Xylaria Primers for Phylogenetic Reconstruction

Alicia Ly, Environmental Studies*

ABSTRACT

Xylaria is a genus of Ascomycete fungi that is comprised of an estimated 400 species. These species are important to the ecosystems in which they inhabit because they contribute to the decomposition of wood, which recycles lignin and cellulose. Though the morphological traits of this genus have been widely studied over the years, their molecular phylogenetics have not been extensively explored due to the historical unavailability of species genomes. Now that complete genomes have become available, there have been an increasing number of ribosomal Internal Transcribed Spacer (ITS) sequence studies in recent years. However, ITS is too variable to consistently construct accurate phylogenetic trees, so *Xylaria* phylogeny is still being researched. This study explores the use of single-copy genes as an alternative to ITS methods. Here we present three primer pairs for different single-copy genes for use during future phylogenetic evaluation of *Xylaria*. These primers, Splicing Factor PRP43 F1054/R1691, MCM5 F409/R1059 and 40S Ripro S3Ae F47/R605, are more reliable than ITS sequences for deeper levels of analysis due to less variability in their sequences. This study focuses on *Xylaria*, while utilizing related species within in the Xylariaceae and Sordariaceae families, providing a promising start to the complete reconstruction of the *Xylaria* phylogenetic tree.

1. INTRODUCTION

Xylaria, the type genus of the family Xylariaceae (Xylariales, Sordariomycetes), consists of hundreds of species from around the world (Rogers 2000). Ascomycetes contribute to the decomposition of wood and are important to the natural recycling of lignin and cellulose in many ecosystems (Whalley 1996). Species are often found living on dead angiospermous wood, but have also been discovered on other substrates, including gymnosperms, non-woody angiosperms, living hosts, soil, litter and stool, and among insects (Peláez et al. 2008). A number of the species are common endophytes, which flourish in the foliage of trees and shrubs (Davis et al. 2003; Peláez et al. 2008). Although they are distributed worldwide, *Xylaria* have been observed most abundantly in tropical zones (Rogers 2000; Whalley 1996). There are also sizeable areas of the world that have not been surveyed for *Xylaria (*Stadler 2010). Further exploration of such areas could lead to great advancements in the understanding of the genus.

*Alicia Ly graduated with honors from the Environmental Science program at the University of Oregon. Her research interests include ecology and species conservation, as well as sustainable technologies. She hopes to contribute to finding creative solutions to environment issues worldwide. Please direct correspondence to alicial@uoregon.edu.
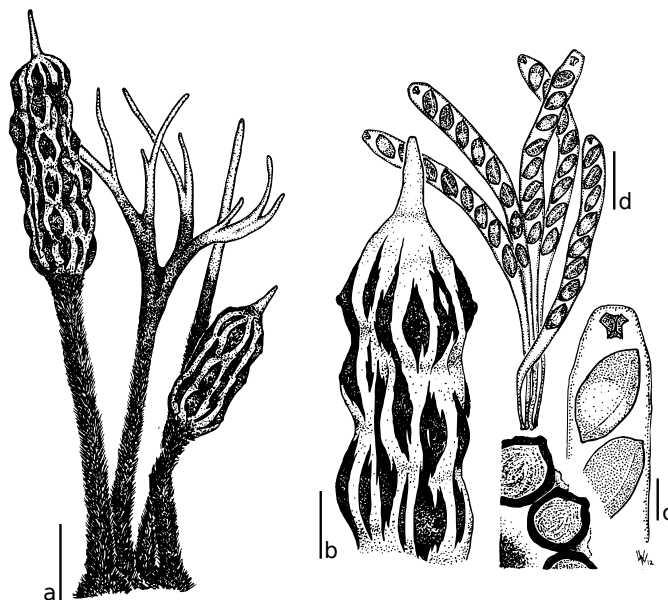
Figure 1. Drawing of one of the target species in this study, *Xylaria apiculata*, by Roo Vandegrift. (Bars: a = 2 mm; b = 1 mm (including stromatal section); c = 10 μm; d = 50 μm).

Basic characteristics that most *Xylaria* exhibit include perithecial (flask-shaped) ascomata and generally erect stromata (Lee et al. 2000; Rogers 2000), shown in the drawing of *Xylaria apiculata* in Figure 1. They also have dark ascospores with a hyaline germination slit and cylindrical, monoseriate asci that have an amyloid apical ring (Lee et al. 2000). However, there is an enormous range of variation between other characteristics of the species (Rogers 2000; Whalley 1996). Though there are diverse traits, there are also a number of species that are virtually indistinguishable using only their gross morphologies (Rogers 1979). Due to this unreliability of gross morphology, many specimens have been identified incorrectly and evaluation of species' relationships has proven to be difficult (Rogers 1979, 2000). In fact, there are approximately 700 named *Xylaria* species, but probably only about 400 genuine species, due to the re-description based on variation in non-informative characters (Ju, personal communication). Unfortunately, the genus has never been monographed. Characterization based solely on morphology has led to irregular, incorrect, and wholly artificial groupings of species within *Xylaria* (Dennis 1956), as well as the family Xylariaceae (Whalley 1996). These challenges indicate the need for identification and evaluation of species' relationships on a molecular level, in addition to studies of morphology.

Xylariaceae also includes, but is not limited to, the widely accepted genera *Daldinia*, *Hypoxylon*, *Nemania* and *Kretzschmaria* (Tang et al. 2009). This study focuses on *Xylaria*, but utilizes species from all of these genera to consider evolutionary history. The analysis also makes use of species from the genus *Neurospora*, another Sordariomycete, but from a different family and order (Sordariaceae, Sordariales) (Cai 2006; Tang et al. 2009). Past studies of *Xylaria* suggest it is a large and complex genus (Hsieh et al. 2010; Peláez et al. 2008), indicating the need for the inclusion of many species in an analysis of phylogenetic relationships. In fact,

previous studies have shown that *Xylaria* is a paraphyletic genus, meaning it is not inclusive of all descendants of a single common ancestor (Hsieh et al. 2010). A number of species, including some *Kretzschmaria* and *Nemania* species, have been observed amongst groups of *Xylaria* in past phylogenetic studies (Hsieh et al. 2010; Stadler 2013). There is a great need to clarify these relationships to properly delineate the genus.

The use of DNA sequences of ribosomal genes is now crucial to confirming the identification of a species (Dentinger et al. 2010). Examining genetic variation of these sequences can allow the differentiation between species that look almost exactly the same. Additionally, sequences of DNA allow the consideration of evolutionary histories more accurately, because they utilize huge data sets of characters for analysis (Lee et al. 2000). By studying the history of speciation, the precise processes of evolution can be assessed and relative time estimates of evolutionary events can be formulated. Findings such as these could lead to important advancements in mycology and evolutionary biology, which has pushed scientists to utilize DNA sequences for most current taxonomic and phylogenetic studies. More informative loci (DNA sequences) for *Xylaria* were never examined previously, because their genomes were not available (Lee et al. 2000). Now that relevant genomes have been published, primer pairs can be designed and used to amplify specific DNA sequences. Primers, unpaired strands of DNA, serve as starting points for DNA synthesis during polymerase chain reactions. Polymerase chain reaction (PCR) is a method that rapidly alternates heating and cooling cycles to amplify a specific DNA sequence (Dentinger et al. 2010; Folmer 1994).

Recently, the use of the nuclear ribosomal DNA internal transcribed spacer (ITS) region has been quite popular for species analysis (Larena et al. 1999). Numerous prominent scientific papers have been published primarily or exclusively using ITS data (Sánchez-Ballesteros 2000; Triebel et al. 2005; Peláez et al. 2008). In fact, a paper by Schoch et al. (2012) claimed it should be "the universal DNA barcode marker for fungi". However, the variability within ITS, which includes intraspecific variation, inhibits this approach from being entirely reliable for phylogenetic inferences (Álvarez and Wendel 2003; Kiss 2012). Furthermore, the variability of the ITS region makes it difficult or impossible to align accurately. This variation has developed because ITS is a freely evolving intron. Although the use of ITS sequences has been very useful for species identification, other methods should be explored for phylogenetic analyses (Álvarez and Wendel 2003). Of the additional primers that have been developed for other Ascomycetes, those that have been tried with *Xylaria* were not successful. For example, β-tubulin was used throughout the studies of *Hypoxylon* species conducted by Hsieh et al. (2005), but β-tubulin is often ineffective with *Xylaria* species because two active copies of the gene are present. It is not possible to consistently sequence the same copy each time it is tested. Therefore, accurate comparisons between genes of different species cannot be made.

