ACADEMIC PRESS

Available online at www.sciencedirect.com

SCIENCE DIRECT®

Molecular Phylogenetics and Evolution 29 (2003) 102–114

MOLECULAR PHYLOGENETICS AND EVOLUTION

www.elsevier.com/locate/ympev

Phylogenetic position of Phthiraptera (Insecta: Paraneoptera) and elevated rate of evolution in mitochondrial 12S and 16S rDNA

Kazunori Yoshizawa[a,*] and Kevin P. Johnson[b]

[a] *Systematic Entomology, Department of Ecology and Systematics, Graduate School of Agriculture, Hokkaido University, Sapporo 060-8589, Japan*
[b] *Illinois Natural History Survey, 607 East Peabody Drive, Champaign, IL 61820-6970, USA*

Received 21 August 2002; revised 16 February 2003

Abstract

Phthiraptera (chewing and sucking lice) and Psocoptera (booklice and barklice) are closely related to each other and compose the monophyletic taxon Psocodea. However, there are two hypotheses regarding their phylogenetic relationship: (1) monophyletic Psocoptera is the sister group of Phthiraptera or (2) Psocoptera is paraphyletic, and Liposcelididae of Psocoptera is the sister group of Phthiraptera. Each hypothesis is supported morphologically and/or embryologically, and this problem has not yet been resolved. In the present study, the phylogenetic position of Phthiraptera was examined using mitochondrial 12S and 16S rDNA sequences, with three methods of phylogenetic analysis. Results of all analyses strongly supported the close relationship between Phthiraptera and Liposcelididae. Results of the present analyses also provided some insight into the elevated rate of evolution in mitochondrial DNA (mtDNA) in Phthiraptera. An elevated substitution rate of mtDNA appears to originate in the common ancestor of Phthiraptera and Liposcelididae, and directly corresponds to an increased G + C content. Therefore, the elevated substitution rate of mtDNA in Phthiraptera and Liposcelididae appears to be directional. A high diversity of 12S rDNA secondary structure was also observed in wide range of Phthiraptera and Liposcelididae, but these structures seem to have evolved independently in different clades.

1. Introduction

Phthiraptera (chewing and sucking lice) and Psocoptera (booklice, barklice, and psocids) have long been recognized to be closely related to each other, and together compose the monophyletic taxon Psocodea (e.g., Kristensen, 1995). Monophyly of Psocodea is strongly supported by specialized hypopharingeal structure (Rudolph, 1982, 1983) and molecular data (Wheeler et al., 2001). In contrast, there are two alternative hypotheses about the phylogenetic position of Phthiraptera. In the traditional taxonomic system, the order Psocoptera is treated as an independent order and the order Phthiraptera is placed as its sister group. Monophyly of Psocoptera is supported by the egg structure and the embryonic orientation (Seeger, 1979). Alternatively, Lyal (1985) extensively investigated the external morphology of Phthiraptera and Psocoptera and concluded that the family Liposcelididae of Psocoptera is the sister group of Phthiraptera, making Psocoptera paraphyletic. However, almost all character states supporting sister group relationship between Phthiraptera and Liposcelididae are loss characters, and Lyal (1985) mentioned that they may have evolved independently by responding to their similar habitat.

Phthiraptera are permanent ectoparasites of birds and mammals. Background of the origin of parasitism in Phthiraptera includes many interesting evolutionary, morphological, and systematic problems (Barker, 1994; Waage, 1979). For example, was the louse ancestor commensal with vertebrates? Was there a single or multiple origin of parasitism in lice? Reliable estimation of the phylogenetic position of Phthiraptera would provide a basis to answer these questions. In the present study, we investigate the phylogenetic position of Phthiraptera based on the mitochondrial 12S and 16S rDNA sequences.

* Corresponding author. Fax: +81-11-706-4939.
E-mail address: psocid@res.agr.hokudai.ac.jp (K. Yoshizawa).

1055-7903/$ - see front matter © 2003 Elsevier Science (USA). All rights reserved.
doi:10.1016/S1055-7903(03)00073-3

Mitochondrial 12S and 16S rDNA were selected for several reasons, including the slow evolutionary rate, the utility in the phylogenetic analyses of higher insect groups, and the existence of universal insect primers and ease of reliable PCR amplification (Simon et al., 1994). Additionally, mitochondrial genomes of Phthiraptera are interesting because of their apparently elevated substitution rates (Simmons and Weller, 2001). Hafner et al. (1994) and Page et al. (1998) reported an elevated rate of substitution in louse mitochondrial genes (COI and cyt-*b*) relative to vertebrate hosts. Johnson et al. (2003) showed an elevated rate of substitution in louse mitochondrial COI gene relative to aphids. However, these studies lacked data of closest relatives of lice (Psocoptera). Therefore, it is not clear whether the phenomenon is unique to lice and when and how the phenomenon originated. The present study is based on two mitochondrial genes and includes wide range of Phthiraptera and Psocoptera, which will provide insights about the origin of the elevated rate of evolution in mitochondrial genes of lice and their relatives.

2. Materials and methods

2.1. Taxon sampling and sequence determination

Specimens stored in 99.5% ethanol were used for the study. Total DNA was extracted from 49 specimens following the methods described by Cruickshank et al. (2001). The samples include one neuropteran (root), three hemipterans (closer outgroup of Psocodea), 27 psocopterans (including six *Liposcelis*), and 18 phthirapterans (including three amblyceran, 14 ischnoceran, and one anopluran lice) (Table 1). Primer sets 12Sai + 12Sbi and 16Sar + 16Sbr (Simon et al., 1994) were used to amplify partial sequences of mitochondrial 12S and 16S rDNA, respectively. Modified 12Sbi primer (1 mer deleted from the 3′ end) was used to amplify 12S rDNA of *Bovicola*. Reaction cycle was 94 °C for 3′ followed by 40 cycles of 94 °C for 30″, 45 °C for 45″, and 72 °C for 60″. Amplified products were purified using PCR Purification Kit (Qiagen) and sequenced by CEQ2000 DNA Analysis System (Beckman Coulter) following manufacturer's protocols. Sequences of *Heterodoxus macropus* (Phthiraptera: Amblycera) and *Triatoma dimidiata* (Hemiptera: Heteroptera) were obtained from GenBank (Table 1).

2.2. Alignment and data evaluation

The difficulties of multiple sequence alignment of louse 12S rDNA were stated by Paterson et al. (2000) and Page et al. (2002). Here, we aligned sequences using Secondary Structure Mode of ClustalX (Thompson et al., 1997). A 12S rDNA alignment for lice is provided by Page et al. (2002) and a 16S rDNA alignment for insects is provided by Buckley et al. (2000). These secondary structure models were used as alignment profiles. Hickson et al. (2000) mentioned that results of 12S rDNA alignment by ClustalW are improved with small gap and gap-extension costs. Thus, we tried alignments with several pairs of Gap:Gap-extension costs (Gap costs 10, 15, and 20; Gap-extension costs 0.1, 1, 3, and 6.66). With high Gap-extension cost (e.g., 15:6.66 which is the default setting of ClustalX), genetic distances of some sequence pairs exceeded 1.0 which must be avoided. With lower Gap-extension costs, many stem regions were recovered in the alignment, and different cost sets provided very similar alignments. Here 10:1 was selected since it recovered the maximum numbers of stem regions. However, with other cost sets such as 20:0.1, results more closely matched secondary structure models for a few stem regions. Thus, such regions were manually edited based on alignments resulting from different costs to make the total alignment a closer match to secondary structure models. Hickson et al. (2000) also mentioned that different costs may be appropriate for different regions, especially where large insertions or deletions occur. A NEXUS file of the aligned sequences is available from the following URL address or by request to the first author. <http://insect3.agr.hokudai.ac.jp/psocid/data/index.html>

Each sequence was divided into two data sets (discussed below). We performed partition homogeneity test (Farris et al., 1994, 1995) to compare the homogeneity of each data set using PAUP 4.0b10 PPC (Swofford, 2002). We also compared signal of each data set by comparing positional congruence of the resulting trees (Estabrook, 1992) using RadCon (Thorley and Page, 2000).

2.3. Phylogenetic analysis

As an initial analysis, equally weighted maximum parsimonious (MP) analysis was performed with 100 TBR replication. However, G + C contents of the present data set are highly variable among taxa (from 43.9% of *Bovicola* to 19.6% of *Hemipsocus*: $\Delta GC = 24.3\%$) which can be a problem for some phylogenetic analysis methods, such as equally weighted parsimony (Galtier and Gouy, 1995). Thus, as an alternative method, neighbor-joining (NJ) analysis was performed using LogDet distances, which can correct the nonstationarity of base composition (Lockhart et al., 1994). The proportion of invariant sites (0.1417) was estimated using maximum-likelihood.

