Distribution of *Frankia* genotypes occupying *Alnus nepalensis* nodules with respect to altitude and soil characteristics in the Sikkim Himalayas

Anindita Khan\(^a\), David D. Myrold\(^b\) and Arvind K. Misra\(^{a,*}\)

\(^a\)Molecular Genetics Laboratory, Department of Botany, North Eastern Hill University, Shillong-793022, Meghalaya, India

\(^b\)Department of Crop and Soil Science, Oregon State University, Corvallis, OR 97331, USA

Correspondence
*Corresponding author, e-mail: arvindkmisra@nehu.ac.in

Received 9 September 2006; revised 28 November 2006
doi: 10.1111/j.1399-3054.2006.00872.x

We investigated the distribution of *Frankia* genotypes with respect to three altitudinal zones in the Sikkim Himalayas. The study was carried out for 90 *Alnus nepalensis* trees at nine different locations from three altitudes from the east and north districts of Sikkim, India. We used a PCR-based technique to amplify the internally transcribed spacer (ITS) region of the *rrn* operon using two primers specific for the distal part of 16S ribosomal RNA (rRNA), ITS and proximal part of 23S rRNA genes of *Frankia*. The PCR products were digested with the restriction enzyme *Rsa*I to generate amplified recombinant DNA restriction analysis (ARDRA) patterns. *Frankia* genotypes were categorized on the basis of 17 different ARDRA patterns. Nucleotide sequences of some representative amplicons were also obtained. Rhizosphere soil samples associated with the 90 trees were collected and analysed for eight different soil properties. In general, soil properties did not differ by site or altitude, and did not correlate with *Frankia* genotypes. *Frankia* community composition was strongly affected by altitude and to a lesser extent by site.

Introduction

The actinomycete *Frankia* is of fundamental and ecological interest for several reasons including its wide distribution, its ability to fix nitrogen and its capacity to nodulate plants belonging to about 24 genera (Lavire and Cournoyer 2003). The nitrogen-fixing symbiosis between *Frankia* and actinorhizal plants contributes greatly to the global nitrogen cycle and plays a significant role in biological nitrogen fixation. Actinorhizal plants also play a major role in agroforestry. Better utilization of this symbiotic relationship requires better understanding about properties, such as its persistence in the soil and competitiveness of related strains, which requires precise strain identification (Normand et al. 1992).

Besides molecular genetics, studies on the molecular ecology of nitrogen-fixing microbes in the rhizosphere must be a component of research about biological nitrogen fixation. We must understand the effect of ecological parameters on the molecular constitution and behaviour of both partners. Because very little is known about the ecology of frankiae independent of their host plants, there is a need to develop increased understanding of the microbial genes, and plant and soil factors involved in rhizosphere interactions.

Earlier investigations had shown genetic diversity in strains of *Frankia* isolated from nodules of *Alnus nepalensis* growing in India (Ganesh et al. 1994). Several workers have utilized the PCR-restriction fragment length

---

**Abbreviations** – ARDRA, amplified recombinant DNA restriction analysis; EDTA, ethylenediaminetetraacetic acid; ITS, internally transcribed spacer; MRPP, multiresponse permutation procedure; NMS, non-metric multidimensional scaling; PCR, polymerase chain reaction; RFLP, restriction fragment length polymorphism; rRNA, ribosomal RNA; SNP, single nucleotide polymorphism.
polymorphism (RFLP) based approach for characterizing Frankia strains (Jamann et al. 1993, Ritchie and Myrold 1999, Rouvier et al. 1996, Varghese et al. 2003a, Verghese and Misra 2000). The PCR-RFLP involves restriction digestion of an amplified segment of DNA to generate restriction patterns. When applied to the PCR amplification of ribosomal genes and subsequent RFLP analysis, this method has been termed amplified recombinant DNA (rDNA) restriction analysis (ARDRA) (Vaneechoutte et al. 1992). The present study was designed to utilize the discriminative power of ARDRA and nucleotide sequences in assessing the altitudinal distribution of Frankia strains in Sikkim Himalayas.

Materials and methods

In the present study, ARDRA was used for discriminating among different Frankia strains. The ribosomal RNA (rRNA) encoding genes with their conserved areas that alternate with highly variable zones provide targets for discriminative studies (Normand et al. 1992, Varghese et al. 2003b).