An alternative to ITS and other methods is the use of single-copy nuclear genes (Álvarez and Wendel 2003; Salvador et al. 2014). Single-copy genes streamline homologous comparisons, have codons that minimize ambiguities in alignments, and can supply copious amounts of data based on 3rd position variability (Álvarez and Wendel 2003). Plus, they have lower likelihood of homoplasy (characteristics that species share, but did not evolve from a common ancestor) than

ITS data (Álvarez and Wendel 2003). Several scientists argue that single-copy genes are the most accurate way to analyze species' histories and are the future of phylogenetic reconstruction (Álvarez and Wendel 2003; Capella-Gutierrez et al. 2014).

Now that *Xylaria* genomic sequences are available, there is an opportunity to design primer pairs specifically for single-copy *Xylaria* genes. With this advancement, my team composed primer pairs using all available Xylariaceae genomes. During this study, these primers were tested on a panel of *Xylaria* and related species in order to determine if any would be informative for accurate evaluation of *Xylaria* evolutionary relationships and reconstruction of their phylogenetic tree.

2. METHODS

2.1 DESIGNING AND TESTING PRIMER SETS

A preliminary list of potential primer sets was assembled using all available Xylariaceae genomes. These included published genomes (three *Hypoxylon*, one *Daldinia*; data provided by the Joint Genomics Institute) and one unpublished genome, *Xylaria hypoxylon*, supplied by Dr. Joseph Spatafora, a professor at Oregon State University. Though no specific formula was used for selection, many factors were considered when designing the primers. Primer pairs that produced long sequences, comprised of approximately 600-700 base pairs, with no significant gaps were selected. Primers possessing more guanine and cytosine bases at the 3' end of the strand were considered advantageous, because they bind more securely to DNA than adenine and thymine. In addition, low levels of degeneracy were selected for. Degenerate sequences contain multiple potential bases at some positions (Compton 1990). Degeneracy in primers increases the likelihood of binding at sites that differ between genomes, but consequently reduces the specificity of the primer. Lastly, potential for binding among/within primers of a single pair was avoided when possible. Considering all of these factors, a list of 30 possible primer pairs that had potential to work with *Xylaria* was created. The list of all primer pairs tested can be found in Table 2, in the appendix.

We began initial tests of these 30 primer pairs using a set of genomic DNA from 4 *Xylaria* species: *X. fabelliforme (cubensis)*, *X. anisopleura*, an undescribed species from Taiwan herein denoted as *X. sp. nov. 1*, and *X. sp. 77080301*. PCRs were run to examine the ability of each primer set to successfully amplify their target genes. These four species were selected because previous studies showed that they were significantly different from one another at the ITS locus. The list of primer pairs that had potential was quickly narrowed down based on these preliminary tests. We then expanded the panel of genomic DNA for the effective primer pairs to include 17 species of *Xylaria* and their relatives. After many PCRs, the primer pairs that yielded the most promising results were selected. A 96-well plate of these samples was sent to Functional Biosciences (Madison, WI) for sequencing.

2.2 SEQUENCE DATA ANALYSIS

Sequences were analyzed (forward and reverse) for 6 single-copy genes from the majority of the trial species. The raw sequences were imported into the program Geneious (Biomatters, Ltd.), where most sequence editing and phylogenetic analysis was conducted. The forward and reverse sequences were trimmed and aligned using the default Geneious alignment, which automatically determines sequence directions. Minor edits were made manually and consensus sequences (with 100% threshold) were subsequently created for each species at each locus. Generating the consensus sequences served as a data checking method to ensure the sequences were as accurate as possible. Many of the MCM7 F9 (forward) sequences were not viable (likely due to the nature of the primer), so the reverse sequences alone were relied on for the analyses.

2.3 PHYLOGENETIC TREES

All of the consensus sequences for each gene were aligned. MUSCLE alignments were used for all multiple alignments, which required me to manually reverse-complement the reverse sequences. These 6 alignments all involved 13-17 species sequences, depending on the success of the amplification and sequencing. We collected additional sequences from published species genomes in order to accurately root the resulting gene trees. A *Daldinia*, 3 *Hypoxylons* and 3 *Neurosporas* were added (data provided by JGI). These references were added to the alignment of the genes that had been sequenced. Then, phylogenetic trees were made from the resulting alignments. The three *Neurospora* species served as the root for all trees created, just as they did in the Xylariaceae phylogenetic studies conducted by Peláez et al. (2008). This outgroup method was also used in the study by Walker et al. (2010). *Neurospora* serves as an ideal outgroup, because they always cluster independently of *Xylaria* species (because they do not belong to the Xylariaceae family), yet still have deep evolutionary relationships with them (Cai et al. 2005).

The phylogenetic trees were constructed using both maximum likelihood (ML) and neighbor-joining (NJ) methods. Maximum likelihood models are regarded as amply robust, because they employ all sequence information and evaluate all different tree topologies (Huelsenbeck 1995). Maximum likelihood trees are built to illustrate the topology that has the highest probability of resulting throughout numerous re-samplings. NJ trees are considered less reliable by much of the scientific community, since the bottom-up clustering method can reduce sequence information. NJ trees were created for comparison, but ML trees were used in later analyses. All 12 of the phylogenetic trees (ML and NJ for all 6 genes) are displayed in Figures 8-19, in the appendix.

For all of the maximum likelihood trees generated in this study, 500 bootstrap samples were used; all phylogenetic trees present bootstrap values at the branch nodes (Figures 8-19). Bootstrap values can be defined as the percentage of times the same node is produced throughout resampling (Felsenstein 1985). The scale for each of the trees is in substitutions per site. In phylogenetic trees such as these, branch lengths represent an amount of genetic change (Sanderson 2002).

To create a single phylogenetic tree based on all of the genetic data, we concatenated (linked together in succession) all sequences and built an alignment. We did this first with all 6 genes that were sequenced, in 17 total species. Later, we created a concatenated alignment using only 4 genes (from 21 species). From these concatenated alignments, we generated new, composite ML trees, which were used in the analysis of the primers' phylogenetic informativeness.

2.4 ANALYSIS OF PHYLOGENETIC INFORMATIVENESS

In order to evaluate phylogenetic informativeness of the primer pairs, the trees had to be converted to ultrametric trees (Walker et al. 2012). These have equal path lengths from the root to the branch tips of the tree. Consequently, the trees represent relative evolutionary time, rather than a relative measure of evolutionary change. The relative timescale is calculated based on the assumption of a constant mutation rate. A nonparametric rate smoothing method (Sanderson 1997) was conducted in the program R (version 3.0.1), to create the ultrametric trees.

The resulting ultrametric trees, along with their corresponding sequence alignments, were input into PhyDesign to assess phylogenetic informativeness of the sequences generated from each primer pair (Lopez-Giraldez and Townsend 2011; Walker et al. 2012; http://phydesign.townsend.yale.edu/). Each output provided a graph of phylogenetic informativeness positioned below an ultrametric tree; these graphs matched directly with the relative time-correlated measure of the ultrametric trees. Phylogenetic informativeness is a unit-less measure of the amount of information per base pair contributing to the overall ultrametric tree. Useful primer pairs provide a large amount of information for the ultrametric tree, so the graphs produced show broad curves with high levels of phylogenetic informativeness, especially towards the species level.

3. RESULTS

3.1 PCR AMPLIFICATION

Of the original 30 primer pairs tested, 10 were abandoned promptly after they proved to be unsuccessful at amplifying 3 of the 4 test species' genes. Meanwhile, 9 primer pairs successfully amplified genes in all test species without any modification. The remaining 11 primer pairs showed potential to be successful after adjustments to the PCR process. After increasing the panel of DNA to 17 species for all of the potentially successful primers, we focused further on only 6 primer pairs that produced the most effective results (Table 1). Each primer set amplified genes in at least 15 of the 17 species. The species that did not amplify were primarily not *Xylaria*, but one of the related species.

Table 1. Top 6 primer pairs that produced the highest amount of amplification of the test species' genes and were selected to be evaluated for their phylogenetic informativeness.

