AEM Accepts, published online ahead of print on 31 August 2007
Appl. Environ. Microbiol. doi:10.1128/AEM.01131-07

Revised to be re-submitted to *Applied and Environmental Microbiology* on 8 Aug. 2007

Bacterial endosymbiont of the slender pigeon louse *Columbicola columbae* allied to endosymbionts of grain weevils and tsetse flies

Takema Fukatsu,[1*] Ryuichi Koga,[1] Wendy A. Smith,[2] Kohjiiro Tanaka,[1] Naruo Nikoh,[3] Kayoko Sasaki-Fukatsu,[1] Kazunori Yoshizawa,[4] Colin Dale,[2] and Dale H. Clayton[2]

Institute for Biological Resources and Functions, National Institute of Advanced Industrial Science and Technology (AIST), Tsukuba 305-8566, Japan[1]; Department of Biology, University of Utah, UT84106, USA[2]; Division of Natural Sciences, University of the Air, Chiba 261-8586, Japan[3]; Department of Ecology and Systematics, Hokkaido University, Sapporo 060-8589, Japan[4]

*Corresponding author. Mailing address: National Institute of Advanced Industrial Science and Technology (AIST), Tsukuba 305-8566, Japan. Phone: 81-29-861-6087. Fax: 81-29-861-6080. E-mail: t-fukatsu@aist.go.jp.

Running title: Endosymbiont of slender pigeon louse

Abstract

The current study focuses on a symbiotic bacterium found in the slender pigeon louse *Columbicola columbae* (Insecta: Phthiraptera). Molecular phylogenetic analyses indicated that the symbiont belongs to the γ-subdivision of the class *Proteobacteria* and is allied to *Sodalis glossinidius*, the secondary symbiont of the tsetse flies *Glossina* spp., and also to the primary symbiont of the grain weevils *Sitophilus* spp. Relative rate tests revealed that the symbiont of *C. columbae* exhibits accelerated molecular evolution in comparison with the tsetse symbiont and the weevil symbiont. Whole mount in situ hybridization was used to localize the symbiont and determine infection dynamics during host development. In 1st and 2nd instar nymphs, the symbionts were localized in the cytoplasm of oval bacteriocytes that formed small aggregates on the both sides of the body cavity. In 3rd instar nymphs, the bacteriocytes migrated to the central body and were finally located in the anterior region of the lateral oviducts, forming conspicuous tissue formations called ovarial ampullae. In adult females, the symbionts were transmitted from the ovarial ampullae to developing oocytes in the ovarioles. In adult males, the bacteriocytes often disappeared without migration. Diagnostic PCR survey of insects collected from Japan, USA, Australia and Argentina detected 96.5% (109/113) infection, with a few uninfected male insects. This study provides the first microbial characterization of a bacteriocyte-associated symbiont from a chewing louse. Possible biological roles of the symbiont are discussed in relation to the host nutritional physiology associated with the feather-feeding lifestyle.

INTRODUCTION

Symbiotic associations with microorganisms are ubiquitous among a diverse array of insects. Some obligate symbionts are mutualistic in nature and contribute to the fitness of their hosts, while others are facultative and may have negative impacts upon host fitness (4, 5, 8). In many of these intimate associations, the symbionts are housed in specialized cells known as bacteriocytes or mycetocytes. Regardless of their obligate or facultative nature, most of these symbiotic microorganisms are vertically transmitted at early stages of oogenesis or embryogenesis, wherein the transmission process is integrated into the life cycle of the host insects (6, 14, 42).

A number of insects live solely on diets that are nutritionally incomplete or difficult to utilize, such as woody materials (hard to digest, low nitrogen), plant sap (few proteins and lipids), vertebrate blood (deficient in B vitamins), and others. In many of these cases, symbiotic microorganisms have been shown to play crucial roles in compensating for these nutritional deficiencies. In termites, for example, gut protozoans and bacteria enable the host to digest cellulose. In addition, some of these bacterial symbionts are involved in nitrogen fixation for the termite host (7, 30). In aphids, the endocellular bacterial symbiont *Buchnera aphidicola* efficiently synthesizes essential amino acids that are lacking in plant phloem sap (12, 39). In tsetse flies, the endocellular bacterial symbiont *Wigglesworthia glossinidia* provides the host with B vitamins that are lacking in vertebrate blood (1, 28).

Chewing lice (Insecta: Phthiraptera), embracing over 4,400 described species in the world, are ectoparasitic insects feeding on avian feather or mammalian skin and skin products (32). The main component of feather and hair is keratin, a protein constituting the intermediate filament of eukaryotic cells, concentrated in hard animal tissues such as feather, hair, nail,

63 scale, beak and horn, and resistant to solubilization, proteolysis and digestion (18, 27).
64 Although protein-rich and potentially nutritious, these hard tissues are difficult to utilize for
65 most animals, with only a few insect groups such as chewing lice (Phthiraptera: Ischnocera,
66 Amblycera), carpet beetles (Coleoptera: Dermestidae), keratin beetles (Coleoptera: Trogidae)
67 and clothes moths (Lepidoptera: Tineidae) having evolved the ability to live on this difficult
68 diet (21, 43). Possibly relevant to the nutritional difficulty, some chewing lice possess a
69 well-developed endosymbiotic association, wherein bacteriocyte-associated symbiotic
70 bacteria migrate to the ovary in a peculiar passage and are vertically transmitted to the
71 oocytes in the maternal body (8, 34). Although these bacteria were visualized in early
72 histological studies, no formal identification has yet been provided by molecular phylogenetic
73 analyses. Previous studies have however demonstrated the presence of facultative
74 endosymbiotic bacteria of the genus *Wolbachia* in many chewing and sucking lice (25) and a
75 diverse assemblage of putative gut bacteria in the chewing lice of pocket gophers (33).
76 In this study, we present the first microbiological characterization of the
77 bacteriocyte-associated symbiotic bacterium in the slender pigeon louse *Columbicola*
78 *columbae* using molecular phylogenetic analyses and histological techniques. Although
79 chewing lice are phylogenetically close to sucking lice (3, 22), the symbiont of *C. columbae*
80 was not related to the symbionts of primate lice *Riesia* spp. (2, 37), but, unexpectedly, was
81 allied to the symbionts of tsetse flies and grain weevils.
82
83 **MATERIALS AND METHODS**
84 **Insect materials**. Samples of the slender pigeon louse, *C. columbae,* used in this study are
85 listed in Table 1. The insects were collected from the rock pigeon, *Columba livia*, and

immediately preserved in acetone (15).

DNA extraction and morphological inspection. Each of the acetone-preserved insects was briefly dried in air, cut into two parts using a razor, and digested in 200 μl of lysis buffer (50 mM Tris-HCl [pH 8.0], 10 mM EDTA, 0.5% SDS, 0.8 mg/ml proteinase K) at 55 ^{o}C overnight. The exoskeleton of the insect was recovered, mounted on a microscope slide, and observed under a light microscope for morphological identification. The lysate was extracted with phenol-chloroform, subjected to ethanol precipitation, and the precipitated DNA was dried and dissolved in 50 μl of TE buffer (20 mM Tris-HCl [pH 8.0], 0.1 mM EDTA).

DNA cloning and sequencing. The DNA samples from individual insects were subjected to PCR amplification of a 1.5 kb segment of the eubacterial *16S rRNA* gene using the primers 16SA1 (5'-AGA GTT TGA TCM TGG CTC AG-3') and 16SB1 (5'-TAC GGY TAC CTT GTT ACG ACT T-3') (17), and a 0.76 kb segment of *fusA* gene using the primers FusAF (5'-CAT CGG CAT CAT GGC NCA YAT HGA-3') and FusAR (5'-CAG CAT CGG CTG CAY NCC YTT RTT-3') (11). The PCR products were subjected to cloning, restriction fragment length polymorphism (RFLP) genotyping, and DNA sequencing as previously described (17).