Study area

This study was conducted in Sikkim, which is situated in northeast India, surrounded by Nepal in the west, China in the north, West Bengal in the south and both Bhutan and China in the east. Sikkim is part of a hot spot of biodiversity, and is situated approximately between 88°E to 89°E and 27°N to 28°N, and ranges from 300 to 8545 m above mean sea level. Nine different locations were selected at three different sites and three different altitude ranges (Table 1).

Sample collection

Nodule and rhizosphere soil samples were collected from 90 randomly selected A. nepalensis trees representing 10 trees from each location. Alder nodules were collected in the month of October. Weeds around the trees were cleared and the soil was carefully excavated to expose the roots. Only fresh, light-brown nodule clusters were collected. Soil samples associated with the alder roots were collected from 0- to 15-cm depth. Two to three drops of toluene were added to the soil samples to stop further microbial activity. Nodule and soil samples were collected in separate plastic bags and brought back to the laboratory. On arrival to the lab, nodule clusters from each tree were washed thoroughly under running tap water, air dried and stored at −20°C for further use. Soil samples were air dried, sieved and stored in fresh plastic bags at room temperature.

Soil analysis

The following soil chemical properties were measured on all samples: pH and electrical conductivity in distilled water (1.25 w/v), organic carbon (Walkley and Black 1934), total Kjeldahl nitrogen (Allen et al. 1974), extractable phosphorus (Allen et al. 1974), available potassium by the neutral ammonium acetate method (Allen et al. 1974), exchangeable calcium by the ammonium acetate extractant method (Hesse 1971) and exchangeable magnesium by direct titration with ethylenediaminetetraacetic acid (EDTA). This multivariate data set of soil chemical properties was evaluated by the multiresponse permutation procedure (MRPP) using PC-Ord software (MJJ Software Design, Gleneden Beach, OR). Individual soil chemical properties were analysed by analysis of variance using a one-way blocked design, with sites as blocks using SAS v. 8.2 (SAS Institute Inc., Cary, NC). When appropriate, the Student–Newman–Keuls test was used to detect differences among altitudes.

DNA isolation

Extraction of total genomic DNA from individual nodule lobes was carried out following the method given by Rouvier et al. (1996) with minor modifications. A single nodule lobe was first sterilized in 30% hydrogen peroxide (H₂O₂) for 1 or 2 min. After that it was washed in autoclaved distilled water three to four times. Then the outer epidermis was peeled off with a sterilized needle and the nodule lobe was transferred to a 1.5-ml microcentrifuge tube. The lobe was then crushed properly and about 350 µl of warm DNA extraction buffer [0.1 M Tris–HCl, 1.4 M NaCl, 0.02 M EDTA, 2% (w/v) cetyltrimethylammonium bromide and 1% (w/v) polyvinyl pyrrolidone] along with 10 µl of 20% (w/v) sodium lauryl sulphate. The homogenate was incubated in a water bath for 1 h at 65°C. After the incubation, the tube was centrifuged at 6081.9 g for

<table>
<thead>
<tr>
<th>Site</th>
<th>Altitude</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Low</td>
<td>Ranipool, East Sikkim; 980 m</td>
</tr>
<tr>
<td></td>
<td>Medium</td>
<td>BSI Complex, Gangtok, East Sikkim; 1710 m</td>
</tr>
<tr>
<td></td>
<td>High</td>
<td>Hanumantok, Gangtok, East Sikkim; 2100 m</td>
</tr>
<tr>
<td>2</td>
<td>Low</td>
<td>Rumtek, East Sikkim; 980 m</td>
</tr>
<tr>
<td></td>
<td>Medium</td>
<td>Rateychu, East Sikkim; 1710 m</td>
</tr>
<tr>
<td></td>
<td>High</td>
<td>Himalayan Zoological Park, Gangtok, East Sikkim; 2200 m</td>
</tr>
<tr>
<td>3</td>
<td>Low</td>
<td>Mangan, North Sikkim; 1020 m</td>
</tr>
<tr>
<td></td>
<td>Medium</td>
<td>Chunthang, North Sikkim; 1700 m</td>
</tr>
<tr>
<td></td>
<td>High</td>
<td>Maltin, North Sikkim; 2150 m</td>
</tr>
</tbody>
</table>
about 7 min at room temperature. The supernatant was transferred to a fresh tube and an equal volume of chloroform–isoamyl alcohol (24:1) was added to it. The tube was centrifuged again at 16060 g for 15 min at room temperature. The upper aqueous layer was transferred to another fresh tube and 2.5 volumes of ice-cold ethanol were added. It was kept overnight at −20°C for precipitation. The next day the tube was centrifuged at 16060 g for 30 min at 4°C. Ethanol was discarded and the DNA precipitate was washed twice with 350 μl of 70% alcohol. Later the DNA precipitate was vacuum dried, dissolved in ultrapure water and stored at −20°C for future use.