Primer	Forward	Reverse
40S Ripro S3Ae	47	605
MCM5	409	1059
Splicing Factor PRP 43	1054	1691
MCM7	665	1209
MCM7	9	595
MCM5	1188	1970

3.2 PHYLOGENETIC ANALYSIS

Comparison of the maximum likelihood and neighbor-joining trees did not show significant conflicts in the relationships between species. We found that neighbor-joining trees showed significantly more polytomy, or unresolved branches. Consequently, the trees would have been difficult to use for further analyses, because their inferences were imprecise. ML trees often resolved many of the nodes, though it is important to note that lower bootstrap values sometimes resulted. We accepted nodes with bootstrap support of 50% or greater. Due to the knowledge of their robust statistical process and the observations of more resolved branches, ML trees were relied on for additional analyses.

3.3 PHYLOGENETIC INFORMATIVENESS

The results from PhyDesign provided a well-defined graph, displaying phylogenetic informativeness levels of each primer pair. The first graph, assessing all 6 sequences from 17 species' genes, showed an unexpected spike towards the tips of the tree (shown in Figure 7, in the appendix). By examining the sequences in Geneious, the team realized that the loci targeted by the two primer pairs showing this irregularity both contained introns. Introns are noncoding regions that are spliced out of mRNA. Thus, those regions can have a huge range of variability, since countless mutations could have surfaced over the course of history without affecting organisms. The variability makes them difficult to align accurately and consistently. The spike in the graph reinforced the conclusion that there was a tremendous amount of variation in the sequences. Though the sequences seemed to indicate a high level of informativeness at the species-level, it was uncertain whether they would remain useful for deeper-level analyses.

The intron sections were removed from the sequences in Geneious and a new alignment and tree was created to repeat the evaluation in PhyDesign. There was a significant reduction in the spike at the branch tips after intron removal (Figure 2). Though the spike was not completely smoothed out, the reduction allowed for a better observation of the shape of an overall broad curve, as expected. The 40S Ripro S3Ae sequences became significantly less informative with the intron removed. However, the intron constituted over 20% of the entire sequence. Thus, removing the intron significantly shortened the sequence, reducing the amount of information it could provide. Additional evaluation demonstrated that it actually has a much higher level of informativeness when the intron was not removed from the sequence.
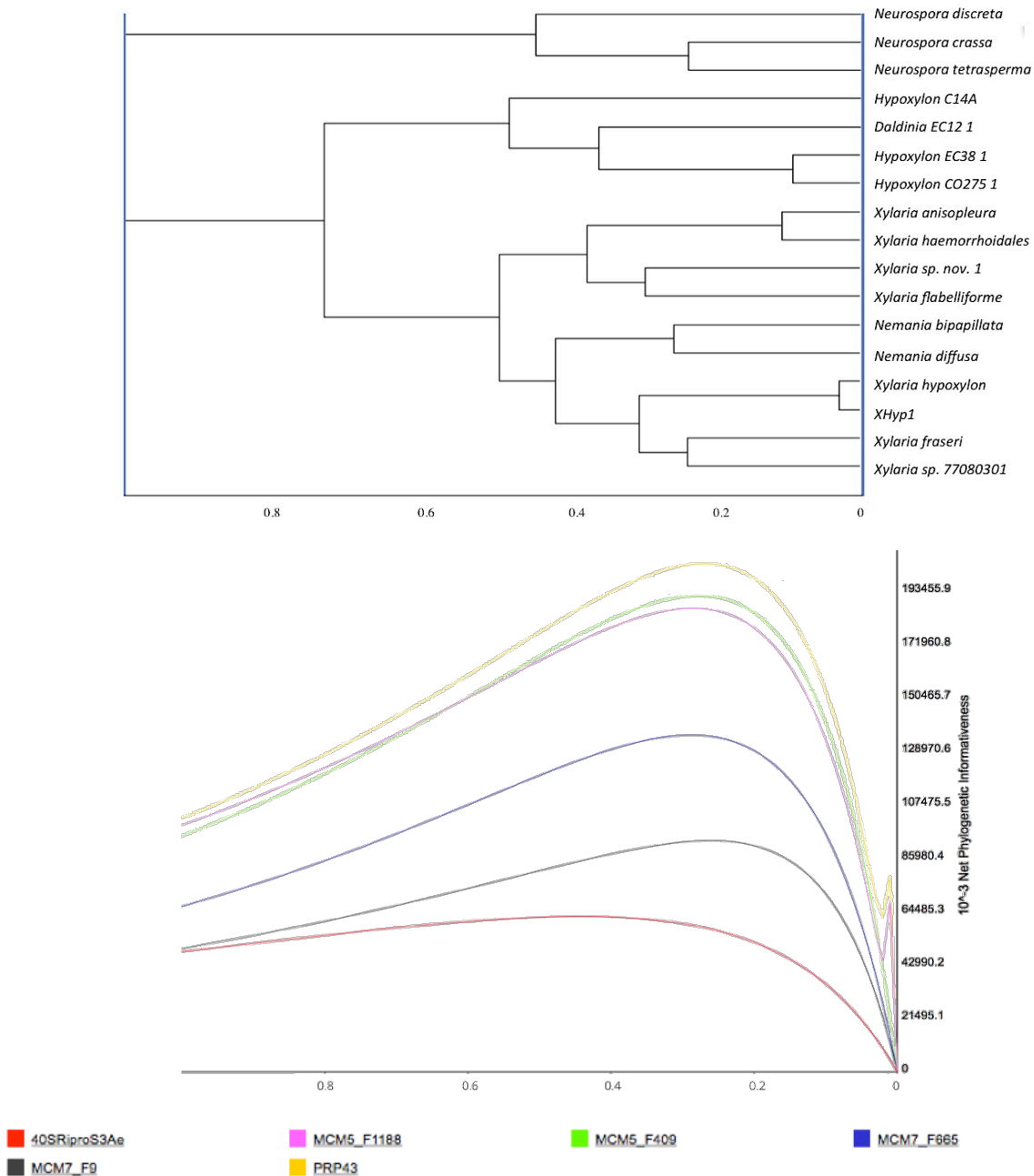
Figure 2. Ultrametric tree (above) and phylogenetic informativeness graph (below) created for all 6 primer pairs with the introns removed from MCM5 F1188/R1970 and 40S Ripro S3Ae F47/R605.

The same method was employed with the 4 primer pairs that were most successful at sequencing genes. This analysis included genes from 21 species (Figure 3). The graph resulted in a broad curve, with no significant irregularities and the general results were unchanged in both analyses. The Splicing factor PRP 43 F1054/R1691 was the most phylogenetically informative primer pair of those tested. The second most informative primer pair was MCM5 F409/R1059.
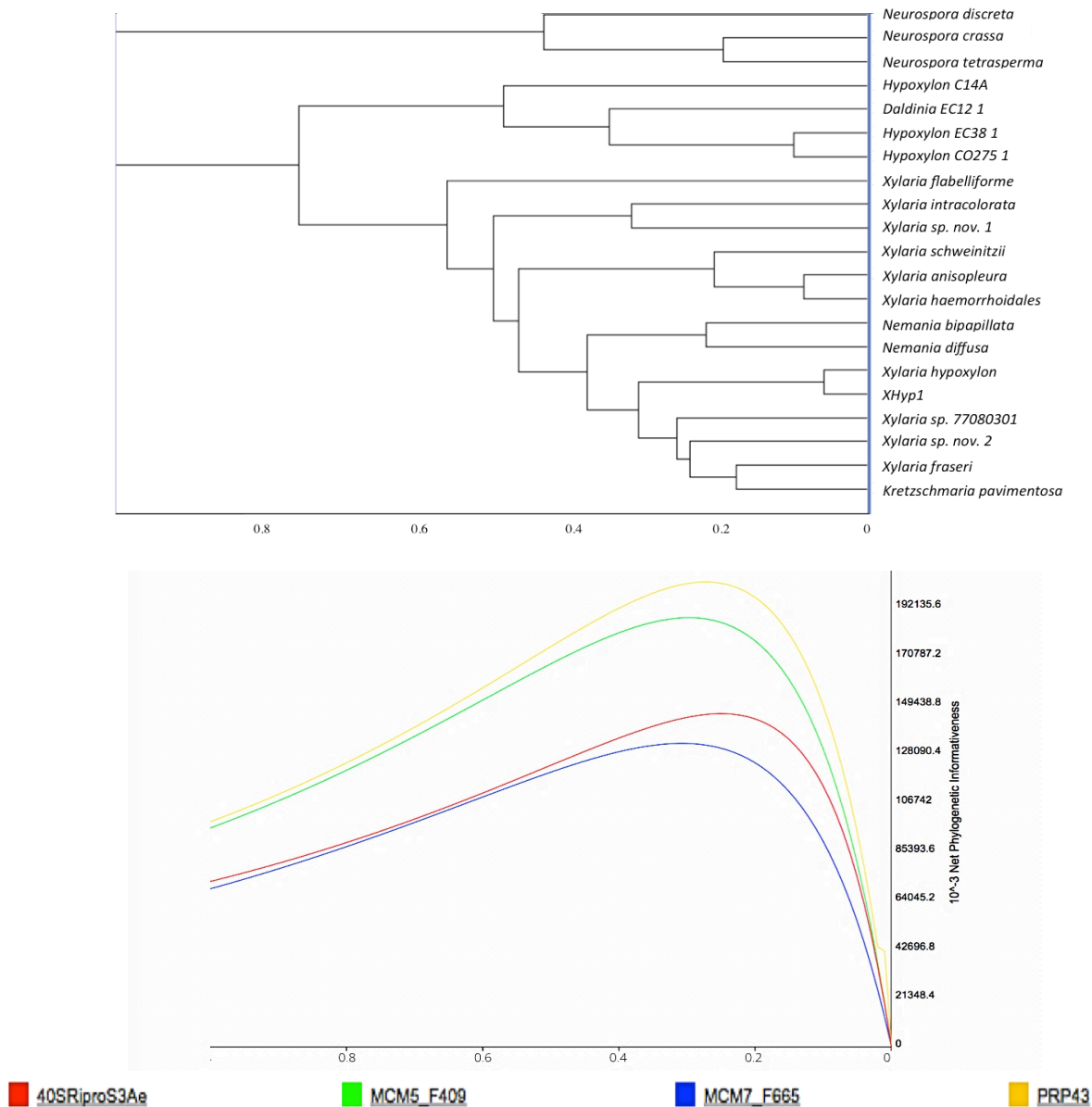
Figure 3. Ultrametric tree (above) and phylogenetic informativeness graph (below) of 4 primer pairs, created with sequences from 21 species' genes.