We also performed maximum-likelihood (ML) analysis, which is relatively robust to the nonstationary data (Galtier and Gouy, 1995). Parameters for ML analysis were estimated using Modeltest 3.06 (Posada and Crandall, 1998). As a result of Modeltest, the GTR + I + G

Table 1
Specimens included in study

No.	Species	Host	Order (Suborder:Family)	GenBank Accession Nos.	
				12S rDNA	16S rDNA
1.	Micromus sp. (Japan)	———	Neuroptera (Planipennia:Hemerobidae)	AY139924	AY139971
2.	Meimuna opalifera	———	Hemiptera (Cicadomorpha:Cicadidae)	AY139921	AY139968
3.	Olarius sp. (Japan)	———	Hemiptera (Fulgolomorpha:Cixiidae)	AY139922	AY139969
4.	Physopelta cincticollis	———	Hemiptera (Heteroptera:Largidae)	AY139923	AY139970
5.	Triatoma dimidiata	———	Hemiptera (Heteroptera:Reduviidae)	AF301594 (Dotson and Beard, 2001)	AF301594 (Dotson and Beard, 2001)
6.	Echmepteryx hageni	———	Psocoptera (Trogiomorpha:Lepidopsocidae)	AY139916	AY139963
7.	Echmepteryx madagascarensis	———	Psocoptera (Trogiomorpha:Lepidopsocidae)	AY139915	AY139962
8.	Neolepolepis occidentalis	———	Psocoptera (Trogiomorpha:Lepidopsocidae)	AY139917	AY139964
9.	Cerobasis guestfalica	———	Psocoptera (Trogiomorpha:Trogidae)	AY139918	AY139965
10.	Lepinotus reticulatus	———	Psocoptera (Trogiomorpha:Trogidae)	AY139920	AY139967
11.	Dorypteryx domestica	———	Psocoptera (Trogiomorpha:Psyllipsocidae)	AY139919	AY139966
12.	Compsocus elegans	———	Psocoptera (Troctomorpha:Compsocidae)	AY139914	AY139961
13.	Stimulopalpus japonicus	———	Psocoptera (Troctomorpha:Amphientomidae)	AY139913	AY139960
14.	Musapsocus sp.	———	Psocoptera (Troctomorpha:Musapsocidae)	AY245855	AY245857
15.	Tapinella sp. (Costa Rica)	———	Psocoptera (Troctomorpha:Pachytroctidae)	AY139902	AY139949
16.	Liposcelis bostrychophila	———	Psocoptera (Troctomorpha:Liposcelididae)	AY139897	AY139944
17.	Liposcelis paeta	———	Psocoptera (Troctomorpha:Liposcelididae)	AY139899	AY139946
18.	Liposcelis decolor (USA)	———	Psocoptera (Troctomorpha:Liposcelididae)	AY139898	AY139945
19.	Liposcelis decolor (EUR)	———	Psocoptera (Troctomorpha:Liposcelididae)	AY139901	AY139948
20.	Liposcelis brunnei	———	Psocoptera (Troctomorpha:Liposcelididae)	AY139900	AY139947
21.	Liposcelis sp.	———	Psocoptera (Troctomorpha:Liposcelididae)	AY245854	AY245856
22.	Archipsocus recens	———	Psocoptera (Psocomorpha:Archipsocidae)	AY139892	AY139939
23.	Hemipsocus chloroticus	———	Psocoptera (Psocomorpha:Hemipsocidae)	AY139910	AY139957
24.	Lichenomima muscosa	———	Psocoptera (Psocomorpha:Myopsocidae)	AY139908	AY139955
25.	Psococerastis nubila	———	Psocoptera (Psocomorpha:Psocidae)	AY139905	AY139952
26.	Loensia fasciata	———	Psocoptera (Psocomorpha:Psocidae)	AY139906	AY139953
27.	Ptycta sp. (Japan)	———	Psocoptera (Psocomorpha:Psocidae)	AY139907	AY139954
28.	Idatenopsocus orientalis	———	Psocoptera (Psocomorpha:Mesopsocidae)	AY139909	AY139956
29.	Haplophalus sp. (Japan)	———	Psocoptera (Psocomorpha:Philotarsidae)	AY139912	AY139959
30.	Epipsocopsis sp. (Taiwan)	———	Psocoptera (Psocomorpha:Epipsocidae)	AY139911	AY139958
31.	Xanthocaecilius sommermanae	———	Psocoptera (Psocomorpha:Caeciliusidae)	AY139903	AY139950
32.	Stenopsocus aphidoformes	———	Psocoptera (Psocomorpha:Stenopsocidae)	AY139904	AY139951
33.	Heterodoxus macropus	Macropus agilis	Phthiraptera (Amblycera:Boopidae)	AF270939 (Shao et al., 2001a,b)	AF270939 (Shao et al., 2001a,b)
34.	Trinoton querquedulae	Anas penelope	Phthiraptera (Amblycera:Menoponidae)	AY139896	AY139943
35.	Myrsidea ishizawai	Zoothera dauma	Phthiraptera (Amblycera:Menoponidae)	AY139895	AY139942
36.	Austromenopon transversum	Larus crassicornis	Phthiraptera (Amblycera:Menoponidae)	AY139894	AY139941

37. Trichodectes canis	Nyctereutes procyonoidides	Phthiraptera (Ischnocera:Trichodectidae)	AY139878	AY139925
38. Trichodectes pinguis	Ursus thibetanus	Phthiraptera (Ischnocera:Trichodectidae)	AY139879	AY139926
39. Bovicola tibialis	Cervus nippon	Phthiraptera (Ischnocera:Trichodectidae)	AY139880	AY139927
40. Columbicola columbae	Columba livia	Phthiraptera (Ischnocera:Philopteridae)	AY139882	AY139929
41. Columbicola sp.	Treron formosae	Phthiraptera (Ischnocera:Philopteridae)	AY139883	AY139930
42. Saemundssonia lari	Larus crassirostris	Phthiraptera (Ischnocera:Philopteridae)	AY139884	AY139931
43. Quadraceps sp.	Larus crassirostris	Phthiraptera (Ischnocera:Philopteridae)	AY139885	AY139932
44. Strigiphilus heterocerus	Strix uralensis	Phthiraptera (Ischnocera:Philopteridae)	AY139886	AY139933
45. Auricotes rotunddus	Ptilinopus occipitalis	Phthiraptera (Ischnocera:Philopteridae)	AY139890	AY139937
46. Campanulotes compar	Columba livia	Phthiraptera (Ischnocera:Philopteridae)	AY139887	AY139934
47. Acidoproctus emersoni	Dendrocygna eytoni	Phthiraptera (Ischnocera:Philopteridae)	AY139888	AY139935
48. Ornithobius cygni	Cygnus cygnus	Phthiraptera (Ischnocera:Philopteridae)	AY139893	AY139940
49. Anatoecus sp.	Anas penelope	Phthiraptera (Ischnocera:Philopteridae)	AY139891	AY139938
50. Aquanirmus sp.	Podoceps poliocephalus	Phthiraptera (Ischnocera:Philopteridae)	AY139889	AY139936
51. Pediculus humanus	Homo sapiens	Phrhiraptera (Anoplura:Pediculidae)	AY139881	AY139928

model was selected (unequal base frequencies: A = 0.3567, C = 0.1082, G = 0.1747, T = 0.3604; six substitution categories: A–C = 1.0223, A–G = 3.5723, A–T = 2.0615, C–G = 0.4125, C–T = 4.7321, G–T = 1.0000; gamma distributions shape parameter = 0.9598; proportion of invariant sites = 0.1417; four rate categories). We searched for the most likely tree by TBR branch swapping using NJ tree as starting point. Bootstrap supports for ML tree were calculated using 100 replicates with NNI branch swapping. All analyses were performed using PAUP 4.0b10 PPC (Swofford, 2002), with gaps treated as missing.

