Supplementary material and methods

Decontamination and *de novo* assembly of *Ridgeia* symbiont genomes

First assembly and data quality assessment

Raw reads were first assembled at Canada's Michael Smith Genome Sciences Centre using Abyss and preliminary quality assessments of the resulting scaffolds performed with Quast (1). All but one of the samples resulted in assemblies of poor quality (N50 < 1000). The very high quantity of scaffolds containing repetitive sequences, characteristic of eukaryote DNA, indicated significant contamination from worm genetic material. The highest quality sample (Ind 11; N50 = 61 758) had a bimodal distribution of GC% and a heterogeneous distribution of k-mer frequencies, evidence of a strong, distinct symbiont signal with fewer contaminating scaffolds than the other assemblies. This assembly was uploaded into the IMG-ER platform for gene calling and decontamination.

Decontamination

The Abyss assembly of Ind 11 was manually curated from potential contaminating scaffolds using a combination of reference dependant and independent methods according to JGI's Single Cell Data Decontamination guide. First, we considered scaffolds of lengths ≥ 1000 bp and used the IMG/ER embedded k-mer Frequency Analysis tool to generate a k-mer plot using the following parameters; window size=1000, fragment step=100, oligomer size=4 bp, minimum variation=10. Scaffolds out or extending out of the main scaffold cluster were tagged as contaminants if their respective gene counts were ≤ 1 (large intergenic sequences are not expected in bacterial genomes which have high gene density) or if they did not contained at least one gene phylogenetically affiliated to the genome of the close relative *Riftia pachyptila* symbiont (2). Contaminant scaffolds were then removed from the genome assembly. Finally, scaffolds smaller than 1000 bp with at least one gene phylogenetically affiliated to the genome of the *Riftia pachyptila* symbiont were added to the curated assembly. The resulting curated assembly GC% frequency distribution was unimodal with a mode at 60% of GC as for the *Riftia* and *Tevnia* symbiont genomes.

De novo assembly

Raw paired-end reads were filtered with Prinseq (3), to remove nucleotides with a quality score inferior to 20 on both ends of each read. Ind 11 reads were then mapped onto the curated assembly using Bowtie2 (4) and the mapped reads were extracted with Samtools (5).

Next, these 'good' reads were assembled with SPAdes (6) using the following k-mer sizes: 26, 55, 67, 85, 89, 95 & 99. Scaffolding was performed with SSPACE-standard (7) and resulted in a high quality assembly of the metagenome of *Ridgeia* symbionts. We will refer to this assembly as the *Ridgeia* 1 symbiont.

Finally, we used the *Ridgeia* 1 symbiont assembly to map and extract the symbiont reads from the Hiseq and Miseq data from the five other samples. These data were pooled to increase the depth of coverage and the subsequent assembly quality. The reads were assembled with SPAdes using the same parameters as for *Ridgeia* 1 symbiont but with the additional k-mer size 127 to account for the longer reads generated by the MiSeq platform. This pooled assembly is referred to as the *Ridgeia* 2 symbiont.

References

1. **Gurevich A**, **Saveliev V**, **Vyahhi N**, **Tesler G**. 2013. QUAST: quality assessment tool for genome assemblies. Bioinformatics **29**:1072–1075.

2. **Robidart JC**, **Bench SR**, **Feldman RA**, **Novoradovsky A**, **Podell SB**, **Gaasterland T**, **Allen EE**, **Felbeck H**. 2008. Metabolic versatility of the *Riftia pachyptila* endosymbiont revealed through metagenomics. Environ Microbiol **10**:727–737.

3. **Schmieder R**, **Edwards R**. 2011. Quality control and preprocessing of metagenomic datasets. Bioinformatics **27**:863–864.

4. **Langmead B**, **Salzberg SL**. 2012. Fast gapped-read alignment with Bowtie 2. Nat Methods **9**:357–359.

5. **Li H**, **Handsaker B**, **Wysoker A**, **Fennell T**, **Ruan J**, **Homer N**, **Marth G**, **Abecasis G**, **Durbin R**, **Subgroup 1000 Genome Project Data Processing**. 2009. The sequence alignment/map format and SAMtools. Bioinformatics **25**:2078–2079.

6. **Bankevich A**, **Nurk S**, **Antipov D**, **Gurevich AA**, **Dvorkin M**, **Kulikov AS**, **Lesin VM**, **Nikolenko SI**, **Pham S**, **Prjibelski AD**, **Pyshkin AV**, **Sirotkin AV**, **Vyahhi N**, **Tesler G**, **Alekseyev MA**, **Pevzner PA**. 2012. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. J Comput Biol **19**:455–477.

7. **Boetzer M**, **Henkel CV**, **Jansen HJ**, **Butler D**, **Pirovano W**. 2011. Scaffolding pre-assembled contigs using SSPACE. Bioinformatics **27**:578–579.

Supplementary tables

Table S1 Accessory genome exclusive to *Ridgeia* symbionts (*Ridgeia* 1 and *Ridgeia* 2 symbiont genome assemblies).

Table S2 Accessory genome exclusive to the East Pacific Rise symbionts (*Tevnia*, *Rifita* 1, and *Rifita* 2 symbiont genome assemblies).

Table S3 Accessory genome found in *Riftia* symbionts (*Rifita* 1 and *Rifita* 2 symbiont genome assemblies).

Table S4 Accessory genome found in 9°N symbionts (*Tevnia* and *Rifita* 2 symbiont genome assemblies but not in *Riftia* 1 symbionts).

Table S5 Genes of particular interest. Modified or complemented from Gardebrecht *et al.* (1).

1. **Gardebrecht A, Markert S, Sievert SM, Felbeck H, Thürmer A, Albrecht D, Wollherr A, Kabisch J, Le Bris N, Lehmann R**. 2011. Physiological homogeneity among the endosymbionts of *Riftia pachyptila* and *Tevnia jerichonana* revealed by proteogenomics. ISME J **6**:766–776.

Back to list of tables

Table S1 Accessory genome exclusive to *Ridgeia* symbionts (*Ridgeia* 1 and *Ridgeia* 2 symbiont genome assemblies). Total size = 95 456 bp; n.a. = Not Annotated; presence of sequence or part of sequence but no genes were called. **(Continued)**

Exclusive sequence Reference contig sequence length Start - End positions (in reference contig)	**Low flanking conservation**	**Locus-tag in *Ridgeia* 1 symbionts**	**Locus-tag in *Ridgeia* 2 symbionts**	**Product**
Exclusive sequence 1 Ga0074115_146 17457 0 - 17457		Ga0074115_1461	Ga0076813_11512	Curli production assembly/transport component CsgG
		Ga0074115_1462	Ga0076813_11513	Zn-binding Pro-Ala-Ala-Arg (PAAR) domain-containing protein, involved in TypeVI secretion
		Ga0074115_1463	Ga0076813_11514	Methyltransferase domain-containing protein
		Ga0074115_1464	Ga0076813_14753	Protein of unknown function (DUF1795)
		Ga0074115_1465	Ga0076813_14752	Hypothetical protein
		Ga0074115_1466	Ga0076813_14751; Ga0076813_16801	Type VI secretion protein, EvpB/VC_A0108 family/Type VI secretion protein, VC_A0107 family
		Ga0074115_1467	Ga0076813_16802	Type VI secretion system secreted protein VgrG
		Ga0074115_1468	Ga0076813_16803	Hypothetical protein
		Ga0074115_1469	Ga0076813_16804	Type VI secretion system protein ImpA
		Ga0074115_14610	Ga0076813_16805	Type VI secretion system protein ImpB
		Ga0074115_14611	Ga0076813_13723	Type VI secretion system protein ImpC
		Ga0074115_14612	Ga0076813_13722	Type VI secretion system secreted protein Hcp
		Ga0074115_14613	Ga0076813_13721; Ga0076813_14484	ImpE protein
		Ga0074115_14614	Ga0076813_14483	Type VI secretion system protein ImpF
		Ga0074115_14615	Ga0076813_14482	Type VI secretion system protein ImpG
		Ga0074115_14616	Ga0076813_14481	Type VI secretion system protein ImpH
Exclusive sequence 2 Ga0074115_125 11590 33173 - 44763	✓	Ga0074115_12524	Ga0076813_10858	WxcM-like, C-terminal
		Ga0074115_12525	Ga0076813_10859	UDP-3-O-[3-hydroxymyristoyl] glucosamine N-acyltransferase
		Ga0074115_12526	Ga0076813_108510	Hypothetical protein
		Ga0074115_12527	Ga0076813_108511	Hydroxymethylglutaryl-CoA synthase
		Ga0074115_12528	Ga0076813_108512	Acyl carrier protein
		Ga0074115_12529	Ga0076813_108513	dTDP-4-amino-4,6-dideoxygalactose transaminase
		Ga0074115_12530	Ga0076813_108514	Glycosyl transferase family 2
		Ga0074115_12531	Ga0076813_108515	Lipopolysaccharide transport system permease protein
		Ga0074115_12532	Ga0076813_108516	ABC-2 type transport system ATP-binding protein
		Ga0074115_12533	Ga0076813_108517	Glycosyl transferases group 1
		Ga0074115_12534	Ga0076813_108518	Hypothetical protein
		Ga0074115_12535	Ga0076813_108519	Glycosyltransferase involved in cell wall bisynthesis
		Ga0074115_12536	Ga0076813_108520	Hypothetical protein
Exclusive sequence 3 Ga0074115_157 8235 0 - 8235		Ga0074115_1571	Ga0076813_10168	Hypothetical protein
		Ga0074115_1572	Ga0076813_10167	Hypothetical protein
		Ga0074115_1573	Ga0076813_10166	Hypothetical protein
		Ga0074115_1574	Ga0076813_10165	Hypothetical protein
		Ga0074115_1575	Ga0076813_10164	TerY-C metal binding domain
		Ga0074115_1576	Ga0076813_10163	Hypothetical protein
		Ga0074115_1577	Ga0076813_10162	Mobile mystery protein A
		Ga0074115_1578	Ga0076813_10162; Ga0076813_16633	Mobile mystery protein B
		Ga0074115_1579	Ga0076813_16632	Fic/DOC family protein
		Ga0074115_15710	Ga0076813_14733; Ga0076813_16631	Hypothetical protein
		Ga0074115_15711	Ga0076813_14732	Predicted nuclease of the RNAse H fold, HicB family
		Ga0074115_15712	Ga0076813_14731	Protein of unknown function (DUF3732)
		Ga0074115_15713	n.a.	Hypothetical protein
Exclusive sequence 4 Ga0074115_147 8311 0 - 8311		n.a.	Ga0076813_142210	Transposase
		Ga0074115_1471	Ga0076813_14229	Protein of unknown function DUF91
		Ga0074115_1472	Ga0076813_14228	Hypothetical protein
		Ga0074115_1473	Ga0076813_14227	Relaxase/Mobilisation nuclease domain-containing protein
		Ga0074115_1474	Ga0076813_14226	Hypothetical protein
		Ga0074115_1475	Ga0076813_14225	Hypothetical protein
		Ga0074115_1476	Ga0076813_14224	Hypothetical protein
		Ga0074115_1477	Ga0076813_14223	Hypothetical protein
		Ga0074115_1478	Ga0076813_14222	tRNA-splicing ligase RtcB
Exclusive sequence 5 Ga0074115_105 7528	✓	Ga0074115_105110	Ga0076813_13394	Hypothetical protein
		Ga0074115_105111	Ga0076813_11781; Ga0076813_14331;	Methyltransferase domain-containing protein

Back to list of tables

Table S1 Accessory genome exclusive to *Ridgeia* symbionts (*Ridgeia* 1 and *Ridgeia* 2 symbiont genome assemblies). Total size = 95 456 bp; n.a. = Not Annotated; presence of sequence or part of sequence but no genes were called. **(Continued)**

