PHYLOGENY AND BIOGEOGRAPHY OF TAXUS (TAXACEAE) INFERRED FROM SEQUENCES OF THE INTERNAL TRANSCRIBED SPACER REGION OF NUCLEAR RIBOSOMAL DNA

JIANHUA LI,1 CHARLES C. DAVIS,2 PETER DEL TREDICI,1 AND MICHAEL J. DONOGHUE3

Abstract. Species of Taxus, particularly Taxus baccata, have long been cultivated as ornamentals in Europe and North America. Ten species of Taxus are generally recognized, but phylogenetic relationships among these species remain unclear. We used sequences of the nuclear ribosomal DNA ITS region to infer phylogenetic relationships. Three of the four New World species form a well-supported clade, within which the Pacific coastal species T. brevifolia is sister to a clade containing T. floridana and T. globosa of northwestern Florida and northeastern Mexico, respectively. Taxus canadensis, which is more widely distributed in eastern North America, appears to be more closely related to Old World taxa than to other New World species. This relationship, though weakly supported in our analysis, is consistent with leaf anatomical features. Taxus chinensis and T. mairei of southeast Asia form a clade, which is sister to a clade containing T. cuspidata of Japan and northeastern China, and T. baccata of Europe and North Africa. Our ITS phylogeny implies that intercontinental disjunctions in Taxus entailed at least two vicariance events: an initial split between the New World T. floridana–T. brevifolia–T. globosa clade and the rest, and a later split separating T. canadensis from the Old World species.

Keywords: biogeography, ITS, phylogeny, Taxaceae, Taxus.

Species of Taxus L. (Taxaceae) are small to medium-sized trees or shrubs; their leaves are arranged on branches in two ranks, and two inconspicuous stomatal bands occur on the underside of the leaf. The seeds are wrapped at least halfway by a conspicuous aril, from orange to deep red in color. The monophyly of Taxus is supported by both morphological (Cheng and Fu, 1978; Hart, 1987; Judd et al., 1999) and molecular evidence (Cheng et al., 2000; Wang and Shu, 2000).

Some species of Taxus, such as the common yew, T. baccata L., have long been cultivated as ornamentals in Europe and North America (Hartzell, 1991). Recently Taxus has attracted public attention owing to the discovery of cancer-inhibitory taxol in the bark of the Pacific yew, T. brevifolia Nuttall (Wani et al., 1971).

Ten species of Taxus are generally recognized, and these are distributed primarily in the Northern Hemisphere (Fig. 1) from as far north as 61° N in Russia to as far south as Sumatra and El Salvador (Farjon, 1998). In the New World, T. canadensis Marshall is widely distributed in northeastern North America from Newfoundland to Manitoba and south to Virginia and Iowa. Taxus floridana Nuttall ex Chapman is restricted to small areas in the Appalachicola River valley in the panhandle of Florida, and T. brevifolia is found along the Pacific coast of North America, including California, Oregon, Washington, British Columbia, and southern Alaska (Price, 1990). Taxus globosa Schlecht. occurs from northeastern Mexico to Guatemala and El Salvador (Hartzell, 1991).

In the Old World, Taxus cuspidata Sieb. et Zucc. is found in Japan, northern China, and eastern Russia, and T. baccata is widely distributed in Europe and North Africa. In Southeast Asia, T. wallichiana Zucc. and T. fuana Li et Fu occur in the eastern Himalayas, and T. chinensis (Pilger) Rehder (= T. sumatrana) and T. mairei Lemee et Lev. are found in central and southern China; the former species (T. chinensis) extends south of the equator to Malaysia.

The taxonomy of Taxus has been difficult and controversial, partly because there are few reliable morphological characters for diagnosing species. Pilger (1903) treated all previously described Taxus species as subspecies of T.
baccata. This treatment has been adopted by a few authors (Voliotis, 1986; Bolsinger and Jaramillo, 1990), but 10 species are more often accepted (Silba, 1984; Krüssmann, 1985; Welch and Haddow, 1993; Li and Fu, 1997; Farjon, 1998).

Phylogenetic relationships among species of Taxus have been implied by various taxonomic treatments but have never been explicitly tested. Both Pilger (1903) and Schneider (1913) treated T. chinensis as a variety of T. cuspidata, implying a close relationship between these two species. Henry (1906), on the other hand, regarded T. chinensis as a variety of T. baccata, whereas Florin (1948) considered it a variety of T. wallichiana. Taxus mairei has been recognized as a variety of either T. chinensis (Cheng and Fu, 1978) or T. wallichiana (Li and Fu, 1997). Silba (1984) treated T. floridana as a variety of T. canadensis.

