Symbiosis, 38 (2005) 109–122
Balaban, Philadelphia/Rehovot

LIBRARY
JAN 2 0 2010
SAINT MARY'S UNIVERSITY
HALIFAX, CANADA
B3H 3C3

Review article.

Use of Population Genetic Structure to Define Species Limits in the *Rhizobiaceae*

BERTRAND EARDLY[1*] and PETER VAN BERKUM[2]
[1]*Penn State Berks-Lehigh Valley College, Berks Campus, Reading, PA 19610, USA, Tel. +1-610-396-6131, Fax. +1-610-396-6024, Email. bde1@psu.edu;*
[2]*Soybean Genomics and Improvement Laboratory, Agricultural Research Service, U.S. Department of Agriculture, Beltsville, MD 20705, USA, Tel. +1-301-504-7280, Fax. +1-301-504-5728, Email. pberkum@ba.ars.usda.gov*

Received June 9, 2004; Accepted July 28, 2004

Abstract

Symbiotic bacteria of the family *Rhizobiaceae* are currently defined using a consensus or polyphasic approach, where emphasis in determining species limits is based on the quantification of overall genotypic and phenotypic similarity. In the first part of this review some of the limitations of this approach are examined. In the second part an alternative population-based approach is considered. The primary assumption underlying this approach is that ecological selection is the dominant force constraining genetic diversity in bacterial populations. Practical methods for assessing the range and extent of this diversity are described, along with an example of how such information has been used to provide evidence for two symbiotic nitrogen-fixing species within the genus *Sinorhizobium*.

Keywords: Symbiotic, *Rhizobiaceae*, polyphasic, systematics, MLST, ecotype, species concept, genetic structure, phylogeny

*The author to whom correspondence should be sent.

0334-5114/2005/$05.50 ©2005 Balaban

1. Introduction

Biological taxonomy is the science of classification of organisms, and consists of three separate, but interrelated areas: classification, nomenclature, and identification. The goal of bacterial taxonomy is to classify and catalogue bacterial species in a way that reflects their evolutionary history. According to Brenner (2001) most new species are defined using a pragmatic polyphasic (consensus) approach that integrates all available genotypic and phenotypic information. Although there are no strict rules or guidelines governing this approach, genotypic similarity currently receives the primary emphasis in most taxonomic decisions. Genotypic similarity is usually measured through DNA-DNA hybridization analyses and through comparisons of *16S* rRNA (*16S*) gene sequences. Phenotypic similarity is usually assessed through comparisons of a variety of morphological, physiological, and biochemical characteristics. Using this information, pairwise strain similarities are calculated as percentages, and strains are grouped using one of several clustering algorithms.

When a bacterium or a cluster of bacteria is proposed as a species, a single representative is chosen as the type strain. This strain serves as both the name-bearer and permanent reference specimen for that species (Brenner et al., 2001). This practice was derived from the Linnaean system of classification where a "type specimen" was chosen for each species to reflect the belief that all earthly items are projections of their perfect, heavenly forms. Even though bacterial speciation occurs in a population context, bacterial taxonomic studies have seldom relied on population genetic analyses to determine species limits. In fact, from 1990 to 2000, approximately 40% of all new bacterial species and genera were based on the characterization of only a single isolate (Christensen et al., 2001). This has apparently been due, at least in part, to the lack of suitable methods for characterizing allelic differences among large numbers of strains. This problem has been addressed recently through the use of a method called multilocus sequencing typing, or MLST (Maiden et al., 1998). This is a genomic indexing method that involves the determination of partial (~500 bp) DNA sequences for approximately ten housekeeping genes within a large set of reference strains.

The goals in this review are to: (i) critically examine current methods used in polyphasic bacterial taxonomy, particularly those that are routinely applied to members of the family *Rhizobiaceae*, (ii) consider the theoretical basis for the use of population genetic structure to define species limits, and (iii) present an example of how population genetic information has been used to define species limits within the family *Rhizobiaceae*.

2. DNA-DNA Hybridization

DNA-DNA hybridization is a method that is based on the ability of double-stranded DNA to reversibly dissociate and reassociate with complementary strands. This property is dependent on both the temperature and the salt concentration of the solution in which the DNA is incubated. When sheared DNAs from two bacteria are denatured, mixed, and then allowed to reassociate, heteroduplexes between complementary strands form if the degree of sequence homology is sufficient. Where sequence homology is insufficient, sheared fragments remain in solution as single-stranded DNA. The relative proportions of single-stranded DNA to heteroduplex DNA is interpreted as an indication of the degree of divergence between the bacteria being compared (Brenner et al., 2001). With this method a species is often defined as a group of strains that share at least 70% total genome homology (Wayne et al., 1987).