Exclusive sequence **Reference contig** **sequence length** **Start - End positions** **(in reference contig)**	**Low flanking conservation**	**Locus-tag in *Ridgeia* 1 symbionts**	**Locus-tag in *Ridgeia* 2 symbionts**	**Product**
121389 - 128917			Ga0076813_13762	
		Ga0074115_105112	Ga0076813_13761; Ga0076813_12832	Hypothetical protein
		Ga0074115_105113	Ga0076813_12831	Hypothetical protein
Exclusive sequence 6 Ga0074115_132 7011 25125 - 32136		Ga0074115_13218	Ga0076813_14266	Carbohydrate binding domain-containing protein
		Ga0074115_13219	Ga0076813_14265	Hypothetical protein
		Ga0074115_13220	Ga0076813_14264	Hypothetical protein
		Ga0074115_13221	Ga0076813_14263	GntR family transcriptional regulator
		Ga0074115_13222	Ga0076813_14262	Tripartite-type tricarboxylate transporter, receptor component TctC
		Ga0074115_13223	Ga0076813_14261; Ga0076813_13331	Tripartite tricarboxylate transporter TctB family protein
Exclusive sequence 7 Ga0074115_109 4694 1942 - 6636	✓	Ga0074115_1094	Ga0076813_16722	CRISPR-associated protein, Csd2 family
		Ga0074115_1095	Ga0076813_16721; Ga0076813_10881; Ga0076813_14464	CRISPR-associated protein Csd1
		Ga0074115_1096	Ga0076813_14463	CRISPR-associated protein, Cas5d family
		Ga0074115_1098	Ga0076813_14462	Hypothetical protein
		Ga0074115_1097	Ga0076813_14461	Hypothetical protein
Exclusive sequence 8 Ga0074115_141 4127 0 - 4127		Ga0074115_1411; Ga0074115_1412	Ga0076813_10531	Hypothetical protein
		Ga0074115_1413	Ga0076813_10532	Hypothetical protein
Exclusive sequence 9 Ga0074115_103 2816 141098 - 143914	✓	Ga0074115_103137	Ga0076813_16812	Type I secretion system ABC transporter, HlyB family
		Ga0074115_103138	Ga0076813_16811; Ga0076813_14381	Hemolysin D
		Ga0074115_103139	Ga0076813_14382	Hypothetical protein
Exclusive sequence 10 Ga0074115_114 2482 43812 - 46294		Ga0074115_11438	Ga0076813_127316	Hypothetical protein
		Ga0074115_11439	Ga0076813_127315	Hypothetical protein
		n.a.	Ga0076813_127314	Hypothetical protein
		Ga0074115_11440	Ga0076813_127313	Integrase core domain-containing protein
		Ga0074115_11441	Ga0076813_127312	HTH-like domain-containing protein
		Ga0074115_11442	Ga0076813_127310	Hypothetical protein
		Ga0074115_11443	Ga0076813_12739	Transposase
Exclusive sequence 11 Ga0074115_170 2271 0 - 2271		Ga0074115_1701	n.a.	Cation diffusion facilitator family transporter
		Ga0074115_1702	Ga0076813_13551	AcrB/AcrD/AcrF family protein
		Ga0074115_1703	Ga0076813_13552	AcrB/AcrD/AcrF family protein
		Ga0074115_1704	n.a.	Hypothetical protein
Exclusive sequence 12 Ga0074115_132 2223 3349 – 5572		Ga0074115_1641	Ga0076813_14711	PKD repeat-containing protein
Exclusive sequence 13 Ga0074115_161 2011 3372 – 5383	✓	Ga0074115_1614	Ga0076813_13865	Transposase
		Ga0074115_1615	Ga0076813_13866	Transposase IS66 family protein
Exclusive sequence 14 Ga0074115_159 1926 0 – 1926		Ga0074115_1591	Ga0076813_14181	Secreted protein containing bacterial Ig-like domain and vWFA domain
Exclusive sequence 15 Ga0074115_123 1831 335 – 2166	✓	Ga0074115_1232	Ga0076813_14196	TIGR02452 family protein
		Ga0074115_1233	Ga0076813_14195	Hypothetical protein
Exclusive sequence 16		n.a.	Ga0076813_14051	Transcriptional regulatory protein, C terminal

Back to list of tables

Table S1 Accessory genome exclusive to *Ridgeia* symbionts (*Ridgeia* 1 and *Ridgeia* 2 symbiont genome assemblies). Total size = 95 456 bp; n.a. = Not Annotated; presence of sequence or part of sequence but no genes were called. **(Continued)**

Exclusive sequence **Reference contig** **sequence length** **Start - End positions (in reference contig)**	**Low flanking conservation**	**Locus-tag in *Ridgeia* 1 symbionts**	**Locus-tag in *Ridgeia* 2 symbionts**	**Product**
Ga0074115_140 1736 0 – 1736		Ga0074115_10577 Ga0074115_10578	Ga0076813_14053 Ga0076813_14052	Carbon-nitrogen hydrolase ubiE/COQ5 methyltransferase family protein
Exclusive sequence 17 Ga0074115_132 1726 0 - 1726		n.a. Ga0074115_1321	Ga0076813_15291 Ga0076813_15292	Transposase Hypothetical protein
Exclusive sequence 18 Ga0074115_180 1396 0 – 1396		Ga0074115_1801 Ga0074115_1802	Ga0076813_11412 Ga0076813_11411	Spherulation-specific family 4 Hypothetical protein
Exclusive sequence 19 Ga0074115_102 1154 129891 – 131045	✓	Ga0074115_102124 Ga0074115_102125	Ga0076813_169035 Ga0076813_169034	Putative transposase Putative transposase
Exclusive sequence 20 Ga0074115_165 903 2388 – 3291		Ga0074115_1654	Ga0076813_10561	Hypothetical protein
Exclusive sequence 21 Ga0074115_126 856 11285 – 12141		Ga0074115_1262 Ga0074115_1263	Ga0076813_11072 Ga0076813_11402	Hypothetical protein Hypothetical protein
Exclusive sequence 22 Ga0074115_143 852 15512 – 16364		Ga0074115_14312 Ga0074115_14313	Ga0076813_15414 Ga0076813_15415	Nitrate/nitrite transport system substrate-binding protein Hypothetical protein
Exclusive sequence 23 Ga0074115_156 628 6057 – 6685	✓	Ga0074115_1567	Ga0076813_13514	Aminoacyl-tRNA editing domain-containing protein
Exclusive sequence 24 Ga0074115_156 591 4884 – 5475	✓	Ga0074115_1566	Ga0076813_13515	Hypothetical protein
Exclusive sequence 25 Ga0074115_118 586 61784 – 62370		Ga0074115_11853	Ga0076813_16776	Chemoreceptor zinc-binding domain-containing protein
Exclusive sequence 26 Ga0074115_112 515 0 - 515		Ga0074115_1121	Ga0076813_16143	Hypothetical protein

Back to list of tables

Table S2 Accessory genome exclusive to the East Pacific Rise symbionts (*Tevnia*, *Rifita* 1, and *Rifita* 2 symbiont genome assemblies). Total size = 141 335 bp; n.a. = Not Annotated; presence of sequence or part of sequence but no genes were called. **(Continued)**

Exclusive sequence Reference contig Sequence length Start - End positions (in reference contig)	Low flanking conservation	Locus-tag in *Tevnia* symbionts	Locus-tag in *Riftia* 1 symbionts	Locus-tag in *Riftia* 2 symbionts	Product
Exclusive sequence 1 AFZB01000005 33791 69865 - 103656		TevJSym_ae00610; TevJSym_ae00620	Rifp1Sym_dq00030	Rifp2Sym_fc00010	Anaerobic dimethyl sulfoxide reductase subunit A
		n.a.	n.a.	Rifp2Sym_fc00020	Hypothetical protein
		TevJSym_ae00630	n.a.	Rifp2Sym_fc00030	Hypothetical protein
		TevJSym_ae00640	Rifp1Sym_dq00020	Rifp2Sym_fc00040	Transcriptional regulator, ArsR family
		TevJSym_ae00650	Rifp1Sym_gl00020	Rifp2Sym_fc00050	Fatty acid hydroxylase
		TevJSym_ae00660	Rifp1Sym_gl00040	Rifp2Sym_fc00060	Uncharacterized membrane protein YdjX, TVP38/TMEM64 family, SNARE-associated domain
		TevJSym_ae00670	Rifp1Sym_gl00060	Rifp2Sym_fc00070	Hypothetical protein
		TevJSym_ae00680	Rifp1Sym_gl00080	Rifp2Sym_fc00080	Uncharacterized protein involved in oxidation of intracellular sulfur
		TevJSym_ae00690	n.a.	n.a.	Adenylate cyclase
		TevJSym_ae00700	n.a.	Rifp2Sym_cv00010	Cyclic nucleotide binding protein
		TevJSym_ae00710	Rifp1Sym_ew00090	Rifp2Sym_cv00020	Hypothetical protein
		TevJSym_ae00720	Rifp1Sym_ew00080	Rifp2Sym_cv00030	Glyoxylase, beta-lactamase superfamily II
		TevJSym_ae00730	Rifp1Sym_ew00070; Rifp1Sym_ew00060; Rifp1Sym_ew00050	Rifp2Sym_cv00040	NADPH-dependent glutamate synthase beta chain
		TevJSym_ae00740	Rifp1Sym_ew00040	Rifp2Sym_cv00050	Pyridine nucleotide-disulphide oxidoreductase
		TevJSym_ae00750	Rifp1Sym_ew00030	Rifp2Sym_cv00060	Hypothetical protein
		TevJSym_ae00760	Rifp1Sym_ew00020	Rifp2Sym_cv00070; Rifp2Sym_cv00080	Hydroxylamine dehydrogenase
		n.a.	Rifp1Sym_ew00010	n.a.	Hypothetical protein
		n.a.	n.a.	Rifp2Sym_cv00090	Hypothetical protein
		TevJSym_ae00770	n.a.	Rifp2Sym_cv00100	Hypothetical protein
		TevJSym_ae00780	n.a.	n.a.	Hypothetical protein
		TevJSym_ae00790	Rifp1Sym_bf00360	Rifp2Sym_cv00110	D-alanyl-lipoteichoic acid acyltransferase DltB, MBOAT superfamily
		TevJSym_ae00800	Rifp1Sym_bf00350	Rifp2Sym_cv00120	Hypothetical protein
		TevJSym_ae00810	Rifp1Sym_bf00340	Rifp2Sym_cv00130	DNA-binding transcriptional response regulator, NtrC family, contains REC, AAA-type ATPase, and a Fis-type DNA-binding domains
		TevJSym_ae00820	Rifp1Sym_bf00330	Rifp2Sym_cv00140	Signal transduction histidine kinase
		TevJSym_ae00830	Rifp1Sym_bf00320	Rifp2Sym_cv00150	ParE-like toxin of type II toxin-antitoxin system
		TevJSym_ae00840	n.a.	Rifp2Sym_cv00160	ParD-like antitoxin of type II toxin-antitoxin system
		TevJSym_ae00850	Rifp1Sym_bf00310	Rifp2Sym_cv00170	Cytochrome b
		TevJSym_ae00860	Rifp1Sym_bf00300	Rifp2Sym_cv00180	Protein of unknown function (DUF1924)
		TevJSym_ae00870	Rifp1Sym_bf00290	Rifp2Sym_cv00190	Dihaem cytochrome c

Back to list of tables

Table S2 Accessory genome exclusive to the East Pacific Rise symbionts (*Tevnia*, *Rifita* 1, and *Rifita* 2 symbiont genome assemblies). Total size = 141 335 bp; n.a. = Not Annotated; presence of sequence or part of sequence but no genes were called. **(Continued)**