Our objectives were to infer phylogenetic relationships within Taxus using sequences of the internal transcribed spacer region of nuclear ribosomal DNA (see Baldwin et al., 1995; Gerandt and Liston, 1999; Kim and Kim, 1999; Li et al., 2000) and to consider the biogeographic implications of these relationships.

**MATERIALS AND METHODS**

**Taxa**

Fifteen samples were included in this study, representing an outgroup, Pseudotaxus chienii (Cheng) Cheng, and all species of Taxus except for T. fuana and T. wallichiana (Table 1). For most of the species, two or more individuals were sampled to check intraspecific variation.

**Molecular Technique**

DNAs were extracted from fresh leaf material using the DNeasy Kit (Qiagen, Santa Clarita, Calif.), following the manufacturer’s instructions. Polymerase Chain Reactions (PCR) were conducted to amplify the nrDNA ITS region in a Perkin-Elmer thermocycler using the primers ITS4 (White et al., 1990) and ITS-LEU (see Baum et al., 1998). Each 25-μL reaction contained 2.5 μl of Taq polymerase reaction buffer, 4 μl of dNTP, 0.8 μl of 50 mM MgCl₂, 1 μl of each of the two primers (10 μM), 50–100 ng of genomic DNA, 2 μl of
Table 1. Taxa included in this study.

<table>
<thead>
<tr>
<th>SPECIES</th>
<th>AA#</th>
<th>SOURCE</th>
<th>GenBank Accession #</th>
</tr>
</thead>
<tbody>
<tr>
<td>Taxus baccata L.</td>
<td>546–89C</td>
<td>Weilheim-Schongan, Germany</td>
<td>AF259287</td>
</tr>
<tr>
<td>T. baccata</td>
<td>372–35A</td>
<td>Mt. Vitosha, Bulgaria</td>
<td>AF259294</td>
</tr>
<tr>
<td>T. brevifolia Nuttall</td>
<td>1087–88A</td>
<td>Marion, Oregon, USA</td>
<td>AF259295</td>
</tr>
<tr>
<td>T. canadensis Marshall</td>
<td>601–58A</td>
<td>Ontario, Canada</td>
<td>AF259296</td>
</tr>
<tr>
<td>T. canadensis</td>
<td>1539–82C</td>
<td>West Roxbury, Massachusetts, U.S.A.</td>
<td>AF259298</td>
</tr>
<tr>
<td>T. chinensis (Pilger) Rehder</td>
<td>1395–60A</td>
<td>Hunnewell Pinetum, Massachusetts, U.S.A.</td>
<td>AF259292</td>
</tr>
<tr>
<td>T. chinensis var. mairei (Lemee et Levi.) Cheng et Fu</td>
<td>333–81A</td>
<td>Hangzhou Botanical Garden, China (cultivated)</td>
<td>AF259291</td>
</tr>
<tr>
<td>T. cuspidata Sieb. et. Zucc</td>
<td>17404C</td>
<td>Japan</td>
<td>AF259289</td>
</tr>
<tr>
<td>T. cuspidata</td>
<td>17404D</td>
<td>Japan</td>
<td>AF259290</td>
</tr>
<tr>
<td>T. floridana Nuttall ex Chapman</td>
<td>1178–89A</td>
<td>Liberty County, Florida, U.S.A.</td>
<td>AF259286</td>
</tr>
<tr>
<td>T. globosa Schlechtendal</td>
<td>905–90A</td>
<td>Nuevo León, Mexico</td>
<td>AF259299</td>
</tr>
<tr>
<td>T. x hunnewelliana Rehder</td>
<td>10760B</td>
<td>Hunnewell Pinetum, Massachusetts, U.S.A.</td>
<td>AF259288</td>
</tr>
<tr>
<td>T. x media Rehder</td>
<td>11267–A</td>
<td>Hunnewell Pinetum, Massachusetts, U.S.A.</td>
<td>AF259297</td>
</tr>
<tr>
<td>T. wallichiana Zucc.</td>
<td>141–88A</td>
<td>Pinetum, Blijdenst (cultivated)</td>
<td>AF259293</td>
</tr>
<tr>
<td>Pseudotaxus chienii (Cheng) Cheng</td>
<td>401–95</td>
<td>Central China</td>
<td>AF259300</td>
</tr>
</tbody>
</table>

