_004553679	YP_004553678	YP_004553677	YP_004553676
Gilvimarinus chinensis (unknown strain)	Chromosome	WP_020211024	WP_020211028	WP_020211029	WP_020211030	WP_020211031
Phenylobacterium zucineum str. HLK1	Chromosome	YP_002131189	YP_002131192	YP_002131193	YP_002131194	YP_002131195
Agrobacterium tumefaciens str. LBA4213 (Ach5)	pTi	AHK05284	AHK05288	AHK05289	AHK05290	AHK05291

Table 1 (Document S1)

Taxon[@]	Locale[~]	VirB4[^]	VirB8[$]	VirB9	VirB10[#]	VirB11[#]
Ochrobactrum anthropi str. ATCC 49188	Chromosome	YP_001369231	YP_001369236	YP_001369237	YP_001369238	YP_001369239
Agrobacterium fabrum str. C58	pAt	NP_396095	NP_396099	NP_396100	NP_396101	NP_396102
Sinorhizobium meliloti str. 1021	pSymA	NP_435962	NP_435958	NP_435957	NP_435956	NP_435955
Bartonella birtlesii str. LL-WM9	Chromosome	EJF74241	EJF74251	EJF74252	EJF74253	EJF74254
Bartonella tribocorum str. CIP 105476	Chromosome	CAK02346	CAK02353	CAK02354	CAK02355	CAK02356
Bartonella grahamii str. as4aup	Chromosome	YP_002972601	YP_002972609	YP_002972610	YP_002972611	YP_002972612
Bartonella grahamii str. as4aup	Chromosome	YP_002972815	YP_002972820	YP_002972821	YP_002972822	YP_002972823
Bartonella quintana str. Toulouse	Chromosome	YP_032622	YP_032626	YP_032627	YP_032628	YP_032629
Bartonella henselae str. Houston-1	Chromosome	YP_034053	YP_034057	YP_034058	YP_034059	YP_034060
Bartonella birtlesii str. LL-WM9	Chromosome	EJF74266	EJF74270	EJF74271	EJF74272	EJF74273
Bartonella grahamii str. as4aup	Chromosome	YP_002972359	YP_002972363	YP_002972364	YP_002972365	YP_002972366
Bartonella tribocorum str. CIP 105476	Chromosome	CAK02005	CAK02009	CAK02010	CAK02011	CAK02012
Pseudomonas putida str. LD209	pLD209	YP_008994870	YP_008994866	YP_008994865	YP_008994864	YP_008994863
Xylella fastidiosa str. 9a5c	pXF51	NP_061663	NP_061668	NP_061669	NP_061670	NP_061671
Aeromonas caviae str. HGB5	pFBAOT6	YP_067837	YP_067832	YP_067831	YP_067830	YP_067829
Stenotrophomonas maltophilia str. EPM1	Chromosome	EMF61225	EMF61232	EMF61231	EMF61230	EMF61229
Xanthomonas campestris pv. *campestris* str. 8004	Chromosome	AAY48705	AAY48698	AAY48699	AAY48700	AAY48701
Yersinia pestis biovar Microtus str. 91001	pCRY	NP_995423	NP_995427	NP_995428	NP_995429	NP_995430
Campylobacter jejuni subsp. *jejuni* 81-176	pTet	YP_063475	YP_063483	YP_063484	YP_063485	YP_063486
Campylobacter coli str. CVM N29710	pN29710-1	YP_008543104	YP_008543112	YP_008543113	YP_008543114	YP_008543115
Vibrio fischeri str. ES114	pES100	YP_207169	YP_207170	YP_207171	YP_207172	YP_207173
Bordetella pertussis str. Tohama I	Chromosome	NP_882289	NP_882292	NP_882293	NP_882294	NP_882295
Rhizosphere of *Medicago sativa* (alfalfa) *	pSB102	NP_361048	NP_361042	NP_361041	NP_361040	NP_361039
Brucella suis str. ATCC 23445	Chromosome	YP_001621919	YP_001621915	YP_001621914	YP_001621913	YP_001621912
Pseudomonas syringae pv. *syringae* str. A2	pPSR1	NP_940726	NP_940730	NP_940731	NP_940732	NP_940733
Erwinia amylovora str. UTRJ2	pEU30	NP_943280	NP_943283	NP_943284	NP_943285	NP_943286
Escherichia coli O25b:H4-ST131 str. EC958	pKC394	YP_003717495	YP_003717500	YP_003717501	YP_003717502	YP_003717503
Salmonella enterica subsp. *enterica* serovar Typhimurium	pR46 (IncN)	NP_511190	NP_511195	NP_511196	NP_511197	NP_511198
Bartonella quintana str. Toulouse	Chromosome	YP_032785	YP_032792	YP_032793	YP_032794	YP_032795

