Probable synonymy of the nitrogen-fixing genus *Azotobacter* and the genus *Pseudomonas*

J. M. Young and D.-C. Park

Landcare Research, Private Bag 92170, Auckland, New Zealand

The relationships of the genus *Azotobacter*, *Azomonas macrocytogenes* and the genus *Pseudomonas* were revealed by comparative analysis of partial 16S rRNA and *atpD*, *carA* and *recA* gene sequences and as concatenated nucleotide and peptide sequences. Sequence similarities of *Azotobacter* species and *Azomonas macrocytogenes* indicated that these may be considered to be synonyms at the molecular level. In addition, these species show an intimate relationship with species of *Pseudomonas*, especially *P. aeruginosa* (the type species of the genus). In terms of the current circumscription of the genus *Pseudomonas*, *Azotobacter* and *Azomonas macrocytogenes* should be considered for amalgamation with *Pseudomonas*. *Azotobacter* and *Azomonas* comprise nitrogen-fixing strains with large pleomorphic cells that form cysts, and peritrichous flagella insertion; characteristics not included in the current circumscription of *Pseudomonas*. The data are discussed in the light of whether lateral transfer of genes could be involved in the determination of significant morphological characteristics, thus leading to a problem that may be encountered more frequently: how to resolve classification of taxa based on conserved sequences with those based on their phenotype. More fundamentally, the results illuminate problems that will increasingly be encountered: by what criteria can taxa be delineated, what are the most appropriate methods for classification, and what are the proper assumptions of bacterial classification?

**INTRODUCTION**

The genus *Pseudomonas* has been circumscribed as a heterogeneous taxon (Palleroni, 1984). Revisions of the genus based on comparative analyses of 16S rRNA gene sequence data have resulted in the transfer of many species previously included in *Pseudomonas*, usually to novel genera proposed in the *Alpha*-,- *Beta*- and *Gamma*-proteobacteria (Kersters et al., 1996). *Pseudomonas* now includes species in a clade in the *Gammaproteobacteria*. Relocation of taxa within *Pseudomonas* (sensu stricto) has been indicated by comparative analyses of 16S rRNA gene sequence data (Moore et al., 1996; Anzai et al., 2000) and on combined data of gyrase B and the rho factor (Yamamoto et al., 2000). Anzai et al. (2000) noted a close relationship between *Pseudomonas*, *Azotobacter* and *Azomonas*. *Azotobacter* and *Azomonas* are specifically circumscribed as genera containing species capable of fixing nitrogen in aerobic conditions (Kennedy & Rudnick, 2005; Kennedy et al., 2005). *Pseudomonas* is not recorded as having this capability (Palleroni, 2005) although this has been reported for a limited number of species, some of which were of doubtful identification (Line, 1997; Vermeiren et al., 1999; Kulakov et al., 2002).

A preliminary comparison of 16S rRNA gene sequences suggested a much closer relationship between *Azotobacter beijerinckii* and *Pseudomonas* species than was previously indicated by their inclusion in separate families of the order *Pseudomonadales* (Palleroni, 1984), now as separate genera in the family *Pseudomonadaceae* (Palleroni, 2005). These considerations led to a re-evaluation of the relationships using partial sequences of the catalytic subunit of the ATP synthase (*atpD*), the small subunit of carbamoyl phosphate synthase (*carA*), recombinase A (*recA*) and the 16S rRNA gene. Sequences from *Azomonas macrocytogenes* (the type species *Azomonas agilis* was not available to us) and *Azotobacter* species were compared with those representing *Pseudomonas*. Representative strains of *Azospirillum* species were included to confirm the classification of this genus also.

**Abbreviations**: ML, maximum-likelihood; MLSA, multilocus sequence analysis; NJ, neighbour-joining.

The GenBank/EMBL/DDBJ accession numbers for the 16S rRNA, *atpD*, *carA* and *recA* gene sequences determined in this study are given in Fig. 1 and Supplementary Figs S1 and S2.

All authority references for names can be found at Euzeby (1997).