4. DISCUSSION

Many past studies have relied on ITS sequences for phylogenetic analyses, which have proven to be extremely variable. However, the recent availability of published Xylariaceae genomes has opened up opportunities to test and develop methods for phylogenetic analysis using single-copy genes. This study developed primer pairs for single-copy genes and proposes them for use during future *Xylaria* phylogenetic analysis.

4.1 PRIMERS ANALYSIS

Pairing the graph with the ultrametric trees provided an opportunity to examine where each primer set provided the most information, corresponding to the depth of branching within the tree. The top primer pairs were informative over a long range, corresponding to majority of the phylogenetic tree. The top primers had broad curves, with peaks around 0.25 on the relative time-correlated x-axis. In comparison, the primer pairs that contained introns originally produced a curve that had a single large spike at the tips of the branches. This reinforces the concept that introns, which are noncoding and highly variable, can be very informative for species-level analyses, but not for deeper inference.

The primer pair for the Splicing Factor PRP 43 gene (F1054: RCTYATCGAGCAGACTTAYCC; R1691:YACCTGCATGAAGAAGCC) resulted in the most successful and informative sequence of all those tested in this study. There was successful amplification in all 17 of the species that we sequenced. It consistently produced the highest curve on the phylogenetic informativeness graphs. The curve also showed that this primer pair provided a wide range of information across the full depth of the tree. We propose this Splicing Factor PRP 43 primer pair as a leading primer set for future *Xylaria* phylogenetic analysis. Examination of the gene's function reveals that the Splicing Factor PRP 43 gene codes for a protein associated with mRNA and is involved in the production of alternative splicing.

The second primer set we present is a section of MCM5 (F409: TGATCAAGTACCTGTVGGAGAGC; R1059:TCTCGCATCTTGTCGAACTCG), which also showed high levels of phylogenetic informativeness. The primer pair produced the second highest curve on the informativeness graph across multiple analyses (Figures 2 and 3). We were able to collect sequences from 15/17 test species. The MCM5 protein is part of the mini-chromosome maintenance complex, which was named from a study that focused on fungal mini-chromosomes. However, it is not limited to mini-chromosomes, as many fungi do not actually have them. MCM5 has a broad, important function within DNA replication in all eukaryotic cells, including human cells (Lei 2005).

The ribosomal associated protein 40S Ripro S3Ae (F47: GCGAACGAYGCCCTCAAGG; R605:TTGACCTTGCGGATGTGAACC) primer pair is the third primer set that is advocated. The sequence contains an intron and removing it greatly reduced the information it provided in the other analysis. This intron, where alignable, appears to provide excellent phylogenetic information towards the tips of the tree, as well as providing some deeper information. It effectively sequenced 15/17 species in the study. The 40S Ripro S3Ae is considered to be an effective primer pair, though further testing is encouraged.

4.2 TREE PRODUCT

A recommended phylogenetic tree was constructed for the species tested in this study. The tree includes sequences from 14 species that were amplified in the lab and 7 species sequences from published works. There were 4 genes that contributed to the creation of the trees. The

phylogenetic and ultrametric versions are depicted in Figure 4. It is important to note that this composite, phylogenetic tree displays higher bootstrap values than many of the trees made by individual genes, because it utilized information from 4 genes sequences, rather than just one.
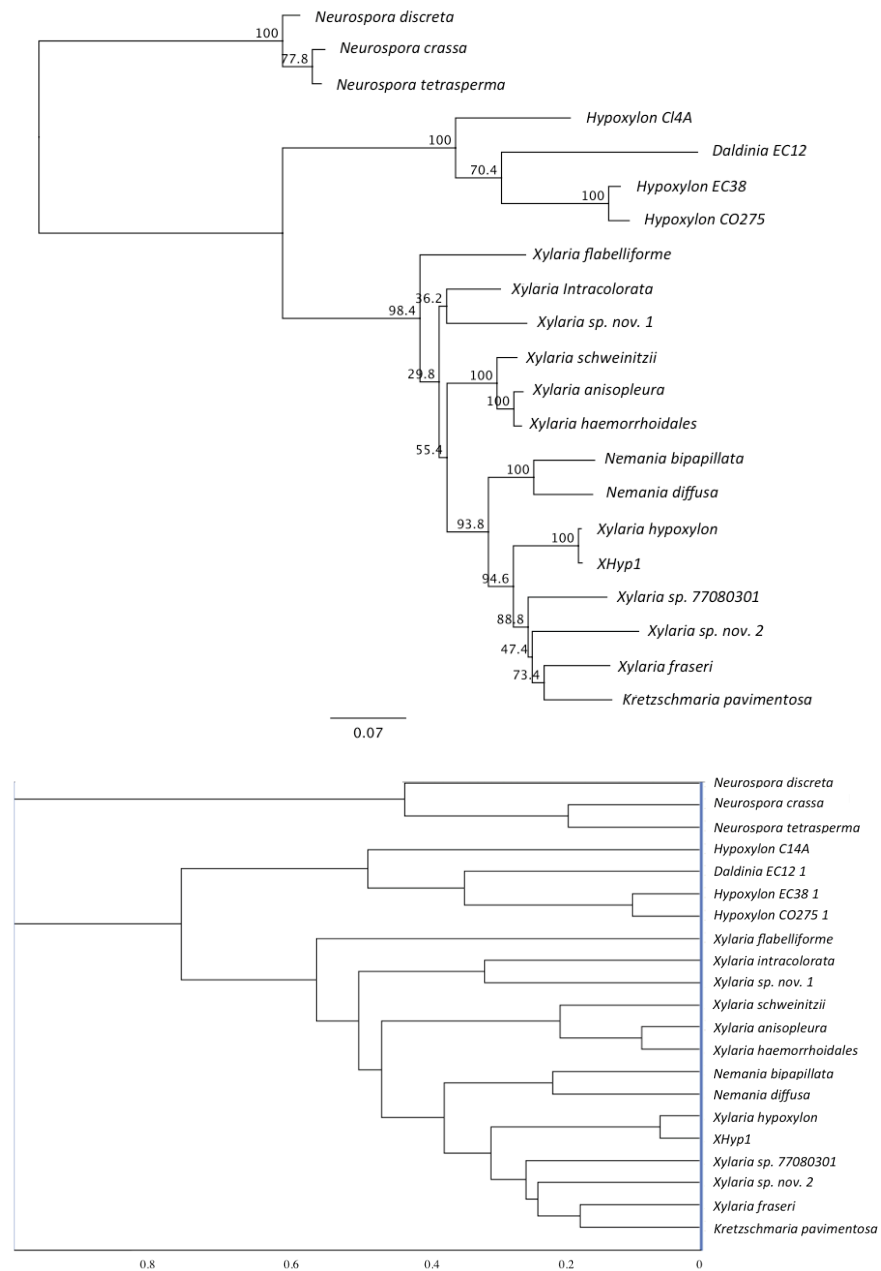

Figure 4. Proposed phylogenetic (above) and ultrametric trees (below) for the tested species. These were created with sequences from 4 genes in 21 *Xylaria* and related species.

