JOURNAL OF BACTERIOLOGY, Mar. 1995, p. 1414–1417
0021-9193/95/$04.00+0

Vol. 177, No. 5

Remarkable N_2-Fixing Bacterial Diversity Detected in Rice Roots by Molecular Evolutionary Analysis of *nifH* Gene Sequences

TOHRU UEDA,* YUKO SUGA, NOBUTAKA YAHIRO, AND TATSUHIKO MATSUGUCHI

Laboratory of Soil Microbiology and Biochemistry, Department of Agricultural Chemistry, Kyushu University, Fukuoka, Japan

Received 11 July 1994/Accepted 4 January 1995

To demonstrate the extent of phylogenetic diversity of diazotrophic bacteria associated with rice roots, we characterized phylogenetically 23 *nifH* gene sequences obtained by PCR amplification of mixed organism DNA extracted directly from rice roots without culturing the organisms. The analyses document the presence of eight novel NifH types, which appear to be a variety of significant components of the diazotrophic community, dominated mainly by proteobacteria.

The method developed by Pace et al. to determine species diversity and composition, using rRNA (or ribosomal DNA) isolated directly from nature has opened a window into the world of unculturable bacteria, although this method has the disadvantage that organisms whose genes have been isolated by the method cannot be studied for any other trait (11, 16). This report extends the method by applying it to a functionally important gene, *nifH*, to determine the importance and diversity of nitrogen-fixing bacteria associated with rice roots. The nitrogenase iron protein gene *nifH* is one of the oldest existing and functioning genes in the history of gene evolution, and the outline of the NifH tree is reported to be largely consistent with the 16S rRNA phylogeny (21, 22). These features prompted us to study the genetic diversity of N_2-fixing bacteria by the molecular evolutionary analysis of *nifH* sequences amplified directly from rice root DNA, because the rice root DNA contains not only plant DNA but also microbial DNA in the roots. In the rice root zone, N_2 fixation is associated with the activity of N_2-fixing heterotrophic bacteria (4, 20). It has been reported that a large percentage of the total aerobic heterotrophic population in the root is diazotrophic (1, 10, 17). However, studies on the rhizospheric N_2-fixing microflora have until now suffered from the use of selective media for counts and isolations, because it is widely believed that only a small percentage of natural prokaryotes may actually be culturable (18, 19). Here we report a study of N_2-fixing bacterial diversity by analysis of *nifH* gene sequences without a cultivation technique.

DNA extraction. Rice, *Oryza sativa* L. cv. *nihonnbare*, was raised under flooded conditions in the Kyushu University Farm. Rice plants taken at the heading stage in September were dug out from a wetland rice field, and the roots were washed to remove the attached soil. The washed roots were cut into segments, frozen with liquid nitrogen, and ground to a fine powder in a mortar and pestle. The fine powder was suspended in extraction buffer (100 mM Tris, 100 mM EDTA, 250 mM NaCl, 100 μg of proteinase K per ml) supplemented with Sarkosyl (1% final concentration) and lysed by incubation at 55°C for 1 h. Treatment of the lysate with RNase A was followed by chloroform extraction and isopropanol precipitation. Crude DNA was purified by phenol extraction, chloroform extraction, and isopropanol precipitation.

PCR amplification of *nifH* genes. The primers for PCR amplification were chosen by careful inspection of the 37 published *nifH* sequences available in the GenBank, EMBL, and DDBJ DNA databases. The sequences of the amplification primers are GCIWTYTAYGGIAARGGIGG for 19F and AAICCRCCRCAIACIACRTC for 407R (where I represents inosine, R represent A or G, W represents A or T, and Y represents C or T). 19F was a putative ATP-binding region. Approximately 390-bp fragments of *nifH* were amplified between nucleotides 19 and 407 (*Azotobacter vinelandii* M20568 numbering) from rice root DNA. The reaction conditions with 100 ng of template DNA were 0.5 min at 94°C, 1 min at 50°C, and 0.5 min at 72°C for 40 cycles. Negative controls (water used instead of DNA) showed no amplification. Then, the products were purified, and the *nifH* fragments were cloned into vector pT7BlueT to construct a *nifH* library by using *Escherichia coli* JM109. Twenty-three cloned *nifH* genes chosen at random from the library were sequenced. Analysis of double-stranded DNA with the universal M13 primer and the T7 primer was done by using the *Taq* Dye Primer Cycle Sequencing Kit and the DNA sequencer model 373A (Applied Biosystems, Foster City, Calif.) in accordance with the manufacturer's directions. We called these clones H-RIC1 to H-RIC23 (from *nifH* in rice roots).

