

**Bacteriocyte-associated gammaproteobacterial
symbionts of the *Adelges piceae/nordmanniana*
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 <i>Bacteria</i> | 8-25 | 52 | (Juretschko <i>et al.</i> , 1998) |
| 1492R | GGY TAC CTT GTT ACG ACT T | 16S rRNA gene, most <i>Bacteria</i> and <i>Archaea</i> | 1492-1510 | 52 | (Loy <i>et al.</i> , 2005) |
| TTL1-16S-70F | CAT CGG AAA GGA GTT TAC TTC | 16S rRNA gene, ' <i>Candidatus</i> Steffania adelgadicola' | 70-90 | 58 | This study |
| TTL1-16S-1267R | GAG GTC CGC TGA CCC TCA | 16S rRNA gene, ' <i>Candidatus</i> Steffania adelgadicola' | 1267-1284 | 58 | This study |
| TTL4-16S-107F | GGA CGG GTG AGT AAT ATT | 16S rRNA gene, ' <i>Candidatus</i> Ecksteinia adelgadicola' | 107-124 | 58 | This study |
| TTL4-16S-1449R | GTA AGT GCC CTC CAA TAC | 16S rRNA gene, ' <i>Candidatus</i> Ecksteinia adelgadicola' | 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, ' <i>Candidatus</i> Steffania adelgadicola' | 347-364 | 64 | This study |
| TTL1-23S-1169R | CCA GCA ACA CTC TCA TGC | 23S rRNA gene, ' <i>Candidatus</i> Steffania adelgadicola' | 1169-1186 | 64 | This study |
| TTL1-23S-347R | GCT CAC AAC TAA CAC ACC | 23S rRNA gene, ' <i>Candidatus</i> Steffania adelgadicola' | 347-364 | n.d. | This study |
| TTL4-23S-260F | GGG ACA GCC CAG AGC TAG | 23S rRNA gene, ' <i>Candidatus</i> Ecksteinia adelgadicola' | 260-277 | 64 | This study |
| TTL4-23S-1169R | GCA ATG CAT ATT TCA CAT | 23S rRNA gene, ' <i>Candidatus</i> Ecksteinia adelgadicola' | 1169-1183 | 64 | This study |
| Host gene targeting 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 probes | | | | | |
| 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, <i>Bacteria</i> 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, <i>Bacteria</i> 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, ' <i>Candidatus</i> Steffania adelgadicola' | 1028-1045 | 30-35 | This study |
| TTL4-608 | CAC GTT AAG CGC AGG GAT | 16S rRNA, ' <i>Candidatus</i> Ecksteinia adelgadicola' | 609-626 | 30-35 | This study |
| TTL4-1264 | GGT TTG CTT ACT CTT GCG | 16S rRNA, ' <i>Candidatus</i> Ecksteinia adelgadicola' | 1265-1282 | 30-35 | This study |
| TTL4-1445 | GCC CTC CAA TAC ATG GTT | 16S rRNA, ' <i>Candidatus</i> Ecksteinia adelgadicola' | 1446-1460 | 30-35 | This study |
| NONEUB | ACT CCT ACG GGA GGC AGC | Control probe complementary to EUB338-I | 338-355 | n.d. | (Wallner <i>et al.</i> , 1993) |

| | | | | | |
|------------------|-------------------------|---|-----------|-------|------------|
| TTL1-1215 | GCA CAC CTT ACA GTG CAC | 23S rRNA, ' <i>Candidatus Steffania adelgadicola</i> ' | 1216-1233 | 10-25 | This study |
| TTL4-877 | AAG CTG GGA TAG CCC CTT | 23S rRNA, ' <i>Candidatus Ecksteinia adelgadicola</i> ' | 878-895 | 10-25 | This study |
| TTL4-967 | GTT GTT TCC CTT TAC ACG | 23S rRNA, ' <i>Candidatus Ecksteinia adelgadicola</i> ' | 968-985 | 10-25 | This study |
| TTL4-1215 | GCA ATC CTC ACG AGT CAC | 23S rRNA, ' <i>Candidatus Ecksteinia adelgadicola</i> ' | 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. nordmanniana* and *A. piceae* (grey background) and additional variable sites among the *A. nordmanniana*/*A. piceae* species complex. All available *ef1alpha* gene sequences of this complex are shown.