**Amplification of genomic DNA by PCR**

Isolated DNA samples were subjected to PCR amplification. Each reaction mixture contained 2.5 μM primer (Microsynth, Switzerland), 2.5 μl of 10× PCR buffer (Bangalore Genei, India), 2.5 μl of 25 μM MgCl₂ (Bangalore Genei, India), 10 μl of 5 mM dNTP mix (Bangalore Genei, India), 0.75 μl of Taq polymerase (3 units μl⁻¹) (Bangalore Genei, India), 1 μl of template DNA and ultrapure water to make the total volume to 25 μl of PCR mix per tube. Each amplification reaction was carried out for 35 cycles using a thermal cycler (Gene Amp PCR 2400, Perkin–Elmer). Each cycle was comprised of 1-min denaturation at 94°C, 1-min annealing at 49°C and 1 min of elongation at 72°C. A hot start was given for 5 min at 94°C and at the end of the run an additional 7-min extension time at 72°C was added to allow complete extension of all DNA double strands. The reaction mixture for amplification was prepared in strictly aseptic conditions under an ultraviolet light. Proper care was taken to prevent contamination by aerosol DNA. The mix was prepared in thin-walled PCR tubes. Amplification of rm internally transcribed spacer (ITS) was carried out using forward primer FGPS 989ac (5′GGGGTCCGT-AAGGTT3′, Bosco et al. 1992) and reverse primer FGPL 2054′ (5′CCGGTTTCCCCATTGG3′, Simonet et al. 1991). Primer 989ac is located in the 16S rRNA gene and is specific to Frankia of the alder-casuarina group (Bosco et al. 1992). This ensures elimination of unwanted bands that may arise because of presence of DNA of organelles in the extracted nodules.

Both isolated and amplified DNA samples were run in a 0.8% agarose gel at 70 V for 1.5 h to visualize the DNA. The gels were stained in ethidium bromide for about 15 min, scanned, and photographed using Bio-Rad GelDoc1000 and Multi Analyst software (version 1.1).

**Amplified rDNA restriction analysis**

The restriction enzyme Rsal (Roche Pharmaceuticals) was selected for restriction digestion of amplicons. Rsal was selected on the basis of a mock digestion using Mac Vector software. All the samples were digested overnight at 37°C. The restriction digestion mixture was prepared with 10 μl of the amplicon, 5 units of the restriction enzyme, 2 μl of appropriate buffer (buffer B) and ultrapure distilled water to make a final volume of 20 μl. The digested samples were run at 40 V for 7–8 h in 4% agarose gels containing ethidium bromide. The gels were scanned and photographed using Bio-Rad GelDoc1000 and Multi Analyst software. Band sizes of amplified and digested products were calculated using Multi Analyst (version 1.1) software.

ARDRA patterns were analysed by cluster analysis using the Sørenson distance measure and group average linkage method to generate cluster groups using a similarity value of 80%. ARDRA patterns were further analysed by the non-metric multidimensional scaling (NMS) ordination method and relationships with soil chemical properties explored using joint plots. The effects of site and altitude on the distribution of ARDRA patterns were evaluated using MRPP. These analyses were carried out using PC-Ord software (MJM Software Design, Gleneden Beach, OR).