Most of the relationships these trees display seem to be supported by past reports. However, one relationship that does not match previous studies is the nesting of the *Daldinia* between *Hypoxylon* species. Earlier studies have consistently found *Hypoxylon* and *Daldinia* species in separate groups (Triebel et al. 2005; Hsieh et al. 2005). But, in this study, the Hypoxylon Cl4A sequence came from a reference, which was known only by endophyte culture and ITS sequences. Samples were never collected from a fruiting body, during which its identification could be confirmed. Thus, we cannot accurately conclude what species it really is. Therefore, our results do not necessarily contradict or disprove any previous studies.

4.3 RECOMMENDATIONS FOR FURTHER STUDY

This project has opened up many opportunities for further study. First, the primer sets proposed should be tested further on additional *Xylaria* species. These primer sets have potential to become widely used gene markers for *Xylaria* and may even work favorably with related species. By testing the top primer pairs further, more sequence data will be collected as well. This sequence data could be used to build larger, wide-ranging phylogenetic trees for the genus or family. Our lab has the potential to add over 20 additional species from Ecuador to the phylogenetic tree built. In addition, we have reached out to other scientists working with *Xylaria* about the project. Their contributions could provide unique species from Taiwan and Gabon. We aim to facilitate a collaborative project to build a robust phylogenetic tree that includes *Xylaria* species from across the world. Dr. George Carroll has designed many more potential primer pairs for *Xylaria* as well. This study could be replicated to search for additional potential primer sets.

4.4 RELATION TO ENVIRONMENTAL SCIENCE

Xylaria species are quite important to the environments in which they inhabit, making further study of the genus valuable to mycology, environmental science, and many other related fields. Their roles as decomposers contribute to the recycling of lignin and cellulose throughout the ecosystems in which they are found. In addition, they play a role in the carbon cycle, because they recycle carbon dioxide back into the atmosphere during their life processes. Studies that investigate the evolutionary history of such an important genus can provide insight into many scientific questions. We can gain a better understanding of the relative ages of the species, diversity within the genus, and patterns of evolution of species that are supported by the decomposition of wood by *Xylaria*. Especially in a world that is currently experiencing major changes in the environment, understanding the diversity within the genus could be vital to preserving *Xylaria* species and other organisms supported by them in the future.

4.5 CONCLUSION

In this study, three primer pairs were proposed for future use during phylogenetic evaluation of *Xylaria*. The methods used to develop and evaluate these can be replicated for many other groups of organisms. Studies of numerous fungi, not just *Xylaria*, have relied solely on ITS data over the years. Finding primers that could be used for single-copy genes in other

organisms could significantly enhance the study of their species relationships. Use of these methods could lead to greater confidence in numerous phylogenetic analyses conducted in the future.
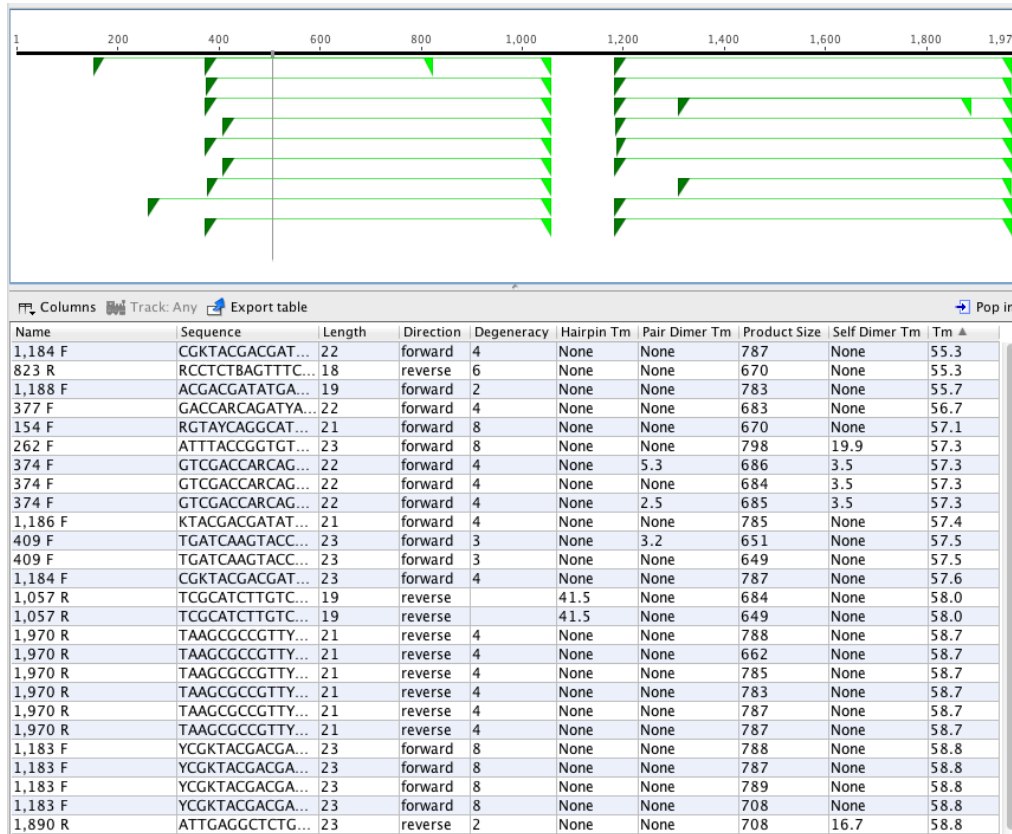
APPENDIX

Name	Sequence	Length	Direction	Degeneracy	Hairpin Tm	Pair Dimer Tm	Product Size	Self Dimer Tm	Tm ▲
1,184 F	CGKTACGACGAT...	22	forward	4	None	None	787	None	55.3
823 R	RCCTCTBAGTTTC...	18	reverse	6	None	None	670	None	55.3
1,188 F	ACGACGATATGA...	19	forward	2	None	None	783	None	55.7
377 F	GACCARCAGATYA...	22	forward	4	None	None	683	None	56.7
154 F	RGTAYCAGGCAT...	21	forward	8	None	None	670	None	57.1
262 F	ATTTACCGGTGT...	23	forward	8	None	None	798	19.9	57.3
374 F	GTCGACCARCAG...	22	forward	4	None	5.3	686	3.5	57.3
374 F	GTCGACCARCAG...	22	forward	4	None	None	684	3.5	57.3
374 F	GTCGACCARCAG...	22	forward	4	None	2.5	685	3.5	57.3
1,186 F	KTACGACGATAT...	21	forward	4	None	None	785	None	57.4
409 F	TGATCAAGTACC...	23	forward	3	None	3.2	651	None	57.5
409 F	TGATCAAGTACC...	23	forward	3	None	None	649	None	57.5
1,184 F	CGKTACGACGAT...	23	forward	4	None	None	787	None	57.6
1,057 R	TCGCATCTTGTC...	19	reverse		41.5	None	684	None	58.0
1,057 R	TCGCATCTTGTC...	19	reverse		41.5	None	649	None	58.0
1,970 R	TAAGCGCCGTTY...	21	reverse	4	None	None	788	None	58.7
1,970 R	TAAGCGCCGTTY...	21	reverse	4	None	None	662	None	58.7
1,970 R	TAAGCGCCGTTY...	21	reverse	4	None	None	785	None	58.7
1,970 R	TAAGCGCCGTTY...	21	reverse	4	None	None	783	None	58.7
1,970 R	TAAGCGCCGTTY...	21	reverse	4	None	None	787	None	58.7
1,970 R	TAAGCGCCGTTY...	21	reverse	4	None	None	787	None	58.7
1,183 F	YCGKTACGACGA...	23	forward	8	None	None	788	None	58.8
1,183 F	YCGKTACGACGA...	23	forward	8	None	None	787	None	58.8
1,183 F	YCGKTACGACGA...	23	forward	8	None	None	789	None	58.8
1,183 F	YCGKTACGACGA...	23	forward	8	None	None	708	None	58.8
1,890 R	ATTGAGGCTCTG...	23	reverse	2	None	None	708	16.7	58.8