| Organism | Sampling location | Genbank Acc. Number | Alignment position ^a | | | | | | | | | |
|--|-------------------------------|---------------------|---------------------------------|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| | | | 203 | 204 | 295 | 331 | 398 | 446 | 509 | 548 | 649 | 650 |
| <i>Adelges nordmanniana</i> | Slovak Republic | EF073235 | T | T | T | T | T | T | A | T | T | C |
| <i>Adelges piceae</i> | Owishead; Maine; USA | EF073239 | A | - | T | T | C | C | A | C | T | C |
| <i>Adelges piceae</i> | Delemont; Switzerland | EF073240 | T | T | T | T | C | C | A | C | T | C |
| <i>Adelges nordmanniana/piceae</i> | Gosau; Austria | HQ668164 | T | T | T | T | T | T | A | T | T | C |
| <i>Adelges nordmanniana/piceae</i> | Grafrath; Germany | HQ668167 | T | - | T | T | C | C | A | C | T | C |
| <i>Adelges nordmanniana/piceae</i> Klausen-Leopoldsdorf | Klausen-Leopoldsdorf; Austria | HQ668165 | T | T | T | C | T | T | A | T | T | T |
| <i>Adelges nordmanniana/piceae</i> Klausen-Leopoldsdorf | Klausen-Leopoldsdorf; Austria | HQ668166 | T | - | C | T | C | C | G | C | A | C |

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

| | | | | | | | | | | | | | | | | | | |
|------------|-------|-------|-----|--------------------------------|-------------|--------|--------|--------|--------|--------|--------|--------|--------|-------|--|---------|---|---|
| STA_A00450 | 76088 | 76945 | orf | RNA polymerase sigma-32 factor | <i>rpoH</i> | 4e-132 | 1e-151 | 7e-143 | 3e-126 | 8e-120 | 9e-123 | 7e-130 | 8e-123 | 2e-04 | | COG0568 | K | Transcription |
| STA_A00460 | 80510 | 81673 | orf | Phosphoglycerate kinase | <i>pgk</i> | 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_A00470 | 81781 | 82860 | orf | Fructose-bisphosphate aldolase | <i>fba</i> | 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_A00480 | 83860 | 83645 | orf | 50S ribosomal protein L31 | <i>rpmE</i> | 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_A00490 | 84060 | 85550 | orf | Primosomal protein N | <i>priA</i> | 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 adelgadicola' 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 [AENX00000000]

Hd, 'Candidatus Hamiltonella defensa 5AT' (*Acyrtosiphon 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*) [NC_007984]

Ba, *Buchnera aphidicola* str. APS (*Acyrtosiphon 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 adelgadicola'

³ blastp with YfiH protein of *E. coli*

⁴ blastx with nucleotide sequence of 'Candidatus Steffania adelgadicola'

⁵ blastp with ClpB protein of *Sodalis glossinidius*

⁶ blastx with nucleotide sequence of 'Candidatus Steffania adelgadicola'

⁷ blastp with UbiB protein of *Sodalis glossinidius*

⁸ blastx with nucleotide sequence of 'Candidatus Steffania adelgadicola'

⁹ blastp with UbiB protein of *Sodalis glossinidius*

¹⁰ blastx with nucleotide sequence of 'Candidatus Steffania adelgadicola'

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.

| | <i>Escherichia coli</i> str. K-12 substr. MG1655 [NC_000913] | | <i>Yersinia pestis</i> KIM 10 [NC_004088] | <i>Mycobacterium leprae</i> Br4923 [NC_011896] | <i>Sodalis glossinidius</i> str. 'morsitans' [NC_007712] | <i>Serratia symbiotica</i> str. Tucson [NZ_AENX00000000] | <i>Candidatus Hamiltonella defensa</i> 5AT' [NC_012751] | <i>Candidatus Steffania adelgidicola</i> [FR872579] | <i>Candidatus Blochmannia pennsylvanicus</i> str. BPEN' [NC_007292] | <i>Wigglesworthia glossinidia</i> [NC_004344] | <i>Baumannia cicadellincola</i> str. Hc [NC_007984] | <i>Buchnera aphidicola</i> str. APS [NC_002528] | <i>Candidatus Carsonella rudii</i> PV' [NC_008512] | <i>Candidatus Sulcia muelleri</i> GWSS' [NC_010118] | <i>Candidatus Hodgkinia cicadicola</i> Dsem' [NC_012960] |
|------------------------------------|--|----------------------|---|--|--|--|---|---|---|---|---|---|--|---|--|
| Phylogenetic affiliation | Gammaproteobacteria | | Actinobacteria | Gammaproteobacteria | | | | | | | | | Bacteriodesetes | Alphaproteob. | |
| Lifestyle | Free-living commensal | Free-living pathogen | Obligate intracellular pathogen | Facultative insect symbionts | | | Obligate insect symbionts | | | | | | | | |
| 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)