The current emphasis on the use of DNA-DNA hybridization to determine the limits of bacterial species has been criticized. These criticisms have focused mainly on: (i) the labor intensiveness of the method and non-transportability of the results, which in practical terms does not allow for sufficient sampling of the genetic diversity in natural populations, (ii) the use of an arbitrary cutoff (70% similarity) to define species limits, and (iii) the conclusion that species defined using this criterion are not real entities and thus may be of limited value in understanding evolutionary dynamics in natural populations (van Berkum et al., 2003). A related criticism, which is particularly relevant to genera within the family *Rhizobiaceae*, is the potential ambiguity that may arise as the result of acquisition (or loss) of large horizontally transferred genomic segments and replicons (e.g., Sullivan et al., 1995).

3. *16S* rRNA (*16S*) Nucleotide Sequence Similarity

Methods for the reconstruction of bacterial phylogeny have undergone dramatic progress with the advent of sequencing analysis of ribosomal RNA (rRNA) genes (Maidak et al., 1994; Olsen et al., 1994). The *16S* rRNA (*16S*) gene is particularly useful for this purpose because it is slowly evolving and its gene product is both universally essential and functionally conserved. Basing estimates of phylogeny on *16S* sequences presupposes that the evolution of the genome and its *16S* genes progress at a constant rate, and that the relationships between different *16S* genes are strictly hierarchical. In other words it is assumed that the *16S* genes are passed from generation-to-generation by vertical descent, and that the *16S* genes (or gene segments) are not shared between lineages by horizontal transfer. This has led to the assumption that

sequence variation within *16S* genes can be used to map the evolutionary paths of entire genomes. From a practical standpoint, this approach requires that either each genome harbor a single copy of the *16S* gene, or that multiple alleles within a given cell are identical.

Although the results from some multilocus sequencing studies give the impression of congruence between *16S* trees and trees based on other loci (Gaunt et al., 2001), cases of discordant gene trees have been reported (e.g., Feil et al., 2001). Other lines of evidence also bring into question the validity of the common assumptions justifying the use of the *16S* gene for phylogenetic analyses. For example, linear discontinuities within *16S* gene sequences are often taken as evidence of recombination among divergent alleles (Eardly et al., 1996; Smith et al., 1999). In a detailed analysis of the ribosomal operons in rhizobia and other α-Proteobacteria, van Berkum et al. (2003) provided convincing evidence that different sections of rRNA genes may yield different phylogenetic signals, suggesting that they may have histories of intragenic recombination involving divergent *16S* alleles. An extreme case is represented by the species *Themomonospora chromogena*, which has two functional *16S* genes, one of which appears to have been acquired through horizontal gene transfer from another species (Wang et al., 1997; Yap et al., 1999). The validity of this explanation has been supported by related observations that heterologous rRNA alleles do not necessarily inhibit protein synthesis in the cell (Asai et al., 1999).

In addition to questions on the validity of the assumptions that support the use of *16S* sequences in taxonomic studies, there are also questions as to whether it is reasonable to set numerical limits on the percentage sequence similarity required for a given strain to be considered a member of a particular species (Fox et al., 1992). Although Stackebrandt and Goebel (1994) concluded that two strains having *16S* sequence similarities less than 97% should belong to separate species, there are reports of *16S* alleles within single genomes that differ at 6.4% of their nucleotide sequence positions (Wang et al., 1997).

4. Phenotypic Analyses

The *Enterobacteriaceae* was the first family of bacteria for which there was a comprehensive set of phenotypic principles established for identification and classification (Brenner, 2001). This classification was based on a range of morphological, biochemical, and ecological characteristics. Numerical taxonomic methods subsequently improved the validity of phenotypic identification by increasing the number of different analyses. Coefficients of similarity were also introduced as a useful means of quantifying the degree of similarity among strains within a species (Sneath and Sokal, 1973). A

threshold of 80% similarity has generally been accepted as a requirement for the assignment of a bacterium to a particular species. In order to be classified as a separate bacterial species, a strain must not only have distinct genotypic characteristics, but it must also have a distinctive phenotype as well (Wayne et al., 1987). A minimal phenotypic description is currently required when proposing a new bacterial species (Gillis et al., 2001). This requirement has been justified by the need to identify bacterial species through the use of routine laboratory diagnostics.

It is evident from the literature that progressively fewer research groups are using the full range of phenotypic methods traditionally used for conventional taxonomic analyses (Gillis et al., 2001). Some possible explanations for this trend are that these analyses are often laborious and difficult to standardize, and it is often unclear as to whether certain traits should be weighted more heavily than others. Considering the potential limitations of the three standard methods currently used in polyphasic analyses, it is clear that alternative approaches are necessary to more accurately assess the organization of genetic diversity in natural bacterial populations.