Exclusive sequence **Reference contig** **Sequence length** **Start - End positions (in reference contig)**	**Low flanking conservation**	**Locus-tag in *Tevnia* symbionts**	**Locus-tag in *Riftia* 1 symbionts**	**Locus-tag in *Riftia* 2 symbionts**	**Product**
		TevJSym_ae00880	Rifp1Sym_bf00280	Rifp2Sym_bo00320	His Kinase A (phospho-acceptor) domain-containing protein
		TevJSym_ae00890	Rifp1Sym_bf00270	Rifp2Sym_bo00310	Two-component system, NtrC family, response regulator AtoC
		TevJSym_ae00900	Rifp1Sym_bf00260	Rifp2Sym_bo00300	Hypothetical protein
		TevJSym_ae00910	Rifp1Sym_bf00250	Rifp2Sym_bo00290	Major Facilitator Superfamily protein
		TevJSym_ae00920	n.a.	Rifp2Sym_bo00280	Cytochrome C'
		TevJSym_ae00930	Rifp1Sym_bf00240	Rifp2Sym_bo00270	Glyceraldehyde-3-phosphate dehydrogenase 3
Exclusive sequence 2 AFZB01000004 24661 114255 - 138916	✓	TevJSym_ad01110	Rifp1Sym_dc00020	n.a.	
		TevJSym_ad01120	Rifp1Sym_cn00190	Rifp2Sym_az00350; Rifp2Sym_az00340; Rifp2Sym_az00330; Rifp2Sym_az00320; Rifp2Sym_az00310	ATP-dependent helicase HrpA
		TevJSym_ad01130	Rifp1Sym_cn00180	Rifp2Sym_az00300	SEC-C motif-containing protein
		TevJSym_ad01140	Rifp1Sym_cn00170	Rifp2Sym_az00290	Protein of unknown function (DUF4124)
		TevJSym_ad01150	Rifp1Sym_cn00160	Rifp2Sym_az00280	Glycosyltransferase involved in cell wall bisynthesis
		TevJSym_ad01160	Rifp1Sym_cn00150	Rifp2Sym_az00270	PilZ domain-containing protein
		TevJSym_ad01170	Rifp1Sym_cn00140	Rifp2Sym_az00260	Protein of unknown function (DUF1631)
		TevJSym_ad01180	Rifp1Sym_cn00130	Rifp2Sym_az00250	Hypothetical protein
		TevJSym_ad01190	Rifp1Sym_cn00120	Rifp2Sym_az00240	HD-like signal output (HDOD) domain, no enzymatic activity
		TevJSym_ad01200	Rifp1Sym_cn00110	Rifp2Sym_az00230	Sulfate adenylyltransferase
		TevJSym_ad01210	Rifp1Sym_cn00100	Rifp2Sym_az00220	Aconitase
		TevJSym_ad01220	Rifp1Sym_cn00090	Rifp2Sym_az00210	Hypothetical protein
		TevJSym_ad01230	Rifp1Sym_cn00080	Rifp2Sym_az00180; Rifp2Sym_az00190; Rifp2Sym_az00200	Short chain dehydrogenase reductase Sdr
		n.a.	Rifp1Sym_cn00070	Rifp2Sym_az00170	Integral membrane protein
		n.a.	Rifp1Sym_cn00060	Rifp2Sym_az00160	Integral membrane protein
		TevJSym_ad01240	Rifp1Sym_cn00050	Rifp2Sym_az00150	Putative membrane protein
Exclusive sequence 3 AFZB01000005 22442 0 - 22442		TevJSym_ae00010	Rifp1Sym_el00010	Rifp2Sym_et00150	Proteic killer suppression protein
		TevJSym_ae00020	Rifp1Sym_el00020	Rifp2Sym_et00140	Addiction module antidote protein, HigA family
		TevJSym_ae00030	Rifp1Sym_el00030; Rifp1Sym_el00040; Rifp1Sym_el00050	Rifp2Sym_et00130; Rifp2Sym_et00120	Toprim domain-containing protein
		TevJSym_ae00040	Rifp1Sym_el00060	Rifp2Sym_et00110	Hypothetical protein
		n.a.	Rifp1Sym_el00070	Rifp2Sym_et00100	Spermidine synthase
		TevJSym_ae00050	Rifp1Sym_el00080	Rifp2Sym_et00090	Phage integrase family protein

Back to list of tables

Table S2 Accessory genome exclusive to the East Pacific Rise symbionts (*Tevnia*, *Rifita* 1, and *Rifita* 2 symbiont genome assemblies). Total size = 141 335 bp; n.a. = Not Annotated; presence of sequence or part of sequence but no genes were called. **(Continued)**

Exclusive sequence **Reference contig** **Sequence length** **Start - End positions** **(in reference contig)**	**Low flanking conservation**	**Locus-tag in *Tevnia* symbionts**	**Locus-tag in *Riftia* 1 symbionts**	**Locus-tag in *Riftia* 2 symbionts**	**Product**
		TevJSym_ae00060	Rifp1Sym_el00090	Rifp2Sym_et00080	Hypothetical protein
		TevJSym_ae00070	Rifp1Sym_el00100	Rifp2Sym_et00070	Hypothetical protein
		TevJSym_ae00080; TevJSym_ae00090	Rifp1Sym_el00110	Rifp2Sym_et00060	Virulence-associated protein I
		TevJSym_ae00100	Rifp1Sym_el00120	Rifp2Sym_et00050	Hypothetical protein
		TevJSym_ae00110	Rifp1Sym_el00130	Rifp2Sym_et00040	Bacteriophage phi gp55-like protein
		TevJSym_ae00120	Rifp1Sym_el00140	Rifp2Sym_et00030	Hypothetical protein
		TevJSym_ae00130	Rifp1Sym_el00150	Rifp2Sym_et00020	Hypothetical protein
		TevJSym_ae00140	Rifp1Sym_el00160	Rifp2Sym_et00010	Hypothetical protein
		TevJSym_ae00150	Rifp1Sym_dx00010	n.a.	Hypothetical protein
		TevJSym_ae00160	Rifp1Sym_dx00020	Rifp2Sym_ee00010	Hypothetical protein
		TevJSym_ae00170	Rifp1Sym_dx00030	n.a.	Hypothetical protein
		TevJSym_ae00180	Rifp1Sym_dx00040; Rifp1Sym_dx00050; Rifp1Sym_dx00060	Rifp2Sym_ee00020; Rifp2Sym_ee00030	Hypothetical protein
		TevJSym_ae00190	Rifp1Sym_dx00070	Rifp2Sym_ee00040	Hypothetical protein
		TevJSym_ae00200	Rifp1Sym_dx00080; Rifp1Sym_dx00090; Rifp1Sym_dx00100	Rifp2Sym_ee00050; Rifp2Sym_ee00060; Rifp2Sym_ee00070	Phage tape measure protein
		n.a.	n.a.	Rifp2Sym_ee00080	Hypothetical protein
		n.a.	Rifp1Sym_dx00110	n.a.	Hypothetical protein
		TevJSym_ae00210	Rifp1Sym_dx00120	Rifp2Sym_ee00090	Hypothetical protein
		TevJSym_ae00220	Rifp1Sym_dx00130	Rifp2Sym_ee00100	Hypothetical protein
		TevJSym_ae00230	Rifp1Sym_dx00140	Rifp2Sym_ee00110	Hypothetical protein
Exclusive sequence 4 AFZB01000004 9533 17066 - 26599	✓	TevJSym_ad00160	Rifp1Sym_co00130	Rifp2Sym_cu00010	Hypothetical protein
		TevJSym_ad00170	Rifp1Sym_co00140	Rifp2Sym_fm00020	Protein of unknown function (DUF323)
		TevJSym_ad00180	Rifp1Sym_co00150; Rifp1Sym_co00160	Rifp2Sym_fm00010	Sulfatase-modifying factor enzyme 1
		TevJSym_ad00190	Rifp1Sym_co00160	n.a.	Hypothetical protein
		TevJSym_ad00200	Rifp1Sym_co00170	Rifp2Sym_fu00090; Rifp2Sym_fu00080	Protein phosphatase
		TevJSym_ad00210	Rifp1Sym_co00180	Rifp2Sym_fu00070	Protein phosphatase ImpM
Exclusive sequence 5 AFZB01000008 7596 63197 - 70793		n.a.	Rifp1Sym_ev00020	n.a.	Aspartate aminotransferase
		TevJSym_ah00640	Rifp1Sym_ev00030	Rifp2Sym_ii00010	N-6 DNA Methylase
		TevJSym_ah00650	Rifp1Sym_ev00040	Rifp2Sym_ii00020; Rifp2Sym_ii00030; Rifp2Sym_mm00020; Rifp2Sym_mm00010	Type I restriction enzyme M protein
		TevJSym_ah00660	Rifp1Sym_ev00050	n.a.	DEAD/DEAH box helicase

Back to list of tables

Table S2 Accessory genome exclusive to the East Pacific Rise symbionts (*Tevnia*, *Rifita* 1, and *Rifita* 2 symbiont genome assemblies). Total size = 141 335 bp; n.a. = Not Annotated; presence of sequence or part of sequence but no genes were called. **(Continued)**

Exclusive sequence Reference contig Sequence length Start - End positions (in reference contig)	**Low flanking conservation**	**Locus-tag in *Tevnia* symbionts**	**Locus-tag in *Riftia* 1 symbionts**	**Locus-tag in *Riftia* 2 symbionts**	**Product**
Exclusive sequence 6 AFZB01000004 7306 0 - 7306	✓	TevJSym_ad00010	Rifp1Sym_ez00030	Rifp2Sym_cu00190	Putative glycosyltransferase
		TevJSym_ad00020	Rifp1Sym_ez00020	Rifp2Sym_cu00180	Asparagine synthase (glutamine-hydrolysing)
		TevJSym_ad00030	Rifp1Sym_ez00010	Rifp2Sym_cu00170	Putative glycosyltransferase
		TevJSym_ad00040	n.a.	Rifp2Sym_cu00160	Glycosyltransferase, family 2
Exclusive sequence 7 AFZB01000038 6555 2424 - 8979	✓	TevJSym_bl00040	Rifp1Sym_ep00040	Rifp2Sym_gt00010	CRISPR-associated protein Cas7/Csd2, subtype I-C/DVULG
		TevJSym_bl00050	Rifp1Sym_ep00050	Rifp2Sym_gt00020	CRISPR-associated protein Csd1
		TevJSym_bl00060	Rifp1Sym_ep00060	Rifp2Sym_gt00030	CRISPR-associated protein Cas5d
		TevJSym_bl00070	Rifp1Sym_ep00070	Rifp2Sym_gt00040	CRISPR-associated endonuclease/helicase Cas3
Exclusive sequence 8 AFZB01000002 5745 158129 - 163874	✓	TevJSym_ab01580; TevJSym_ab01590	Rifp1Sym_bb00150	Rifp2Sym_ga00040	Putative transmembrane protein
		TevJSym_ab01600	Rifp1Sym_bb00140	Rifp2Sym_dh00160	Endoglucanase A
		TevJSym_ab01610	Rifp1Sym_bb00130	Rifp2Sym_dh00150	Putative transcriptional regulator
		n.a.	n.a.	Rifp2Sym_dh00140	Hypothetical protein
		TevJSym_ab01620	Rifp1Sym_bb00120	Rifp2Sym_dh00130	Hypothetical protein
		TevJSym_ab01630	Rifp1Sym_bb00110	Rifp2Sym_dh00120	Transposase insF for insertion sequence IS3
		TevJSym_ab01640	Rifp1Sym_bb00100	Rifp2Sym_dh00110	Integrase catalytic region
Exclusive sequence 9 AFZB01000037 3882 17256 - 21138	✓	TevJSym_bk00210	Rifp1Sym_cv00080	Rifp2Sym_bj00120	Adenosylhomocysteinase
		n.a.	n.a.	Rifp2Sym_bj00110	Hypothetical protein
		TevJSym_bk00220	Rifp1Sym_cv00090	Rifp2Sym_bj00100	Cobalt-zinc-cadmium efflux system protein
Exclusive sequence 10 AFZB01000045 3514 12768 - 16282	✓	TevJSym_bs00140	Rifp1Sym_gd00030	Rifp2Sym_da00120	Lipopolysaccharide transport system permease protein
		TevJSym_bs00150	Rifp1Sym_gd00020	Rifp2Sym_da00130	Lipopolysaccharide transport system ATP-binding protein
Exclusive sequence 11 AFZB01000064 2784 0 - 2784		TevJSym_cl00010	n.a.	Rifp2Sym_bz00010	Sodium/proton antiporter, NhaA family
		TevJSym_cl00020	Rifp1Sym_gw00010	Rifp2Sym_bz00020	Hypothetical protein
Exclusive sequence 12 AFZB01000018 2422 32751 - 35173		TevJSym_ar00310	Rifp1Sym_bh00100	n.a.	Hypothetical protein
		TevJSym_ar00320	n.a.	Rifp2Sym_ab00280	Hypothetical protein
Exclusive sequence 13 AFZB01000005 1918 119131 - 121049		TevJSym_ae01170	Rifp1Sym_bf00030	Rifp2Sym_bo00030	Hypothetical protein
Exclusive sequence 13 AFZB01000005 1918 119131 - 121049		TevJSym_ae01180; TevJSym_ae01190	Rifp1Sym_bf00020	Rifp2Sym_bo00020	Cytochrome C, class I