**Note:** AA stands for the Arnold Arboretum collections. Voucher specimens are deposited in A.

DMSO, 1 unit of Taq polymerase, and sterilized distilled water. The thermocycler program included 94°C for 2 minutes, 30 cycles of PCR reactions, each with a 30-second denaturation at 94°C, 60-second annealing at 50°C, and 90-second extension at 72°C. An additional 7-minute extension was performed after the PCR cycles. We conducted electrophoresis of the PCR products in a 1% agarose gel to check for the presence of a single band. The Qiagen PCR purification kit was used to clean PCR products (Santa Clarita, Calif.) for direct cycle sequencing, as described elsewhere (Li et al., 1999).

**Sequence Editing and Phylogenetic Analysis**

For each sample, the ITS region was sequenced using three primers, ITS4, ITS-LEU, and 5.8Sgy (Liston et al., 1996). Sequences were read using an ABI 377 Automated Sequencer (Applied Biosystems, La Jolla, Calif.) and were then edited using Sequencher 3.0 (Gene Codes Corp., Ann Arbor, Mich.). The edited sequences were imported into PAUP* (version 4.0b4a; Swofford, 2000) and were aligned by eye. The delimitation of ITS-1, 5.8S, and ITS-2 was determined by comparing published sequences in the GenBank (#U77958). *Pseudotaxus chienii*, which appears to be the sister species of *Taxus* (Cheng et al., 2000; Wang and Shu, 2000), was included in parsimony analyses for rooting purposes. All characters were weighted equally, and character states were treated as unordered. Gaps were treated as the fifth character state. Parsimony analyses were carried out using heuristic tree searches in PAUP*; search options included TBR branch swapping, simple sequence addition, MULPARS on, and steepest descent off. To evaluate branch support, we conducted 100 replicates of bootstrap analysis (Felsenstein, 1985) and obtained decay indices (Bremer, 1988; Donoghue et al., 1992).

**Estimate of Divergence Time**

Maximum likelihood ratio (MLR) tests were conducted using PAUP* following Baum et al.
(1998), to determine whether a molecular clock could be rejected. For these analyses we used a reduced dataset, with just one sequence for each species, so as to minimize the occurrence of polytomies. Several models were applied, including F81, F84, and HKY (see Baum et al., 1998). The fossil record of *Taxus* and *Pseudotoxus* allowed us to calibrate a base substitution rate. Fossils of *Taxus* date back to the middle Jurassic, and *Pseudotoxus* has been recorded from the late Cretaceous (Florin, 1963). We presume that the two lineages diverged at least by the middle Jurassic (~165 mya).

**RESULTS**

**Sequence Characteristics**

We were able to obtain sequences of the entire ITS region of *Taxus* and *Pseudotoxus* using the three sequencing primers. The sequences have been submitted to GenBank (Table 1) and the data matrix and trees are available in TreeBASE (http://www.harvard.edu/treebase). The length of the ITS region ranged from 1127 to 1129 base pairs (bp) in *Taxus* and was 1108 bp in *Pseudotoxus*. The length of 5.8S in all sequences was 145 bp, and that of ITS-2 was 205 bp in all species except *T. canadensis*, which was one bp shorter. Most length variation of the ITS region was found in ITS-1, which was from 777 to 779 bp long in *Taxus* and 760 bp in *Pseudotoxus*. The alignment resulted in a data set of 1147 characters, including 29 gaps, 26 of which were between *Taxus* and *Pseudotoxus*. Of the 1147 sites, 252 were variable and 37 were parsimony informative. Sequence divergence was about 19% between *Taxus* and *Pseudotoxus*. Among samples of *Taxus*, sequence divergences were much lower, ranging from 0% to 1.87%, with an average of 0.97% (Table 2). There was little intraspecific variation in the ITS sequences (<0.09%). The G+C content averaged 59% over all of the species, with little variance among sequences; the average G+C content was 51% in 5.8S, 59% in ITS-2, and 61% in ITS-1.