Molecular phylogenetic analysis. The DNA sequences were subjected to molecular phylogenetic analysis together with the sequences of related γ-proteobacteria that exhibited high BLAST scores in the DNA database search. A multiple alignment of the sequences was generated using the program Clustal W (40). Aligned nucleotide sites containing a gap were removed from the data set, and the final alignment was inspected and corrected manually. Neighbor joining (NJ) trees, with 1,000 bootstrap resamplings, were constructed using Clustal W (40). Kimura's two parameter model was used for correction of multiple substitutions (23).

Maximum parsimony (MP) trees, with 1,000 bootstrap resamplings, were generated by the program MEGA 3.1 (24). For finding the MP trees, the Close-Neighbor-Interchange algorithm was used. An initial tree was generated by random addition of the sequences. Maximum likelihood (ML) trees were constructed by the program TREE-PUZZLE 5.2 (38), wherein supporting values for internal nodes were inferred by 1,000 puzzling steps. As a nucleotide substitution model, the HKY+Γ+Inv model was used. We tested several different substitution models and confirmed that the differences of the substitution models did not lead to any discrepancies in the tree topologies supported with high bootstrap values.

Relative rate test. A relative rate test, based on genetic distances estimated under the Kimura's two parameter model (23), was performed using the program RRTree (35). For *16S rRNA* gene sequences, 1,444 unambiguously aligned nucleotide sites were subjected to the analysis. For *fusA* gene sequences, 499 unambiguously aligned nucleotide sites at 1st and 2nd codon positions were analyzed, while nucleotide sites at 3rd codon positions were omitted from the analysis due to saturated nucleotide substitutions.

Whole mount fluorescent in situ hybridization (wFISH). An oligonucleotide probe specific to the *16S rRNA* sequence from *C. columbae*, AlexaFluor555-CcolSol427R (Al555-5'-CAT CGC CTT CCT CCC AGT CG-3'), was used for whole-mount fluorescent in situ hybridization (wFISH). After being decapitated to facilitate infiltration of reagents, the acetone-preserved insects were fixed in Carnoy's solution (chloroform-ethanol-acetic acid [6:3:1]) for two days. Subsequently the insects were incubated with 6% H_2O_2 in ethanol for two weeks to quench the autofluorescence of insect tissues. The insects were thoroughly washed, and equilibrated with a hybridization buffer (20 mM Tris-HCl [pH 8.0], 0.9 M NaCl, 0.01% sodium dodecyl sulfate, 30% formamide), and the probe and SYTOX green were

added at final concentrations of 10 nM and 5 μM, respectively. After an overnight incubation, the samples were thoroughly washed in 20 mM Tris-HCl [pH 8.0], 0.9 M NaCl, 0.01% sodium dodecyl sulfate, and observed under an epifluorescent microscope (Axiophoto; Carl Zeiss) and a laser confocal microscope (PASCAL5; Carl Zeiss). To confirm specific detection of the symbionts, a series of control experiments, namely no-probe control, RNase digestion control and competitive suppression control with excess unlabelled probe, were conducted as previously described (36).

Nucleotide sequence accession numbers. The nucleotide sequences determined in this study have been deposited in the DDBJ/EMBL/GenBank nucleotide sequence databases under accession numbers AB303382-AB303387 and EU021695-EU021697 (also see Table 1, Fig. 1 and Fig. 2).

RESULTS

Bacterial *16S rRNA* gene sequences from *C. columbae*. From all the insect samples collected in Japan, USA, Australia and Argentina, nearly identical *16S rRNA* gene sequences, 1,478 bp in size and exhibiting 99.9-100% sequence identities to each other, were identified. For each of the samples, more than ten clones of the *16S rRNA* gene segment showed identical RFLP patterns, indicating a single bacterial species dominant in the insects. A BLAST search clearly showed that the sequence belongs to the Enterobacteriaceae in the γ-*Proteobacteria*. In the DNA databases, we found several high score hits including the secondary symbiotic bacteria *Sodalis glossinidius* from the tsetse flies *Glossina* spp. (ex. AY861701; 96.5% sequence identity) and the primary symbiotic bacteria from the grain weevils *Sitophilus* spp. (ex. AF005235; 96.0% sequence identity).

Phylogenetic placement of the symbiont of *C. columbae* based on *16S rRNA* gene sequences. These *16S rRNA* gene sequences were subjected to molecular phylogenetic analysis together with the sequences of related γ-proteobacteria that exhibited high BLAST scores in the DNA database search (Fig. 1). The bacterial sequences from different *C. columbae* populations formed a monophyletic group with nearly 100% statistical support, constituting a distinct lineage in the γ-subclass of the *Proteobacteria*. The sequences also formed a monophyletic group together with the sequences of the tsetse symbionts and the weevil symbionts, which also garnered close to 100% statistical support.

Phylogenetic placement of the symbiont of *C. columbae* based on a protein-coding gene. From DNA samples from insects collected in Australia and USA, we cloned and sequenced a 760 bp segment of *fusA* gene encoding elongation factor G, a bacterial ribosomal translocase (11). Molecular phylogenetic analysis also showed that the *fusA* sequences from *C. columbae* formed a clade with the sequences from the tsetse symbiont and the weevil symbiont (Fig. 2).

Accelerated molecular evolution in the symbiont of *C. columbae*. On the phylogenetic trees (Figs. 1 and 2), the lineage of the symbionts of *C. columbae* exhibited remarkably elongated branches in comparison with the lineages of the tsetse symbionts and the weevil symbionts. Thus, we performed relative rate tests based on genetic distances between the gene sequences. The evolutionary rate of the *16S rRNA* gene sequence in the lineage of the symbionts of *C. columbae* was 3.1 times and 2.7 times faster than those in the lineages of the tsetse symbionts and the weevil symbionts, respectively. In both cases, the differences were highly significant ($P < 0.001$) (Table 2A). The evolutionary rate of the *fusA* gene sequence in the lineage of the symbionts of *C. columbae* was 25 times and 22 times faster than those in

178 the lineages of the tsetse symbionts and the weevil symbionts, respectively (Table 2B) ($P <$
179 0.01).

180 **AT content of *16S rRNA* gene sequences of the symbiont of *C. columbae***. The *16S*
181 *rRNA* gene sequences derived from the symbiont of *C. columbae* were determined to be
182 45.6-45.7% AT, which were not significantly different from the *16S rRNA* sequences of the
183 tsetse symbionts, the weevil symbionts, and other free-living γ-proteobacteria (Fig. 1).

184 **wFISH of the symbiont in *C. columbae***. In order to investigate *in vivo* localization and
185 infection dynamics of the symbiont, nymphs and adults of *C. columbae* at different
186 developmental stages were subjected to wFISH.

187 **General localization in nymphal and adult insects**. Figure 3 shows wFISH detection of
188 the symbiont in the whole body of *C. columbae* at different developmental stages. In 1st, 2nd
189 and 3rd instar nymphs, aggregates of bacteriocytes were located on the both sides of the body
190 cavity (Fig. 3A-F). In a part of late 3rd instar nymphs and all adult females, the symbiont
191 signals in the lateral body cavity disappeared, and the symbiont cells were localized in the
192 ovary (Fig. 3G-I). In adult males, some individuals possessed the symbiont-harboring
193 bacteriocytes in the lateral body cavity (Fig. 3J) while other individuals exhibited few signals
194 of the symbiont (Fig. 3K).

