NOTE

A *recA* gene phylogenetic analysis confirms the close proximity of *Frankia* to *Acidothermus*

Joëlle Maréchal,¹ Béatrice Clement,¹ Renaud Nalin,¹ Corinne Gandon,¹ Stéphanie Orso,¹ Jelena H. Cvejic,² Maud Bruneteau,¹ Alison Berry³ and Philippe Normand¹

Author for correspondence: Philippe Normand. Tel: +33 4 72 43 13 77. Fax: +33 4 72 43 12 23.
e-mail: normand@biomserv.univ-lyon1.fr

¹ Ecologie Microbienne,
UMR CNRS 5557,
Université Claude-Bernard,
Lyon I, 69622 cedex,
Villeurbanne, France

² Laboratoire de Biochimie
des Micro-organismes,
Université Louis-Pasteur,
4 rue Blaise-Pascal,
Strasbourg 67070, France

³ Department of
Environmental
Horticulture, University of
California, Davis,
CA 95616, USA

The closer proximity of *Frankia* and *Acidothermus cellulolyticus* relative to the morphologically close *Geodermatophilus* found previously was confirmed by resequencing the *rrs* gene of *Acidothermus cellulolyticus* and the housekeeping gene, *recA*. The diagnostic sugar 2-O-methyl-D-mannose was detected only in *Frankia*, while hopanoid lipids were present at high levels in both *Acidothermus* and *Frankia*.

Keywords: *Acidothermus*, *Frankia*, *Geodermatophilus*, hopanoid, *recA*

Frankia is one of the few bacterial genera to have developed a symbiosis with plants that extends the range of soil conditions suitable for growth. The nitrogen-fixing actinorhizal symbiosis allows 24 genera of plants belonging to eight dicotyledonous plant families to colonize soils that are poor in nitrogen, e.g. sandy beaches, mine spoils, burned forests or glacial moraines (Benson & Silvester, 1993). The variety of the plants concerned has generated discussion about whether the symbiosis has appeared once and been subsequently lost in several lineages, or if several independent appearances occurred (Soltis *et al.*, 1995).

Discussion on the evolutionary origin of the symbiosis has also been fuelled by recent phylogenetic reconstructions using the *rrs* (also designated the 16S rRNA) gene. The first such work (Hahn *et al.*, 1989) concluded that *Geodermatophilus* was the closest neighbour of *Frankia*, which was consistent with the fact that they shared a common morphological feature (the presence of multilocular sporangia) and a common habitat (soil). Later, as the number of sequences in databases increased, another study of the genus *Frankia* (Normand *et al.*, 1996) came to the conclusion that the closest neighbour of the genus *Frankia* was not *Geodermatophilus* (with a genetic distance of 5.3–7.5% substitution/site between *rrs* genes) but was a recently

described cellulolytic bacterium from an acidic hot spring in Yellowstone National Park; this bacterium, *Acidothermus cellulolyticus* (Mohagheghi *et al.*, 1986), was lying at a genetic distance of 4.8–6.5% (Normand *et al.*, 1996).

Rainey & Stackebrandt (1993), as part of their work on the reorganization of the phylogeny of the actinomycete line of descent, determined the *rrs* sequence of *A. cellulolyticus* and came to the conclusion that this bacterium was phyletically part of the actinomycetes although it was not closely related to any known bacterium. The high number of undetermined positions in the *A. cellulolyticus rrs* sequence could be construed as an indication that errors had been made elsewhere. Thus, a resequencing of the *rrs* gene was thought to be appropriate. Furthermore, since lateral transfer of genes has been demonstrated in bacteria (Hakenbeck *et al.*, 1998; Tamanai-Shacoori *et al.*, 1995), analysis of another ubiquitous but unlinked housekeeping gene was deemed useful to provide an independent confirmation of the unexpected proximity of *Acidothermus* to *Frankia*. Of the various genes recently used for that purpose, e.g. *gyrB*, *rpoB* and *tuf*, we selected *recA* (encoding a protein involved in repairing damaged DNA in the SOS regulon) because it is ubiquitous, highly conserved and has a phylogeny congruent, on the whole, with that of the *rrs* gene (Lloyd & Sharp, 1993; Eisen, 1995). Finally, strains belonging to the genus *Frankia* have been shown to have two diagnostic phenotypic features [i.e. the sugar

The GenBank/EMBL/DBJ accession number for the 16S rRNA sequence (with 21 corrections and 71 'N') described in this work is AJ007290.

