Bacteriocyte-associated gammaproteobacterial symbionts of the *Adelges piceae/nordmannianae* complex (Hemiptera: Adelgidae)

Elena R. Toenshoff et al.

Supplementary Information

Table S1: Primer and probes used in this study.

Short name ^a	Sequence (5´-3´)	Specificity	Target site	Annealing temperature/ Formamide concentration	Reference
16S and 23S rRNA	primer		•		
616V	AGA GTT TGA TYM TGG CTC	16S rRNA gene, most Bacteria	8-25	52	(Juretschko <i>et al.</i> , 1998)
1492R	GGY TAC CTT GTT ACG ACT T	16S rRNA gene, most Bacteria and Archaea	1492-1510	52	(Loy et al., 2005)
TTL1-16S-70F	CAT CGG AAA GGA GTT TAC TTC	16S rRNA gene, 'Candidatus Steffania adelgidicola'	70-90	58	This study
TTL1-16S-1267R	GAG GTC CGC TGA CCC TCA	16S rRNA gene, 'Candidatus Steffania adelgidicola'	1267-1284	58	This study
TTL4-16S-107F	GGA CGG GTG AGT AAT ATT	16S rRNA gene, 'Candidatus Ecksteinia adelgidicola'	107-124	58	This study
TTL4-16S-1449R	GTA AGT GCC CTC CAA TAC	16S rRNA gene, 'Candidatus Ecksteinia adelgidicola'	1450-1466	58	This study
255f	AGT AGY GGC GAG CGA A	23S rRNA gene, <i>Bacteria</i>	241-255	50	(Lane, 1991)
1930R	CGA CAA GGA AYT TCG CTA C	23S rRNA gene, <i>Bacteria</i>	1930-1948	50	(Hunt <i>et al.</i> , 2006)
TTL1-23S-346F	GGT GTG TTA GTT GTG AGC	23S rRNA gene, 'Candidatus Steffania adelgidicola'	347-364	64	This study
TTL1-23S-1169R	CCA GCA ACA CTC TCA TGC	23S rRNA gene, 'Candidatus Steffania adelgidicola'	1169-1186	64	This study
TTL1-23S-347R	GCT CAC AAC TAA CAC ACC	23S rRNA gene, 'Candidatus Steffania adelgidicola'	347-364	n.d.	This study
TTL4-23S-260F	GGG ACA GCC CAG AGC TAG	23S rRNA gene, 'Candidatus Ecksteinia adelgidicola'	260-277	64	This study
TTL4-23S-1169R	GCA ATG CAT ATT TCA CAT	23S rRNA gene, 'Candidatus Ecksteinia adelgidicola'	1169-1183	64	This study
Host gene targetin	ig primer				
911	TTT CTA CAA ATC ATA AAG ATA TTG G	Mitochondrial cytochrome c oxidase subunit 1 (<i>col</i>), various eukaryotes		50	(Guryev <i>et al.</i> , 2001)
912	TAA ACT TCA GGG TGA CCA AAA AAT CA	Mitochondrial cytochrome c oxidase subunit 1 (<i>col</i>), various eukaryotes		50	(Guryev <i>et al.</i> , 2001)
AdelEF1F1	GTA CAT CCC AAG CCG ATT GT	Partial nuclear elongation factor 1-alpha (<i>ef1alpha</i>), various Hemiptera		61	(Havill <i>et al.</i> , 2007)
AdelEF1R2	CTC CAG CTA CAA AAC CAC GA	Partial nuclear elongation factor 1-alpha (<i>ef1alpha</i>), various Hemiptera		61	(Havill <i>et al.</i> , 2007)
Oligonucleotide p	robes				
EUB338-I	GCT GCC TCC CGT AGG AGT	16S rRNA, most <i>Bacteria</i>	338-355 ^b	10-50	(Amann <i>et al.</i> , 1990)
EUB338-II	GCA GCC ACC CGT AGG TGT	16S rRNA, Bacteria not covered by probe EUB338-I, e.g. many <i>Planctomycetes</i>	338-355 ^b	10-60	(Daims <i>et al.</i> , 1999)
EUB338-III	GCT GCC ACC CGT AGG TGT	16S rRNA, Bacteria not covered by probe EUB338-I, e.g. many <i>Verrucomicrobia</i>	338-355 ^b	10-60	(Daims <i>et al.</i> , 1999)
TTL1-1027	GTC ACA GAG TCC CCT AAG	16S rRNA, 'Candidatus Steffania adelgidicola'	1028-1045	30-35	This study
TTL4-608	CAC GTT AAG CGC AGG GAT	16S rRNA, 'Candidatus Ecksteinia adelgidicola'	609-626	30-35	This study
TTL4-1264	GGT TTG CTT ACT CTT GCG	16S rRNA, 'Candidatus Ecksteinia adelgidicola'	1265-1282	30-35	This study
TTL4-1445	GCC CTC CAA TAC ATG GTT	16S rRNA, 'Candidatus Ecksteinia adelgidicola'	1446-1460	30-35	This study
NONEUB	ACT CCT ACG GGA GGC AGC	Control probe complementary to EUB338-I	338-355	n.d.	(Wallner et al., 1993)

