APPLIED AND ENVIRONMENTAL MICROBIOLOGY, Nov. 2004, p. 6444–6452 Vol. 70, No. 11
0099-2240/04/$08.00+0 DOI: 10.1128/AEM.70.11.6444–6452.2004

Diversity and Distribution of *Frankia* Strains Symbiotic with *Ceanothus* in California

Brian Oakley,[1]* Malcolm North,[2] Jerry F. Franklin,[1] Brian P. Hedlund,[3]† and James T. Staley[3]

College of Forest Resources[1] and Department of Microbiology,[3] University of Washington, Seattle, Washington, and Sierra Nevada Research Center, Department of Environmental Horticulture, University of California, Davis, California[2]

Received 29 January 2004/Accepted 9 June 2004

***Frankia* strains symbiotic with *Ceanothus* present an interesting opportunity to study the patterns and causes of *Frankia* diversity and distribution within a particular host infectivity group. We intensively sampled *Frankia* from nodules on *Ceanothus* plants along an elevational gradient in the southern Sierra Nevada of California, and we also collected nodules from a wider host taxonomic and geographic range throughout California. The two sampling scales comprised 36 samples from eight species of *Ceanothus* representing six of the seven major biogeographic regions in and around California. The primary objective of this study was to use a quantitative model to test the relative importance of geographic separation, host specificity, and environment in influencing the identity of *Ceanothus Frankia* symbionts as determined by ribosomal DNA sequence data. At both sampling scales, *Frankia* strains symbiotic with *Ceanothus* exhibited a high degree of genetic similarity. *Frankia* strains symbiotic with *Chamaebatia* (*Rosaceae*) were within the same clade as several *Ceanothus* symbionts. Results from a classification and regression tree model used to quantitatively explain *Frankia* phylogenetic groupings demonstrated that the only significant variable in distinguishing between phylogenetic groups at the more local sampling scale was host species. At the regional scale, *Frankia* phylogenetic groupings were explained by host species and the biogeographic province of sample collection. We did not find any significant correspondence between *Frankia* and *Ceanothus* phylogenies indicative of coevolution, but we concluded that the identity of *Frankia* strains inhabiting *Ceanothus* nodules may involve interactions between host species specificity and geographic isolation.**

Frankia strains are gram-positive, filamentous, N-fixing bacteria that form symbiotic relationships with 24 genera of host plants representing 8 families (2, 3). These actinorhizal plants are found throughout the world and serve an important function in many ecosystems as the major source of N, commonly a limiting nutrient (33, 34). Three main groups of *Frankia* have been classified based on host plant family: an *Alnus* (*Betulaceae*)/*Myricaceae*/*Casuarinaceae* group, an *Eleagnaceae*/*Rhamnaceae* group, and a third group which includes *Frankia* symbiotic with actinorhizal plants in the families *Coriariaceae*, *Datiscaceae*, and *Rosaceae* and the genus *Ceanothus* (*Rhamnaceae*) (9, 27, 35). These groupings only loosely correspond to host plant phylogenetic relationships, but they are generally consistent with the *Frankia* infection mode and phenotype (35). *Frankia* strains within a single host group also exhibit some degree of genetic heterogeneity, yet the patterns and causes of this heterogeneity are less well understood.

For several reasons, *Frankia* strains symbiotic with *Ceanothus* present an interesting opportunity to study the patterns and causes of *Frankia* diversity within a particular host-group. First, *Frankia* strains from the seven actinorhizal genera in the family *Rhamnaceae* all group with *Frankia* strains symbiotic with *Eleagnaceae*, with the sole exception of *Frankia* strains symbiotic with *Ceanothus*, which belong to the *Coriariaceae*/*Datiscaceae*/*Rosaceae* clade (27). Second, *Ceanothus* offers a useful model with which to discern the relative influences of plant host and environment on *Frankia* identity, because the genus contains many species that occur sympatrically with other actinorhizal genera across a wide of range of environments. Finally, although *Ceanothus* is an ecologically important component of many western coniferous forests and range lands (10a), the ecology of the *Ceanothus-Frankia* symbiosis remains poorly studied.

Ceanothus contains two distinct subgenera, *Cerastes* (22 species) and *Ceanothus* (33 species), and is found throughout North America, but the center of its distribution is in California, where 40 species occur (10a, 22). Some researchers have found little diversity among *Frankia* strains symbiotic with *Ceanothus* and have suggested that the host infectivity group as a whole is depauperate (5), but the patterns and causes of this remain unclear. In a systematic sampling of nodules collected from nine *Ceanothus* species in Oregon, differences among *Frankia* strains were attributed to geography, particularly elevation, rather than host plant specificity (32). Similarly, Jeong and Myrold (19) found that *Frankia* strains nodulating *Ceanothus* were very similar to one another but concluded that the differences that did exist were more closely related to geographic location than to host species. However, in a southern California chaparral ecosystem, the diversity of *Ceanothus* symbionts was much higher and appeared to be related to host species; 12 distinct DNA fingerprint patterns based on repetitive sequence (rep)-PCR methods were found from nodules collected from six *Ceanothus* species (24). Because California is the center of *Ceanothus* distribution, it may also contain a

* Corresponding author. Present address: Department of Microbiology, Box 357242, University of Washington, Seattle, WA 98195. Phone: (206) 543-6646. Fax: (206) 543-8297. E-mail: boakley@u.washington.edu.

† Present address: Department of Biological Sciences, University of Nevada—Las Vegas, Las Vegas, NV 89154.

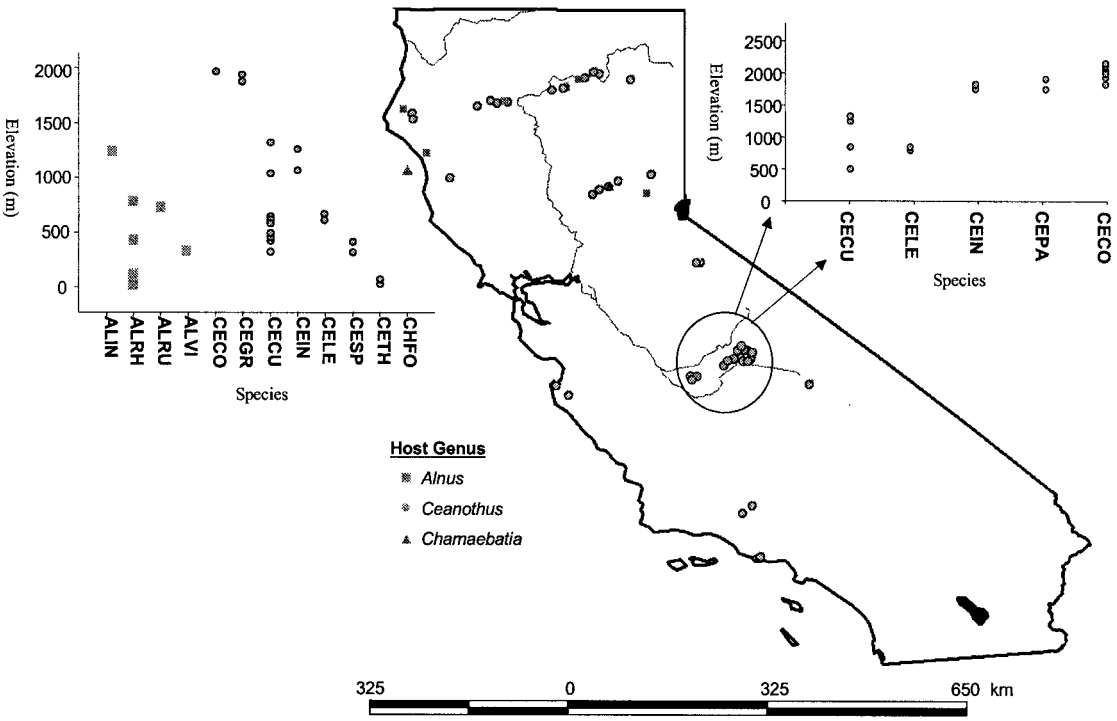

FIG. 1. Geographic locations, elevations, and host plants from which nodules were collected. The graph on the right summarizes host species and elevations from the intensive regional sampling of *Ceanothus* nodules, while the graph on the left summarizes the sampling of three host genera across California. Species names are listed in Table 2. Samples represent five of the six major geographic regions of the California Floristic Province and one region of the Great Basin Floristic Province (per reference 14).

higher diversity of *Ceanothus*-infective *Frankia* than other study regions.