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

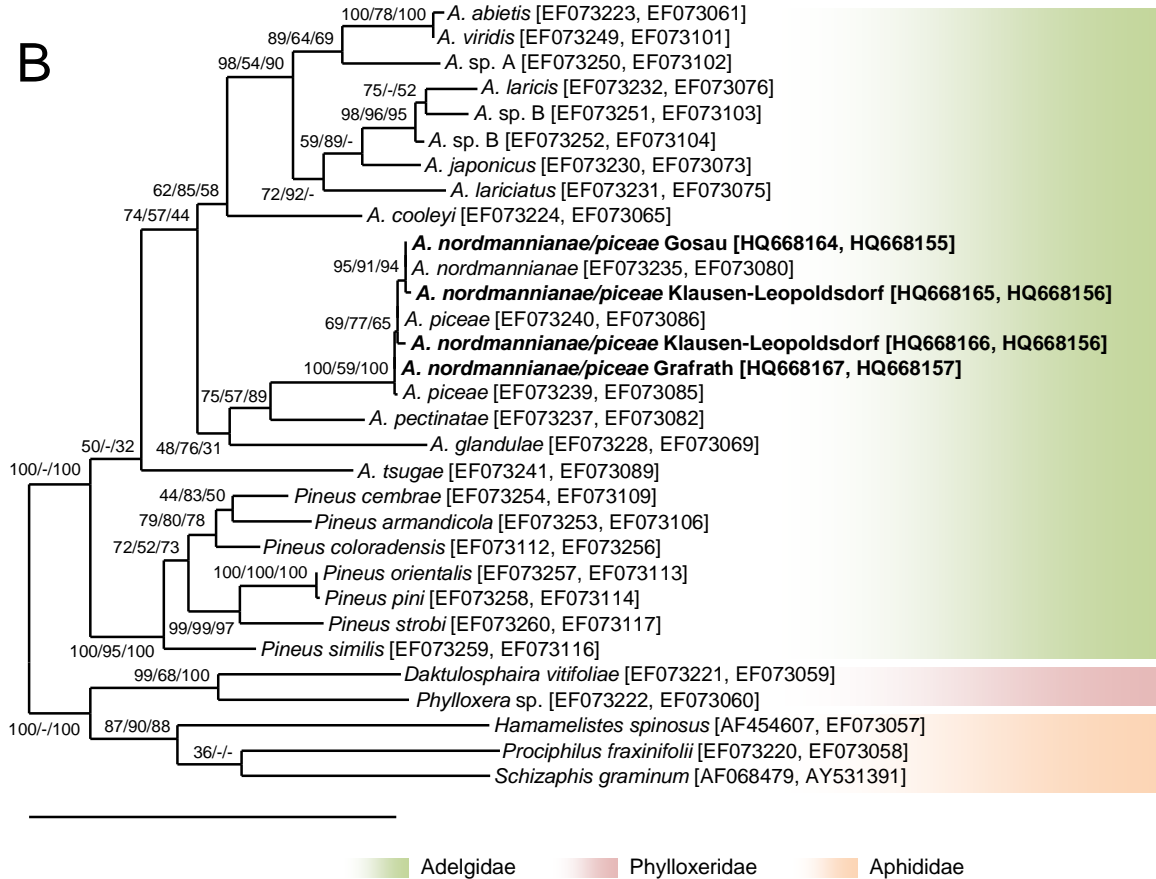


Figure S1. Adelgids analyzed in this study. (A) Adelgid exules (a wingless parthenogenetic life stage) feeding on a branch of silver fir (*Abies alba*). **(B)** Phylogenetic affiliation of the studied adelgids with the *A. nordmanniae/piceae* complex based on a concatenated dataset of *ef1alpha* and *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.