TTL1-1215	GCA CAC CTT ACA GTG CAC	23S rRNA, 'Candidatus Steffania adelgidicola'	1216-1233	10-25	This study
TTL4-877	AAG CTG GGA TAG CCC CTT	23S rRNA, 'Candidatus Ecksteinia adelgidicola'	878-895	10-25	This study
TTL4-967	GTT GTT TCC CTT TAC ACG	23S rRNA, 'Candidatus Ecksteinia adelgidicola'	968-985	10-25	This study
TTL4-1215	GCA ATC CTC ACG AGT CAC	23S rRNA, 'Candidatus Ecksteinia adelgidicola'	1216-1233	10-25	This study

^a Short probe name used in this study or in the references ^b EUB338-I, EUB338-II, and EUB338-III were applied simultaneously to target most *Bacteria*

Table S2: *ef1alpha* single nucleotide polymorphisms differentiating *A. nordmannianae* and *A. piceae* (grey background) and additional variable sites among the *A. nordmannianae/A. piceae* species complex. All available *ef1alpha* gene sequences of this complex are shown.

Organiam	Sampling	Genbank Acc.				Α	ignmen	t positio	n ^a			
Organism	location	Number	203	204	295	331	398	446	509	548	649	650
Adelges nordmannianae	Slovak Republic	EF073235	Т	Т	Т	Т	Т	Т	A	Т	Т	С
Adelges piceae	Owlshead; Maine; USA	EF073239	А	-	Т	Т	С	С	А	С	Т	С
Adelges piceae	Delemont; Switzerland	EF073240	Т	т	Т	Т	С	С	A	С	Т	С
Adelges nordmannianae/piceae	Gosau; Austria	HQ668164	Т	Т	Т	Т	Т	Т	A	Т	Т	С
Adelges nordmannianae/piceae	Grafrath; Germany	HQ668167	Т	-	Т	Т	С	С	А	С	Т	С
Adelges nordmannianae/piceae Klausen-Leopoldsdorf	Klausen- Leopoldsdorf; Austria	HQ668165	т	Т	Т	С	т	т	А	т	т	т
Adelges nordmannianae/piceae Klausen-Leopoldsdorf	Klausen- Leopoldsdorf; Austria	HQ668166	т	-	С	Т	С	С	G	548 649 63 T T T C T C C T C T T C T T C T T C C T C C T C C A C	С	

^anucleotide position based on the *ef1alpha* gene sequence of *Adelges nordmannianae* (EF073235)

Supplementary Table S3: List of genes encoded on the genome fragment of 'Candidatus' Steffania adelgidicola' and the classification of their proteins into COG categories. Comparison of proteins of 'Candidatus' Steffania adelgidicola' to proteins of other insect symbionts and free-living bacteria.

