Phylogenetic Position of *Xenoturbella bocki* and Hemichordates *Balanoglossus carnosus* and *Saccoglossus kowalevskii* Based on Amino Acid Composition or Nucleotide Content of Complete Mitochondrial Genomes

Kenji Sorimachi[1,2], Teiji Okayasu[3], Yoshifumi Ebara[4], Emiko Furuta[5] & Shuji Ohhira[6]

[1] Life Science Research Center, Higashi-Kaizawa, Takasaki, Gunma, Japan

[2] Educational Support Center, Dokkyo Medical University, Mibu, Tochigi, Japan

[3] Center of Medical Informatics, Dokkyo Medical University, Mibu, Tochigi, Japan

[4] Terao Child Clinic, Terao, Takasaki, Gunma, Japan

[5] The Research Institute for Comparative Immunology, Saitama, Japan

[6] Laboratory for International Environmental Health, Dokkyo Medical University, Mibu, Tochigi, Japan

Correspondence: Kenji Sorimachi, Life Science Research Center, Higashi-Kaizawa, Takasaki, Gunma 370-0041, Japan. Tel: 81-27-352-2955. E-mail: kenjis@jcom.home.ne.jp

Abstract

To clarify the phylogenetic position of *Xenoturbella bocki* and hemichordates (*Balanoglossus carnosus* and *Saccoglossus kowalevskii*), which occupy an ambiguous evolutionary position between vertebrates and invertebrates, we expanded a pre-existing baseline vertebrate data set by incorporating randomly selected invertebrates. Based on nucleotide content calculated from complete mitochondrial genomes, invertebrates were classified into two groups, high C/G and low C/G, with vertebrates placed within the high-C/G invertebrate group. *X. bocki* and hemichordates (*B. carnosus* and *S. kowalevskii*) also fell into the high-C/G invertebrate group. We also analyzed amino acid composition and nucleotide content of complete mitochondrial genomes or 16S rRNA genes using Ward's clustering method and neighbor-joining. In the resulting phylogenetic trees, invertebrates are classified into high C/G and low C/G categories, and vertebrates are divided into terrestrial and aquatic groups. In addition, *X. bocki* and representative hemichordates *B. carnosus* and *S. kowalevskii* appear to be more closely related to vertebrates than to invertebrates.

Keywords: amino acid composition, complete mitochondrial genome, vertebrate, invertebrate, hemichordate, *Xenoturbella*, phylogenetic tree, evolution

1. Introduction

Evolution can be regarded as a branching process, whereby populations are altered over time and may speciate into separate branches. Approximately 150 years ago, Charles Darwin and Alfred Wallace established the concept of biological evolution by natural selection. Based on this theory, Charles Darwin and Ernst Haeckel developed the idea of phylogenetic trees as a visualization tool to aid readers' understanding of biological evolution.

Because evolution generally occurs over long periods of time and hence cannot be observed directly, scientists must construct phylogenies to infer evolutionary relationships among present-day organisms. Two different approaches — cladistics and phonetics — have been used to construct phylogenetic trees; the differences between these methods are beyond the scope of our study and are not described here.

To construct phylogenetic trees representing organismal relationships, similarities or differences in organismal traits are analyzed using implicit or explicit mathematical models. These traits can be morphological or molecular. Morphological characters include anatomical characteristics of extant organisms and changes in the fossil record; the latter data are limited and tend to be ambiguous (Cobbett et al., 2007). Molecular data, including protein and DNA sequences, have been widely used for constructing phylogenetic trees (Zuckerkandl

& Pauling, 1962; Dayhoff et al., 1977; Sogin et al., 1986; Doolittle & Brown, 1994; Maizels & Weiner, 1994; DePouplana et al., 1998; Woese & Fox, 1977; Weisburg et al., 1991). Complete genome sequence data are not readily applied to phylogenetic tree construction, however, because of the enormous number of nucleotides constituting many different genes among organisms. We have recently shown that ratios of amino acid composition and nucleotide content are useful for genome research because they can characterize entire genomes (Sorimachi et al., 2001). Based on such indices, we were able to generate phylogenetic trees in which vertebrates were classified into two groups, terrestrial and aquatic, even though the analyzed organisms were chosen randomly without pre-defined criteria (Sorimachi et al., 2013a). Similar classification results were also obtained in the previous study when deliberately selected samples were added to the pre-existing baseline data set (Sorimachi & Okayasu, 2013). Our previously developed method is thus applicable to phylogenetic tree construction. In this study, we therefore used amino acid composition and nucleotide content of complete mitochondrial genomes to ascertain the phylogenetic position of *Xenoturbella bocki* and hemichordates, which occupy an ambiguous evolutionary position between vertebrates and invertebrates.

2. Materials and Methods

Mitochondrial genome data were obtained from the National Center for Biotechnology Information (NCBI, http://www.ncbi.nlm.nih.gov/sites). In an earlier study, organisms were chosen according to the alphabetical order of their scientific names without considering their characteristics (Sorimachi et al., 2013a), because cluster analyses are significantly influenced by different sampling. Vertebrate species included 39 species from a previous study (Sorimachi et al., 2013a) to which 24 other vertebrates had been subsequently intentionally added to evaluate phylogenetic tree reconstructions (Sorimachi & Okayasu, 2013). For the current study, we added 58 invertebrate species, which were randomly chosen based solely on the alphabetical order of their scientific names. Cluster analysis has five methods based on differences in calculation procedures to estimate the distance between two samples; Ward's, nearest neighbor, furtherest neighbor, group average and centroid methods. Ward's method has been widely used, and we used this method to classify various organisms (112 bacteria, 15 archaea and 18 eukaryotes) (Okayasu & Sorimachi, 2009). In our recent study (Sorimachi & Okayasu, 2013), cetaceans (*Balaenoptera musculus*, *Sousa chinensis*, and *Phocoena phocoena*) form a small cluter that is closely related to hippopotamus. This placement is consistent with results of other studies (Catesy, Hayashi, Cronin, & Arctander, 1996; Ursing & Arnason, 1998), and supports the evolution of cetaceans from terrestrial vertebrates within the mammalian evolution. Based on these evidences, Ward's method has been mainly used in the present study. Cluster analyses were carried out on the data using Ward's (Ward, 1963) and neighbor-joining (Saitou & Nei, 1987) methods. Nucleotide contents of coding and non-coding regions of mitochondrial genomes were compared with the content of their complete corresponding single-strand DNA and normalized to 1 (G + C + T + A = 1). Predicted amino acid compositions of mitochondrial genome coding regions were estimated (Sorimachi et al., 2001; Sorimachi & Okayasu, 2004). Calculations were performed using Microsoft Excel (version 2003). Classifications based on Ward's clustering method (Ward, 1963) were conducted using commercially purchased-multivariate software developed by ESUMI (Tokyo, Japan).