Back to list of tables

Table S2 Accessory genome exclusive to the East Pacific Rise symbionts (*Tevnia*, *Rifita* 1, and *Rifita* 2 symbiont genome assemblies). Total size = 141 335 bp; n.a. = Not Annotated; presence of sequence or part of sequence but no genes were called. **(Continued)**

Exclusive sequence Reference contig Sequence length Start - End positions (in reference contig)	**Low flanking conservation**	**Locus-tag in *Tevnia* symbionts**	**Locus-tag in *Riftia* 1 symbionts**	**Locus-tag in *Riftia* 2 symbionts**	**Product**
Exclusive sequence 14 AFZB01000005 1748 63277 - 65025		TevJSym_ae00560	Rifp1Sym_dq00090	Rifp2Sym_ep00040	Transcriptional regulatory protewin ZraR
Exclusive sequence 15 AFZB01000077 1739 0 - 1739		TevJSym_cx00010	Rifp1Sym_db00080	Rifp2Sym_jp00010	Transposase
Exclusive sequence 16 AFZB01000033 1378 27912 - 29290		TevJSym_bg00220	Rifp1Sym_bk00090	Rifp2Sym_bh00120	KAP P-loop domain containing protein
		TevJSym_bg00230	Rifp1Sym_bk00080	Rifp2Sym_bh00110	KAP family P-loop domain-containing protein
		TevJSym_bg00240	Rifp1Sym_bk00070	Rifp2Sym_bh00100	Hypothetical protein
Exclusive sequence 17 AFZB01000037 1027 10832 - 11859	✓	TevJSym_bk00130	Rifp1Sym_cv00020	Rifp2Sym_bj00190	Integrase core domain-containing protein
Exclusive sequence 18 AFZB01000016 984 79511 - 80495		TevJSym_ap00810	Rifp1Sym_fn00070	Rifp2Sym_fv00020	Putative transposase
Exclusive sequence 19 AFZB01000029 392 3174 - 3566		TevJSym_bc00050	Rifp1Sym_al00320	Rifp2Sym_aw00250	Transposase DDE domain-containing protein

Back to list of tables

Table S3 Accessory genome found in *Riftia* symbionts (*Rifita* 1 and *Rifita* 2 symbiont genome assemblies). Total size = 77 303 bp; n.a. = Not Annotated; presence of sequence or part of sequence but no genes were called; n.d. = Not Detected; sequence absent from the assembly. **(Continued)**

Exclusive sequence	Reference contig	Sequence length	Start - End (in reference contig)	Locus-tag in *Riftia* 1 symbionts	Locus-tag in *Riftia* 2 symbionts	Product
Exclusive sequence 1	AFOC01000021	39573	0 - 39573	Rifp1Sym_au00010	Rifp2Sym_gm00020	Type II restriction/modification system, DNA methylase subunit YeeA
				Rifp1Sym_au00020	Rifp2Sym_gm00010	Protein of unknown function (DUF4263)
				Rifp1Sym_au00030	Rifp2Sym_hf00010	TIGR02687 family protein
				Rifp1Sym_au00040	Rifp2Sym_hf00020	Transposase IS200 like
				Rifp1Sym_au00050	n.a.	Transposase IS200 like
				Rifp1Sym_au00060	Rifp2Sym_hf00030; Rifp2Sym_hv00040	ATP-dependent Lon protease
				Rifp1Sym_au00070	Rifp2Sym_hv00020	Hypothetical protein
				Rifp1Sym_au00080	n.a.	Uncharacterized protein YydD, contains DUF2326 domain
				Rifp1Sym_au00090	n.a.	HicB family protein
				Rifp1Sym_au00100	n.a.	Hypothetical protein
				Rifp1Sym_au00110	n.a.	Conjugal transfer mating pair stabilization protein TraG
				Rifp1Sym_au00120	n.a.	Conjugative transfer pilus assembly protein TraH
				Rifp1Sym_au00130	n.a.	Conjugal transfer pilus assembly protein TraF
				Rifp1Sym_au00140	n.a.	Type-F conjugative transfer system mating-pair stabilization protein TraN
				Rifp1Sym_au00150	n.a.	Conjugal transfer pilus assembly protein TrbC
				Rifp1Sym_au00160	n.a.	Conjugal transfer pilus assembly protein TraU
				Rifp1Sym_au00170	n.a.	Patatin-like phospholipase
				Rifp1Sym_au00180	n.a.	Conjugal transfer pilin signal peptidase TrbI
				Rifp1Sym_au00190	n.a.	Type-F conjugative transfer system protein (TrbI_Ftype)
				Rifp1Sym_au00200	n.a.	Conjugal transfer ATP-binding protein TraC
				Rifp1Sym_au00210	n.a.	Type IV conjugative transfer system lipoprotein TraV
				Rifp1Sym_au00220	n.a.	Thiol:disulfide interchange protein DsbC
				Rifp1Sym_au00230	n.a.	Conjugal transfer pilus assembly protein TraB
				Rifp1Sym_au00240	n.a.	Conjugal transfer pilus assembly protein TraK
				Rifp1Sym_au00250	n.a.	Conjugal transfer pilus assembly protein TraE
				Rifp1Sym_au00260	Rifp2Sym_ew00060	Conjugal transfer pilus assembly protein TraA
				Rifp1Sym_au00270	Rifp2Sym_ew00050	Outer membrane protein OmpA
	AFOC01000086	15845	0 - 15845	Rifp1Sym_dg00010	Rifp2Sym_ew00020	ABC-2 type transporter
				Rifp1Sym_dg00020	Rifp2Sym_ew00010	Hypothetical protein
				Rifp1Sym_dg00030	Rifp2Sym_kv00030	Putative cysteine desulfurase
				Rifp1Sym_dg00040	Rifp2Sym_kv00020	Hypothetical protein

Back to list of tables

Table S3 Accessory genome found in *Riftia* symbionts (*Rifita* 1 and *Rifita* 2 symbiont genome assemblies). Total size = 77 303 bp; n.a. = Not Annotated; presence of sequence or part of sequence but no genes were called; n.d. = Not Detected; sequence absent from the assembly. **(Continued)**

Exclusive sequence	Reference contig	Sequence length	Start - End (in reference contig)	Locus-tag in *Riftia* 1 symbionts	Locus-tag in *Riftia* 2 symbionts	Product
				Rifp1Sym_dg00050	Rifp2Sym_kv00010	Hypothetical protein
				Rifp1Sym_dg00060	Rifp2Sym_kj00010	Hypothetical protein
				Rifp1Sym_dg00070	Rifp2Sym_kj00020	Hypothetical protein
				Rifp1Sym_dg00080	n.d.	Hypothetical protein
				Rifp1Sym_dg00090	n.a.	Ribosomal large subunit pseudouridine synthase C
				Rifp1Sym_dg00100	Rifp2Sym_lw00010	Hypothetical protein
				Rifp1Sym_dg00110	n.a.	Hypothetical protein
				Rifp1Sym_dg00120	Rifp2Sym_me00010	Hypothetical protein
				Rifp1Sym_dg00130	n.a.	Putative plasmid stabilization system protein
				Rifp1Sym_dg00140	n.d.	Hypothetical protein
				Rifp1Sym_dg00150	n.d.	Hypothetical protein
				Rifp1Sym_dg00160	n.a.	Hypothetical protein
	AFOC010000153	5054	0 - 5054	Rifp1Sym_fv00010	Rifp2Sym_ja00010	Hypothetical protein
				Rifp1Sym_fv00020	Rifp2Sym_ja00020	Hypothetical protein
				Rifp1Sym_fv00030	Rifp2Sym_kh00010	Hypothetical protein
				Rifp1Sym_fv00040	n.a.	YbaK/prolyl-tRNA synthetase associated region
Exclusive sequence 2	AFOC010000112	9453	0 - 9453	Rifp1Sym_eg00010	Rifp2Sym_jz00010	CRISPR-associated protein Cas1
				Rifp1Sym_eg00020	n.a.	CRISPR system Cascade subunit CasE
				Rifp1Sym_eg00030	n.a.	CRISPR system Cascade subunit CasD
				Rifp1Sym_eg00040	Rifp2Sym_hp00010	CRISPR system Cascade subunit CasC
				Rifp1Sym_eg00050	Rifp2Sym_hp00030	CRISPR system Cascade subunit CasA
				Rifp1Sym_eg00060	Rifp2Sym_mb00010	CRISPR-associated endonuclease/helicase Cas3
Exclusive sequence 3	AFOC010000148	5239	0 - 5239	Rifp1Sym_fq00010	Rifp2Sym_fa00040	Protein of unknown function (DUF1788)
				Rifp1Sym_fq00020	Rifp2Sym_fa00030; Rifp2Sym_fa00020	Hypothetical protein
				Rifp1Sym_fx00010	n.a.	Transcriptional regulator, AlpA family
				Rifp1Sym_fx00020	n.a.	Integrating conjugative element relaxase, PFL_4751 family
	contig00270	2139	0 - 2139	n.d.	Rifp2Sym_iy00020	Conjugative coupling factor TraD, SXT/TOL

Back to list of tables

Table S4 Accessory genome found in 9°N symbionts (*Tevnia* and *Rifita* 2 symbiont genome assemblies but not in *Riftia* 1 symbionts). Total size = 75 723 bp;n.a. = Not Annotated; presence of sequence or part of sequence but no genes were called. **(Continued)**