Figures showing the comparative analysis of individual *atpD*, *carA* and *recA* nucleotide and peptide sequences of *Azomonas*, *Azotobacter* and *Pseudomonas* species using the maximum-likelihood and neighbour-joining algorithms are available as supplementary material with the online version of this paper.
METHODS

Strains used. Strains of Azomonas, Azospirillum, Azotobacter and Pseudomonas species were obtained from the International Collection of Microorganisms from Plants (ICMP), Landcare Research, Auckland. Information on these strains is given at the ICMP website: http://www.landcareresearch.co.nz/research/biodiversity/fungiprog/icmp.asp

DNA extraction. Strains were cultured on nutrient agar (Difco) incubated at 27 °C for 2–3 days. Genomic DNA was extracted directly from colony growth on nutrient agar using phenol and CTAB (hexadecyltrimethylammonium bromide) (Ausubel et al., 1987).

Amplification and sequencing. The primers and conditions used for PCR amplification of atpD, carA and recA and 16S rRNA gene sequences as given by Young & Park (2007) were followed. PCR products were column-purified with a Roche High Pure PCR Product Purification kit. In cases where more than one DNA band occurred on the gel, the DNA fragment of the correct size was extracted using the PCR purification kit. Purified products were cycle sequenced with the appropriate primers using BigDye Terminator Ready Reaction Mix v3.1 (ABI) and sequences were obtained in both directions with an ABI 3100 Avant Genetic Analyzer. Sequences were assembled and edited with Sequencher 4.5 (Gene Codes Corp.).

Sequence analysis. All 16S rRNA, atpD, carA and recA gene sequences for Azomonas, Azospirillum, Azotobacter and Pseudomonas species included in this study have been registered in GenBank. GenBank accession numbers for the 16S rRNA genes are shown in Fig. 1. GenBank accession numbers for individual gene sequences are shown in Supplementary Figs S1 and S2 in IJSEM Online. The outgroups for alignments of the 16S rRNA, atpD, carA and recA genes were the type strain of Escherichia coli (ICMP 15663T) and Xylella fastidiosa clone 9a5c, obtained directly from GenBank.

Nucleotide sequences were aligned with CLUSTAL_X 1.83 with default parameters and both ends of each alignment were trimmed to the following sizes: 16S rRNA gene, 1386 positions; atpD, 793 positions; carA, 630 positions; recA, 598 positions.

Peptide sequences for atpD, carA and recA were determined using GeneDoc, aligned using CLUSTAL_X and edited manually, again using GeneDoc, to ensure that alignment gaps in protein-coding genes did not cause errors in the amino acid sequence. Calculated peptide sequence lengths for atpD, carA and recA were 263, 209 and 198 amino acids respectively.

Individual nucleotide and peptide sequences were compared using PAUP* version 4.0b10 with the neighbour-joining (NJ) algorithm. For the distance measure, an uncorrected ‘p’ distance and Jukes–Cantor model were used.

The atpD, carA and recA sequences were concatenated and trimmed to give a total length of 2070 positions, which resulted in a concatenated peptide length of 670 amino acids.

Concatenated sequences were treated using Gblocks 9.1b (http://molevol.ibmb.csic.es/Gblocks.html) (Castresana, 2000), which eliminates poorly aligned positions and divergent regions of an alignment of DNA sequences; positions which may not be homologous or may have been saturated by multiple substitutions. Because the atpD sequences for Pseudomonas putida and Pseudomonas stutzeri were not available, a non-stringent option was used to allow inclusion of the sequences from other strains. Concatenated nucleotide and peptide sequences were also analysed using the NJ algorithm as described above and maximum-likelihood (ML) using PHYML Online (Guindon et al., 2005) with the default models for nucleotides (HYK85) and for peptides (Blosum62).