195 **Localization of the symbiont in lateral aggregates of bacteriocytes in 1st-2nd instar**
196 **nymphs**. In 1st and 2nd instar nymphs, oval bacteriocytes were 10-15 μm in longer diameter,
197 5-7 μm in shorter diameter, and their cytoplasm was observed to be packed with symbiotic
198 bacteria (Fig. 4A and B). These bacteriocytes were found in groups, usually on the both sides
199 of the 3rd and 4th abdominal segments (Fig. 3A-C). The clustered bacteriocytes were
200 arranged linearly just beneath the hypodermis of the abdominal segments (Fig. 4C and D).

Migration of bacteriocytes to lateral oviducts in 3rd instar nymphs. In 3rd instar nymphs, localization of the bacteriocytes showed a drastic change. At the beginning, some of the bacteriocytes in the lateral body cavity were found outside the aggregates, apparently migrating toward the central body region (Fig. 3E and F; Fig. 4E). The more bacteriocytes participated in the migration, the further disintegration of the aggregates proceeded (Fig. 4E-G). Finally, all the bacteriocytes were located at the anterior region of the lateral oviducts, and formed specialized tissue formations for symbiont transmission, so-called ovarial ampullae (8, 34) (Fig. 4H and I).

Vertical transmission of the symbiont from ovarial ampullae to oocytes in adult females. In adult females, the symbiont cells were vertically transmitted from the well-developed ovarial ampullae (Fig. 3H and I; Fig. 4J) to the posterior pole of oocytes in the ovarioles (Fig. 4K-M). The symbionts from the ovarial ampullae passed through follicle cells and reached the posterior pole of oocytes (Fig. 4L), where a specific region was densely infected with the symbiont cells (Fig. 4M).

Prevalence of the symbiont in worldwide populations of *C. columbae*. We examined individuals of *C. columbae* collected from Japan, USA, Australia and Argentina by diagnostic PCR for the symbiont infection. First, 53 insects from 6 populations were inspected without sexing, which revealed 94.3% (50/53) infection. Next, we inspected 30 adult females and 30 adult males from 2 populations, and found 100% (30/30) infection in females and 96.7% (29/30) infection in males. In total, 109 of 113 insects examined were infected with the symbiont, indicating an infection frequency of 96.5% (Table 1).

DISCUSSION

Over 70 years ago, early investigators reported that some chewing lice harbor bacteriocyte-associated endosymbiotic bacteria (8, 34). Since then, however, the nature of the symbiosis has been elusive. To our knowledge, this study provides the first microbiological characterization of a bacteriocyte-associated symbiont from a chewing louse.

Morphologically, chewing lice, consisting mainly of bird lice such as *C. columbae*, are thought to be related to sucking lice, consisting exclusively of mammalian lice such as human lice *Pediculus humanus* and *P. capitis* (46). Recent molecular phylogenetic analyses confirmed that the clade of sucking lice is actually nested in a clade of chewing lice (3, 22). However, we found that the endosymbiont of *C. columbae* was not closely related to the bacterial endosymbiont found in the human lice *Riesia* spp. (Fig. 1). Hence, their symbiotic bacteria are likely of independent evolutionary origins, reflected by the symbiotic organs of different types, namely the highly specialized organ, called the stomach disc, in human lice (13, 34, 37) versus the loosely associated bacteriocytes in *C. columbae* (Figs. 3 and 4) (34). The difference might be relevant to their distinct ecological niches and nutritional physiology: sucking lice persist exclusively on a diet of vertebrate blood whereas chewing lice feed on a keratin-rich diet composed primarily of feather or skin (32). Not only in sucking lice but also in chewing lice, cytology, localization and infection dynamics of their endosymbionts are extremely diverse (8, 34), corroborating the idea that their symbiotic associations are of independent evolutionary origins. Interestingly, there are several peculiar features that are shared between the endosymbiotic systems of sucking lice and chewing lice, such as the symbiont migration to the ovary at the 3rd instar and the specialized tissue formations for symbiont transmission called ovarial ampullae (Figs. 3 and 4) (8, 13, 34, 37). These shared features are perhaps indicative of some common developmental and evolutionary bases

underlying their endosymbiotic systems. In order to better understand the complexities of endosymbiosis in these systems, more intensive surveys are needed to characterize the host and symbiont diversity.

In vivo localization and infection dynamics of the symbiont of *C. columbae* were described in detail in the pioneering work of Ries (34). Our wFISH results (Figs. 3 and 4) were totally concordant with these early histological descriptions. Here we point out an enigmatic phenomenon that Ries (34) and ourselves observed consistently in the migration of the bacteriocytes from the lateral body cavity into the ovary. In 3rd instar nymphs of *C. columbae*, individual bacteriocytes begin to migrate toward the central body region (Fig. 3E and F; Fig. 4E-G), arriving at the anterior region of the lateral oviducts (Fig. 3G; Fig. 4I), whereupon ovarial ampullae are formed for symbiont transmission to the oocytes (Fig. 3H and I; Fig. 4J-M). It should be noted that the bacteriocytes in the lateral body cavity (Fig. 4A-D), the migrating bacteriocytes in the central body region (Fig. 4F and G) and the bacteriocytes located inside the lateral oviducts (Fig. 4I) look very similar cytologically. In the migration process, neither disintegrating bacteriocytes nor extracellular symbiont cells are observed, and thus it seems that the whole bacteriocytes somehow pass through the wall of the oviducts and gain entry into the ovarial cavity (34). The mechanism of the whole cell penetration into the ovary is intriguing, and should be pursued by more detailed histological examinations in the future.

Molecular phylogenetic analyses revealed that the symbiont of *C. columbae* is closely related to the secondary symbiont of the tsetse flies, *S. glossinidius*, and also to primary symbiont of the grain weevils (Figs. 1 and 2). The phylogenetic proximity of the symbionts is somewhat puzzling, since chewing lice, tsetse flies and grain weevils represent different

insect orders Phthiraptera, Diptera and Coleoptera, respectively. To account for the sporadic distribution of the *Sodalis*-allied endosymbionts, there may have been horizontal transfer between the distant insect lineages some time in the evolutionary past, although biological connections between these insects are difficult to imagine. Recently, a new member of *Sodalis*-allied symbiont was identified from a hippoboscid fly *Craterina melbae* (29). It therefore seems likely that the host insect range of this symbiont clade is much broader than previously envisioned.

Recent molecular evolutionary analyses have suggested that the lifestyle of obligate insect symbionts has strongly affected their genome evolution, causing AT-biased nucleotide composition, accelerated rate of molecular evolution and significant genome reduction. These peculiar genetic traits are hypothesized to be the consequence of attenuated purifying selection due to small population size and frequent transmission bottlenecks, which are associated with the lifestyle of vertically transmitted symbionts (20, 45). The symbiont of *C. columbae* exhibited significantly faster molecular evolutionary rates in *16S rRNA* and *fusA* gene sequences than the tsetse symbiont and the weevil symbiont (Table 2), suggesting that these population genetic parameters might be strongly affected in the symbiont of *C. columbae*. To better understand this phenomenon, further studies are required to determine the age of the association between *Columbicola* spp. and their bacterial symbionts. It should be noted that the endosymbioses in the tsetse flies and the grain weevils are presumably of relatively recent origins. The eroded genome of *S. glossinidius* suggests recent transition of the bacterial lifestyle from free-living to endosymbiotic (41). While many of the weevils of the family Dryophthoridae are associated with an ancient symbiont lineage of the genus *Nardonella*, only the grain weevils *Sitophilus* spp. are associated with the *Sodalis*-allied

symbiont, which suggests a later replacement of the endosymbiotic associates in the ancestor of the weevil genus (26). Interestingly the AT-contents of *16S rRNA* gene sequences were not different among the related louse, tsetse and weevil symbionts (Fig. 1). The genome size of the tsetse symbiont *S. glossinidius* is known to be 4,171,146 bp (41), whereas the genome size of the weevil symbiont is estimated to be around 3.0 Mb (9). It will therefore be of interest to determine the genome size of the symbiont of *C. columbae*.