3. Results

3.1 Relationships Based on Nucleotide Content

When nucleotide contents calculated from coding regions of complete mitochondrial genomes were plotted against those of entire genomes, invertebrates were classified into two groups: high C/G and low C/G (Figure 1). *X. bocki* and hemichordates (*Balanoglossus carnosus* and *Saccoglossus kowalevskii*) were located in the region where vertebrates and invertebrates overlapped. Vertebrates widely overlapped with high C/G invertebrates; on the plot, they were positioned in the upper right portion of the high-C/G invertebrate distribution (Figure 1).

When relationships based on nucleotide content were examined between homonucleotides using linear regression in a previous study (Sorimachi et al., 2013b), linear regression lines with high regression coefficients were obtained. Two distinct lines corresponding to invertebrates and vertebrates were observed. Because these two lines form a "wedge-shape" when superimposed, we concluded that vertebrates and invertebrates diverged from a common ancestor (Sorimachi et al., 2013b), and that extant organisms diverged from a single origin (Sorimachi, 2010). *X. bocki* and the hemichordates were located between these two lines (Figure 2).

3.2 Cluster Analysis of Invertebrates

Based on nucleotide content from complete mitochondrial genomes, the 58 randomly chosen invertebrates (Sorimachi & Okayasu, 2008) were analyzed using Ward's method (Ward, 1963). Three major clusters were obtained (Figure 3). Similar results were obtained using predicted amino acid composition, although the groups were not as well-differentiated (data not shown); this may be because the high- and low C/G invertebrate

classification is based on nucleotide relationships. In the tree generated using nucleotide content, one major cluster, consisting of Echinodermata, Mollusca, Arthropoda, and Placozoa — all high-C/G invertebrates except for *Trichoplax adhaerens* — is divided into two sub-clusters. The hemichordates *B. carnosus* and *S. kowalevskii* are found in one of these sub-clusters, with the deuterostome *X. bocki* falling into the other (Figure 3). The second major cluster, all low-C/G invertebrates, consists of Cnidaria, Arthropoda, Nematoda, and Platyhelminthes. Chordata (*Branchiostoma belcheri*) also falls into this major cluster, grouped with Platyhelminthes. The third major cluster consists of three sub-clusters. One of the three sub-clusters comprises Amoebozoa, Protozoa, Loukozoa, Mycetozoa, and Nematoda, all low-C/G invertebrates except for *Dictyostelium citrinum*. The second sub-cluster, containing high-C/G invertebrates, consists of Mollusca and Arthropoda. The final sub-cluster contains low-C/G organisms — Arthropoda, Protista (*Tetrahymena pyriformis*), and Nematoda (*Strelkovimermis spiculatus* and *Romanomermis spiculatus*), as well as high-C/G invertebrates.

Of seven species pairs belonging to the same genera, three (*Bactrocera*, *Reticulitermes*, and *Acanthaster*) remained joined in the phylogenetic tree; the other four species pairs (*Tigriopus*, *Dictyostelium*, *Drosophila*, and *Hemichordata*) were divided between separate clusters.

Similar classifications were obtained when nucleotide content calculated from coding and non-coding regions of complete mitochondrial genomes was analyzed.
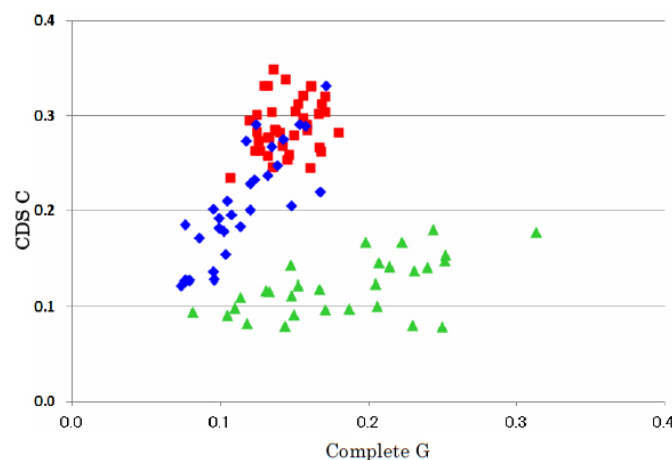
Figure 1. Relationships based on nucleotide content of complete mitochondrial genomes. Values are normalized to 1 (G + C + T + A = 1). Closed blue squares, closed green triangles, and closed red squares represent high-C/G invertebrates, low-C/G invertebrates, and vertebrates, respectively
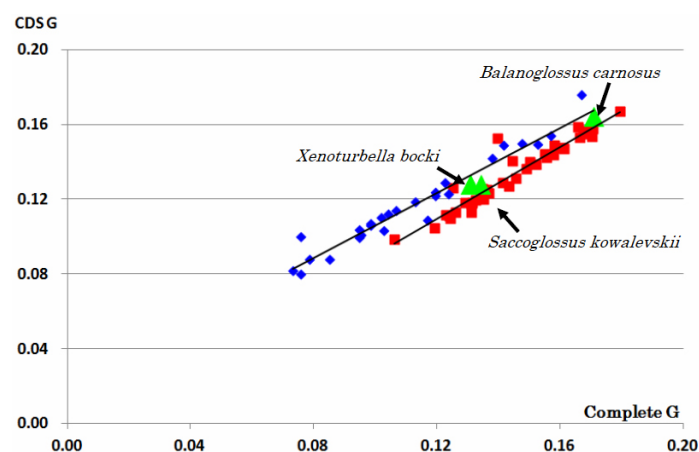
Figure 2. Nucleotide relationships based on normalized vertebrate (red squares) and high-C/G invertebrate (blue squares) valuesThe G nucleotide content in the coding region (vertical axis) is plotted against G content in a complete single DNA strand (horizontal axis). Closed green triangles represent *Xenoturbella* and Hemichordata

