Symbiosis, 22 (1997) 59–65
Balaban, Philadelphia/Rehovot

Xenorhabdus and *Photorhabdus*: Are they Sister Genera or are their Members Phylogenetically Intertwined?

E. STACKEBRANDT[1*], R.-U. EHLERS[2], and F.A. RAINEY[1]
[1]*DSMZ-German Collection of Microorganisms and Cell Cultures GmbH, Mascheroder Weg 1b, 38124 Braunschweig, Germany. Tel. +49-531-2616352, Fax. +49-531-2616418, E-mail. erko@gbf-braunschweig.de; and*
[2]*Abteilung Biotechnologie und Biologischer Pflanzenschutz, Institut für Phytopathologie, Christian-Albrechts-Universität, Klausdorfer Str. 28–36, 24223 Klausdorf, Germany*

Received May 15, 1996; Accepted July 25, 1996

Abstract

The evolution of an organism can be considered as the sum or mean of the evolution of its genes. Thus, the restriction of the phylogenetic analysis to a single molecule, such as the most widely used 16S rDNA, highlights the evolution of the molecule but not necessarily the evolution of the organisms from which the molecule was isolated for analysis. For a few groups of bacteria more than only a single conservative molecule has been analysed; most of these data sets agree with each other and even provide more insight into the evolution and phylogeny of the organisms. For the majority of bacterial taxa, however, genealogical patterns are exclusively based on a single molecule and support for the order of lineages must be derived from phenotypic properties of its members. As additional factors have been identified that influence the topology of the branching pattern, the stability of a phylogenetic branching pattern need to be assessed carefully before taxonomic conclusions are made.

Keywords: *Xenorhabdus*, *Photorhabdus*, phylogeny, 165 rDNA

*The author to whom correspondence should be sent.

0334-5114/97/$05.50

1. Introduction

The genus *Photorhabdus* gen. nov. and *P. luminescens* comb. nov has recently been described to encompass *Xenorhabdus luminescens* (Boemare et al., 1993). Strains of this species were considered to be sufficiently unrelated to *Xenorhabdus nematophilus*, *X. poinarii*, *X. bovienii* and *X. beddingii* to warrant individual genus status. This conclusion was based on low interspecies DNA reassociation values and phenotypic and chemotaxonomic properties. Previous 16S rRNA cataloguing data on *X. nematophilus* and *X. (Photorhabdus) luminescens* showed these two species to form an individual branch within the family Enterobacteriaceae (Ehlers et al., 1988). As no other Xenorhabdus species were analysed in that study, the question as to whether or not the two species might belong to two different genera could not be answered. In this communication we discuss the problems of unambiguous phylogenetic placement of organisms when one of the groups under investigation appears to be subjected to a higher rate of evolution than the phylogenetic nearest neighbors.

2. Methods

The almost complete 16S rDNA sequence was determined for all but one of the validly described species of *Xenorabdus* and *Photorhabdus* (Rainey et al., 1995) and the sequences were compared to the existing 16S rDNA database of members of the Enterobacteriaceae and other members of the gamma subclass of *Proteobacteria* (Maidak et al., 1994). Dissimilarity values were transformed into phylogenetic distance values that compensate for multiple substitutions at any given site in the sequence (Jukes and Cantor, 1969). It must be kept in mind that the number of possible characters at a particular site is limited and that only contemporary sequences are available. As a consequence, the true extent of exchanges, namely evolutionary events, cannot be fully recognised and are underestimated. The number of unrecognized evolutionary events increase not only with the dissimilarity values but also in case where one subset of sequences is subjected to tachytelic (fast evolving) evolution. Phylogenetic trees were reconstructed by applying additive treeing methods, such as Neighbor Joining (Saitou and Nei, 1987) and the algorithm of De Soete (1983) using the corrected dissimilarity values. One thousand bootstrap values of 920 polymorphic sites were calculated to test the stability of the Neighbor-Joining tree.

3. The Generation and Interpretation of Phylogenetic Trees

Most readers of scientific publications that contain phylogenetic trees generated on the basis of nucleic acid sequences, especially 16S rDNA or rRNA, are often not aware of the fact that tree topologies are influenced by a significant number of variables, examples of which are listed below. While topologies are to some extend influenced by intrinsic features of treeing algorithms and the nucleotide sequence, other factors can be introduced by the operator.

The influence of algorithms (Li and Graur, 1991)

(i) All available treeing methods are based on assumptions (for example, that sites evolve independently) which may not be true with real data.

(ii) Programs usually do not perform tests of all possible tree topologies and therefore may fail to find the optimal tree. The reason being the enormous number of possible tree topologies (about 8×10^{20} different trees can be generated from as little as 20 sequences) and the expense in computing time.

(iii) Since most programs add and treat the data according to their input order to save computing time a bias of the topology should be eliminated by changing the order of entries randomly and performing several runs.

The influence of the nucleotide sequence (Woese et al., 1991)

(i) In general, every new sequence which is different from those in the existing data set confers additional data on the information content of individual sequence positions as well as on the characteristics defining phylogenetic groups. The inclusion of this additional information may change the tree locally or even globally.

