Analyses of ribosomal DNA internal transcribed spacer sequences from *Juglans nigra* and leaf-associated fungi in Zoar Valley, NY

by

Vincent Ragozine

Youngstown State University

Submitted in Partial Fulfillment of the Requirements

for the Degree of

Masters of Science
Analyses of ribosomal DNA internal transcribed spacer sequences from Juglans nigra and leaf-associated fungi in Zoar Valley, NY

Vincent K. Ragozine

I hereby release this thesis to the public. I understand that this thesis will be made available from the OhioLINK ETD Center and the Maag Library Circulation Desk for public access. I also authorize the University or other individuals to make copies of this thesis as needed for scholarly research.

Signature:

____________________________________________________________
Vincent K. Ragozine, Student     Date

Approvals:

____________________________________________________________
Dr. Carl Johnston, Thesis Advisor     Date

____________________________________________________________
Dr. Thomas Diggins, Committee Member     Date

____________________________________________________________
Dr. Heather Lorimer, Committee Member     Date

____________________________________________________________
Peter J. Kasvinsky, Dean of School of Graduate Studies and Research     Date
Abstract: A genetic analysis of samples of *Juglans nigra* (Black Walnut) from three different locations in Zoar Valley, New York, was conducted. Nuclear ribosomal DNA (nrDNA) intragenic spacer regions (ITS1, ITS2) were PCR amplified along with the 5.8S ribosomal RNA (rRNA) gene. Consensus sequence alignment of *J. nigra* DNA samples from Zoar Valley showed sequence variation (both base additions and substitutions) between samples. It is likely that at least some of the base substitutions in the consensus sequence are not an artifact of the method, and are different from the published sequence for *J. nigra*. This indicates the method has potential for examining within species variation for different populations of *J. nigra*.

A survey of fungi associated with the phyllosphere of two elm species native to Zoar Valley, NY, *Ulmus americana* and *Ulmus rubra*, was conducted on samples recovered from Zoar Valley, NY. Fungi were identified by sequencing cloned DNA of PCR amplified ribosomal DNA (rDNA) extracted from leaf tissue. Probable endophytes were identified (*Phoma, Coprinellus*), but the majority of fungi detected (*Cryptococcus, Ampelomyces, Colletotrichum*) were most likely parasites. Multiple genera of fungi were detected in single leaf tissue samples.
**Introduction:** DNA was isolated and sequenced from leaf samples collected in old growth forest in Zoar Valley, in hopes of forming a preliminary colonial history of *Juglans nigra* in Zoar Valley, and detection of specific hereditary variations in sequence within populations. This combined thesis was a byproduct of the universal eukaryotic ribosomal DNA primers (ITS1 and ITS4) used in this study. The original intention was to isolate and PCR replicate plant DNA only for further downstream application. As research methods progressed to DNA sequencing, it became apparent after the completion of numerous sequencing reactions, that fungal DNA was being PCR amplified and cloned along with plant DNA. Fungi from either within the leaves or upon their surface was being DNA extracted along with the leaves themselves, and was present in the DNA extracts to be PCR amplified.

Interestingly, only fungal DNA was extracted and PCR amplified from the leaves of *Ulmus americana* (American elm) and *Ulmus rubra* (Slippery elm), while no plant sequences were obtained for the elms. Conversely, no fungal DNA was sequenced from the *Juglans nigra* (Black walnut) leaf samples. PCR reactions performed on elm-leaf DNA extracts favored amplification of fungal DNA, rather than elm DNA. It was decided to use the fungal sequences we obtained in a second project.
Table of Contents:

Chapter I: Identification and comparison of ribosomal DNA internal transcribed spacer sequences (rDNA ITS) isolated from *Juglans nigra* leaf tissue collected from Zoar Valley, NY

I. List of Figures, Tables, and Appendices vi
II. Background 1
III. Methods 6
IV. Results and Discussion 8
V. Conclusion 21
VI. Future Research Implications 22
VII. References 23

Chapter II: A molecular identification of fungal ribosomal DNA isolated from *Ulmus americana* and *Ulmus rubra* leaf tissue samples from Zoar Valley, NY using ribosomal DNA internal transcribed spacer sequences (rDNA ITS)

I. List of Figures and Tables vi
II. Background 26
III. Methods 29
IV. Results and Discussion 30
V. Conclusion 37
VII. References 37
### Chapter I:

I. Table 1  
   Page: 8

II. Table II  
   Page: 10

III. Table III  
   Page: 11

IV. Table IV  
   Page: 13

V. Table V  
   Page: 15

VI. Table VI  
   Page: 16

VII. Table VII  
   Page: 17

VIII. Figure I  
   Page: 3

IX. Figure II  
   Page: 4

X. Figure III  
   Page: 7

XI. Figure IV  
   Page: 12

XII. Figure V  
   Page: 12

XIII. Figure VI  
   Page: 14

XIV. Figure VII  
   Page: 14

XV. Figure VIII  
   Page: 15

XVI. Figure IX  
   Page: 16

XVII. Figure X  
   Page: 17

XVIII. Figure XI  
   Page: 17

XIX. Figure XII  
   Page: 19

XX. Appendix I (Methods)  
   Page: 40

XXI. Appendix II  
   Page: 45

XXII. Appendix III  
   Page: 57

XXIII. Appendix IV  
   Page: 59

XXIV. Appendix V  
   Page: 62

XXV. Appendix VI  
   Page: 65

XXVI. Appendix VII  
   Page: 68

XXVII. Appendix VIII  
   Page: 71

### Chapter II:

I. Table I  
   Page: 30

II. Table II  
   Page: 32
Title: Identification and comparison of ribosomal DNA internal transcribed spacer sequences (rDNA ITS) isolated from *Juglans nigra* leaf tissue collected from Zoar Valley, NY

**Background:** The genomes of all living organisms contain certain genes responsible for fundamental biochemical functions. These genes can be sequenced, aligned, and analyzed to study phylogenetic relationships, even among morphologically indistinguishable but otherwise distinct species (Hillis and Moritz 1990). This level of genetic variability is central to population biology because the amount of variability directly influences the evolutionary potential for a species or populations (Shaal and Learn 1988).

Nuclear ribosomal DNA (nrDNA) is a strong tool because it is ubiquitous in all organisms. Also, it is in relatively high copy number to chromosomal DNA, making it more accessible. Ribosomal DNA has been used to systemically evaluate and construct genetic histories, known as phylogenies, through many taxonomically diverse groups of plants, fungi, and animals. The nrDNA can be viewed as units, the 18S, 5.8S, and 26S ribosomal genes, and the two internal transcribed spacers, ITS1 and ITS2, located between the 18S and 26S coding regions. The ribosomal genes have a much slower rate of sequence change than do the internal transcribed spacer regions (Suh et al. 1993).

Ribosomal DNA has been very useful in the study of plant evolutionary biology. The DNA conserved sequences that code for the 18S, 5.8S, and 26S ribosomal subunits have provided information on phylogenetic relationships among the species within a genus, and have also illuminated higher level relationships (Shaal and Learn 1988).
intergenic spacer (ITS) regions of ribosomal DNA are highly variable, with variation occurring within populations and in individuals of the same population (Shaal and Learn 1988).

The internal transcribed spacers (ITS1 and ITS2) and 5.8S rRNA gene of nuclear ribosomal DNA were sequenced and analyzed to determine genetic heredity in the angiosperm species *Juglans nigra*. The ITS1-5.8S-ITS2 stretch of eukaryotic nrDNA (each <300bp) can be readily amplified by PCR and sequenced using universal primers (Baldwin et al 1995) (see Figure 1). The ITS region is known to undergo rapid concerted evolution (Linder et al. 2000). Differences in ITS sequence between species can mostly be attributed to point mutations acquired over evolutionary time (Baldwin et al 1995). The ITS sequences have proven to be useful determining taxonomic relationships among many species of angiosperms, and species can be readily distinguished through sequence variation (Jobes and Thien 1997, Baldwin et al 1995).
Previous studies (Gonzalez-Lamothe et al. 2002, Potter et al. 2001) have successfully utilized ribosomal DNA sequence, specifically the ITS1 and ITS2 spacer regions, to answer systematic questions and determine phylogenetic heredity. For example the 18S, 28S ribosomal genes, as well as the ITS1 – 5.8S- ITS2 region of the fungus Spilocaea oleagina (responsible for peacock leaf spot disease in olive trees) was sequenced to determine its identity. The phylogenetic classification of S. oleagina was previously undecided among mycologists. It was determined from nrDNA sequence data that S. oleagina belongs to the class Dothideomycetes, and is an anamorphic phase of a yet unidentified Venturia species (Gonzalez-Lamothe et al. 2002). In a plant example, ITS spacer sequences from the plant Arabidopsis suecica were PCR-amplified, cloned and sequenced, confirming that this allopolyploid species contains two distinct types of ITS sequences, one from A. thaliana and the other from C. arenosa, confirming they are the putative parents of A. suecica (O’Kane et al. 1996). In another landmark study, Paradox, a hybrid walnut cultivar [J. regia and J. californica] important to the California walnut industry, was analyzed for gene flow from other North American walnut species.
[J. major, J. hindsii and J nigra]. ITS data showed that among various walnut industry sources of the Paradox strain, there was considerable genetic contribution from all North American species in at least some of these samples (Potter et al. 2001).

New York State’s Zoar Valley was chosen for sampling due to its intact riparian ecosystem and healthy distribution of J. nigra (Black Walnut). The woodlands of Zoar Valley are highly varied and diverse, and collectively meet all objective criteria for eastern old growth forest (Diggins and Kershner 2005). Zoar Valley is the most intact forest-gorge landscape in the western New York region, having the largest area of virgin and secondary old growth forest, and contiguous climax forest. For these reasons and others, Zoar Valley has been studied scientifically and has become the subject of multiple conservation efforts.
B.

**Figure 2: Location** of Zoar Valley in Western New York (A) and sample study areas within (B).

*Juglans nigra* [Black walnut, a dicotyledonous tree species of the family *Juglandaceae*] is an important tree economically, for both its edible nut and in its use in commercial wood production (Stanford et al. 2000). Black walnut was chosen as the focus of this study because it is well established inside Zoar Valley, both in old growth stands and in younger forest. Minimal *J. nigra* population structure exists outside Zoar Valley due to heavy logging (Diggins and Kershner 2005). This made *J. nigra* an ideal species to analyze for genetic variation within an isolated population since it is safe to assume minimal gene flow between *J. nigra* populations inside of Zoar Valley with populations outside.

Within Zoar Valley, three sites were strategically selected for sampling. The first two sites, South Branch Floodplain and Lookout Point Terrace, lie within the undisturbed forested valley, and are separated geographically by about 1 kilometer. The third site, Valentine’s Flats Plantation, lies along the western rim of Zoar Valley within a kilometer of South Branch Floodplain and Lookout Point Terrace and consists of Black Walnut trees planted by the State of New York in the late 1960s.
Genetic analysis of *J. nigra* between Zoar Valley’s old growth stands, developing forests, and Valentine Flats Plantation was performed using gene sequence alignment. Using ITS sequence analysis enabled the comparison of multi-aged, spatially distant tree stands throughout Zoar Valley. This knowledge was applied to determine a preliminary colonial history of the *J. nigra* population in Zoar Valley, expose genetic identity or variation inside of each distinct population, and link trees in young forest with their parent trees in old growth if possible. In all, some insight into the population dynamics of this species in Zoar Valley was provided, yielding a glimpse into the history, heredity, and future implications for these species.

**Methods:** Samples were obtained from Zoar Valley on October 4th, 2005. Samples were collected from three distinct areas of Zoar Valley, Lookout Point Terrace, South Branch Floodplain, and Valentine Flats Plantation. Samples of *J. nigra* leaf tissue were first identified morphologically, then collected and stored in deep freeze at -80 C° until DNA extraction was performed.

**DNA extraction:** Total DNA was extracted using a revised CTAB method of Doyle and Doyle (1987).

**PCR:** Extracted DNA was amplified with ITS specific primers. Forward primer used was **ITS1** [TCCGTAGGTGAACCTGCGG] and the reverse primer used was **ITS4** [TCCTCCGCTTATTGATATGC] of White et al. (1990). PCR performed on PTC-200 DNA engine.
**Cloning:** Purified PCR products were cloned using the TOPO TA Cloning Kit from Invitrogen (cat# K4530-20) or StrataClone PCR cloning kit from Stratagene (cat#240205)

**Restriction Digest:** A restriction digest using EcoR1 was used to screen clones before sequencing.

![EcoR1 digest revealing restriction fragments of approximately 700bp, indicative of the PCR product generated using the universal primers ITS1 and ITS4 of White et al. (1990)](image)

**Figure 3:** EcoR1 digest revealing restriction fragments of approximately 700bp, indicative of the PCR product generated using the universal primers ITS1 and ITS4 of White et al. (1990)

**DNA Sequencing:** Sequencing reactions were carried out on purified cloning extracts using Beckman Coulter Sequencing Kit, sequenced using primers M13 (-20, -47) Forward and M13 Reverse, with sites provided in the cloning vector.

**Alignment:** Sequences were aligned to determine identity or variation in ITS region of each species using computer programs. NCBI nucleotide BLAST (http://www.ncbi.nlm.nih.gov/BLAST/) was used to identify sequenced DNA and
ClustalW (http://www.align.genome.jp/sit-bin/clustalw) was used to construct a multiple alignment of the sequences.