Candidatus S	Steffania	adelgio	licola'			Ec	Sg	Ss	Hd	Вр	Wg	Bc	Ba	Cr				
locus tags	start	stop	type	gene discription	gene name										remarks	COGs	code	discription
STA_A00010	907	1	orf	Ribosomal RNA small subunit methyltransferase H	rsmH	6e-109	6e-122	4e-105	4e-96	4e-90	5e-76	6e-101	1e-86	2e-04		COG0275	M	Cell envelope biogenesis, outer membrane
STA_A00020	2981	2310	orf	Uracil-DNA glycosylase	ung	2e-87	2e-102	7e-85	6e-79	5e-81	1e-50	3e-90	2e-73	2.6		COG0692	L	DNA replication, recombination and repair
STA_A00030	4730	3771	orf	Acetyl-coenzyme A carboxylase carboxyl transferase subunit	accA	5e-164	1e-176	4e-166	2e-149	3e-146	3e-144	4e-154	2.1	0.19		COG0825	_	Lipid metabolism
STA_A00040	8224	4742	orf	DNA polymerase III subunit alpha	dnaE	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	5e-122		COG0587	L	DNA replication, recombination and repair
STA_A00050	8300	8623	pseudogene	Ribonuclease HII*	rnhB*	2e-85	7e-113	7e-89	1e-80	6e-55	3.9	0.15	0.65	2.5	1	COG0164	L	DNA replication, recombination and repair
						0.023	0.003	0.14	1.3	1.3	0.77	2.8	1.1	0.044	2			
STA_A00060	10062	8878	orf	Lipid-A-disaccharide synthase	lpxB	2e-154	8e-169	2e-157	2e-132	5e-135	3e-155	0.006	0.82	0.38		COG0763	M	Cell envelope biogenesis, outer membrane
STA_A00070	10820	10032	orf	Acyl-[acyl-carrier-protein]UDP-N-acetylglucosamine O-	lpxA	1e-117	2e-133	2e-112	8e-86	9e-111	6e-105	2.0	0.020	0.76		COG1043	M	Cell envelope biogenesis, outer membrane
STA_A00080	11280	10825	orf	(3R)-hydroxymyristoyl-[acyl-carrier-protein] dehydratase	fabZ	7e-70	2e-79	1e-69	2e-66	5e-74	3e-08	6e-67	5e-10	7.0		COG0764		Lipid metabolism
STA_A00090	12452	11430	orf	UDP-3-O-[3-hydroxymyristoyl] glucosamine N-acyltransferase	lpxD	1e-153	2e-168	4e-152	3e-131	6e-138	7e-116	1e-08	5e-11	0.022		COG1044	M	Cell envelope biogenesis, outer membrane
STA_A00100	12953	12456	orf	Chaperone protein skp	skp	2e-39	1e-67	2e-57	2e-28	1e-45	3e-32	1e-59	1.7	6.5		COG2825	M	Cell envelope biogenesis, outer membrane
STA_A00110	15491	13065	orf	Outer membrane protein assembly factor yaeT	yaeT	0.0	0.0	0.0	0.0	0.0	1e-143	0.0	3e-103	0.010		COG4775,	M	Cell envelope biogenesis, outer membrane
					_											COG0729		
STA_A00120	16878	15529	orf	Regulator of sigma E protease	rseP	1e-178	0.0	4e-04	2e-125	8e-126	4e-89	1e-139	0.025	0.15		COG0750	M	Cell envelope biogenesis, outer membrane
STA_A00130	17755	16898	orf	Phosphatidate cytidylyltransferase	cdsA2	5e-94	3e-108	1e-86	1e-88	2e-71	6e-38	1e-81	4.1	0.13		COG0575		Lipid metabolism
STA_A00140	18506	17763	orf	Undecaprenyl pyrophosphate synthase	uppS	7e-100	3e-117	1e-101	5e-87	1e-95	4e-95	1e-101	3e-80	2.1		COG0020		Lipid metabolism
STA_A00150	19902	18706	orf	1-deoxy-D-xylulose 5-phosphate reductoisomerase	dxr	5e-147	2e-174	1e-149	4e-145	3e-132	2e-131	2e-153	3 3e-122	0.62		COG0743		Lipid metabolism
STA_A00160	22196	20790	orf	6-phosphogluconate dehydrogenase, decarboxylating	gnd	0.0	0.0	0.0	0.0	0.0	0.037	0.0	0.0	2e-39		COG0362	G	Carbohydrate transport and metabolism
STA_A00170	24086	22920	orf	P-protein	pheA	4e-142	2e-164	9e-134	0.075	2e-69	0.40	0.007	9e-97	0.83		COG0077,	E	Amino acid transport and metabolism
																COG1605		
STA_A00180	24957	24226	orf	UPF0169 lipoprotein yfiO	yfiO	8e-91	1e-101	2e-86	2e-78	1e-51	3e-25	2e-54	2e-34	2.3		COG4105	R	General function prediction only
STA_A00190	25088	26068	orf	Pseudouridine synthase	rluD	4e-123	6e-139	3e-123	1e-108	3e-91	2e-65	1e-98	5e-79	3e-15		COG0564	J	Translation, ribosomal structure and biogenesis
STA_A00200	26065	26796	pseudogene	UPF0124 protein yfiH*	yfiH*	5e-145	9e-21	0.11	1.8	1.4	0.63	4.1	0.34	0.62	3	COG1496	S	Function unknown
						3e-06	0.44	2.6	0.45	0.31	0.14	9.8	1.0	5.8	4			
STA_A00210	26937	29232	pseudogene	Chaperone protein clpB*	clpB*	0.0	0.0	0.0	0.0	0.0	1e-04	0.0	1e-05	9e-08	5	COG0542	0	Posttranslational modification, protein turnover, chaperones
						1e-160	3e-178	2e-73	7e-134	6e-60	0.71	9e-145	4.0	0.65	6			
STA_Ar0010	30685	32195	rRNA	16S rRNA gene														
STA_At0010	32250	32322	tRNA	GlutRNA	()	0.0	0.0	0 00	0.0	0.0	4 470		0.00	1.0		0000111	_	
STA_A00220	32564	33952	orf	Fumarate hydratase class II	fumC	0.0	0.0	8e-86	0.0	0.0	1e-1/8	0.0	2e-09	1.3		COG0114	C	Energy production and conversion
STA_A00230	34852	33995	ort	Bifunctional protein folD	toID	9e-130	6e-142	1e-74	2e-115	2e-113	2e-109	1e-124	4e-105	0.13		COG0190	н	Coenzyme metabolism