Among chewing lice, Ries (34) made histological observations of bacterial endosymbionts from the genera *Columbicola*, *Sturnidoecus*, *Goniocotes*, *Campanulotes*, *Colocerus*, *Goniodes*, *Anaticola*, *Turdinirmus*, *Kelerinirmus* and *Brueelia*, while no endosymbionts were detected from the family Trichodectidae. Our ongoing work will determine whether these chewing lice harbor the *Sodalis*-allied endosymbionts in common with *C. columbae* or other unrelated symbiotic bacteria.

Diagnostic PCR surveys demonstrated that the symbiont consistently exhibited high infection frequencies in natural populations of *C. columbae* worldwide (Table 1). Considering the prevalence of the symbiont infection (Table 1) and the highly developed endosymbiotic devices such as bacteriocytes and ovarial ampullae (Figs. 3 and 4), it seems likely that the symbiont does play some important biological roles for the host insect. In the grain weevils *Sitophilus* spp., the primary symbiont contributes to growth and fecundity of the host insects (19). In the tsetse flies *Glossina* spp., biological roles of the secondary symbiont *S. glossinidius* have been obscured by the presence of the primary symbiont *W. glossinidia* (41, 44). In *C. columbae*, although biological roles of the symbiont are currently unknown, the well-developed endosymbiotic system might be relevant to the feather-feeding lifestyle and physiology of the insect. The main component of feather is keratin, a hard protein resistant to

316 solubilization, proteolysis and digestion (21, 43). Hence, the symbiont might possibly be
317 involved in keratin digestion, although the non-intestinal localization of the symbiont (Figs. 3
318 and 4) is not favorable to the hypothesis. Alternatively, the symbiont might contribute to the
319 host insect nutritionally. Amino acid composition of keratin is conspicuously biased (18, 27),
320 and the symbiont might compensate for this bias. Feather might also be devoid of vitamins
321 and other trace nutrients, which could be supplied by the symbiont. We are currently carrying
322 out physiological studies using symbiotic and aposymbiotic insects, and genomic studies of
323 the symbiont to provide insights into these biological aspects of the endosymbiosis in the
324 chewing louse.

325 Of 113 individuals examined in this study, only 4 insects were diagnosed as negative of
326 the symbiont (Table 1). The diagnostic PCR results of sexed individuals (Table 1) and the
327 FISH results of adult males (Fig. 3) suggest that these symbiont-free insects are probably
328 males. Male-specific absence of symbiont infection has been reported from aphids, coccids
329 and other insects (8, 16), which might be relevant to the fact that males do not contribute to
330 vertical transmission of the symbiont to the next generation. It should be noted, however, that
331 the samples of *C. columbae* are field-collected and thus may contain old insects and unhealthy
332 insects, from which the symbiont infection could be accidentally lost irrespective of their sex.

333 Exceptionally among insect endosymbionts, the tsetse symbiont *S. glossinidius* is
334 culturable in cell-free media (10), making the bacterium a unique model for studies of
335 insect-microbe symbiosis (31). For example, it was experimentally demonstrated that, like
336 many pathogenic bacteria, the symbiont recruits the type III secretion system for invasion to
337 the host cells (11). In addition, the ability to maintain *S. glossinidius* in pure culture greatly
338 facilitated genome sequencing of the symbiont (41). Considering the phylogenetic affinity to

the tsetse symbiont, the symbiont of *C. columbae* would, in addition to the primary symbiont of the grain weevils, provide further insights into how endosymbiotic associations could evolve from parasitism, through commensalism, and ultimately toward mutualism. At present, attempts are underway to culture the *C. columbae* endosymbiont using the techniques established previously for *S. glossinidius*.

ACKNOWLEDGMENTS

We thank Takeshi Wada and Nobutaka Urano for their help with sampling of *C. columbae*. The authors are supported by NSF awards DEB-0614565 (to D.C. and C.D.) and EF-58501127 (to C.D.)

REFERENCES

1. Akman, L., A. Yamashita, H. Watanabe, K. Oshima, T. Shiba, M. Hattori, and S. Aksoy. 2002. Genome sequence of the endocellular obligate symbiont of tsetse flies, *Wigglesworthia glossinidia*. Nat. Genet. 32: 402-407.

2. Allen, J. M., D. L. Reed, M. A. Perotti, and H. R. Braig. 2007. Evolutionary relationships of "*Candidatus* Riesia spp.," endosymbiotic enterobacteriaceae living within hematophagous primate lice. Appl. Environ. Microbiol. 73:1659-1664.

3. Barker, S. C., M. Whiting, K. P. Johnson, and A. Murrell. 2003. Phylogeny of the lice (Insecta, Phthiraptera) inferred from small subunit rRNA. Zool. Scripta 32: 407-414.

4. Bourtzis, K., and T. A. Miller. 2003. Insect symbiosis. CRC press, Boca Raton, F. L.

5. Bourtzis, K., and T. A. Miller. 2006. Insect symbiosis II. CRC press, Boca Raton, F. L.

362 6. Braendle, C., T. Miura, R. Bickel, A. W. Shingleton, S. Kambhampati, and D. L. Stern.
363 2003. Developmental origin and evolution of bacteriocytes in the aphid-*Buchnera* symbiosis.
364 PLoS Biol. 1:70-76.
365 7. Breznak, J. A., and A. Brune. 1994. Role of microorganisms in the digestion of
366 lignocellulose by termites. Ann. Rev. Entomol. 39: 453-487.
367 8. Buchner, P. 1965. Endosymbiosis of animals with plant microorganisms. Interscience, New
368 York, N.Y.
369 9. Charles, H., G. Condemine, C. Nardon, and P. Nardon. 1997. Genome size characterization
370 of the principal endocellular symbiotic bacteria of the weevil *Sitophilus oryzae*, using pulsed
371 field gel electrophoresis. Insect Biochem. Mol. Biol. 27: 345-350.
372 10. Dale, C., and I. Maudlin 1999. *Sodalis* gen. nov. and *Sodalis glossinidius* sp. nov., a
373 microaerophilic secondary endosymbiont of the tsetse fly *Glossina morsitans morsitans*. Int. J.
374 Syst. Bacteriol. 49: 267-275.
375 11. Dale, C., G. R. Plague, B. Wang, H. Ochman, and N. A. Moran. 2002. Type III secretion
376 systems and the evolution of mutualistic endosymbiosis. Proc. Natl. Acad. Sci. USA 99:
377 12397-12402.
378 12. Douglas, A. E. 1998. Nutritional interactions in insect-microbial symbioses: aphids and
379 their symbiotic bacteria *Buchnera*. Ann. Rev. Entomol. 43:17-37.
380 13. Eberle, M. W., and D. L. McLean. 1983. Observation of symbiote migration in human
381 body lice with scanning and transmission electron microscopy. Can. J. Microbiol. 29:755-762.
382 14. Frydman, H. M., J. M. Li, D. N. Robson, and E. Wieschaus. 2006. Somatic stem cell niche
383 tropism in *Wolbachia*. Nature 441: 509-512.
384 15. Fukatsu, T. 1999. Acetone preservation: a practical technique for molecular analysis. Mol.

385 Ecol. 8: 1935-1945.

386 16. Fukatsu, T. and H. Ishikawa 1992. Soldier and male of an eusocial aphid *Colophina arma*
387 lack endosymbiont: implications for physiological and evolutionary interaction between host
388 and symbiont. J. Insect Physiol. 38: 1033-1042.