Table 1 (Document S1)

Taxon [@]	**Locale** [~]	**VirB4** [^]	**VirB8** [$]	**VirB9**	**VirB10** [#]	**VirB11** [#]
Bartonella henselae str. Houston-1	Chromosome	YP_034262	YP_034269	YP_034270	YP_034271	YP_034272
Bartonella birtlesii str. LL-WM9	Chromosome	EJF76413	EJF76425	EJF76426	EJF76427	EJF76428
Bartonella grahamii str. as4aup	Chromosome	YP_002972703	YP_002972716	YP_002972717	YP_002972718	YP_002972719
Bartonella tribocorum str. CIP 105476	Chromosome	CAK02481	CAK02496	CAK02497	CAK02498	CAK02499

[@] Taxa are listed from top to bottom as they appear in the phylogeny. Asterisk denotes ‘Rhizosphere of *Medicago sativa* (alfalfa)’, for which the laboratory host is *Sinorhizobium meliloti* strain FP2. All sequence IDs are NCBI GenBank protein accession numbers.

[~] Unless plasmids are specified in the NCBI annotation, ‘chromosome’ is listed. For the *rvh* T4SS, see text for description of *rvh-I* and *rvh-II*.

[^] Underlined numbers refer to VirB3/VirB4 fusions. The asterisk refers to a missing gene encoding an RvhB4a homolog for “*Candidatus* Midichloria mitochondrii” str. IricVA.

[$] For *Anaplasma phagocytophilum* str. HZ (*rvh-II*) a homolog from str. HZ2 (YP_008332568) was used since it contains a full length N-terminal sequence.

[#] For *rvh-II*, asterisks denote the RvhB10 and RvhB11 proteins that were used for *rvh-I* only (see text for details).