Fig. 1. Phylogenetic RecA (left) and 16S rRNA (right) trees. The RecA amino acid distances (180 amino acids were considered) were corrected for multiple substitutions according to Nei (1987). The neighbour-joining method (Saitou & Nei, 1987) was used to derive a tree. The distances are expressed as substitutions per site and those groupings also obtained by parsimony (Kluge & Farris, 1969) are indicated by a 'P' beside the bootstrap (Felsenstein, 1985). Results are expressed as percentages. For comparative purposes, the 16S rRNA gene was used to derive a phylogenetic tree with distances computed according to Kimura (1980), using the neighbour-joining method. The distances are expressed as substitutions per site and the same code is used to characterize groupings. The accession numbers for the sequences used are given beside the species/strain names. In the case of ArI3, the sequence used was that of ACoN24d in the same genomic species, while in the case of HR40, the sequence used was that of HR27-14 at a 50% level of DNA-DNA hybridization (Fernandez *et al.*, 1989). The sequence of the 16S rRNA gene of *Frankia* strain ULFA was not determined because it had been found previously (Normand *et al.*, 1996) that the *Elaeagnus*-infective strains, like ULFA, have (at most) six substitutions over the whole of the 16S rRNA gene. The thick grey branches in the 16S rRNA tree are those for which the topology differs from that in the RecA tree. Bars, 0.02 substitutions/site.

2-*O*-methyl-D-mannose (Mort *et al.*, 1983) and large amounts of hopanoid lipids relative to the total cell lipid content] that distinguish members of this genus from most other bacteria described to date (Berry & Kleemann, 1994; Berry *et al.*, 1991). Consequently, these compounds were looked for in the *Frankia* phyletic neighbours.

Frankia strains were grown in liquid propionate BAP medium (Murry *et al.*, 1984). *A. cellulolyticus* was grown at 55 °C according to Mohagheghi *et al.* (1986), while *Geodermatophilus obscurus* subsp. *obscurus* was grown on both solid and liquid ATCC no. 172 media according to Luedemann (1968). Genomic DNA was released from thermal-shocked (three liquid nitrogen/boiling water passages) sonicated cells without further purification. Amplification and sequencing of the *rrs* gene was done as described previously (Normand *et al.*, 1996). The signature sequences defined by Stackebrandt *et al.* (1997) to differentiate family *Acidothermaceae* from other families in the class *Actinobacteria* remain valid. The distances obtained (not shown) and the topology of the resulting tree (Fig. 1) did not change significantly from those obtained before (Normand *et al.*, 1996), *Acidothermus* remaining the closest neighbour to *Frankia*, with *Geodermatophilus obscurus* subsp. *obscurus* being further away. These results were confirmed by the maximum-likelihood and parsimony methods. Thus, none of the GenBank sequences from cultivated micro-organisms is closest to *Frankia*, although it must be kept in mind that unisolated bacteria detected in the

rhizosphere of *Alnus viridis* are closer to *Frankia* (Normand & Chapelon, 1997).