**Results and Discussion:** Identity of the sequenced DNA was assessed by sequence comparison within NCBI Genbank.

**Table 1.** List of single run DNA sequences from trees morphologically identified as *J. nigra*

<table>
<thead>
<tr>
<th>Sample # and primer set used (Forward or Reverse)</th>
<th>Collection Site within Zoar Valley</th>
<th>Closest match in NCBI Nucleotide Blast Search</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 (Forward)*</td>
<td>South Branch Floodplain</td>
<td><em>Juglans nigra</em></td>
</tr>
<tr>
<td>1 (Reverse)*</td>
<td>South Branch Floodplain</td>
<td><em>Juglans nigra</em></td>
</tr>
<tr>
<td>1 (Forward) (-47)**</td>
<td>South Branch Floodplain</td>
<td><em>Juglans nigra</em></td>
</tr>
<tr>
<td>3 (Forward)*</td>
<td>Lookout Point Terrace</td>
<td><em>Juglans major</em></td>
</tr>
<tr>
<td>3 (Reverse)*</td>
<td>Lookout Point Terrace</td>
<td><em>Juglans nigra</em></td>
</tr>
<tr>
<td>5 (Forward) (-47)**</td>
<td>Valentines Flat Plantation</td>
<td><em>Juglans microcarpa</em></td>
</tr>
<tr>
<td>5 (Reverse)*</td>
<td>Valentines Flat Plantation</td>
<td><em>Juglans nigra</em></td>
</tr>
<tr>
<td>11 (Forward)*</td>
<td>Lookout Point Terrace</td>
<td><em>Juglans major</em></td>
</tr>
<tr>
<td>11 (Reverse)*</td>
<td>Lookout Point Terrace</td>
<td><em>Juglans nigra</em></td>
</tr>
<tr>
<td>15 (Forward)*</td>
<td>South Branch Floodplain</td>
<td><em>Juglans nigra</em></td>
</tr>
<tr>
<td>15 (Reverse)*</td>
<td>South Branch Floodplain</td>
<td><em>Juglans nigra</em></td>
</tr>
</tbody>
</table>

* Forward or reverse primer supplied by Invitrogen cloning kit (cat# K4530-20)

**Sequence Identification of Single Run Forward and Reverse Reactions:** Results from the Genbank sequence identity comparison (Table 1) showed that all single run sequence samples belonged to the *Juglans* genera. Samples #3, 5, and 11 showed variation within
the same sample, identifying better with other *Juglans* sp. than *J. nigra* when sequenced using the forward primer. However when sequenced with the reverse primer, all were most identical to *J. nigra* sequences in NCBI Genbank.

A total of five plasmids, from five different trees, were successfully sequenced with the forward and reverse primers. The inconsistent species identity observed when comparing the forward and reverse sequence of samples #3, 5, and 11 cannot be interpreted as variation within the same sample, since both the forward and reverse sequencing reactions were run from plasmid DNA extracted from the same clone.

It has been shown that reforestation by planting within a species’ native range is an example of human mediated gene transport, and if trees in off-site plantations cross with those in native populations, diversity may increase in the next generation, although with negative consequences for local adaptation (Ledig 1992). The possible transport of non-native genes occurred with the planting of Valentine’s Flats Plantation, and this may have further increased any genetic diversity in the adjacent *J. nigra* populations in Zoar Valley. This explanation, again, however unlikely (due to the relatively short time period between the planting of Valentine’s flats and the amount of genetic variation observed) could account for the variation observed in the Valentines Flats and Lookout Point samples. *Juglans* species are renowned for their ability to form hybrids (Potter et al. 2001).

Referring to Table 1, if *J. microcarpa* (5 –Forward (-47)) or *J. major* (3 –Forward and 11 –Forward), or perhaps a hybrid of these two species was introduced to Zoar
Valley in the past, then it is possible that these genes were introduced into the Zoar Valley *J. nigra* population. If this happened, hybridized walnut trees were then created. They might be morphologically identical to the native *J. nigra* population, however revealing their true identity only in DNA analysis. (*J. microcarpa* and *J. major* are native to the southwestern U.S.). This possibility of ITS hybrids allows for the discrepancy observed between the forward and reverse primers. For further study of this observed variation, we must construct consensus sequences from all sequencing reactions for each sample, and then compare these consensus sequences with others in the sample set and with those published in NCBI GENBANK. The consensus sequences help to rule out error as a cause of variation.

**Assembling and Analyzing the Consensus Sequence:** Consensus sequences were constructed from the multiple sequence alignment of all single runs for each sample. These consensus sequences were constructed to represent the most highly observed sequence configuration, after alignment and comparison, for all sequenced plasmids. The consensus sequence was then subjected to an NCBI BLAST assay and the results were as follows:

<table>
<thead>
<tr>
<th>Tree DNA Sample</th>
<th># of Sequencing Runs</th>
<th>Consensus sequence constructed</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>3</td>
<td>yes</td>
</tr>
<tr>
<td>3</td>
<td>2</td>
<td>Yes</td>
</tr>
<tr>
<td>5</td>
<td>2</td>
<td>Yes</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>---</td>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td>11</td>
<td>2</td>
<td>Yes</td>
</tr>
<tr>
<td>15</td>
<td>2</td>
<td>No</td>
</tr>
</tbody>
</table>

Consensus sequence nucleotides are listed either in capital letters (“A”) or lowercase letters bracketed by parenthesis (“(a/t)”). Capital letters stand for undisputed consensus nucleotides, while lowercase letters are found where there is an inconsistency between the single run sequences, with both nucleotides listed.

**Table 3. Consensus Sequence and NCBI Blast best match of sample 1, with highest identity segment of overlap from Clustalw multiple alignment shown below**

**Sample 1 Consensus Sequence (799 bp)**

| CAACTTCGCCCTTTCCGTAGGTGAACCTGCAGGATCAATTGTGCACTTGAGGTGAACCTGCGGAACATGTAATAAATACCCCTCTCTTCTCCCAGAAACGGGAGGGGAGGAGCAACGTGACATTGGCGTTGCGTGTTGGTGTTCCTTCTACATTAACAGACTCTCGGCAACGGAGATATCCTGTCGCTCTCAGATCGATGAAAGAAGCTAGCGAATGCGGATTACCGTGCCCAGCCGTACAGAGATGAGGAAGGGCAGGATTCGCGGCCGCTAAATTCAATTCAGCCCTATAGTGAGTCGTATTACAATTCACTGGCGTA |

**NCBI BLAST Best Match:** Juglans nigra isolate 836 18S ribosomal RNA gene

**E Value:** 0.0
The sample 1 consensus sequence is 95% (512/538 bp) identical to *J. nigra*.

Three sequences were used to construct this consensus sequence, two forward and one reverse (Figure 4). This sequence was then assessed for matches in a BLAST assay (Figure 5).

The segment of the consensus sequence used in the BLAST assay shown in Figure 5 has also been highlighted in Table 3. Figure 5 shows five distinct differences between the consensus sequence and the GENBANK closest match. The differences
observed are all base additions to the consensus sequence. None of the single base
additions are conserved in all three of the single run sequences, making it unlikely that
this is real sequence variation (Figure 4). These extra bases may have been erroneously
added to the sample DNA sequence during the editing of the sequence chromatograms,
and therefore are a likely byproduct of this protocol.

Table 4. Consensus Sequence and NCBI Blast best match for sample 3, with highest identity segment
of overlap from ClustalW multiple alignment shown below

<table>
<thead>
<tr>
<th>Sample 3 Consensus Sequence (808 bp)</th>
</tr>
</thead>
<tbody>
<tr>
<td>GTGAACCTGCGGAAGGATCATTGTGATCACCCTGTGCCCAGCAG</td>
</tr>
<tr>
<td>AACGACCTGTGACACATGTAATAACCTTCTGGGTGGGGGTGT</td>
</tr>
<tr>
<td>AATGCCCCCTTCACCACAAAAACCGGTTGAGGGCAGCTTGAGA</td>
</tr>
<tr>
<td>TTTGCCCACCTTGCTCTCTGTTGTGGTGGTCTACTCTTGTT</td>
</tr>
<tr>
<td>CCCTTCCTGGATGAACAAAGGAACCCCGCCTGCTGCTGACCA</td>
</tr>
<tr>
<td>(a/g)GGAACTTAA(a/c)CA(a/g)GGAGGTAACCACGAGGGCGGCC</td>
</tr>
<tr>
<td>CCGGAACCGGTTGTCCTGTGTTGGACGGTGTGGTCCTTAC</td>
</tr>
<tr>
<td>CAAGATACATAACGGACTCTCTGGCCCAACGGATATCTCGTCTCTC</td>
</tr>
<tr>
<td>GCATCTGCTGAAAGAAGCGGGATAATCTGCTTGTGGTGTT</td>
</tr>
<tr>
<td>GAATTGCGAAATCTCGCGGACATCGAGTCTTTGTCAGCAG</td>
</tr>
<tr>
<td>GTTGCCGCCGGAAGCATTCCGGCAGGCGGACCGAGCCCTTGCTGGG</td>
</tr>
<tr>
<td>TGTCACGCATCTGGCCCAACCCCAACACTTCTTGCTGCTTGC</td>
</tr>
<tr>
<td>GC(g/c)(g/c)GGTGGCGGG(g/a)(g/a)A(g/a)(a/t)ACATTGCGCCTGCC</td>
</tr>
<tr>
<td>GCGCTTTTCTGCTCGGTAGCTCAATGAGTAGTCTCGCTACCA</td>
</tr>
<tr>
<td>GAGCGCCACACGACATCGGTTGAGGAGAACCCTCGTCGACCCG</td>
</tr>
<tr>
<td>TCGTGTGTCGCCCGCTGCTGTGAAGGTGCTCCTGACCCTATT</td>
</tr>
<tr>
<td>GTGTCTTCTCTGCGACTCTAACATCGGACCCAGTCAAGCGG</td>
</tr>
<tr>
<td>GGATTACCGCGCTGAATTTAAGCA TATCAATAGCGGAGGAAA</td>
</tr>
<tr>
<td>GGGCGAATTCGTATTAAAACCATGCAG</td>
</tr>
</tbody>
</table>

NCBI Best Match: Juglans nigra isolate 836 18S ribosomal RNA gene
E Value: 0.0
Identitiy: 95%
The sample 3 consensus sequence is 95% (717/748 bp) identical to *J. nigra*.

Two sequences were used to construct this consensus sequence, one forward and one reverse (Figure 6). This sequence was then assessed for matches in a BLAST assay (Figure 7).

The segment of the consensus sequence used in the BLAST assay shown in Figure 7 has also been highlighted in Table 4. Figure 7 shows one distinct difference between the consensus sequence and the GENBANK closest match. The variant observed in this segment is found in both the forward and reverse runs of this sequence, and is found in an otherwise highly conserved stretch of DNA sequence between the consensus sequence and the Genbank entry (Figure 6). This single nucleotide difference has the possibility of being a true variation. More sequences must be processed for sample 3 to determine if this is in fact real.
Table 5. Consensus Sequence and NCBI Blast best match for sample 5, with highest identity segment of overlap from Clustalw multiple alignment shown below

<table>
<thead>
<tr>
<th>Sample 5 Consensus Sequence (833 bp)</th>
</tr>
</thead>
<tbody>
<tr>
<td>TAGCGCACGTGGAATTGTAATACGACTCATAATAGGGTTTCGAATTGAATTATTAGCGGCCGC GAATTCGCCCTTTCCTCGCTTATTGATATGCTTAAATT CAGCGGGTAATCCCGCCTGAC CTGGGGTTCGAGATGAGTCCGAAG ACGACACAATAGGTCGAGGAGACCACTTCAACA GCGACGGGCGACACA CGACCGGTCACGAGGGTTTTCTCAACCACGATTGCTGCGCTCCGCTGCGCTG CCTAGGACTCTATTAGGCTAACCACGCAAGCAAGCCACGCGGAGGCC AATGTCTTCCCCCCGCAACAGCGTAAAGGATGTGTTGGGTTGGG GasCGAGTACCTGGACACCC AGGCAGACGTGCCCCTCGGGCCGAATGGCTTCTCG GGCACAACTTGCCTTCAAAGACTCGATGATTCCGGGATTGTCG A/g) ATTCA(a/c)(a/c)(a/c) C(a/c)A(g/a)GTATCGCATTTCGCTACGCTTCTCAGT GCAGGAAGCCGAAATATCCGTTCCCAGAGGTGTATGTATCATGG TAAAGACGTCGTACCAAGCAGACGGCCTGCTTCCGGGCGCCCAGT/a)GT (g/t) ACTCC(c/t)TGGTAAAGGTCTCCTGGCGAGACCACCGGC(c/g)GGGGCTACT TGTTCGATCGGGAAGGGAACGAGAGGATCGACCACCACACGAGGGCG AGGGGGCAGAATCTCCAACGTGC(c/t) TGGGA(a/g)GGGGCC(t/g)GTAC(a/c)CC CCAGGAA(g/a)GTATTATTACATGTCCTTCCGCTGCTG</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>NCBI Best Match: Juglans nigra isolate 834 18S ribosomal RNA gene</th>
</tr>
</thead>
<tbody>
<tr>
<td>E Value: 0.0</td>
</tr>
<tr>
<td>Identity: 96%</td>
</tr>
</tbody>
</table>

Figure 8. Segment of consensus sequence from Sample 5 multiple alignment of single runs

<table>
<thead>
<tr>
<th>5 Reverse</th>
<th>CGAATATCCGTTGGCCGAGAGTCGCTATGTATCATGGTAAAGACGTCACCAA 125</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>CGAATATCCGTTGGCCGAGAGTCGCTATGTATCATGGTAAAGACGTCACCAA 538</td>
</tr>
<tr>
<td></td>
<td>*********************************************************</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>5 Reverse</th>
<th>CGACACGCACACCGTTTCGCGGGCGCCCGGA 155</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>CGACACGCACACCGTTTCGCGGGCGCCCGGT 568</td>
</tr>
<tr>
<td></td>
<td>**********************************</td>
</tr>
</tbody>
</table>
The sample 5 consensus sequence is 96% (604/627 bp) identical to *J. nigra*.

Two sequences were used to construct this consensus sequence, one forward and one reverse (Figure 8). This sequence was then assessed for matches in a BLAST assay (Figure 9).

