Journal of Nematology 24(1):140–147. 1992.

Taxonomic Affinities and Intra- and Interspecific Variation in *Bursaphelenchus* spp. as Determined by Polymerase Chain Reaction

KAREN BECKENBACH,[1] MICHAEL J. SMITH,[1,2] AND JOHN M. WEBSTER[1]

Abstract: Identification of closely related nematode species or races can be very difficult when diagnostic characters are plastic and overlapping. In this study we describe the use of polymerase chain reaction technology and direct DNA sequencing on 19 populations of *Bursaphelenchus* spp. to help understand their taxonomic relationships. The 5′ end of the heat shock 70A gene from *Caenorhabditis elegans* was used as the target DNA sequence because it contains both coding and noncoding regions. The results indicate that the 19 populations could be divided into five types within *B. xylophilus* and four types within *B. mucronatus*. On a larger scale, the data revealed three distinct groups, representing *B. xylophilus* from North America and Japan, *B. mucronatus* from Japan, and "*B. mucronatus*" from Europe. There is sufficient difference between the European and Japanese "*B. mucronatus*" groups to warrant their consideration as separate species.

Key words: Bursaphelenchus mucronatus, B. xylophilus, Caenorhabditis elegans, DNA, gene, heat shock, molecular phylogeny, nematode, pinewood nematode, polymerase chain reaction, species complex, taxonomy.

Bursaphelenchus xylophilus and *B. mucronatus,* members of the pinewood nematode species complex (PWNSC), are economically important to lumber trading nations because of their pathogenicity to pines and some related genera. These nematode species are biologically interesting because of their life cycles, variable pathogenicities, and wide geographic distribution of different infraspecific forms (15,21). The uncertainty in taxonomic distinction among some members within this complex has led to "super species" (5), "supra species" (6), and "species complex" (21) descriptors. Recently, the relative taxonomic status of some populations within this complex was partially clarified by demonstrating genotypic differences among isolates from closely and widely separated geographic populations with DNA probes (1,28).

Identification of closely related species is difficult, especially if samples are small or if diagnostic characters are environmentally influenced or overlapping in nature. In many instances, analysis of DNA sequences from the species in question can clarify taxonomic relationships and provide an accurate identification based on a fundamental character (2). Until recently, it was difficult and time consuming to analyze and compare DNA sequences for a large number of populations, especially for small individuals such as some nematode species. With the development of the polymerase chain reaction (PCR) (18,22), it is now possible to obtain and analyze DNA sequence variation rapidly within and between populations, and single nucleotide differences between individual nematodes can be located and identified. Comparison of the frequency and location of nucleotide differences between genomes reveals the degree of genomic diversity and, hence, the taxonomic affinities of a given group (2,13,27). The objective of this study was to differentiate among populations of the PWNSC from a wide geographic range on the basis of their relative genetic affinities. By so doing, it should be

Received for publication 7 June 1991.

[1] Department of Biological Sciences, Simon Fraser University, Burnaby, Vancouver, British Columbia, V5A 1S6, Canada.

[2] Institute of Molecular Biology and Biochemistry, Simon Fraser University.

We thank Donna Bartlett, Susan Chow, and Paul Lowenstein for stock maintenance, and Forestry Canada, the Science Council of British Columbia, and the Natural Sciences and Engineering Research Council (A4679) for financial support. We also thank T. P. Snutch and D. L. Baillie for providing us with the *C. elegans* 70A heat shock gene clone pCes462. We acknowledge the perceptive and helpful comments of the reviewers.

possible to further clarify the taxonomic relationships within this species complex.

Materials and Methods

Nineteen populations of the PWNSC, eleven of *B. xylophilus* from North America and Japan, and eight of *B. mucronatus* from Europe and Japan (Table 1), were maintained and cultured as described (28).

Cloning, screening, and sequencing strategies: Genomic DNA was isolated from at least 100 μl of nematodes from each population (28). To construct a genomic phage library, DNA from the Ibaraki population of *B. xylophilus* was digested with Eco RI and ligated into the Eco RI site of lambda Zap (Stratagene, La Jolla, CA) following the protocol provided with this vector. The ligation reaction was packaged into phage using Gigapack Gold (Stratagene). These phage were grown in *E. coli* strain BB4 on NZY plates containing 10 μg/ml tetracycline at 37 C.