A

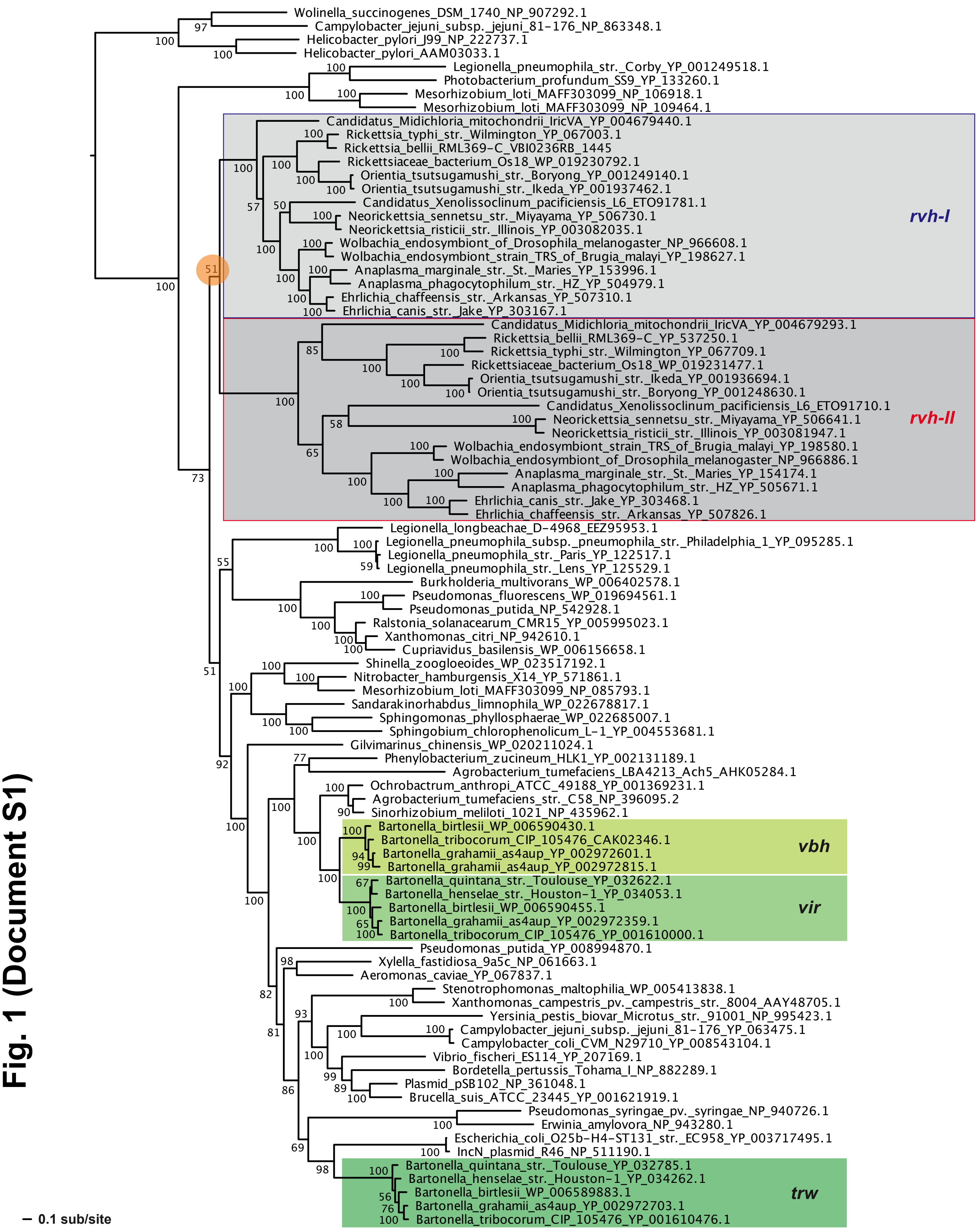

Fig. 1 (Document S1)