**Figure 9. Sample 5 NCBI BLAST result (partial)**

```
S5     497  CGAAATATCCGTTGCCGAGAGTCG G TTATGTATCATGGTAAAGACGTCACCAACGACACG  556
      ||| |||||||||||||||||||| |||||||||||||||||||||||||||||||||||
Gen    340  CGAATATCCGTTGCCGAGAGTCG - TTATGTATCATGGTAAAGACGTCACCAACGACACG  282
S5     557  CACACCGTTTCCGGGGCGCCCGT
      |||||||||||||||||||||||||
Gen    281  CACACCGTTTCCGGGGCGCCCGT
```

The segment of the consensus sequence used in the BLAST assay shown in Figure 9 has also been highlighted in Table 5. Figure 9 shows two distinct differences between the consensus sequence and the GENBANK closest match. One difference observed was a base addition to the consensus sequence. This single base addition was not conserved in both of the single run sequences, making it unlikely that this is real sequence variation (Figure 8). The other difference is a substitution of an “A” for a “G”. This substitution is observed in the reverse run but not in the forward, and therefore is likely an editing mistake or an incorrect call by the sequencer.

**Table 6. Consensus Sequence and NCBI Blast best match for sample 11, with highest identity segment of overlap from Clustalw multiple alignment shown below**

<table>
<thead>
<tr>
<th>Sample 11 Consensus Sequence (796 bp)</th>
</tr>
</thead>
<tbody>
<tr>
<td>TCCTCTGTTTAAACCAATTCGCCCTTTCCAGCTTATGATATGCTTAAATTCA</td>
</tr>
<tr>
<td>GGGG GTAATCCCGCTGACCTGGGGGTCGCTGATGAGGTCGCAAGAAGACACAC</td>
</tr>
<tr>
<td>ATAGGGTGCG AGGGACACCTTCAACAGCGACGCCGACACACGGGGNGNTACAGA</td>
</tr>
<tr>
<td>GGGTTTCCTAAACCC CGATTTGTCGTTGGCGCTCGCTGCCTTTAGGACTACCTTTTAGGCC</td>
</tr>
<tr>
<td>TAACCCGCGGCAAAGCGCAGGAGGGCAATGTCTTTCCCGCACCCAGCACAGCG</td>
</tr>
<tr>
<td>GATAAGAAGTGGGGGGGGGTCGGGCAACCCAGCGGACGCACG</td>
</tr>
<tr>
<td>GCCCTCGCCGAAATGCGCTGCTGGCGCACTTGC GTTCAAGACTC(g/t) TATGAT</td>
</tr>
</tbody>
</table>
The sample 11 consensus sequence is 96% (629/653 bp) identical to *J. nigra*. Two sequences were used to construct this consensus sequence, one forward and one reverse (Figure 10). The consensus sequence was then assessed for matches in a BLAST assay (Figure 11).
The segment of the consensus sequence used in the BLAST assay shown in Figure 11 has also been highlighted in Table 6. Figure 11 shows six distinct differences between the consensus sequence and the GENBANK closest match. All these differences can be attributed to the editing process and subsequent construction of the consensus sequence. They are all extra base additions (Figure 11) and are due to having mismatched bases between the two single run sequences (forward and reverse), and having to include both possibilities in the consensus. This problem may have been avoided had there been at least one more single run sequence, since there would be a greater number of sequences to draw the consensus from, thereby avoiding the problem of having to include two possible bases in the sequence. In the example shown in Figure 10 the \textit{11 (reverse)} sequence is identical to the Genbank acquired sequence shown in Figure 11. It is therefore the \textit{11 (Forward)} sequence that appears to have been inaccurate in this instance.

\textbf{Table 7. Consensus Sequence and NCBI Blast best match for sample 15}

<table>
<thead>
<tr>
<th>Sample 15 Consensus Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>No significant similarity is observed in the multiple alignment of the Forward and reverse reactions of sample 15 (from Appendix 2)</td>
</tr>
</tbody>
</table>

The reason for the lack of similarity of sample 15 is not completely understood. It is likely that it is a combination of poor sequence quality to begin with, as well as editing error. Looking at the NCBI BLAST of both the forward and reverse reactions for sample 15 Appendix 2), it is apparent that this sample did not have the high sequence
resolution of the other samples. Contamination of the plasmid or during the sequencing of sample 15 may have also played a part.

Consensus Sequence Multiple Alignment

Figure 12. Partial multiple alignment for all sample consensus sequences with J. nigra sequence from Genbank included (from Appendix) [Additions (red)] [Substitutions (blue)] [Deletions (green)]

Figure 12 depicts the multiple alignment of the sample consensus sequences and J. nigra sequence taken from Genbank. Overall, the alignment pattern is comprised of short spans of high identity sequence of varying length, interrupted intermittently by sequence variation in one or more of the sample sequences, in the form of a nucleotide base addition, substitution, or deletion.

The majority of the sequence discrepancy is in the form of additions. As mentioned earlier, the likely culprit for these additions is human error during the
sequence editing process, or sequencer error. This theory is supported in that most of these additions were only observed in one of the sample sequences at any certain point throughout the sequence. Looking at Figure 12, when additions did happen in more than one sample at the same time in that sequence, they were sometimes conserved. It is those base additions that are conserved in more than one sequence that hold the best chances of actually being sample variation, and should be further studied.

Nucleotide substitutions were less frequent in the multiple alignment. They can also be byproducts of the sequence editing process. Several substitutions were visible in the first and second lines of Figure 12. These divergences of sequence may be indicative of true variation since they were conserved in both the forward and reverse reactions that went into the consensus sequence, and therefore should be examined further in future studies.

Only one deletion was observed. Located in the first line of Figure 12, sample 1, it appears that the nucleotide base had been shifted to the left somehow, since the missing nucleotide was listed as an addition in the adjacent space.

**Comparison by location within Zoar Valley:** All samples sequenced were positively identified as being of the species Juglans nigra. No clear discrepancy or pattern of genetic variance was able to be discerned between sample locations.
Conclusion: The primary goals of this study- to sequence multiple J. nigra subjects within 3 distinct populations inside Zoar Valley, and identify these sequences utilizing examples published in Genbank- were accomplished.

The use of a consensus sequence is of the utmost importance when sequencing DNA because each single run sequencing run can yield slightly different results. An example of this phenomenon was observed in this study, using a consensus sequence allowed for more accurate DNA sequencing. In the initial NCBI BLAST of the single run sequences (Table 1), one sequence was most identical to *J. microcarpa*, 5 (Forward -47), and two were most identical to *J. major* 3 (Forward) and 11 (Forward), instead of all the sequences being most identical to *J. nigra* as assumed they would be. The results of the consensus sequence analysis however showed that although one of the single run sequences may have been more identical to another *Juglans* species, when these single runs were used to construct a consensus sequence, that sequence was most identical to the *J. nigra* sequences in a BLAST assay. All the consensus sequences were matched with *J. nigra* sequences that have been submitted to and published Genbank.

While these results are very descriptive of the rDNA ITS regions of the sampled trees, the number of single run sequences used to construct the consensus sequences was minimal, with three sequences for sample 1 and two for the remaining samples. An increased number of sequencing and identification trials should be conducted before placing any of the sampled trees definitively under the classification of *J. nigra*, and will help determine which, if any, sequence variations observed in this study are real.

The results fell short of answering questions in key areas, such as determining a preliminary colonial history of *J. nigra* in Zoar Valley, and identification of specific
inherited variations in sequence within populations. While variation was observed, the results remain inconclusive because the total number of sequences was low, therefore further analysis is needed.

Many samples were lost during their processing from leaf tissue to DNA sequences, particularly during the methods of PCR, cloning, and sequencing. Perhaps if more samples had been successfully sequenced, the results of this study would provide more powerful evidence for the possible variation observed in the sample sequences from this study.

**Future Research Implications:** Three persistently problematic methods were: the consistent production of PCR products, efficiency in manufacture of positive (containing PCR product insert) clones, and reliable sequencing reactions.

PCR product was not consistently produced from DNA extracts. Many samples were contaminated, resulting in the amplification of fungal DNA instead of plant DNA, because the ITS primers used are universal for eukaryotes. Therefore mixes of plant and fungi DNA resulted in competition for the primers, reducing the yield of the desired plant DNA PCR product. Increased vigilance and focus on detail during collection of samples seemed to prevent this problem. Also, it is theorized that many failed reactions were the result of the ITS1 and ITS4 primers themselves. These primers are fairly short, and had there been any variation between the primer sequence and target DNA sequence, steric hinderance may have prevented the reliable annealing of the primers.

During cloning, the primary problem experienced was in generation of clones that contained a complete insert. This was believed to be due to free nucleotides present in the PCR product that were being preferentially inserted into the plasmid, rendering
plenty of clones, but that when subjected to EcoR1 digest, showed no insert. PCR cleanup kits were tried and did not rectify the problem. Another possible cause of the problem was contamination of the cloning kit chemicals. Two cloning kits were used (see Methods), and fresh supplies were purchased, but the problem remained.

Sequencing problems were in two parts: reliably generating sequence, and generating full sequences. Generation of sequences required careful quantification of the concentration of plasmid DNA and primers, and many reactions likely failed early on during this research due to incorrect quantification of DNA concentrations. As to generating relatively lengthy sequences, it seemed that low resolution of sequencing reaction products by the sequencer may have, in part with slightly miscalculated reagents, contributed to the generation of short fragments. It is recommended for the future that smaller PCR fragments be generated. Use of the ITS2 and ITS3 primers (White et al. 1990), located on either end of the 5.8S ribosomal gene, along with the ITS1 and ITS4 primers, would produce such shorter PCR products.

References:


Title: A molecular identification of fungal ribosomal DNA isolated from *Ulmus americana* and *Ulmus rubra* leaf tissue samples from Zoar Valley, NY using ribosomal DNA internal transcribed spacer sequences (rDNA ITS)

**Background:**

The entire, living plant leaf is known as the phyllosphere, and it includes the surface and interior. The plant leaf surface is a complex terrestrial habitat containing a wide variety of microorganisms including bacteria, filamentous fungi, and yeast (Carroll et al. 1977, Levetin & Dorsey 2006). Fungal endophytes are fungi that inhabit the tissue of plants without causing visible disease symptoms (Schulz & Boyle 2005). This is contrasted with the term “fungal epiphyte”, which refers to organisms living on the outside of the plant (Wilson 1995). Phylloshpere fungi are those that found on the surface, and within living leaves (Petrini 1991).

In woody perennials, fungal endophytes are thought to protect the plants in which they live by one or more mechanisms (antibiosis, mycoparasitism, induced resistance, competitive exclusion), and are thought to develop from environmental and background inoculum, and are not transferred from generation to generation (Johnson & Whitney 1989; Crozier *et al.* 2006). Some endophytic fungi have been shown to effectively antagonize herbivores and pathogenic fungi (Carroll, 1988; Clay, 1988). Defensive mutualisms involving the protection of host plants by animals, primarily ants, are well known but may be far less common than defensive mutualisms with fungi (Clay 1988).

New York State’s Zoar Valley was chosen for sampling due to its intact riparian ecosystem healthy distribution of the two elm species of interest. The woodlands of Zoar
Valley are highly varied and diverse, and collectively meet all objective criteria for
eastern old growth forest (Diggins and Kershner 2005). No known previous analysis of
leaf-associated fungi has been conducted in these woodlands on any native tree species.
This study represents a preliminary survey of the phyllospheric fungi associated with the
native elm species of Zoar Valley, from several distinct woodland areas within. Little is
known concerning the common fungal species inhabiting the phyllosphere of non-
economically important eastern woodland broadleaf tree species such as elms. A previous
paper from Levetin and Dorsey (2006) analyzed the contribution of *Ulmus americana*
leaf surface fungi to the air spora. Samples taken from leaves were cultured and identified
morphologically and then compared with cultures grown from spores collected from the
air.

Elms, in particular *U. americana* (American elm), have been an unmistakable
landmark of culture in the North American continent. Elms are highly valued shade trees,
once found lining city blocks from coast to coast (Hubbes 1999). Elm trees still thrive in
Zoar Valley today, although their populations have been decimated elsewhere due to
Dutch Elm Disease, a catastrophic infection caused by the fungus *Ophiostoma ulmi*,
which usually results in death of the tree.

In 2003 the American Phytopathological society called for an increase in funding
for research of many plant associated microbes, stating the importance of genetic
screening and identification of plant-associated pathogenic microbes in disease
prevention. Included in its list of “high priority species” for study were *Ophiostoma
novo-ulmi*, as well as *Candidatus phytoplasma ulmi*, a phytoplasma that causes Elm
Yellows. Analysis of phyllospheric communities may help identify endophytic
associations in elm trees in Zoar Valley, shedding light on the existence of certain colonizing fungi that help to prevent pathogenicity of other microbes.

Numerous studies analyzing temperate and tropical trees have showed that endophytes represent an important and quantifiable component of fungal biodiversity (Levetin and Dorsey 2006, Clay 1988, Arnold et al. 2001).

**Methods:** Samples were obtained from Zoar Valley over fall of 2005 and spring of 2006. Samples were collected from 5 distinct areas of Zoar Valley, Lookout Point Terrace, South Branch Floodplain, Burchfield Terrace, Skinny Dip Terrace, and Skinny Dip Streamside. Samples of *U. Americana* leaf tissue were first positively identified by analysis of leaf structures and bark, then collected and stored in deep freeze at -80 C° until DNA extraction was performed.

**DNA extraction:** Total DNA was extracted using a revised CTAB method of Doyle and Doyle (1987).

**PCR:** Extracted DNA is amplified with ITS specific primers. Primers used were ITS1 and ITS5 (forward) and ITS4 (reverse) of White et al. (1990). PCR performed on PTC-200 DNA engine.