The Ibaraki genomic library was screened with a ^{32}P-labeled plasmid clone, pCes462, containing the hsp70A gene from *Caenorhabditis elegans* (25). Hybridizations were performed at 62 C in 5× SSPE (1× SSPE = 0.18 M NaCl, 10 mM $(Na_{1.5})PO_4$, 1 mM Na_2EDTA, pH 7.0), 0.3% sodium dodecyl sulfate (SDS), and 5× Denhardt's solution (1× Denhardt's = 0.02% w/v bovine serum albumin, Ficoll 400, and polyvinylpyrrolidone 40) (4). Nitrocellulose filters were washed at 45 C in 2× SSPE + 0.2% SDS, air dried, and exposed to Kodak X-OmatK X-ray film. Six of the plaques that gave the strongest autoradiographic signal were selected, amplified, and again probed with pCes462. The recombinant phage that reacted most intensely in this second screen was chosen, and the Bluescript plasmid was induced to excise from the phage. The phagemids produced were used to infect *E. coli* strain BB4, which was plated on NZY plates containing 50 μg/ml ampicillin (Stratagene).

Table 1. The name, source, and sequence type number of populations of *Bursaphelenchus xylophilus* and *B. mucronatus*.

Population	Sequence type number†	Habitat	Origin	Source‡
		B. xylophilus		
Q52A	1	woodchips	Canada (Quebec)	I
Q1426§	1	Balsam fir	Canada (Quebec)	I
mm§	2	*Monochamus* spp	Canada (BC)	II
St. John	3	woodchips	Canada (New Brunswick)	I
St. William§	3	nursery	Canada (Ontario)	I
British Columbia	3	woodchips	Canada	I
Alberta	4	unknown	Canada	I
MsP4§	4	unknown	USA (Missouri)	III
BxUJa	5	unknown	USA (Georgia)	III
Vermont La	5	Larch	USA	VI
Ibaraki§	5	Black pine	Japan	VII
		B. mucronatus		
Norway	6	unknown	Norway	VIII
French 1§	6	unknown	France	IV
French 2	6	unknown	France	IV
Smaland	6	unknown	Sweden	IX
F-88-287	6	woodchips	Finland	V
F-88-237§	7	woodchips	Finland	V
F-88-286§	8	woodchips	Finland	V
RB§	9	unknown	Japan	III

† The sequence type number indicates sequence identity in the gene region studied.

‡ Source abbreviations I = R. V. Anderson, II = J. R. Sutherland, III = R. I. Bolla, IV = G. de Guiran, V = D. Trotter, VI = S. Halik, VII = Y. Mamiya, VIII = D. G. McNamara, IX = C. Magnusson.

§ These populations are the representatives used in the phylogenetic analysis.

Plasmid DNA was isolated by the alkaline lysis method (16). The ends of the cloned insert, pBx10, were sequenced (24) with Sequenase (United States Biochemical, Cleveland, OH).

To obtain the sequence of the entire clone, two strategies were used. First, clone pBx10, which contained a 1.3-kb Eco RI fragment, was restricted with Hind III and Eco RI, and the resulting fragments were subcloned into pVZ I (9). These subclones (named pBx11 and pBx12), containing 0.68-kb and 0.62-kb Eco RI and Hind III fragments, respectively, were sequenced and aligned with both the parental clone (pBx10) and the *C. elegans* hsp70A gene (25). The second method involved the use of nested deletions (8). The clone pBx10 was restricted with Kpn I and Sal I, treated with Exonuclease III and SI nuclease, ligated, and transformed into competent (17) *E. coli* strain JM 83. Deletions of about 0.2 kb, 0.5 kb, and 0.7 kb were chosen and sequenced.