The *rrs* gene has become the golden standard of bacterial phylogeny, more than 23 000 sequences having been determined so far with no sign of saturation yet. Numerous works have sought to confirm surprising results obtained with it using housekeeping protein genes such as *rpoB* (Mollet *et al.*, 1998), *gyrB* (Yamamoto & Hayarama, 1996), *tuf* (Cousineau *et al.*, 1992) and *recA*. The general conclusion that emerges from such comparisons is that although there are slight modifications, such as the non-monophyly of *Actinobacteria* (Stackebrandt *et al.*, 1997) with low G+C Gram-positives (Eisen, 1995; Galtier & Gouy, 1994), the overall topology remains. A set of primers was selected in conserved areas of the *recA* gene: primer FGPR-476 (5'-GAR-TCN-TCN-GGN-AAR-ACN-AC-3') corresponds to the conserved motif ESSGKTT at coordinates 68–74 of the *Streptomyces lividans* sequence, (accession no. X76076) and primer FGPR-477 (5'-GGN-GCN-ACY-TTR-TTY-TTR-AC-3') corresponds to the complement of the conserved motif VKNKVAP (at coordinates 247–253 of the *S. lividans* sequence). Amplifications were performed according to Mullis & Faloona (1987) and the optimal temperature was 53 °C. An amplicon of about 550 nucleotides was expected, corresponding to 50% of the whole gene. Sequencing was done on an automatic ABI sequencer (Applied Biosystems) using the amplification primers. Nearly full-length double-stranded sequences were

obtained for all strains tested. These have been deposited in EMBL and the accession numbers are given in Fig. 1. GenBank was scanned for related sequences using the algorithm BLAST (Altschul *et al.*, 1997) and the related sequences (*S. lividans*, *Mycobacterium tuberculosis*, *Corynebacterium glutamicum* and *Bacillus subtilis*) were included in the analyses. Sequences were aligned using CLUSTAL X (Thompson *et al.*, 1997). Indel-containing regions were excluded from the analyses. Matrix pairwise comparisons of nucleic acid sequences were corrected for multiple base substitutions according to Kimura (1980). Pairwise evolutionary distances between encoded polypeptides were computed using the Poisson correction for multiple substitutions (Nei, 1987). Phylogenetic trees were constructed by the neighbour-joining (Saitou & Nei, 1987), maximum-likelihood (Felsenstein, 1981) and parsimony (Kluge & Farris, 1969) methods. A bootstrap confidence analysis was performed on 1000 replicates to determine the reliability of the distance tree topologies obtained (Felsenstein, 1985). Graphic representations of the resulting trees was made using NJPLOT and PHYLO_WIN software (Perrière & Gouy, 1996). The nucleic acid distances between *Frankia* sequences (data not shown) were 0.8–19.5% substitutions/site and 1.9–10%. *A. cellulolyticus* is closest to the different *Frankia* strains, with DNA distances ranging from 25.3 to 28.9% and amino acid distances ranging from 9.3 to 15.8%. The tree obtained when all available *Actinobacteria* sequences were included and *Bacillus subtilis* was used as the outgroup is given in Fig. 1. It shows that the *A. cellulolyticus* strain and the five *Frankia* strains form a cluster but that *G. obscurus* is not closer to *Frankia* than either *Streptomyces* or *Mycobacterium*. This difference in topology between the *rrs* and *recA* trees was found to occur with nucleic acid sequences and with all phylogenetic reconstruction methods used. The two *Elaeagnus*-infective *Frankia* strains were clustered together, as were the two *Alnus*-infective strains and the *Casuarina*-infective strain. The topology of the nucleic acid distance tree obtained was similar to that obtained with amino acid distances except that the *M. tuberculosis* sequence was found within the *Frankia*–*Acidothermus* cluster. Study of the *recA* gene thus confirms the close proximity of *A. cellulolyticus* to *Frankia*. A lack of congruence was noted between the *rrs* and *recA* tree topologies (Fig. 1) with regard to the position of *Geodermatophilus* and that of the pair *Mycobacterium*–*Corynebacterium*. This may be due to the short length of the *recA* gene studied (550 nucleotides or about half of the gene), which can be misleading especially in a highly conserved gene.

Whole-cell sugars were determined on alditol acetate derivatives according to Sawardeker *et al.* (1965) on 8 mg lyophilized cells. Gas-phase chromatography was performed on Intersmat 120FL equipment with a silica SP2380 (Supelco) capillary (0.025 mm × 30 m) column and a flame-ionization detector. Large amounts of 2-*O*-methyl-D-mannose, as well as rhamnose, fucose,

ribose, arabinose, xylose, mannose, galactose, glucose and inositol, were found in the two *Frankia* strains studied. Of the two *Frankia* phyletic neighbours, a faint peak barely above the detection threshold was found only in *Geodermatophilus* grown on solid medium (and not in *Geodermatophilus* from liquid cultures). No 2-*O*-methyl-D-mannose was detected at all in *Acidothermus*. The sugar 2-*O*-methyl-D-mannose has not been found as a major component in the two phyletic neighbours analysed. Its diagnostic status for strains belonging to the genus *Frankia* (Mort *et al.*, 1983) thus remains valid.