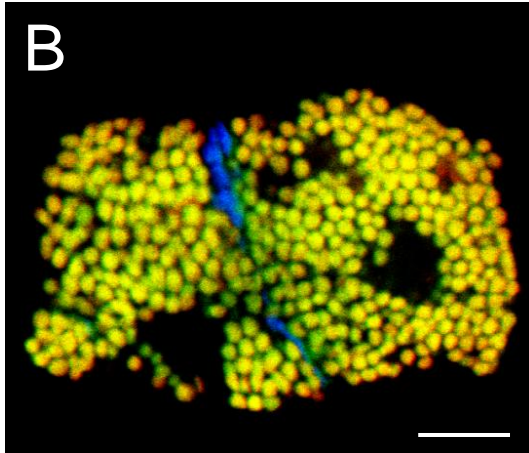
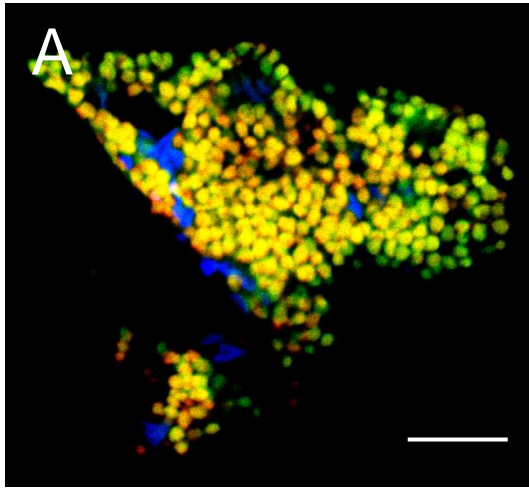


Figure S2. In situ identification of '*Candidatus Steffania adelgicola*' and '*Candidatus Ecksteinia adelgicola*' in bacteriocytes of further *A. nordmanniana/piceae* populations. Oligonucleotide probes targeting the 16S were used in FISH experiments for the assignment of the 16S rRNA gene sequences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

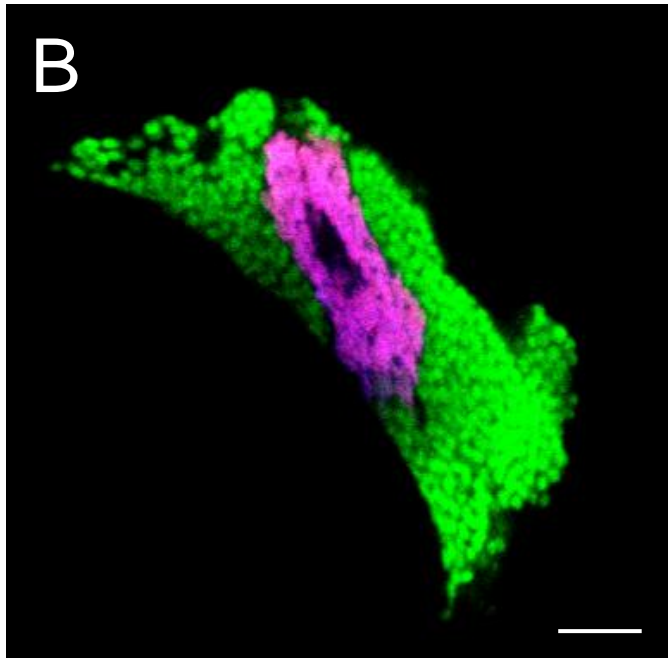
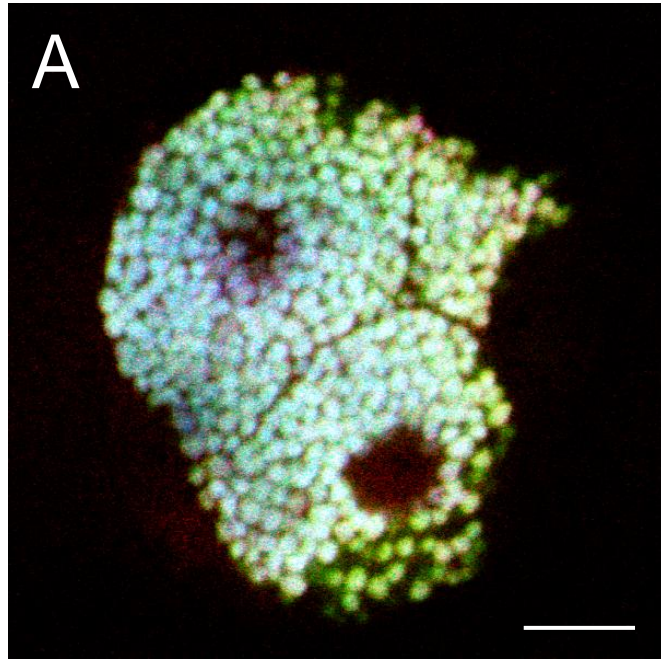