Polymerase chain reaction: To prepare specific PCR primers, the sequence of pBx10 was compared to the hsp70A and hsp70C genes (Fig. 1), the two most closely related genes in the *C. elegans* gene family (10). By this means we insured that only the

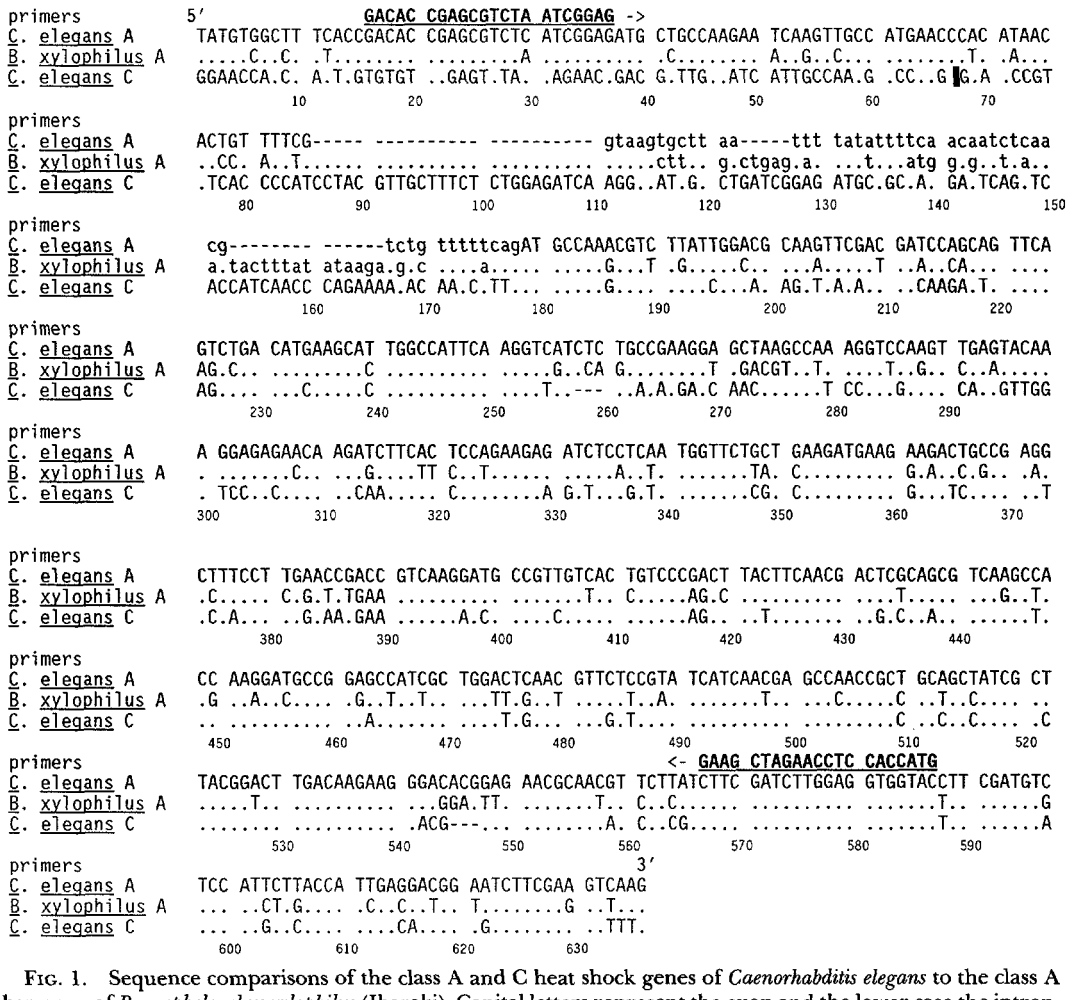

FIG. 1. Sequence comparisons of the class A and C heat shock genes of *Caenorhabditis elegans* to the class A hsp gene of *Bursaphelenchus xylophilus* (Ibaraki). Capital letters represent the exon and the lower case the intron. The heavy line (▌) in the class C gene represents an intron position. (The intron has been removed from the C gene to facilitate the alignment of the sequences.) The first line indicates the position of the primers and the arrow indicates the direction of DNA synthesis in the polymerase chain reaction.

hsp70A gene homologs were amplified. Primer 1 extended from nucleotide position 136 to 158 and had 21/22 nucleotide identity with the *C. elegans* hsp70A gene but only 6/22 with the hsp70C gene. Primer 2, position 662 to 683, was identical with the hsp70A and the hsp70C genes. These oligonucleotide primers were synthesized in an Applied Biosystems (Foster City, CA) 391 DNA Synthesizer.