3. Results

3.1. Alignment and data evaluation

Using the Secondary Structure Mode of ClustalX based on Gap:Gap-extension costs = 10:1, most of the stem regions of 12S and 16S rDNA were well aligned and correspond to secondary structure models. Page et al. (2002) mentioned that some stem and loop regions of louse 12S rDNA (39, 42, and 47) are highly variable in structure. In the present alignment, long insertion/deletion (indels) were detected in these regions, and alignment that correspond to secondary structure models for these regions could not be obtained. Similarly, long indels (more than 20 bp) were detected in some regions of 16S rDNA (75, 81, 84, and 88), and the resulted alignment for this regions do not correspond to secondary structure models. Page et al. (2002) showed that highly variable loop and stem regions of louse 12S rDNA are phylogenetically less informative. Thus, we divided the data into the following four partitions and examined their behavior: well-aligned regions (12S and 16S) (308 and 425 bp); highly variable regions having long indels (12S-indel and 16S-indel) (183 and 152 bp).

We compared the homogeneity of the signal from each data set by partition homogeneity tests (Farris et al., 1994, 1995) with 1000 replicates. Significant heterogeneity was detected in all comparisons ($P < 0.01$). However, the trees estimated from 12S and 16S were very similar (trees not shown). Thus, we then compared the Positional Congruence (Estabrook, 1992) between NJ trees resulting from 12S and 16S data sets, and the comparison indicated a high topological congruence (86.2%). Again, trees estimated from 12S-indel and 16S-indel were not congruent with trees derived from 12S, 16S nor 12S + 16S (<50%). These results supported combination of 12S and 16S data sets while excluding the highly variable regions. The combined well-aligned data set includes 733 bp. The followings are average ML pairwise distances from outgroup to each taxon: 0.465 (Hemiptera); 0.455 (Psocoptera excl. Liposcelididae and *Tapinella*); 0.921 (Liposcelididae); and 1.007 (Phthiraptera).

Although the above examinations supported exclusion of highly variable regions, we also retained a data matrix including all data sets for phylogenetic analyses to assess the results of including these regions. This data matrix includes 1068 bp.

3.2. Phylogenetic analyses

First we analyzed the data set excluding highly variable regions because of possible noise included in these regions. All analyses yielded very similar topologies

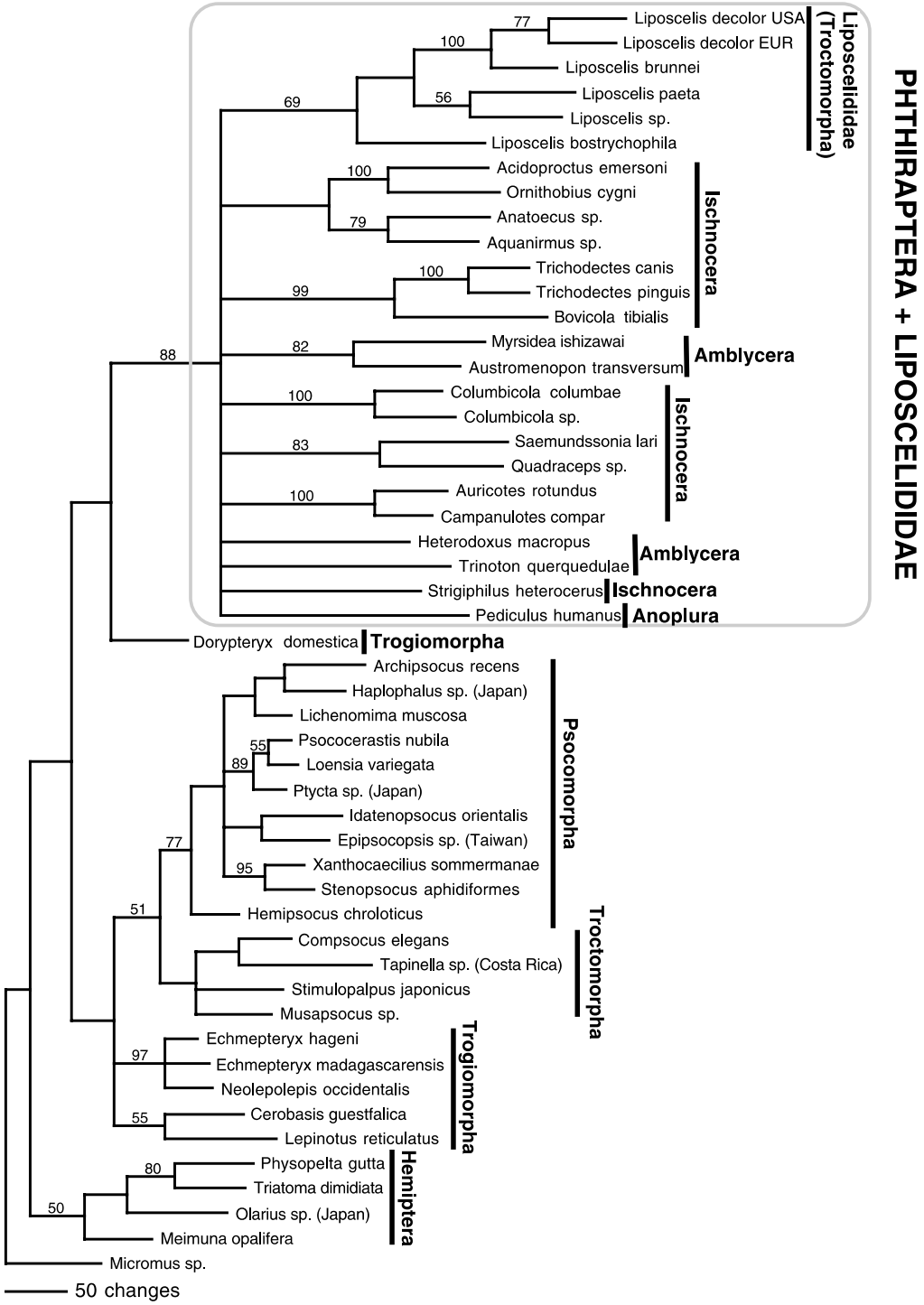

Fig. 1. Strict consensus of nine equally parsimonious trees estimated from the data set excluding highly variable regions (length = 5387; CI = 0.245). Branch lengths are proportional to reconstructed changes. Numbers next to branches indicate nodes supported in >50% of bootstrap replicates.