Exclusive sequence	**Reference contig name**	**sequence length**	**Start – End (in reference contig)**	**Low Flanking conservation**	**Locus-tag in *Tevnia* symbionts**	**Locus-tag in *Riftia* 2 symbionts**	**Product**
Exclusive sequence 1	AFZB01000005	16398	22436 - 38834		TevJSym_ae00240	Rifp2Sym_ee00120	Phage protein, HK97 gp10 family
					TevJSym_ae00250	n.a.	Hypothetical protein
					TevJSym_ae00260	n.a.	Hypothetical protein
					TevJSym_ae00270	n.a.	Hypothetical protein
					TevJSym_ae00280	n.a.	Serine-rich adhesin-like protein for platelets precursor
					TevJSym_ae00290	Rifp2Sym_er00020	Hypothetical protein
					TevJSym_ae00300	Rifp2Sym_er00030	Peptidase family M23
					TevJSym_ae00310	Rifp2Sym_er00040	Hypothetical protein
					TevJSym_ae00320	Rifp2Sym_er00050	Hypothetical protein
					TevJSym_ae00330	Rifp2Sym_er00060	Mu-like prophage I protein
					TevJSym_ae00340	Rifp2Sym_er00080; Rifp2Sym_er00070	Antitoxin YefM
					TevJSym_ae00350	Rifp2Sym_er00090	Hypothetical protein
					TevJSym_ae00360	Rifp2Sym_er00110; Rifp2Sym_er00100	Hypothetical protein
					TevJSym_ae00370	Rifp2Sym_bw00170	Hypothetical protein
					TevJSym_ae00380	Rifp2Sym_bw00160	Hypothetical protein
					TevJSym_ae00390	Rifp2Sym_bw00150	Hypothetical protein
					TevJSym_ae00400	Rifp2Sym_bw00140	Hypothetical protein
Exclusive sequence 2	AFZB01000002	2860	126914 - 129774	✓	TevJSym_ab01270	Rifp2Sym_bw00010	Hypothetical protein
					TevJSym_ab01280	n.a.	Hypothetical protein
					TevJSym_ab01290	n.a.	Vault protein inter-alpha-trypsin domain-containing protein
					TevJSym_ab01300	n.a.	Transposase InsO and inactivated derivatives
Exclusive sequence 3	AFZB01000050	9786	0 - 9786		TevJSym_bx00010	n.a.	ATPase

Back to list of tables

Table S4 Accessory genome found in 9°N symbionts (*Tevnia* and *Rifita* 2 symbiont genome assemblies but not in *Riftia* 1 symbionts). Total size = 75 723 bp;n.a. = Not Annotated; presence of sequence or part of sequence but no genes were called. **(Continued)**

					TevJSym_bx00020	n.a.	Putative ATPase involved in DNA repair
					TevJSym_bx00030	n.a.	Hypothetical protein
Exclusive sequence 4	AFZB01000057	6657	0 - 6657		TevJSym_ce00010	n.a.	T/G mismatch-specific endonuclease
					TevJSym_ce00020	n.a.	DNA (cytosine-5)-methyltransferase 1
					TevJSym_ce00030	n.a.	KamA family protein
Exclusive sequence 5	AFZB01000036	6411	26656 - 33067	✓	TevJSym_bj00220	Rifp2Sym_ht00010	Hemolysin-type calcium-binding repeat-containing protein
					TevJSym_bj00230	Rifp2Sym_ag00010	Hemolysin D
					TevJSym_bj00240	Rifp2Sym_jc00010	Leukotoxin translocation ATP-biding protein/toxin secretion ABC transporter ATP-binding
Exclusive sequence 6	AFZB01000022	5730	0 - 5730		TevJSym_av00010	Rifp2Sym_he00010	Integron integrase
Exclusive sequence 7	AFZB01000006	1847	18500 - 20347	✓	TevJSym_af00170	Rifp2Sym_de00140	Membrane-bound lytic murein transglycosylase F
					TevJSym_af00180	Rifp2Sym_de00150	Phospholipase A1
					TevJSym_af00190	Rifp2Sym_de00160	tRNA(adenine34) deaminase
					TevJSym_af00200	n.a.	GMP synthase (glutamine-hydrolysing)
Exclusive sequence 8	AFZB01000083	1367	0 - 1367		n.a.	Rifp2Sym_km00020	Group II intron, maturase-specific domain
					n.a.	Rifp2Sym_km00020	IS3 family transposase, orfB
Exclusive sequence 9	AFZB01000002	5754	173150 - 178904		TevJSym_ab01730	Rifp2Sym_jq00010	ATP-dependent RNA helicase RhlE
					TevJSym_ab01740	n.a.	Iron complex outermembrane recepter protein
Exclusive sequence 10	AFZB01000045	3703	0 - 3703		TevJSym_bs00010	Rifp2Sym_da00010	Integrase/recombinase XerD
					TevJSym_bs00020	Rifp2Sym_da00020	Resolvase, N terminal domain
Exclusive sequence 11	AFZB01000035	3521	1360 - 4881		TevJSym_bi00030	Rifp2Sym_hr00010	Peptide methionine sulfoxide reductase msrA/msrB

Back to list of tables

Table S4 Accessory genome found in 9°N symbionts (*Tevnia* and *Rifita* 2 symbiont genome assemblies but not in *Riftia* 1 symbionts). Total size = 75 723 bp;n.a. = Not Annotated; presence of sequence or part of sequence but no genes were called. **(Continued)**

					TevJSym_bi00040	Rifp2Sym_hr00020	Hypothetical protein
					TevJSym_bi00050	Rifp2Sym_hr00030; Rifp2Sym_fg00010	Ribonucleoside-triphosphate reductase class III catalytic subunit
Exclusive sequence 12	AFZB01000068	3381	0 - 3381		TevJSym_cp00010	n.a.	Tertatricopeptide TPR_2 repeat protein
Exclusive sequence 13	AFZB01000005	2120	79186 - 81306		TevJSym_ae00700	Rifp2Sym_cv00010	MMPL family protein
Exclusive sequence 14	AFZB01000079	1665	0 - 1665		TevJSym_cz00010	Rifp2Sym_jr00010	Filamentation induced by cAMP protein Fic
Exclusive sequence 15	AFZB01000064	1541	0 - 1541		TevJSym_cl00010	Rifp2Sym_ic00010	Sodium/proton antiporter, NhaA family
Exclusive sequence 16	AFZB01000064	1487	2936 - 4423		TevJSym_cl00030	n.a.	Ribosomal subunit interface protein
					TevJSym_cl00040	Rifp2Sym_ko00010; Rifp2Sym_ko00020; Rifp2Sym_ko00030	LysR family transcriptional regulator, transcriptional activator of nhaA
Exclusive sequence 17	AFZB01000001	1211	135783 - 136993	✓	TevJSym_aa01480	Rifp2Sym_bk00140	Nitrate/nitrite transport system substrate-binding protein
					TevJSym_aa01490	Rifp2Sym_bk00150	ANTAR domain-containing protein
Exclusive sequence 18	AFZB01000016	284	12994 - 13278		TevJSym_ap00160	Rifp2Sym_is00010	LTXXQ motif family protein

Back to list of tables

Table S5 Genes of particular interest. Modified or complemented from Gardebrecht *et al.* (1). n.d.: Not Detected; no sequences were found in this genome assembly; n.a.: Not Annotated; orthologous sequence was found but no genes were detected in this genome assembly. **(Continued)**

1. **Gardebrecht A, Markert S, Sievert SM, Felbeck H, Thürmer A, Albrecht D, Wollherr A, Kabisch J, Le Bris N, Lehmann R**. 2011. Physiological homogeneity among the endosymbionts of *Riftia pachyptila* and *Tevnia jerichonana* revealed by proteogenomics. ISME J **6**:766–776.

				Locus-tag in Genbank accessions				
Categorie	**Sub-Categorie**	**Gene**	**Function**	***Ridgeia* 1 symbionts**	***Ridgeia* 2 Symbionts**	***Tevnia* symbiont**	***Riftia* 1 symbiont**	***Riftia* 2 symbiont**
1. Sulfur metabolism		aprA	Adenylylsulfate reductase, alpha subunit	Ga0074115_10324	Ga0076813_14621	TevJSym_ag00880	Rifp1Sym_ej00110	Rifp2Sym_bz00020 ; Rifp2Sym_bz00030
		aprB	Adenylylsulfate reductase, beta subunit	Ga0074115_10323	Ga0076813_14622	TevJSym_ag00890	Rifp1Sym_ej00100	Rifp2Sym_bz00040
		dsrA	Dissimilatory sulfite reductase, alpha subunit	Ga0074115_12121	Ga0076813_15802	TevJSym_aw00210	Rifp1Sym_am00340	Rifp2Sym_aa00430 ; Rifp2Sym_aa00440
		dsrB	Dissimilatory sulfite reductase, beta subunit	Ga0074115_12120	Ga0076813_15803	TevJSym_aw00190	Rifp1Sym_am00350	Rifp2Sym_aa00410
		sopT	ATP sulfurylase	n.d.	n.d.	TevJSym_ad01200	Rifp1Sym_cn00110	Rifp2Sym_az00240
2. Carbon metabolism	2.1. Calvin Benson Bassham cycle	cbbM	Ribulose 1,5 bisphosphate carboxylase/oxygenase	Ga0074115_14812	Ga0076813_14651	TevJSym_aj00630	Rifp1Sym_at00130	Rifp2Sym_bi00210
		gapA	Glyceraldehyde 3 phosphate dehydrogenase	Ga0074115_12924	Ga0076813_16395	TevJSym_an00400	Rifp1Sym_dp00070	Rifp2Sym_ar00120
		pgk	Phosphoglycerate kinase	Ga0074115_12925	Ga0076813_1392	TevJSym_an00390	Rifp1Sym_dp00060	Rifp2Sym_ar00110
		prkB	Phosphoribulokinase	Ga0074115_1484	Ga0076813_15275	TevJSym_aj00540	Rifp1Sym_at00030	Rifp2Sym_bf00040 ; Rifp2Sym_bf00050
		aclA	ATP citrate lyase, alpha subunit	Ga0074115_13717	Ga0076813_13921	TevJSym_az00170	Rifp1Sym_bt00020	Rifp2Sym_br00010 ; Rifp2Sym_aa000101
	2.2. TCA cycle	aclB	ATP citrate lyase, beta subunit	Ga0074115_13718	Ga0076813_14412	TevJSym_az00180	Rifp1Sym_bt00010	Rifp2Sym_br00020
		acnA	Aconitate hydratase/aconitase A	Ga0074115_13715	Ga0076813_13923	TevJSym_az00150	Rifp1Sym_bt00040	Rifp2Sym_aa00030
		icd	Isocitrate dehydrogenase [NADP]	Ga0074115_13716	Ga0076813_13922	TevJSym_az00160	Rifp1Sym_bt00030	Rifp2Sym_aa00020
		korA	2 oxoglutarate oxidoreductase, alpha subunit 2	Ga0074115_13621	Ga0076813_14941	TevJSym_az00290	Rifp1Sym_dm00080	Rifp2Sym_br00130
				Ga0074115_102185	Ga0076813_169112	TevJSym_bb00050	Rifp1Sym_aa00580	Rifp2Sym_an00050
				Ga0074115_11633	Ga0076813_13526	TevJSym_ar00630	Rifp1Sym_bp00050	Rifp2Sym_ab00620
		korB	2 oxoglutarate oxidoreductase, beta subunit 2	Ga0074115_13622	Ga0076813_14942	TevJSym_az00280	Rifp1Sym_dm00070	Rifp2Sym_br00120
				Ga0074115_102184	Ga0076813_169111	TevJSym_bb00040	Rifp1Sym_aa00590	Rifp2Sym_an00040

Back to list of tables

Table S5 Genes of particular interest. Modified or complemented from Gardebrecht *et al.* (1). n.d.: Not Detected; no sequences were found in this genome assembly; n.a.: Not Annotated; orthologous sequence was found but no genes were detected in this genome assembly. **(Continued)**

1. **Gardebrecht A, Markert S, Sievert SM, Felbeck H, Thürmer A, Albrecht D, Wollherr A, Kabisch J, Le Bris N, Lehmann R**. 2011. Physiological homogeneity among the endosymbionts of *Riftia pachyptila* and *Tevnia jerichonana* revealed by proteogenomics. ISME J **6**:766–776.