(ii) As it cannot be assumed that the rates of changes even at homologous positions within the rRNA species is similar in phylogenetically equivalent taxa, comprehensive phylogenetic analyses of remotely related sequences should always be based on complete sequences. Obviously, complete sequences provide the maximum of available phylogenetic information.

(iii) However, the inclusion of all positions of a sequence does not warrant optimal results, i.e., reflection of the true course of evolution. It is the level of relatedness one wants to determine that decides about the selection of which nucleotide stretch to use. Invariant and conserved residues indicate homology and are useful for the alignment of the sequences, but confer little phylogenetic information. As individual positions of the other nucleotides change at different rates, phylogenetic analyses of the other stretches can be used to

cover a wide range of phylogenetic distances. More variable positions define closer related groups, but are not informative for lower levels of phylogenetic relationships. Rapidly changing positions report on more recent events, but due to multiple base changes (homoplasy) their inclusion in the analysis may add random noise especially concerning the deeper branches.

(iv) Another potential source of treeing artefacts applying distance and parsimony methods is compositional bias. The G+C contents of rRNAs in general vary within a narrow margin between 53 and 55 mol%. The values are usually higher in molecules from thermophilic bacteria (up to 63%) but not exclusively so. However, the purine content is rather stable among bacterial rRNAs, allowing certain programs such as transversion analysis to compensate for this bias. Unfortunately the information content of the sequences is reduced due to the two character states (purines and pyrimidines), and therefore, there is an increase of random identity.

The influence of the operator

(i) Subtrees can be rooted by including sequences from moderately related organisms which group deeper than the group under investigation. In order to recognise artefacts which may be again due to "false" sequence identities between the outgroup and part of the ingroup organisms, outgroup sequences should be included from different phylogenetic entities.

(ii) Besides the outgroup references, the addition of any new homologous sequence to an existing data set influences the derived tree and may change its topology.

(iii) Lineages represented by a single organism often cannot be stably positioned in phylogenetic trees. The addition of related sequences usually stabilises the branching point of the respective lineage. The addition of incomplete or incorrect sequences may however produce artefacts and reduce the accuracy of the tree.

(iv) The reduction of sequence information to either partial sequences of one to a few hundred nucleotides or to only variable regions significantly reduces the stability of phylogenetic trees. This information is suitable for the purpose of identification but fails to reflect phylogenies (Stackebrandt et al., 1992). Although most changes are indeed located in the variable and hypervariable regions the number of changes occurring in the more conserved stretches is significant.

4. The Phylogenetic Relatedness of *Xenorhabdus* and *Photorhabdus*

The phylogenetic tree presented during the Cost Workshop in Debrecen, May 1994, was based on the first 500 nucleotides of the type strains of the species of *Xenorhabdus* and *Photorhabdus*. As indicated during the symposium and explained above it is not surprising that the position of these two genera within the tree of Enterbacteriaceae change when the analysis is based on almost complete sequences of 1500 nucleotides [see point 3 (iv)]. Similarity values for the *Xenorhabdus* species range between 96. and 97.7 %, while the values for these organisms and *P. luminescens* range between 94.1 and 95.5%. These symbiotic bacteria and other members of Enterobacteriaceae show similarity values between 92.3 and 95.2%. *Vibrio parahaemolyticus*, selected as the outgroup reference for the enterics, has about 4 to 6% lower similarity values.

5. Determination of the Phylogenetic Position between Species of *Xenorhabdus* and *Photorhabdus*

Both distance matrix methods used in this study place members of *Xenorhabdus* and *Photorhabdus* on a single line of descent separate from other members of the family Enterobacteriaceae. This clustering is supported by 100% bootstrap values and by several nucleotides found to be unique for these organisms when compared to other members of Enterobacteriaceae. It is interesting to note that the G+C content of 16S rDNA of both genera of the symbionts of entomopathogenic nematodes is about 2 to 3% higher than those of their phylogenetic neighbors. The four *Xenorhabdus* species are related at the 65% confidence level. While *X. bovienii* and *X. poinarii* form a phylogenetically stable pair (93%), all other combinations of pairwise relationships are less significant. This is in contrast to results of DNA reassociation values which found *X. nematophilus* and *X. beddingii* to be 31% related. *Proteus vulgaris* is the nearest neighbor of these nematode symbionts.

In contrast to these data which would support the description of Photorhabdus as a genus separate from *Xenorhabdus* (Boemare et al., 1993), the majority of phylogenetic trees generated from the sequence data do not indicate these two genera to be sister groups. From a phylogenetic point of view only those taxonomic entities should be treated as individual genera when the type species of either genus is placed outside the radiation of the other genus. Surprisingly, even the neighbor joining tree, for which the bootstrap values indicate the coherency of the *Xenorhabdus* species to the exclusion of *P. luminescens*, places the branching point of this organism within the radiation of *Xenorhabdus* species. In order to exclude factors raised in point 3 (i), (ii), and

(iv) the analysis was performed using almost complete sequences, and by changing the number and origin of reference and outgroup organisms. Only in one tree, in which sequences of *Buchnera*, which are symbionts of aphids (Munson et al., 1991), enterobacteria and two rooting sequences were included, *P. luminescens* was placed outside the radiation of *Xenorhabdus* species and next to *Proteus vulgaris*. In all other trees ,based on different combinations of sequences, *P. luminescens* was found about equidistantly related to the four *Xenorhabdus* species. While in each tree *X. nematophilus* and *X. beddingii* were found to cluster together, the position of *X. poinarii* and *X. bovienii* changed with the composition of the reference and outgroup organisms.