STA_At0020	35050	35123	tRNA	Arg tRNA		0- 00	5- 00	F - 00	0- 70	5.0	4.4	0.50	0.00	0.44		0004000	D	One and for the and letter and
STA_A00240	36028	35444	orr	UPFU2U/ protein	yibR sfa7	36-80	56-98	56-83	20-76	5.6	4.4	0.56	0.93	0.11		COG1896	R	General function prediction only
STA_A00250	20270	20575	orf	Putative lipid A core:curface polymor ligace	maal	1.2	62 120	0.70	0.09	1.0	0.75	1.3	2.0	1.0		0003207	м	Call anvelope biogenesis, outer membrane
STA_A00200	40401	20520	orf	lipopolycoccharido biosynthesis protoin	wall/2	0.01	10.150	70.120	2.1	2.2	0.52	2.3	4.7	2.0		0003307	IVI	Cell envelope biogenesis, outer membrane
STA_A00270	40491	42420	orf	Putative lipopolysaccharide diversities protein	rfaR	30-13	50-164	10-110	60-07	0.46	0.55	0.17	2.2	2.0		0000438	м	Cell envelope biogenesis, outer membrane
STA_A00200	41014	42423	orf	Dibudroxy-acid debudratase	ilvD	0.0		0.0	0.028	0.40	0.00	0.33	0.0	40-83		COG0129	FG	Amino acid transport and metabolism [E]
01A_A00230	44323	43073	011	Dinydroxy-acid denydratase	1100	0.0	0.0	0.0	0.020	0.0	0.01	0.55	0.0	40-03		0000123	L, U	Carbohydrate transport and metabolism [C]
STA 400200	45010	44097	orf	Pronchod chain amino acid aminotransforaça	ilvE	10 159	50 160	0.12	0.027	10 125	12	20.09	27	40.04		0000115	с U	Amino acid transport and metabolism [C]
STA_A00300	45919	44907	011	Brancheu-chain-amino-aciu aminotransierase	IIVE	46-120	5e-169	0.12	0.037	Te-135	1.5	26-00	2.1	46-04		COGUIIS	Е, П	Coonsume transport and metabolism [E],
OTA 400040	40405	45005	4	A satala state suggitures is sugger a A satall suburit	36 . A. A	0- 00	0- 00	0- 45	4.4	4- 00	0.44	7.0	4.0			0000440	-	Coercyrrie transport and metabolism [ri]
STA_A00310	46195	45965	off	Acetolactate synthase isozyme 2 small subunit	IIVM itro	6e-20	26-30	36-15	1.1	4e-20	0.44	7.3	1.8	nss		COG0440	E	Amino acid transport and metabolism
STA_A00320	47841	46195	оп	Acetolactate synthase isozyme 2 large subunit	IIVG	0.0	0.0	5e-22	56-04	0.0	0.15	0.051	46-122	96-96		COG0028	Е, Н	Amino acid transport and metabolism [E],
								_										Coenzyme transport and metabolism [H]
STA A00330	48083	47982	orf	IIVGMEDA operon leader peptide	IVL	nss	0.39	nss	3.4	3.8	nss	nss	8.4	4.7		no COG	12.11	
STA_A00340	50001	49042	orr	Birunctional protein birA	DIFA	1e-115	9e-137	7e-118	4e-92	26-68	16-49	1e-71	0.94	0.33		COG0340 [H],	к, н	Coenzyme metabolism [H],
																COG1654 [K]		Transcription [K]
STA_A00350	51038	49998	orf	UDP-N-acetylenolpyruvoylglucosamine reductase	murB	2e-116	5e-139	2e-107	4e-101	1e-100	2e-94	8e-117	2e-87	1.6		COG0812	M	Cell envelope biogenesis, outer membrane
STA_A00360	51626	52105	orr	tRNA-specific adenosine deaminase	tadA	16-49	2e-62	56-53	1e-40	5e-42	4e-40	4e-43	6e-39	3.3		COG0590	F, J	Nucleotide transport and metabolism [F],
					-	4 450	5 400	4 450	0 405	0.44	0 100	0 100	0 100	0.00		0000540	-	I ranslation, ribosomal structure and biogenesis [J]
STA_A00370	53744	52800	orf	Aspartate carbamoyltransterase	pyrB	46-150	56-162	1e-150	36-105	2e-14	2e-120	2e-139	2e-128	0.36		COG0540	F	Nucleotide transport and metabolism
STA_A00380	56360	54996		Lysine-sensitive aspartokinase 3	iysC	0.0	0.0	0.0	16-43	26-41	66-30	0.015	46-41	36-15		0000527	E	Amino acid transport and metabolism
STA_A00390	56960	58609	orf	Glucose-6-phosphate isomerase	pgi	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	3.5	7	COG0166	G	Carbonydrate transport and metabolism
31A_A00400	01/38	03355	pseudogene		UDID	20.129	10.0	0.27	10.0	20.09	70 77	2.033	0.2	9.4	0	000001	к	General function prediction only
STA A00/10	67142	65601	orf	Trk system potassium uptake protein trkH	trkH	0.0	0.0	0.0	0.0	16	4e-167	0.001	4.0	1.053	0	COG0168	Р	Inorganic ion transport and metabolism
STA A00410	60171	68/170	orf		fro	30-87	70-106	10-86	50-74	2.0	0.22	10-74	7.3	0.052	<u> </u>	COG0543	н́с	Coenzyme metabolism [H]
517_700420	33171	30470	011			36-07	100100	10-00	36-74	2.0	0.22	10-74	2.0	0.002		0000343	11, 0	Energy production and conversion [C]
STA Ar0020	70122	72600	-DNIA	225 rBNA gono										<u> </u>				
STA_A10020	73125	73240				<u> </u>	<u> </u>		<u> </u>			+	-					
STA A00/20	73807	73352	orf	Dnak sunnressor protein	dksA	26-78	30-80	3e-78	26-68	30-56	3.0	16-71	16-64	5.1		COG1734	т	Signal transduction mechanisms
STA A00430	74700	74647	nseudogene	Cold shock-like protein cspC*	csnC*	2e-34	26-35	16-34	46-29	2e-34	96-30	16-20	36-33	8e-19	٩	COG1278	ĸ	Transcription
0.7(_700-140	. 4133		pesudogene		0000	3e-05	1e-05	8e-06	7e-05	4e-06	3e-04	3e-04	5e-07	0.44	10	5551210		