3.3 Phylogenetic Analysis of Vertebrates and High-C/G Invertebrates Using Ward's Method

In our previous study (Sorimachi et al., 2013a), vertebrates were classified into terrestrial and aquatic groups based on cluster analysis of amino acid composition using Ward's method. This classification was corroborated through analysis of 16S rRNA gene sequences using neighbor-joining. In this study, we consequently used Ward's method to analyze amino acid composition of a sample set consisting of 39 vertebrate species and 31 high-C/G invertebrates (Figure 4). In the resulting tree, vertebrates and invertebrates correspond to the two major groups, with the exception of the hemichordates and *X. bocki*, which are placed with the vertebrates. *X. bocki* is found in one of the two vertebrate sub-clusters, which comprises mammals and some reptiles, and is most closely related to reptiles. The hemichordates belong to the sub-cluster consisting of fish, amphibians, reptiles, and birds; within that sub-cluster, they are closely related to birds and reptiles. Hagfish (*Eptatretus burgeri*) is placed in the cluster consisting of mammals, consistent with our previous results (Sorimachi et al., 2013a). This placement may be due to its controversial characteristics (Janvier, 2010).

We also analyzed this sample set using nucleotide content calculated from complete mitochondrial genome data, but similarly clear classifications were not obtained (data not shown).
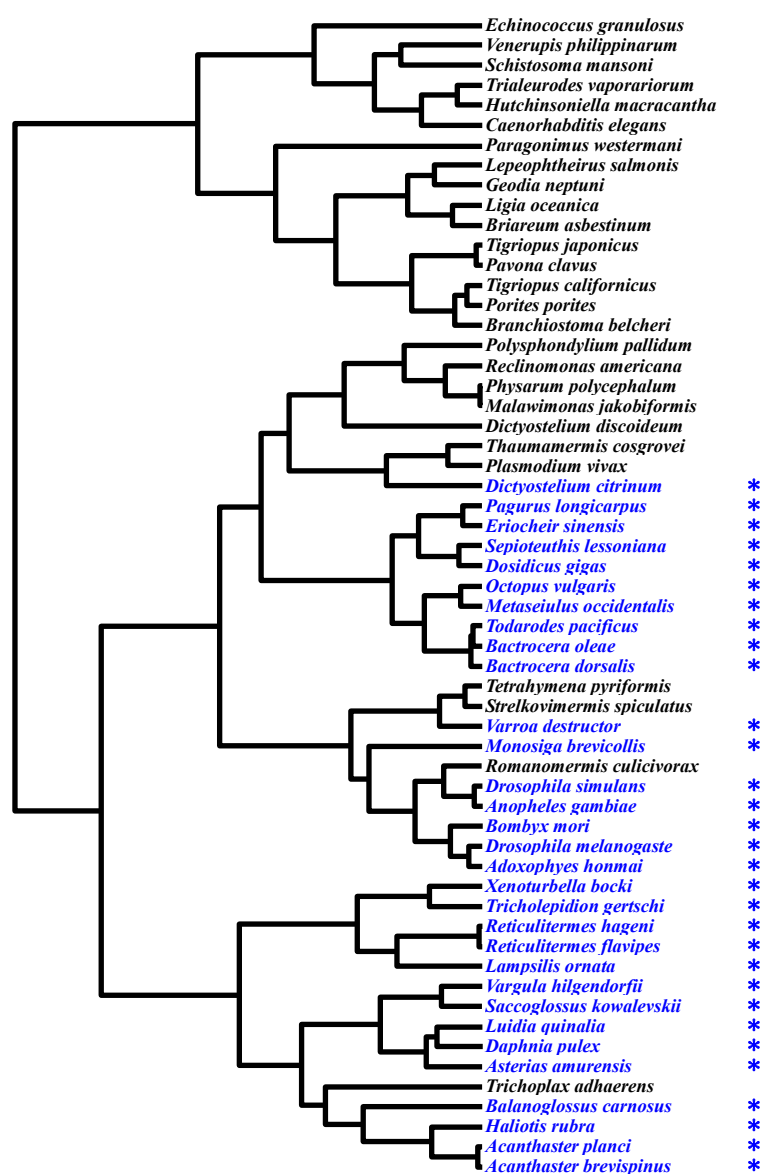
Figure 3. Phylogenetic tree generated using Ward's cluster analysis method (Ward, 1963) from nucleotide content of complete mitochondrial genomes. High-C/G and low-C/G invertebrates are indicated in blue (*) and black (null), respectively
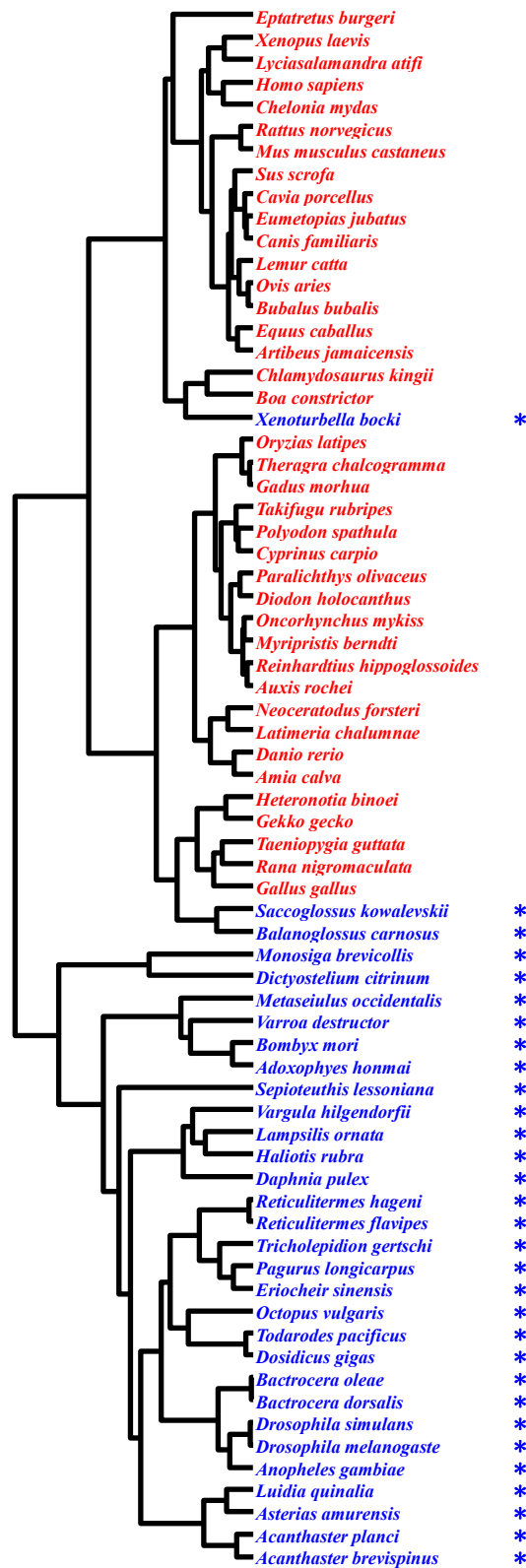
Figure 4. Phylogenetic tree generated using Ward's cluster analysis method (Ward, 1963) from predicted amino acid composition of complete mitochondrial genomes of 31 invertebrates and 39 vertebrates. High-C/G invertebrates and vertebrates in indicated in blue (*) and red (null), respectively
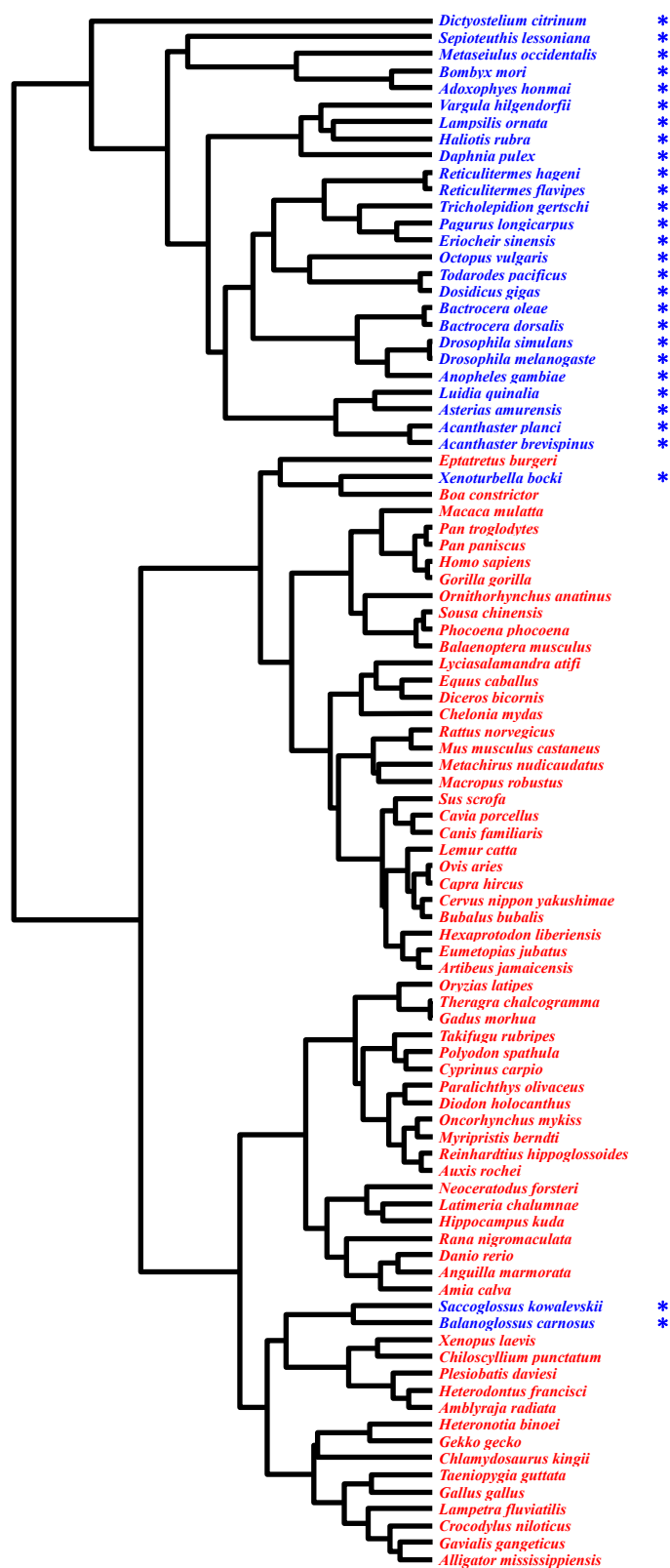
Figure 5. Phylogenetic tree generated using Ward's cluster analysis method (Ward, 1963) from predicted amino acid composition of complete mitochondrial genomes of 29 invertebrates and 63 vertebrates. High-C/G invertebrates and vertebrates are shown in blue (*) and red (null), respectively
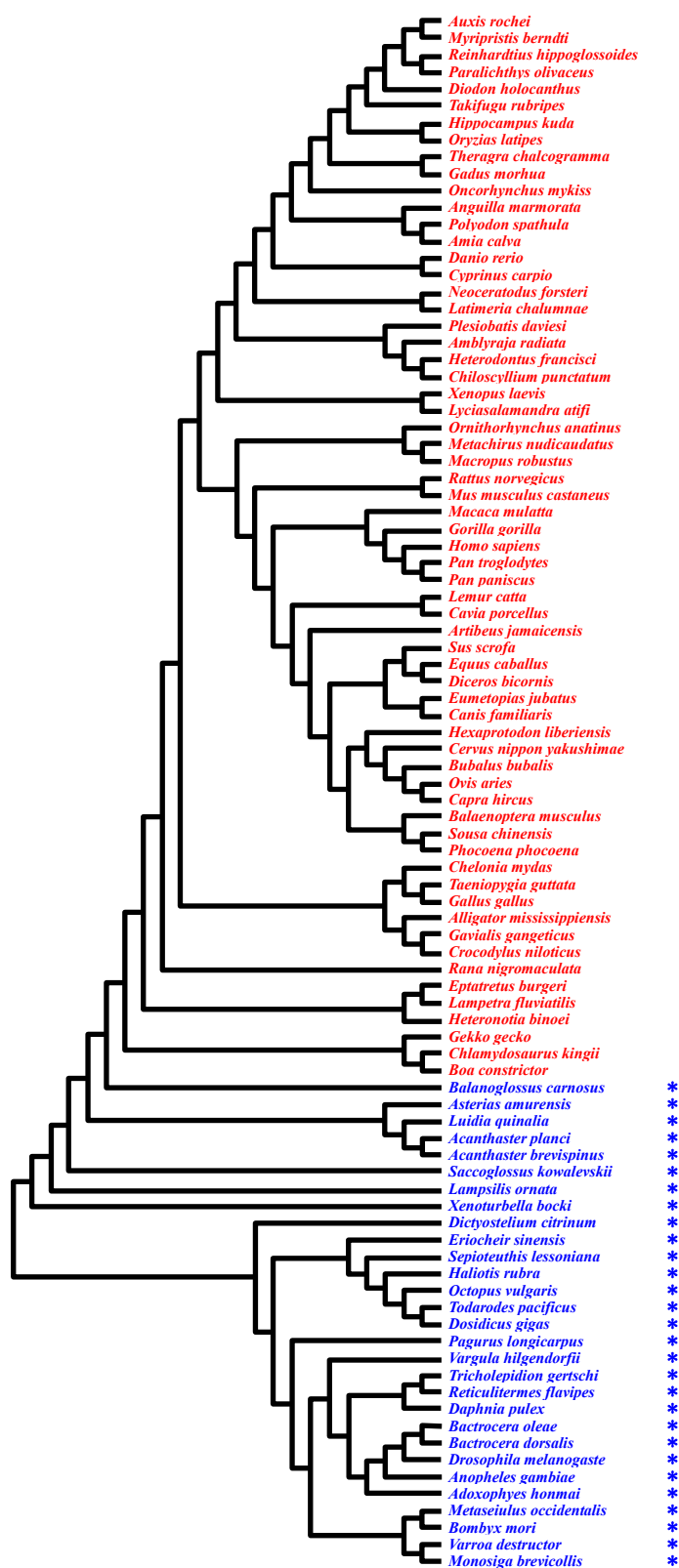
Figure 6. Phylogenetic tree generated using neighbor-joining (Saitou & Nei, 1987) from predicted amino acid composition of complete mitochondrial genomes. High-C/G invertebrates and vertebrates are designated by blue (*) and red (null), respectively