**Phylogenetic Relationships**

Parsimony analysis generated four trees of 307 steps (CI = 0.96; RI = 0.81); Fig. 2 shows one of the four trees. The other three trees differed from Fig. 2 in either the unresolved E clade or the partially resolved H clade, where *T. x media* was sister to a polytomy of other three accessions of *T. baccata*. Three North American species form a well-supported clade (labeled B in Fig. 2; bootstrap = 85%, decay = 2), with *T. brevifolia* sister to a C clade containing *T. floridana* and *T. globosa*. The D clade, containing the Old World species and *T. canadensis* of the New World, is weakly supported (bootstrap = 53%, decay = 1). Relationships among three lineages within the D clade are unresolved in the strict consensus tree of the four trees. *Taxus chinensis* and *T. mairei* form a clade (F) with weak support (bootstrap = 68%, decay = 1). The two samples of *T. canadensis* are strongly united (bootstrap = 98%, decay = 4). The G clade, containing *T. cuspidata* and *T. baccata* (H; bootstrap = 94%, decay = 2) is weakly supported (bootstrap = 60%, decay = 1).

**Estimated Divergence Time**

MLR tests using different evolutionary models did not reject the null hypothesis of clock-like evolution (all P > 0.2). On the basis of sequence divergence (ca. 19%) and fossil evidence for the divergence of *Taxus* and *Pseudotoxus* by at least 165 mya, we estimated a rate of change of 5.76 × 10^{-10} base substitutions per site per year. Based on this rate, the estimated times of divergence for the clades were 9.2–16.15 (A), 6.16–6.94 (B), 0.78 (C), 4.6–10.76 (D), 6.94–10.76 (E), 7.7 (F), and 3.13–4.6 (G) mya, respectively (Fig. 2).

**DISCUSSION**

**Sequence Length Variation**

In angiosperms, the ITS region, including the 5.8S gene, is often about 700 bp long (Baldwin et al., 1995). In other seed plants, in contrast, the ITS region varies greatly, from about 1000 bp to several thousand bp (Liston et al., 1996, 1999). Nevertheless, as in some conifers (Liston et al., 1996; Klein et al., unpubl.), the length of the ITS region within *Taxus* varies only within a small range—from 1127 to 1129 bp. In conifers, the ITS-1 region is several times longer than the ITS-2 region (Klein and Li, unpubl.); in *Taxus*, ITS-1 sequences are about four times longer than ITS-2.

**Phylogenetic Relationships**

Interspecific relationships within *Taxus* have been implied in various taxonomic treatments. Silba (1984) treated *T. floridana* as a variety of *T. canadensis*, suggesting a close relationship between these two species. In the ITS phylogenetic tree, however, *T. canadensis* did not fall into the B clade; instead it is weakly allied with the Old World species (Fig. 2). A similar
<table>
<thead>
<tr>
<th>SPECIES</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>12</th>
<th>13</th>
<th>14</th>
<th>15</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Taxus baccata 546–89C</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2. T. baccata 372–35A</td>
<td>0.09</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3. T. brevifolia 1087–88A</td>
<td>1.42</td>
<td>1.51</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4. T. canadensis 601–58A</td>
<td>0.89</td>
<td>0.98</td>
<td>1.42</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5. T. canadensis 1539–82C</td>
<td>0.98</td>
<td>1.07</td>
<td>1.51</td>
<td>0.09</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6. T. chinensis 1395–60A</td>
<td>1.15</td>
<td>1.24</td>
<td>1.69</td>
<td>1.15</td>
<td>1.24</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7. T. mairei 333–81A</td>
<td>1.15</td>
<td>1.24</td>
<td>1.51</td>
<td>1.15</td>
<td>1.24</td>
<td>0.89</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8. T. cuspidata 17404C</td>
<td>0.44</td>
<td>0.53</td>
<td>1.15</td>
<td>0.62</td>
<td>0.71</td>
<td>0.89</td>
<td>0.89</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9. T. cuspidata 17404D</td>
<td>0.36</td>
<td>0.44</td>
<td>1.06</td>
<td>0.53</td>
<td>0.62</td>
<td>0.80</td>
<td>0.80</td>
<td>0.09</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10. T. floridana 1178–89A</td>
<td>1.60</td>
<td>1.69</td>
<td>0.71</td>
<td>1.60</td>
<td>1.69</td>
<td>1.86</td>
<td>1.69</td>
<td>1.33</td>
<td>1.24</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>11. T. globosa 905–90A</td>
<td>1.60</td>
<td>1.69</td>
<td>0.80</td>
<td>1.60</td>
<td>1.69</td>
<td>1.87</td>
<td>1.69</td>
<td>1.33</td>
<td>1.24</td>
<td>0.09</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>12. T. hunnewelliana 10760B</td>
<td>0.53</td>
<td>0.62</td>
<td>1.24</td>
<td>0.71</td>
<td>0.80</td>
<td>0.98</td>
<td>0.98</td>
<td>0.27</td>
<td>0.18</td>
<td>1.42</td>
<td>1.42</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>13. T. media 11267-A</td>
<td>0.00</td>
<td>0.09</td>
<td>1.24</td>
<td>0.71</td>
<td>0.80</td>
<td>0.98</td>
<td>0.98</td>
<td>0.27</td>
<td>0.18</td>
<td>1.42</td>
<td>1.42</td>
<td>0.36</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>14. T. baccata 141–81A</td>
<td>0.00</td>
<td>0.09</td>
<td>1.42</td>
<td>0.89</td>
<td>0.98</td>
<td>1.15</td>
<td>1.15</td>
<td>0.44</td>
<td>0.36</td>
<td>1.60</td>
<td>1.60</td>
<td>0.53</td>
<td>0.00</td>
<td></td>
<td></td>
</tr>
<tr>
<td>15. Pseudotaxus chienii 401–95</td>
<td>19.00</td>
<td>18.91</td>
<td>19.01</td>
<td>18.81</td>
<td>18.82</td>
<td>18.91</td>
<td>18.81</td>
<td>18.63</td>
<td>18.72</td>
<td>18.91</td>
<td>18.88</td>
<td>18.90</td>
<td>18.83</td>
<td>19.00</td>
<td></td>
</tr>
</tbody>
</table>