**Cloning:** Purified PCR products were cloned using the TOPO TA Cloning Kit from Invitrogen (cat# K4530-20) or StrataClone PCR cloning kit from Stratagene (cat#240205)

**Restriction Digest:** A restriction digest using EcoR1 was used to screen clones before sequencing.
**Sequencing:** Sequencing reactions were carried out on purified cloning extracts using Beckman Coulter Sequencing Kit, sequenced using primers M13 (-20, -47) Forward and M13 Reverse, with sites provided in the cloning vector. Other sequences were from purified plasmid generated in this lab and sequenced at the Ohio State University Plant Microbe genomics facility using the same technique. Sequencing reactions were conducted at Ohio State University on a 3730 DNA analyzer from Applied Biosystems, Inc.

**Alignment:** Sequences are aligned to determine identity or variation in ITS region of each species using computer programs, such as NCBI nucleotide BLAST (http://www.ncbi.nlm.nih.gov/BLAST/) and ClustalW (http://www.align.genome.jp/sit-bin/clustalw).

**Results and Discussion:**

Ten fungal samples were recovered from *Ulmus Americana* and 15 from *Ulmus rubra* leaves. Multiple fungal species were recovered from the same leaf sample in several cases. Sample 43 yielded seven different species of fungus.

**Table 1.** Identification of fungi using PCR amplified DNA extracted from elm leaves within Zoar Valley

<table>
<thead>
<tr>
<th>Tree species fungal sample was isolated from</th>
<th>Sample # and ITS ribosomal primer set used to sequence</th>
<th>Collection Site within Zoar Valley</th>
<th>Closest match in NCBI Nucleotide Blast Search And % Identity</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>U. americana</em></td>
<td>9 -Reverse</td>
<td>South Branch Floodplain</td>
<td>Phlebia radiata 90%</td>
</tr>
<tr>
<td><em>U. americana</em></td>
<td>12 -Forward</td>
<td>Lookout Point Terrace</td>
<td>Phoma sp. 90%</td>
</tr>
<tr>
<td>Specimen</td>
<td>Section</td>
<td>Location</td>
<td>Species</td>
</tr>
<tr>
<td>--------------</td>
<td>---------</td>
<td>-----------------------------</td>
<td>--------------------------</td>
</tr>
<tr>
<td><em>U. americana</em></td>
<td>12</td>
<td>Lookout Point Terrace</td>
<td>Phoma sp.</td>
</tr>
<tr>
<td><em>U. americana</em></td>
<td>63</td>
<td>South Branch Floodplain</td>
<td>Coprinellus sp.</td>
</tr>
<tr>
<td><em>U. rubra</em></td>
<td>50</td>
<td>Skinny Dip Terrace</td>
<td>Uncultured leaf litter</td>
</tr>
<tr>
<td><em>U. americana</em></td>
<td>55</td>
<td>Skinny Dip Streamside</td>
<td>Coprinellus sp.</td>
</tr>
<tr>
<td><em>U. americana</em></td>
<td>31D</td>
<td>Burchfield terrace</td>
<td>Coprinellus sp.</td>
</tr>
<tr>
<td><em>U. americana</em></td>
<td>31D</td>
<td>Burchfield terrace</td>
<td>Coprinellus sp.</td>
</tr>
<tr>
<td><em>U. americana</em></td>
<td>31E</td>
<td>Burchfield terrace</td>
<td>Coprinellus sp.</td>
</tr>
<tr>
<td><em>U. rubra</em></td>
<td>43i #4</td>
<td>Skinny Dip Terrace</td>
<td>Colletotrichum truncatum</td>
</tr>
<tr>
<td><em>U. rubra</em></td>
<td>43i #5</td>
<td>Skinny Dip Terrace</td>
<td>Colletotrichum truncatum</td>
</tr>
<tr>
<td><em>U. rubra</em></td>
<td>43s #2</td>
<td>Skinny Dip Terrace</td>
<td>Cryptococcus sp.</td>
</tr>
<tr>
<td><em>U. rubra</em></td>
<td>43s #4</td>
<td>Skinny Dip Terrace</td>
<td>Phaeosphaeria sp.</td>
</tr>
<tr>
<td><em>U. rubra</em></td>
<td>43s #5</td>
<td>Skinny Dip Terrace</td>
<td>Ampelomyces sp.</td>
</tr>
<tr>
<td><em>U. rubra</em></td>
<td>43s #6</td>
<td>Skinny Dip Terrace</td>
<td>Gyoerffyella sp.</td>
</tr>
<tr>
<td><em>U. rubra</em></td>
<td>43i #4</td>
<td>Skinny Dip Terrace</td>
<td>Colletotrichum truncatum</td>
</tr>
<tr>
<td><em>U. rubra</em></td>
<td>43i</td>
<td>Skinny Dip Terrace</td>
<td>Didymella bryoniae</td>
</tr>
<tr>
<td><em>U. rubra</em></td>
<td>43ii</td>
<td>Skinny Dip Terrace</td>
<td>Phoma glomerata</td>
</tr>
<tr>
<td><em>U. rubra</em></td>
<td>43s#2</td>
<td>Skinny Dip Terrace</td>
<td>Cryptococcus sp.</td>
</tr>
<tr>
<td><em>U. rubra</em></td>
<td>43s#4</td>
<td>Skinny Dip Terrace</td>
<td>Colletotrichum truncatum</td>
</tr>
<tr>
<td><em>U. rubra</em></td>
<td>43s#6</td>
<td>Skinny Dip Terrace</td>
<td>Colletotrichum truncatum</td>
</tr>
<tr>
<td><em>U. rubra</em></td>
<td>45#1</td>
<td>Skinny Dip Terrace</td>
<td>Ampelomyces humuli</td>
</tr>
</tbody>
</table>
Background on genetically identified genera (Listed in order of highest to lowest # unique sequences identified per genera):

*Phoma*: In this study, we recovered three unique sequences identified from three different leaf samples (Table 1 and Table 2). Samples containing *Phoma* were from Lookout Point Terrace and Skinny Dip Terrace. *Phoma* is a large genus of anamorphic fungi in the form class Coelomycetes that is characterized by conidia formation in a
pycnidium, and their conidia are believed to be dispersed throughout a tree via a rain splash mechanism. *Phoma* species are documented endophytes, and have been frequently found in a wide array of plants species, including several types of cacti in Arizona, beech and giant dogwood in Japan, as well as American elm (Suryanarayanan et al. 2005, Osono and Mori 2003, Osono et al. 2004, Levetin and Dorsey 2006). *Phoma* species likely exist as endophytes in the trees sampled.

**Coprinellus:** Three unique sequences were detected in 3 different leaf tissue samples (Table 1 and Table 2). *Coprinellus* was found to be well distributed throughout the sampled areas of this study, with samples from South Branch Floodplain, Skinny Dip Streamside, and Birchfield Terrace (Figure 1). Species of this Basidiomycete genus have shown to be common endophytes in the stems and pods of agriculturally grown *Theobroma cacao*, the tree whose fermented and dried beans are used to produce chocolate (Crozier et al. (2006). *Coprinellus* has also shown promise as an agricultural agent in Chinese cabbage, effective for suppressing soil-borne pathogens, presenting new possibilities for biological control of vegetable diseases (Nakasaki et al. 2007). It is probable that *Coprinellus* exists in *Ulmus sp*. of Zoar Valley as an endophyte, but further studies would be needed to confirm.

**Cryptococcus:** Two unique samples of *Cryptococcus* were found in one geographic area, Skinny Dip Terrace. One species, *C. neoformans*, is a pathogenic yeast, and the most common fungal cause of meningitis in patients with AIDS (Litvinseva et al. 2007). This yeast opportunistically infects humans and other animals. *Cryptococcus* species have
been associated with soil, animal droppings, and other organic materials. Endophytic or pyloospheric colonization is unlikely. Presence of Cryptococcus in leaf samples is probably from an animal source.

**Ampelomyces:** Two unique sequences identified as *Ampelomyces* were detected in two samples from Skinny Dip Terrace. This genus includes many species of mycoparasitic fungi that cause powdery mildews. The typical *Ampelomyces* fungus infects and forms pycnidia inside of fungal hyphae of other fungi. Cells of this parasite therefore grow inside of the host, causing pathogenesis and death (Rotem et al. 1999). *Ampelomyces* species have been closely associated with apple shoots and aerial parts of 13 other flowering plant species. *Ampelomyces* has been shown to play an integral role in “bud bursting” of certain plants, such as apple trees in Holland (Szentivanyi and Kiss 2003). The presence of *Ampelomyces* in *Ulmus* samples likely indicates infection of a separate host fungus present in the leaf sample.

**Colletotrichum:** One unique sequence identified as *Colletotrichum* was isolated from Skinny Dip Terrace. Species of this genus have been observed to exist in the phyllosphere of giant dogwood trees in Japan (Osono et al. 2004) *Colletotrichum* species have caused increasing numbers of opportunistic human infections in recent years, usually in the immunocompromised HIV and transplant patients. The genus *Colletotrichum* is one of the most important genera of plant pathogens because of the diverse variety of economically important plants it colonizes (Cano et al. 2004). *Colletotrichum* species cause economically significant diseases of plants (generally
known as anthracnoses) that affect cereals and grasses, legumes, vegetables, and perennial crops, including fruit trees. *Colletotrichum* likely existed as a pathogen in the *Ulmus sp.* samples from Zoar Valley.

**Phlebia:** *Phlebia* was found to exist in one sample from South Branch Floodplain. It is classified within class Basidiomycota, which is known for its abilities to degrade lignin. It has been found that proteosomal degradation upon nitrogen and carbon starvation is possibly involved in the regulation of ligninolytic activities in these wood decaying fungi (Staszczak 2007). These fungi are currently the subject of numerous microbial ecology studies for their lignin degrading abilities. *Phlebia* is not known for endophytic relationships and may be associated with soil contamination of specimens.

**Phaeosphaeria:** One sequence identified as *Phaeosphaeria* was isolated from Skinny Dip Terrace. Many fungi placed taxonomically in the genus *Phaeosphaeria* were once found in the genus *Leptosphaeria*, and morphological distinctions of classification between these genera are often blurry. Some characteristics of *Phaeosphaeria* are production of ascospores with a distinguishing perispore, and the induction of Stagonospora leaf blotch diseases in cereals (Ueng et al. 2003). Phaeosphaeria exhibit pathogenesis of certain plants, and it is possible that it was existing parasitically with elm samples.

**Gyoerffyella:** One *Gyoerffyella* sequence was isolated from Skinny Dip Terrace. This genus of fungi is hypothesized to be a colonizer of the leaf phyllosphere. Species of this hyphomycete genus have been observed in rainwater collected after draining from forest
canopies in British Columbia (Gonczol and Revay 2006). Not much information is available on the species of this genus, although they are found extensively throughout North America and Europe. It is probable that *Gyroerffyella* colonized the phyllosphere of elm samples from Zoar Valley as an epiphyte, evidenced by its documented presence in rainwater collections.

**Didymella:** One *Didymella* sequence was detected in a sample from Skinny Dip Terrace. Species of this genus, such as *D. bryoniae*, are known plant pathogens. Plants affected by *Didymella* infection are wide ranging and include many economically important species, such as wheat, watermelon, pumpkin, cucumber and squash. One associated disease is Gummy Stem Blight. Appearance of spots on the leaves, petioles and stems are a typical sign of infection which usually become pale brown or gray in color. Gummy exudates may occur from cracks, especially in watermelon and pumpkin. Severe infection often results in death of the plant (Ferreira and Boley 1992). Didymella therefore may have existed as a pathogen in *Ulmus* species.

Skinny Dip Terrace showed the greatest variety of genera in our samples, with 7 of 9 total genera identified in this location. This resulted because sample 43, from Skinny Dip Terrace, was sequenced 12 times. 7 genera of fungi were detected in sample 43 alone. Most of the fungi implicated as pathogenic were taken from this sample.

*Phoma* and *Coprinellus*, the most likely of all genera detected to exist as endophytes, were relatively well dispersed in sampling areas, identified in 2 and 3 out of
five sample sites respectively. This finding supports the idea that these fungi are endophytes within elm species, since they were present in multiple samples and areas.

Conclusion:

This study represents a preliminary survey of phyllosphere fungi associated with *U. americana* and *U. rubra* in Zoar Valley. Genera detected were wide ranging and likely occupy various roles in their association with the trees they inhabit.

Future research should utilize a method of extraction to distinguish between endophytes and epiphytes, such as washing the leaf surface, then separately identifying fungi found in the wash from those isolated in leaf tissue. This will improve the understanding of how a fungus is associated with the plant, and enable the researcher to better hypothesize a plant-fungus relationship scenario.

Another improvement would be to increase the sample size. This would not only yield more genera of leaf-associated fungi for all locations, but would better characterize the phyllospheric ecosystems of *Ulmus sp.* within each location. This information could be vital in diagnosis and prevention of plant pathogens in the future.

References:


Appendix 1: Methods

DNA Extraction:
Johnston Lab CTAB DNA Extraction Protocol (Reference: Doyle and Doyle, 1987; and Cullings 1992)
Revised December 11th, 2006

1. Final preparation of CTAB buffer. Must use within 5-7 days. Add polyvinylpyrrolidone (PVP) and B-mercaptoethanol in 0.04 and 0.005 volumes respectively. Stir gently to dissolve.

<table>
<thead>
<tr>
<th>CTAB Buffer</th>
<th>PVP</th>
<th>B-merc</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.5 ml</td>
<td>0.02g</td>
<td>2.5 ul</td>
</tr>
<tr>
<td>5 ml</td>
<td>0.2g</td>
<td>25 ul</td>
</tr>
</tbody>
</table>

2. Weigh out 40 mg of frozen plant tissue. Avoid using tissue samples that are discolored or show obvious freeze-thaw damage.

3. Grind tissue with glass pestle in liquid nitrogen cooled mortar
   a. Pestles must be acid-washed and rinsed well with DI water before reuse.
   b. Nitrogen gas will condense into liquid in the mortar. Allow sample to dry of liquid nitrogen, before adding CTAB, by removing the mortar from its base and placing on lab bench for several minutes.