The PCR reactions contained 5 μl 10× reaction buffer (10× = 100 mM Tris at pH 8.3, 500 mM KCl, 15 mM $MgCl_2$, 0.1% [w/v] gelatin), 8 μl dNTPs (1.25 μM each nucleotide), 2.6 μl of each primer (10 μM), 28.6 μl sterile H_2O, 2 units Taq Polymerase (Perkin Elmer Cetus, Norwalk, CT), and 3.0 μl template DNA (about 1–10 ng total) (18,22). About 20–30 μl of mineral oil were added to each reaction solution to prevent evaporation. After several tests, the following PCR conditions were selected: Cycle 1) denature at 95 C for 90 seconds, anneal at 57 C for 30 seconds, and extend at 72 C for 60 seconds; Cycle 2) denature at 95 C for 45 seconds, anneal at 57 C for 30 seconds, extend at 72 C for 60 seconds, and repeat cycle 2 forty times; Cycle 3) denature at 95 C for 45 seconds, anneal at 57 C for 30 seconds, and extend at 72 C for 10 minutes. This regime generally produced a single product of about 570 base pairs, which was purified on Centricon 30 columns (Amicon, Danvers, MA). It should be noted that the template DNA in the PCR reactions is a sample from a population of nematodes from each area collected; thus the PCR product may represent the genome that reflects the most abundant sequence type in the population. To ensure that nucleotide differences were truly present, each sample was amplified and sequenced two to four times.

Sequencing and analysis of PCR product: The 5′ end containing the intron, which was expected to be more divergent than the coding sequence, of each PCR-amplified product was sequenced according to protocol two from Kessing (12), modified by replacement of water in the extension step with a 1:50 dilution of the extension mix. The DMSO method (T. P. Snutch, pers. comm.) in conjunction with the Sequenase kit was used on some of the PCR products to reduce secondary structure. In the latter case, the 1:5 dilution of the extension mix was increased to a 1:50 dilution.

The sequences were analyzed with an eyeball sequence editor, ESEE (3), an alignment program, FastA (19), bootstrap and parsimony from Felsenstein's Phylip 3.3 (7), and neighbor-joining analysis (23) using NJTREE (version 2.0; N. Saitou and L. Jin, pers. comm.) and NJDRAW (version 1.0, N. W. H. Ferguson, pers. comm.).

Results

The 11 *B. xylophilus* populations fell into five sequence types based on the sequence identity of the analyzed fragment (Table 1). Similarly, *B. mucronatus* fell into four sequence types. The sequences were analyzed using one representative sequence of each population type. Sequence type 1 was represented by Q1426, type 2 by mm, type 3 by St. William, type 4 by MsP4, type 5 by Ibaraki, type 6 by French 1, type 7 by F-237, type 8 by F-286, and type 9 by RB.

The alignment of the five *B. xylophilus* sequence types, the four *B. mucronatus* types, and the *C. elegans* sequence is shown in Figure 2. The two most divergent types of *B. xylophilus* (sequence divergence of 5.9% [189/201]) were type 1, containing two populations from Quebec, Canada, and type 5, containing two populations from the United States and one from Japan. The sequence variation among the *B. mucronatus* groups shows two extremes. Within the European populations, a maximum sequence divergence of 1.5% was observed, whereas between the European and Japanese sequence types it was 23.5%.

The sequence divergence observed across the entire 201 base pairs between *B. xylophilus* and the European and Japanese *B. mucronatus* was 14.6% and 18.9%. If divergence in the coding sequence alone was considered, these percentages dropped to 5.7% and 0.8%. The divergence between the European and Japanese *B. mucronatus*