Whole-cell lipids were determined according to Bligh & Dyer (1959) from 0.1 g lyophilized and sonicated cells. Total extracted lipids were dried, weighed and subjected to H₅IO₆/NaBH₄ treatment before derivatization by acetylation. Hopanoid lipids were identified by GC/MS and comparison with synthetic standards (Rohmer *et al.*, 1984). Quantification of total hopanoid lipids was performed using peak area integration. Large amounts of hopanoid lipids were found in the two *Frankia* strains studied (7.5% w/w and 22% w/w of whole-cell lipids in Ar13 and in ULFA, respectively) as well as in *Acidothermus*, in which 15% (w/w) of the total lipid content was hopanoid. In *Geodermatophilus*, the amounts found (3% w/w) were much lower.

According to the molecular clock calibration of Moran *et al.* (1993), the genus *Frankia* (which has 4.8–6.5% divergence in the *rrs* gene) would have shared a common ancestor with *A. cellulolyticus* 250–300 million years ago. This time estimate would predate the emergence of the earliest actinorhizal plants, [i.e. *Myrica* (Maggia & Bousquet, 1994), which is estimated (from studies of the pollen record) to have appeared 100–110 million years ago in the Cenomanian period (Macdonald, 1977)] but would follow the appearance of the first land plants 400 million years ago. The presence of large amounts of hopanoid lipids in a *Frankia* phyletic neighbour may indicate that they were also present in the common ancestor and that their protective function, the purpose of which remains unknown, later evolved towards protection of nitrogenase from oxygen diffusion in *Frankia* (Berry & Kleemann, 1994) and protection of *A. cellulolyticus* from the acidic thermal environment in which it grows. Large amounts of hopanoid lipids, such as those reported for *Frankia* and *Acidothermus*, have been found so far only in *Zymomonas mobilis*, which grows at high alcohol concentrations (Hermans *et al.*, 1991), and in *Alicyclobacillus acidocaldarius*, which also grows in thermal acidic springs (Poralla *et al.*, 1980). The presence of multiple cellulases in *A. cellulolyticus* highlights the ability of this hypothetical ancestor to thrive in the cellulose-rich rhizosphere of primitive land plants. The presence of comparatively low levels of cellulose in the interface capsule that surrounds *Frankia* in symbiotic plant cells has already been described (Berg, 1990) and could be due to the ability of *Frankia* to degrade cellulose (Safo-Sampah &

Torrey, 1988). This aspect could be studied again in the light of the close relationship between *Frankia* and *A. cellulolyticus* confirmed by the present study.