Figure S3. In situ identification of '*Candidatus Steffania adelgadicola*' and '*Candidatus Ecksteinia adelgadicola*'. Oligonucleotide probes targeting the 16S and 23S rRNA were used in FISH experiments to assign 16S and 23S rRNA sequence types to the two different symbiont populations in the adelgid bacteriome. Probe names and sequences are listed in Supplemental Table S1. **(A)** '*Candidatus Steffania adelgadicola*' 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 symbiont-specific probes appears white. **(B)** '*Candidatus Ecksteinia adelgadicola*' 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 symbiont-specific probes appears purple, '*Candidatus Steffania adelgadicola*' appears green. Bars represent 10 μm .

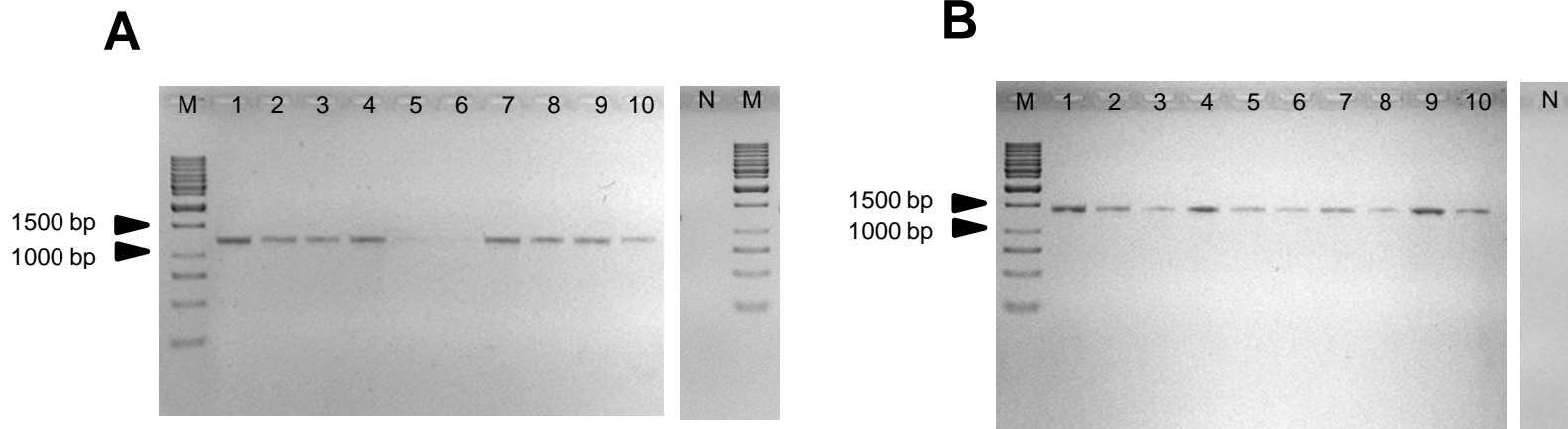


Figure S4. Screening of ten single adelgid individuals (Gosau) by PCR using symbiont specific 16S rRNA gene primers for the presence of **(A)** ,*Candidatus* Steffania adelgicola' and **(B)** ,*Candidatus* Ecksteinia adelgicola' (Supplementary Table S1). Lanes 1 to 10 indicate the presence of PCR products 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