RESULTS

16S rRNA gene sequences

The comparative analysis of 16S rRNA gene sequences using the NJ algorithm produced a tree (Fig. 1) with sequences representing Azospirillum (Azospirillum brasilense and

---

Fig. 1. Comparative analysis of 16S RNA gene sequences of Azomonas (Azm.), Azospirillum (Azm.), Azotobacter (Azt.) and Pseudomonas species, using the NJ algorithm. Bar, number of substitutions per site. Accession numbers of sequences are held in GenBank. Nucleotide sequences marked with an asterix (*) are derived from type strains. Sequences EF100149* and EF100150* refer to Azospirillum brasilense (ATCC 29145) and Azospirillum lipoferus (8672 and ATCC 29707), respectively.
**Azospirillum lipoferum** as outliers to a complex cluster representing *Pseudomonas*, *Azomonas* and *Azotobacter*. Sequences representing *Pseudomonas* formed two groups: group 1, comprising *P. aeruginosa*, *P. fluorescens*, *P. mendocina*, *P. putida*, *P. resinovorans* and *P. stutzeri*; and group 2, comprising *P. agarici*, *P. aureofaciens*, *P. aurantiaca*, *P. chlororaphis*, *P. cichorii*, *P. fluorescens*, *P. syringae* and *P. tolaasii*. Sequences representing *Azotobacter* and *Azomonas macrocytogenes* were grouped together within rRNA group 1 of *Pseudomonas*, being closest to *P. aeruginosa* and *P. resinovorans*.

**Individual atpD, carA and recA gene and peptide sequences**

Nucleotide sequences of *atpD*, *carA* and *recA* were obtained for all strains except those of *Azospirillum* species. The online version of this paper contains supplementary figures showing results of analyses of individual nucleotide and peptide sequences based on ML (Supplementary Fig. S1) and NJ (Supplementary Fig. S2) algorithms.

**atpD sequences.** Nucleotide sequences divided into two discrete groups. All sequences representing *Azotobacter* and *Azomonas macrocytogenes* were present in one group together with *P. aeruginosa* and *P. resinovorans*. The remaining sequences of *Pseudomonas* species formed a second group. Peptide sequences produced a tree in which sequences of *Azotobacter* and *Azomonas macrocytogenes* were outliers to *Pseudomonas*.

**carA sequences.** Nucleotide sequences of *Azomonas macrocytogenes* were outliers to a complex *Azomonas–Azotobacter–Pseudomonas* structure in which *Azotobacter* sequences formed a discrete group embedded among *Pseudomonas* sequences. Peptide sequences produced a tree in which *Pseudomonas* was divided into the two groups expressed by the 16S rRNA gene, with sequences representing *Azotobacter* and *Azomonas macrocytogenes* included occurring separately in rRNA group 1.

**recA sequences.** Nucleotide sequences of *Azomonas macrocytogenes* were outliers to a complex *Azotobacter–Pseudomonas* structure in which *Azotobacter* species sequences formed a discrete group embedded within the *Pseudomonas* complex, adjacent to *P. aeruginosa*. The peptide tree has a different structure, with *Azomonas macrocytogenes* included with *Azotobacter* sequences and *P. stutzeri* as nearest neighbour. These are embedded in sequences identified with rRNA group 1.

There was considerable minor variability between tree structures of individual *Pseudomonas atpD*, *recA* and *carA* nucleotide and peptide sequences, whether analysed using NJ or ML, but in general terms, individual nucleotide and peptide trees were similar (see Supplementary Figs S1 and S2 in IJSEM Online). For *Azomonas macrocytogenes*, the *atpD* sequence was similar to those of *Azotobacter*, but *carA* and *recA* nucleotides were outliers to those of *Azotobacter–Pseudomonas* and the *carA* peptide was nearest neighbour to *P. stutzeri* and *P. fluorescens* sequences. The *recA* peptide was grouped with all *Azotobacter* sequences.

Removal of *Azomonas macrocytogenes* nucleotide and peptide sequences from the analyses or removal of all *atpD* data resulted in trees with nitrogen-fixing species embedded in *Pseudomonas*, with *P. aeruginosa* as nearest neighbour (data not shown).