389 17. Fukatsu, T, and N. Nikoh. 1998. Two intracellular symbiotic bacteria from the mulberry
390 psyllid *Anomoneura mori* (Insecta, Homoptera). Appl. Environ. Microbiol. 64: 3599-3606.

391 18. Gillespie, J. M., and M. J. Frenkel. 1974. The diversity of keratins. Comp. Biochem.
392 Physiol. 47B: 339-346.

393 19. Heddi, A, and P. Nardon. 2005. *Sitophilus oryzae* L.: a model for intracellular symbiosis
394 in the Dryophthoridae weevils (Coleoptera). Symbiosis 39: 1-11.

395 20. Hosokawa, T., Y. Kikuchi, N. Nikoh, M. Shimada, and T. Fukatsu. 2006. Strict
396 host-symbiont co-speciation and reductive genome evolution in insect gut bacteria. PLoS Biol.
397 4: e337.

398 21. Hughes, J., and A. P. Vogler. 2006. Gene expression in the gut of keratin-feeding clothes
399 moths (Tineola) and keratin beetles (Trox) revealed by subtracted cDNA libraries. Insect
400 Biochem. Mol. Biol. 36: 584-592.

401 22.Johnson, K. P., K. Yoshizawa, and V. S. Smith. 2004. Multiple origins of parasitism in lice.
402 Proc. R. Soc. Lond. B 271: 1771-1776.

403 23. Kimura, M. 1980. A simple method for estimating evolutionary rates of base substitutions
404 through comparative studies of nucleotide sequences. J. Mol. Evol. 16:111-120.

405 24. Kumar, S., K. Tamura, and M. Nei. 2004. MEGA3: integrated software for molecular
406 evolutionary genetics analysis and sequence alignment. Briefings in Bioinformatics 5:
407 150-163.

25. Kyei-Poku, G. K., D. D. Colwell, P. Coghlin, B. Benkel, and K. D. Floate 2005. On the ubiquity and phylogeny of *Wolbachia* in lice. Mol. Ecol. 14: 285-294.

26. Lefevre, C, H. Charles, A. Vallier, B. Delobel, B. Farrell, and A. Heddi. 2004. Endosymbiont phylogenesis in the Dryophthoridae weevils: Evidence for bacterial replacement. Mol. Biol. Evol: 21: 965-973.

27. Marshall, R. C., D. F. G. Orwin, and J. M. Gillespie. 1991. Structure and biochemistry of mammalian hard keratin. Electron Microsc. Rev. 4: 47-83.

28. Nogge, G. 1982. Significance of symbionts for the maintenance of an optimal nutritional state for successful reproduction in hematophagus arthropods. Parasitology 82: 299-304.

29. Novakova, E., and V. Hypsa. 2007. A new *Sodalis* lineage from bloodsucking fly *Craterina melbae* (Diptera, Hippoboscoidea) originated independently of the tsetse flies symbiont *Sodalis glossinidius*. FEMS Microbiol Lett. 269: 131-135.

30. Ohkuma, M. 2003. Termite symbiotic systems: efficient bio-recycling of lignocellulose. Appl. Microbiol. Biotechnol. 61: 1-9.

31. Pontes, M. H., and C. Dale 2006. Culture and manipulation of insect facultative symbionts. Trends Microbiol. 14: 406-412.

32. Price, R. D., R. A. Hellenthal, R. L. Palma, K. P. Johnson, and D. H. Clayton. 2003. The chewing lice: World checklist and biological overview. Illinois Natural History Survey Special Publication 24 (x + 501 pp.), Champaign, IL.

33. Reed, D. L., and M. S. Hafner. 2002. Phylogenetic analysis of bacterial communities associated with ectoparasitic chewing lice of pocket gophers: a culture-independent approach. Microbial Ecol. 44: 78-93.

34. Ries, E. 1931. Die Symbiose der Laüse und Federlinge. Z. Morphol. Ökol. Tiere

20:233-367.

35. Robinson-Rechavi, M., and D. Huchon. 2000. RRTree: Relative-rate tests between groups of sequences on a phylogenetic tree. Bioinformatics 16:296-297.

36. Sakurai, M., R. Koga, T. Tsuchida, X.-Y. Meng, and T. Fukatsu. 2005. *Rickettsia* symbiont in the pea aphid *Acyrthosiphon pisum*: novel cellular tropism, effect on host fitness, and interaction with the essential symbiont *Buchnera*. Appl. Environ. Microbiol. 71:4069–4075.

37. Sasaki-Fukatsu, K., R. Koga, N. Nikoh, K. Yoshizawa, S. Kasai, M. Mihara, M. Kobayashi, T. Tomita, and T. Fukatsu. 2006. Symbiotic bacteria associated with stomach discs of human lice. Appl. Environ. Microbiol. 72: 7349-7352.

38. Schmidt, H.A., K. Strimmer, M. Vingron, and A. Haeseler. 2002. TREE-PUZZLE: maximum likelihood phylogenetic analysis using quartets and parallel computing. Bioinformatics 18: 502-504.

39. Shigenobu, S., H. Watanabe, M. Hattori, Y. Sakaki, and H. Ishikawa. 2000. Genome sequence of the endocellular bacterial symbiont of aphids *Buchnera* sp APS. Nature 407: 81-86.

40. Thompson, J. D., D. G. Higgins, and J. J. Gibson. 1994. Clustal W: improving the sensitivity of progressive multiple alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res. 22:4673-4680.

41. Toh, H., B. L. Weiss, S. A. H. Perkin, A. Yamashita, K. Oshima, M. Hattori, and S. Aksoy. 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.

42. Veneti, Z., M. E. Clark, T. L. Karr, C. Savakis, and K. Bourtzis. 2004. Heads or tails: Host-parasite interactions in the *Drosophila-Wolbachia* system. Appl. Environ. Microbiol. 70:

5366-5372.

43. Waterhouse, D. F. 1957. Digestion in insects. Ann. Rev. Entomol. 2: 1-18.

44. Weiss, B. L., R. Mouchotte, R. V. M. Rio, Y. N. Wu, Z. Y. Wu, A. Heddi, and S. Aksoy. 2006. Interspecific transfer of bacterial endosymbionts between tsetse fly species: Infection establishment and effect on host fitness. Appl. Environ. Microbiol. 72: 7013-7021.

45.Wernegreen, J. J. 2002. Genome evolution in bacterial endosymbionts of insects. Nat. Rev. Genet. 3: 850-861.

46. Yoshizawa, K., and K. P. Johnson. 2006. Morphology of male genitalia in lice and their relatives and phylogenetic implications. Syst. Entomol. 31: 350-361.

FIGURE CAPTIONS

FIG. 1. Molecular phylogenetic analysis on the basis of *16S rRNA* gene sequences of the symbiont of *C. columbae* and allied γ-proteobacteria. A NJ tree inferred from 1,363 unambiguously aligned nucleotide sites is shown; MP and ML analyses gave essentially the same results (data not shown). Statistical support values higher than 70% are indicated at the nodes in the order of NJ/MP/ML. Sequence accession numbers are shown in brackets. AT contents of the sequences are shown in parentheses.

FIG. 2. Molecular phylogenetic analysis on the basis of *fusA* gene sequences of the symbiont of *C. columbae* and allied γ-proteobacteria. A NJ tree inferred from 499 unambiguously aligned nucleotide sites is shown; MP and ML analyses gave essentially the same results (data not shown). Statistical support values higher than 70% are indicated at the nodes in the order

477 of NJ/MP/ML. Sequence accession numbers are shown in brackets.