In our previous studies, additional vertebrates were deliberately selected to aid evaluation of reconstructed phylogenetic trees and were added to a set of randomly chosen vertebrates from an earlier study (Sorimachi & Okayasu, 2013). The number of vertebrate species in the expanded data set was increased to 63 in this fashion. Predicted amino acid compositions of these samples were analyzed using Ward's method, resulting in classification into three major clusters (Figure 5). High-C/G invertebrates form one major cluster, with vertebrates classified into two clusters. One vertebrate cluster consists of a large mammalian group and a small cluster that includes hagfish (*E. burgeri*), snake (*Boa constrictor*), and *X. bocki*; and the other comprises a large fish sub-cluster and various other sub-clusters including hemichordates, amphibians, reptiles, and birds. The Hemichordata are closely related to amphibians and reptiles. In this analysis, the hemichordates were separated from *X. bocki*, belonging to different clusters. All of these species fall into vertebrate groups, however, and are separate from invertebrates.

3.4 Neighbor-Joining Analysis Using Complete Mitochondrial Genome Nucleotide Sequences

Using nucleotide sequences of complete mitochondrial genomes, neighbor-joining analysis generated a different topology, as shown in Figure 6. Two groups, vertebrates and invertebrates, are evident, although clearly distinct major clusters were not obtained. *X. bocki* and hemichordates *B. carnosus* and *S. kowalevskii* are placed among the invertebrates, but do not form a separate cluster. They are apparently intermediate between vertebrates and invertebrates.

3.5 Neighbor-Joining Analysis Using 16S rRNA Gene Sequences

When 16S rRNA gene sequences were analyzed, the hemichordates and *X. bocki* were not found in the major cluster containing most of the invertebrates (Figure 7). According to the phylogenetic tree, these three species are closely related to one another, and also to *Lampsilis ornata* and *Metaseiulus occidentalis*. Hemichordata and *X. bocki* are more closely related to vertebrates than to invertebrates.

3.6 Phylogenetic Analysis of Vertebrates and Invertebrates

Using predicted amino acid composition from complete mitochondrial genomes, low-C/G invertebrates were also examined. We added 29 low-C/G invertebrate species to the data set consisting of 29 high-C/G invertebrates and 63 vertebrates. Using Ward's method, vertebrates and invertebrates were completely classified into two major clusters (Figure 8). In the resulting tree, *X. bocki* and Hemichordata (*S. kowalevskii* and *B. carnosus*) are placed into the vertebrate group. Hemichordata belongs to a fish sub-cluster, while *X. bocki* is grouped with *Chlamydosaurus kingii* and *B. constrictor* in the mammalian sub-cluster. Hagfish (*E. burgeri*) falls into the invertebrate group, in contrast to its placement among the vertebrates in our other analyses (Figures 4-7) and in our previous studies (Sorimachi et al., 2013a; Sorimachi & Okayasu, 2013).

Among the 7 species pairs in the set of 58 invertebrates analyzed, only three were paired in the phylogenetic tree generated from analysis of invertebrates alone (Figure 3). When 63 vertebrate species were analyzed along with 58 invertebrates, all 7 species pairs were found in the phylogenetic tree (Figure 8), even though increased sampling increases coincidental similarities and leads to worsening clustering results. These results support the appropriateness of Ward's method using amino acid composition. The results of cluster analysis of the data set consisting of 58 invertebrate and 63 vertebrate species are consistent with our previous results (Sorimachi et al., 2013a; Sorimachi & Okayasu, 2013).