**Nucleotide sequencing**

Samples were chosen for nucleotide sequence analysis of Frankia to represent different locations and different ARDRA patterns. For each pattern only one or two samples were sequenced. Selected samples were cloned using the pGEM TEasy (Promega, Madison, WI) and sequenced using BigDye Terminator v. 3.1 Cycle Sequencing kit and ABI Prism3730 Genetic Analyzer at the Centre for Genomics and Bioinformatics (Oregon State University). We could not sequence samples from low altitudes at sites 1 and 3 despite repeated attempts. The sequences obtained have been deposited in GenBank (accession numbers: DQ988956–DQ988982).

The nucleotide sequences obtained were individually fed into the GenBank for a Basic Local Alignment Search Tool search using the web site http://ncbi.nlm.nih.gov. The retrieved sequences along with the original sequences were then aligned using multiple sequence alignment program CLUSTAL W [version 1.83 (http://align.genome.jp/)]. The online tool PHYLIP 3.66 (http://evolution.genetics. washington.edu/phylip.html) was used to construct a phylogenetic tree using DNA parsimony analysis.

**Results**

**Soil chemical properties**

The collective analysis of soil chemical properties using MRPP did not show a significant effect of altitude, nor
Table 2. Chemical characteristics of the soils used in this study. Data are means ± standard errors (EC, electrical conductivity; C, organic carbon; N, total Kjeldhal nitrogen; P, available phosphorus; K, available potassium; Ca, exchangeable calcium; Mg, exchangeable magnesium; low altitude, 900–1000 m; middle altitude, 1400–1710 m; high altitude, above 2000 m).

<table>
<thead>
<tr>
<th>Site</th>
<th>Altitude</th>
<th>Soil property</th>
<th>pH</th>
<th>EC (meq/100 g)</th>
<th>C (mg/kg)</th>
<th>N (mg/kg)</th>
<th>P (%)</th>
<th>K (kg/ha)</th>
<th>Ca (meq/100 g)</th>
<th>Mg (meq/100 g)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Low</td>
<td>4.83 ± 0.12</td>
<td>0.06 ± 0.01</td>
<td>19.8 ± 2.0</td>
<td>0.50 ± 0.07</td>
<td>12.4 ± 0.4</td>
<td>277 ± 72</td>
<td>2.44 ± 0.47</td>
<td>1.62 ± 0.27</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Medium</td>
<td>5.16 ± 0.19</td>
<td>0.08 ± 0.02</td>
<td>22.7 ± 1.3</td>
<td>0.55 ± 0.10</td>
<td>9.7 ± 1.7</td>
<td>643 ± 128</td>
<td>2.70 ± 0.99</td>
<td>1.89 ± 0.92</td>
<td></td>
</tr>
<tr>
<td></td>
<td>High</td>
<td>5.04 ± 0.09</td>
<td>0.06 ± 0.01</td>
<td>15.5 ± 2.0</td>
<td>1.04 ± 0.30</td>
<td>19.6 ± 2.7</td>
<td>437 ± 122</td>
<td>1.63 ± 0.45</td>
<td>1.04 ± 0.19</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>Low</td>
<td>6.79 ± 0.32</td>
<td>0.49 ± 0.22</td>
<td>17.8 ± 1.3</td>
<td>0.34 ± 0.09</td>
<td>14.9 ± 2.0</td>
<td>193 ± 96</td>
<td>2.66 ± 0.59</td>
<td>1.05 ± 0.31</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Medium</td>
<td>4.87 ± 0.06</td>
<td>0.05 ± 0.01</td>
<td>16.0 ± 1.5</td>
<td>0.43 ± 0.15</td>
<td>17.6 ± 1.0</td>
<td>175 ± 44</td>
<td>1.60 ± 0.40</td>
<td>1.10 ± 0.27</td>
<td></td>
</tr>
<tr>
<td></td>
<td>High</td>
<td>5.29 ± 0.14</td>
<td>0.03 ± 0.00</td>
<td>11.1 ± 1.1</td>
<td>0.07 ± 0.01</td>
<td>13.3 ± 1.7</td>
<td>189 ± 39</td>
<td>1.38 ± 0.72</td>
<td>0.51 ± 0.12</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>Low</td>
<td>5.31 ± 0.11</td>
<td>0.08 ± 0.01</td>
<td>16.9 ± 1.3</td>
<td>0.61 ± 0.13</td>
<td>13.3 ± 1.8</td>
<td>371 ± 101</td>
<td>1.93 ± 0.43</td>
<td>1.17 ± 0.17</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Medium</td>
<td>5.56 ± 0.09</td>
<td>0.27 ± 0.01</td>
<td>22.9 ± 0.7</td>
<td>0.83 ± 0.10</td>
<td>13.8 ± 0.4</td>
<td>281 ± 48</td>
<td>7.04 ± 0.38</td>
<td>1.57 ± 0.11</td>
<td></td>
</tr>
<tr>
<td></td>
<td>High</td>
<td>5.68 ± 0.10</td>
<td>0.07 ± 0.01</td>
<td>17.1 ± 1.7</td>
<td>0.46 ± 0.19</td>
<td>11.8 ± 0.7</td>
<td>207 ± 58</td>
<td>1.69 ± 0.54</td>
<td>0.75 ± 0.11</td>
<td></td>
</tr>
</tbody>
</table>