STA_A004	50 7608	8 76945	orf	RNA polymerase sigma-32 factor	rpoH	4e-132	1e-151	7e-143	3e-126	8e-120	9e-123	7e-130	8e-123	2e-04	COG0568	K	Transcription
STA_A004	60 8051	0 81673	orf	Phosphoglycerate kinase	pgk	3e-178	0.0	7e-177	6e-163	2e-158	3e-101	3e-147	5e-123	1.1	COG0126	G	Carbohydrate transport and metabolism
STA_A004	70 8178	1 82860	orf	Fructose-bisphosphate aldolase	fba	9e-163	0.0	3e-164	4e-158	2e-152	4e-131	5e-152	3e-133	1.5	COG0191	G	Carbohydrate transport and metabolism
STA_A004	80 8386	0 83645	orf	50S ribosomal protein L31	rpmE	5e-28	1e-30	4e-29	4e-23	1e-25	4e-19	3e-27	1e-26	0.001	COG0254	J	Translation, ribosomal structure and biogenesis
STA_A004	90 8406	0 85550	orf	Primosomal protein N	priA	1e-160	0.0	4e-170	1e-148	1.7	5e-06	7e-151	8e-74	0.078	COG1198	L	DNA replication, recombination and repair

* To search for 'Candidatus Steffania adelgidicola' pseudogene homologs the respective S. glossinidius or E. coli proteins were used (see remarks).