				Locus-tag in Genbank accessions				
Categorie	**Sub-Categorie**	**Gene**	**Function**	***Ridgeia* 1 symbionts**	***Ridgeia* 2 Symbionts**	***Tevnia* symbiont**	***Riftia* 1 symbiont**	***Riftia* 2 symbiont**
				Ga0074115_11632	Ga0076813_13525	TevJSym_ar00620	Rifp1Sym_bp00060	Rifp2Sym_ab00610
		korD	2 oxoglutarate oxidoreductase, delta subunit	Ga0074115_13620	Ga0076813_15152	TevJSym_az00300	Rifp1Sym_dm00090	Rifp2Sym_br00140
		korG	2 oxoglutarate oxidoreductase, gamma subunit	Ga0074115_13623	Ga0076813_14943	TevJSym_az00270	Rifp1Sym_dm00060	Rifp2Sym_br00110
		gltA2	citrate synthase I	Ga0074115_1551	Ga0076813_11841	TevJSym_bc00420	Rifp1Sym_ax00310	Rifp2Sym_aq00120
		maeB	malate dehydrogenase	Ga0074115_1552	Ga0076813_11901	TevJSym_bc00430	Rifp1Sym_ax00320 ; Rifp1Sym_ax00330	Rifp2Sym_aq001101 ; Rifp2Sym_aq00100
		sdhA	Succinate dehydrogenase, flavoprotein subunit2	Ga0074115_13138	Ga0076813_15895	TevJSym_ah01000	Rifp1Sym_dd00140	Rifp2Sym_ez00030
				Ga0074115_1555	Ga0076813_14861	TevJSym_cj00010	Rifp1Sym_ax00360	Rifp2Sym_aq000701 ; Rifp2Sym_aq00060
		sdhB	succinate dehydrogenase iron sulfur protein2	Ga0074115_13139	Ga0076813_15896	TevJSym_ah01010	Rifp1Sym_dd00150	Rifp2Sym_ez00020
				Ga0074115_1556	Ga0076813_14862	TevJSym_cj00020	Rifp1Sym_ax00370	Rifp2Sym_aq00050
		sdhC	succinate dehydrogenase membrane anchor 2	Ga0074115_13137	Ga0076813_15894	TevJSym_ah00990	Rifp1Sym_dd00130	Rifp2Sym_ez00040
				Ga0074115_1554	Ga0076813_15132	TevJSym_bc00450	Rifp1Sym_ax00350	Rifp2Sym_aq00080
		sdhD	succinate dehydrogenase cyt b subunit	Ga0074115_1553	Ga0076813_11902	TevJSym_bc00440	Rifp1Sym_ax00340	Rifp2Sym_aq00090
3. Nitrogen metabolism		narG	Respiratory nitrate reductase, alpha chain	Ga0074115_10332	Ga0076813_150210	TevJSym_ag00800	Rifp1Sym_ak00460	Rifp2Sym_bt00080
		narH	Respiratory nitrate reductase, beta chain	Ga0074115_10333	Ga0076813_15028	TevJSym_ag00790	Rifp1Sym_ak00450	Rifp2Sym_bt00090
		narJ	Respiratory nitrate reductase, delta chain	Ga0074115_10334	Ga0076813_15027	TevJSym_ag00780	Rifp1Sym_ak00440	Rifp2Sym_bt00100
		narI	Respiratory nitrate reductase, gamma chain	Ga0074115_10335	Ga0076813_15026	TevJSym_ag00770	Rifp1Sym_ak00430	Rifp2Sym_bt00110
		napD	Periplasmic nitrate reductase, subunit NapD	Ga0074115_1444	Ga0076813_15362	TevJSym_al00660	Rifp1Sym_ao00250	Rifp2Sym_bd00210
		napA	Periplasmic nitrate reductase, subunit NapA	Ga0074115_1443	Ga0076813_15361	TevJSym_al00650	Rifp1Sym_ao00240 ; Rifp1Sym_ao00230	Rifp2Sym_bd00220
		napG	Periplasmic nitrate reductase, subunit NapG	Ga0074115_1442	n.d.	TevJSym_al00640	Rifp1Sym_ao00220	Rifp2Sym_bd00230

Back to list of tables

Table S5 Genes of particular interest. Modified or complemented from Gardebrecht *et al.* (1). n.d.: Not Detected; no sequences were found in this genome assembly; n.a.: Not Annotated; orthologous sequence was found but no genes were detected in this genome assembly. **(Continued)**

1. **Gardebrecht A, Markert S, Sievert SM, Felbeck H, Thürmer A, Albrecht D, Wollherr A, Kabisch J, Le Bris N, Lehmann R**. 2011. Physiological homogeneity among the endosymbionts of *Riftia pachyptila* and *Tevnia jerichonana* revealed by proteogenomics. ISME J **6**:766–776.

				Locus-tag in Genbank accessions				
Categorie	**Sub-Categorie**	**Gene**	**Function**	***Ridgeia* 1 symbionts**	***Ridgeia* 2 Symbionts**	***Tevnia* symbiont**	***Riftia* 1 symbiont**	***Riftia* 2 symbiont**
		napH	Periplasmic nitrate reductase, subunit NapH	Ga0074115_1441	n.d.	TevJSym_al00630	Rifp1Sym_ao00210	Rifp2Sym_bd00240
		napB	Periplasmic nitrate reductase, subunit NapB	Ga0074115_1762	Ga0076813_1782	TevJSym_al00620	Rifp1Sym_ao00200	Rifp2Sym_bd00240
		napC	Periplasmic nitrate reductase, subunit NapC	Ga0074115_12753	Ga0076813_1783	TevJSym_al00610	Rifp1Sym_ao00190	Rifp2Sym_bd00260
		nirN	Cytochrome cd1/nitrite reductase, subunit NirN	Ga0074115_10211	Ga0076813_13753	TevJSym_bo00090	Rifp1Sym_cp00140	Rifp2Sym_db00140
		nirJ	Cytochrome cd1/nitrite reductase, subunit NirJ	Ga0074115_10212	Ga0076813_12089	TevJSym_bo00100	Rifp1Sym_cp00120 ; Rifp1Sym_cp00130	Rifp2Sym_db00110
		nirH	Cytochrome cd1/nitrite reductase, subunit NirH	Ga0074115_10213	Ga0076813_12088	TevJSym_bo00110	Rifp1Sym_cp00110	Rifp2Sym_db00100
		nirG	Cytochrome cd1/nitrite reductase, subunit NirG	Ga0074115_10215	Ga0076813_12086	TevJSym_bo00130	Rifp1Sym_cp00090	Rifp2Sym_db00080
		nirL	Cytochrome cd1/nitrite reductase, subunit NirL	Ga0074115_10216	Ga0076813_12085	TevJSym_bo00140	Rifp1Sym_cp00080	Rifp2Sym_db00070
		nirD	Cytochrome cd1/nitrite reductase, subunit NirD	Ga0074115_10217	Ga0076813_12084	TevJSym_bo00150	Rifp1Sym_cp00070	Rifp2Sym_db00060
		nirF	Cytochrome cd1/nitrite reductase, subunit NirF	Ga0074115_10218	Ga0076813_12083	TevJSym_bo00160	Rifp1Sym_cp00060	Rifp2Sym_db00050
		nirC	Cytochrome cd1/nitrite reductase, subunit NirC	Ga0074115_10219	Ga0076813_12082	TevJSym_bo00170	Rifp1Sym_cp00050	Rifp2Sym_db00040
		nirT	Cytochrome cd1/nitrite reductase, subunit NirT	Ga0074115_10221	Ga0076813_10286	TevJSym_bo00190	Rifp1Sym_cp00030	Rifp2Sym_db00010
		nirS	Cytochrome cd1/nitrite reductase, subunit NirS	Ga0074115_10222	Ga0076813_10285	TevJSym_bo00200	Rifp1Sym_cp00020	Rifp2Sym_gn00030
		norB	Nitric oxide reductase subunit B	Ga0074115_10224	Ga0076813_10283	TevJSym_cv00020	n. a.	Rifp2Sym_jh00010
		norC	Nitric oxide reductase subunit C	Ga0074115_10225	Ga0076813_10282	TevJSym_cv00010	n. a.	Rifp2Sym_jh00030
		nosZ	Nitrous oxide reductase	Ga0074115_11648	Ga0076813_16471	TevJSym_ac00130	Rifp1Sym_ag00460	Rifp2Sym_au00130 ; Rifp2Sym_au00140
4. Storage compounds		cphA	Cyanophycin synthase	Ga0074115_1582	Ga0076813_12881	TevJSym_ay00040	Rifp1Sym_ah00040	Rifp2Sym_at00040
5. Oxidative stress response		ahpC	Alkyl hydroperoxide reductase	Ga0074115_1298	Ga0076813_1752	TevJSym_ah00580	Rifp1Sym_ai00550	Rifp2Sym_af00050
		sodB	Superoxide dismutase [Fe]	Ga0074115_10415	Ga0076813_192	TevJSym_an00790	Rifp1Sym_eb00070	Rifp2Sym_dw00080

Back to list of tables

Table S5 Genes of particular interest. Modified or complemented from Gardebrecht *et al.* (1). n.d.: Not Detected; no sequences were found in this genome assembly; n.a.: Not Annotated; orthologous sequence was found but no genes were detected in this genome assembly. **(Continued)**

1. **Gardebrecht A, Markert S, Sievert SM, Felbeck H, Thürmer A, Albrecht D, Wollherr A, Kabisch J, Le Bris N, Lehmann R**. 2011. Physiological homogeneity among the endosymbionts of *Riftia pachyptila* and *Tevnia jerichonana* revealed by proteogenomics. ISME J **6**:766–776.

				Locus-tag in Genbank accessions				
Categorie	**Sub-Categorie**	**Gene**	**Function**	***Ridgeia* 1 symbionts**	***Ridgeia* 2 Symbionts**	***Tevnia* symbiont**	***Riftia* 1 symbiont**	***Riftia* 2 symbiont**
6. Secretion systems	6.1. Secretion across the inner membrane	secA	Preprotein translocase subunit SecA	Ga0074115_10789	Ga0076813_142814	TevJSym_aj00210	Rifp1Sym_bc00080	Rifp2Sym_ax00280
		secB	Preprotein translocase subunit SecB	Ga0074115_10285	Ga0076813_14246	TevJSym_at00230	Rifp1Sym_cl00120	Rifp2Sym_as00170
		secD	Preprotein translocase subunit SecD	Ga0074115_1295	Ga0076813_13813	TevJSym_an00610	Rifp1Sym_cf00160	Rifp2Sym_ar00340
		secE	Preprotein translocase subunit SecE	Ga0074115_12212	n.d.	TevJSym_bu00110	Rifp1Sym_di00120	Rifp2Sym_dg00010
		secF	Preprotein translocase subunit SecF	Ga0074115_1296	Ga0076813_13814	TevJSym_an00600	Rifp1Sym_cf00150	Rifp2Sym_ar00330
		secG	Preprotein translocase subunit SecG	n.d.	n.d.	TevJSym_ap00640	Rifp1Sym_bl00190	Rifp2Sym_eh00010
		secY	Preprotein translocase subunit SecY	Ga0074115_1496	Ga0076813_11654	TevJSym_aa00220	Rifp1Sym_bd00210	Rifp2Sym_ce00130
		tatA	Twin arginine translocation protein TatA	Ga0074115_10885	Ga0076813_150318	TevJSym_bh00280	Rifp1Sym_cq00210	Rifp2Sym_cb00170
		tatB	Twin arginine translocation protein TatB	Ga0074115_10884	Ga0076813_150319	TevJSym_bh00290	Rifp1Sym_cq00220	Rifp2Sym_cb00180
		tatC	Twin arginine translocation protein TatC	Ga0074115_10883	Ga0076813_150320	TevJSym_bh00300	Rifp1Sym_cq00230	Rifp2Sym_cb00190
	6.3. Secretion across the outer membrane	exeA	General secretion pathway protein A 2	Ga0074115_11430	Ga0076813_15682	TevJSym_ab01170	Rifp1Sym_aa00370	Rifp2Sym_an00280
				Ga0074115_1393	Ga0076813_11442	TevJSym_bg00130	Rifp1Sym_bk00170	Rifp2Sym_bh00210
				Ga0074115_1331	Ga0076813_11735	TevJSym_bb00270	Rifp1Sym_aa00370	Rifp2Sym_an00280
							Rifp1Sym_cw00170	
				Ga0074115_11330	Ga0076813_168710	TevJSym_ai00590	Rifp1Sym_cw00160	Rifp2Sym_cf00120
				Ga0074115_11430	Ga0076813_15682	TevJSym_ab01170	Rifp1Sym_ee00070	Rifp2Sym_dq00020
		xpsD	General secretion pathway protein D	Ga0074115_12041	Ga0076813_12179	TevJSym_ay00270	Rifp1Sym_ah00230	Rifp2Sym_do00140
		outE	General secretion pathway protein E	Ga0074115_1392	Ga0076813_11452	TevJSym_bg00140	Rifp1Sym_bk00160	Rifp2Sym_bh00200
		xcpR	General secretion pathway protein E	Ga0074115_12421	Ga0076813_15101	TevJSym_av00370	Rifp1Sym_ci00020	Rifp2Sym_be00150
		xcpS	General secretion pathway protein F	Ga0074115_12420	Ga0076813_14619	TevJSym_av00380	Rifp1Sym_ci00030	Rifp2Sym_be00160
				Ga0074115_1391	Ga0076813_11451	TevJSym_bg00150	Rifp1Sym_bk00150	Rifp2Sym_bh00190
	6.4. Type VI secretion system		Zn-binding Pro-Ala-Ala-Arg (PAAR) domain, involved in TypeVI secretion	Ga0074115_1462	Ga0076813_11513	n.d.	n.d.	n.d.