**NOTE:** Species names are followed by the accession numbers of the Arnold Arboretum
relationship has been recently suggested by a study of stomatal structure in *Taxus* (Spjut, 2000). Forcing *T. canadensis* into a clade with *T. floridana* requires two extra steps. Additional data are needed to test this novel hypothesis.

*Taxus chinensis* has been treated as a variety of either *T. wallichiana* (Florin, 1948), *T. baccata* (Henry, 1906), or *T. cuspidata* (Pilger, 1903; Schneider, 1913). In our ITS trees, *T. baccata* and *T. cuspidata* form a weakly supported clade (G), with *T. chinensis* more distantly related.

*Taxus mairei* has been considered a variety of either *T. chinensis* (Cheng and Fu, 1978) or *T. wallichiana* (Li and Fu, 1997). In our ITS trees, *T. mairei* and *T. chinensis* form a clade. *Taxus wallichiana* was not available for this study.

Rehder (1925) described *Taxus × hunnewelliana* as a probable hybrid between *T. canadensis* and *T. canadensis*.
and *T. cuspidata*. In this study we sampled one of the living specimens Rehder used in the original description of this putative hybrid (Arnold Arboretum accession #10760B). We did not find polymorphic sites indicative of hybridization between *T. canadensis* and *T. cuspidata*, and this sample clustered with *T. cuspidata*, separate from *T. canadensis*.

**Biogeographic Implications**

We found low within-species sequence divergences (<0.09%), and with only one exception, divergences between species are at least twice as great as within species (Table 2). Low sequence divergence between *T. floridana*, from Florida, and *T. globosa*, from Mexico and Central America, suggests very recent separation between the lineages in these regions and is consistent with treating these populations as belonging to the same species. Disjunctions between eastern North America and eastern Mexico are well known (e.g., *Liquidambar* and *Viburnum*; see Burnham and Graham, 1999), and in at least some of these cases, levels of sequence divergence are also low (e.g., *Hamamelis* [Li et al., 2000] and *Cercis* [Davis et al., unpubl. data]). Further study of such cases may indicate recent vicariance events simultaneously affecting a number of lineages.

Curiously, *Taxus floridana* and *T. globosa* appear not to be directly related to the eastern North American species, *T. canadensis*. Instead, these species are united quite strongly with the western North American *T. brevifolia*. If the rooting of our trees is correct, the four New World species do not form a clade, and *T. canadensis* is actually more closely related to Old World species. Unfortunately, the exact relationships of *T. canadensis* remain unresolved. It is possible, on the basis of our analyses, that *T. canadensis* is most closely related to Asian species. In contrast, a direct connection to the European species, *T. baccata*, seems less likely.

*Taxus* has a rich and deep fossil history (Florin, 1963). If our age estimates are at all accurate, the modern species of *Taxus* stem from a rather recent common ancestor. In view of the apparent great age of *Taxus*, our results suggest either a long period of stasis prior to the radiation that resulted in the modern species, or perhaps earlier periods of diversification followed by the extinction of lineages. Until *Taxus* fossils can be incorporated in phylogenetic analyses, their bearings on these hypotheses will be difficult to interpret.

**LITERATURE CITED**