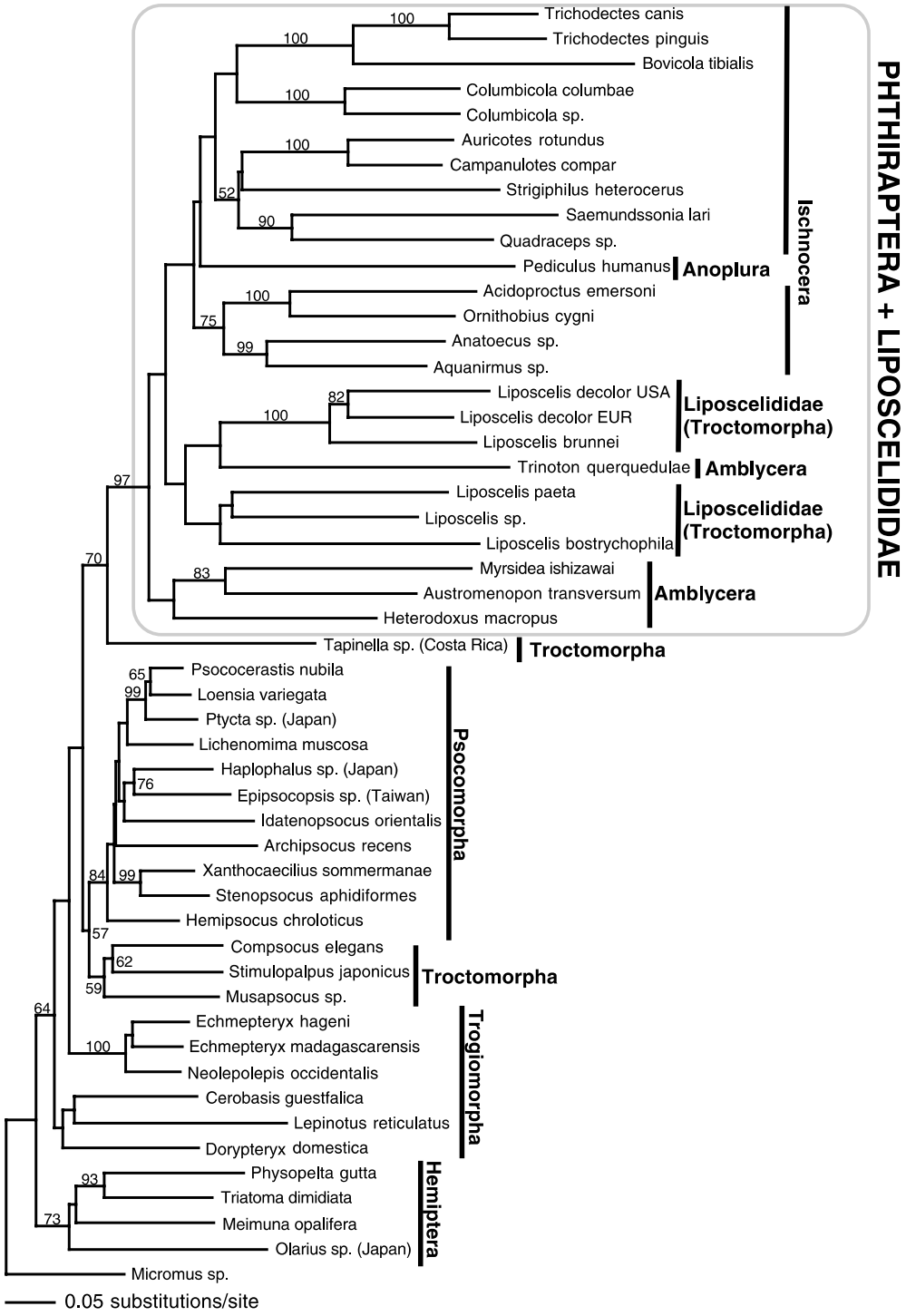

Fig. 2. Neighbor-joining tree estimated from the data set excluding highly variable regions. Branch lengths are proportional to reconstructed distances. Numbers next to branches indicate nodes supported in >50% of bootstrap replicates.

(Figs. 1–3). Monophyly of Phthiraptera + Liposcelididae was strongly supported by all analyses (bootstrap values: MP 88%; NJ 97%: ML 86%). In addition to bootstrapping, we also tested the robustness of Phthiraptera + Liposcelididae clade by comparing the best ML tree estimated above with an alternative ML tree estimated with constrained monophyly of Psocoptera. As a result, the best tree (Fig. 3) was considered to be significantly better than the alternative tree (tree not shown: $-\ln = 20491.528$) by Shimodaira and Hasegawa (1999) test ($P = 0.024$).

All analyses supported the monophyly of Hemiptera, Psocodea, and Psocomorpha. However, monophyly of Phthiraptera and Amblycera was not recovered by all

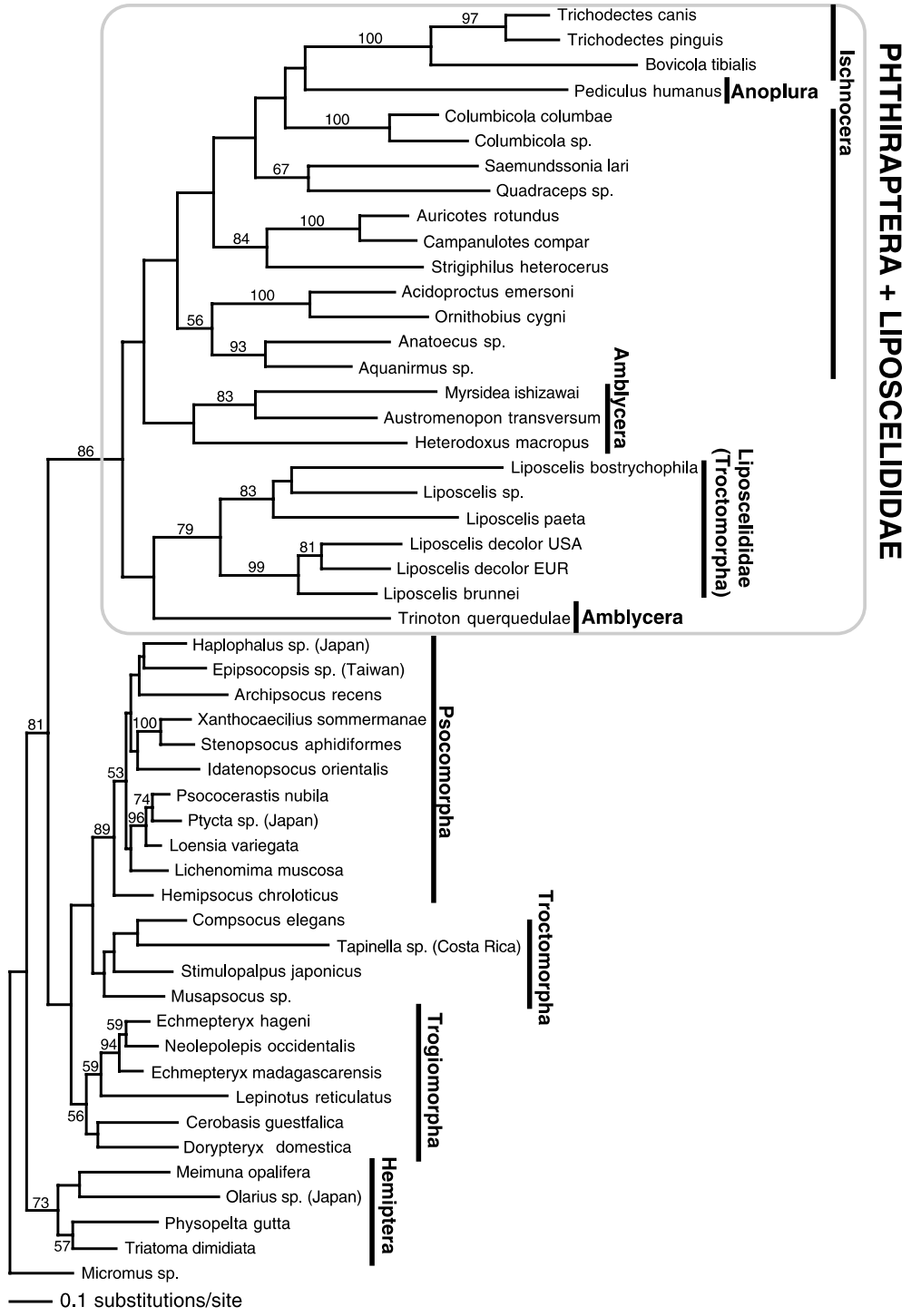

Fig. 3. Maximum-likelihood tree estimated from the data set excluding highly variable region ($-\ln = 20471.478$). Branch lengths are proportional to reconstructed distances. Numbers next to branches indicate nodes supported in >50% of bootstrap replicates.

analyses, with Liposcelididae falling with these groups in these trees. Monophyly of Ischnocera + Anoplura was supported by NJ and ML analyses, but monophyly of Ischnocera was not recovered by any analysis, with *Pediculus* falling inside Ischnocera.

We also analyzed the data set including highly variable regions. Results of these analyses (trees not shown) were nearly identical with those obtained from the previous data set. Phthiraptera + Liposcelididae clade was also supported. In contrast, bootstrap support for some clades, particularly those within Psocomorpha, improved upon the inclusion of these highly variable regions. Additionally, using NJ analysis, monophyly of Trogiomorpha was recovered by inclusion of highly

variable regions which was not the case with the well-aligned data set (Fig. 2).

3.3. Elevated rate of evolution in mtDNA

Branch lengths of the resulting trees clearly indicated an elevated rate of substitution in louse mtDNA (Fig. 4), as reported by Hafner et al. (1994), Page et al. (1998), and Johnson et al. (2003). An elevated rate of substitution in mtDNA was observed throughout Phthiraptera. Additionally, an elevated rate was also observed in Liposcelididae and to some extent *Tapinella*. There were no significant differences in the branch lengths between Phthiraptera and Liposcelididae ($P = 0.40$ [ML] by Mann–Whitney U test). Mean branch length of Phthiraptera + Liposcelididae estimated by ML analysis was about three times longer than that of Psocoptera excluding Liposcelididae and *Tapinella*. Differences of branch lengths between Phthiraptera + Liposcelididae and Psocoptera (excl. Liposcelididae and *Tapinella*) were significant ($P < 0.001$). In contrast, branch lengths of Psocoptera (excl. Liposcelididae and *Tapinella*) were not significantly different from those of Hemiptera ($P = 0.35$ [ML]).

Divergence of 12S rDNA secondary structure reported in Phthiraptera (Page et al., 2002) was here identified in a loop region of Liposcelididae. In *Liposcelis brunnea* and the USA sample of *L. decolor*, very long insertions were detected in the loop region between helix 36 and 38. Estimated lengths of the loop region were 35 in *L. brunnea* and 39 in *L. decolor* (USA), whereas those in the other samples of *Liposcelis*, Psocoptera, Hemiptera, and Neuroptera were 3–11. Furthermore, an additional helix (helix 37 of Page et al., 2002) could be detected in the loop region of *L. brunnea* and *L. decolor* (USA). Sequences and structures of the helix were entirely different from each other (Fig. 5). In other regions of Liposcelididae, such long insertions and additional helixes could not be detected. Existence of helix 37 was also detected in four genera and four species of Phthiraptera in the present study (Fig. 5), on top of the five species in four genera reported by Page et al. (2002).