4. Discussion

In our previous study (Sorimachi & Okayasu, 2003), we demonstrated that a universally representative phylogenetic tree is not currently available because of various factors affecting phylogenetic results. Increased sampling and reductions in the number of character states increase coincidental similarities, leading to worsening clustering results. Conversely, factors that reduce coincidental similarities yield better phylogenetic constructions. When amino acid composition was analyzed using two independent clustering methods, vertebrates and invertebrates were well-differentiated, even when the number of vertebrates was increased from 32 to 63 (Figures 3-8). In addition, hemichordates and *X. bocki* were placed with vertebrates, and were separated from invertebrates. Based on these results, *X. bocki* and hemichordates *B. carnosus* and *S. kowalevskii* are more closely related to vertebrates than to invertebrates.
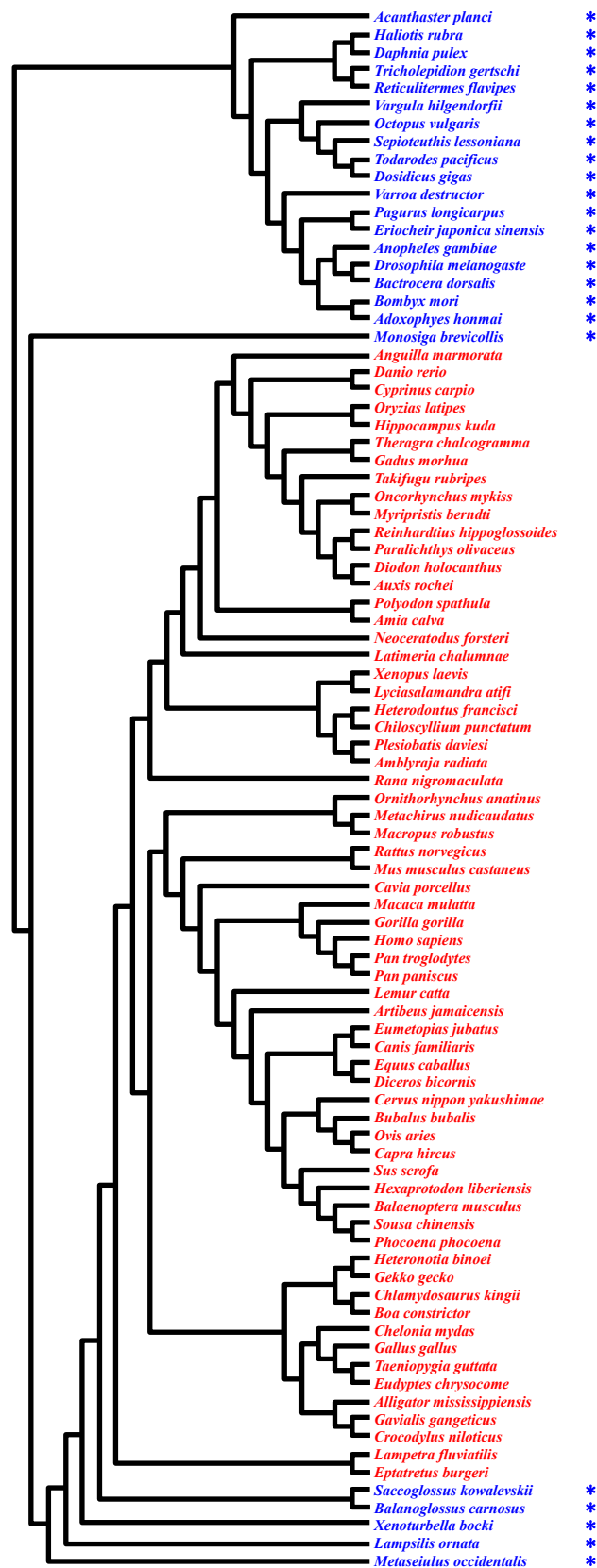
Figure 7. Phylogenetic tree generated using neighbor-joining (Saitou & Nei, 1987) from 16S rRNA gene sequences. High-C/G invertebrates and vertebrates are shown in blue (*) and red (null), respectively
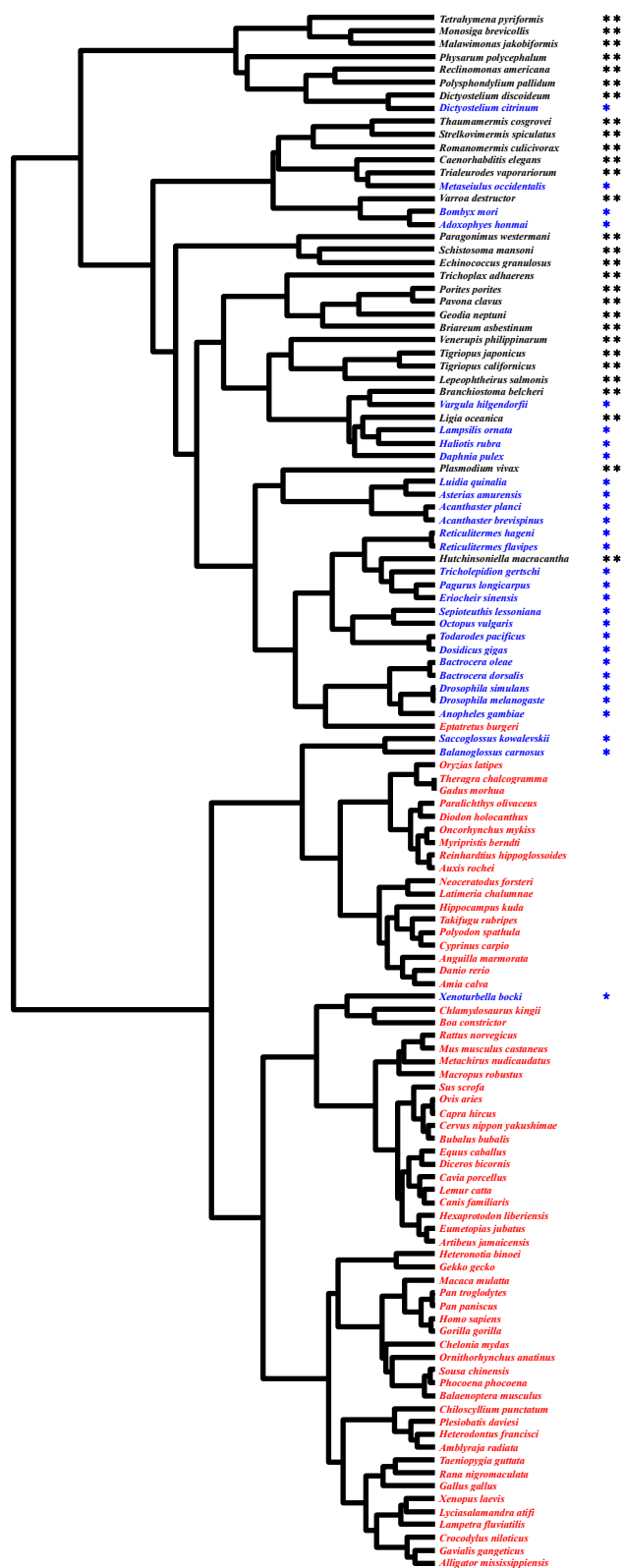
Figure 8. Phylogenetic tree generated using Ward's cluster analysis method (Ward, 1963) from predicted amino acid composition of complete mitochondrial genomes of 58 invertebrates and 63 vertebrates.High-C/G invertebrates, low-C/G invertebrates, and vertebrates are indicated in blue (*), black (**), and red (null), respectively