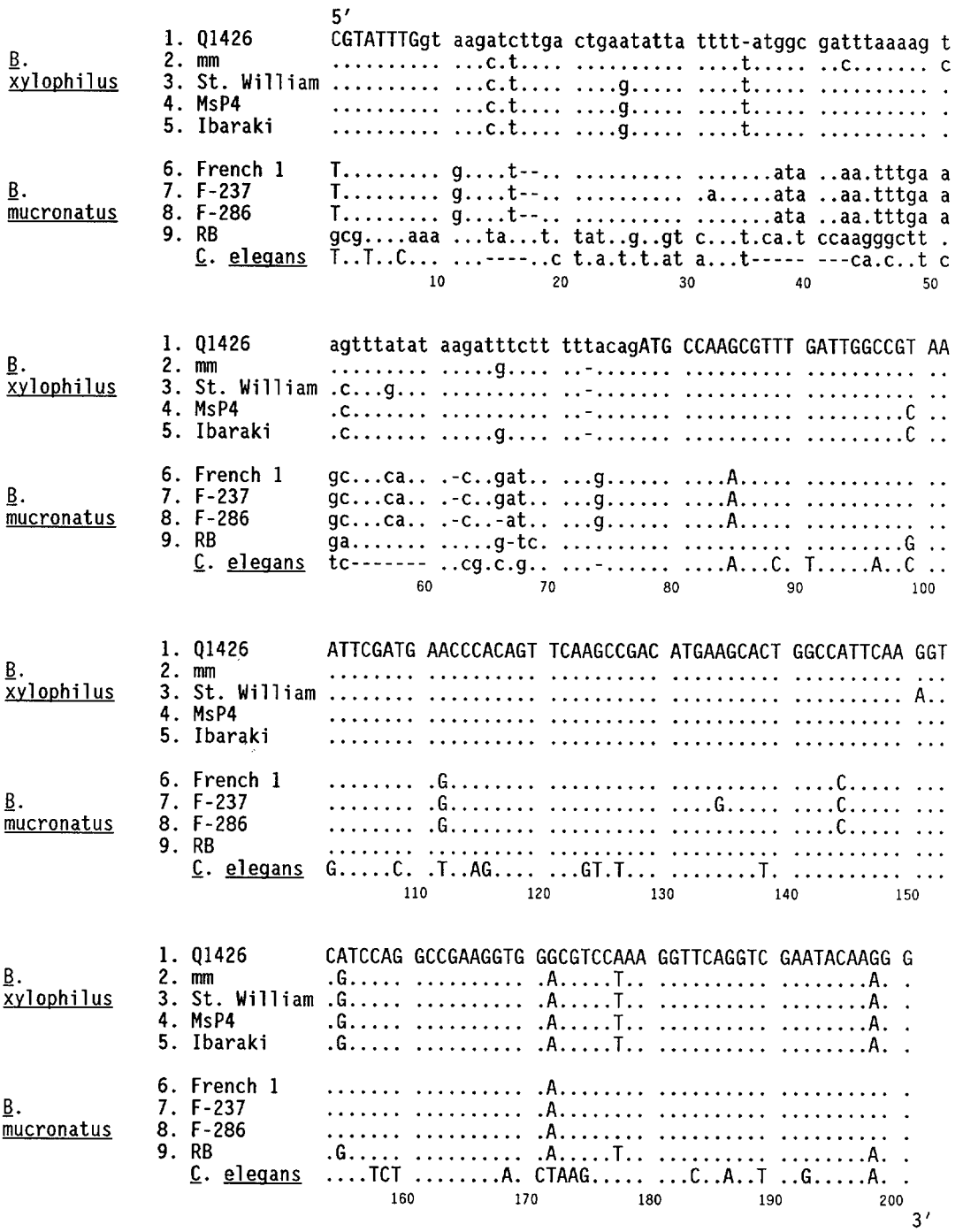

FIG. 2. Alignment of the 5′ end of the heat shock gene for each representative sequence type from *Bursaphelenchus xylophilus, B. mucronatus,* and *Caenorhabditis elegans.* Capital letters represent the exon, lower case letters the intron, and the dots indicate sequence homology. See Table 1 for the complete list of populations in each sequence type.

was 23.5% across all 201 base pairs and 5.7% in the coding region.