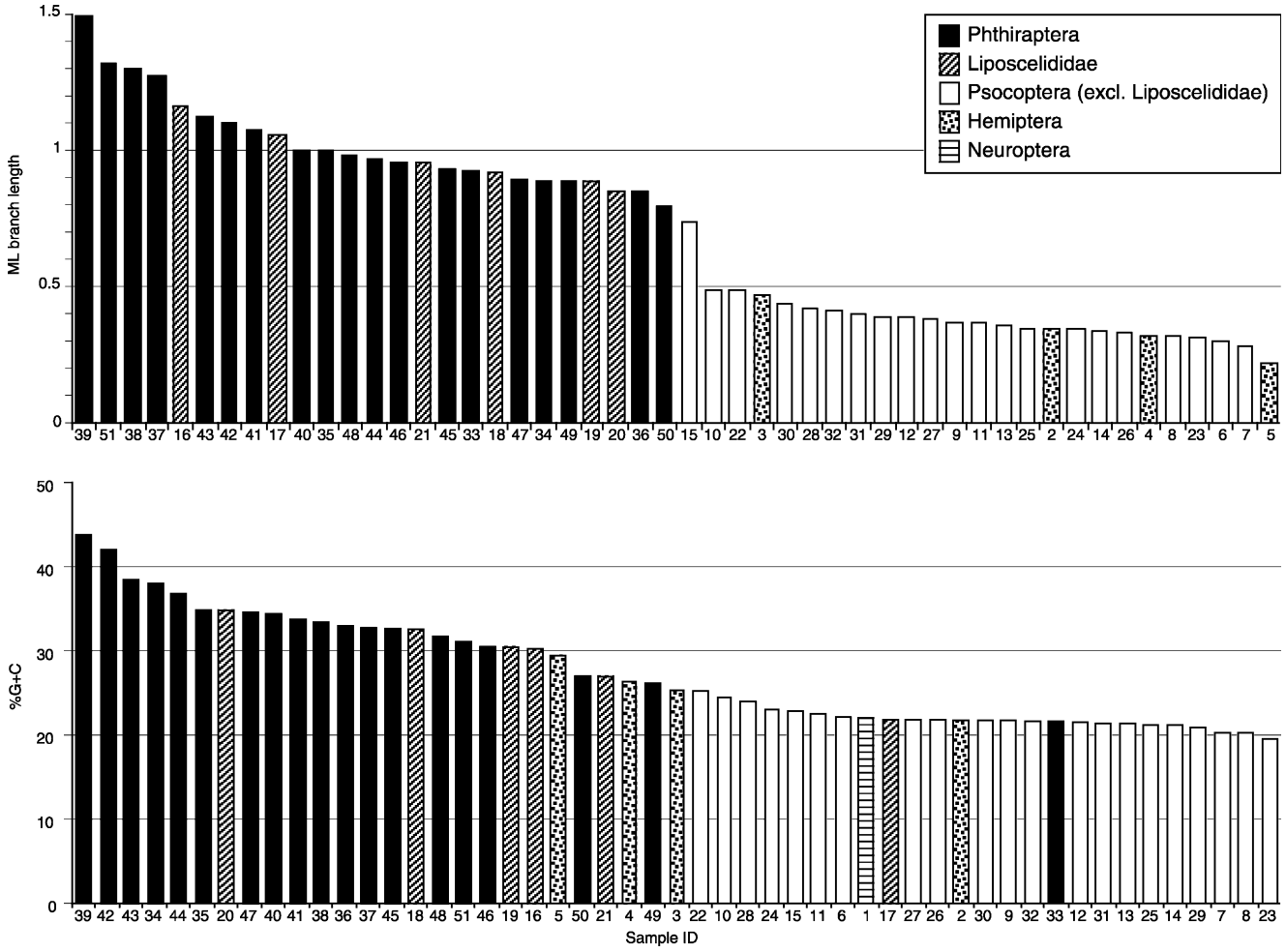

Fig. 4. Maximum-likelihood branch length from outgroup node to tip of branch based on Fig. 3 (top), and G + C contents of each sample (bottom). Major higher categories are differentiated by color patterns. Sample ID (X axis) corresponds to that in Table 1.

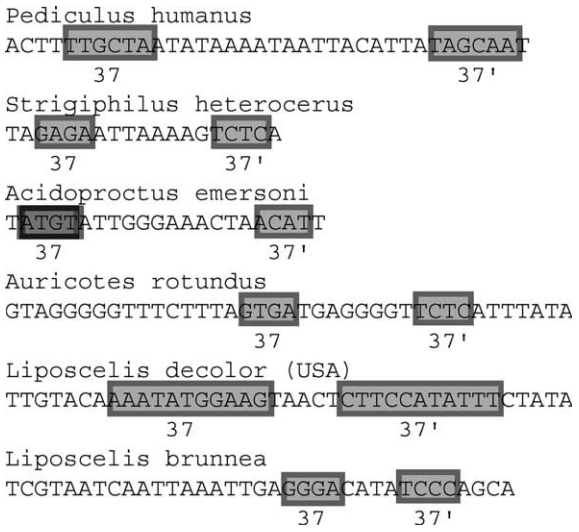

Fig. 5. Estimated structure of helix 37.

Although exact positions and structures could not be determined, long insertions were detected somewhere between helix 39 and 42 of *Austromenopon transversum* and two *Trichodectes* species, and between helix 34′ and 33′ of *H. macropus*. These sequences include two amblyceran species, although an amblyceran species examined by Page et al. (2002) and three other species examined here had the typical insect structure. Long indels were also detected in some 16S rDNA regions. However, these highly variable regions of 12S and 16S rDNA were very poorly aligned and thus, structural analysis of these regions awaits further study.

Related to the large divergences of the mtDNA of lice and liposcelidids, significant differences in G + C contents among Phthiraptera + Liposcelididae, the other Psocoptera, and Hemiptera were also detected (Fig. 4). By Mann–Whiteny *U* test, no significant differences of G + C contents between Phthiraptera (21.6–43.9%: mean 33.5%) and Liposcelididae (21.9–34.7%: mean 29.5%) could be detected ($P = 0.07$). However G + C contents of Psocoptera (19.6–25.2%: mean 22.0%) were significantly lower than those of Phthiraptera ($P < 0.001$) and Liposcelididae ($P = 0.003$). Differences between Psocoptera and Hemiptera were also significant ($P = 0.011$), and G + C contents of Hemiptera (21.8–29.5%: mean 25.8%) were slightly higher than those of Psocoptera. Higher G + C contents in louse cyt-*b* in comparison with some other insect orders was also reported by Simmons and Weller (2001).

4. Discussion

4.1. Alignment and data evaluation

Using the Secondary Structure Mode of ClustalX, most of the stem regions of 12S and 16S rDNA can be identified reliably. In contrast, some highly variable regions are only poorly aligned. As shown by Page et al. (2002), some stem and loop regions of rDNA contain little phylogenetic signal and, if this is the case, elimination of these regions should improve the phylogenetic accuracy (Olsen and Woese, 1993; Swofford et al., 1996; Yang, 1998). Elimination of data is sometimes criticized as being too arbitrary (Gatesy et al., 1993), or as eliminating useful phylogenetic information (Lee, 2001). However, as examined above, we cannot identify any data homogeneity nor topological congruence between conserved and highly variable regions. Judging from the ambiguity of alignment, the highly variable regions seem to contain little consistent phylogenetic information over the entire tree.

The data set including highly variable regions yields basically similar trees and provides higher bootstrap supports for several clades, especially for Psocomorpha (MP 77 → 86%; NJ 84 → 96%; and ML 89 → 91%). Monophyly of Trogiomorpha is also strengthened by inclusion of highly variable regions (NJ < 50 → 69%; ML 56 → 80%). Monophyly of these taxa is also strongly supported by morphology (Yoshizawa, 2002, pers. obs.). As examined above, some highly variable regions were very poorly aligned mainly because long indels are observed in some phthirapteran and liposcelidid samples. These results strongly suggest that highly variable regions include considerable amount of phylogenetic information for Psocoptera. In contrast, possible negative effects of inclusion of highly variable regions are observed. For example, the present data set includes two samples of *L. decolor*, and they compose a monophyletic group across all analyses. However, bootstrap supports for *L. decolor* are apparently weakened by inclusion of highly variable regions (MP 77 → 64%; NJ 82 → 73%; and ML 81 → 73%), probably because long indels are observed only in the USA sample of this species. This result indicates that these regions have only a limited usefulness for uncovering the higher level phylogenetic relationships in Psocodea.