Neighbor-joining analysis of 16S rRNA gene sequences separated *X. bocki* and hemichordates from the invertebrate group and placed them within the vertebrates, whereas *L. ornata* and *M. occidentalis* were distinct from invertebrates (Figure 6). *X. bocki* and Hemichordata are obviously more closely related to vertebrates than invertebrates, although some exceptions were observed. In addition, neighbor-joining using predicted amino acid composition from mitochondrial genomes suggested that *X. bocki* and Hemichordata are closely related to vertebrates (Figure 7). Based on these results, *X. bocki* and hemichordates (*B. carnosus* and *S. kowalevskii*) are closely related to vertebrates.

The taxonomic position of the genus *Xenoturbella* has long been uncertain (Westblad, 1949). It was once thought to be closely related to mollusks (Noren & Jondelius, 1997), but this conclusion was due to DNA contamination from mollusks, a food source for *Xenoturbella* (Bourlat et al., 2006; Israelsson & Budd, 2006). The genus currently comprises its own phylum, and a close relationship with Acoelomorpha is strongly supported based on both morphological and molecular data (Lundkin, 1998; Raikova et al., 2000; Hejnol et al., 2009). A sister group relationship with echinoderms and hemichordates was also recently reported (Philippe et al., 2011). In our study, in contrast, *Xenoturbella* is closely allied with reptiles (*B. constrictor*) and hagfish (*E. burgeri*) (Figure 5), or to reptiles (*B. constrictor* and *C. kingii*) (Figure 4). These results are based on Ward's method using amino acid composition. According to cluster analyses using mixtures of vertebrates and invertebrates, *Xenoturbella* is more closely related to vertebrates than to invertebrates, falling into a large cluster consisting of all mammals. In addition, *Xenoturbella* is separated from another large cluster consisting of all fish, although hemichordates belong to this cluster. Ultimately, *Xenoturbella* is closely related to neither Echinodermata nor Hemichordata in our present study.

5. Conclusion

Among the invertebrates, the Hemichordata are the closest extant phylogenetic relatives of chordates, and have been described as a sister group to echinoderms such as the sea urchin (Philippe et al., 2011). In our study, consistent with this classification, hemichordates fall into a cluster consisting of echinoderms, mollusks, and arthropods (Figure 2). *Xenoturbella* is also contained in this cluster. Cluster analyses using samples of vertebrates and invertebrates, however, revealed that hemichordates are more closely related to vertebrates than to invertebrates, and that they are phylogenetically distant from *X. bocki* (Figures 4-8). Based on these results, both hemichordates and *Xenoturbella* are more closely related to vertebrates than to invertebrates, but do not fully belong in either group.

Refernces

Bourlat, S. J., Juliusdottir, T., Lowe, C. J., Freeman, R., Arnowicz, J., Kirschner, M., … Telford, M. J. (2006). Deuterostome phylogeny reveals monophletic chordates and the new phylum Xenoturbellida. *Nature, 444*, 85-88. http://dx.doi.org/10.1038/nature05241

Cobbett, A., Wilkinson, M., & Wills, M. (2007). Fossils impact as hard as living taxa in parsimony analyses of morphology. *Systems Biol., 17*, 753-766. http://dx.doi.org/10.1080/10635150701627296

Dayhoff, M. O., Park, C. M., & McLaughlin, P. J. (1977). Building a phylogenetic trees: cytochrome C. In M. O. Dayhoff (Ed.), *Atlas of protein sequence and structure* (Volume 5, pp. 7-16). Washington D. C.: National Biomedical Foundation.

DePouplana, L., Turner, R. J., Steer, B. A., & Schimmel, P. (1998). Genetic code origins: tRNAs older than their synthetases? *Proc. Natl. Acad. Sci. USA, 95*, 11295-11300. http://dx.doi.org/10.1073/pnas.95.19.11295

Doolittle, W. F., & Brown, J. R. (1994). Tempo, mode, the progenote, and the universal root. *Proc. Natl. Acad. Sci. USA, 91*, 6721-6728. http://dx.doi.org/10.1073/pnas.91.15.6721

Gatesy, J., Hayashi, C., Cronin, M. A., & Arctander, P. (1996). Evidence from milk caseinvgenes that cetaceans are close relatives of hipoppotamid artiodactyls. *Mol. Biol. Evol., 13*, 954-963. http://dx.doi.org/10.1093/oxfordjournals.molbev.a025663