Ec, Escherichia coli K-12 substr. MG1655 [NC_000913] Sg, Sodalis glossinidius str. 'morsitans' [NC_007712]

Ss, Serratia symbiotica str. Tucson [AENX0000000]

Hd, 'Candidatus Hamiltonella defensa 5AT' (*Acyrthosiphon pisum*) [NC_012751] Bp, '*Candidatus* Blochmannia pennsylvanicus str. BPEN' [NC_007292]

Wg, Wigglesworthia glossinidia endosymbiont of Glossina brevipalpis [NC_004344]

Bc, Baumannia cicadellinicola str. Hc (Homalodisca coagulata) [MC_007984] Ba, Buchnera aphidicola str. APS (Acyrthosiphon pisum) [NC_002528]

Cr, 'Candidatus Carsonella ruddii PV' [NC_008512]

nss = no significant similarity

¹ blastp with RnhB protein of Sodalis glossinidius

² blastx with nucleotide sequence of 'Candidatus Steffania adelgidicola'

³ blastp with YfiH protein of *E. coli*

⁴ blastx with nucleotide sequence of ' Candidatus Steffania adelgidicola'

⁵ blastp with ClpB protein of Sodalis glossinidius

⁶ blastx with nucleotide sequence of 'Candidatus Steffania adelgidicola'

⁷ blastp with UbiB protein of Sodalis glossinidius

⁸ blastx with nucleotide sequence of ' Candidatus Steffania adelgidicola'

⁹ blastp with UbiB protein of Sodalis glossinidius

¹⁰ blastx with nucleotide sequence of ' Candidatus Steffania adelgidicola'

Table S4. Genome features of 'Candidatus Steffania adelgidicola' and selected free-living bacteria and insect symbionts.

Genome features according to the NCBI Genome database. Values for 'Candidatus' Steffania adelgidicola' are based on a 85 kb genome fragment.

	Escherichia coli str. K-12 subser	Versinia pestis KIM 10.5	Mycobacterium leprae Br4923 INC-011896	Sodalis glossinidius str. m.	Serratia symbiotica str. Turo.	Candidatus Hamiltonella des.	Candidatus Steffania ador.	Candidatus Block.	Wigglesworthia or.	Baumannia cicadellinico. Inc	Buchnera aphidicola str. Hc INC 001984]	Candidatus Carsonella nucional	Candidatus Sulcia muelleri GWSS [,]	Candidatus Hodgkinia cicadicola Ds _{em} ,
Phylogenetic affiliation	Gammaprot	eobacteria	Actinobacteria				Gamma	proteobact	eria				Bacteriodetes	Alphaproteob.
Lifesstyle	Free-living commensal	Free-living pathogen	Obligate intracellular pathogen	Facultat	ive insect sy	/mbionts				OI	oligate inse	ect symbior	nts	
Age of symbiosis ^b [my]				n.d.	~90	n.d.	n.d.	50	>40	100	180	120	>270	190
Genome size [bp]	4,639,675	4,600,755	3,268,071	4,171,146	2,573,085	2,110,331	n.d.	791,654	697,724	686,194	640,681	159,662	245,530	143,795
GC content [%]	50	47	57	54	49	40	39.1	29	22	33	26	16	22	58
% Coding	85	82	49	50.9 ^a	56	80	59.3	76	86	85	86	93	91	90
Structural RNAs	175	100	50	91	72	85	>5	44	40	46	36	31	36	19
CDS	4145	4049	1604	2432	2157	2094	n.d.	610	611	595	564	182	227	169
Pseudogenes/ pseudogenes per 10 kbp	179/0.386	54/0.117	1116/3.415	972 ^a /2.33	550/2.138	187/0.886	5/0.584	4/0.051	0/0	10/0.146	1/0.016	0/0	1/0.041	0/0
Average CDS length	941.4	931.7	998.4	857.6	668	806.24	805	948.9	982.1	980.3	976.9	815.9	984.3	765.8
Average spacer length	167.9	204.5	1039.1	857.6	524.9	201.6	909.3	299.7	159.9	173	159	61.4	97.4	85.1