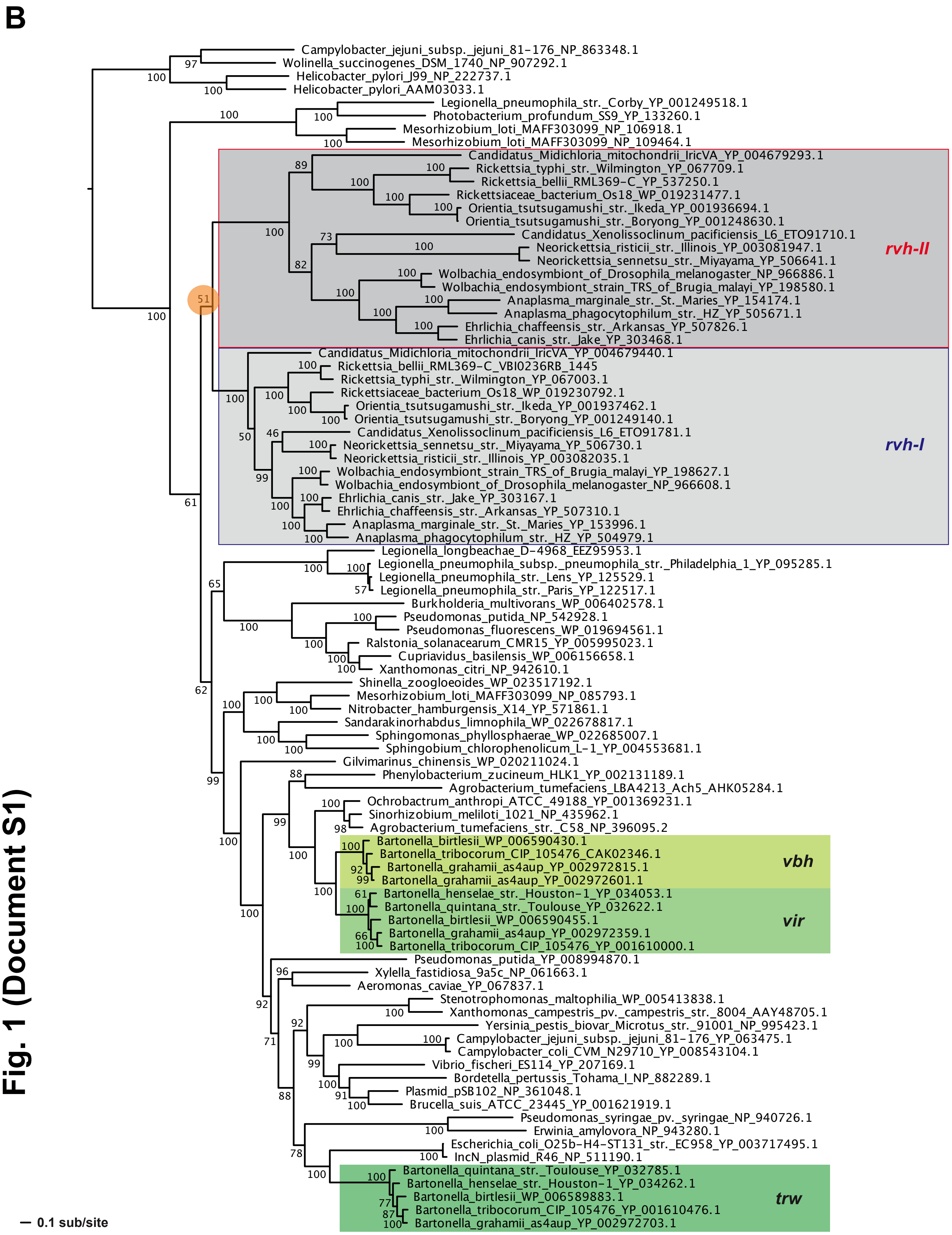

Fig. 1 (Document S1)

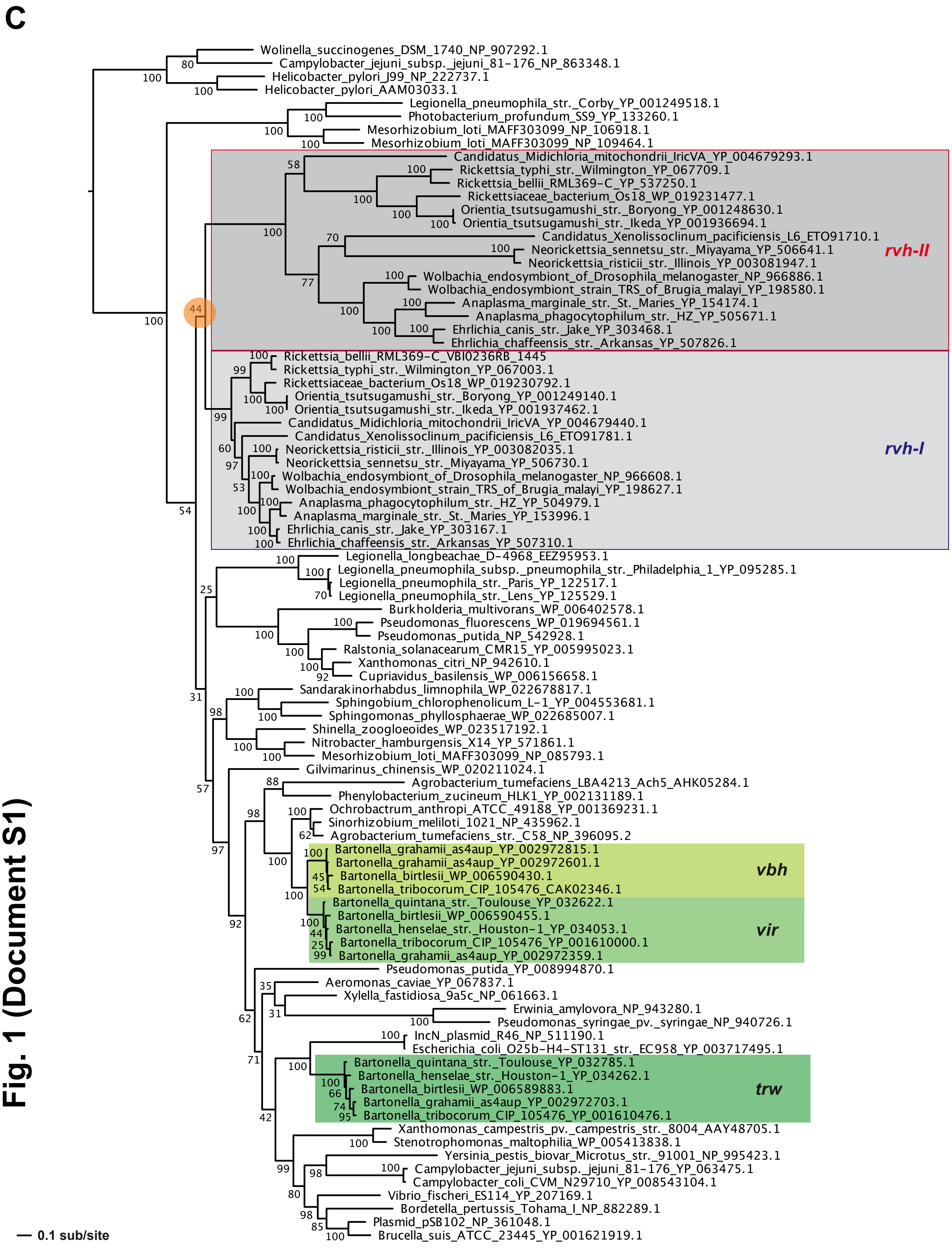

Fig. 1 (Document S1)