**Concatenated sequences**

Analyses of concatenated nucleotide and peptide data treated using Gblocks indicated that <2% of the sequences might be poorly aligned or hypervariable regions. These regions were considered to be so small that all sequence data were included. Analyses of concatenated sequences are shown in Fig. 2. The concatenated nucleotide and peptide sequences analysed using the NJ algorithm both produced trees in which *Azotobacter* species and *Azomonas macrocytogenes* were outliers to *Pseudomonas* species. The ML algorithm applied to nucleotide and peptide sequences resulted in trees in which *Azotobacter* and *Azomonas macrocytogenes* lay within the *Pseudomonas* ‘aeruginosa’ cluster. The tree comparing nucleotide sequences was not very well resolved, but *Azotobacter* and *Azomonas macrocytogenes* sequences adjoined *P. aeruginosa* and *P. resinovorans*. The peptide tree reproduced the ‘aeruginosa’ and ‘fluorescens’ clusters, with *Azotobacter* and *Azomonas macrocytogenes* sequences adjoining *P. aeruginosa* and *P. resinovorans*.

**DISCUSSION**

Three 16S rRNA gene sequences from species representing the genus *Azospirillum*, *Azospirillum brasilense* and *Azospirillum lipoferum*, were outliers to all other sequences, including *E. coli*. This is consistent with the placement of *Azospirillum* in the family *Rhodospirillaceae* in the *Alphaproteobacteria* (Pfennig & Trüper, 2005), therefore unrelated to *Pseudomonas* and the other nitrogen-fixing genera under consideration, and the genus is not considered further in this study.

The analysis of 16S rRNA gene sequences of *Pseudomonas* produced the two groups corresponding to the ‘aeruginosa’ cluster and ‘fluorescens’ cluster reported previously (Hilario et al., 2004; Moore et al., 1996). [As well as comparing 16S rRNA gene sequences, Hilario et al. (2004) concatenated the rRNA gene together with three genes, without appreciating that the rRNA gene has regions that are continuously conserved, whereas degeneracy in codons means that genes are more variable and therefore these different forms of sequence make incommensurable contributions when they are concatenated.] The nucleotide and peptide sequence analyses reported here also gave general support for this subdivision. However, these studies have included relatively small representations of
the approximately 110 species now included in the genus (Euzeby, 1997). Where larger numbers have been compared using different nucleotides for comparison, this differentiation is not so obvious and may be obliterated or result in subdivision into more groups (Anzai et al., 2000; Ait Tayeb et al., 2005; Yamamoto et al., 2000). The subgeneric groups of Moore et al. (1996) are used here only as points of reference for discussion of analyses of the other sequences (Young et al., 1992).

Sequences of Azotobacter species consistently formed a single heterogeneous group within the ‘aeruginosa’ cluster, usually very close to P. aeruginosa and P. resinovorans. The nucleotide carA and recA sequences of Azomonas macro-

Figure 2. Comparative analysis of concatenated atpD, carA and recA nucleotide and peptide sequences of Azomonas (Azm.), Azotobacter (Azt.) and Pseudomonas species, using the NJ algorithm (a) and the ML algorithm (b). Accession numbers of strains held in the International Collection of Microorganisms from Plants are specified. Bars, number of substitutions per site.
cytogenes were outliers to Pseudomonas but the derived peptide sequences were embedded in Pseudomonas. This can be attributed to the similarity between individual sequences of these genes and significant differences between nucleotides and peptides due to degeneracy in the second and third bases of codons. The influence of Azomonas macrocytogenes sequences appears to account for the different tree structures of the concatenated sequences in which Azotobacter and Azomonas macrocytogenes sequences are outliers to those of Pseudomonas. In general, however, individual and concatenated nucleotide and peptide sequence data with the 16S rRNA gene indicate a close relationship between the three genera, with Azotobacter species and Azomonas macrocytogenes being members of Pseudomonas.