478

479 **FIG. 3**. General localization of the symbiont in nymphs and adults of *C. columbae*. (A) 1st
480 instar nymph; (B), (C) 2nd instar nymphs; (D)-(G) 3rd instar nymphs; (H), (I) Female adults;
481 (J), (K) male adults. Red and green signals indicate symbiont cells and host nuclei,
482 respectively, while cuticles often exhibit red signals due to autofluorescence. Arrowheads,
483 hybridization signals of bacteriocytes in the body cavity; arrows, hybridization signals
484 associated with lateral oviducts, ovarial ampullae and/or oocytes.

485

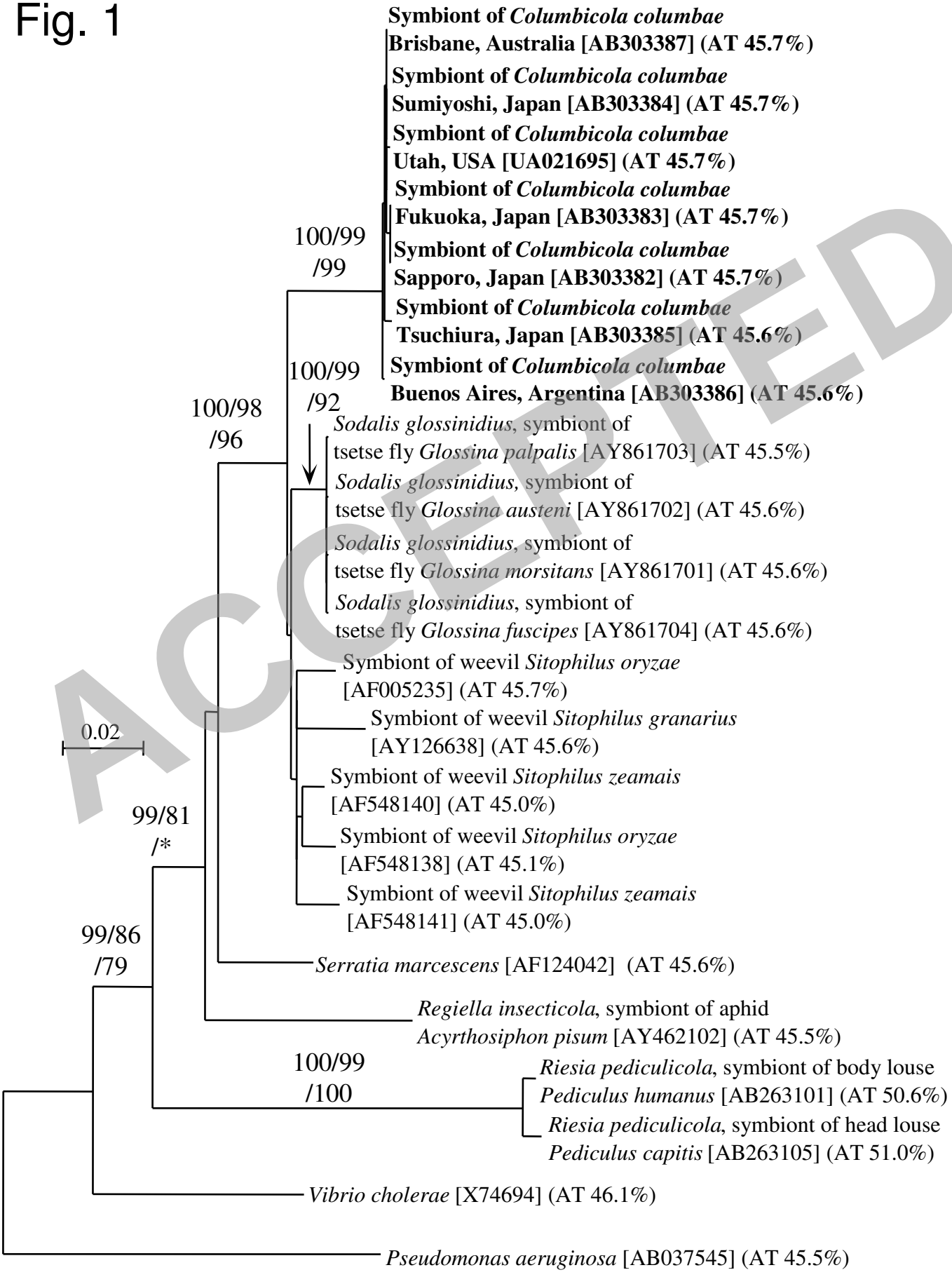
486 **FIG. 4**. Localization of the symbiont in nymphs and adults of *C. columbae*. (A) An aggregate
487 of bacteriocytes in the lateral body cavity of a 1st instar nymph. Cytoplasm of oval
488 bacteriocytes is full of rod-shaped symbiont cells. (B) An aggregate of bacteriocytes in the
489 lateral body cavity of a 2nd instar nymph. The number of bacteriocytes increases, and the area
490 occupied by the aggregate stretches longitudinally. (C) Distribution of bacteriocytes beneath
491 the hypodermis of the abdominal segment in a 1st instar nymph. (D) Z-axis reconstruction of
492 distribution of bacteriocytes through the plane D shown in C. (E) Migration process of
493 bacteriocytes in a 3rd instar female nymph. (F), (G) Enlarged images of the migrating
494 bacteriocytes. (H) Localization of bacteriocytes in lateral oviducts and formation of ovarial
495 ampullae in a 3rd instar female nymph. (I) An enlarged image of an ovarial ampulla of the 3rd
496 instar female nymph. (J) An enlarged image of an ovarial ampulla of a female adult. (K) An
497 ovarial ampulla associated with a mature oocyte in an ovariole. (L) Infection process of
498 symbiont cells from an ovarial ampulla to the posterior pole of an oocyte through follicle cells.
499 (M) Symbiont cells localized in the posterior region of an oocyte. Abbreviations: bc,

500 bacteriocytes; cu, cuticle; fc, follicle cells; mo, mature oocyte; oa, ovarial ampullae; yo,

501 young oocyte.