In this study, we used two different sampling schemes to better understand factors associated with patterns of *Frankia* diversity within the *Ceanothus* host specificity group. First, we intensively sampled *Frankia* strains forming nodules on *Ceanothus* plants within a relatively small geographic area along an elevational gradient in the southern Sierra Nevada of California, and second, we less intensively sampled nodules from a wider geographic and taxonomic range. We collected nodules from *Ceanothus*, *Alnus*, and *Chamaebatia* (*Rosaceae*) at the scale of major biogeographic regions throughout California. Together, the two sampling scales represent 36 samples from eight species of *Ceanothus* from six of the seven major biogeographic regions of California. The primary objective of this study was to test the relative importance of geographic separation, host specificity, and environment in determining the identity of *Ceanothus Frankia* symbionts.

MATERIALS AND METHODS

Study area and field sampling methods. We first conducted intensive sampling of multiple *Ceanothus* species at a spatial scale of tens of kilometers on the western slope of the southern Sierra Nevada. We focused on this area because of previous research indicating a high diversity of *Ceanothus*-infective *Frankia* occurring at a local scale in the region (24). Nodules from five species of *Ceanothus* were collected along a 1,500-m elevational transect (Fig. 1). Following this initial sampling, we conducted less-intensive sampling from *Ceanothus* and two other host genera (*Alnus* and *Chamaebatia*) from six distinct biogeographic regions representative of an area of several thousand square kilometers within and around California (14) (Fig. 1). When possible, we collected nodules from host plants that were growing within several meters of each other. This paired sampling was replicated six times (Table 1). Detailed information on host plants is given in Table 2.

When nodules were found, we recorded their depth from the soil surface and took soil samples within the immediate vicinity of the nodules. A handheld global positioning system (GPS) unit was used to record the location of each sample with an accuracy of ±5 m. Nodules and soil were kept on ice during transport back to the laboratory, where soil samples were analyzed for moisture content and pH.

Greenhouse experiment. To further examine the relative influence of host species and environment on *Frankia* identity, we also conducted an experiment in which *Ceanothus cordulatus* seedlings were grown in native soil in a greenhouse environment in Seattle. *C. cordulatus* seeds collected from the Teakettle Experimental Forest in the southern Sierra Nevada of California (see reference 28 for a detailed description) were soaked in boiling water, allowed to cool for 30 min, and then given a 3-month cold stratification treatment at 4°C. Seedlings were then transferred into soil collected from the field from depths of 10 to 25 cm in the immediate root environments of nine different *Ceanothus* plants. For each soil sample, three replicate seedlings were inoculated, and for each replicate, three dilutions were performed (ratios of field-collected soil to sterilized potting soil, 1:1, 1:10, and 1:100), for a total of 162 seedlings. Seedlings were maintained with N-free watering for 6 months, after which nodules were harvested and *Frankia* DNA was extracted, amplified, and sequenced.

Laboratory methods. (i) DNA extractions. DNA was extracted directly from root nodules by using a protocol modified from the work of Baker and Mullin (4) and Ritchie and Myrold (32). Immediately after return from the field, nodules were rinsed with sterile distilled water (dH_2O), placed in sterile sample cups, and washed with ca. 100 ml of sdH_2O and 1% Tween 20 on a shaker for 20 to 30 min. Nodule clusters were then rinsed with sdH_2O followed by 70% ethanol, air dried, and stored at −20°C until further use.

Single nodule lobes were peeled by using a dissecting microscope, surface sterilized with 95% ethanol, and frozen in liquid N immediately prior to DNA extraction. Nodule lobes were crushed to a fine powder by using a sterile plastic pestle and were then incubated at 65°C for 30 min in 600 μl of CTAB buffer (2% cetyltrimethylammonium bromide, 100 mM Tris [pH 8.0], 20 mM EDTA, 1.4 M NaCl). DNA was extracted twice with equal volumes of 24:1 chloroform–isoamyl alcohol and precipitated by addition of 1 volume of ice-cold isopropanol, a 30-min incubation at −70°C, and a 30-min 14,000 × *g* centrifugation at 4°C.

TABLE 1. Number of host plants, by geographic location and host genus, from which nodules were collected[a]

Geographic region	Subregion	No. of plants			
		Alnus	*Ceanothus*	*Chamaebatia*	Total
California Floristic Province					
Cascade range (CaR)	Cascade range high (CaRH)	1	3	1	5
Central western (CW)	South Coast range (SCoR)		2		2
Northwestern (NW)	Klamath ranges (KR)	1	4		5
	North Coast ranges (NCoR)	2	2		4
Sierra Nevada (SN)	Sierra Nevada foothills (SNF)	1	5		5
	High Sierra Nevada (SNH)	1	15		17
Southwestern (SW)	Transverse ranges (TR)		3		3
Great Basin Floristic Province, Modoc plateau (MP)	Modoc plateau (MP)	1	2		3
Total		7	36	1	44

[a] As many nodules as possible were collected from each plant in the field, and at least one *Frankia* sequence was obtained from each plant.

DNA was precipitated a second time by resuspension in 50 μl of 10:1 TE (10 mM Tris [pH 8.0], 1 mM EDTA) followed by addition of 1/4 volume of 10 M ammonium acetate, 1 volume of ice-cold isopropanol, a 30-min incubation at −70°C, and a 30-min 14,000 × *g* centrifugation at 4°C. DNA pellets were washed with 70% ethanol, dried, and resuspended in 50 μl of 10:0.1 TE. To remove compounds inhibitory to the PCR, DNA was further purified by addition of 1 volume of PEG-NaCl (20% polyethylene glycol [molecular weight, 8000], 2.5 M NaCl), a 15-min incubation at 37°C, a 15-min 14,000 × *g* centrifugation, and two washes and 5-min centrifugations with 80% ethanol. Purified DNA pellets were dried, resuspended in 10:0.1 TE, and stored at −20°C.

(ii) PCRs. We amplified a 2,098-bp region of the *Frankia* genome that includes the 3′ end of the 16S rRNA gene, the intergenic spacer (IGS) region, and a portion of the 23S rRNA gene by using primers 1649F (5′-GATTGGGACGA AGTCGT-3′) and 2309R (5′-ATCGCATGCCTACTACC-3′) (32). Although the IGS region by itself has been used previously to distinguish among *Frankia* strains, we also included the 5′ end of the 23S rRNA gene, which contains a variable region that has been used to increase strain-level discrimination (15). Amplification reactions were performed with an optimized buffer (10 mM Tris-HCl [pH 9.2], 1.5 mM $MgCl_2$, 7.5 mM KCl), 0.2 mM each deoxynucleoside triphosphate, 0.2 μM each primer, 2 U of *Taq* polymerase, 1 μl of template DNA, and sterile dH_2O for a reaction volume of 25 μl. The thermal cycling program consisted of a 2-min denaturation at 94°C; 35 cycles of denaturation at 94°C for 45 s, annealing at 53°C for 45 s, and extension at 72°C for 90 s; and a final 72°C extension for 5 min. PCR products were visualized on a 1% agarose gel.