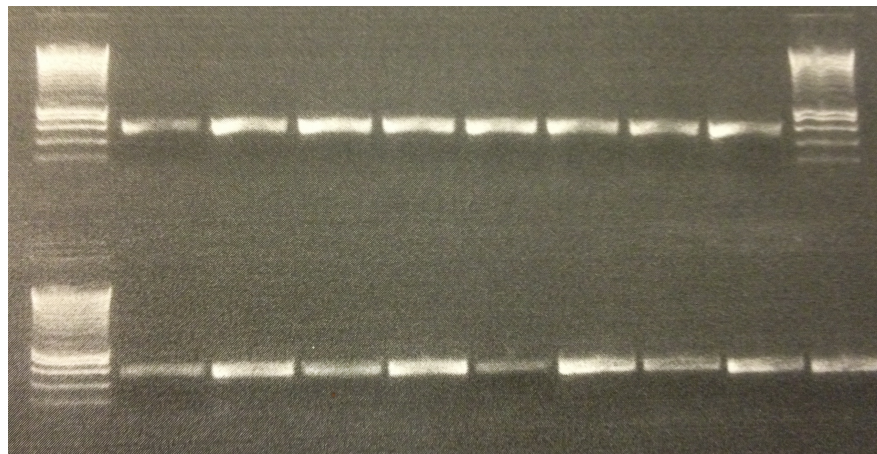**Figure 5.** Image showing the selection of primer pairs in Geneious.

Figure 6. Sample Gel, showing result of gel electrophoresis of PRP 43 F1054/R1691 primer pair.

Table 2. List of all potential primer pairs tested in this study.

Primer	Forward	Reverse
Phosphatidylinositol UDP-GlcNAc transferase PIG-A	88	761
Phosphatidylinositol UDP-GlcNAc transferase PIG-A	500	1299
Transcription regulatory protein pro 1	711	1324
MCM2	599	1249
DNA helicase RAD3	938	1656
RNA helicase PRP 16	1007	1756
Splicing factor PRP 43	127	776
MCM5	154	823
DNA polymerase gamma	1618	2275
MCM7	9	595
Transcription regulatory protein pro 1	1168	1917
MCM2	1291	1930
DNA helicase RAD3	82	683
RNA helicase PRP 16	372	1081
RNA helicase PRP 16	1732	2518
MCM5	1188	1970
DNA polymerase gamma	106	852
DNA polymerase gamma	938	1707
MCM7	200	961
Transcription regulatory protein pro 1	381	1020
MCM2	32	644
Splicing factor PRP 43	900	1595
Splicing factor PRP 43	1054	1691
GTP-binding-protein YPT1	16	727
MCM5	409	1059
MCM7	665	1209
Nucleosome assembly protein	150	817
Nucleosome assembly protein	478	1221
40S ripro S3Ae	47	605
40S ripro S3Ae	187	707

Table 3. Specifications for Splicing Factor PRP 43 primer pair.

Name	Sequence	Dir.	Prod. Size	Length	Degen.	Tm	Hrpn Tm	Pri-dm	Slf-dm
1,054 F	RCTYATCGAGCAGACTTAYCC	F	638	21	8	57	None	None	None
1,691 R	YACCTGCATGAAGAAGCC	R	638	18	2	60	None	None	None

Table 4. Specifications for MCM5 primer pair.

Name	Sequence	Dir.	Prod. Size	Length	Degen.	Tm	Hrpn Tm	Pri-dm	Slf-dm
409 F	TGATCAAGTACCTGTVGGAGAGC	F	651	23	3	58	None	3	None
1,059 R	TCTCGCATCTTGTCGAACTCG	R	651	21		61	42	3	None

Table 5. Specifications for 40S Ripro S3Ae primer pair.