The sequence relationships were analyzed with bootstrap, parsimony, and neighbor-joining methods. Parsimony and multiple runs of bootstrap produced five similar trees. The variations among trees result from slight sequence divergence

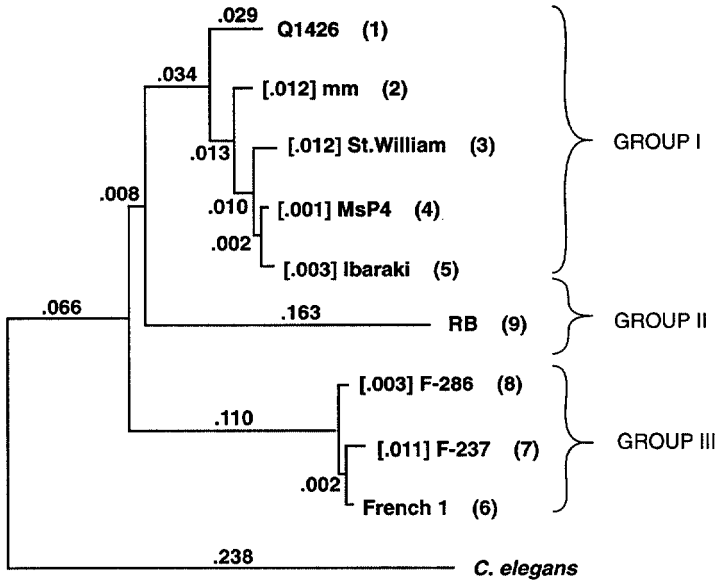

FIG. 3. The phylogenetic relationships among five population types of *Bursaphelenchus xylophilus* and four population types of *B. mucronatus*, with *C. elegans* included in the analysis as an outgroup. This analysis, using neighbor-joining in conjunction with NJTREE and NJDRAW, was performed on the sequences derived from the PCR product. Numbers 1 through 9 were assigned to each of the terminal points. Group numbers I, II, and III, corresponding to *B. xylophilus*, the Japanese *B. mucronatus*, and the European *B. mucronatus*, were assigned to each major branch. The decimal numbers on the tree branches are the Jukes–Cantor corrected sequence divergence estimates (11).

among types 3, 4, and 5 (represented by St. William, MsP4, and Ibaraki in the *B. xylophilus* branch) and types 6, 7, and 8 (representing the European *B. mucronatus* branch). Figure 3 contains a representative neighbor-joining tree. All analyses strongly support three distinct branches or groups of *Bursaphelenchus*—Group I, *B. xylophilus* from North America and Japan; Group II, *B. mucronatus* from Japan; and Group III, European *B. mucronatus*. The maximum sequence divergence of values within Group I (5.9%) and Group III (1.5%) are within the range of sequence variation documented for the Adh gene from populations of *Drosophila melanogaster* (14).

DISCUSSION

The sensitivity of the PCR method has enabled us to distinguish between some populations from a range of geographic locations within *B. xylophilus* and *B. mucronatus*. The PCR sequences presented here essentially represent the most abundant sequence present in a pool of 500–1,000 individuals, which in turn may provide the most common sequence from each geographic location sampled. The one exception is the mm isolate, which represents a very small collection from a single *Monochamus* beetle.

The groupings of *B. xylophilus* encourage speculation about the geographic origin of some of the populations as explained by lumber trade routes or by continuity of nematode habitat. Sequence types 4 and 5 are only 0.5% divergent (200 of 201 nucleotides), suggesting within-population variation of a contiguous eastern North American group. This cluster, which also contains a population from Japan, may indicate a common origin of these five populations and tends to support an earlier hypothesis (29) that Japan was infested originally by a North American population. The mm isolate (type 2) was collected from a *Monochamus* spp. beetle in British Columbia and is 3.5% different from the next closest sequence type, which is type 3. Type 3 contains another

isolate from British Columbia, as well as populations from central and eastern Canada. The sequence difference between the two isolates from British Columbia may indicate a native or local western population represented by the mm isolate versus an introduced population, the B.C. isolate. The two type 1 populations from Quebec are identical with each other but are 5.4% divergent from the next most similar group and probably represent an original, undisturbed population in Quebec. Further speculation about population types and geographic origins would require sampling from more locations.