5. A Theoretical Basis for the Consideration of Population Genetic Structure in the Bacterial Species Definition

In highly sexual animal and plant populations, genetic exchange is the dominant force of cohesion within a species (Avise, 2000; Mayr, 1963). In contrast, in bacterial populations, genetic exchange does not appear to be a dominant cohesive force. This is due to the fact that in bacteria genetic recombination is not coupled to reproduction, so traits are usually inherited vertically. Furthermore, genetic recombination in bacteria is relatively rare and can also be highly promiscuous. Upon the basis of these observations Cohan (1996) concluded that the function of genetic exchange among bacteria differs from its function among eukaryotes, and that the evolutionary consequences of genetic exchange – that zoologists and botanists take for granted – do not extend to bacteria. He contended that since genetic exchange does not significantly hinder adaptive divergence in bacteria, it cannot provide a sound basis for their classification. The question remains however, if genetic exchange does not hinder adaptive divergence in bacteria, what are the forces that do?

Genetic variation is fundamental to Darwinian evolution. Point mutations, genomic rearrangements, and horizontal gene transfer are all important driving forces in creating variants that are subject to genetic drift and Darwinian selection (Arber, 1993). By their nature, point mutations usually result in small changes that often require long periods of time to produce distinct bacterial

lineages. In contrast, mutations resulting from genomic rearrangements and lateral gene transfers can result in almost instantaneous reticulate (netlike) microbial phylogenies. This variety of mutational mechanisms provides a reasonable explanation for the genetic responsiveness of bacteria to rapidly changing environmental pressures.

The three primary processes by which bacteria are able to share genetic information are transformation, conjugation, and transduction (Ochman et al., 2000; Zgur-Bertok, 1999). However not all bacterial species are capable of sharing genetic information by all of these mechanisms. This is likely due to differences in mismatch repair systems and other forms of sexual isolation (Radman and Wagner, 1993; Matic et al., 1995; Majewski et al., 2000), and probably explains why rates of recombination in different bacterial species vary so widely. For example, in an MLST analysis of the genetic variation among several housekeeping genes in *Neisseria meningitidis, Escherichia coli,* and *Streptococcus pneumoniae,* Feil et al. (2001) concluded that a single nucleotide was 10- to 80- fold more likely to change as a result of recombination than as a result of mutation (depending on the particular species).

Such results have lead to the recognition that a range of different population genetic structures exist in bacteria. For example in *Mycobacterium tuberculosis,* where recombination rates are low and allelic variation within the core genome is relatively limited, most natural isolates appear to be very similar, reflecting a strongly clonal pattern of inheritance (Feil, 2003). However at the other extreme in *Helicobacter pylori,* recombination rates are relatively high and allelic combinations among isolates of a population appear to be at linkage equilibrium (or panmixis). In these populations most of the isolates possess unique genotypic combinations, and it is difficult to discern discrete clonal clusters or groups (Suerbaum et al., 1998). Between these extremes are a number of species that are characterized by a limited number of widespread, predominant genotypes (or clones) that coexist with a large number of relatively rare genotypes. The predominant clones in these groups tend to belong to much larger clusters of related genotypes whose alleles differ by only one or a few point mutations. Cohan (2001) has proposed that these clonal complexes, which he calls "ecotypes", have the quintessential characteristics of biological species. If these complexes are to be used as taxonomic units (i.e., species) one of the main challenges will be to recognize real discontinuities in the patterns of variation between them. It also seems likely that some traditional bacterial "species" will need to be reclassified as genera, or perhaps as clusters of related species.

This type of species definition can be thought of as an ecological species concept, in which a species is defined as a group of organisms that exploit the same ecological niche (van Valen, 1976). The assumption underlying this concept is that if two sets of organisms occupy the same ecological niche, at the

same place and at the same time, then stochastic processes will inevitably result in the displacement of one of the groups by the other. The fine-scale dynamics that mediate the development and/or demise of such incipient ecological species have been described (Cohan, 2001). Theoretically, the longevity of a particular sequence cluster (or ecotype) will be limited, because over time it is likely that it will be out-competed by an adaptive mutant from within the population. This type of periodic selection is sometimes referred to as a "selective sweep" (Guttman and Dykhuizen, 1994). The net effect of selective sweeps would be to purge a population of nearly all its genetic diversity, conferring an appearance of genetic cohesion among the survivors. As Cohan (2001) pointed out, the resulting ecotypes that are successful in occupying a particular niche share many properties with eukaryotic species.