Name	Sequence	Direction	Product Size	Length	Degeneracy	Tm	Hairpin Tm	Primer-dimer Tm	Self dimer Tm
47 F	GCGAACGAYGCCCTCAAGG	forward	559	19	2	59	40	None	None
605 R	TTGACCTTGCGGATGTGAACC	reverse	559	21		61	None	None	None
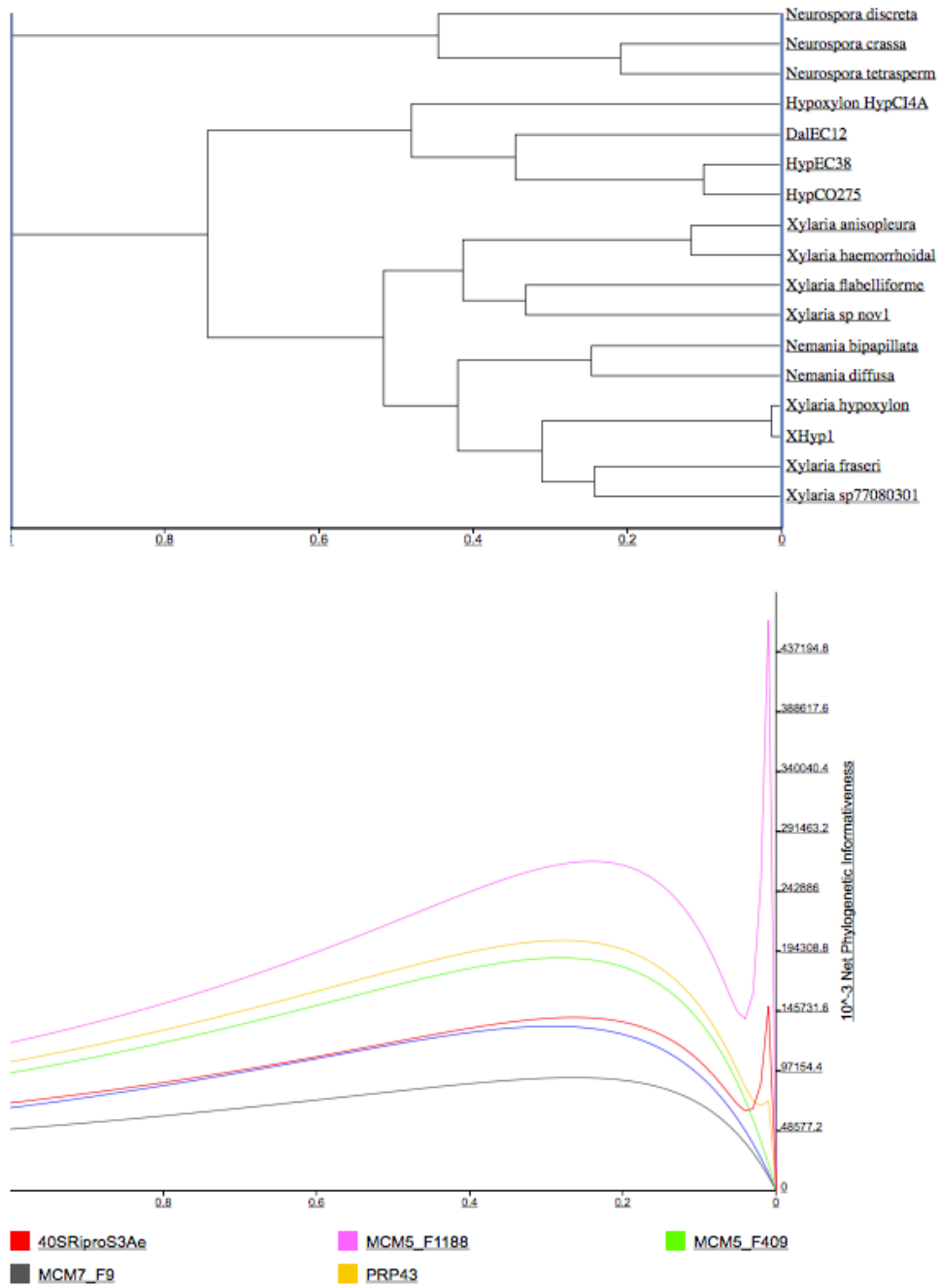
Figure 7. Ultrametric tree (above) and phylogenetic informativeness (below) of 6 Primer Pairs (with introns present).
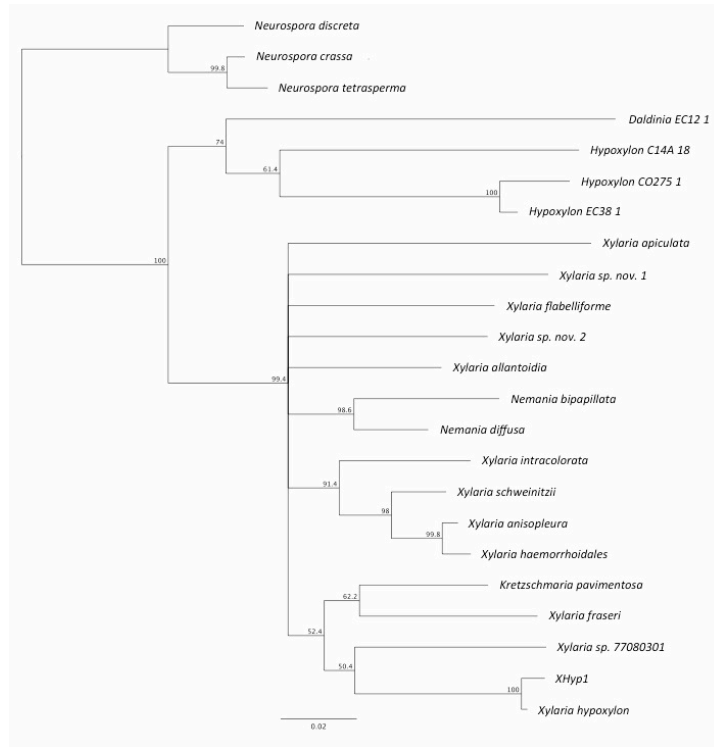
Figure 8. 40S Ripro S3Ae F47/R595 Neighbor-Joining Tree
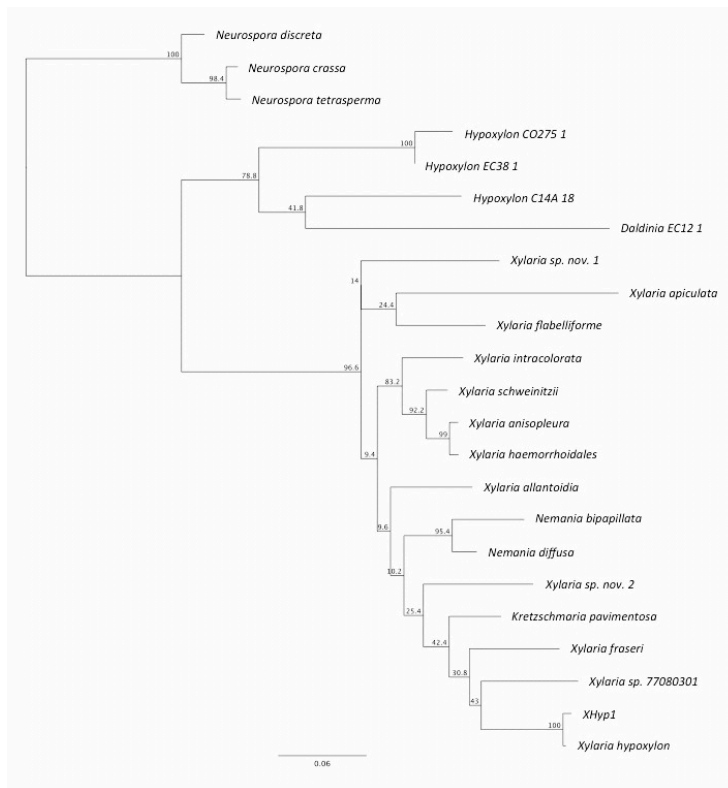
Figure 9. 40S Ripro S3Ae F47/R595 Maximum Likelihood Tree
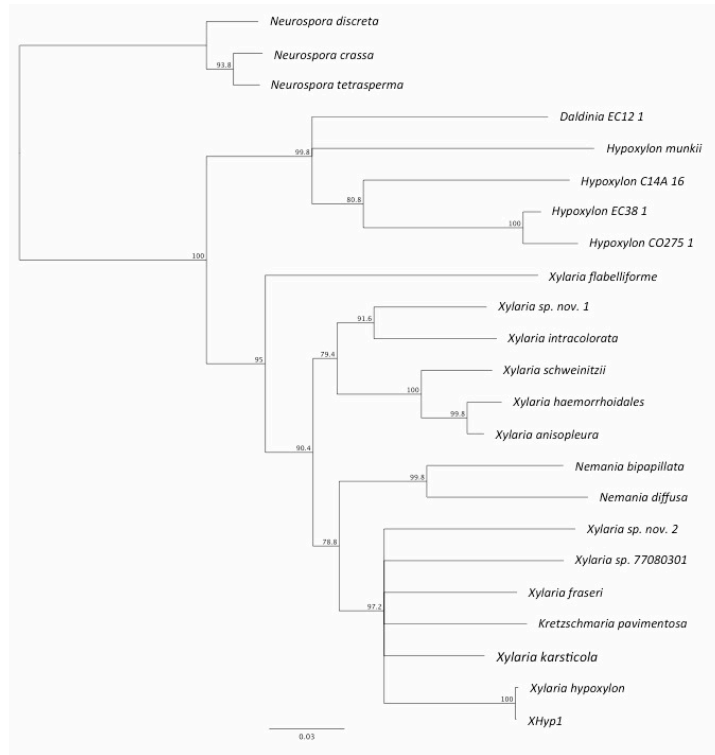