4. Place freshly ground tissue in sterile microfuge tubes and add 500 ul of CTAB buffer.

5. Invert tube 5-10 times and incubate samples at 55 degrees C overnight.

6. Add 500 ul of 24:1 Chloroform: Iso Amyl Alcohol and mix well by shaking tubes.

7. Centrifuge 10 minutes at 13,000 RPM.
   a. Following centrifugation should have 3 layers in tube: top: aqueous phase, middle: protein debri, and bottom: chloroform.
   b. Aqueous phase contains DNA, pipette off quickly into a fresh microfuge tube.

8. Estimate the volume of the collected aqueous phase.

9. Add 0.1 volumes of cold 7.5 M ammonium acetate and 0.6 volumes of cold isopropanol (using combined volumes of aqueous layer + ammonium acetate). Ammonium acetate and isopropanol should be kept in the -20 freezer just prior to use.

10. Mix well. Place in -20 degrees C freezer for one hour to overnight.

11. Centrifuge for 3 minutes at 13,000 RPM

12. DNA pellet should be visible. Pour or pipette off supernatant, careful not to lose pellet. Pellet may vary in color from light brown to creamy white.

13. Add 700 ul of cold 70% ethanol and mix. Centrifuge at 13,000 RPM for one minute.

14. Pour off liquid. Add 700 ul of cold 95% ethanol and mix. Centrifuge one minute.

15. Carefully pour of liquid, being sure not to lose pellet.

16. Dry pellet by inverting on Kim-wipe for an hour or until dry.

17. Re-suspend sample in 100 ul of TE Buffer. Place in freezer -20 degrees C to store.
Stock Solutions:

CTAB Buffer: for 1 liter
- 100 ml of 1 M Tris, pH 8.0
- 280 ml of 5 M NaCl
- 40 ml of 0.5 M EDTA
- 20 g of CTAB (Cetyltrimethyl ammonium bromide)

1 M Tris, pH 8.0: 1 liter
- 121.1 g Tris
- 700 ml mqH2O
  Dissolve Tris, bring volume to 900 ml
  pH to 8.0 with concentrated HCl (~50 ml)
  Bring to 1 liter

0.5 M EDTA, pH 8.0: 1 liter
- 186.12 g of EDTA
- 750 ml mqH2O
  Add approximately 20 g of NaOH pellets until EDTA dissolves (~pH 8.0)

5 M NaCl: 1 liter
- 292.2 g of NaCl
- 700 ml of mqH2O
  Dissolve and bring volume to 1 L

TE Buffer: 1 liter
- 10 mM 10 ml of 1 M Tris, pH 8.0
- 1 mM 2 ml of 0.5 M EDTA


Generating PCR Fragments
Johnston Lab – PCR protocol for nrDNA Internal Transcribed Spacer (ITS) sequence amplification using primers ITS1 and ITS4 of White et al, 1990.

Plant DNA extracted and stored -20 degrees C in TE buffer until PCR set-up

Following Ingredient list is for a 50ul PCR reaction

<table>
<thead>
<tr>
<th>Ingredient</th>
<th>Conc. Used</th>
<th>Volume Used</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sterile water</td>
<td>-</td>
<td>34.25ul</td>
</tr>
<tr>
<td>Immobuffer</td>
<td>10X</td>
<td>5ul</td>
</tr>
<tr>
<td>MgCl2</td>
<td>50mM</td>
<td>2.5ul</td>
</tr>
<tr>
<td>ITS1</td>
<td>25X</td>
<td>2ul</td>
</tr>
<tr>
<td>ITS4</td>
<td>25x</td>
<td>2ul</td>
</tr>
<tr>
<td>dNTPs</td>
<td>40mM</td>
<td>1.25ul</td>
</tr>
<tr>
<td>Immolase (Taq)</td>
<td>5 Units/ul</td>
<td>0.5ul</td>
</tr>
<tr>
<td>DNA</td>
<td>1/100 dil.</td>
<td>2.5ul</td>
</tr>
<tr>
<td><strong>TOTAL</strong></td>
<td></td>
<td><strong>50ul</strong></td>
</tr>
</tbody>
</table>

PCR cycle “IT” program:
1. 95 C for 7 min.
2. 94 C for 20s
3. 54.9 C for 30s
4. 72 C for 30s
5. Repeat 2-4 40 times
6. 72 C for 5 min.

Gels run on Agarose, 1.5% in TAE buffer

Cloning of fresh PCR product:
TOPO TA Cloning Kit for Sequencing (Invitrogen cat. No. K4530-20)
Recombinant colonies are analyzed with EcoRI restriction digest and selected for sequencing. Requires fresh PCR product with A-overhangs for the Topoisomerase-1 enzyme to efficiently ligate the PCR product. Non-recombinant colonies are selected through life cycle termination protein integrated into the vector, only activated during ligation of the vector without an insert.

StrataClone PCR cloning kit from Stratagene (Stratagene cat. No. 240206)
Uses the Topoisomerase-1 enzyme therefore requires fresh PCR product with A-overhangs. Cre recombinase gene in activated in vectors that ligate without the insert, therefore clones must be plated on X-gal/ampicillin plates (50ug/ml), and blue/white selection for recombinant clones.

Plasmid Isolation:
Plasmids are isolated from the cloned cells using the alkaline plasmid screen. Plasmid DNA is isolated from an overnight culture of Luria Broth plus 50ug/ml of Ampicillin. The plasmid DNA was eluted into sterile water instead of TE buffer in preparation for sequencing.

1. Inoculate using a flame sterilized needle, 5ml of LB broth with antibiotics in a sterile culture tube with a single recombinant colony. For quick screening, grow cells in 1.5 ml eppendorf tube. (plasmid must be high copy #)
2. Grow cells in 37°C incubator overnight with shaking (200rpm)
3. For culture tubes, spin down at 3,000 rpm for 5 minutes, for eppendorf tubes 1 minute. Pour off the supernatant.
4. Resuspend cells in 150ul of P1 (15 mM Tris pH8, 10mM EDTA) + 10ug/ml RNAse A.
5. Add 150ul of P2 (0.2N NaOH 1%SDS). Mix genly by inverting 3 times.
6. Add 150ul of P3 (3M KOAc pH 5.5). Mix by gently inverting tube 5-10 times.
7. Remove white precipitate by centrifugation at 12,000 rpm for 10 minutes.
8. Carefully transfer the supernatant to a fresh 1.5ml eppendorf tube.
10. Remove the aqueous phase (top) to a new 1.5ml eppendorf tube. This phase contains plasmid DNA.
11. Add two volumes of 95% ethanol. Incubate for 30 minutes at -20°C.
12. Spin in the microcentrifuge for 10 minutes at 12,000 rpm. Decant the supernatant and wash the pellet once with cold 70% ethanol.
13. Air dry the pellet by inverting the tube over a Kimwipe. Resuspend in 30ul of TE buffer and store at -20°C. The plasmid DNA is now ready for downstream applications like EcoR1 digestion and sequencing.

EcoR1 digestion of cloned DNA
Per single reaction:
Water 6.8ul
EcoR1 buffer 1ul
EcoR1 0.2ul
Miniprep 2ul
Combine in tube, being sure to keep enzyme at -20°C while out of freezer, and add the enzyme to the tube last. Next incubate for two hours at 37°C.

*For double digests (2 enzymes at once) use 0.2ul of each enzyme and 6.6ul of sterile water.

Sequencing:
Beckman Coulter sequencing protocol were followed in use of GenomeLab Dye Terminator Cycle sequencing with Quick Start Kit. Following the generation of nested DNA fragments, precipitation, and pellet formation, the pellets were left dry overnight until sequencing on the next morning. This was to avoid freeze-thaw complications regarding the sample loading solution (SLS).
Appendix 2: Sequence Data Comparison with Nucleotide Analysis Via Genbank

The following sequences are those of the Zoar Valley *J. nigra* samples specified in the title, underscored by their closest match, as determined through utilization the NCBI BLAST tool (see Methods). Sample #1 Forward has been labeled {} with the intention of using it as a key to interpreting the BLAST data.

**SAMPLE #1 Forward** {SAMPLE}

> gi|18028823|gb|AF338492.1|AF338492 {ACCESSION #} Juglans nigra isolate 836 18S ribosomal RNA gene, partial sequence; internal transcribed spacer 1, 5.8S ribosomal RNA gene and internal transcribed spacer 2, complete sequence; and 26S ribosomal RNA gene, partial sequence {PUBLISHED SEQUENCE IDENTITY}
Length=750

Score = 1026 bits (555), Expect = 0.0
Identities = 604/631 (95%), Gaps = 11/631 (1%)
Strand=Plus/Minus

*{for the purpose of this paper, the only identity value was utilized. Query represents the research derived sample DNA, while Sbjct (Subject) represents the Genbank published DNA sequence. Dots in the Sbjct represent identical nucleotide match with the Query}*

<table>
<thead>
<tr>
<th>Query</th>
<th>82</th>
</tr>
</thead>
<tbody>
<tr>
<td>CCGCTTAATTGATATGCTTAAATCAGGGGGTAATCCGCTGACCTGGGGTCGCGATGGT</td>
<td>141</td>
</tr>
<tr>
<td>Sbjct</td>
<td>749</td>
</tr>
<tr>
<td>............................................................</td>
<td>690</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Query</th>
<th>142</th>
</tr>
</thead>
<tbody>
<tr>
<td>AGAGTCGCAAAGAAGCACAAATAGGGTCAGAGCACCCTCACACGCGACGGGCGGACACAC</td>
<td>201</td>
</tr>
<tr>
<td>Sbjct</td>
<td>689</td>
</tr>
<tr>
<td>............................................................</td>
<td>630</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Query</th>
<th>202</th>
</tr>
</thead>
<tbody>
<tr>
<td>GACGGGTCACGAGGGTTTCTCAACCACCGGATTTGTGCTGGCGGCTGCTGGCTTTGCTAGGACTCAC</td>
<td>261</td>
</tr>
<tr>
<td>Sbjct</td>
<td>629</td>
</tr>
<tr>
<td>............................................................</td>
<td>570</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Query</th>
<th>262</th>
</tr>
</thead>
<tbody>
<tr>
<td>TTTTGGCTAACCAGCGACGCAAAGCAGCACGGAGGCCAATGTCTTTGCACCCGACACCACCGC</td>
<td>321</td>
</tr>
<tr>
<td>Sbjct</td>
<td>569</td>
</tr>
<tr>
<td>............................................................</td>
<td>510</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Query</th>
<th>322</th>
</tr>
</thead>
<tbody>
<tr>
<td>AGCGTAAGAGTGTTTGGGGTTGGGCAACGATGCGTGACCCAGGACGACGTCCCTC</td>
<td>381</td>
</tr>
<tr>
<td>Sbjct</td>
<td>509</td>
</tr>
<tr>
<td>............................................................</td>
<td>450</td>
</tr>
</tbody>
</table>
SAMPLE #1 Reverse

Juglans nigra isolate 836 18S ribosomal RNA gene, partial sequence; internal transcribed spacer 1, 5.8S ribosomal RNA gene and internal transcribed spacer 2, complete sequence; and 26S ribosomal RNA gene, partial sequence

Length=750

Score = 780 bits (422), Expect = 0.0
Identities = 448/461 (97%), Gaps = 8/461 (1%)
Strand=Plus/Plus
SAMPLE #5 Foreward

Juglans microcarpa isolate 108 18S ribosomal RNA gene, partial sequence; internal transcribed spacer 1, 5.8S ribosomal RNA gene and internal transcribed spacer 2, complete sequence; and 26S ribosomal RNA gene, partial sequence

Length=735

Score = 1094 bits (592),  Expect = 0.0
Identities = 690/749 (92%), Gaps = 21/749 (2%)
Strand=Plus/Minus

Query  90
CCGCTTATTGATATGCTTAAATTCAGCGGGTAATCCCGCCTGACCTGGGGTCGCGATGGT  149
Sbjct  734
............................................................  675

Query  150
AGAGTCGCAAGAACGACACAATAGGGTCGAGGAGCACCTTCACAGCGACGGGCGACACAC  209
Sbjct  674
............................................................  615

Query  210
GACGGGTCAAGGAGGTTTCTCAACCACCGATTGTCGTCGCTGCTGCTGCTACGACTCAC  269
Sbjct  614
............................................................  555
SAMPLE #5 Reverse

> gi|18028823|gb|AF338492.1|AF338492 Juglans nigra isolate 836 18S ribosomal RNA gene, partial sequence; internal transcribed spacer 1, 5.8S ribosomal RNA gene and internal transcribed spacer 2, complete sequence; and 26S ribosomal RNA gene, partial sequence
SAMPLE #11 Forward

Juglans major isolate 870 18S ribosomal RNA gene, partial sequence; internal transcribed spacer 1, 5.8S ribosomal RNA gene and internal transcribed spacer 2, complete sequence; and 26S ribosomal RNA gene, partial sequence

Length=747

Score = 870 bits (471), Expect = 0.0
Identities = 480/484 (99%), Gaps = 1/484 (0%)
Strand=Plus/Plus

Query 45
CCGTAGGTGAACCTGCGGAAGGATCATTGTCGATACCTGCCCAGCAGAACGACCTGTGAA 104
Sbjct 19
............................................................ 78

Query 105
CATGTAATAACCTTCTGGGTGGGGGTGTAATGCCCGCCCTCCCAAAAAACGGTTGGGAGGGC 164

SAMPLE #11 Reverse

Juglans nigra isolate 836 18S ribosomal RNA gene, partial sequence; internal transcribed spacer 1, 5.8S ribosomal RNA gene and internal transcribed spacer 2, complete sequence; and 26S ribosomal RNA gene, partial sequence

Length=750

Score = 968 bits (524), Expect = 0.0
Identities = 537/545 (98%), Gaps = 3/545 (0%)
Strand=Plus/Minus
SAMPLE #11 Reverse II