502

Fig. 1



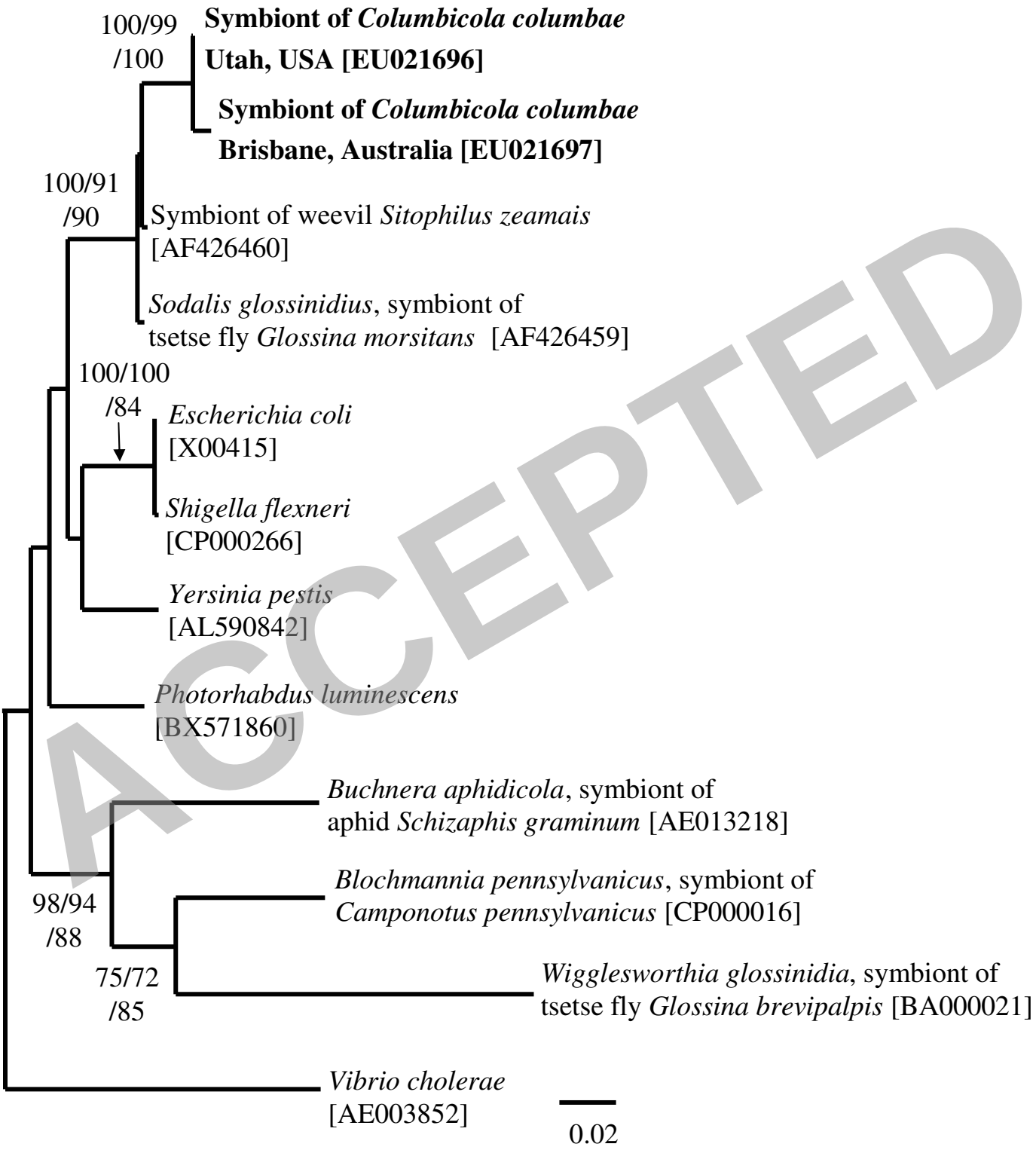
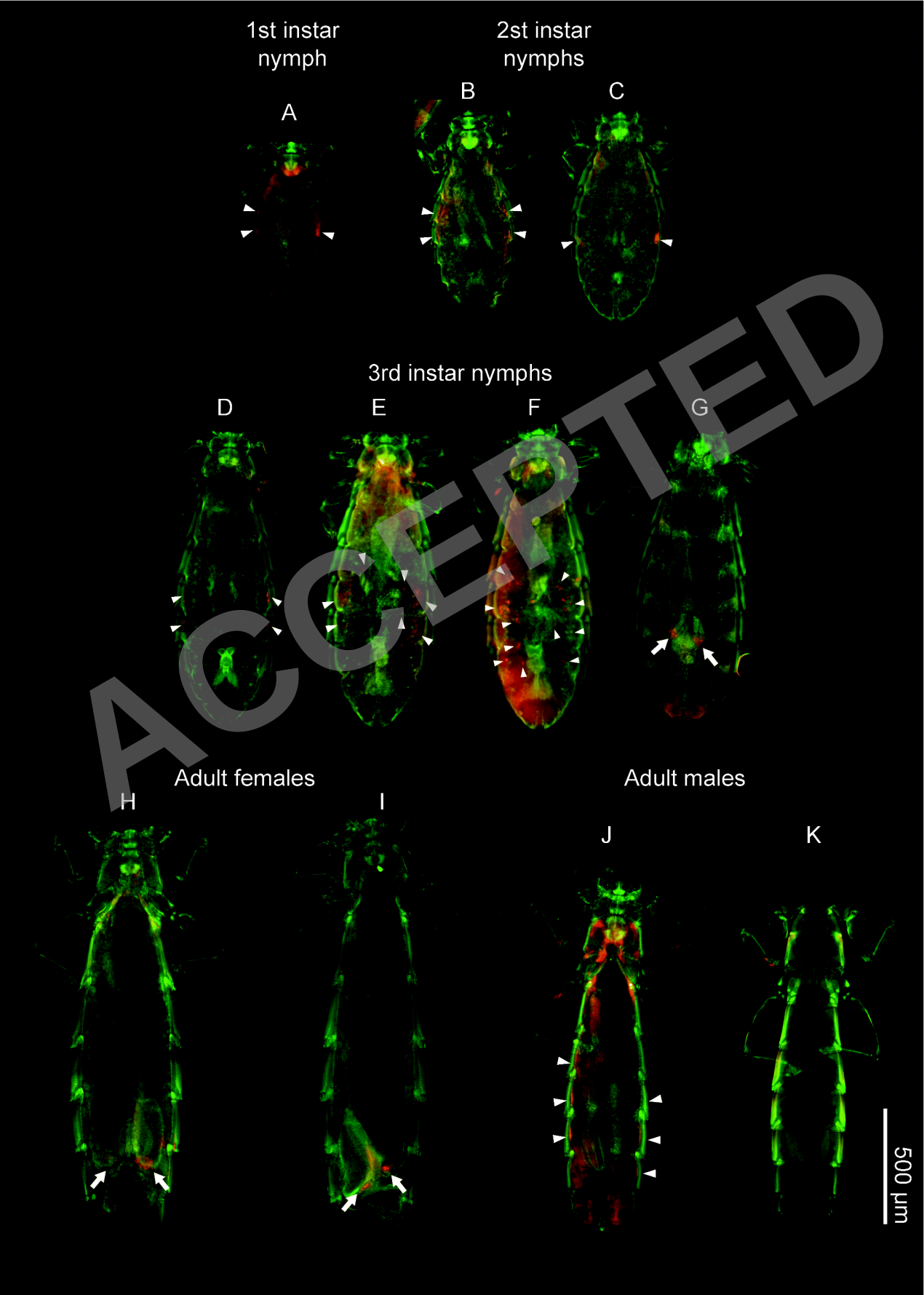


Fig. 2


1st instar nymph
2st instar nymphs
A
B
C
3rd instar nymphs
D
E
F
G
Adult females
Adult males
H
I
J
K
500 µm