The European *B. mucronatus* populations are extremely similar to one another, with only three of 201 nucleotides different among these populations. This is consistent with the continuity of European forests. On the other hand, the current distribution may be very recent, with insufficient time for sequence divergence to have accumulated. The *B. mucronatus* population from Japan is quite distinct from the European populations. This distinction agrees with cross-hybridization experiments in our laboratory (20), in which a French and a Japanese population of *B. mucronatus* did not freely interbreed, but contradicts the report (6) of interbreeding of another European and Japanese population. This conflict may result from the use of different populations, because we have observed limited interbreeding between some populations of *B. xylophilus* and the French and Japanese population of *B. mucronatus*.

Interestingly, the data presented here show a sequence divergence of 18.9% between *B. xylophilus* and the Japanese *B. mucronatus*, 14.6% between *B. xylophilus* and the European *B. mucronatus*, and 23.5% between the European and Japanese *B. mucronatus*. This suggests that gene flow has not occurred between the European and Japanese populations of *B. mucronatus* as well as between each of them and the North American populations of *B. xylophilus*, supporting the de Guiran and Bruguier (6) "circle of species" theory.

Abad et al. (1), using three, six-base cutter restriction enzymes in conjunction with the heterologous *unc-22* gene probe from *C. elegans*, concluded that the European and Japanese populations of *B. mucronatus* are the same species and are distinct from *B. xylophilus*. This contradiction with our results may be due in part to the failure of the intergeneic regions of the *unc-22* gene probe to bind to the heterologous species and thus miss possible variation. Also, restriction fragment length polymorphisms tend to underrepresent the actual sequence variation that may be present because this method samples relatively few nucleotides (26). Hence Abad et al. (1) could easily have missed significant sequence variation between the European and Japanese populations of *B. mucronatus*, leading to their conclusion that these populations are probably the same species.

We believe that separate species status for the Japanese and European *B. mucronatus* populations is justified and an extensive comparative morphological study of individuals from representative populations is necessary as a prelude to verifying such species status. As a corollary, the descriptors "super species," "supra species," or "species complex" for *Bursaphelenchus* species would become redundant.

Note added in proof

The GenBank accession numbers for the sequences presented in this paper are as follows: BC, M84996; BxUJa, M84997; F-237, M84998; F-286, M84999; F-287, M85000; French 1, M85001; MsP4, M85002; Norway, M85003; Q1426, M85004; Smaland, M85005; St. John, M85006; St. William, M85007; Vermont La, M85008; French 2, M85009; Q52A, M85010; Ibaraki, M85011; mm, M85012; RB, M85013; Alberta, M85014.

Literature Cited

1. Abad, P., S. Tares, N. Bruguier, and G. de Gurian. 1991. Characterization of the relationships in the pinewood nematode species complex (PWNSC) (*Bursaphelenchus* spp.) using a heterologous *unc-22* DNA probe from *Caenorhabditis elegans*. Parasitology 102:303–308.

2. Brunk, C. F., R. W. Kahn, and L. A. Sadler. 1990. Phylogenetic relationships among *Tetrahymena* species determined using polymerase chain reaction. Journal of Molecular Evolution 30:290–297.

3. Cabot, E. L., and A. T. Beckenbach. 1989. Simultaneous editing of multiple nucleic acid and protein sequences with ESEE. Computer Applications in the Biosciences 5:233–234.

4. Davis, R. W., D. Botstein, and J. R. Roth. 1980. Advanced Bacterial Genetics. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory.

5. de Guiran, G., and A. Boulbria. 1986. Lé nématode des pins. Caractéristiques de la souche francaise et risque d'introduction et d'extension de *Bursaphelenchus xylophilus* en Europe. OEPP/EPPO Bulletin 16:445–452.

6. de Guiran, G., and N. Bruguier. 1989. Hybridization and phylogeny of pinewood nematode, (*Bursaphelenchus* spp.). Nematologica 34:321–330.

7. Felsenstein, J. 1989. PHYLIP—Phylogeny inference package (version 3.2). Cladistics 5:164–166.

8. Henikoff, S. 1984. Unidirectional digestion with exonuclease III creates targeted breakpoints for DNA sequencing. Gene 28:351–359.

9. Henikoff, S., and M. K. Eghtedarzadeh. 1987. Conserved arrangement of nested genes at the *Drosophila Gart* locus. Genetics 117:711–725.