Juglans nigra isolate 836 18S ribosomal RNA gene, partial sequence; internal transcribed spacer 1, 5.8S ribosomal RNA gene and internal transcribed spacer 2, complete sequence; and 26S ribosomal RNA gene, partial sequence

Length=750

Score = 926 bits (501), Expect = 0.0
Identities = 537/559 (96%), Gaps = 7/559 (1%)
Strand=Plus/Minus
Query 70
GCTTATTGATATGCTTAAATTCAACCGGTAATCCCGCCTGACCTGGGGTCGCGATGGTAG 129
Sbjct 747
............................................................ 688

Query 130
AGTCGCAAGAACGACACAATAGGGTCGAGGAGCACCTTCACAGCGACGGGCGACACACGA 189
Sbjct 687
............................................................ 628

Query 190
CGGGNTCACGAGGGTTTCTCAACCCAGATTGTGCGTGCGCTCGCTGCTAGCCTCAGT 249
Sbjct 627
............................................................ 569

Query 250
TTTAGCTAACCAGGAGCAAAAGCGCAGGGAGCCAAATGTCTTCCCCGCACCCCNACA 309
Sbjct 568
............................................................GC... 509

Query 310
GCGTAAGAAGTGTTTGGGGTTGGGGCAACGATGCGTGACACCCAGGCAGACGTGCCCTCG 369
Sbjct 508
............................................................ 449

Query 370
GCCGAAATGGCTGGCGGCGAACTTGCGTTCAAAGACTCTNATGATTCGCGGGATTCTGCA 429
Sbjct 448
............................................................-G.  390

Query 430
ATTCAACCAAGATATCGCATTTTGCCTACTCTCTTCTATCGATGCNAGAGCGAGATATC 489
Sbjct 389
............................................................-G.  332

Query 490
CGTTGCCGAGCTGTTTATGTATCATGGTAAAGACGTTACCAACGACACCCGTT 549
Sbjct 331
............................................................C.  273

Query 550
TCCGGNMGCCGTGGTTACTCCTCTGTATATCGTAAAGACGTTACCAACGCACACCCGT 608
Sbjct 272
........................GC...................................GG...G...GG-  214

Query 609  CATTGTNCATCGGNAAG 627
Sbjct 213 ......T.G......G...  195

SAMPLE #3 Forward

>gi|18028815|gb|AF338484.1|AF338484  Juglans major isolate 870 18S ribosomal RNA gene, partial sequence; internal transcribed spacer 1, 5.8S ribosomal RNA gene and internal transcribed spacer 2, complete sequence; and
26S ribosomal RNA gene, partial sequence
Length=747

Score = 850 bits (460), Expect = 0.0
Identities = 479/491 (97%), Gaps = 1/491 (0%)
Strand=Plus/Plus

Query 2
AACCNGCGGAAGGATCATTGTCGATACCTGCCCAGCAGAACGACCTGTGAACATGTAATA 61
Sbjct 28
................................................................. 87

Query 62
ACCTTCTGGTGTTGGTGTAATGCCACTCTCCCAAAAAACGGTTGGAGGCACTGAGA 121
Sbjct 88
.................................................................. 147

Query 122
TTTGCCCACTGCTCCTCGTGTGTGGTTGGTGATCCTCTCGTTCCCTCCTCCGATCGACA 181
Sbjct 148
................................................................. 207

Query 182
ACGACAACCGGCGGCTCGCAAGAACTTAAACAAAGGATGAAACGCGGCGCCCG 241
Sbjct 208
................................................................. 267

Query 242
GAAACGCTGTGCTGTTGACGTCTTTACCAAGATACATAACGACTCTCGGCAAC 301
Sbjct 268
................................................................. 327

Query 302
GGTAATCTCGGCTCGCAGATGGAAGCGTAAAGATTCACTGCTTGGTGTAATT 361
Sbjct 328
.................................................................. 387

Query 362
GCAGAATCCCGCAATCATCGATGTCTCTCGTGTGATGACGCTGCGCCGAAATGCCATTCGCAAC 421
Sbjct 388
.................................................................. 447

Query 422
AGGGCACCCTGGCTTGTCACGCATCGTTGCCGCTCCTCCTTGGCGCCGACGCATTCGCCG 481
Sbjct 448
.........................................................CC.......................T..T 507

Query 482
TNCNCGGGTG 492
Sbjct 508 .G.G-...... 517

SAMPLE #3 Reverse

> gi|18028823|gb|AF338492.1|AF338492  Juglans nigra isolate 836 18S
ribosomal RNA gene, partial sequence;
internal transcribed spacer 1, 5.8S ribosomal RNA gene
and internal transcribed spacer 2, complete sequence; and
26S ribosomal RNA gene, partial sequence
Length=750

Score = 824 bits (446), Expect = 0.0
Identities = 481/498 (96%), Gaps = 9/498 (1%)
Strand=Plus/Minus

Query  31
CCGTTATTGTATATGCTTTAAATTGACGCCGTAGTAATCCCCTGACCTGGGGTCGATGGT  90
Sbjct  749
............................................................  690

Query  91
AGAGTGCAGAAGACACACAAATAGGGTGCAGACACCTCAGACGGCAGCAGCCCACAC  150
Sbjct  689
............................................................  630

Query  151
GACGGGTTCAGAAGGTGTTTCTCAACCACGGATGTCGTCGGTGGTGTTGCTGCCTAGGACTCAC  210
Sbjct  629
............................................................  570

Query  211
TTTTAGGCTAACCCGGAGCAAAACGGCGCGGGAGCAATGATCTTCTCCCGCACCACCGC  270
Sbjct  569
............................................................  510

Query  271
AGCGTAAGAAGTGGTTGGGGTGGGAACGATGCGTGACACCCAGGCAGGCGTGCCCTC  330
Sbjct  509
.............................................................A 450

Query  331
GGCCGAATGGCTTCGGGCAGCAACCTTGCGTTCAAAGACTCGATGATTCGCGGGATTCTGC  390
Sbjct  390
.............................................................A 391

Query  391
AANNCACCACAGTGATATCGCTTTGGGACAGATGCAGATGCCACCCAGCGGTGCACC  450
Sbjct  390
..TT....-..............................C 334

Query  451
TCCCTTGCGAAGATGTTATTATATGACGTTGTTAAGACGTCCTCCCAANCGACACCGCCACC  510
Sbjct  333
.............................................................A..-..-.A..-.. 276

Query  511
CGTTTTCNCGGCGCCG  528
Sbjct  275 ----.G--------  261

SAMPLE #6 Forward
>
>gi|17065874|emb|AJ251683.1|BAL251683 Betula alba 18S rRNA gene, 5.8S rRNA gene, 25S rRNA gene, internal transcribed spacer 1 (ITS1) and internal transcribed spacer
2 (ITS2)
Length=686

Score = 756 bits (409), Expect = 0.0
Identities = 451/471 (95%), Gaps = 9/471 (1%)
Strand=Plus/Minus

Query
GCTTAAATTGACGCGGGTATGGCTCGACCTGGGTCGCTGGGACGCTGCCTGGGC
Sbjct
686

C...........................................A..... 627

Query
GACACAGGACGGTCAAAGGACGACGACGAGGCAGACGCACGACGCACGACGAGC

Sbjct
626

........................................G........ 567

Query
GTTTGTCACACCGATTGCTGGGGCCGCCTGGCCGAGCTCGCTTTTGCCCGCAACCG

Sbjct
566

..........................A............................A..... 507

Query
CATGCTGACGCTACGGGAGGCCATCTGCCCCACAGGCCCCCTCGTCCCTTGAAG

Sbjct
506

............................................................ 447

Query
GAGATGGGGTTGGGGGCAACGGTTGACACCCAGGCAGACGTGCCCTCGGCAGGTGG

Sbjct
446

............................................................ 387

Query
CTNCGGCCGCAACTTGGCTCAAGACTNCATGATTTCGCGGGATCTGCAATTACACC

Sbjct
386

..........................-............................ 328

Query
AAGTATCCGATTTNCGGATCGATTCAGCGAGCAGATATCCGCTTGGCTG

Sbjct
327

............................................................ 273

Query
CGAGAGGCGGTTTGGTTTAGCAAGATTTCCGCTCCCGCAGGCAGACACACC

Sbjct
272

..........................TC........TA-------------A------------- 225

15 Forward

Juglans nigra isolate 836 18S ribosomal RNA gene, partial sequence;
internal transcribed spacer 1, 5.8S ribosomal RNA gene and internal transcribed spacer 2, complete sequence; and
26S ribosomal RNA gene, partial sequence
Length=750

Score = 617 bits (334), Expect = 1e-173
Identities = 510/581 (87%), Gaps = 67/581 (11%)
Strand=Plus/Minus

15 Reverse

Juglans nigra isolate 836 18S ribosomal RNA gene, partial sequence;
internal transcribed spacer 1, 5.8S ribosomal RNA gene
and internal transcribed spacer 2, complete sequence; and
26S ribosomal RNA gene, partial sequence
Length=750

Score = 558 bits (302), Expect = 6e-156
Identities = 371/399 (92%), Gaps = 25/399 (6%)
Strand=Plus/Plus

Query 51  TTTCCGTAGGTAAACTGCGGAAGGATC-TTGTCGATACCTGCCCAGACAGAACGACCTG  109

15 Reverse

gb|AF338492.1|AF338492  Juglans nigra isolate 836 18S ribosomal RNA gene, partial sequence;
internal transcribed spacer 1, 5.8S ribosomal RNA gene
and internal transcribed spacer 2, complete sequence; and
26S ribosomal RNA gene, partial sequence
Length=750

Score = 558 bits (302), Expect = 6e-156
Identities = 371/399 (92%), Gaps = 25/399 (6%)
Strand=Plus/Plus

Query 51  TTTCCGTAGGTAAACTGCGGAAGGATC-TTGTCGATACCTGCCCAGACAGAACGACCTG  109
Appendix 3: Consensus sequence multiple alignment with \textit{J. nigra} sequence acquired from Genbank

\begin{verbatim}
11              ---------------------------------------GGCAATTG--AATATGCGGCC 19
j               AAAGTCGTAACAAGGTTTCCGTAGGTGAACCTGCGGAAGGATCATTGTCGATACCTGCCC 60
3               ------------------------GTGAACCTGCGGAAGGATCATTGTCGATACCTGCCC 36
1               ------------------------------------------------------------
5               ------------------------------------------------------------
11              -GCAATTCGCCTT-----CCCGTAG--GTGAACCTGCGGAAGGATCATTGTCGATACCTG 71
j               AGCAGAACGACCTGTGAACATGTAATAATAACCTTCTGGGTGGGGGTGTAATGCCCCCTC 120
3               AGCAGAACGACCTGTGAACATGTAA---TAACCTTCTGGGTGGGGGTGTAATGCCCCCTC 93
1               ------------------------------------------------------------
5               ------------------AGACCGACCCTGGTGAACATGTAATAATACCTCTTTCCTGGTG 42
11              CCCAGCAACGGTTGGGAGGGCACGTTGAGATTTGCCCACTGCTCCTCGTGTGTGGGTTGG 131
j               CCAGAAAACGGTTGGGAGGGCACGTTGAGATTTGCCCACTGCTCCTCGTGTGTGG-TTGG 179
3               CCAAAAAACGGTTGGGAGGGCACGTTGAGATTTGCCCACTGCTCCTCGTGTGTGG-TTGG 152
1               ------------------------------------------------------------
5               TGGGGGGTGTACCAGGCCCCCTTCCCAAGGCACGTTGAGATTTGCCCCCTGCCCTCGT 102
11              GCCAAGG-AACTTAAACAAGGAG---TAACCACGGG-CGCCCA---GGAAACGG-TGTGC 242
j               GCCAAGG-AACTTAAACAAGGAG---TAACCACGGG-CGCCCA---GGAAACGG-TGTGC 282
\end{verbatim}
Appendix 4: Sample 1 consensus sequence and BLAST result

reverse
1 CAACCTCCGCTTCCGGTTGGAACCTGCGGAAGGATCATTGTCGATACCTGCCCAGCA 60
1For

reverse
1 TGCTAACGACCTGTGAACATGTAATAATAACCTTCTGGGTGGGGGTGTAATGCCCCCTCC 120
1For

reverse
1 GTGTCGATCCTCTCGT---CCTCCCCGATCGCCTAAATGAAACCCCCCGGTCGCTGCCC A 110
1For

reverse
1 GGGACCTAAAACAAGGAGTAACCACGGGCGCCCCCGGGAAACGCGCTGCTGC-CTG 168
1For

reverse
1 GTGACGTCTTTACCATGATACATAACGACTCTCGGCAACGGATATCTCGTGCTCTCATC 227
1For

reverse
1 CGATGAAGAACGTAGCGAAATGCGATACTTGGTGTGAATTGCAGAATCCCGCGAATCATC 287
1For

reverse
1 AGTCTTTGAACGCAAGTTGCGCCCGAAGCCATTCGGCCGAGGGC---ACGTCT---GCCTGG 344
1For

reverse
1 GGTGTCACGCATCGTTGC---CCTCCCCGATCGCCTAAATGAAACCCCCCGGTCGCTGCCC A 403
1For

reverse
1 GACATTGGCCTCCCGTGCGCTTTTGCTCGCGGTTAGCCTAAAAGTGAGTCCTAGGCGACG 463
1For

reverse
1 GGGGGAAAGACCAATTG---------------------------------------------- 338
reverse AGCGCCACGACAATCGGTGTTGAGAAACCTTCGTAACCCCTCTGCTGGTGCTGGCCCGTCGCT 523
1
1For

reverse GTGAAGGTTGCTGTCGACCCTATTGCTGTTTCCTGACTACATCCATGCCGACCCAGG 583
1
1For

reverse TCAGGGGGATTACCCGCTGAATTTAAGCATATCAATAAGCAGGAGGAAAGGGCGATTCG 643
1
1For

reverse GGCCGCTAAAATCTCAATCAGCCTCTATAGTGCTATTACATTCTACCTGCCGTA 698
1
1For

>Sample 1 consensus
CAACTTCGCCCTTTCGTAAGGTAACCTGCGGAAGGATCATCATGTGCATTGTCGATACCTGC
CCACGACTGCTAACCAGACCTGTGAAACATGTAAATAAAACCTTCTGGGTGGGGG
TGTAATGCCCCCTCCAGAAACGGG(t/g)GGGAGGGG(c/g)CAACGTTGAGATTGGC
CCACGACTGCTC(t/c)TCGGTGTG(g/t)GGTTGGGTCGATCCTCTCGTTCCCT(t/c)CC
CGATCG(a/g)ACAATGAACCCCCGGCGCGGTCTGCGCCAAGG(a/g)ACTTAAAA
CAAGGAGTGAAACCAGGCGCCCGCCCGG(3/g)AAACGGTGTGCGTGCTGGTG
ACGTCTTTACCATGACCTACGACCTCCTCAGGCAACCGATATCCTGTGCTCTCT
GCATCGATGGAAGAACGTAAGCGGAATGCGATACCTTGCTGTAAGTACGAATCT
CCGCGAATCATCGAGTCTTTGAGCAAGTTGCGCCCGAAGCCATTCGGCCG
AGGAGCAGCTG(t/c)TCGGCCTGGGTGCATGACCTGTTGCGCCCGACCCCAAAC
ACTTCTTACGCTGTGCGGGGGTGCGGCCGTGGGGAAGACGCGCAATCGGTGTG
AGAAAACCCCTCGTGACCCCTGTGTGTCGTCGCCCGCTGCTGTAAGGTTGCTGTCCTCG
ACCTATTGTTGCTGCTCTTGCGACTCTACCATCGCGACCCAGTGACGGCGG
ATTACCCCGCTGAATTTAAGCATATCATAAAGCGGAGGAAAGGCGATTCGCG
GCCGCTAAATCTCAATTCGACCTCTATAGTGACTGTCGATACCTAACATCGCG