When amplification was insufficient for sequencing, PCR products were cloned (Invitrogen Topo TA cloning kit), and PCR was performed using M13 primers with plasmid DNA extracted from clones screened for ligation and transformation.

(iii) Sequencing. Following successful amplification, excess oligonucleotides were cleaned from PCR products, and sequencing reactions were performed in a 10-μl volume by using the BigDye Ready Reaction Mix (Perkin-Elmer), 200 to 500 ng of DNA, and 3.2 pmol of primer. Primers used for PCR and two internal sequencing primers designed for this study (5′-AATAATTCCCGTGTGCCAC C-3′ and 5′-TTACGCACTCTTTCAAGGGTGG-3′) were used for sequencing. Unincorporated dyes were cleaned from the sequencing products by using Sephadex G-50 columns, and samples weresubmitted to the University of Washington Biochemistry DNA Sequencing Facility.

Statistical methods. (i) Phylogenetic analysis of DNA sequences. We performed three separate phylogenetic analyses. We first compared sequences of the full ca. 2-kb region of DNA amplified by PCR, which included the IGS and a variable portion of the 23S rRNA gene (15), by using two different data sets. First, we used sequences only from the intensive regional sampling of *Ceanothus* symbionts, and second, we analyzed samples collected from all of California, which included seven *Alnus* symbionts and one symbiont from *Chamaebatia foliolosa*. Finally, to compare our samples to previously published sequences of *Frankia* from *Ceanothus* nodules (19, 32), we restricted the analysis to the 482-bp IGS region.

For each analysis, distance matrices were calculated for 1,000 bootstrapped replicates of sequence alignments by using the Kimura 2-parameter model for transversions and transitions with the TreeCon program (36). A consensus tree was constructed from 1,000 bootstrapped neighbor-joining trees and viewed in TreeView (30).

(ii) Prediction of climate parameters and statistical analyses. Because the precise location of each sample was known from GPS data collected in the field, we were able to predict several climate parameters and then use these predicted values as independent variables in a statistical model to determine their relative importance in explaining variability among *Frankia* strains. We used data from a climate model called PRISM (parameter-elevation regressions on independent slopes model) (12, 13) to predict January, August, and annual mean temperatures, as well as August and annual precipitation. The PRISM climate model combines digital elevation models with 30 years of data from ca. 3,000 meteorological stations in California to interpolate a response surface for temperature and precipitation on the basis of latitude, elevation, slope, aspect, and distance from the coastline (12, 13). We chose January and August as months in which temperature and precipitation reach extremes and therefore may potentially be biotically limiting factors.

Each of the predicted climate parameters was then used as an independent variable in a CART (classification and regression tree) model with the phylogenetic grouping of each *Frankia* strain as the dependent variable. CART is a nonparametric classification method analogous to multiple regression in which data are classified by successively subdividing into increasingly homogeneous groupings. At each partitioning step, the new groups of data are examined separately along the multivariate axes of the predictor variable and further partitioned (6). This process, called binary recursive partitioning, provides a quantitative comparison of the relative importance of each independent variable for grouping the dependent variable (6, 37). Regression trees are generally more robust than regression models in dealing with categorical data. We used predicted climate parameters, geographic region, subregion, soil pH, latitude, elevation, and previously established phylogenetic groupings of *Ceanothus* (17) as independent variables in the CART analyses.

RESULTS

Regional differences among *Ceanothus* symbionts. Phylogenetic analyses based on sequences of the IGS and 23S rRNA gene indicated that *Frankia* strains symbiotic with *Ceanothus* were very homogeneous but still formed two distinct clades (Fig. 2). The greatest pairwise sequence difference between *Ceanothus* symbionts we collected was eight transitions or transversions and two insertions or deletions. About one-third of the heterogeneity was concentrated in the IGS region, with the other two-thirds in the ca. 1.5-kb 23S region that we sequenced.

Despite this homogeneity, the bootstrap analysis did support the division of *Ceanothus* symbionts collected in the southern Sierra Nevada into two statistically distinct clades (Fig. 2). Alternate tree construction methods yielded consistent topologies. *Frankia* strains belonging to these two groups were different from one another in several respects. There was no overlap among the species assemblages of the two groups:

TABLE 2. Attributes of host plants for *Frankia* strains analyzed in this study

Host genus and species[a]	GenBank accession no.	Elevation (m)	Soil pH	Geographic region[b]	Subregion[b]
Alnus glutinosa (strain AcN14a)	M88466[c]	NA[d]	NA	NA	NA
Alnus incana[1]	AY627709	1,307	6.45	CaR	CaRH
Alnus rhombifolia[2]	AY627710	70	5.97	NW	NCoR
Alnus rhombifolia	AY627711	164	6.53	NW	NCoR
Alnus rhombifolia	AY627712	472	6.43	SN	SNF
Alnus rhombifolia	AY627713	837	7.31	MP	MP
Alnus rubra	AY627714	777	6.27	SN	SNH
Alnus viridis	AY627715	390	6.91	NW	KR
Ceanothus cordulatus	AF050761[e]	808	NA	OR	NA
Ceanothus cordulatus	AY627716	1,885	NA	SN	SNH
Ceanothus cordulatus	AY627717	1,951	NA	SN	SNH
*Ceanothus cordulatus**	AY627718	2,012	NA	SN	SNH
*Ceanothus cordulatus**	AY627719	2,012	NA	SN	SNH
*Ceanothus cordulatus**	AY627720	2,012	NA	SN	SNH
Ceanothus cordulatus	AY627721	2,022	6.3	SN	SNH
Ceanothus cordulatus[3]	AY627722	2,022	NA	SN	SNH
Ceanothus cordulatus[3]	AY627723	2,022	NA	SN	SNH
Ceanothus cordulatus	AY627724	2,039	NA	SN	SNH
Ceanothus cordulatus	AY627725	2,045	NA	SN	SNH
Ceanothus cuneatus	AF050762[e]	91	NA	OR	NA
Ceanothus cuneatus	AY627726	365	7.15	NW	KR
Ceanothus cuneatus	AY627727	471	6.5	CW	SCoR
Ceanothus cuneatus	AY627728	481	6.25	SN	SNF
Ceanothus cuneatus	AY627729	488	6.22	CW	SCoR
Ceanothus cuneatus	AY627730	637	6.3	CaR	CaRH
Ceanothus cuneatus	AY627731	655	6.63	NW	KR
Ceanothus cuneatus	AY627732	677	6.23	NW	KR
Ceanothus cuneatus	AY627733	875	NA	SN	SNF
Ceanothus cuneatus	AY627734	1,080	6.91	MP	MP
Ceanothus cuneatus	AY627735	1,250	NA	SN	SNH
Ceanothus cuneatus	AY627736	1,303	NA	SN	SNH
Ceanothus cuneatus	AY627737	1,365	6.01	MP	MP
Ceanothus cuneatus	AY627738	1,645	NA	SN	SNH
Ceanothus cuneatus	AY627739	1,929	6.29	SN	SNH
Ceanothus greggii	AY627740	1,999	6.65	SW	TR
Ceanothus integerrimus	AF050763[e]	366	NA	OR	NA
Ceanothus integerrimus[4]	AY627741	1,113	6.14	CaR	CaRH
Ceanothus integerrimus[1]	AY627742	1,307	6.81	CaR	CaRH
Ceanothus integerrimus[6]	AY627743	1,768	NA	SN	SNH
Ceanothus integerrimus	AY627744	1,811	NA	SN	SNH
Ceanothus leucodermis	AY627745	666	6.71	NW	KR
Ceanothus leucodermis	AY627746	710	5.84	SN	SNF
Ceanothus leucodermis[5]	AY627747	914	NA	SN	SNF
Ceanothus leucodermis[5]	AY627748	914	NA	SN	SNF
Ceanothus parvifolius[6]	AY627749	1,768	NA	SN	SNH
Ceanothus parvifolius	AY627750	1,929	NA	OR	SNH
Ceanothus prostratus	AF050764[e]	1,372	NA	OR	NA
Ceanothus pumilus	AF050765[e]	549	NA	OR	NA
Ceanothus spinosus	AY627751	354	7.91	SW	TR
Ceanothus spinosus	AY627752	468	6.81	SW	TR
Ceanothus thyrsifolius[2]	AY627753	88	6.43	NW	NCoR
Ceanothus thyrsifolius	AY627754	102	6.65	NW	NCoR
Ceanothus velutinus	AF050766[e]	366	NA	OR	NA
Ceanothus velutinus	AF050767[e]	549	NA	OR	NA
Ceanothus velutinus	AF050768[e]	1,219	NA	OR	NA
Ceanothus velutinus	AF036901[f]	1,250	NA	OR	NA
Chamaebatia foliolosa[4]	AY627755	1,113	6.55	CaR	CaRH