Back to list of tables

Table S5 Genes of particular interest. Modified or complemented from Gardebrecht *et al.* (1). n.d.: Not Detected; no sequences were found in this genome assembly; n.a.: Not Annotated; orthologous sequence was found but no genes were detected in this genome assembly. **(Continued)**

1. **Gardebrecht A, Markert S, Sievert SM, Felbeck H, Thürmer A, Albrecht D, Wollherr A, Kabisch J, Le Bris N, Lehmann R**. 2011. Physiological homogeneity among the endosymbionts of *Riftia pachyptila* and *Tevnia jerichonana* revealed by proteogenomics. ISME J **6**:766–776.

				Locus-tag in Genbank accessions				
Categorie	**Sub-Categorie**	**Gene**	**Function**	***Ridgeia* 1 symbionts**	***Ridgeia* 2 Symbionts**	***Tevnia* symbiont**	***Riftia* 1 symbiont**	***Riftia* 2 symbiont**
			putative component of the type VI protein secretion system	Ga0074115_1466	Ga0076813_14751 ; Ga0076813_16801	n.d.	n.d.	n.d.
		ImpA	type VI secretion-associated protein, ImpA family	Ga0074115_1469	Ga0076813_16804	n.d.	n.d.	n.d.
		ImpB	type VI secretion protein, VC_A0107 family	Ga0074115_14610	Ga0076813_16805	n.d.	n.d.	n.d.
		ImpC	type VI secretion protein, EvpB/VC_A0108 family	Ga0074115_14611	Ga0076813_13723	n.d.	n.d.	n.d.
		Hcp	Type VI protein secretion system component Hcp (secreted cytotoxin)	Ga0074115_14612	Ga0076813_13722	n.d.	n.d.	n.d.
		ImpF	type VI secretion system lysozyme-like protein	Ga0074115_14614	Ga0076813_14483	n.d.	n.d.	n.d.
		ImpG	type VI secretion protein, VC_A0110 family	Ga0074115_14615	Ga0076813_14482	n.d.	n.d.	n.d.
		ImpH	type VI secretion protein, VC_A0111 family	Ga0074115_14616	Ga0076813_14481	n.d.	n.d.	n.d.
7. Horizontal gene transfer	7.1. F Type conjugative plasmid	traG	IncF plasmid conjugative transfer protein TraG	n.d.	n.d.	n.d.	Rifp1Sym_au00110	n.a.
		traH	IncF plasmid conjugative transfer pilus assembly protein	n.d.	n.d.	n.d.	Rifp1Sym_au00120	n.a.
		traF	IncF plasmid conjugative transfer pilus assembly protein	n.d.	n.d.	n.d.	Rifp1Sym_au00130	n.a.
		traN	IncF plasmid conjugative transfer protein TraN	n.d.	n.d.	n.d.	Rifp1Sym_au00140	n.a.
		trbC	IncF plasmid conjugative transfer protein TrbC	n.d.	n.d.	n.d.	Rifp1Sym_au00150	n.a.
		traU	IncF plasmid conjugative transfer pilus assembly protein	n.d.	n.d.	n.d.	Rifp1Sym_au00160	n.a.
		pat17	Patatin like phospholipase A2	n.d.	n.d.	n.d.	Rifp1Sym_au00170	n.a.
		traC	IncF plasmid conjugative transfer pilus assembly protein	n.d.	n.d.	n.d.	Rifp1Sym_au00200	n.a.
		traV	Conjugative transfer protein TraV	n.d.	n.d.	n.d.	Rifp1Sym_au00210	n.a.
		traB	IncF plasmid conjugative transfer pilus assembly protein	n.d.	n.d.	n.d.	Rifp1Sym_au00230	n.a.
		traK	IncF plasmid conjugative transfer pilus assembly protein	n.d.	n.d.	n.d.	Rifp1Sym_au00240	n.a.
		traE	Plasmid like conjugative transfer protein	n.d.	n.d.	n.d.	Rifp1Sym_au00250	n.a.

Back to list of tables

Table S5 Genes of particular interest. Modified or complemented from Gardebrecht *et al.* (1). n.d.: Not Detected; no sequences were found in this genome assembly; n.a.: Not Annotated; orthologous sequence was found but no genes were detected in this genome assembly. **(Continued)**

1. **Gardebrecht A, Markert S, Sievert SM, Felbeck H, Thürmer A, Albrecht D, Wollherr A, Kabisch J, Le Bris N, Lehmann R**. 2011. Physiological homogeneity among the endosymbionts of *Riftia pachyptila* and *Tevnia jerichonana* revealed by proteogenomics. ISME J **6**:766–776.

				Locus-tag in Genbank accessions				
Categorie	**Sub-Categorie**	**Gene**	**Function**	***Ridgeia* 1 symbionts**	***Ridgeia* 2 Symbionts**	***Tevnia* symbiont**	***Riftia* 1 symbiont**	***Riftia* 2 symbiont**
			TraE					
7. Horizontal gene transfer	7.2. Transposases		IS5Y transposase	n.d.	n.d.	TevJSym_bc00030	Rifp1Sym_al00300	Rifp2Sym_aw00230
			Transposase IS66	Ga0074115_1614	Ga0076813_13865	TevJSym_el00010	n.d.	Rifp2Sym_mp00010
				Ga0074115_1615	Ga0076813_13866	n.a	n.a	n.a
			Transposase	Ga0074115_102126	Ga0076813_169033	TevJSym_aq00210	Rifp1Sym_ae00170	Rifp2Sym_bm00130
				Ga0074115_10991	Ga0076813_1302	TevJSym_ar00260	Rifp1Sym_bh00140	Rifp2Sym_ab00230
				Ga0074115_1654	Ga0076813_10561	TevJSym_ck00010	Rifp1Sym_db00090	Rifp2Sym_gs00010
				Ga0074115_12166	Ga0076813_16881	n.d.	n.d.	n.d.
				Ga0074115_11443	Ga0076813_12739	n.d.	n.d.	n.d.
				Ga0074115_1652	Ga0076813_10563	n.d.	n.d.	n.d.
				Ga0074115_102124	Ga0076813_169035	n.d.	n.d.	n.d.
				Ga0074115_1651	Ga0076813_10564	n.d.	n.d.	n.d.
				Ga0074115_102125	Ga0076813_169034	n.d.	n.d.	n.d.
			IS5 transposase	n.d	n.d	TevJSym_bc00040	Rifp1Sym_al00310	Rifp2Sym_aw00240
			Transposase insF for insertion sequence IS3	n.a	n.a	TevJSym_ab01630	Rifp1Sym_bb00110	Rifp2Sym_dh00120
			Transposase for IS1668	n.d	n.d	TevJSym_bf00240	Rifp1Sym_bg00180	Rifp2Sym_bl00030
			Transposase, IS4 family	n.a	n.a	TevJSym_cx00010	Rifp1Sym_db00080	Rifp2Sym_jp00010
8. Possibly involved in host infection	8.1. Adherence by type IV pili/fimbriae	fimT	Type IV fimbrial biogenesis protein FimT	Ga0074115_11736	Ga0076813_159923	TevJSym_am00340	Rifp1Sym_aq00380	Rifp2Sym_ak00170
				Ga0074115_12515	Ga0076813_14637	TevJSym_bs00040	Rifp1Sym_et00070	Rifp2Sym_da00030
		fimV	Type IV pilus assembly protein FimV	Ga0074115_10978	Ga0076813_15952	TevJSym_ar00130 ; TevJSym_ar00120	Rifp1Sym_dy00010 ; Rifp1Sym_dy00020	Rifp2Sym_ab00080
		pilB	Type IV fimbrial assembly protein PilB	Ga0074115_11359	Ga0076813_139111	TevJSym_ai00290	Rifp1Sym_az00200	Rifp2Sym_ba00260
				Ga0074115_11426	Ga0076813_14567	TevJSym_ab01140	Rifp1Sym_ee00030	Rifp2Sym_dq00060
		pilE	Type IV pilus biogenesis protein PilE	Ga0074115_11742	Ga0076813_159917	TevJSym_am00410	Rifp1Sym_an00070	Rifp2Sym_ak00100

Back to list of tables

Table S5 Genes of particular interest. Modified or complemented from Gardebrecht *et al.* (1). n.d.: Not Detected; no sequences were found in this genome assembly; n.a.: Not Annotated; orthologous sequence was found but no genes were detected in this genome assembly. **(Continued)**

1. **Gardebrecht A, Markert S, Sievert SM, Felbeck H, Thürmer A, Albrecht D, Wollherr A, Kabisch J, Le Bris N, Lehmann R**. 2011. Physiological homogeneity among the endosymbionts of *Riftia pachyptila* and *Tevnia jerichonana* revealed by proteogenomics. ISME J **6**:766–776.