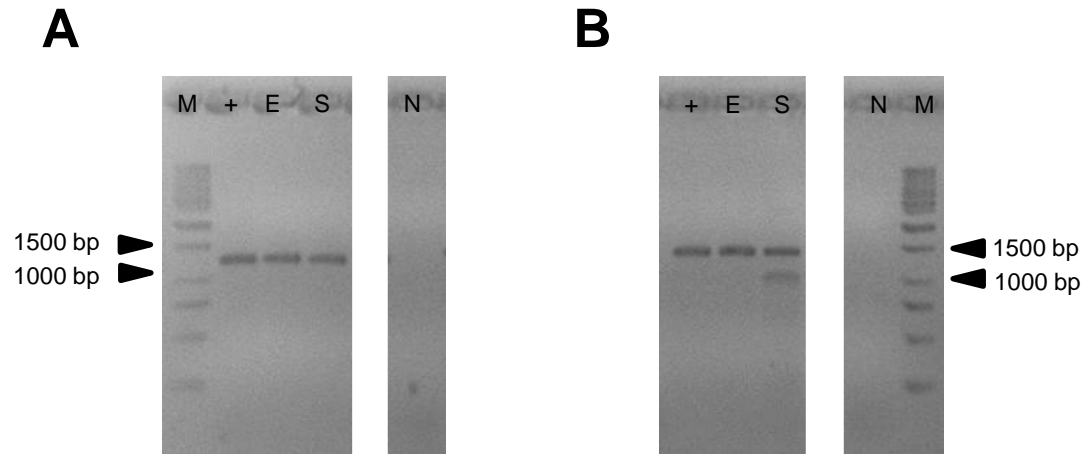


Figure S5. Screening for the presence of both symbionts in different life stages. PCR was performed using symbiont specific primer for the symbionts **(A)** ,*Candidatus Steffania adelgadicola* and **(B)** ,*Candidatus Ecksteinia adelgadicola* (Supplementary Table S1). M = molecular weight marker, E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