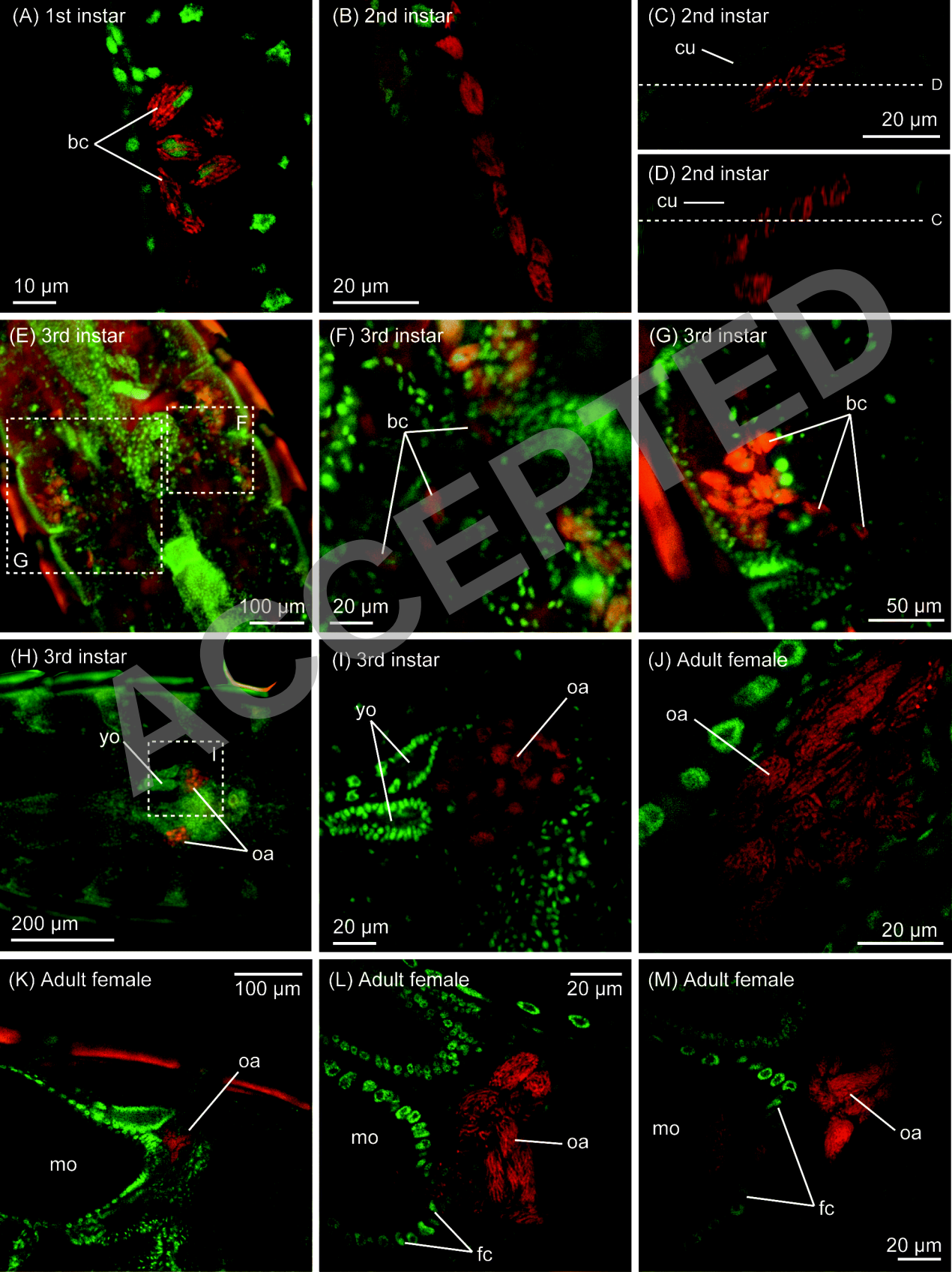

(A) 1st instar
bc
10 µm
(B) 2nd instar
20 µm
(C) 2nd instar
cu
D
20 µm
(D) 2nd instar
cu
C
(E) 3rd instar
F
G
100 µm
(F) 3rd instar
bc
20 µm
(G) 3rd instar
bc
50 µm
(H) 3rd instar
yo
I
oa
200 µm
(I) 3rd instar
oa
yo
20 µm
(J) Adult female
oa
20 µm
(K) Adult female
100 µm
oa
mo
(L) Adult female
20 µm
mo
oa
fc
(M) Adult female
mo
oa
fc
20 µm

Table 1. Samples of *C. columbae* examined in this study, results of diagnostic PCR detection, and sequence accession numbers.

Sample code	Collection locality	Collection date & collector[1]	Infection frequency[2]				*16S rRNA* gene accession no.	*fusA* gene accession no.
			No sexing[3]	Female	Male	Total		
FKK99	Ropponmatsu, Fukuoka,	27 Aug. 1999	-	-	-	-	AB303383	
	Japan	KY	-	-	-	-		
SPR06	Sapporo, Hokkaido,	16 Aug. 2006	100%	100%	100%	100%	AB303382	
	Japan	KY	(10/10)	(10/10)	(10/10)	(30/30)		
SMY06	Sumiyoshi, Osaka,	19 Oct. 2006	90%	-	-	90%	AB303384	
	Japan	TW	(9/10)	-	-	(9/10)		
TTR06	Tsuchiura, Ibaraki,	10 Nov. 2006	100%	-	-	100%	AB303385	
	Japan	KSF & TF	(10/10)	-	-	(10/10)		
NNW07	Naniwa, Osaka,	5 Mar. 2007	-	100%	95%	97.5%		
	Japan	NU	-	(20/20)	(19/20)	(39/40)		
BNS06	Buenos Aires,	19 Oct. 2006	100%	-	-	100%	AB303386	
	Argentina	TF	(10/10)	-	-	(10/10)		
UTH98	Utah,	29 Jun. 1998	100%	-	-	100%	UA021695	
	USA	DC	(3/3)	-	-	(3/3)		
UTH99	Utah,	1999	-	-	-	-		UA021696
	USA	DC	-	-	-	-		
BRB02	Brisbane	2002	-	-	-	-		UA021697
	Australia	DC	-	-	-	-		
BRB07	Brisbane,	16 Feb. 2007	80%	-	-	80%	AB303387	
	Australia	TF	(8/10)	-	-	(8/10)		
Total			94.3%	100%	96.7%	96.5%		

(50/53) (30/30) (29/30) (109/113)

[1]DC, Dale Clayton; KSF, Kayoko Sasaki-Fukatsu; KY, Kazunori Yoshizawa; NU, Nobutaka Urano; TF, Takema Fukatsu; TW, Takeshi Wada.

[2]For example, 90% (9/10) means 90% infection, with 9 infected insects per 10 insects examined by diagnostic PCR.

[3]Including female adults, male adults and nymphs.

Table 2. Relative-rate tests for comparing the molecular evolutionary rate of *16S rRNA* gene and *fusA* gene between the symbionts of *C. columbae*, the symbionts of tsetse flies, and the symbionts of grain weevils.

Gene	Lineage1	Lineage2	Outgroup	K1[1]	K2[2]	Difference of distance[3]	Rate ratio[4]	*P*-value[5]
(A) *16SrRNA* gene								
	Symbionts of *C. columbae*[6]	Symbionts of tsetse flies[7]	*S. marcescens*[8]	0.027	0.009	0.018	3.1	0.00059
	Symbionts of *C. columbae*[6]	Symbionts of grain weevils [9]	*S. marcescens*[8]	0.026	0.010	0.016	2.7	0.00092
(B) *fusA* gene								
	Symbiont of *C. columbae* [10]	Symbiont of tsetse fly[11]	*E. coli*[12]	0.025	0.001	0.024	25.0	0.0020
	Symbiont of *C. columbae* [10]	Symbiont of grain weevil[13]	*E. coli* [12]	0.022	0.001	0.021	22.0	0.0034

[1]Estimated mean distance between lineage 1 and the last common ancestor of lineages 1 and 2.

[2]Estimated mean distance between lineage 2 and the last common ancestor of lineages 1 and 2.

[3]K1-K2.

[4]K1/K2.

[5]*P*-value was generated using the program RRTree (35).

[6]Symbionts of *C. columbae* from Sapporo, Japan (AB303382) and Buenos Aires, Argentina (AB303386).

[7]*Sodalis glossinidius,* symbionts of tsetse flies *G. morsitans* (AY861701) and *G. palpalis* (AY861703).

[8]*Serratia marcescens* (AF124042).

[9] Symbionts of grain weevils *S. oryzae* (AF005235) and *S. zeamais* (AF005235).

[10] Symbionts of *C. columbae* from Utah, USA (UA021696) and from Brisbane, Australia (UA021697).

[11] *Sodalis glossinidius*, symbiont of tsetse fly *G. morsitans* (AF426459).

[12]*Escherichia coli* (X00415).

[13] Symbiont of grain weevil *S. zeamais* (AF426460).