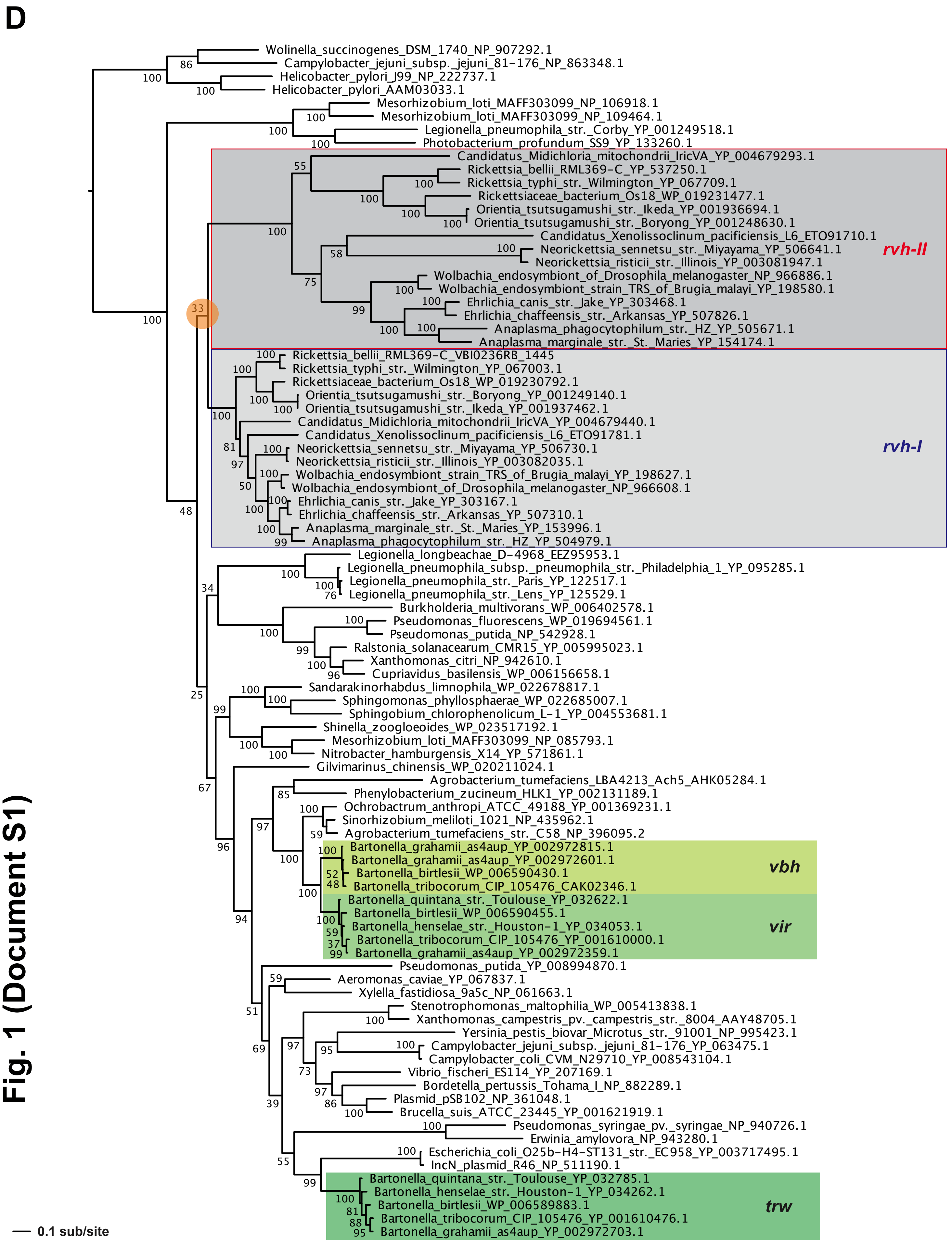

Fig. 1 (Document S1)

Fig. 1 (Document S1)