				Locus-tag in Genbank accessions				
Categorie	**Sub-Categorie**	**Gene**	**Function**	***Ridgeia* 1 symbionts**	***Ridgeia* 2 Symbionts**	***Tevnia* symbiont**	***Riftia* 1 symbiont**	***Riftia* 2 symbiont**
		pilG	twitching motility protein PilG	Ga0074115_14215	Ga0076813_11134	TevJSym_bn00040	Rifp1Sym_cb00050	Rifp2Sym_cj00200
		pilH	twitching motility protein PilH	Ga0074115_14214	Ga0076813_11965	TevJSym_bn00030	Rifp1Sym_cb00040	Rifp2Sym_cj00210
		pilM	Type IV pilus biogenesis protein PilM	Ga0074115_102194	Ga0076813_12694	TevJSym_bb00140	Rifp1Sym_aa00490	Rifp2Sym_an00140
		pilN	Type IV pilus biogenesis protein PilN	Ga0074115_102195	Ga0076813_12695	TevJSym_bb00150	Rifp1Sym_aa00480	Rifp2Sym_an00150
				Ga0074115_12033	Ga0076813_113526	TevJSym_ay00340	Rifp1Sym_ah00310	Rifp2Sym_do00050
		pilQ	Type IV pilus biogenesis protein PilQ	Ga0074115_102198	Ga0076813_12698	TevJSym_bb00180	Rifp1Sym_aa00450	Rifp2Sym_an00180
		pilR	Type IV fimbriae expression regulatory protein PilR	Ga0074115_11363	Ga0076813_10808	TevJSym_ai00250	Rifp1Sym_az00160	Rifp2Sym_ba00220
		pilS	Sensor protein PilS	Ga0074115_11364	Ga0076813_10807	TevJSym_ai00240	Rifp1Sym_az00150	Rifp2Sym_ba00210
		pilT	Twitching mobility protein PilT	Ga0074115_11726	Ga0076813_12482	TevJSym_am00210	Rifp1Sym_aq00280	Rifp2Sym_ak00290
		pilV	Type IV pilus modification protein PilV	Ga0074115_11737	Ga0076813_159922	TevJSym_am00350	Rifp1Sym_aq00390	Rifp2Sym_ak00160
		pilY	Type IV fimbrial biogenesis protein PilY	Ga0074115_11740	Ga0076813_159924	TevJSym_am00390	Rifp1Sym_an00030 ; Rifp1Sym_an00040 ; Rifp1Sym_an00050	Rifp2Sym_ak00130 ; Rifp2Sym_ak00120
		pilZ	Type IV pilus assembly PilZ	Ga0074115_11818	Ga0076813_12504	TevJSym_bc00220	Rifp1Sym_ax00100	Rifp2Sym_aq00310
		TadD	Flp pilus assembly protein TadD, contains TPR repeats	Ga0074115_12218	Ga0076813_16271	TevJSym_bt00010	Rifp1Sym_fu00030	Rifp2Sym_ct00190
		pilO	Tfp pilus assembly protein PilO	Ga0074115_102196	Ga0076813_12696	TevJSym_bb00160	Rifp1Sym_aa00470	Rifp2Sym_an00160
		pilX	Tfp pilus assembly protein PilX	Ga0074115_11739	Ga0076813_159920	TevJSym_am00380	Rifp1Sym_an00020	Rifp2Sym_ak00140
		pilP	Tfp pilus assembly protein PilP	Ga0074115_102197	Ga0076813_12697	TevJSym_bb00170	Rifp1Sym_aa00460	Rifp2Sym_an00170
	8.2. Adherence by flagellum	mshL	Pilus (MSHA Type) biogenesis protein MshL	Ga0074115_1395	Ga0076813_13261	TevJSym_bg00100	Rifp1Sym_bk00190	Rifp2Sym_bh00230
		flaG	Flagellin protein FlaG	Ga0074115_1433	Ga0076813_12133	TevJSym_aa01600	Rifp1Sym_fp00070	Rifp2Sym_bk00240
		fleN	Flagellar synthesis regulator FleN	Ga0074115_12148	Ga0076813_12351	TevJSym_aw00490	Rifp1Sym_am00080	Rifp2Sym_aa00690
		flgA	Flagellar basal body P ring formation protein FlgA	Ga0074115_13818	Ga0076813_14258	TevJSym_aa01200	Rifp1Sym_bw00270	Rifp2Sym_aj00140

Back to list of tables

Table S5 Genes of particular interest. Modified or complemented from Gardebrecht *et al.* (1). n.d.: Not Detected; no sequences were found in this genome assembly; n.a.: Not Annotated; orthologous sequence was found but no genes were detected in this genome assembly. **(Continued)**

1. **Gardebrecht A, Markert S, Sievert SM, Felbeck H, Thürmer A, Albrecht D, Wollherr A, Kabisch J, Le Bris N, Lehmann R**. 2011. Physiological homogeneity among the endosymbionts of *Riftia pachyptila* and *Tevnia jerichonana* revealed by proteogenomics. ISME J **6**:766–776.

				Locus-tag in Genbank accessions				
Categorie	**Sub-Categorie**	**Gene**	**Function**	***Ridgeia* 1 symbionts**	***Ridgeia* 2 Symbionts**	***Tevnia* symbiont**	***Riftia* 1 symbiont**	***Riftia* 2 symbiont**
		fldG	Flagellar hook protein FlgG	Ga0074115_13812	Ga0076813_14252	TevJSym_aa01270	Rifp1Sym_bw00200	Rifp2Sym_aj00080
		flgE	Flagellar hook protein FlgE	Ga0074115_13814	Ga0076813_14254	TevJSym_aa01250	Rifp1Sym_bw00220	Rifp2Sym_aj00100
		flgF	Flagellar basal body rod protein FlgF	Ga0074115_13813	Ga0076813_14253	TevJSym_aa01260	Rifp1Sym_bw00210	Rifp2Sym_aj00090
		flgH	Flagellar L ring protein FlgH	Ga0074115_13811	Ga0076813_14251	TevJSym_aa01280	Rifp1Sym_bw00190	Rifp2Sym_aj00070
		flhA	Flagellar biosynthesis protein FlhA	Ga0074115_12150	Ga0076813_16232	TevJSym_aw00510	Rifp1Sym_am00060	Rifp2Sym_aa00720
		fliA	RNA polymerase sigma factor for flagellar operon	Ga0074115_12147	Ga0076813_14524	TevJSym_aw00480	Rifp1Sym_am00090	Rifp2Sym_aa00680
		fliD	Flagellar hook associated protein 2	Ga0074115_1432	Ga0076813_12134	TevJSym_aa01610	Rifp1Sym_fp00060	Rifp2Sym_bk00250
		fliG	Flagellar motor switch protein FliG	Ga0074115_1197	Ga0076813_12727	TevJSym_ac00750	Rifp1Sym_cx00180	Rifp2Sym_cm00100
		fliJ	Flagellar protein FliJ	Ga0074115_1194	Ga0076813_12724	TevJSym_ac00720	Rifp1Sym_ck00160	Rifp2Sym_cm00060
		fliM	Flagellar motor switch protein FliM	Ga0074115_12157	Ga0076813_12633	TevJSym_aw00580	n.d.	Rifp2Sym_aa00790
		flgB	Flagellar basal-body rod protein FlgB	Ga0074115_13817	Ga0076813_14257	TevJSym_aa01210	Rifp1Sym_bw00260	Rifp2Sym_aj00130
		FliL	Flagellar basal body-associated protein FliL	Ga0074115_1347	Ga0076813_155326	TevJSym_ah00540	Rifp1Sym_ai00510	Rifp2Sym_af00090
				Ga0074115_12158	Ga0076813_12632	TevJSym_aw00590	Rifp1Sym_ct00020	Rifp2Sym_aa00800
		FliO	flagellar biosynthetic protein FliO	Ga0074115_12155	Ga0076813_12635	TevJSym_aw00560	Rifp1Sym_am00010	Rifp2Sym_aa00770
		GldG	ABC-type uncharacterized transport system involved in gliding motility, auxiliary component	Ga0074115_1173	Ga0076813_12682	TevJSym_ai00040	Rifp1Sym_aq00050	Rifp2Sym_ba00010
		flgC	flagellar basal-body rod protein FlgC	Ga0074115_13816	Ga0076813_14256	TevJSym_aa01220	Rifp1Sym_bw00250	Rifp2Sym_aj00120
		FlgG	flagellar basal-body rod protein FlgG, Gram-negative bacteria	Ga0074115_13812	Ga0076813_14252	TevJSym_aa01270	Rifp1Sym_bw00200	Rifp2Sym_aj00080
		FlgI	Flagellar basal body P-ring protein FlgI	Ga0074115_13810	Ga0076813_13251	TevJSym_aa01290	Rifp1Sym_bw00180	Rifp2Sym_aj00060
		FlgL	flagellar hook-associated protein 3	Ga0074115_1387	Ga0076813_1246	TevJSym_aa01320	Rifp1Sym_bw00140	Rifp2Sym_aj00030
		FlhF	flagellar biosynthetic protein FlhF	Ga0074115_12149	Ga0076813_12352	TevJSym_aw00500	Rifp1Sym_am00070	Rifp2Sym_aa00710
		motA	Flagellar motor rotation protein motA	Ga0074115_12143	Ga0076813_1553	TevJSym_aw00430	Rifp1Sym_am00130	Rifp2Sym_aa00640
				Ga0074115_12447	Ga0076813_164510	TevJSym_av00100	Rifp1Sym_av00240	Rifp2Sym_ci00070
		motB	Flagellar motor rotation protein motB	Ga0074115_10496	Ga0076813_13374	TevJSym_bi00240	Rifp1Sym_cj00080	Rifp2Sym_ai00430
				Ga0074115_12448	Ga0076813_15591	TevJSym_av00090	Rifp1Sym_av00250	Rifp2Sym_ci00060
				Ga0074115_12142	Ga0076813_1552	TevJSym_aw00420	Rifp1Sym_am00140	Rifp2Sym_aa00630

Back to list of tables

Table S5 Genes of particular interest. Modified or complemented from Gardebrecht *et al.* (1). n.d.: Not Detected; no sequences were found in this genome assembly; n.a.: Not Annotated; orthologous sequence was found but no genes were detected in this genome assembly. **(Continued)**

1. **Gardebrecht A, Markert S, Sievert SM, Felbeck H, Thürmer A, Albrecht D, Wollherr A, Kabisch J, Le Bris N, Lehmann R**. 2011. Physiological homogeneity among the endosymbionts of *Riftia pachyptila* and *Tevnia jerichonana* revealed by proteogenomics. ISME J **6**:766–776.

				Locus-tag in Genbank accessions				
Categorie	**Sub-Categorie**	**Gene**	**Function**	***Ridgeia* 1 symbionts**	***Ridgeia* 2 Symbionts**	***Tevnia* symbiont**	***Riftia* 1 symbiont**	***Riftia* 2 symbiont**
	8.3. Cell attachment	FliK	Flagellar hook length control protein	Ga0074115_11949	Ga0076813_12961	TevJSym_ac01170	Rifp1Sym_da00090	Rifp2Sym_ay00290
		hlyIII	Hemolysin III like protein	Ga0074115_10985	Ga0076813_11011	TevJSym_ar00200	Rifp1Sym_bh00190	Rifp2Sym_ab00180
		TlyC	Putative Hemolysin C	Ga0074115_1751	Ga0076813_14012	TevJSym_bc00400	Rifp1Sym_ax00290	Rifp2Sym_aq00130
		hlyD	Type I secretion protein, HlyD family	Ga0074115_101107	Ga0076813_16091	TevJSym_ao00270	Rifp1Sym_ac00300	Rifp2Sym_dc00070
				Ga0074115_10993	Ga0076813_1304	TevJSym_ar00280	Rifp1Sym_bh00120	Rifp2Sym_ab00250
				Ga0074115_1373	Ga0076813_11203	TevJSym_az00030	Rifp1Sym_bt00160	Rifp2Sym_aa00180
		kamA	Lysine 2,3 aminomutase	Ga0074115_12934	Ga0074115_11543	TevJSym_an00300	Rifp1Sym_ey00040	Rifp2Sym_ar00010
		cvpA	Colicin V production protein	Ga0074115_12444	Ga0076813_16457	TevJSym_av00130	Rifp1Sym_av00210	Rifp2Sym_ci00100
			Outer membrane adhesin like protein	Ga0074115_13915	Ga0076813_12205	TevJSym_ae00280	Rifp1Sym_gp00020	Rifp2Sym_fz00070
				Ga0074115_11649	Ga0076813_16472	TevJSym_ac00140	Rifp1Sym_ag00450	Rifp2Sym_au00150
			Hyalin	Ga0074115_11167	Ga0076813_11761	TevJSym_bf00250	Rifp1Sym_bg00190	Rifp2Sym_bl00010 ; Rifp2Sym_bl00020
			Fibronectin Type III domain protein	Ga0074115_10729	Ga0076813_12454	TevJSym_au00400	Rifp1Sym_bo00250	Rifp2Sym_al001101 ; Rifp2Sym_al00120
			Cell wall associated biofilm protein	Ga0074115_1301	Ga0076813_12431	TevJSym_ay00010	Rifp1Sym_ah00020	Rifp2Sym_at000101 ; Rifp2Sym_at00020