None of the 23 analyzed H-RIC sequences is identical to a published sequence. When the 23 sequences were aligned by visual inspection, several regions were found to be conserved in all sequences considered, whereas others were found to be divergent (data not shown). The three conserved cysteine residues (Cys-39, Cys-86, and Cys-98; *A. vinelandii* M20568 numbering), which correspond to the ligands for the iron-sulfur cluster, and conserved Arg-101, necessary for reversible inactivation through ADP ribosylation in *A. vinelandii* (2), were found to be conserved in all sequences. The location of sequence changes and the highest error rate reported for *Taq* DNA polymerase (1 in 400 after 30 amplification cycles) (12) suggest that errors and artefacts cannot account for a significant proportion of the observed variation. In other words, it appeared to be certain that each sequence comes from a different *nifH* gene of N_2-fixing bacteria in rice roots.

Phylogenetic analysis. All known 37 different *nifH* gene and

* Corresponding author. Mailing address: Laboratory of Soil Microbiology and Biochemistry, Department of Agricultural Chemistry, Kyushu University, Hakozaki 6-10-1, Fukuoka, 812 Japan. Phone: 092-641-1101, ext. 6421. Fax: 092-632-1960.

6 chlorophyll-iron protein gene sequences were extracted from the GenBank, EMBL, and DDBJ DNA databases by using the ODEN system (5). The deduced partial amino acid sequences were then aligned with each other to maximize the sequence similarity by using the ODEN multiple-sequence alignment program (5) and by visual inspection. In estimating evolutionary distances, we discarded the regions common to the 66 sequences in which deletions were observed even for 1 sequence, since there is no way to assess the rate and size of deletions during the course of evolution. Omission of the regions left us with 106 comparable amino acid sites. To estimate evolutionary distances between amino acid sequences, we used Poisson correction. By using these estimates, molecular evolutionary trees were constructed from the resulting distance matrix by the neighbor-joining (NJ) method, which does not require the assumption of a constant substitution rate (13), and the unweighted-pair group method using arithmetic averages, which requires the assumption of a constant substitution rate (8, 14), with the MEGA package (6) as shown in Fig. 1. In these analyses the chlorophyll-iron proteins were used as outgroups. The reliability of the tree nodes was analyzed by using the bootstrapping program (3, 6). The percentages of 500 bootstrap resamplings that support each topological element more than 50% are indicated in Fig. 1.

As shown in Fig. 1A and B, for the most part both trees have the same topology, except for the positions of *Frankia* spp. and cyanobacteria. These data suggest that the main points of the topology of this tree may be reliable and that the iron protein family evolved at a relatively constant rate. Furthermore, in agreement with the findings of Young (21, 22), the topology of the tree of NifH protein sequences derived by the NJ method is similar to that of 16S rRNA sequences (7) in several points. (i) Both NJ analyses clustered archaea with gram-positive bacteria, cyanobacteria, and then proteobacteria, successively (though slightly) more distantly related in a rooted tree. (ii) The evolutionary distances between the branching points leading to the clusters visible in both trees are very small for the major clusters of the bacteria. (iii) Cyanobacteria are closely related to gram-positive bacteria. (iv) As for the division of gram-positive bacteria, its two subdivisions by high and low G+C contents together do not form a monophyletic cluster. (v) Both trees derived by the NJ method can be used to assign bacterial species to a given division because the appurtenance of each bacterial species to a particular cluster is stable.

Of course, it must be noted that some of the branching orders of the groups differ to some extent. For example, gram-positive bacteria with low G+C contents are associated with the delta subdivision of proteobacteria. In such case there may be some possibility that horizontal gene transfer might have occurred, as suggested by Normand and Bousquet (9). However, although some of the branching orders of the groups differ to some extent, it is important that the outline of the groups established and the deeper-branching order are largely congruent. Therefore, it can be assumed that a bacterium with a sequence of the *nifH* gene which falls in some known cluster might be related to the bacteria in that cluster.