[a] Asterisks indicate nodules formed in a greenhouse in Seattle on seedlings grown in field-collected soil, as described in the text. [1,2,3,4,5,6], sympatric paired samples.
[b] Abbreviations for geographic regions and subregions are given according to Table 1. OR, Oregon.
[c] From reference 21.
[d] NA, not available.
[e] From reference 32.
[f] From reference 19.

Frankia strains from *C. cordulatus, Ceanothus integerrimus*, and *Ceanothus parvifolius* were most similar to one another and constituted phylogenetic group A, while strains from *Ceanothus cuneatus* and *Ceanothus leucodermis* constituted group B (Fig. 2). The two phylogenetic groups differed in other important respects as well. The mean elevation of samples in group A (range, 1,741 to 2,015 m) was significantly different ($P < 0.05$) from the mean elevation for group B (range, 862 to 1,620 m) as determined by simple *t* tests.

Because host plant species distribution is also a function of

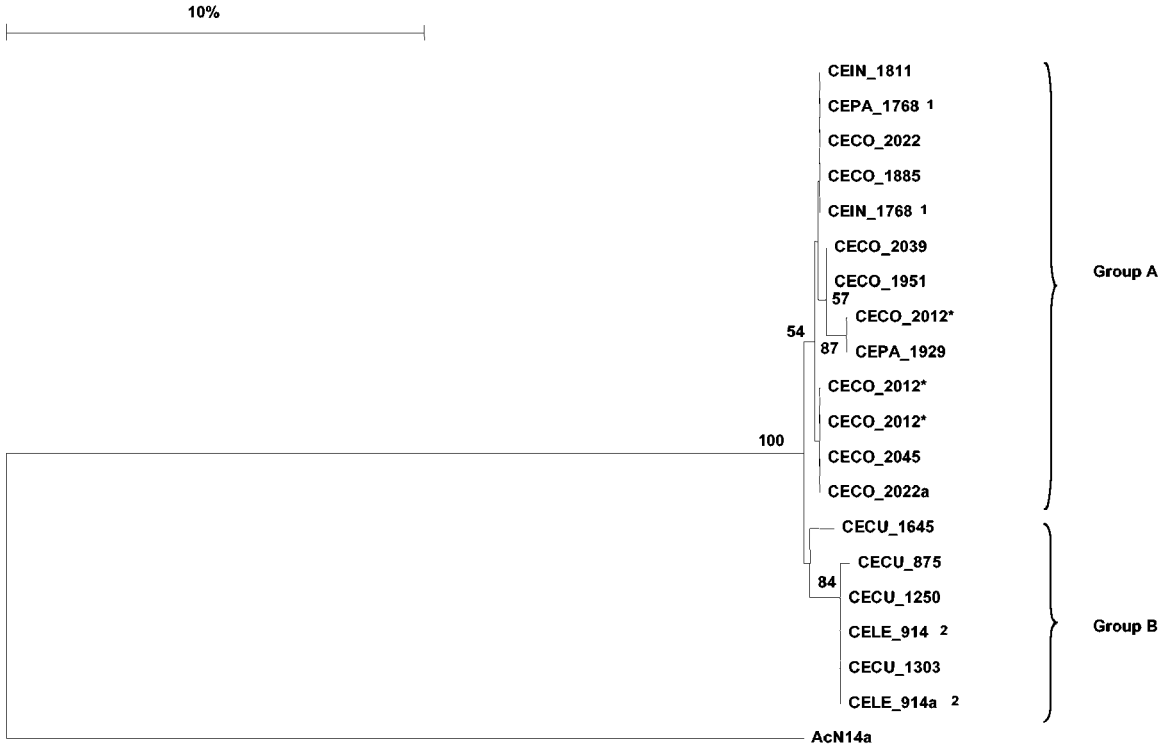

FIG. 2. Neighbor-joining tree based on IGS and 23S sequence data from *Frankia* strains symbiotic with *Ceanothus* collected from the intensive regional sampling in the southern Sierra Nevada. The tree is rooted with a sequence from the *Alnus* symbiont strain AcN14a (21). Bootstrap values represent 1,000 replicates; values are shown as percentages only when they exceed 50%. Branch lengths correspond to sequence differences as indicated by the scale bar. Designations at terminal nodes consist of a four-letter host species code (CEIN, *C. integerrimus*; CEPA, *C. parvifolius*; CECO, *C. cordulatus*; CECU, *C. cuneatus*; CELE, *C. leucodermis*) followed by the elevation (in meters) where the plant was growing. Details of sample origins are presented in Table 2. *, nodules formed in a greenhouse in Seattle on *C. cordulatus* seedlings grown in field-collected soil as described in the text. [1,2], sympatric paired samples.

elevation, we used a quantitative model to help discern the relative importance of these and other variables that may influence *Frankia* identity and distribution. When the CART model was used to explain the *Frankia* phylogenetic groupings using host species, elevation, latitude, and the five climate variables predicted from the PRISM model (January, August, and annual mean temperatures; August and annual precipitation) as independent variables, the only significant variable in distinguishing between the two phylogenetic groups was host species (data not shown).

Greenhouse results. Sequences from nodules formed on *C. cordulatus* seedlings grown in a greenhouse in Seattle were in the same clade as *Frankia* strains collected from the same host species in various locations in California (phylogenetic group A [Fig. 2]).

Comparison of *Alnus*, *Ceanothus*, and *Chamaebatia* symbionts collected throughout California. When the taxonomic and geographic extents of sampling were expanded to include eight *Ceanothus* taxa from six of the seven major biogeographic regions in California, we still found few sequence differences among *Frankia* strains symbiotic with *Ceanothus*, but we were able to distinguish at least three phylogenetic groups (Fig. 3).