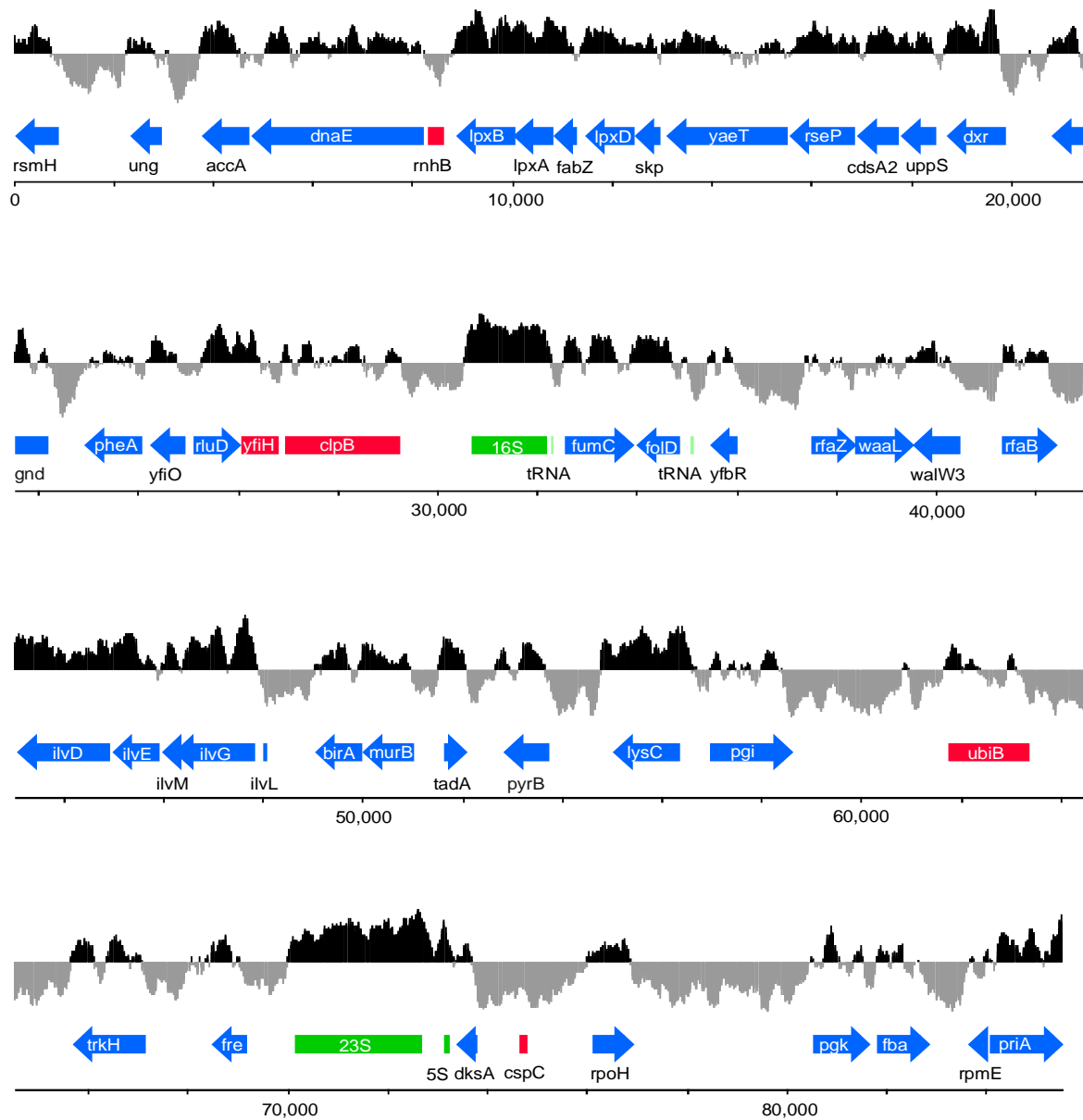


Figure S6. Genome fragment of ‘*Candidatus Steffania adelgadicola*’. An 85 kb contig reconstructed from three fosmid clones is shown. Predicted CDSs are indicated in blue, rRNA and tRNA genes in green, pseudogenes in red. Deviations from the average G+C content (39.1%) are displayed in black and grey,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

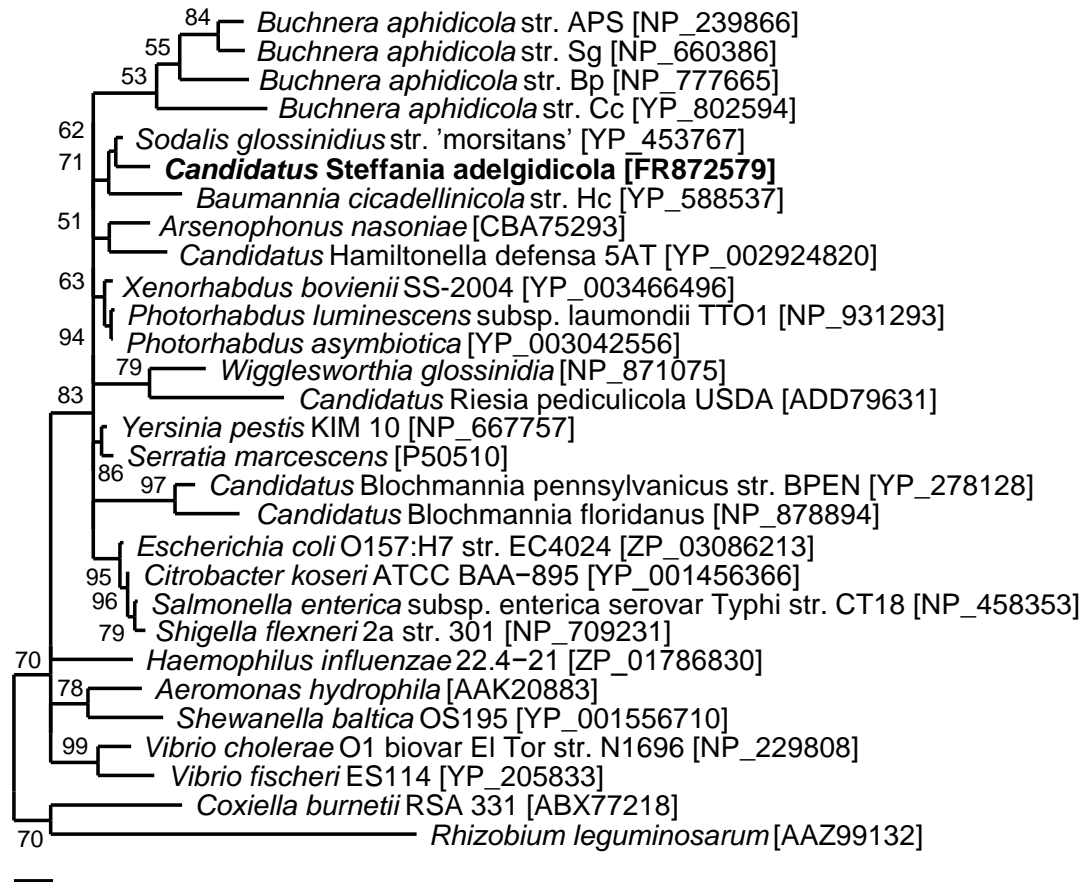


Figure S7. Phylogenetic relationships of 'Candidatus Steffania adelgidicola' based on protein sequences of RpoH. TREEPUZZLE tree is shown. Bars, 10% estimated evolutionary distance.

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; e-pub 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, Footitt 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. *Annu 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.