10. Heschl, M. F. P., and D. L. Baillie. 1989. Characterization of the hsp70 multigene family of *Caenorhabditis elegans*. DNA 8:233–243.

11. Jukes, T. H., and C. R. Cantor. 1969. Evolution of protein molecules. Pp. 21–132 *in* H. N. Munro, ed. Mammalian protein metabolism, vol. 3. New York: Academic Press.

12. Kessing, B., H. Coon, A. Martin, C. McIntosh, W. O. McMilliam, and S. Palumbi. 1989. The simple fool's guide to PCR, p. 13, method 2. Available from S. Palumbi, Department of Zoology, University of Hawaii.

13. Kocher, T. D., W. K. Thomas, A. Meyer, S. V. Edwards, S. Paabo, F. X. Villablanca, and A. C. Wilson. 1989. Dynamics of mitochondrial DNA evolution in animals: Amplification and sequencing with conserved primers. Proceedings of the National Academy of Sciences USA 86:6196–6200.

14. Kreitman, M. 1983. Nucleotide polymorphism at the alcohol dehydrogenase locus of *Drosophila melanogaster*. Nature 304:412–417.

15. Mamiya, Y. 1987. Origin of the pinewood nematode and its distribution outside the United States. Pp. 59–65 *in* M. J. Wingfield, ed. Pathogenicity of the pinewood nematode. St. Paul: American Phytopathological Society Press.

16. Maniatis, T., E. G. Fritsch, and J. Sambrook. 1982. Molecular cloning, a laboratory manual. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory.

17. Morrison, D. A. 1979. Transformations and preservation of competent bacterial cells by freezing. Methods in Enzymology 68:326–331.

18. Mullis, K. B. 1990. The unusual origin of the polymerase chain reaction. Scientific American 262(4):56–65.

19. Pearson, W. 1990. Rapid and sensitive sequence comparison with FastP and FastA. Methods in Enzymology 183:63–98.

20. Riga, E., K. Beckenbach, and J. M. Webster. 1992. Taxonomic relationships of *Bursaphelenchus xylophilus* and *B. mucronatus* based on interspecific and intraspecific cross-hybridization and DNA analysis. Fundamental and Applied Nematology, in press.

21. Rutherford, T. A., Y. Mamiya, and J. M. Webster. 1990. Nematode-induced pine wilt disease: Factors influencing its occurrence and distribution. Forest Science 36:145–155.

22. Saiki, R. K., D. H. Gelfand, S. Stoffel, S. J. Scharf, R. Higuchi, G. T. Horn, K. B. Mullins, and H. A. Erlich. 1988. Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase. Science 239:487–491.

23. Saitou, N., and M. Nei. 1987. The neighbor-joining method: A new method for reconstructing phylogenetic trees. Molecular Biology and Evolution 4:406–425.

24. Sanger, F., S. Nicklen, and A. R. Coulson. 1977. DNA sequencing with chain-terminating inhibitors. Proceedings of the National Academy of Sciences USA 74:5463–5467.

25. Snutch, T. P. 1984. Molecular and genetic analysis of the heat shock response of *Caenorhabditis elegans*. Ph.D. thesis, Simon Fraser University, Burnaby, British Columbia.

26. Thomas, W. K., and A. T. Beckenbach. 1989. Variation in Salmonid mitochondrial DNA: Evolutionary constraints and mechanisms of substitutions. Journal of Molecular Evolution 29:233–245.

27. Vigilant, L., R. Pennington, H. Harpending, T. D. Kocher, and A. C. Wilson. 1989. Mitochondrial DNA sequences in single hairs from a southern African population. Proceedings of the National Academy of Sciences USA 86:9350–9354.

28. Webster, J. M., R. V. Anderson, D. L. Baillie, L. Beckenbach, J. Curran, and T. A. Rutherford. 1990. DNA probes for differentiating isolates of the pinewood nematode species complex. Revue de Nématologie 13:255–263.

29. Wingfield, M. J., R. A. Blanchette, and T. H. Nicholls. 1984. Is the pinewood nematode an important pathogen in the United States? Journal of Forestry 82:232–235.