Comparison of the 12S and 16S data sets raises other questions. Treatment of multiple data sets, partitioning or combining, is usually judged by whether each data set is consistent with the same phylogeny or not (Bull et al., 1993). If data sets contain different phylogenetic signals, then the data sets should be analyzed separately. Alternatively, if we can be sure that data sets have experienced the same phylogenetic history, then the data sets can be combined and analyzed simultaneously.

As examined above, the partition homogeneity test suggests a significant heterogeneity between 12S and 16S data sets ($P = 0.003$). In contrast, topologies of NJ trees estimated from 12S and 16S indicate that each data set contains similar phylogenetic signals. As shown in Fig. 6, divergence of 12S rDNA is greater than that of 16S rDNA. Additionally, substitution of 12S rDNA appears

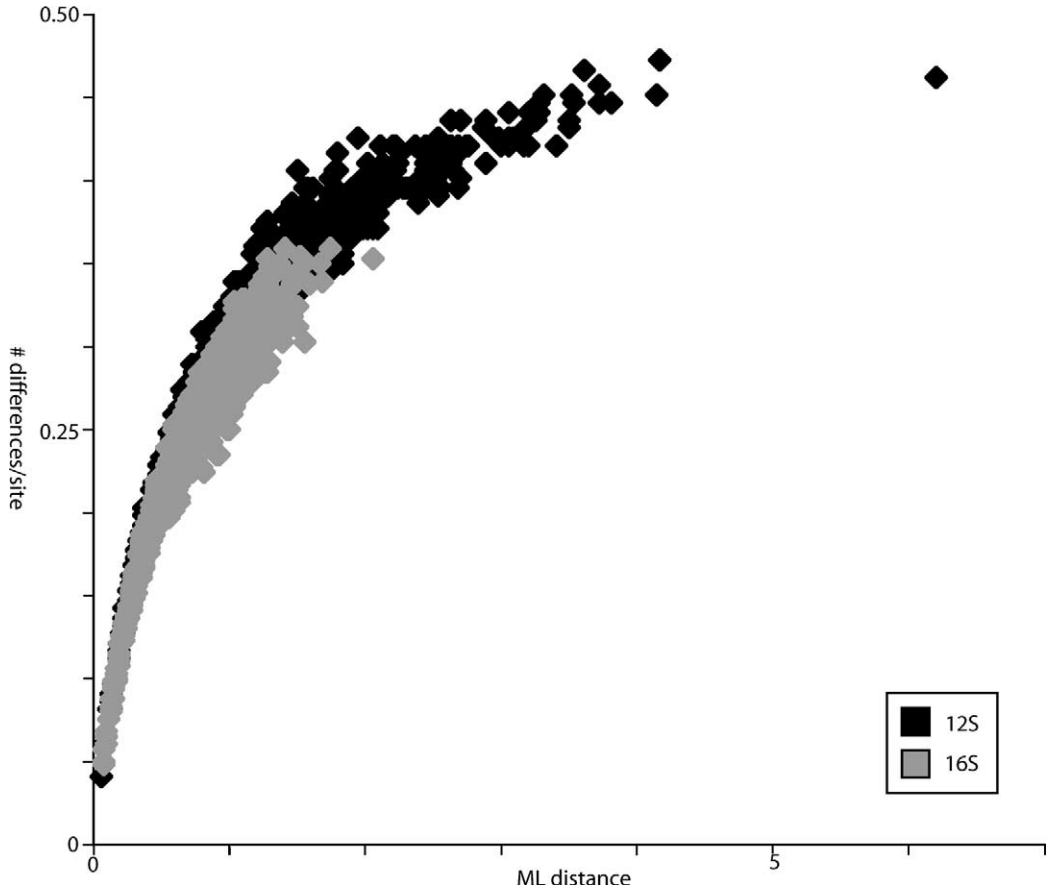

Fig. 6. Plot of number of substitution per site against maximum-likelihood distance for overall pairwise divergence in 12S and 16S rDNA (highly variable regions excluded).

to be almost saturated, while 16S rDNA does not. Therefore, significant heterogeneity between 12S and 16S data sets is probably due to the different evolutionary rate of 12S and 16S rDNA rather than the different phylogenetic signal (Barker and Lutzoni, 2002; Darlu and Lecointre, 2002; Johnson et al., 2002). Topological congruence of 12S and 16S indicates that both data sets are consistent with the same underlying phylogeny and supports combining of data sets.

4.2. Phylogenetic relationships

Results of the present analyses are highly congruent with external results. For example, monophyly of Hemiptera, Psocodea, Trogiomorpha, and Psocomorpha is supported by the present analyses and is also strongly supported by morphology (Rudolph, 1982, 1983; Yoshizawa and Saigusa, 2001; Yoshizawa, 2002, pers. obs.). This comparison suggests that the results obtained from the present analyses are reasonable, and mitochondrial rDNA sequences provide useful information for the higher systematics of Psocodea.

All the present analyses strongly supports Lyal's (1985) hypothesis of a Phthiraptera + Liposcelididae clade. Our results suggest that all the loss morphological character states shared by Phthiraptera and Liposcelididae are their synapomorphies, although Lyal (1985) mentioned the possibility of their independent origin resulting from similar selective pressures. Also, as mentioned by Lyal (1985), apomorphic egg structure and embryonic orientation observed in Psocoptera should be considered to reverse to the plesiomorphic state in Phthiraptera.

As suggested from morphology (Lyal, 1985), monophyly of Ischnocera + Anoplura is supported by the present analyses. In contrast, monophyly of Ischnocera is not supported since *Pediculus* always falls inside Ischnocera, but its position is highly unstable across analyses. Monophyly of Ischnocera is supported by analyses of other genes, notably 18S rDNA (Johnson and Whiting, 2002).

The sister group of the Liposcelididae + Phthiraptera clade is unstable across analyses: *Dorypteryx* (MP), *Tapinella* (NJ), or all Psocoptera excluding Liposcelididae (ML). In the present taxonomic system based on morphology (e.g., Smithers, 1996), a closer relationship between *Tapinella* and Liposcelididae is supported. However, further molecular and morphological data are

required to clearly identify the sister taxon of Phthiraptera + Liposcelididae.

4.3. Elevated rate of evolution in mtDNA

Elevated rate of substitution in mtDNA of Phthiraptera has been reported by Hafner et al. (1994) [COI], Page et al. (1998) [cyt-*b*], and Johnson et al. (2003) [COI]. Hafner et al. (1994) mentioned that this elevated rate is possibly caused by different generation times between lice and their vertebrate hosts. Additionally, Page et al. (1998) mentioned other possibilities causing accelerated molecular evolution: (1) adaptive evolution related to ectoparasitic lifestyle and (2) low louse population size caused in part by the restricted mode of transmission between hosts. Simmons and Weller (2001) suggested that increased substitution rate in honeybee mtDNA is probably a result of endothermy. Shao et al. (2001b) showed an increased rate of gene rearrangement in mtDNA in Thysanoptera, Psocoptera, and Phthiraptera, noting that the elevated rate of substitution may be correlated with an increased rate of gene rearrangement.

Results of the present analyses show that the elevated rate of substitution in mtDNA originated in the common ancestor of Phthiraptera + Liposcelididae (and maybe also *Tapinella*). Origin of increased gene rearrangement in mtDNA is considered to be older than that of Psocodea, and the common ancestor of Phthiraptera + Liposcelididae is assumed to be free living insect. Therefore, the origin of the elevated substitution rate does not directly correspond to either the increased rate of gene rearrangement or the origin of parasitism.

Elevated substitution rates in louse and liposcelidid mtDNA may be, in part, related to their short generation time. The usual number of instars in Psocoptera is six whereas Liposcelididae and Phthiraptera typically have only four instars (Marshall, 1981; New, 1987). However, unusually long *Tapinella* branch cannot be explained by its generation time because the number of instars of a species of *Tapinella* (*T. africana*) is known to be six in females (Tsutsumi, 1962). The branch length of *Meimuna opalifera* (Cicadomorpha) is almost same with that of the other hemipterans, although cicadas have a generally long generation time.