Relationship of Azotobacter and Azomonas

The description of the genus Azotobacter (Kennedy et al., 2005) is of coccoid to rod-shaped cells that are up to 2.0 μm in diameter and 4 μm in length, but can be filamentous, up to 60 μm in length; if motile then by peritrichous flagella, and are capable of fixing nitrogen in aerobic conditions. Cells can form cysts. Thompson & Skerman (1981) proposed a separate genus for Azotobacter paspali, as Azorhizophilus paspali, but Kennedy et al. (2005) did not recognize this proposal and included the species in Azotobacter. The description of Azomonas (Kennedy & Rudnick, 2005) is of ellipsoidal to rod-shaped cells that are more than 2 μm in diameter, usually 2.5–3.5 μm in length, motile by peritrichous, lophotrichous or polar flagella, produce water-soluble fluorescent pigments, and are capable of fixing nitrogen under aerobic conditions. A consideration of the phenotypic determinative characteristics of the genera supports the synonymy of these genera as indicated by the sequence data. An investigation of these two genera, with a wider selection of strains, including the type species of Azomonas (Azomonas agilis), and perhaps based on other genes, might support their synonymy. If they are established as a single genus, Azotobacter would take priority over Azomonas.

Synonymy of Azotobacter with Pseudomonas?

At the molecular level, the similarity of Azotobacter species housekeeping gene and 16S rRNA gene sequences with those of Pseudomonas point to an intimate relationship between these taxa. Rediers et al. (2004), who studied a large selection of genes, concluded that Azotobacter vinelandii was very closely related to Pseudomonas. Kennedy et al. (2005) also noted the close relationship of these genera. Insofar as the peptide trees give expression to phenotypic relationships these also confirm the intimate relationship of the genera although these are not always congruent with the nucleotide trees. Although Palleroni (2005) did not describe nitrogen fixation as a characteristic in Pseudomonas, Chan et al. (1994) suggested that this capability could be widespread in the genus. That it has not been confirmed widely is perhaps because there have been few comprehensive investigations. It was reported for P. stutzeri (Chan et al., 1994; Vermeiren et al., 1999). Barraquio et al. (1986) and Kulakov et al. (2002) attributed a nitrogen-fixing capability to P. saccharophila, but this species has been transferred to a new and unrelated genus, Pelomonas (Xie & Yokota, 2005). Line (1997) recorded P. acidivorans, P. alcaligenes and P. aureofaciens as nitrogen-fixing species, but identifications were made using determinative methods that might not have given reliable identifications. Jenni et al. (1989) claimed nitrogen-fixing capability for strains of P. pseudoalva, but their identification should also be confirmed.

The consideration of the question of synonymy of the two genera hinges on the structure of Pseudomonas itself. As presently defined, the sequence data reported here could be considered strong circumstantial evidence for synonymy, although comparative analyses of the 16S rRNA gene (Anzai et al., 2000), 16S rRNA gene and rpoB (Ait Tayeb et al., 2005), 16S rRNA gene (Moore et al., 1996) and gyrB and rpoD (Yamamoto et al., 2000) hint at subgeneric species groupings that could imply subdivision of the genus. If further studies showed that Pseudomonas should be represented more properly by a small number of species closely associated with P. aeruginosa then Azotobacter might be retained as a sister genus. While it is sometimes claimed that phenotypic data correlate with sequence data, a careful reading suggests that this is generally not the case. Comparisons have been made using a number of phenotypic datasets: ribotyping (Brosch et al., 1996), fatty acid content of whole-cell hydrolysates and phospholipid fractions (Vancanneyt et al., 1996a), SDS-PAGE of whole-cell protein (Vancanneyt et al., 1996b), westprinting (Tesar et al., 1996) and fatty acid profiling (Stead, 1992). None of these datasets indicate coherent relationships between themselves, and none indicate relationships that correspond to those indicated by sequence data. There is no substantive evidence at present to suggest that Pseudomonas should be represented more properly by a small number of species closely associated with P. aeruginosa, the condition necessary to support retention of Azotobacter as a separate genus.