6. Application of a Population-based Approach to the Systematic Analysis of *Sinorhizobium* sp.

In a previous study Eardly et al. (1990) examined the genetic structure in a collection of 232 strains representing the symbiotic nitrogen-fixing soil bacterium *Sinorhizobium* (*Rhizobium*) *meliloti*. Allelic variation among 14 chromosomal enzyme genes was assessed by multilocus enzyme electrophoresis (MLEE). Like MLST, MLEE is a form of genomic indexing, but was used prior to the development of the sequence-based MLST method. Most of the strains in the *S. meliloti* study originated from Southwest Asia, which is the center of origin of their primary host genus, *Medicago*. Fifty distinct multilocus genotypes (enzyme electrophoretic types or ETs) were identified. These multilocus genotypes were clustered into two primary phylogenetic divisions separated at a genetic distance of 0.76, as revealed by an analysis using unweighted pair group method algorithm (UPGMA) (Fig. 1). By the criterion of genetic differentiation conventionally applied in defining species limits among members of the family *Enterobacteriaceae* and certain other bacteria, it was concluded that the two primary divisions represented distinct evolutionary species. Since the type strain for the species *S. meliloti* was identified as member of division A (ET 1, strain ATCC 9930), it was inferred that this group of bacteria represented the named species *S. meliloti*. Subsequent comparative analyses, including *16S* rRNA, DNA-DNA hybridization, and phenotypic analyses of representative strains from both divisions confirmed the affiliation of the type strain with the division A strains, and also the genotypic and distinctiveness of the division B strains (Rome et al., 1996a,b). As a result it was proposed that the division B strains and similar isolates from other sources represent the new species, *S. medicae* (Rome et al., 1996b).

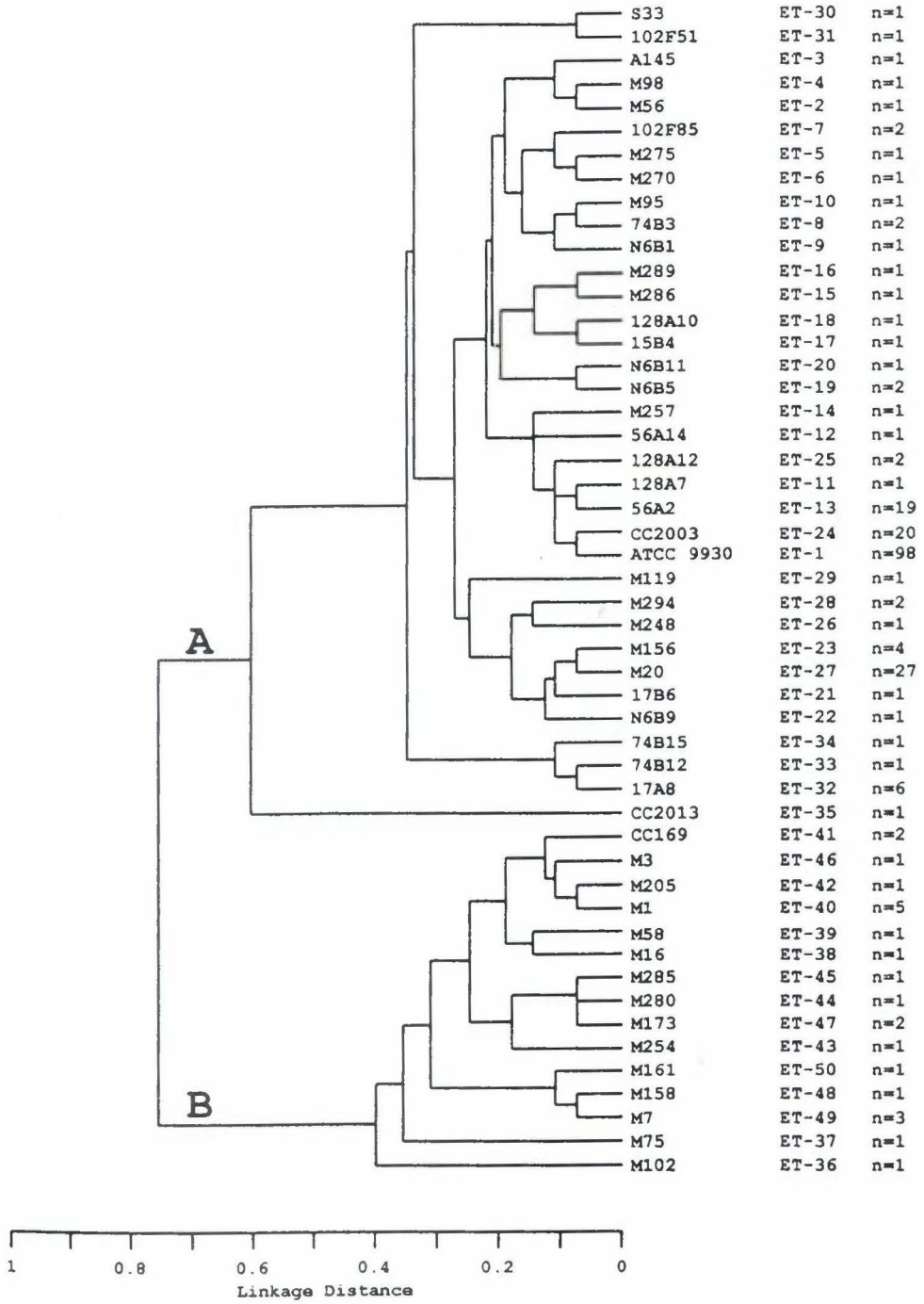

Figure 1. See legend on next page.