The high G + C contents observed in Phthiraptera and Liposcelididae provide different interpretation. We have shown that both the substitution rate and the G + C content of mtDNA are significantly increased in Phthiraptera and Liposcelididae. Thus, these two phenomena are considered to have originated in their common ancestor and are probably strongly correlated with each other. Therefore, the elevated substitution rate of mtDNA in Phthiraptera and Liposcelididae appears to be directional. Both evolutionary trends would not be expected to result from a shorter generation time.

The rate of mtDNA evolution is also known to correlate with the thermal habit of an organism, and the mtDNA of warm-blooded vertebrates evolves much faster than that of cold-blooded ones (e.g., Rand, 1994). In insects, the elevated substitution rate observed in honeybee mtDNA is considered to be a result of their endothermy and high metabolic rate (Simmons and Weller, 2001). All lice parasitize warm-blooded animals, and thus, their habitats are considered to be much warmer than those of the other insects such as psocids and hemipterans. As also mentioned by Simmons and Weller (2001), this warm environment may affect the metabolic rate and the increased evolutionary tempo of louse mtDNA. High G + C contents seem to provide further support for this hypothesis, because G–C bonds are stronger and more stable than A–T bonds (Zucker, 1989; Zucker et al., 1991). Thus, higher G + C contents would be adoptive under high temperatures. However, as already mentioned above, increased substitution rate and increased G + C contents are considered to have evolved in non-parasitic ancestor of Phthiraptera + Liposcelididae. The liposcelidid species used in the present analyses inhabit human environments, and thus, their rates may be affected by a warm environment. To test this, species of *Liposcelis* not known to inhabit continually warm environments need to be examined.

Nadler et al. (1990) suggested that the lice may undergo founder events with each initial infection of a juvenile gopher. This succession of founder events could account for the accelerated substitution rate of louse mtDNA (Ohta, 1972, 1987; Page et al., 1998). Additionally, slightly deleterious mutations are expected to be fixed at relatively higher rate in such small population (Kimura, 1983). Insect mtDNA is known to be AT-rich (Simon et al., 1994), and AT-transversion bias, which is responsible for the AT-rich base composition of insect mtDNA (Tamura, 1992), is reported (Dowton and Austin, 1997). If an AT-rich base composition is maintained by natural selection, a higher G + C contents would be slightly deleterious for insect mtDNA. If this is the case, then the G + C contents should increase in small populations compared to large ones (Ohta, 1973, 1992). Therefore, an increased frequency of founder events could be responsible for both increased substitution rates and higher G + C contents. An incleased rate of nonsynonymous substitution in louse cyt-*b* was reported (Simmons and Weller, 2001) which is also considered to be slightly deleterious and possibly a result of repeated founder events. Again, liposcelidids are free-living insects, and thus, such an increased frequency of founder events is difficult to be assumed for liposcelidids.

In addition to the elevated substitution rate and gene rearrangement, divergence of louse 12S rDNA secondary structure is notable. Page et al. (2002) reported high variability of lice 12S rDNA structure, including the

presence of unique helix 37. In the present study, helix 37 is detected in non-phthirapteran samples for the first time, i.e., *Liposcelis brunnea* and the USA sample of *L. decolor*. Such high variability of 12S rDNA structure cannot be detected in other psocopteran and hemipteran samples. Thus, as suggested by Page et al. (2002), the diversity of 12S rDNA structure seems symptomatic of an elevated rate of substitution in mtDNA. Ohta (1972, 1974) showed that modification of the secondary structure of tRNA is slightly deleterious. Therefore, a high diversity of 12S rDNA structure may also be explained as fixation of slightly deleterious mutations, as in the case of elevated substitution rates and high G + C contents. Our study also newly detects helix 37 in *Pediculus humanus*, *Strigiphilus heterocerus*, *Acidoproctus emersoni*, and *Auricotes rotundus*. The *Pediculus* sequence is the first evidence of helix 37 in Anoplura, because Page et al. (2002) noted that anopluran sequences they examined display typical insect structure. Distribution of samples having helix 37 on the resulted phylogenetic trees clearly shows the repeated independent evolution of helix 37. As mentioned by Page et al. (2002), secondary structure of louse and liposcelidid 12S rDNA contains little phylogenetic information. This result is even the case within closely related groups. Helix 37 is detected in only two of six *Liposcelis* samples. Even between these two liposcelidid samples, sequences and structures of helix 37 are quite different. Furthermore, among two samples of *L. decolor*, only the USA sample possesses helix 37. These results suggest repeated evolution of helix 37, even among closely related species.

4.4. Systematics of psocodea

The present analyses strongly suggest that the order Psocoptera is paraphyletic. Therefore, in a classification that reflects monophyletic groupings, the order Psocoptera in the present sense must be rejected. However, the MP and ML trees indicate a possibility that Psocoptera may be maintained as monophyletic group by excluding Liposcelididae. Another possibility is to recognize three orders in Psocodea: (1) Trogiomorpha, (2) Psocomorpha + Troctomorpha (excl. Lipsocelididae), and (3) Phthiraptera + Liposcelididae, which does not conflict with any of the present results, barring ambiguity in the phylogenetic position of *Tapinella* (Pachytroctidae). In either case, Liposcelididae would likely be transferred to the order Phthiraptera and should be recognized as a fifth suborder of lice.

Acknowledgments

We thank E.L. Mockford and Z. Kucerova for specimens and identification; T. Ito, M.A. Iwasa, R. Matsumoto, K. Mizota, T. Nakamura, H. Sakai, S. Sakurai, B.J. Sinclair, and M. Sueyoshi for specimens; H. Takahashi, S. Hikosaka, and Y. Chin for support to collect lice; D.C. Morris and two anonymous reviewers for helpful comments on the manuscript. K.Y. thanks K. Odagiri, S. Nishida, E. Hasegawa, and K. Futami for assistance in experiments; H. Koike and Y. Saito for allowing to use their molecular laboratories. This project was supported by JSPS research fellowship (PD) and Grant-in Aid from JSPS (13740486) to K.Y. and NSF Grant DEB-0107891 to K.P.J.

References

Barker, S.C., 1994. Phylogeny and classification, origins, and evolution of host associations of lice. Int. J. Parasitol. 24, 1285–1291.

Barker, F.K., Lutzoni, F.M., 2002. The utility of the incongruence length difference test. Syst. Biol. 51, 625–637.

Buckley, T.R., Simon, C., Flook, P.K., Misof, B., 2000. Secondary structure and conserved motifs of the frequently sequenced domains IV and V of the insect mitochondrial large subunit rRNA gene. Insect Mol. Biol. 9, 565–580.

Bull, J.J., Huelsenbeck, J.P., Cunningham, C.W., Swofford, D.L., Waddell, P.J., 1993. Partitioning and combining data in phylogenetic analysis. Syst. Biol. 42, 384–397.

Cruickshank, R.H., Johnson, K.P., Smith, V.S., Adams, R.J., Clayton, D.H., Page, R.D.M., 2001. Phylogenetic analysis of partial sequences of elongation factor 1α identifies major groups of lice (Insecta: Phthiraptera). Mol. Phylogenet. Evol. 19, 202–215.

Darlu, P., Lecointre, G., 2002. When does the incongruence length difference test fail. Mol. Biol. Evol. 19, 432–437.

Dotson, E.M., Beard, C.B., 2001. Sequence and organization of the mitochondrial genome of the Chagas disease vector, *Triatoma dimidiata*. Insect Mol. Biol. 10, 205–215.

Dowton, M., Austin, A.D., 1997. Evidence for AT-transversion bias in wasp (Hymenoptera: Symphyta) mitochondrial genes and its implications for the origin of parasitism. J. Mol. Evol. 44, 398–405.

Estabrook, G.F., 1992. Evaluating undirected positional congruence of individual taxa between two estimates of the phylogenetic tree from a group of taxa. Syst. Biol. 41, 172–177.

Farris, J.S., Kallersjo, M., Kluge, A.G., Bult, C., 1994. Testing significance of congruence. Cladistics 10, 315–320.

Farris, J.S., Kallersjo, M., Kluge, A.G., Bult, C., 1995. Constructing a significance test for incongruence. Syst. Biol. 44, 570–572.

Galtier, N., Gouy, M., 1995. Inferring phylogenies from DNA sequences of unequal base compositions. Proc. Natl. Acad. Sci. USA 92, 11317–11321.

Gatesy, J., DeSalle, R., Wheeler, W., 1993. Alignment-ambiguous nucleotide sites and the exclusion of systematic data. Mol. Phylogenet. Evol. 2, 152–157.