If Azotobacter species were transferred to Pseudomonas as presently defined, they would be included in an ‘aeruginosa’ cluster (Anzai et al., 2000; Moore et al., 1996; Yamamoto et al., 2000). Because of their very close proximity to P. aeruginosa, the type species, they would be considered to be authentic members of the genus. Key generic phenotypic determinative characteristics of Pseudomonas recorded by Palleroni (2005) that differentiate them from Azotobacter are their size (0.5–1.0 × 1.5–5.0 μm), cell morphology and motility (by one or more polar flagella; rarely non-motile). Apparently, no studies have been made to discover whether the diffusible fluorescent pigments produced by Pseudomonas and Azotobacter are chemically similar. Transfer would require emendation of the generic description to include peritrichous flagella insertion, cell pleomorphism and nitrogen-fixing capability.
The diversity of these phenotypic characteristics, morphology, flagella insertion and nitrogen-fixing capability, associated with a population of bacteria sharing core housekeeping genes, invites the consideration that they are peripheral to the core molecular structures of these bacteria. The form of flagella insertion and cell size are generally regarded as distinct generic, determinative characteristics, but these, and the capacity to fix nitrogen, may be adaptive characteristics that are stably integrated but nevertheless arise from transmissible elements. It has been clear for some time that the transfer of genetic material between unrelated bacteria is a widespread phenomenon, though less clear is how regularly such transfer leads to phenotypic expression (Young, 2001). If expression is common, as indicated here, the implications for bacterial systematics are far-reaching. The assumption of current classification is still that there will be some correspondence between bacterial taxa established on the basis of comparative analyses of conserved sequences, and their phenotypic circumscriptions (Murray et al., 1990; Stackebrandt et al., 2002). However, reticulated evolution (Karlin et al., 1997; Rivera et al., 1998; Doolittle, 1999) may be so pervasive that such correspondence fails too often to give support to this concept.

The genes were chosen for this study when such selections were largely arbitrary, on the basis that they were involved in some deep-seated metabolic activity and therefore were unlikely to be hypervariable; influenced by trivial adaptive or random mutation. In a study of bacterial species for which whole chromosomes have been sequenced, Zeigler (2003) compared 32 individual gene sequences as representatives of available chromosomes. Of those tested, genes such as recN, lig, dnaX, glyA and cysS gave best results, with atpD (23rd), recA (29th) and 16S rRNA (32nd) being relatively unrepresentative of the chromosome. atpD, carA and recA were used because primers could be developed from the limited data available at the time (Gaunt et al., 2001; Weir et al., 2004). In future, it should be possible to choose genes that better represent the taxa under consideration. Notwithstanding the caution indicated here, a conclusion of this study, that there is a significant lack of correspondence between relationships indicated by conserved metabolic genes and key phenotypic determinative characteristics, is not likely to be contradicted by a refinement of gene selection.

These data also reinforce a point that has been noted in the past, that tree structures are highly dependent on sequence selection and on the choice of tree-building algorithm (Ludwig & Schleifer, 1999). A consideration of recent issues of IJSEM shows that NJ is still the preferred algorithm for the interpretation of sequence data. More rapid computational methods now permit the application of other analyses. With this done, however, it is not clear what are appropriate criteria for deciding which algorithm is to be preferred. The tendency is probably to choose algorithms that produce trees that conform to the taxonomic inclinations of authors. Natural selection acts at the molecular level to refine protein efficiency. Identical peptides can be encoded by different nucleotide sequences because of the redundancy in coding of the second and third bases of codons. This is shown by the greater variability shown in nucleotide trees compared with peptide trees (see Supplementary Figs S1 and S2 in IJSEM Online). Peptide sequences are the units of selection. Nucleotide sequences are the vehicles for transfer of the functional structure between generations. This leaves open the question as to which sequences, nucleotides or peptides, are most relevant in the consideration of the relationships of organisms.