As shown in Fig. 1, the analysis indicated the presence of eight novel NifH types (H-RIC3,12; H-RIC1 cluster; H-RIC20; H-RIC17; H-RIC21; H-RIC14,19; H-RIC7 cluster; and H-RIC15), which appear to be a variety of significant components of the rice rhizosphere diazotrophic community. These nitrogen-fixing bacteria may be located on the root surface (the rhizoplane) or inside the root. Among these eight groups, there are two major clusters: H-RIC1 and H-RIC7. The H-RIC1 cluster, with eight clones, branched within the delta subdivision of proteobacteria, *Desulfovibrio gigas*. These data indicate a high degree of diversity of *nifH* gene lineages among the related delta subdivision of proteobacteria. A second large cluster (the H-RIC7 cluster) of lineages related to the gamma subdivision of proteobacteria was also observed. Seven clones belong to this cluster. Although the H-RIC7 cluster was most similar to *Klebsiella pneumoniae*, genetic distances suggest that these sequences may not belong to this genus. Within those clusters we have observed unexpected genetic variability. We speculate that some of this sequence diversity may result from the essentially clonal nature of bacterial reproduction which correlated with habitat variability. We also speculate that some of this sequence diversity may result from the presence of multiple copies of *nifH* genes, since some of N_2-fixing bacteria have alternative nitrogenase genes and several copies of *nifH* (2). For example, *Clostridium pasteurianum* has six copies of *nifH* genes including an alternative nitrogenase gene (*anfH*).

In addition, H-RIC17, H-RIC21, H-RIC14, and H-RIC19 might also belong to the gamma subdivision of proteobacteria. It is worthwhile to note that H-RIC15 is closely related to the *Azotobacter nifH* family. Although many reports have indicated the widespread presence of *Azospirillum* spp., members of the alpha subdivision of proteobacteria (1, 10), we could not find *nifH* clones other than H-RIC20 in the alpha subdivision of proteobacteria in Fig. 1. Therefore, it remains to be clarified whether the genus *Azospirillum* is the predominant N_2-fixing bacterium in the rice rhizosphere. However, of course, it must be noted that choosing 23 clones randomly does not ensure that every unique clone will be found. This procedure ensures that abundant clones will be favored and that rare clones may be missed.

Moreover, we could not amplify *nifH* genes from crude DNA extracted from soil around the rice roots by using a freeze-thaw procedure (15). Because we could amplify 16S rRNA genes from the soil DNA and the necessary positive controls were performed, this appeared to be a real result and not a technical problem (data not shown). These data suggest that considerable amounts of N_2-fixing bacteria exist on the root surface or inside the root, compared with the soil around the rice roots. Furthermore, we also tried to amplify *nifH* segments from DNA extracted from maize roots and soybean roots (not nodules), but we could not amplify *nifH* from these DNA preparations (data not shown). These data suggest that many more N_2-fixing bacteria exist in wetland rice roots than in the dryland plant roots. This could be the cause of higher maintenance of soil fertility in flooded soil.

It must be noted here that the existence of *nifH* genes does not always mean the activity of nitrogenase, since this enzyme is regulated at both pre- and posttranslational levels (2). It must also be noted that the distribution of clones in final PCR products may not match the distribution of all *nifH* genes in the original rice root DNA because of differences in amplification efficiency. Hybridization analysis using specific oligonucleotide probes to RNA or DNA isolated from the rice roots will be needed to study how prevalent these organisms are in the roots. The sequences described here include several regions suitable for use as specific probes. In addition, a similar study using *nifD* sequences instead of *nifH* sequences is now in progress and will be published elsewhere in the near future.

Nucleotide sequence accession numbers. Sequences from the clones reported here have been submitted to the DDBJ nucleotide sequence database under accession no. D26284 to D26306.

FIG. 1. Phylogeny of NifH protein sequences, including rice clones, by NJ method (A) and unweighted-pair group method with arithmetic averages (B). The database accession numbers are indicated after the bacterial names. The percentages of 500 bootstrap resamplings that support each topological element more than 50% are indicated on the left.

We thank Y. Ina (National Institute of Genetics) for the sequence analysis program and kind advice, M. Kumar (Pennsylvania State University) for providing the sequence analysis program, M. Sakai (Kyushu University) for valuable advice, and T. Motizuki (Kyushu University) for providing rice plants.