Hafner, M.S., Sudman, P.D., Villablanca, F.X., Spradling, T.A., Demastes, J.W., Nadler, S.A., 1994. Disparate rates of molecular evolution in cospeciating hosts and parasites. Science 265, 1087–1090.

Hickson, R.E., Simon, C., Perrey, S.W., 2000. The performance of several multiple sequence alignment programs in relation to secondary-structure features for a rRNA sequence. Mol. Biol. Evol. 17, 530–539.

Johnson, K.P., Cruickshank, R.H., Adams, R.J., Smith, V.S., Page, R.D.M., Clayton, D.H., 2003. Dramatically elevated rate of mitochondrial substitution in lice (Insecta: Phthiraptera) Mol. Phylogenet. Evol. 26, 231–242.

Johnson, K.P., Weckstein, J.D., Witt, C.C., Faucett, R.C., Moyle, R.G., 2002. The perils of using host relationships in parasite

taxonomy: phylogeny of the *Degeeriella* complex. Mol. Phylogenet. Evol. 23, 150–157.

Johnson, K.P., Whiting, M.F., 2002. Multiple genes and the monophyly of Ischnocera (Insecta: Phthiraptera). Mol. Phylogenet. Evol. 22, 101–110.

Kimura, M., 1983. The Neutral Theory of Molecular Evolution. Cambridge University Press, Cambridge.

Kristensen, N.P., 1995. Forty years' insect phylogenetic systematics. Hennig's "Kritische Bemerkungen..." and subsequent developments. Zool. Beit. N. F. 36, 83–124.

Lee, M.S.Y., 2001. Unalignable sequences and molecular evolution. Trends Ecol. Evol. 16, 681–685.

Lockhart, P.J., Steel, M.A., Hendy, M.D., Penny, D., 1994. Recovering evolutionary trees under a more realistic model of sequence evolution. Mol. Biol. Evol. 11, 605–612.

Lyal, C.H.C., 1985. Phylogeny and classification of the Psocodea, with particular reference to the lice (Psocodea: Phthiraptera). Syst. Entomol. 10, 145–165.

Marshall, A.G., 1981. The Ecology of Ectoparasitic Insects. Academic Press, London.

Nadler, S.A., Hafner, M.S., Hafner, J.C., Hafner, D.J., 1990. Genetic differentiation among chewing louse populations (Mallophaga: Trichodectidae) in a pocket gopher contact zone (Rodentia: Geomyidae). Evolution 44, 942–951.

New, T.R., 1987. Biology of the Psocoptera. Oriental Insects 21, 1–109.

Ohta, T., 1972. Population size and rate of evolution. J. Mol. Evol. 1, 305–314.

Ohta, T., 1973. Slightly deleterious mutant substitutions in evolution. Nature 246, 96–98.

Ohta, T., 1974. Mutational pressure as the main cause of molecular evolution and polymorphism. Nature 252, 351–354.

Ohta, T., 1987. Very slightly deleterious mutations and the molecular clock. J. Mol. Evol. 26, 1–6.

Ohta, T., 1992. The nearly neutral theory of molecular evolution. Annu. Rev. Ecol. Syst. 23, 263–286.

Olsen, G.J., Woese, C.R., 1993. Ribosomal RNA: a key to phylogeny. FASEB J. 7, 113–123.

Page, R.D.M., Cruickshank, R., Johnson, K.P., 2002. Louse (Insecta: Phthiraptera) mitochondrial 12S rRNA secondary structure is highly variable. Insects Mol. Biol. 11, 361–369.

Page, R.D.M., Lee, P.L.M., Becher, S.A., Griffiths, R., Clayton, D.H., 1998. A different tempo of mitochondrial DNA evolution in birds and their parasitic lice. Mol. Phylogenet. Evol. 9, 276–293.

Paterson, A.M., Wallis, G.P., Wallis, L.J., Gray, R.D., 2000. Seabird and louse coevolution: complex histories revealed by 12S rRNA sequences and reconciliation analysis. Syst. Biol. 49, 383–399.

Posada, D., Crandall, K.A., 1998. MODELTEST: testing the model of DNA substitution. Bioinformatics 14, 817–818.

Rand, D.M., 1994. Thermal habit, metabolic rate, and the evolution of mitochondrial DNA. Trends Ecol. Evol. 9, 125–131.

Rudolph, D., 1982. Occurrence, properties, and biological implications of the active uptake of water vapour from the atmosphere in the Psocoptera. J. Insect Physiol. 28, 111–121.

Rudolph, D., 1983. The water–vapour uptake system of the Phthiraptera. J. Insect Physiol. 29, 15–25.

Seeger, W., 1979. Spezialmerkmale an Eihüllen und Embryonen von Psocoptera im Vergleich zu anderen Paraneoptera (Insecta): Psocoptera als monophyletische Gruppe. Stutt. Beit. Natur. (A) 329, 1–57.

Shao, R., Campbell, N.J.H., Barker, S.C., 2001a. Numerous gene rearrangements in the mitochondrial genome of the wallaby louse, *Heterodoxus macropus* (Phthiraptera). Mol. Biol. Evol. 18, 858–865.

Shao, R., Campbell, N.J.H., Schmidt, E.R., Barker, S.C., 2001b. Increased rate of gene rearrangement in the mitochondrial genomes of three orders of hemmipteroid insects. Mol. Biol. Evol. 18, 1828–1832.

Shimodaira, H., Hasegawa, M., 1999. Multiple comparisons of log-likelihoods with applications to phylogenetic inference. Mol. Biol. Evol. 16, 1114–1116.

Simon, C., Frati, F., Beckenbach, A., Crespi, B., Liu, H., Flook, P., 1994. Evolution, weighting, and phylogenetic utility of mitochondrial gene sequences and a compilation of conserved polymerase chain reaction primers. Ann. Entomol. Soc. Am. 87, 651–701.

Simmons, R.B., Weller, S.J., 2001. Utility and evolution of cytochrome *b* in insects. Mol. Phylogenet. Evol. 20, 196–210.

Smithers, C.N., 1996. Psocoptera. In: Wells, A. (Ed.), Psocoptera, Phthiraptera, Thysanoptera, Vol. 26. CSIRO, Melbourne, pp. 1–79.

Swofford, D.L., 2002. PAUP*. Phylogenetic analysis using parsimony (* and other methods). Version 4, Sinauer Assoc., Sunderland, MA.

Swofford, D.L., Olsen, G.J., Waddell, P.J., Hillis, D.M., 1996. Phylogenetic inference. In: Hillis, D.M. et al. (Eds.), Molecular systematics. Sinauer, Sunderland, MA.

Tamura, K., 1992. The rate and pattern of nucleotide substitution in *Drosophila* mitochondrial DNA. Mol. Biol. Evol. 9, 814–825.

Thompson, J.D., Gibson, T.J., Plewniak, F., Jeanmougin, F., Higgins, D.G., 1997. The ClustalX windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucleic Acids Res. 25, 4876–4882.

Thorley, J.L., Page, R.D.M., 2000. RadCon: phylogenetic tree comparison and consensus. Bioinformatics 16, 486–487.

Tsutsumi, C., 1962. Studies on the domiciliary psocids of Japan (I). Jpn. J. Sanit. Zool. 13, 190–196.

Waage, J.K., 1979. The evolution of insect/vertebrate associations. Biol. J. Linn. Soc. 12, 187–224.

Wheeler, W.C., Whiting, M., Wheeler, Q.D., Carpenter, J.M., 2001. The phylogeny of the extant hexapod orders. Cladistics 17, 113–169.

Yang, Z., 1998. On the best evolutionary rate for phylogenetic analysis. Syst. Biol. 47, 125–133.

Yoshizawa, K., 2002. Phylogeny and higher classification of suborder Psocomorpha (Insecta: Psocodea: 'Psocoptera'). Zool. J. Linn. Soc. 136, 371–400.

Yoshizawa, K., Saigusa, T., 2001. Phylogenetic analysis of paraneopteran orders (Insecta: Neoptera) based on forewing base structure, with comments on monophyly of Auchenorrhyncha (Hemiptera). Syst. Entomol. 26, 1–13.

Zucker, M., 1989. Computer prediction of RNA structure. Methods Enzymol. 180, 262–288.

Zucker, M., Jaeger, J.A., Turner, D.H., 1991. A comparison of optimal and suboptimal RNA secondary structures predicted by free energy minimization with structures determined by phylogenetic comparison. Nucleic Acid Res. 19, 2707–2714.