For nodules collected from *Alnus* and *Ceanothus*, the identity of the *Frankia* strain was clearly dependent on the genus of the host plant, even when the plants occupied virtually the same location. Nodules collected from *Ceanothus* and *Alnus* plants growing sympatrically (within 5 m of each other) showed much larger differences than did any nodules from *Ceanothus*, regardless of location (Fig. 3). However, *Frankia* phylogenetic groupings were not always distinguished by host genus. *Frankia* strains collected from *Chamaebatia* formed a single clade with several *Ceanothus* symbionts, including one from a sympatric *Ceanothus* plant (Fig. 3).

The variability among *Frankia* strains symbiotic with *Alnus* (group 4) was much greater than the variability within or between groups 1, 2, and 3 (*Ceanothus* and *Chamaebatia* symbionts) (Fig. 3). Pairwise sequence differences among *Frankia* strains within group 4 ranged from 4 to 144 transitions or transversions, while among *Frankia* strains in groups 1, 2, and 3, the range was 0 to 7 for the *Ceanothus* symbionts. The *Chamaebatia foliolosa* symbiont had 7 to 14 transitions or transversions compared to the *Ceanothus* symbionts.

Correlates of intragroup strain differences. Although several variables were significantly different for the three phylogenetic groups of *Ceanothus* and *Chamaebatia* symbionts as determined by simple analysis of variance and posthoc *t* tests, in the CART analysis, only host species and geographic region were significant. Phylogenetic group 3 was distinguished from groups 1 and 2 on the basis of host species, and groups 1 and 2 were distinguished by geographic region (Fig. 4). Phylogenetic groups with a single member were not allowed for the CART analysis, but within group 3, samples CEGR_1999 and CHFO_1113 were distinct from other members of this group.

Despite this evidence of host specificity, there was little

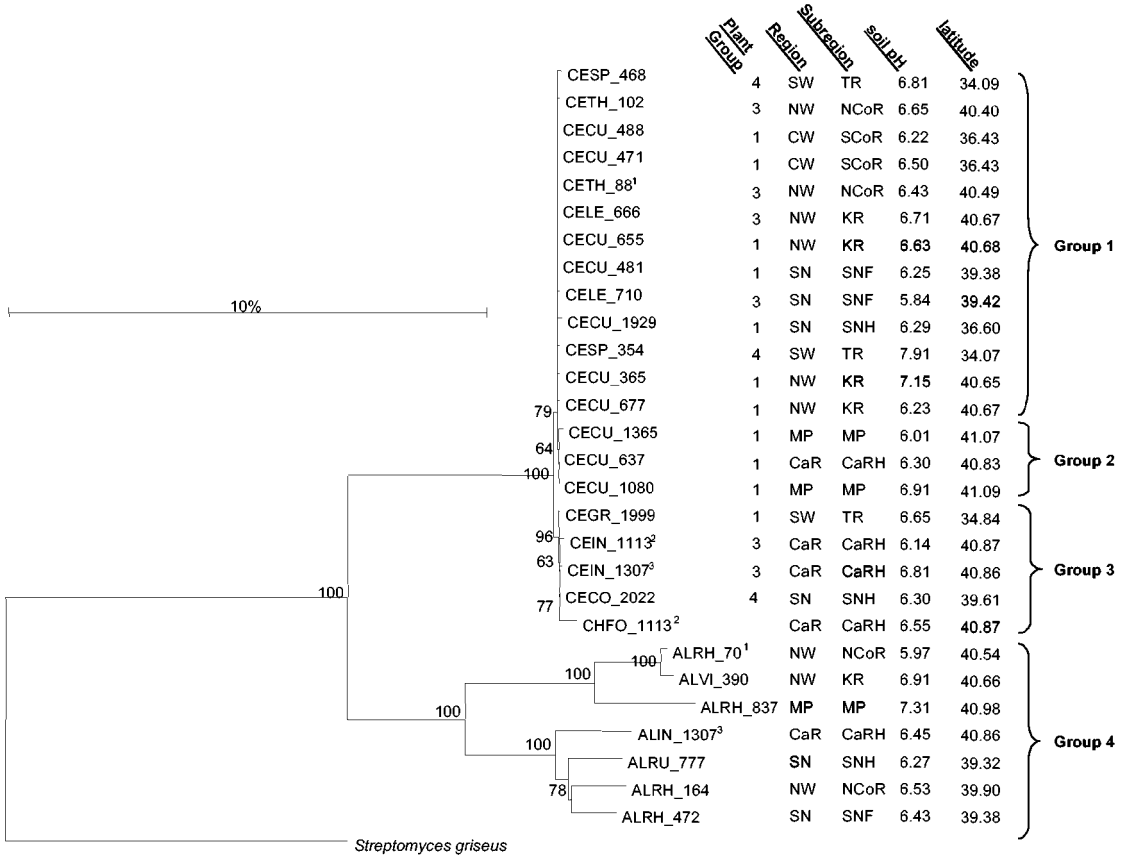

FIG. 3. Neighbor-joining tree based on IGS and 23S sequence data from *Frankia* strains symbiotic with *Ceanothus*, *Alnus*, and *Chamaebatia* collected from six major biogeographic regions throughout California. Sequence from *Streptomyces griseus* was used to root the tree. Bootstrap values represent 1,000 replicates; only values above 50% are shown. Branch lengths correspond to sequence differences as indicated by the scale bar. Designations at terminal nodes consist of a four-letter host species code followed by the elevation (in meters) where the plant was growing. Plant groups are based on a molecular phylogeny of *Ceanothus* (17), with groupings inferred on the basis of morphological characters for *C. greggi* (CEGR), *C. leucodermis*, and *C. spinosus* (CESP). *Frankia* phylogenetic groups were assigned for the purposes of statistical analysis, and so groups with a membership of <3 were not created even when supported by bootstrapping. [1,2,3], sympatric paired samples. CETH, *C. thyrsifolius*; CHFO, *Chamaebatia foliolosa*; ALRH, *Alnus rhombifolia*; ALVI, *Alnus viridis*; ALIN, *Alnus incana*; ALRU, *A. rubra*. Other host species codes are as explained in the legend to Fig. 2.

correspondence between *Frankia* and *Ceanothus* phylogenies. *Ceanothus* phylogenies based on molecular data (17) and morphological data (22) were both used as independent variables in separate CART analyses, but neither was significant in explaining *Frankia* phylogenetic groupings.

Comparison to previously published sequences. When we compared strains representative of the breadth of host species and geographic regions sequenced in this study to previously published sequences, we found that *Ceanothus* strains from Oregon (19, 32) were most similar to each other but formed a single clade that was generally consistent with the host species groupings identified above. Based on a comparison of the 482-bp IGS region, *Ceanothus* symbionts from Oregon were more similar to one another than to any of our samples but were contained within phylogenetic group I of *Frankia* strains from *Ceanothus velutinus*, *C. cuneatus*, *Ceanothus pumilus*, *C. integerrimus*, *C. cordulatus*, *Ceanothus prostratus*, *C. parvifolius*, *Ceanothus greggii*, and *Chamaebatia foliolosa* growing in both California and Oregon (Fig. 5). Phylogenetic group II was composed of *Frankia* symbionts from *C. cuneatus*, *C. leucodermis*, *Ceanothus spinosus*, and *Ceanothus thyrsifolius* (Fig. 5). *Alnus* strain AcN14a (21), originally isolated from *Alnus gluti-*