Although one or more combined gene analyses may give what appear to be robust separations into distinct groups, there is no objective method of calibration to determine whether such groups merit recognition as distinct taxa (Young, 2001; Gevers et al., 2005). For comparisons using DNA–DNA reassociation, the arbitrary value of 70% reassociation was set to discriminate individual species. An attempt to calibrate 16S rRNA gene sequence data with DNA–DNA reassociation data was made by Stackebrandt & Goebel (1994). Using data available then, they demonstrated that strains that had 16S rRNA gene sequence similarity values lower than 97% were invariably from different species based on reassociation data. However, this co-relationship is of limited use because many species based on reassociation have 16S tRNA gene sequence similarities greater than 99%. Using other genes or concatenated sequences does not give even this limited calibration. Determination of taxonomic divisions can only be made by some arbitrary criterion, such as a specified DNA–DNA reassociation value, or by the subjective decision of an author. This problem applies to all datasets gathered by any means.

General taxonomic considerations

In recent times, bacterial classification, especially at the generic level, has been based largely on comparative analyses of a single molecular sequence (usually the 16S rRNA gene) using a single algorithm (usually NJ). This approach has been considered to provide a key to the systematization of relationships at higher taxonomic levels in classifications that aim to demonstrate inferred historical, ancestral relationships; viz., phylogenies (Stackebrandt & Woese, 1984; Woese, 1987; Vandamme et al., 1996). Underpinning such phylogenetic classification is the assumption that change in these, necessarily conserved, sequences occurs with clock-like regularity (Mindell & Thacker, 1996). It has been clear for some time, however, that such interpretations are oversimple and sometimes misleading. Although the rRNA gene is ubiquitous, this is not of itself the basis for asserting that comparative analyses of this sequence alone indicate reliable relationships of organisms (Young, 2001; Young et al., 1992) or that any other gene can be assumed, a priori, to reflect relationships with accuracy.
Increasingly it has come to be understood that molecular analyses can be refined by using some form of consensus based on several genes. Such analyses have been restricted in the past by the technical difficulties of developing primers and of producing large numbers of sequences, which have now largely been overcome. The concept of multi-locus sequence typing (MLST) (Maiden et al., 1998) (now called multi-locus sequence analysis; MLSA) involves systematic selection of a number of genes to represent the chromosome, and analysis of these genes or their peptide products as concatenated sequences (Gevens et al., 2005). Clearly, studies using properly conducted MLSA will give improved taxonomic resolution in future systematics sequence studies.

The goal of MLSA is to reflect the overall similarities and differences in molecular structures of organisms. It is not often noticed that this approach is overtly phenetic (Sneath, 1989) and is explicitly in contrast with the phylogenetic approach. Phenetic classification is based on the overall similarities and differences of organisms. In practice, until now investigation has been confined to a relatively small proportion of the ‘totality’ of an organism, but the intention is clear. Phenetic classification is natural and evolutionary in the sense that it gives expression to bacterial groupings determined by the interplay of natural selection on the phenotype as determined by the genome. It explicitly does not aim to show ancestral relationships, accepting the possibility that apparently diverse taxa may be the product of rapid changes between phylogenetically closely related organisms, or that the ancestrally distant related organisms may appear to be closely related by convergent events that have left no molecular trace. Any dataset can be interpreted as phenetic or phylogenetic classification. Applied as phylogenetic classification, it makes the additional, unverifiable and historical, claim that the analyses indicate reliable ancestral relationships. Unverifiable, because without a fossil record there is no method by which phylogenies can be calibrated (Wilkins, 2002). Attempts at phylogenetic classification based on macro-morphological structures were abandoned as it came to be understood that such classifications were not adequately robust and were open to the logical contradiction that structures could not be construed as responding to change with time when they were necessarily highly conserved by natural selection (Davis & Hayward, 1963). Phylogenetic interpretations based on molecular morphological structures, with the assumed correlation between molecular structural change and time, have the same logical objection.

**ACKNOWLEDGEMENTS**

The Marsden Fund of the Royal Society of New Zealand under contract 97-LAN-LFS-0011 and the New Zealand Foundation for Research, Science and Technology provided financial support. H. M. Harman and N. Waipara are thanked for critically reading the manuscript.

**REFERENCES**


The Proteobacteria, phylogenetic issues and applications. DNA-DNA reassociation and 16S rRNA sequence analysis in the Syst Appl Microbiol.