Although clustering algorithms, such as the UPGMA algorithm used to generate Fig. 1, are useful to identify closely related genotypes, they provide no clear information as to the likely founding genotypes for the populations, nor do they show likely patterns of evolutionary descent within the clusters. This problem has been addressed by implementing an algorithm called eBURST, which is capable of using multilocus genotypic information to provide a realistic model of clonal expansion within natural bacterial populations (Feil et al., 2004). The complex and somewhat arbitrary branching patterns in Fig. 1 were transformed by eBURST into discrete clusters, where distinct patterns of descent could be recognized (Fig. 2). From this analysis a large clonal complex founded by the isolates representing ET 1 was evident. This complex included the type strain for the species. In this context a clonal complex is defined as a group of strains in which each strain shares most of its alleles with the other members of that group. The prevalence of the ET 1 isolates in the dataset is reflected by the larger area of the blue circle in the eBURST diagram. Linked to this founding clone are cofounding clones (noted in yellow) that have diversified into secondary clonal complexes. These contain progressively more allelic differences when compared to the founding clone. Also evident in the diagram are four completely unlinked clonal complexes representing clusters that were observed on the deeper branches of the UPGMA dendogram in Fig. 1. Included among these is the independent clonal complex representing ETs (e.g., ET 47) of the proposed species *S. medicae*. It should be noted, however, that both the unlinked clones and other unlinked clonal complexes in Fig. 2 provide no information as to their genetic distances from the other complexes in the eBURST diagram.

Recent advances in sequencing technology have made it possible to apply these methods to the analysis of nucleotide sequence data. This multilocus sequence typing (MLST) approach has several advantages over the enzyme electrophoresis method described above (Feil and Spratt, 2001). Not only can

See figure on previous page.

Figure 1. The linkage distance among *Medicago*-nodulating rhizobia derived from variation in the electrophoretic mobilities of 14 enzyme loci among 232 strains according to Eardly et al. (1990). A matrix of the strain ID and the electrophoretic type (ET), followed by the allele labels for each, was used in START (Sequence Type Analysis and Recombinational Tests, version 1.05) program to generate the UPGMA dendrogram. The program was written by Keith Jolley, University of Oxford.