References

- Altschul, S. F., Madden, T. L., Schäffer, A. A., Zhang, J., Zhang, Z., Miller, W. & Lipman, D. J. (1997). Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res* **25**, 3389–3402.
- Benson, D. R. & Silvester, W. B. (1993). Biology of *Frankia* strains, actinomycete symbionts of actinorhizal plants. *Microbiol Rev* **57**, 293–319.
- Berg, R. H. (1990). Cellulose and xylans in the interface capsule in symbiotic cells of actinorrhizae. *Protoplasma* **159**, 35–43.
- Berry, A. M., Moreau, R. A. & Jones, A. D. (1991). Bacteriohopanetetrol: abundant lipid in *Frankia* cells and in nitrogen-fixing nodule tissue. *Plant Physiol* **95**, 111–115.
- Berry, A. M. & Kleemann, G. (1994). Hopanoid lipids in relation to nitrogen fixation: function in *Frankia* and occurrence in other diazotrophic organisms. In *Nitrogen Fixation with Non-legumes*, pp. 191–200. Edited by N. A. Hegazi, M. Fayez & M. Monib. Cairo: American University in Cairo Press.
- Bligh, E. G. & Dyer, W. J. (1959). A rapid method of total lipid extraction and purification. *Can J Biochem Physiol* **37**, 911–917.
- Cousineau, B., Cerpa, C., Lefebvre, J. & Cedergren, R. (1992). The sequence of the gene encoding elongation factor Tu from *Chlamydia trachomatis* compared with those of other organisms. *Gene* **120**, 33–41.
- Eisen, J. A. (1995). The RecA protein as a model molecule for molecular systematic studies of bacteria: comparison of trees of RecAs and 16S rRNAs from the same species. *J Mol Evol* **41**, 1105–1123.
- Felsenstein, J. (1981). Evolutionary trees from DNA sequences: a maximum-likelihood approach. *J Mol Evol* **17**, 368–376.
- Felsenstein, J. (1985). Confidence limits on phylogenies: an approach using the bootstrap. *Evolution* **39**, 783–791.
- Fernandez, M. P., Meugnier, H., Grimont, P. A. D. & Bardin, R. (1989). Deoxyribonucleic acid relatedness among members of the genus *Frankia*. *Int J Syst Bacteriol* **39**, 424–429.
- Galtier, N. & Gouy, M. (1994). Molecular phylogeny of Eubacteria: a new multiple tree analysis method applied to 15 sequence data sets questions the monophyly of Gram-positive bacteria. *Res Microbiol* **145**, 531–541.
- Hahn, D., Lechevalier, M. P., Fischer, A. & Stackebrandt, E. (1989). Evidence for a close phylogenetic relationship between members of the genera *Frankia*, *Geodermatophilus*, and “*Blastococcus*” and emendation of the family Frankiaceae. *Syst Appl Microbiol* **11**, 236–242.
- Hakenbeck, R., Konig, A., Kern, I., van der Linden, M., Keck, W., Billot-Klein, D., Legrand, R., Schoot, B. & Gutmann, L. (1998). Acquisition of five high- M_r penicillin-binding protein variants during transfer of high-level beta-lactam resistance from *Streptococcus mitis* to *Streptococcus pneumoniae*. *J Bacteriol* **180**, 1831–1840.
- Hermans, M. A. F., Neuss, B. & Sahm, H. (1991). Content and composition of hopanoids in *Zymomonas mobilis* under various growth conditions. *J Bacteriol* **173**, 5592–5595.
- Kimura, M. (1980). A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. *J Mol Evol* **16**, 111–120.
- Kluge, A. G. & Farris, J. S. (1969). Quantitative phyletics and the evolution of anurans. *Syst Zool* **18**, 1–32.
- Lloyd, A. T. & Sharp, P. M. (1993). Evolution of the *recA* gene and the molecular phylogeny of bacteria. *J Mol Evol* **37**, 399–407.
- Luedemann, G. M. (1968). *Geodermatophilus*, a new genus of the Dermatophilaceae (Actinomycetales). *J Bacteriol* **96**, 1848–1858.
- Macdonald, A. D. (1977). Myricaceae: floral hypothesis for *Myrica gale* and *Comptonia*. *Can J Bot* **55**, 2636–2651.
- Maggia, L. & Bousquet, J. (1994). Molecular phylogeny of the actinorhizal Hamamelidae and relationships with host promiscuity towards *Frankia*. *Mol Ecol* **3**, 459–467.
- Mohagheghi, A., Grohmann, K., Himmel, M., Leighton, L. & Updegraff, D. M. (1986). Isolation and characterization of *Acidothermus cellulolyticus* gen. nov., sp. nov., a new genus of thermophilic, acidophilic cellulolytic bacteria. *Int J Syst Bacteriol* **36**, 435–443.
- Mollet, C., Drancourt, M. & Raoult, D. (1998). Determination of *Coxiella burnetii* *rpoB* sequence and its use for phylogenetic analysis. *Gene* **207**, 97–103.
- Moran, N. A., Munson, M. A., Baumann, P. & Ishikawa, H. (1993). A molecular clock in endosymbiotic bacteria is calibrated using the insect hosts. *Proc R Soc Lond B* **253**, 167–171.
- Mort, A., Normand, P. & Lalonde, M. (1983). 2-O-Methyl-D-mannose, a key sugar in the taxonomy of *Frankia*. *Can J Microbiol* **29**, 993–1002.
- Mullis, K. B. & Faloona, F. A. (1987). Specific synthesis of DNA in vitro via a polymerase-catalyzed chain reaction. *Methods Enzymol* **155**, 335–350.
- Murry, M. A., Fontaine, M. S. & Torrey, J. G. (1984). Growth kinetics and nitrogenase induction in *Frankia* sp. HFP ArI3 grown in batch culture. *Plant Soil* **78**, 61–78.
- Nei, M. (1987). *Molecular Evolutionary Genetics*. New York: Columbia University Press.
- Normand, P. & Chapelon, C. (1997). Direct characterization of *Frankia* and of close phyletic neighbors from an *Alnus viridis* rhizosphere. *Physiol Plant* **99**, 722–731.
- Normand, P., Orso, S., Cournoyer, B., Jeannin, P., Chapelon, C., Dawson, J., Evtushenko, L. & Misra, A. K. (1996). Molecular phylogeny of the genus *Frankia* and related genera and emendation of the family Frankiaceae. *Int J Syst Bacteriol* **46**, 1–9.
- Perrière, G. & Gouy, M. (1996). WWW-Query: An on-line retrieval system for biological sequence banks. *Biochimie* **78**, 364–369.
- Poralla, K., Kannenberg, E. & Blume, A. (1980). A glycolipid containing hopane isolated from the acidophilic, thermophilic *Bacillus acidocaldarius*, has a cholesterol-like function in membranes. *FEBS Lett* **113**, 107–110.
- Rainey, F. A. & Stackebrandt, E. (1993). Phylogenetic evidence for the classification of *Acidothermus cellulolyticus* into the subphylum of actinomycetes. *FEMS Microbiol Lett* **108**, 27–30.
- Rohmer, M., Bouvier-Nave, P. & Ourisson, G. (1984). Distribution of hopanoid triterpenes in prokaryotes. *J Gen Microbiol* **130**, 1137–1150.
- Safo-Sampah, S. & Torrey, J. G. (1988). Polysaccharide-hydrolyzing enzymes of *Frankia* (Actinomycetales). *Plant Soil* **112**, 89–97.
- Saitou, R. R. & Nei, M. (1987). The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol Biol Evol* **44**, 406–425.