Fig. 1. The top gel photo is an example of the amplified recombinant DNA restriction analysis (ARDRA) patterns found at one location (site 2, middle elevation). The numbers represent the 10 nodule samples from this site and the ARDRA patterns are designated with the prefix ‘P’; ‘U’ represents the uncut fragment and ‘M’ the marker lanes. The bottom figure is a diagrammatic representation of the 17 ARDRA patterns (designated with the prefix ‘P’) observed in this study.
were there differences among sites; however, some site-specific significant differences among the chemical properties of soils from different altitudes were observed ($P < 0.05$). The soil chemistry of the high altitude location at site 1 differed from the lower altitude locations, each altitude had a unique soil chemistry at site 2 and the soil chemistry of the middle altitude differed from those of the other altitudes at site 3. These altitudinal differences in soil chemistry were not associated with differences in individual soil chemical properties because of relatively high sample-to-sample variation (Table 2). Only magnesium concentration showed a significant difference among altitudes; being lowest at high altitudes and highest at middle altitudes ($P < 0.05$).

**ARDRA diversity and composition**

We observed three to six fragments in an individual ARDRA, with a total of 20 different-sized fragments among all samples. A total of 17 distinct ARDRA patterns were detected (Fig. 1, Table 3). Some ARDRA patterns were common among locations (e.g. ARDRA 13 and 14 were found at low elevations at all sites) but many were found at only a single location. At a given location, the number of ARDRA patterns varied from one to four. Across sites, there was a trend towards fewer ARDRA patterns as altitude increased. Site 3 had more different ARDRA patterns than either other site.

Cluster analysis allowed us to group closely related ARDRA patterns, resulting in seven cluster groups (Table 3). Many of the trends among sites and altitudes observed with ARDRA patterns were preserved when using cluster groups. There were some interesting exceptions, however. For example, the four ARDRA patterns observed at the medium altitude of site 3 were part of a single cluster group, whereas the three ARDRA patterns observed at the high altitude of site 1 remained as three separate cluster groups. Overall, cluster groups corresponded well to altitudinal differences: groups A, B, C and D were found only at high altitude, groups B and E were found only at middle altitudes and groups F and G were found only at low altitude.

The patterns that can be observed in Table 3 agree with those found by NMS ordination (Fig. 2). More than 75% of the variation in the ARDRA data were represented on the first two axes and clearly showed differences in the *Frankia* found at the different locations. At each site, the composition of the *Frankia* community varied by altitude and MRPP analysis showed that these differences were significant in each case ($P < 0.05$). MRPP analysis also showed that these altitudinal differences in *Frankia* communities were significant across all sites collectively, similar to the association found between altitudes and cluster groups. In general, MRPP analysis showed that the *Frankia* community at site 3 was significantly different than that of the other two sites, which were similar. This site pattern was not as clear based on the use of cluster groups.