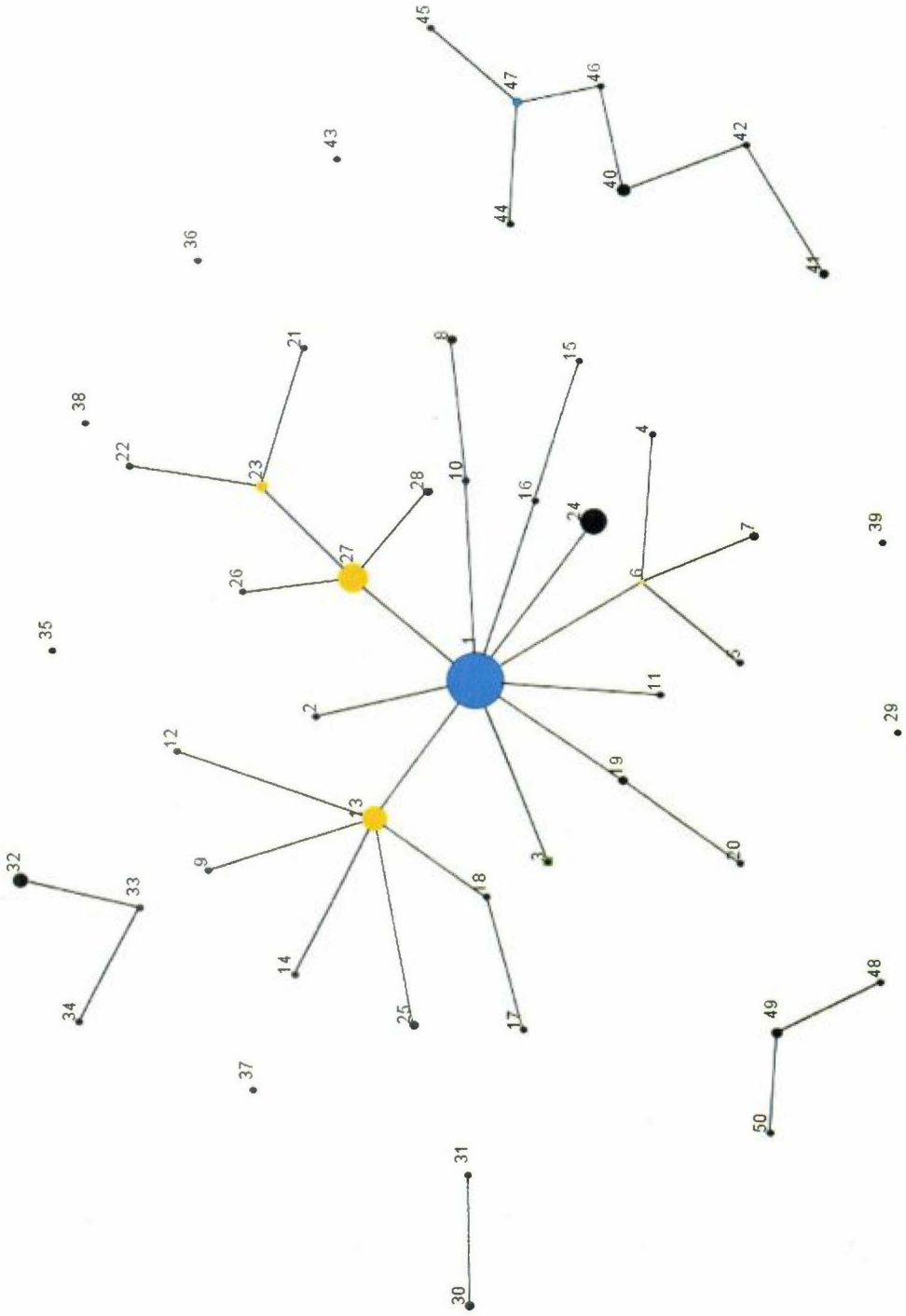

Figure 2. See legend on next page.

MLST provide a fine-scale model of clonal expansion in natural bacterial populations, but it can also provide fully transportable datasets that allow the identification and quantification of mutation and recombination events within a population. MLST analyses have already been reported for several pathogenic bacterial species (Feil, 2003) and are currently being applied in the analysis of the *Sinorhizobium* (*Rhizobium*) strains from the MLEE study described above (van Berkum et al., 2004). The goal of this work is to characterize the allelic variation within and among the two proposed *Sinorhizobium* (*Rhizobium*) species through an analysis of a core set of ten chromosomal genes in all of the strains. With this information it should be possible to verify the founders of the clonal complexes in Fig. 2 and examine the degree to which homologous recombination occurs within the core genes of these species.

Accurate measurements of the rates of recombination within a species are particularly important for population-based phylogenetic analyses, because high rates of recombination can obscure phylogenetic signals produced by individual genes. For example, if the recombination rates within a population are high, and this is not recognized prior to the analysis of a multilocus dataset, then a clustering algorithm could artificially force hierarchical grouping patterns that would be more accurately portrayed as networks (Maynard Smith et al., 1993). Fortunately this can be recognized during MLST analysis by examining patterns of congruence among individual gene trees (Dykhuizen and Green, 1991). In previous studies where this approach was used, with species known to have a basically clonal population structure (such as *Escherichia coli*), gene trees were produced having high levels of congruence. In corresponding studies involving species known to have a panmictic population structure, such as the naturally transformable species *Neisseria meningitidis*, individual gene trees were produced that had low levels of congruence (Zhou et al., 1997). It would therefore be expected that the phylogenetic trees reconstructed from the core genes of this species would not provide a valid approximation of their true phylogenetic history.

See figure on prevoius page.

Figure 2. Clonal complex of *Medicago*-nodulating rhizobia derived from variation in the electrophoretic mobilities of 14 enzyme loci among 232 strains according to Eardly et al. (1990). A matrix of the electrophoretic type (ET) followed by the allele labels for each was used in eBURST (Feil et al., 2004) to generate a diagram of the evolutionary patterns among the strains. The sizes of the circles are related to the numbers of strains within each ET. The founder and co-founder genotypes are colored blue and yellow, respectively. Distances indicated between ETs by the connecting lines are arbitrary.

7. Summary

An *ad hoc* committee for the re-evaluation of the species definition in bacteriology met recently (February 2002) in Gent, Belgium, to discuss the potential value of population-based genetic analyses to bacterial systematics (Stackebrandt et al., 2002). The impetus for this meeting was the recognition that innovative new methods, such as MLST, can provide new opportunities for the elucidation of bacterial genomic relatedness at both the inter- and intra-specific level. One of the recommendations of this committee was that microbiologists be encouraged to base their species descriptions on more than one strain. A second recommendation was that investigators be encouraged to propose new species on the basis of innovative new genomic methods, such as MLST, using strain collections for which DNA-DNA hybridization data is also available.

REFERENCES

Arber, W. 1993. Evolution of prokaryotic genomes. *Gene* **135**: 49–56.

Asai, T., Zaprojets, D., Squires, C., and Squires, C.L. 1999. An *Escherichia coli* strain with all chromosomal rRNA operons inactivated: complete exchange of rRNA genes between bacteria. *Proceeding of the National Academy of Sciences USA* **96**: 1971–1976.

Avise, J.C. 2000. *Phylogeography: The History and Formation of Species*. Harvard University Press, Cambridge University Press, Cambridge, 464 pp.