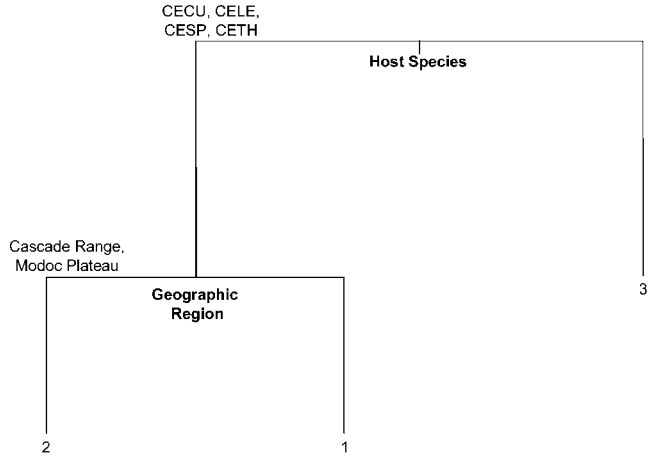

FIG. 4. Classification tree of relative importance of variables used to explain the phylogenetic groupings of *Frankia* strains associated with *Ceanothus* and *Chamaebatia* shown in Fig. 3. Values at terminal nodes are the phylogenetic groups from Fig. 3. Of the 12 independent variables used in the analysis (host species, elevation, geographic region, geographic subregion, soil pH, latitude, plant phylogenetic grouping, and five predicted climate parameters for each sample location as described in the text), only host species and geographic region were significant.

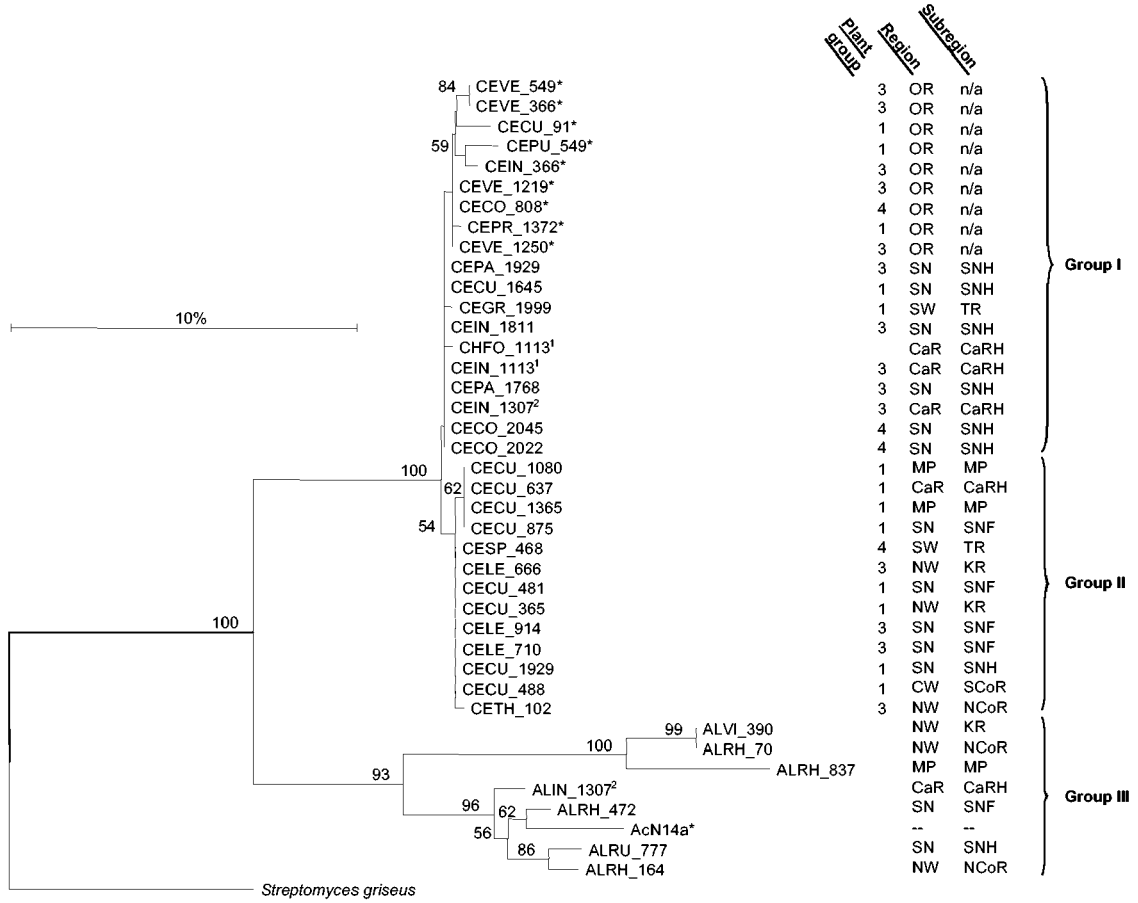

FIG. 5. Neighbor-joining tree based on IGS sequence data comparing previously published sequences (19, 32) with those for a subset of *Frankia* strains analyzed in this study that were chosen to be representative of the geographic regions covered. Previously published sequences (19, 32) are marked by asterisks. The tree is rooted with a sequence from *Streptomyces griseus*. Bootstrap values represent 1,000 replicates; only values above 50% are shown. Branch lengths correspond to sequence differences as indicated by the scale bar. Designations at terminal nodes consist of a four-letter host species code followed by the elevation (in meters) where the plant was growing. Plant groups are based on a molecular phylogeny of *Ceanothus* (17), with groupings inferred on the basis of morphological characters for *C. greggi*, *C. leucodermis*, *C. parvifolius*, and *C. spinosus* (CESP). Identical superior numbers indicate sympatric paired samples. CEVE, *C. velutinus*; CEPU, *C. pumilus*; CEPR, *C. prostratus*; CETH, *C. thyrsifolius*. Other host species codes are as explained in the legends to Fig. 2 and 3.

nosa, was more similar to some of our *Alnus* symbionts than they were to each other.

DISCUSSION

Variability among *Frankia* strains within a given host-specificity group has been documented previously, but the causes of this diversity have not been comprehensively explored. Several factors correlating with intra-host group *Frankia* heterogeneity have been proposed, including elevation (32), soil C content (25), and soil pH (16). Some of these have correlated well with differences among *Frankia* strains, but it has proven difficult to isolate the relative importance of these factors. In this study, we considered four main factors influencing variability among *Frankia* strains within and between host specificity groups: host plant identity, geographic location, climate, and soil pH. It is difficult to determine which of these factors may be the strongest driver because all are confounded to some degree. For example, climate and soil conditions often vary by geographic location and affect which host plants are present. To solve this problem, we attempted to isolate the effect of each of these factors by using statistical and climate models, conducting a greenhouse experiment, and applying inferences from the models to a broad geographic and host plant sampling representative of California and neighboring regions.

When paired samples belonged to different host specificity groups, the differences between the two *Frankia* strains were dramatic, consistent, and directly attributable to the identity of the host. This result is in agreement with previous phylogenetic classifications of *Frankia* strains, which have consistently been found to form distinct clades corresponding to host plant taxonomy at the family level or above (5, 9, 10, 27). *Frankia* strains that nodulate *Alnus* (*Betulaceae*) form a clade with *Frankia* strains symbiotic with plants in the *Casuarinaceae* and *Myricaceae* (27), and *Ceanothus* symbionts are most similar to *Frankia* strains from nodules of members of the *Rosaceae*, *Coriariaceae*, and *Datiscaceae* (5, 8, 9, 10, 27).