^a modified using the publication of the *Sodalis glossinidius* genome (Toh et al, 2006)

^o (Burke & Moran 2011; Gosalbes et al, 2010; Moran et al, 2008)



Figure S1. Adelgids analyzed in this study. (A) Adelgid exules (a wingless parthenogenetic life stage) feeding on a branch of silver fir (*Abies alba*). **(B)** Phylogenetic affiliation of the studied adelgids with the *A. nordmanniae/piceae* complex based on a concatenated dataset of *ef1alpha* and *col*. A maximum likelihood tree (PhyML) based on nucleotide sequence alignments is shown. Maximum likelihood bootstrap values (1000 resamplings), TREEPUZZLE support values and maximum parsimony bootstrap values (1000 resamplings) are indicated at the internal nodes. Selected members of the Phylloxeridae and the Aphididae were used as outgroup. GenBank accession numbers for *ef1alpha* and *col* genes, respectively, are given in square brackets. Bar, 10% estimated evolutionary distance.



Figure S2. In situ identification of 'Candidatus Steffania adelgidicola' and 'Candidatus Ecksteinia adelgidicola' in bacteriocytes of further A. nordmannianae/piceae populations. Oligonucleotide probes targeting the 16S were used in FISH experiments for the assignment of the 16S rRNA gene sequences to the two different symbionts in the adelgid bacteriome. Probe names and sequences are listed in Supplemental Table S1. 'Candidatus Steffania adelgidicola' labeled with a specific probe labeled in Cy3 (red), 'Candidatus Ecksteinia adelgidicola' labeled with specific probes labeled in Cy5 (blue) and the general bacterial probe mix labeled with FLUOS (green) were used; the combined signal from bacterial probe labeled with FLUOS (green) and the 'Candidatus Steffania adelgidicola-specific probe appears yellow and the 'Candidatus Ecksteinia adelgidicola'-specific probes appears blue-green. to increase signal intensity three different 16S rRNA probes were applied simultaneously for 'Candidatus Ecksteinia adelgidicola'. Adelgid populations sampled in (A) Grafrath, Germany and (B) Klausen-Leopoldsdorf, Austria are shown. Bars represent 10 µm.



Figure S3. In situ identification of '*Candidatus* Steffania adelgidicola' and '*Candidatus* Ecksteinia

adelgidicola'. Oligonucleotide probes targeting the 16S and 23S rRNA were used in FISH experiments to assign 16S and 23S rRNA sequence types to the two different symbiont populations in the adelgid bacteriome. Probe names and sequences are listed in Supplemental Table S1. (A) 'Candidatus Steffania adelgidicola' labeled with specific probes targeting the 23S rRNA labeled in Cy3 (red) and the 16S rRNA labeled in Cy5 (blue); the combined signal from bacterial probe labeled with FLUOS (green) and symbiontspecific probes appears white. (B) 'Candidatus Ecksteinia adelgidicola' labeled with specific probes targeting the 23S rRNA labeled in Cy3 (red) and the 16S rRNA labeled in Cy5 (blue); to increase signal intensity three different 16S or 23S rRNA probes were applied simultaneously. The green signal from the single bacterial probe is very dim, hence the combined signal from bacterial and symbiontspecific probes appears purple, 'Candidatus Steffania adelgidicola' appears green. Bars represent 10 µm.