- Sawardeker, J. J., Sloneker, J. H. & Jeanes, A. R. (1965). Quantitative determination of mono-saccharides as their alditol acetates by gas-liquid chromatography. *Anal Biochem* **12**, 1602–1604.
- Soltis, D. E., Soltis, P. S., Morgan, D. R., Swensen, S. M., Mullin, B. C., Dowd, J. M. & Martin, P. G. (1995). Chloroplast gene sequence data suggest a single origin of the predisposition for symbiotic nitrogen fixation in angiosperms. *Proc Natl Acad Sci USA* **92**, 2647–2651.
- Stackebrandt, E., Rainey, F. A. & Ward-Rainey, N. L. (1997). Proposal for a new hierarchic classification system, *Actinobacteria* classis nov. *Int J Syst Bacteriol* **47**, 479–491.
- Tamanai-Shacoori, Z., Arturo, M., Pommepuy, M., Mamez, C. & Cormier, M. (1995). Conjugal transfer of natural plasmids between *Escherichia coli* strains in sterile environmental water. *Curr Microbiol* **30**, 155–160.
- Thompson, J. D., Gibson, T. J., Plewniak, F., Jeanmougin, F. & Higgins, D. G. (1997). The CLUSTAL-X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Res* **25**, 4876–4882.
- Yamamoto, S. & Harayama, S. (1996). Phylogenetic analysis of *Acinetobacter* strains based on the nucleotide sequences of *gyrB* genes and on the amino acid sequences of their products. *Int J Syst Bacteriol* **46**, 506–511.