Brenner, D.J., Staley, J.T., and Krieg, N.R. 2001. Classification of prokaryotic organisms and the concept of bacterial speciation. In: *Bergey's Manual of Systematic Bacteriology*. Boone, D.R. and Castenholz, R.W., eds. Springer-Verlag, New York, pp. 27–31.

Christensen, H., Bisgaard, M., Frederiksen, W., Mutters, R., Kuhnert, P., and Olsen, J.E. 2001. Is characterization of a single isolate sufficient for valid publication of a new genus or species? Proposal to modify Recommendation 30b of the Bacteriological Code (1990 Revision). *International Journal of Systematic Bacteriology* **51**: 2221–2225.

Cohan, F.M. 1996. The role of genetic exchange in bacterial evolution. *ASM News* **62**: 631–636.

Cohan, F.M. 2001. Bacterial species and speciation. *Systematic Biology* **50**: 513–524.

Dykhuizen, D.E. and Green, L. 1991. Recombination in *Escherichia coli* and the definition of a biological species. *Journal of Bacteriology* **173**: 7257–7268.

Eardly, B.D., Materon, L., Smith, N.H., Johnson, D.A., Rumbaugh, M.D., and Selander, R.K. 1990. Genetic structure of natural populations of the nitrogen-fixing bacterium *Rhizobium meliloti*. *Applied and Environmental Microbiology* **56**: 187–194.

Eardly, B.D., Wang, F.-S., and van Berkum, P. 1996. Corresponding 16S rRNA gene segments in Rhizobiaceae and Aeromonas yield discordant phylogenies. *Plant and Soil* **186**: 69–74.

Feil, E.J. 2003. How stable are the core genes of bacterial pathogens? *ASM News* **6**: 234–238.

Feil, E.J., Holmes, E.C., Bessen, D.E., Chan, M.-S., Day, N.P.J., Enright, M.C., Goldstein, R.,

Hood, D.W., Kalia, A., Moore, C.E., Zhou, J., and Spratt, B.G. 2001. Recombination within natural populations of pathogenic bacteria: short-term empirical estimates and long-term phylogentic consequences. *Proceedings of the National Academy of Sciences USA* **98**: 182–187.

Feil., E.J., Li, B.C., Aanensen, D.M., Hanage, W.P., and Spratt, B.G. 2004. eBURST: Inferring patterns of evolutionary descent among clusters of related bacterial genotypes from multilocus sequencing typing data. *Journal of Bacteriology* **186**: 1518–1530.

Feil, E.J. and Spratt, B.G. 2001. Recombination and the population structures of bacterial pathogens. 2001. *Annual Reviews of Microbiology* **55**: 561–590.

Fox, G.E., Wisotzkey, J.D., and Jurtshuk, P. 1992. How close is close: 16S rRNA sequence identity may not be sufficient to guarantee species identity. *International Journal of Systematic Bacteriology* **42**: 166–170.

Gaunt, M.W., Turner, S.L., Rigottier-gois, L., Lloyd-Macgilp, S.A., and Young, J.P.W. 2001. Phylogenies of atpD and recA support the small subunit rRNA-based classification of rhizobia. *International Journal of Systematic and Evolutionary Microbiology* **51**: 2037–2048.

Gillis, M., Vandamme, P., De Vos, P., Swings, J., and Kersters, K. 2001. Polyphasic taxonomy. In: *Bergey's Manual of Systematic Bacteriology*. Boone, D.R. and Castenholz, R.W., eds. Springer-Verlag, New York, pp. 43–48.

Guttman, D.S. and Dykhuizen, D.E. 1994. Clonal divergence in *Escherichia coli* as a result of recombination, not mutation. *Science* **266**: 1380–1383.

Maidak, B.L., Larsen, N., McCaughey, M.J., Overbeek, R., Olsen, G.J., Fogel, K., Blandy, J., and Woese, C.R. 1994. The ribosomal database project. *Nucleic Acids Research* **22**: 3485–3487.

Maiden, M.C.J., Bygraves, J.A., Feil, E., Morelli, G., Russell, J.E., Urwin, R., Zhang, Q., Zhou, J., Zurth, K., Caugant, D.A., Feavers, I.M., Achtman, M., and Spratt, B.G. 1998. Multilocus sequence typing: a portable approach to the identification of clones within populations of pathogenic microorganisms. *Proceedings of the National Academy of Sciences USA* **95**: 3140–3145.

Majewski, J., Zawadzki, P., Pickerill, P., Cohan, F.M., and Dowson, C.G. 2000. Barriers to genetic exchange between bacterial species: *Streptococcus pneumoniae* transformation. *Journal of Bacteriology* **182**: 1016–1023.

Matic, I., Rayssiguier, C., and Radman, M. 1995. Interspecies gene exchanges in bacteria: the role of SOS and mismatch repair systems in evolution of species. *Cell* **80**: 507–515.

Maynard Smith, J., Smith, N.H., O'Rourke, M., and Spratt, B.G. 1993. How clonal are bacteria? *Proceedings of the National Academy of Sciences USA* **90**: 4384–4388.