Because of the homogeneity that we found among the *Frankia* strains within the *Ceanothus*/*Chamaebatia* host group, it was much more difficult to discern the relative importance of host, environment, and geographic location. For example, the

mean elevation, predicted mean annual temperature, and predicted mean annual precipitation of phylogenetic group I were all significantly different from those of group II, but the host species assemblages of the two groups also had little overlap. Because host species distribution and climate are also confounded with geographic location, we first used the PRISM climate model to predict climatic variables for each sample location (effectively normalizing for latitude, elevation, slope and aspect) and then used the CART model to identify the factors that were most significant in distinguishing between *Frankia* phylogenetic groups. In the CART model, host species was the most significant variable in distinguishing between the phylogenetic groups within the *Ceanothus*/*Chamaebatia* clade. Based on these results, we concluded that even within a host specificity group, variation among *Frankia* strains is primarily a function of host species identity.

We expected to find a high diversity of *Frankia* strains symbiotic with *Ceanothus*, because California is the center of distribution of the genus (22) and contains a wide range of climate zones. In the only previous study in California of the molecular diversity of *Frankia* strains inhabiting *Ceanothus* nodules, Murry et al. (24) identified 12 unique *Frankia* strains from six *Ceanothus* species collected from the Santa Monica mountains. Despite the fact that some of our samples overlapped geographically and taxonomically, our results do not confirm this degree of diversity, although this may be attributable to the greater resolution of the rep-PCR method used by Murry et al. (24).

Benson et al. (5) also found low diversity of *Frankia* strains symbiotic with *Ceanothus* in the eastern United States and suggested that *Frankia* strains in this host specificity group are depauperate in general. Indeed, Clawson et al. (7) found low overall diversity of *Coriaria* symbionts (members of the same host specificity group as *Ceanothus* symbionts) in New Zealand. Ritchie and Myrold (32) drew similar conclusions in their study of *Ceanothus* nodules in Oregon, and when we included their sequence data in the analysis, their samples were not distinct from our phylogenetic group I.

The reasons for this low diversity remain unclear; however, the time of divergence of *Ceanothus* relative to other actinorhizal genera may be important. Molecular clock studies estimate the time of divergence of the two subgenera within *Ceanothus* as 18 to 39 million years ago, and evidence for recent speciation supports the high genetic similarity of many *Ceanothus* species (17). In contrast, *Myrica* and *Alnus* symbionts are highly diverse, and the two genera are among the most ancient of the actinorhizal plants, first appearing in the fossil and pollen records 80 to 110 million years ago (11, 23).

If geographic location has an important effect on *Frankia* variation within a host specificity group, the observed phylogenetic patterns should cluster into groups which correlate with spatial location. In the CART analysis of statewide samples, geographic location was a secondary (but still significant) variable in explaining the *Frankia* phylogenetic groupings. Variations among *Frankia* strains within a host group that correlate with geographic location have been found previously (4, 19, 32), yet it is difficult to infer the exact mechanisms of geographic influence, which could involve reproductive isolation due to past glaciation, local competitive dynamics, environmental differences among different regions (controlled for in this case by use of the PRISM model), or dispersal patterns. *Frankia* dispersal is poorly understood, but known *Frankia* dispersal vectors include birds (31) and water (1). *Frankia* appears to be able to persist saprophytically in soils without a host plant (see, e.g., references 20, 25, and 26), and because it can remain viable when air dried (29), wind-borne dispersal may also be common.

In the present study, *C. cuneatus* samples from the Modoc Plateau and the Cascade Range (phylogenetic group 2) were significantly different from the other samples as indicated by the phylogenetic and CART analyses. *C. cuneatus* has one of the widest distributions of any *Ceanothus* species and is found throughout California (14). Several subspecies of *C. cuneatus* exist, including several in northern California (14), but sample collections in this study were not identified to the subspecies level. If these three samples are indeed somehow genetically distinct from other *C. cuneatus* plants, this would strengthen the conclusion of host specificity as the most important factor associated with variation among *Frankia* strains.

Greenhouse experiment. By collecting *Frankia* strains from sympatric host pairs and normalizing for latitude and geographic region by use of the PRISM model, we concluded that host identity was more important than environment in determining *Frankia* identity, but we also sought an experimental test of the importance of host species versus environment. *Frankia* strains that formed nodules with *C. cordulatus* under climate-controlled conditions in the greenhouse in Seattle grouped with *Frankia* strains inhabiting nodules from the same host collected from California. Consistent with the CART model, the experiment suggests that climate alone is a less important influence on the diversity of *Frankia* strains forming nodules than the plant species and soils that are present locally.

The evidence presented here for host specificity as the main driver of differences among *Frankia* strains collected from *Ceanothus* does not necessarily provide support for a coevolutionary relationship. There was only limited correspondence between the phylogenetic relationships of the *Frankia* strains we collected and previously established *Ceanothus* phylogenies constructed on the basis of morphological (22) or molecular (17) data (not shown). Because host plant and *Frankia* phylogenies also correspond only loosely for the main host specificity groups defined to date (27), it has been suggested that the actinorhizal symbiosis has originated several times in evolutionary history (18, 35), which could explain the anomalous grouping of *Frankia* strains symbiotic with *Ceanothus* relative to other actinorhizal genera belonging to the *Rhamnaceae*. Within *Ceanothus*, we did not find evidence unequivocally supporting coevolution but concluded that *Frankia* differentiation within *Ceanothus* involves both host specificity and geographic isolation. Because distributions of host plants are often confounded with geographic location, further sampling of sympatric hosts belonging to different clades within the same host specificity group and elucidations of the distribution mechanisms of *Frankia* strains are clearly warranted.

ACKNOWLEDGMENTS

Financial support for this research was provided by a cooperative research agreement between the USDA Forest Service Pacific Southwest (PSW) Research Station and the University of Washington. Additional funding was provided by a PSW competitive grant to M.N. and

B.B.O. and an award from the Northwest Scientific Association to B.B.O.

We thank Cheryl Jenkins, David Myrold, and Nancy Ritchie for much helpful advice and assistance.

REFERENCES

1. **Arveby, A. S., and K. Huss-Danell.** 1988. Presence and dispersal of infective *Frankia* in peat and meadow soils in Sweden. Biol. Fertil. Soils **6:**39–44.
2. **Baker, D. D., and C. R. Schwintzer.** 1990. Introduction, p. 3–11. *In* C. R. Schwintzer and J. D. Tjepkema (ed.), The biology of *Frankia* and actinorhizal plants. Academic Press, Inc., San Diego, Calif.
3. **Baker, D. D., and B. C. Mullin.** 1992. Actinorhizal symbioses, p. 259–292. *In* G. Stacey, R. H. Burris, and H. J. Evans (ed.), Biological nitrogen fixation. Chapman and Hall, New York, N.Y.
4. **Baker, D. D., and B. C. Mullin.** 1994. Diversity of *Frankia* nodule endophytes of the actinorhizal shrub *Ceanothus* as assessed by RFLP patterns from single nodule lobes. Soil Biol. Biochem. **26:**547–552.
5. **Benson, D. R., D. W. Stephens, M. L. Clawson, and W. B. Silvester.** 1996. Amplification of 16S rRNA genes from *Frankia* strains in root nodules of *Ceanothus griseus*, *Coriaria arborea*, *Coriaria plumosa*, *Discaria toumatou*, and *Purshia tridentata*. Appl. Environ. Microbiol. **62:**2904–2909.
6. **Chambers, J. M., and T. J. Hastie.** 1992. Statistical models in S. Wadsworth & Brooks/Cole Advance Books & Software, Pacific Grove, Calif.
7. **Clawson, M. L., D. R. Benson, S. C. Resch, D. W. Stephens, and W. B. Silvester.** 1997. Typical *Frankia* infect actinorhizal plants exotic to New Zealand. N. Z. J. Bot. **35:**361–367.
8. **Clawson, M. L., M. Caru, and D. R. Benson.** 1998. Diversity of *Frankia* strains in root nodules of plants from the families *Eleagnaceae* and *Rhamnaceae*. Appl. Environ. Microbiol. **64:**3539–3543.
9. **Clawson, M. L., and D. R. Benson.** 1999. Natural diversity of *Frankia* strains in actinorhizal root nodules from promiscuous hosts in the family *Myricaceae*. Appl. Environ. Microbiol. **65:**4521–4527.
10. **Clawson, M. L., J. Gawronski, and D. R. Benson.** 1999. Dominance of *Frankia* strains in stands of *Alnus incana* subsp. *rugosa* and *Myrica pensylvanica*. Can. J. Bot. **77:**1203–1207.