>Sample 1 consensus Fasta
CAACTTCGCCCTTTCGTAAGGTAACCTGCGGAAGGATCATCATGTGCATTGTCGATACCTGC
CCACGACTGCTAACCAGACCTGTGAAACATGTAAATAAAACCTTCTGGGTGGGGG
TGTAATGCCCCCTCCAGAAACGGG(t/g)GGGAGGGG(c/g)CAACGTTGAGATTGGC
CCACGACTGCTC(t/c)TCGGTGTG(g/t)GGTTGGGTCGATCCTCTCGTTCCCT(t/c)CC
CGATCG(a/g)ACAATGAACCCCCGGCGCGGTCTGCGCCAAGG(a/g)ACTTAAAA
CAAGGAGTGAAACCAGGCGCCCGCCCGG(3/g)AAACGGTGTGCGTGCTGGTG
ACGTCTTTACCATGACCTACGACCTCCTCAGGCAACCGATATCCTGTGCTCTCT
GCATCGATGGAAGAACGTAAGCGGAATGCGATACCTTGCTGTAAGTACGAATCT
CCGCGAATCATCGAGTCTTTGAGCAAGTTGCGCCCGAAGCCATTCGGCCG
AGGAGCAGCTG(t/c)TCGGCCTGGGTGCATGACCTGTTGCGCCCGACCCCAAAC
ACTTCTTACGCTGTGCGGGGGTGCGGCCGTGGGGAAGACGCGCAATCGGTGTG
AGAAAACCCCTCGTGACCCCTGTGTGTCGTCGCCCGCTGCTGTAAGGTTGCTGTCCTCG
ACCTATTGTTGCTGCTCTTGCGACTCTACCATCGCGACCCAGTGACGGCGG
ATTACCCCGCTGAATTTAAGCATATCATAAAGCGGAGGAAAGGCGATTCGCG
GCCGCTAAATCTCAATTCGACCTCTATAGTGACTGTCGATACCTAACATCGCG

60
GAACGTAGCGAAATGCAGATACTTTGGTGTGAATTGCAAGAATCCCGCGAATCATCGAGTCTTTGAACGCAAGTTGCGCCCGAAGCCATTCGGCCGAGGGCCACGTCGTTCGGCCTGGGGTGTCAACGCATCGTTGCCAACCCCAACCCAAACACTTCTTACGCTGTGCGGGAAGCCAGCAACATCGGTGGTTGAGAAACCCCTCTGTAACCCGTCGTGTCGTGCCCGTCGCTGTGAAGGTGCTCCTCGACCCTATTGTGTCGTTCTTGCGAATCTACATCGGCGACCCCAGGTACGGCGGATTACCGCTGATTTAAGCATATCAATAAGCGGAGGAAGGGCGATTCGCGGCCGCTAAATTC

Closest Genbank match

gb|AF338492.1|AF338492  Juglans nigra isolate 836 18S ribosomal RNA gene, partial sequence; internal transcribed spacer 1, 5.8S ribosomal RNA gene and internal transcribed spacer 2, complete sequence; and 26S ribosomal RNA gene, partial sequence
Length=750

Sort alignments for this subject sequence by:

Percent identity

<table>
<thead>
<tr>
<th>Query start position</th>
<th>Subject start position</th>
</tr>
</thead>
<tbody>
<tr>
<td>Score = 828 bits (448), Expect = 0.0</td>
<td>Identities = 512/538 (95%), Gaps = 24/538 (4%)</td>
</tr>
<tr>
<td>Strand=Plus/Plus</td>
<td></td>
</tr>
</tbody>
</table>

Query 12  TTTCCTGATTGAGATCCCTCGAGACCTGCACTATTGCGATCTCCCGCCAGCATGCTAACGACC
| Sbjct 16  TTTCCCTGAGATCCCTCGCCACCTGTAATCTTGGCTGTGGTGGTGTTGGCCTGTA

Query 72  TGTGAACATGTAATAATAACCTTCTGGGTGGGGGTGTAATGCCCTCCCACGAAACGGG
| Sbjct 73  TGTGAACATGTAATAATAACCTTCTGGGTGGGGGTGTAATGCCCTCCCACGAAACGGG

Query 132  GGGGAGGCCACGTGGATCTGCCCACGTTCTCCTCCCTGGTGTTGGTGGTGTTGGTGTTGGTGCTGATC
| Sbjct 133  TGGGA-GG-GACGTTGAGATCTGCCCTCCCACGTTCTCCTCCCTGGTGTTGGTGGTGTTGGTGCTGATC

Query 192  CTCTGTCTCCCTCCCGATCGAGACCTGCACTATTGCGATCTCCCGCCAGCATGCTAACGACC
| Sbjct 186  CTCTGTCTCCCTCCCGATCGAGACCTGCACTATTGCGATCTCCCGCCAGCATGCTAACGACC

Query 252  TTAAACACAGGAAGATACCCACACCGGCCGCCGCCAGGAACCGGCTGCGCTGCGCTGGTGACG
| Sbjct 242  TTAAACACAGGAAGATACCCACACCGGCCGCCGCCAGGAACCGGCTGCGCTGCGCTGGTGACG

Query 312  TCTCTGTCCCTCGAGACCTGCACTATTGCGATCTCCCGCCAGCATGCTAACGACC
| Sbjct 298  TCTCTGTCCCTCGAGACCTGCACTATTGCGATCTCCCGCCAGCATGCTAACGACC

Query 372  AGAACGTAGCGAAATGCAGATACTTTGGTGTGAATTGCAAGAATCCCGCGAATCATCGAGTCTTTGAACGCAAGTTGCGCCCGAAGCCATTCGGCCGAGGGCCACGTCGTTCGGCCTGGGGTGTCAACGCATCGTTGCCAACCCCAACCCAAACACTTCTTACGCTGTGCGGGAAGCCAGCAACATCGGTGGTTGAGAAACCCCTCTGTAACCCGTCGTGTCGTGCCCGTCGCTGTGAAGGTGCTCCTCGACCCTATTGTGTCGTTCTTGCGAATCTACATCGGCGACCCCAGGTACGGCGGATTACCGCTGATTTAAGCATATCAATAAGCGGAGGAAGGGCGATTCGCGGCCGCTAAATTC
| Sbjct 357  AGAACGTAGCGAAATGCAGATACTTTGGTGTGAATTGCAAGAATCCCGCGAATCATCGAGTCTTTGAACGCAAGTTGCGCCCGAAGCCATTCGGCCGAGGGCCACGTCGTTCGGCCTGGGGTGTCAACGCATCGTTGCCAACCCCAACCCAAACACTTCTTACGCTGTGCGGGAAGCCAGCAACATCGGTGGTTGAGAAACCCCTCTGTAACCCGTCGTGTCGTGCCCGTCGCTGTGAAGGTGCTCCTCGACCCTATTGTGTCGTTCTTGCGAATCTACATCGGCGACCCCAGGTACGGCGGATTACCGCTGATTTAAGCATATCAATAAGCGGAGGAAGGGCGATTCGCGGCCGCTAAATTC

Query 432  TTGAACGTAGCGAAATGCAGATACTTTGGTGTGAATTGCAAGAATCCCGCGAATCATCGAGTCTTTGAACGCAAGTTGCGCCCGAAGCCATTCGGCCGAGGGCCACGTCGTTCGGCCTGGGGTGTCAACGCATCGTTGCCAACCCCAACCCAAACACTTCTTACGCTGTGCGGGAAGCCAGCAACATCGGTGGTTGAGAAACCCCTCTGTAACCCGTCGTGTCGTGCCCGTCGCTGTGAAGGTGCTCCTCGACCCTATTGTGTCGTTCTTGCGAATCTACATCGGCGACCCCAGGTACGGCGGATTACCGCTGATTTAAGCATATCAATAAGCGGAGGAAGGGCGATTCGCGGCCGCTAAATTC
| Sbjct 417  TTGAACGTAGCGAAATGCAGATACTTTGGTGTGAATTGCAAGAATCCCGCGAATCATCGAGTCTTTGAACGCAAGTTGCGCCCGAAGCCATTCGGCCGAGGGCCACGTCGTTCGGCCTGGGGTGTCAACGCATCGTTGCCAACCCCAACCCAAACACTTCTTACGCTGTGCGGGAAGCCAGCAACATCGGTGGTTGAGAAACCCCTCTGTAACCCGTCGTGTCGTGCCCGTCGCTGTGAAGGTGCTCCTCGACCCTATTGTGTCGTTCTTGCGAATCTACATCGGCGACCCCAGGTACGGCGGATTACCGCTGATTTAAGCATATCAATAAGCGGAGGAAGGGCGATTCGCGGCCGCTAAATTC
Appendix 5: Sample 3 consensus sequence and BLAST result

Reverse
3
GTGAACCTGCGGAAGGATCATTGTCGATACCTGCCCAGCAGAACGACCTGTGAACATGTA 60

Reverse
3
ATAACCTTCTGGGTGGGGGTGTAATGCCCCCTCCCAAAAAACGGTTGGGAGGGCACGTTG 120

Reverse
3
AGATTTGCCCACTGCTCCTCGTGTGTGGTTGGTCGATCCTCTCGTTCCCTTCCCGATCGA 180

Reverse
3
------------------------GCCAGGGAACTTAACCAGGGAGGTAACCACGGGGCG 36

Reverse
3
ACAACGAACCCCGGCGCGGTCTGCGCCAAGGAACTTAAACAAGGAG-TAACCACGGG-CG 238

Reverse
3
CCCCCGGGAACCGGTTGCGTGGTGTGGGGACGTCTTTACCAAGATACATAACGAC 96

Reverse
3
CCCG--GGAAACGG-TGTGCG-TGTCG-TTGGTGACGTCTTTACCAAGATACATAACGAC 293

Reverse
3
TCTCGGCAACGGATATCTCGGCTCTCGCATCGATGAAGAACGTAGCGAAAATGCGATACT 156

Reverse
3
TCTCGGCAACGGATATCTCGGCTCTCGCATCGATGAAGAACGTAGCGAAA-TGCGATACT 352

Reverse
3
TGGTG--TGAATTGCAGAATCCCGCGAATCATCGAGTCTTTGAACGCAAGGTTGCGCCCG 216

Reverse
3
TGGTG--TGAATTGCAGAATCCCGCGAATCATCGAGTCTTTGAACGCAAG-TTGCGCCCG 409

Reverse
3
AAGCCATTCGCCGAGGGGACGCGCTGCCTGGGTGTCACGCATCGTTGCCCCAACCCCAAA 276

Reverse
3
AAGCCATTCGCCGAGGGGAGGCGCCCTGCGCTTGCGAGGACGACATCGTGCCCCCAACCCCAA 469
Sample 3 consensus sequence for Thesis

GTGAACCTGCGGAAGGATCATTGTCGATACCTGCCCAGCAG
AACGACCTGTTGAAACATGTAATAAACCTCTTCGGTGAGGGGATG
ATATGCCCCCTCCCAAACACCGTTGGGAGGGGATCGT
TTTGCCCATGCTCTGCTGTGTTGTTGAGTCACTCGCTTGCTTGCTTGAGGCTAG
CTCGTGACCCGTCGTGTGTCGCCCGTCGCTGTGAAGGTGCTCCTCGACCCTATTGCTG
TTCTTGCGACTCTACCATCGCGACCCCAGGTCAGGCGGGATTACCCGCTGAATTTAAGCA
TATCAATAACGGGAGAAGGCGCAATCGGTTTAAAACAAATGCA

Fasta:

>sample 3 consensus
GTGAACCTGCGGAAGGATCATTGTCGATACCTGCCCAGCAG
AACGACCTGTTGAAACATGTAATAAACCTCTTCGGTGAGGGGATG
ATATGCCCCCTCCCAAACACCGTTGGGAGGGGATCGT
TTTGCCCATGCTCTGCTGTGTTGTTGAGTCACTCGCTTGCTTGCTTGAGGCTAG
CTCGTGACCCGTCGTGTGTCGCCCGTCGCTGTGAAGGTGCTCCTCGACCCTATTGCTG
TTCTTGCGACTCTACCATCGCGACCCCAGGTCAGGCGGGATTACCCGCTGAATTTAAGCA
TATCAATAACGGGAGAAGGCGCAATCGGTTTAAAACAAATGCA
<table>
<thead>
<tr>
<th>Query</th>
<th>Sbjct</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>10</td>
</tr>
<tr>
<td>59</td>
<td>70</td>
</tr>
<tr>
<td>118</td>
<td>130</td>
</tr>
<tr>
<td>178</td>
<td>190</td>
</tr>
<tr>
<td>238</td>
<td>246</td>
</tr>
<tr>
<td>298</td>
<td>299</td>
</tr>
<tr>
<td>358</td>
<td>358</td>
</tr>
<tr>
<td>418</td>
<td>414</td>
</tr>
<tr>
<td>478</td>
<td>474</td>
</tr>
<tr>
<td>538</td>
<td>527</td>
</tr>
<tr>
<td>598</td>
<td>587</td>
</tr>
<tr>
<td>658</td>
<td></td>
</tr>
</tbody>
</table>

Closest Genbank Match

%gb|AF338487.1|AF338487 Juglans microcarpa isolate 108 18S ribosomal RNA gene, partial sequence; internal transcribed spacer 1, 5.8S ribosomal RNA gene and internal transcribed spacer 2, complete sequence; and 26S ribosomal RNA gene, partial sequence
Length=735

Score = 1190 bits (644), Expect = 0.0
Identities = 717/748 (95%), Gaps = 26/748 (3%)
Strand=Plus/Plus
Appendix 6: Sample 5 consensus sequence and BLAST result

reverse
5
TAGCGCACGTGGAATTGTAATACGACTCACTATAGGGTTCGAATTGAATTTAGCGGCCGC 60

reverse
5
GAATTCGCCCTTTCCTCCGCTTATTGATATGCTTAAATTCAGCGGGTAATCCCGCCTGAC 120

reverse
5
CTGGGGTCGCGATGGTAGAGTCGCAAGAACGACACAATAGGGTCGAGGAGCACCTTCACA 180

reverse
5
GCGACGGGCGACACACGACGGGTCACGAGGGTTTCTCAACCACCGATTGTCGTTGCGCTC 240

reverse
5
GTCGCCCTAGGAACCTACTTTTTAGGTTAACCCGAGGACACCAACTTGGGCTCAGGCTAAT 300

reverse
5
TCCCCGCACCACGCACAGCGTAAGAAGTGTTTGGGGTTGGGGCAACGATGCGTGACACCC 360

reverse
5
AGGCAGACGTGCCCTCGGCCGAATGGCTTCGGGCGCAACTTGCGTTCAAAGACTCGATGA 420

reverse
5
GTTCGCGGGGA-TTCTGCAATTCNACACCAAGTATCGCATTTCGCTACGTTCTTCATCGAT 478

reverse
5
GCGAGAGCCGAAATATCCGTTGCCGAGAGTCGGTATGTATCATGGTAAAGACGTCACCAA 538

reverse
5
CGACACGCACACCGTTTCCGGGGCGCCCGAGGTGACTCCCTGGGTAAGTTCCTTGGCGCA 597

reverse
5
GACCGCGCGGGGTTACTTGGCTAAGGAGAGAGGTACGACCGACCACACA 656

reverse
5
CGAGGGGCAGGGGGCAAATCTCAACGTGCCCTCCCAACGTTTTCTGGGAGGGGGCAGGT 711

reverse
5
ACACCCCCACCCAGGAAGGGTATTATTACATGTTCACAGGTCGGTCGTTCTAAATGGGTCAGG 769
Sample 5 consensus sequence

TAGCGCACGTGGAATTGTAATACGACTCACTATAGGGTTCGAATTGAATTTAGCGGCCGC
GAATTCGCCCTTCTCCGCTGGTATTAGTATGGTTCTTTATTACATTCAGCGGGTACGACACTCTACAG
GCCGCACGGGCAGACACGACACCCAGCTTACCGGCAATGAGATGTACAGTACTTTAGGCTAACCGCGAGCA
CAAGCGCACGGGAGGCCAATGTCTTCCCCGCACCCCGCACAGCGTAAGAAGTGTTTGGGGTTGGGCAACGATGCGTG
ACACCGAGCGAGACGTGCCCTCGGCCGAATGGCTTCGGGCGCAACTTGCGTTCAAAGACTCGATGATTCGCGGGGAT
T(c/t)/(t/c) TGC(a/g) ATTCA(a/c)(a/c)(a/c) C(a/c)A(g/a) GTATCGCATTTCGCTACGTTCTTCATCGAT
GCGAGAGCCGAAATATCCGTTGCCGAGAGTCGtTATGTATCATGGTAAAGACGTCACCAACGACACGCACACCGTTTCC
GGGGCGCCCG(t/a)GGGTGTACTCTCATGGGCAGACCGCGC(c/g)GGGGTTCATTGTTCGATCGGGAAGGGAACGAGA
GGATCGACCACCACACAGGCAAGGGCCAGGGGCAAATCTCAACGTGC(c/t)TGGGA(a/g)GGGGCC(t/g)GTAC(a/c)
CCCCCA(c/a)CCAGGAA(g/a)GGTATTATTACATGTTC(a/c)CAGGGTCGGTCT

Closest Genbank Match

>gb|AF338491.1|AF338491  Juglans nigra isolate 834 18S ribosomal RNA gene, partial sequence; internal transcribed spacer 1, 5.8S ribosomal RNA gene and internal transcribed spacer 2, complete sequence; and 26S ribosomal RNA gene, partial sequence
Length=750

Score = 1018 bits (551), Expect = 0.0
Identities = 604/627 (96%), Gaps = 17/627 (2%)
Strand=Plus/Minus

Query  77   CCGTTATAGTATGTTAATTTTCGTACGACCAATAGGGTTCGACAGCTCCTACGGCT

Sbjct  749  CCGTTATAGTATGTTAATTTTCGTACGACCAATAGGGTTCGACAGCTCCTACGGCT

Query  137   AGAGTGCAAAAGACACCAATAGGGTTCGAGGACACCTTTACACGACAGGCGACACAC

Sbjct  689  AGAGTGCAAAAGACACCAATAGGGTTCGAGGACACCTTTACACGACAGGCGACACAC

66
Appendix 7: Sample 11 consensus sequence and BLAST result

Reverse
11  TCCTCTGTTTAAACCAATTCGCTCCTCTCCAGCTTATTGATATGCTTAAATTCAGCGG  60

Reverse
11  GTAATCCGCCTGACCTGGGGTCGCGATGGTAGAGTCGCAAGAACGACACAATAGGTTCG  120

Reverse
11  AGGAGCACCTTCACATGCACCGGACACAGGNTACGAGGGTATCTCAACACC  180

Reverse
11  GATTGTCCGTGCCGTGTCGGTCTAGTCACTTATTGCAACGGGAGCAAAACCGCA  240

Reverse
11  ----------------TCCCCGCACCCGCAACACGTAAGAAGGTGGGTTGGGGGCCG  300

Reverse
11  ****************** ***********

Reverse
11  ACGATGCGTGACACCCAGGCAGACGTGCCCTCGGCCGGAATGGCTTCGGGCGCAACTTGC  359

Reverse
11  **************************************

Reverse
11  GTTCAAAGACTCG-ATGATTCGCGGGATTCTGCAATTCACACCAAGTATCGCATTTTCGC  419

Reverse
11  ************ **********************************************

Reverse
11  CTACGTTTCTTATCGATGCGAGAGCGAGATATCCGTTGCCGAGAGTCGTATGTATCA  479

Reverse
11  ***************** ****************

Reverse
11  TGGTAAAGACGTCACCAACGACACGCACACCGTTTCCGGGGCGCCCGTGGTTACTCCTTG  539

Reverse
11  ******************* ******************************* ********

Reverse
11  TTTAAGTTTGCACGAGACCCAGGGTTATGTTGATCGAGGACACAGGTTGAGTATGAC  594

Reverse
11  ***************** ***   ****************

Reverse
11  GGATCGACCAACCACACACGAGGAGCAGTGGGCAAATCTCAACGTGCCCTCCCAACCGTT  607

Reverse
11  ---------------

Reverse
11  TTGGAGGGGGGATTACACCCCCCAACCCCAAGGGTTATGTTGATCGAGGACACAGGTTGAGTATGAC  607

Reverse
11  GCTGGGCAGGTATCGACAATGATCCTTCCGCCAGGTTCACCTACGGGAAGGCGAATTGCGG  520

Reverse
11  CCGCATATTCAATGCC  537
Sample 11 consensus for thesis:

```
TCCTCCTGTTTAAACCAATTCGCCCTTTCCTCCAGCTTATTGATATGCTTAAATTCAGCGG
GTAATCCCGCCTGACCTGGGGTCGCGATGGTAGAGTCGCAAGAACGACACAATAGGGTCG
AGGAGCACCTTCACAGCGACGGGCGACACACGACGGGNTCACGAGGGTTTCTCAACCAC
GATTGTCGTGGCGCTCGTCGCCTAGGACTCACTTTTAGGCTAACCGCGAGCAAAAGCGCACGGGAGGCCAATGTCTTCC
CCGCACCCCGCACAGCGTAAGAAGTGTT(t/g) GGGGTTGGGGC(a/g)ACGATGCGTGACACCCAGGCAGACGTGCCCTCGGCCGGAATGGCTTCGGGCGCAACTTGC
GTCAAAGACTC(g/t)TATGATTCGCGGGATTCTGCAATTCACACCAAGTATCGCA TTT(c/g)GC
CTACGTTCTTCATCGATGCGAGAGCCGAGATATCCGTTGCCGAGAGTCGTTtATGTATCATGGTAAAGACGT(c/t)ACCAA
CGACACGCACCCCGGCGGCCGCGCTGGTACTCCTTGTTTAAGTTCCTTGGCGCA(g/a) ACC(g/o)(g/c)(g/c)CCGGGGTTCATTGTTC(g/c) ATCGGGGAAGGGAACGAcgA(a/g) GATC(g/c)ACCA(a/c)
CCACACACGAGGAGCAAGTGCGCATAATCTCAACTGCCCTCCAACCGTGT
GCTGGGCAAGTATCGCAAAATGATCCTCCAGCTTATTGATATGCTTAAATTCAGCGG
```

Sample 11 Fasta

```
TCCTCCTGTTTAAACCAATTCGCCCTTTCCTCCAGCTTATTGATATGCTTAAATTCAGCGGGTAATCCCGCCTGACCTGGGGTCGCGATGGTAGAGTCGCAAGAACGACACAATAGGGTCGAGGAGCACCTTCACAGCGACGGGCGACACACGA
```

Closest Genbank Match:

```
>gb|AF338492.1|AF338492
Juglans nigra isolate 836 18S ribosomal RNA gene, partial sequence;
internal transcribed spacer 1, 5.8S ribosomal RNA gene
and internal transcribed spacer 2, complete sequence; and
26S ribosomal RNA gene, partial sequence
Length=750
```

```
Sort alignments for this subject

<table>
<thead>
<tr>
<th>identity</th>
<th>E value</th>
<th>Score</th>
<th>Percent</th>
</tr>
</thead>
<tbody>
<tr>
<td>start position</td>
<td>Query start position</td>
<td>Subject</td>
<td></td>
</tr>
</tbody>
</table>

Score = 1051 bits (569), Expect = 0.0
Identities = 629/653 (96%), Gaps = 23/653 (3%)
Strand=Plus/Minus

Query 34  GCTTATTGATATGCTTAAATTCAGCGGTAATCCCGCCTGACCTGGGGTCGCGATGGTAG  93
  |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   |||||||||   ||||||##

69
Appendix 8: Sample 15 consensus sequence and BLAST result

reverse  ------------------TGCATTCACACCAAGGTTATCGC--ATTTCGCTACGGTTTCT 40
15 ACTATATCATCATATCGGTGGCAATGATATAGACGGCCGGAATTTCGCCCTTTCCTCC 60

---TCATCGATGCAGAGACCAGAGATATCTGCTCCGAGGAGGTTCGCTTATGCA 90
*** *   *   **     ***  *******       **

reverse  --TCATCGATGCG---AGAGCCGA--GATATCGTTGCCCCAGA--GAGTTGCTTATGCA 148
15 GGTGCGCAAGAGACACACATAGGGTGCTGGGAGCAGAACCTCAACAGAGGACAGAGAGA 179

***   **  ***   * *  * *  * * **    ***    * **  * *   **

reverse  TCATGGTAAAGACGTT-CACCAACGACACGCACACCGTTTTCCGGGGTCGTCCCGT-TGG 148
15 GAGTCGCAAGAACGACACAATAGGGTCGAGGAGACAACCTTCAACAGACGAAACGGGACGA 179

***    ** *   *     ****   * *  * * ***  * **   **  *** **

reverse  TTTACTCCTTGGTTTAA----GGTTCCCTTG-GTCGAGACCAGCGGCCGGGTTTCATGTT 203
15 CAACACAGCGGTTCAACAGGGTTGTTTCTCAACAGCAGTGTCTGGCGCCT-----CG 234

***   * ***  *  * *     * *  *   * *  ** *          *  **

reverse  TCCCTTTAGTGGAGGAGGAGGAGGTCGACCAACCACCCAGAGAGA-ATCGTGGGCAAAT 306
15 GCCAATTGGCTGGTTTTCCCCGCAGCAGCAAGAGGTAAAGAGGTTTGGGT--GAGTGGGCAAT 321

**   ** *     **

reverse  ------------------------------------------------------------- 331
15 GCCGAGAGTCGTTAATGCATCATGGTAAAGAACGTCACCCAACGAACACGCCACCACCGT 593

reverse  ------------------------------------------- 382
15 TTCCCGGGCGCCGTGGGTAACTCCTTGGTTAAGTCCTGGCGAC 636

No significant similarity