Mayr, E. 1963. *Animal Species and Evolution*. Harvard Univ. Press, Cambridge, 811 pp.

Ochman, H., Lawrence, J.G., and Groisman, E.A. 2000. Lateral gene transfer and the nature of bacterial innovation. *Nature* **405**: 299–304.

Olsen, G.J. Woese, C.R., and Overbeek, R. 1994. The winds of (evolutionary) change: breathing new life into microbiology. *Journal of Bacteriology* **176**: 1–6.

Radman, M. and Wagner, R. 1993. Mismatch recognition in chromosomal interactions and speciation. *Chromosoma* **102**: 369–373.

Rome, S., Brunel, B., Normand, P., Fernandez, M., and Cleyet-Marel, J.-C. 1996a. Evidence that two genomic species of *Rhizobium* are associated with *Medicago truncatula*.

Archives of Microbiology **165**: 285–288.

Rome, S., Fernandez, M.P., Brunel, B., Normand, P., and Cleyet-Marel, J.-C. 1996b. *Sinorhizobium medicae* sp. nov., isolated from annual *Medicago* spp. *International Journal of Systematic Microbiology* **46**: 972–980.

Smith, N.H., Holmes, E.C., Donovan, G.M., Carpenter, G.A., and Spratt, B.G. 1999. Networks and groups within the genus *Neisseria*: analysis of argF, recA, rho, and 16S rRNA sequences from human *Neisseria* species. *Molecular Biology and Evolution* **16**: 773–783.

Sneath, P.H.A. and Sokal, R.R. 1973. *Numerical Taxonomy. The Principles and Practice of Numerical Classification.* W.H. Freeman, San Francisco. 573 pp.

Stackebrandt, E. and Goebel, B.M. 1994. Taxonomic note: a place for DNA-DNA reassociation and 16S rRNA sequence analysis in the present species definition in bacteriology. *International Journal of Systematic Microbiology* **44**: 846–849.

Stackebrandt, E., Frederiksen, W., Garrity, G.M., Grimont, P.A.D., Kämpfer, P., Maiden, M.C.J., Nesme, X., Roselló–Mora, R., Swings, J., Trüper, H.G., Vauterin, L., Ward, A.C., and Whitman, W.B. 2002. Report of the ad hoc committee for the re-evaluation of the species definition in bacteriology. *International Journal of Systematic and Evolutionary Microbiology* **52**: 1043–1047.

Suerbaum, S., Maynard Smith, J., Bapumia, K., Morelli, G., Smith, N.H., Kunstmann, E., Dyrek, I., and Achtman, M. 1998. Free recombination within *Helicobacter pylori*. *Proceedings of the National Academy of Sciences USA* **95**: 12619–12624.

Sullivan, J.T., Patrick, H.N., Lowther, W.L., Scott, D., and Ronson, C.W. 1995. Nodulating strains of *Rhizobium loti* arise through chromosomal symbiotic gene transfer in the environment. *Proceedings of the National Academy of Sciences USA* **92**: 8985–8989.

van Berkum, P., Elia, P., and Eardly, B. 2004. unpublished data.

van Berkum, P., Terefework, Z., Paulin, L.,Suomalainen, S., Lindström, K., and Eardly, B.D. 2003. Discordant phylogenies within the *rrn* loci of rhizobia. *Journal of Bacteriology* **185**: 2988–2998.

van Valen, L. 1976. Ecological species, multispecies, and oaks. *Taxon* **25**: 233–239.

Wang, Y., Zhang, Z., and Ramanan, N. 1997. The actinomycete *Thermobispora bispora* contains two distinct types of transcriptionally active 16S rRNA genes. *Journal of Bacteriology* **179**: 3270–3276.

Wayne, L.G., Brenner, D.J., Colwell, R.R., Grimont, P.A.D., Kandler, O., Krichevsky, M.I., Moore, L.H., Moore, W.E.C., Murray, R.G.E., Stackebrandt, E., Starr, M.P., and Truper H.G. 1987. Report of the ad hoc committee on reconciliation of approaches to bacterial systematics. *International Journal of Systematic Microbiology* **37**: 463–464.

Yap, W.H., Zhang, Z., and Wang, Y. 1999. Distinct types of rRNA operons exist in the genome of the actinomycete *Thermomonospora chromogena* and evidence for horizontal transfer of an entire rRNA operon. *Journal of Bacteriology* **181**: 5201–5209.

Zhou, J., Bowler, L.D., and Spratt, B.G. 1997. Interspecies recombination, and phylogenetic distortions, within the glutamine synthetase and shikimate dehydrogenase genes of *Neisseria meningitidis* and commensal *Neisseria* species. *Molecular Microbiology* **23**: 799–812.

Zgur-Bertok, D. 1999. Mechanisms of horizontal gene transfer (Review). *Folia Biologica Prague* **45**: 91–96.