10a. **Conard, S. G., A. E. Jaramillo, K. Cromack, Jr., and S. Rose.** 1985. The role of the genus *Ceanothus* in western forest ecosystems. USDA Forest Service General Technical Report PNW-182. USDA Pacific Northwest Research Station, Portland, Oreg.

11. **Crane, P. R.** 1989. Early fossil history and evolution of the *Betulaceae*, p. 87–116. *In* P. R. Crane and S. Blackmore (ed.), Evolution, systematics, and fossil history of the *Hamamelidae*, vol 2. 'Higher' *Hamamelidae*. Oxford University Press, New York, N.Y.
12. **Daly, C., R. P. Neilson, and D. L. Phillips.** 1994. A statistical-topographic model for mapping climatological precipitation over mountainous terrain. J. Appl. Meteorol. **33:**140–158.
13. **Daly, C., G. H. Taylor, and W. P. Gibson.** 1997. The PRISM approach to mapping precipitation and temperature. *In* Proceedings of the 10th American Meteorological Society Conference on Applied Climatology. American Meteorological Society, Boston, Mass.
14. **Hickman, J. C. (ed.).** 1993. The Jepson manual: higher plants of California. University of California Press, Berkeley, Calif.
15. **Hönerlage, W., D. Hahn, K. Zipp, J. Zeyer, and P. Normand.** 1994. A hypervariable 23S rRNA region provides a discriminating target for specific characterization of uncultured and cultured *Frankia*. Syst. Appl. Microbiol. **17:**433–443.
16. **Huguet, V., J. McCray Batzli, J. F. Zimpfer, P. Normand, J. O. Dawson, and M. P. Fernandez.** 2001. Diversity and specificity of *Frankia* strains in nodules of sympatric *Myrica gale*, *Alnus incana*, and *Shepherdia canadensis* determined by *rrs* gene polymorphism. Appl. Environ. Microbiol. **67:**2116–2122.
17. **Jeong, S. C., A. Liston, and D. D. Myrold.** 1997. Molecular phylogeny of the genus *Ceanothus* (*Rhamnaceae*) using *rbcL* and *ndhF* sequences. Theor. Appl. Genet. **94:**852–857.
18. **Jeong, S.-C., N. J. Ritchie, and D. D. Myrold.** 1999. Molecular phylogenies of plants and *Frankia* support multiple origins of actinorhizal symbioses. Mol. Phylogenet. Evol. **13:**493–503.
19. **Jeong, S.-C., and D. D. Myrold.** 1999. Genomic fingerprinting of *Frankia* microsymbionts from *Ceanothus* copopulations using repetitive sequences and polymerase chain reactions. Can. J. Bot. **77:**1220–1230.
20. **Jeong, S.-C., and D. D. Myrold.** 2001. Population size and diversity of *Frankia* in soils of *Ceanothus velutinus* and Douglas-fir stands. Soil Biol. Biochem. **33:**931–941.
21. **Johnson, D. A.** 1992. Ph.D. thesis. University of Ottawa, Ottawa, Canada.
22. **McMinn, H. E.** 1942. *Ceanothus*. Part II. A systematic study of the genus *Ceanothus*. The Santa Barbara Botanic Garden, Santa Barbara, Calif.
23. **Muller, J.** 1984. Significance of fossil pollen for angiosperm history. Ann. Missouri Bot. Garden **74:**419–443.
24. **Murry, M. A., A. S. Konopka, S. D. Pratt, and T. L. Vandergon.** 1997. The use of PCR-based typing methods to assess the diversity of *Frankia* nodule endophytes of the actinorhizal shrub *Ceanothus*. Physiol. Plant. **99:**714–721.
25. **Myrold, D. D., and K. Huss-Danell.** 1994. Population dynamics of *Alnus*-infective *Frankia* in a forest soil with and without host trees. Soil Biol. Biochem. **26:**533–540.
26. **Nalin, R., P. Normand, and A. Domenach.** 1997. Distribution and N_2-fixing activity of *Frankia* strains in relation to soil depth. Physiol. Plant. **99:**732–738.
27. **Normand, P., S. Orso, B. Cournoyer, P. Jeannin, C. Chapelon, J. Dawson, L. Evtushenko, and A. K. Misra.** 1996. Molecular phylogeny of the genus *Frankia* and related genera and emendation of the family *Frankiaceae*. Int. J. Syst. Bacteriol. **46:**1–9.
28. **North, M., B. Oakley, J. Chen, H. Erickson, A. Gray, A. Izzo, D. Johnson, S. Ma, J. Marra, M. Meyer, K. Purcell, T. Rambo, D. Rizzo, B. Roath, and T. Schowalter.** 2002. Vegetation and ecological characteristics of mixed-conifer and red-fir forests at the Teakettle Experimental Forest. USDA Forest Service General Technical Report PSW-GTR-186. USDA Pacific Southwest Research Station, Albany, Calif.
29. **Oremus, P. A.** 1980. Occurrence and infective potential of the endophyte of *Hippophaë rhamnoides* L. ssp. *rhamnoides* in coastal sand-dune areas. Plant Soil **56:**533–540.
30. **Page, R. D. M.** 1996. TREEVIEW: an application to display phylogenetic trees on personal computers. Comput. Appl. Biosci. **12:**357–358.
31. **Paschke, M. W., and J. O. Dawson.** 1993. Avian dispersal of *Frankia*. Can. J. Bot. **71:**1128–1131.
32. **Ritchie, N. J., and D. D. Myrold.** 1999. Geographic distribution and genetic diversity of *Ceanothus*-infective *Frankia* strains. Appl. Environ. Microbiol. **65:**1378–1383.
33. **Schwintzer, C. R., and J. D. Tjepkema.** 1990. The biology of *Frankia* and actinorhizal plants. Academic Press, Inc., San Diego, Calif.
34. **Stacey, G., R. H. Burris, and H. J. Evans.** 1992. Biological nitrogen fixation. Chapman and Hall, New York, N.Y.
35. **Swensen, S. M.** 1996. The evolution of actinorhizal symbioses: evidence for multiple origins of the symbiotic association. Am. J. Bot. **83:**1503–1512.
36. **Van de Peer, Y., and R. De Wachter.** 1993. TREECON: a software package for the construction and drawing of evolutionary trees. Comput. Appl. Biosci. **9:**177–182.
37. **Venables, W. N., and B. D. Ripley.** 1994. Modern applied statistics with S-Plus. Springer-Verlag, New York, N.Y.