Figure S4. Screening of ten single adelgid individuals (Gosau) by PCR using symbiont specific 16S rRNA gene primers for the presence of **(A)** ,*Candidatus* Steffania adelgidicola' and **(B)** ,*Candidatus* Ecksteinia adelgidicola' (Supplementary Table S1). Lanes 1 to 10 indicate the presence of PCR products of the correct size (1197 bp and 1342 bp, respectively) in all of the screened individuals. N = negativ control (no DNA added), M = molecular weight marker



Figure S5. Screening for the presence of both symbionts in different life stages. PCR was performed using symbiont specific primer for the symbionts (A) ,*Candidatus* Steffania adelgidicola' and (B) ,*Candidatus* Ecksteinia adelgidicola' (Supplementary Table S1). M = molecular weight marker, E = exules, S = sexuparae, + = positive control, N = negative control (no DNA added).



Figure S6. Genome fragment of '*Candidatus* Steffania adelgidicola'. An 85 kb contig reconstructed from three fosmid clones is shown. Predicted CDSs are indicated in blue, rRNA and tRNA genes in green, pseudogenes in red. Deviations from the average G+C content (39.1%) are displayed in black and grey, respectively. Scale bar in bp.



Figure S7. Phylogenetic relationships of '*Candidatus* Steffania adelgidicola' based on protein sequences of RpoH. TREEPUZZLE tree is shown. Bars, 10% estimated evolutionary distance.

References

- Amann RI, Binder BJ, Olson RJ, Chisholm SW, Devereux R, Stahl DA (1990). Combination of 16S rRNA-targeted oligonucleotide probes with flow cytometry for analyzing mixed microbial populations. *Appl Environ Microbiol* **56**: 1919-1925.
- Burke GR, Moran NA (2011). Massive genomic decay in *Serratia symbiotica*, a recently evolved symbiont of aphids. *Genome Biol Evol* 2011; epub ahead of print 25 January 2011, doi: 10.1093/gbe/evr002
- Daims H, Bruhl A, Amann R, Schleifer KH, Wagner M (1999). The domain-specific probe EUB338 is insufficient for the detection of all *Bacteria*: development and evaluation of a more comprehensive probe set. *Syst Appl Microbiol* **22**: 434-444.
- Gosalbes MJ, Latorre A, Lamelas A, Moya A (2010). Genomics of intracellular symbionts in insects. Int J Med Microbiol 300: 271-278.
- Guryev V, Makarevitch I, Blinov A, Martin J (2001). Phylogeny of the genus *Chironomus* (Diptera) inferred from DNA sequences of mitochondrial cytochrome b and cytochrome oxidase I. *Mol Phylogenet and Evol* **19**: 9-21.
- Havill NP, Foottit RG, von Dohlen CD (2007). Evolution of host specialization in the Adelgidae (Insecta: Hemiptera) inferred from molecular phylogenetics. *Mol Phylogenet Evol* 44: 357-370.
- Hunt DE, Klepac-Ceraj V, Acinas SG, Gautier C, Bertilsson S, Polz MF (2006). Evaluation of 23S rRNA PCR primers for use in phylogenetic studies of bacterial diversity. *Appl Environ Microbiol* **72**: 2221-2225.
- Juretschko S, Timmermann G, Schmid M, Schleifer KH, Pommerening-Roser A, Koops HP *et al* (1998). Combined molecular and conventional analyses of nitrifying bacterium diversity in activated sludge: *Nitrosococcus mobilis* and *Nitrospira*-like bacteria as dominant populations. *Appl Environ Microbiol* **64:** 3042-3051.
- Lane DJ (1991). 16S/23S rRNA sequencing. Wiley: Chichester.
- Loy A, Schulz C, Lucker S, Schopfer-Wendels A, Stoecker K, Baranyi C *et al* (2005). 16S rRNA gene-based oligonucleotide microarray for environmental monitoring of the betaproteobacterial order "*Rhodocyclales*". *Appl Environ Microbiol* **71**: 1373-1386.
- Moran NA, McCutcheon JP, Nakabachi A (2008). Genomics and evolution of heritable bacterial symbionts. Annl Rev Gen 42: 165-190.
- Toh H, Weiss BL, Perkin SAH, Yamashita A, Oshima K, Hattori M et al (2006). Massive genome erosion and functional adaptations provide insights into the symbiotic lifestyle of *Sodalis glossinidius* in the tsetse host. *Genome Res* **16**: 149-156.
- Wallner G, Amann R, Beisker W (1993). Optimizing fluorescent in situ hybridization with rRNA-targeted oligonucleotide probes for flow cytometric identification of microorganisms. *Cytometry* **14**: 136-143.