REFERENCES

1. **Bally, R., D. Thomas-Bauzon, T. Heulin, J. Balandreau, C. Richard, and J. D. Ley.** 1983. Determination of the most frequent N_2-fixing bacteria in a rice rhizosphere. Can. J. Microbiol. **29:**881–887.
2. **Dean, D. R., and M. R. Jacobson.** 1992. Biochemical genetics of nitrogenase, p. 763–834. *In* G. Stacy, R. H. Burris, and H. J. Evans (ed.), Biological nitrogen fixation. Chapman and Hall, New York.
3. **Felsenstein, J.** 1985. Confidence limits on phylogenies: an approach using the bootstrap. Evolution **39:**783–791.
4. **Hirota, Y., T. Fujii, Y. Sano, and S. Iyama.** 1978. Nitrogen fixation in the rhizosphere of rice. Nature (London) **276:**416–417.
5. **Ina, Y.** 1991. ODEN: molecular evolutionary analysis system for DNA and amino acid sequences, version 1.1.1. DNA Research Center, National Institute of Genetics, Mishima, Japan.
6. **Kumar, S., K. Tamura, and M. Nei.** 1993. MEGA: molecular evolutionary genetics analysis, version 1.01. Pennsylvania State University, University Park.
7. **Neefs, J.-M., Y. V. Peer, P. D. Rijk, S. Chapelle, and R. D. Wachter.** 1993. Compilation of small ribosomal subunit RNA structures. Nucleic Acids Res. **21:**3025–3049.
8. **Nei, M.** 1975. Molecular population genetics and evolution. North-Holland, Amsterdam.
9. **Normand, P., and J. Bousquet.** 1989. Phylogeny of nitrogenase sequences in *Frankia* and other nitrogen-fixing microorganisms. J. Mol. Evol. **29:**436–447.
10. **Oyaizu-Masuchi, Y., and K. Komagata.** 1988. Isolation of free-living nitrogen-fixing bacteria from the rhizosphere of rice. J. Gen. Appl. Microbiol. **34:**127–164.
11. **Pace, N. R., D. A. Stahl, D. J. Lane, and G. J. Olsen.** 1986. The analysis of natural microbial populations by ribosomal RNA sequences. Adv. Microbiol. Ecol. **9:**1–55.
12. **Saiki, R. K., D. H. Gelfand, S. Stoffel, S. J. Scharf, R. Higuchi, G. T. Horn, K. B. Mullis, and H. A. Ehrlich.** 1988. Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase. Science **239:**487–491.
13. **Saitou, N., and M. Nei.** 1987. The neighbor-joining method: a new method for reconstructing phylogenetic trees. Mol. Biol. Evol. **4:**406–425.
14. **Sneath, P. H. A., and R. R. Sokal.** 1973. Numerical taxonomy. Freeman, San Francisco.
15. **Tsai, Y. L., and B. H. Olson.** 1991. Rapid method for direct extraction of DNA from soil and sediments. Appl. Environ. Microbiol. **57:**1070–1074.
16. **Ward, D. M., M. M. Bateson, R. Weller, and A. L. Ruff-Roberts.** 1992. Ribosomal RNA analysis of microorganisms as they occur in nature. Adv. Microb. Ecol. **12:**219–286.
17. **Watanabe, I., and W. Barraquio.** 1979. Low levels of fixed nitrogen required for isolation of free-living N_2-fixing organisms from rice roots. Nature (London) **277:**565–566.
18. **Wayne, L. G., D. J. Brenner, R. R. Colwell, P. A. D. Grimont, O. Kandler, M. I. Krichevsky, L. H. Moore, W. E. C. Moore, R. G. E. Murray, E. Stackebrandt, M. P. Starr, and H. G. Trüper.** 1987. Report of the Ad Hoc Committee on Reconciliation of Approaches to Bacterial Systematics. Int. J. Syst. Bacteriol. **37:**463–464.
19. **Wolfe, R. S.** 1981. Foreword, p. v–vi. *In* M. P. Starr, H. Stolp, H. G. Trüper, A. Balows, and H. G. Schlegel (ed.), The Procaryotes, vol. I. Springer-Verlag, Berlin.
20. **Yoshida, T., and R. R. Ancajas.** 1973. The fixation of nitrogen in the rice rhizosphere. Soil Biol. Biochem. **5:**153–155.
21. **Young, J. P. W.** 1992. Phylogenetic classication of nitorogen-fixing organisms, p. 43–86. *In* G. Stacy, R. H. Burris, and H. J. Evans (ed.), Biological nitrogen fixation. Chapman and Hall, New York.
22. **Young, J. P. W.** 1993. Molecular phylogeny of rhizobia and their relatives, p. 587–592. *In* R. Palacios, J. Mora, and W. E. Newton (ed.), New horizons in nitrogen fixation. Kluwer Academic Publishers, London.