Hejnol, A., Obst, M., Stamatakis, A., Ott, M., Rouse, G. W., Edgecombe, G. D., … Dunn, C. W. (2009). Assessing the root of lilaterian animals with scalable phylogenomic methods. *Proc. Royl. Soc. B, 276*, 426-4270. http://dx.doi.org/10.1098/rspb.2009.0896

Israelsson, O., & Budd, G. E. (2006). Egg and embryos in *Xenoturbella* (Phylum uncertain) are not ingested prey. *Dev. Gen. Evol., 215*, 358-363. http://dx.doi.org/10.1007/s00427-005-0485-x

Janvier, P. (2010). Micro RNAs revive old views about jawless vertebrate divergence and evolution. *Proc. Natl. Acad. Sci. USA, 107*, 19137-19138. http://dx.doi.org/10.1073/pnas.1014583107

Lundin, K. (1998). The epidermal ciliary roots of *Xenoturbella* bocki (Xenoturbellida) revisited: new support for a possible kinship with the Acoelomorpha (Platyhelminthes). *Zoologica Scripta, 27*, 263-270. http://dx.doi.org/10.1111/j.1463-6409.1998.tb00440.x

Maizels, N., & Weiner, A. M. (1994). Phylogeny from function: evidence from the molecular fossil record that tRNA originated in replication, not translation. *Proc. Natl. Acad. Sci. USA, 91*, 6729-6734. http://dx.doi.org/10.1073/pnas.91.15.6729

Noren, M., & Jondelius, U. (1997). Xenoturbell'a molluscan relatives. *Nature, 390*, 31-32. http://dx.doi.org/10.1038/36242

Okayasu, T., & Sorimachi, K. (2009). Organisms can essentially be classified according to two codon patterns. *Amino Acids, 36*, 261-271. http://dx.doi.org/10.1007/s00726-008-0059-0

Philippe, H., Brinkmann, H., Copley, R. R., Morz, L. L., Nakano, H., Poustka, A. J., … Telford, M. J. (2011). Acoelomorph flatworms are deuterostomes related to Xenoturbella. *Nature, 470*, 255-258. http://dx.doi.org/10.1038/nature09676

Raikova, O. I., Reuter, M., Jondelius, U., & Gustafsson, M. K. S. (2000). An immunocytochemical and ultrastructural study of the nervous and mucular systems of *Xenoturbella west blladi* (Bilateria inc. sed.). *Zoomorphology, 120*,107-118. http://dx.doi.org/10.1007/s004350000028

Saitou, N., & Nei, M. (1987). The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol. Biol. Evol., 4*, 406-425.

Sogin, M. L., Elwood, H. J., & Gunderson, J. H. (1986). Evolutionary diversity of eukaryotic small subunit rRNA genes. *Proc. Natl. Acad. Sci. USA, 83*, 1383-1387. http://dx.doi.org/10.1073/pnas.83.5.1383

Sorimachi, K. (2009). Evolution from primitive life to *Homo sapiens* based on visible genome structures: the amino acid world. *Natural Science, 1*, 107-119. http://dx.doi.org/10.4236/ns.2009.12013

Sorimachi, K. (2010). Genomic data provides simple evidence for a single origin of life. *Natural Science, 2*, 519-525. http://dx.doi.org/10.4236/ns.2010.25065

Sorimachi, K., Itoh, T., Kawarabayasi, Y., Okayasu, T., Akimoto, K., & Niwa, A. (2001). Conservation of basic pattern of cellular amino acid composition during biological evolution and the putative amino acid composition of primitive life forms. *Amino Acids, 21*, 393-399. http://dx.doi.org/10.1007/s007260170004

Sorimachi, K., & Okayasu, T. (2004). Classification of eubacteria based on their complete genome: where does Mycoplasmataceae belong? *Proc. R. Lond. B (Suppl.), 271*, S127-S130. http://dx.doi.org/10.1098/rsbl.2003.0141

Sorimachi, K., & Okayasu, T. (2008). Universal rules governing genome evolution expressed by linear formulas. *Open Genom. J., 1*, 33-43.

Sorimachi, K., & Okayasu, T. (2013). Phylogenetic tree construction based on amino acid composition and nucleotide content of complete vertebrate mitochondrial genomes. *IOSR J. Pharm., 3*, 51-60.

Sorimachi, K., Okayasu, T., Ohhira, S., Masawa, N., & Fukasawa, I. (2013a). Natural selection in vertebrate evolution under genomic and biosphere biases based on amino acid content: primitive vertebrate hagfish (Eptatretus burger). *Natural Science, 5*, 221-227. http://dx.doi.org/10.4236/ns.2013.52033

Sorimachi, K., Okayasu, T., Ohhira, S., Fukasawa, I., & Masawa, N. (2013b). Evidence for the independent divergence of vertebrate and high C/G ratio invertebrate mitochondria from the same origin. *Natural Science, 4*, 479-483. http://dx.doi.org/10.4236/ns.2012.47064

Ursing, B. M., & Arnason, U. (1998). Analyses of mitochondrial genomes strongly support a hipoppotumus-whale clade. *Proc. Royal Soc. B., 265*, 2251-2255. http://dx.doi.org/10.1098/rspb.1998.0567

Ward, J. H. (1963). Hierarchic grouping to optimize an objective function. *J. Amer. Statistic. Assoc., 58*, 236-244. http://dx.doi.org/10.1080/01621459.1963.10500845

Weisburg, W. G., Barns, S. M., Pelletier, D. A., & Lane, D. J. (1991). 16S ribosomal DNA amplification for phylogenetic study. *J. Bacteriol., 173*, 697-703.

Westblad, E. (1949). *Xenoturbella bocki* n.g., n.sp., a peculiar, primitive Tubellarian type. *Arkiv for Zoologi, 1*, 3-29.

Woese, C. R., & Fox, G. E. (1977). Phylogenetic structure of the prokaryotic domain: the primary kingdoms 2. *Proc. Natl. Acad. Sci. USA, 74*, 5088-5090. http://dx.doi.org/10.1073/pnas.74.11.5088

Zuckerkandl, E., & Pauling, L. B. (1962). Molecular disease, evolution, and genetic heterogeneity. In M. Kasha & B. Pullman (Eds.), *Horizons in Biochemistry* (pp. 189-225). New York: Academic Press.

Copyrights