The influence of points raised under 2 (i) and (iv) can most likely be excluded. The position of *P. luminescens* within the radiation of *Xenorhabdus* species is rather stable and not greatly influenced by either the treeing algorithm, reference sequences or the base composition of rDNA which is the same for all entomopathogenic symbionts. The only explanation we have at present for the inability to determine a phylogenetically stable place for P. luminescens is the possibility that this species evolves slightly more rapidly than its phylogenetic neighbors. Support for this hypothesis originates from the finding that the similarity values for this bacterium and the reference organism is about 2% lower than those found for *Xenorhabdus* species and the reference organisms (which, as outlined above, should be very similar in comparison to outgroup sequences). At the nucleotide level this idea is supported by the presence of several absolutely unique positions which are conserved for members of the gamma subclass of *Proteobacteria*.

The question whether *Photorhabdus* and *Xenorhabdus* constitute two separate genera, as described by Boemare et al. (1993), or members of the same genus can at present not be answered. The inter- and intrageneric DNA-DNA similarity values are so low for the majority of species that this approach offers no help. The same is true for the 16S rDNA similarity values and the topology of phylogenetic trees which fail to convincingly place these organisms in the phylogenetic trees. The strategy that should be considered next is to increase the stability of the tree by adding more sequences of members of the two species [see point 3 (iii)], especially more strains of *P. luminescens*. If, as expected, members of *P. luminescens* are slightly more rapidly evolving, then this should be reflected and detectable in a greater variation of those nucleotide positions considered to be conserved in neighboring taxa.

A recent study on the 16S rDNAs of a large number of *P. luminescens* strains and the type strains of all valid *Xenorhabdus* species, including *X. japonicus*, has impressively demonstrated the influence of strain numbers and selection of strains on the topology of the 16S rDNA tree (Szallas et al., 1996). This tree now shows a clear and statistically significant separation of the two genera

Photorhabdus and *Xenorhabdus* with all members of the two genera forming phylogenetically coherent entities.

REFERENCES

Boemare, N.E., Akhurst, R.J., and Mourant, R.G. 1993. DNA relatedness between *Xenorhabdus* spp. (*Enterobacteriaceae*), symbiotic bacteria of entomopathogenic nematodes, and a proposal to transfer *X. luminescens* to a new genus, *Photorhabdus* gen.nov. *International Journal of Systematic Bacteriology* **43**: 249–255.

De Soete, G. 1983. A least squares algorithm for fitting additive trees to proximity data. *Psychometrika* **48**: 621–626

Ehlers, U.-R., Wyss, U., and Stackebrandt, E. 1988. 16S rRNA cataloging and the phylogenetic position of *Xenorhabdus*. *Systematic and Applied Microbiology* **10**: 121–125.

Jukes, T.H. and Cantor, C.R. 1969. Evolution of protein molecules. In: *Mammalian Protein Metabolism*, H.N. Munro, ed. Academic Press, New York, pp. 21–132.

Maidak, B.L., Larsen, N., McCaughey, J., Overbeek, R., Olsen, G.J., Fogel, K., Blandy, J., and Woese, C.R. 1994. The ribosomal database project. *Nucleic Acids Research* **22**: 8483–3487.

Li, W.-H. and Graur, D. 1991. *Fundamentals of Molecular Evolution*. Sinauer Associates, Inc., Sunderland, MA.

Munson, M.A., Baumann, P., and Kinsey, M.G. 1991. *Buchnera* gen. nov.and *Buchnera aphidicola* sp. nov., a taxon consisting of the mycetocyte-associated, primary endosymbionts of aphids. *International Journal of Systematic Bacteriology* **41**: 566–568.

Rainey, F.A., Ehlers, R.U., and Stackebrandt, E. 1995. Inability of the polyphasic approach to systematics to determine the relatedness of the genera *Xenorhabdus* and *Photorhabdus*. *International Journal of Systematic Bacteriology* **45**: 379–381.

Saitou, N. and Nei, M. 1987. The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Molecular Biological Evolution* **4**: 406–425.

Stackebrandt, E., Liesack, W., and Witt, D. 1992. Ribosomal RNA and ribosomal DNA sequence analyses. *Gene* **115**: 255–260.

Woese, C.R., Achenbach, L . Rouviere, P., and Mandelco, L. 1991. Archaeal phylogeny: reexamination of the phylogenetic position of *Archaeoglobus fulgidus* in the light of certain composition-induced artifacts. *Systematic and Applied Microbiology* **14**: 364–371.