The use of joint plot analysis did not show any correlations between soil chemical properties and *Frankia* community composition ($P > 0.05$). This suggested that variations in soil chemical properties among sites and altitudes did not consistently correspond to shifts in

| Table 3. Distribution of *Frankia* among sites based on amplified recombinant DNA restriction analysis (ARDRA) patterns. ARDRA cluster groups were defined at a level of 80% similarity. |
|-------------|--------|----------|----------|
| Site | Altitude | ARDRA patterns | ARDRA groups |
| 1 | Low | 13, 14 | F |
| | Medium | 6 | B |
| | High | 1, 2, 3 | A, B, C |
| 2 | Low | 13, 14, 15 | F |
| | Medium | 6, 7, 8 | B |
| | High | 4 | D |
| 3 | Low | 13, 14, 16, 17 | F, G |
| | Medium | 9, 10, 11, 12 | E |
| | High | 4, 5 | D |

![Fig. 2. Non-metric multidimensional scaling ordination of *Frankia* based on amplified recombinant DNA restriction analysis patterns. Data are means for each site and altitude; error bars represent 1 standard error.](image-url)
ARDRA patterns or other measures of Frankia community composition.

Sequence and phylogenetic analysis

Sequence analysis of the 16S–23S rRNA ITS region revealed variations among samples of different sites. Generally we found that the distal part of 16S rRNA gene was highly conserved as no differences were observed in the initial part of the amplicon. At high altitude, site 1, there were several single nucleotide polymorphisms (SNPs) especially in the middle part of the amplicon. Two specific regions with high density of SNPs were between 814–822 and 863–899 bp. For high altitude, site 2, two specific regions with high density of SNPs were between 800–822 and 863–899 bp. At high altitude, site 3, several SNPs were detected even in the distal part of the 16S rRNA gene, which was unexpected as this portion of the gene is generally considered highly conserved. The distribution of SNPs along the rest of the regions was more or less similar to the other cases.

Cluster analysis using nucleotide sequence data generally tended to confirm the relationship among site and altitude found with ARDRA patterns. For example, Sample 13 of high altitude, site 2, showed a close relationship with Sample 65 from high altitude, site 3 (Fig. 3). Samples 13 and 65 are in ARDRA group D and both are ARDRA pattern 4. These two samples were in turn related to Sample 1 (ARDRA group A, ARDRA pattern 1) from high altitude, site 1. These three samples formed a separate

![Fig. 3. Phylogenetic tree based on internally transcribed spacer sequence data. Sequences obtained in this study are designated with an L, M or H (for low, middle or high altitude), followed by 1, 2 or 3 (for the three sites), followed by S and the sample number.](image-url)
found altitude and/or geographic location to be related to the composition of the Frankia community (Dai et al. 2004, Igual et al. 2006, Jeong and Myrold 1999, Oakley et al. 2004, Ritchie and Myrold 1999). Unfortunately, it is difficult to infer the causative factors behind the influence of altitude and geographic location. In this particular study, the lack of relationship between Frankia genotypes and soil chemical properties—even those such as pH, organic matter and calcium that some prior studies have found correlated to Frankia populations or genotypes (Burleigh and Dawson 1994, Crannell et al. 1994, Smolander and Sundman 1987)—implicates microclimatic factors, such as temperature and moisture, as major determinants of the distribution of Frankia genotypes.

Acknowledgements – This work was partly supported by Department of Science and Technology, Government of India grant [DST/INT/US(NSF-RPO)4/2000] to Prof. A. K. Misra. Ms Anindita Khan acknowledges financial support obtained from the University Grants Commission-Departmental Special Assistance programme of the Department of Botany, North Eastern Hill University, Shillong, India. We appreciate the assistance of Drs R. S. Tripathi, R. R. Yarwood and David Biate and the reviewer comments. This material is based on work supported by the National Science Foundation under grant no. 0096793.

References


Dawson JO, Kowalski DG, Dart PJ (1989) Variation with soil depth, topographic position and host species in the capacity of soils from an Australian locale to nodulate Casuarina and Allocasuarina seedlings. Plant Soil 118: 1–11


Smolander A (1990) Frankia populations in soils under different tree species—with special emphasis on soils under Betula pendula. Plant Soil 121: 1–10


Varghese R, Chauhan VS, Misra AK (2003b) Hypervariable spacer regions are good sites for developing specific PCR-RFLP markers and PCR primers for screening actinorhizal symbionts. J Biosci 48: 437–442