Figure 10. MCM5 F409/R1059 Neighbor-Joining Tree
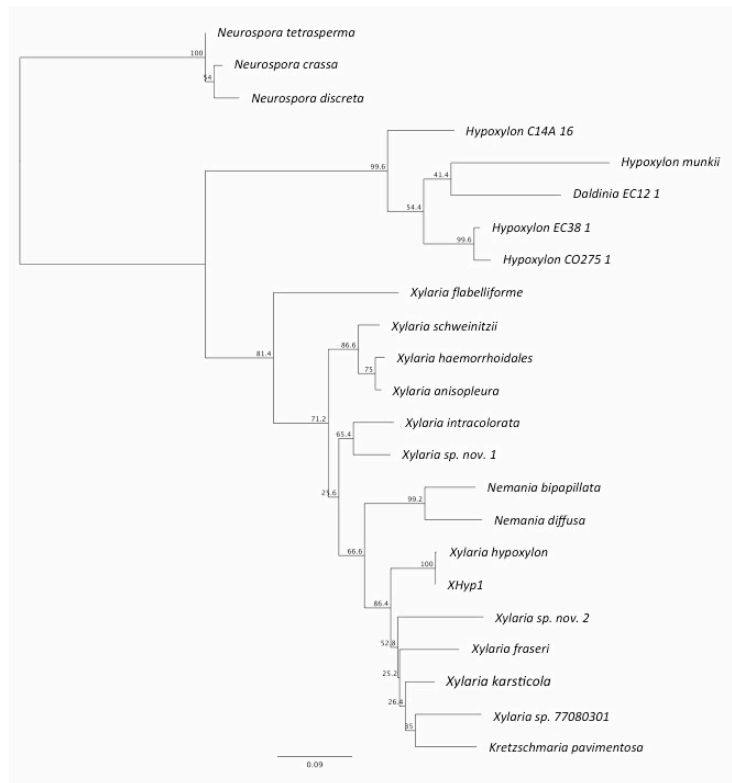
Figure 11. MCM5 F409/R1059 ML
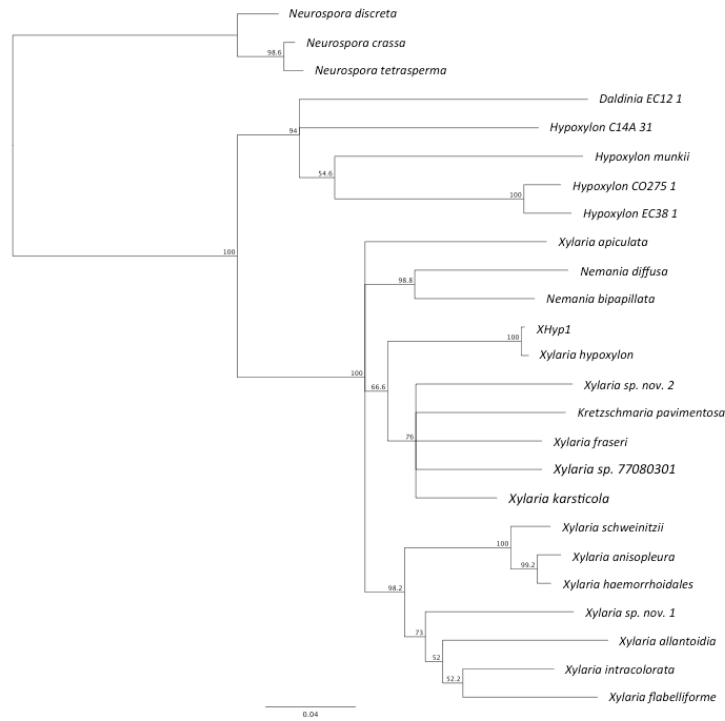
Figure 12. Splicing Factor PRP 43 F1054/R1691 Neighbor-Joining Tree
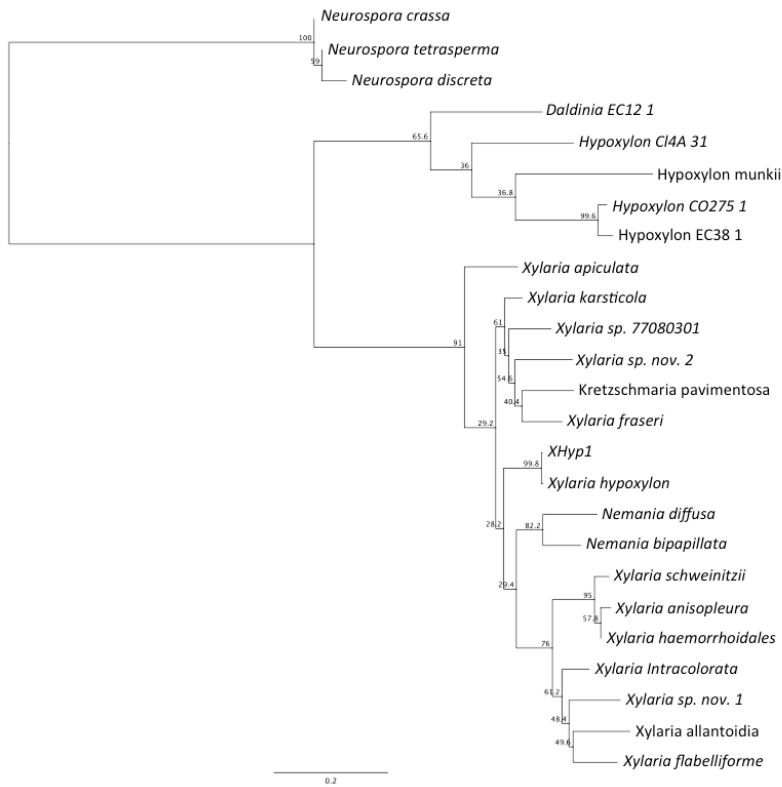
Figure 13. Splicing Factor PRP 43 F1054/R1691 Maximum Likelihood Tree
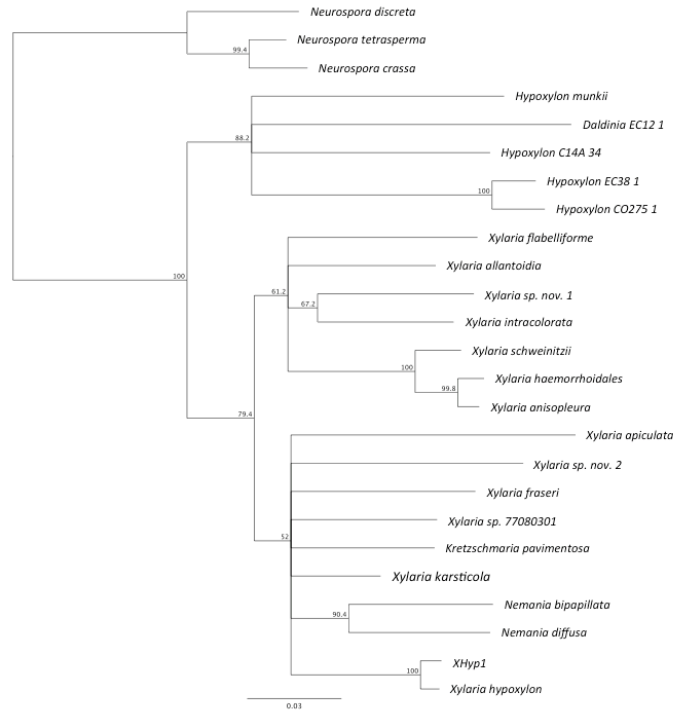
Figure 14. MCM7 F665/R1209 Neighbor-Joining Tree
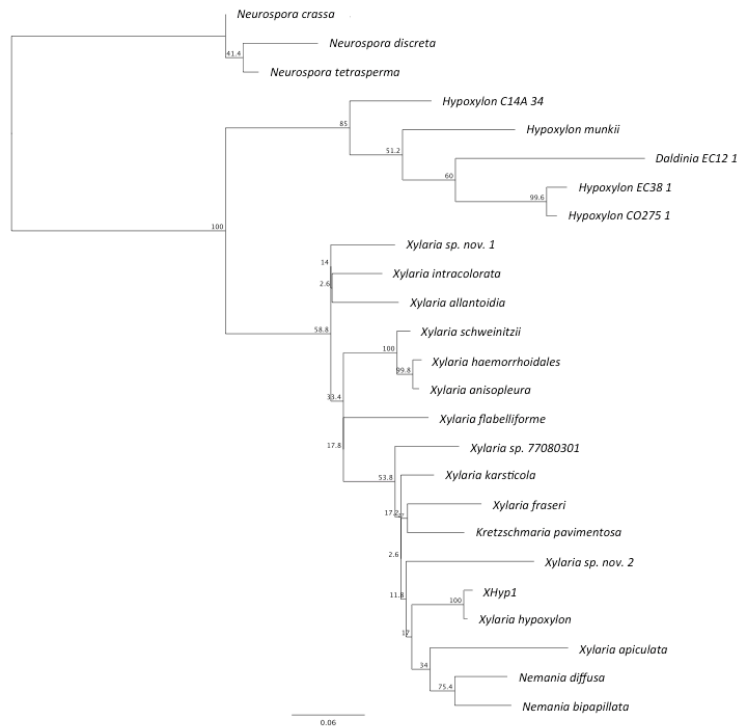
Figure 15. MCM7 F665/R1209 Maximum Likelihood Tree
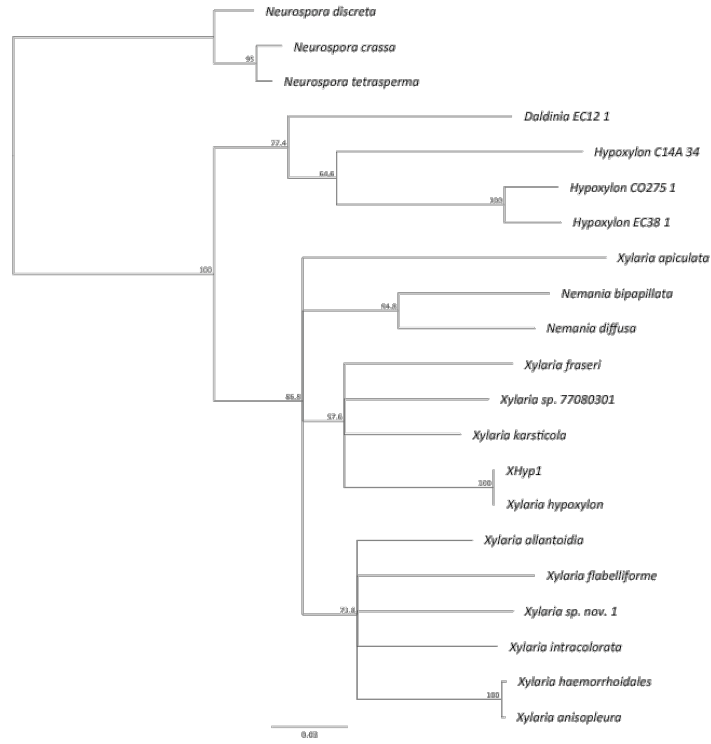
Figure 16. MCM7 F9/R595 Neighbor-Joining Tree
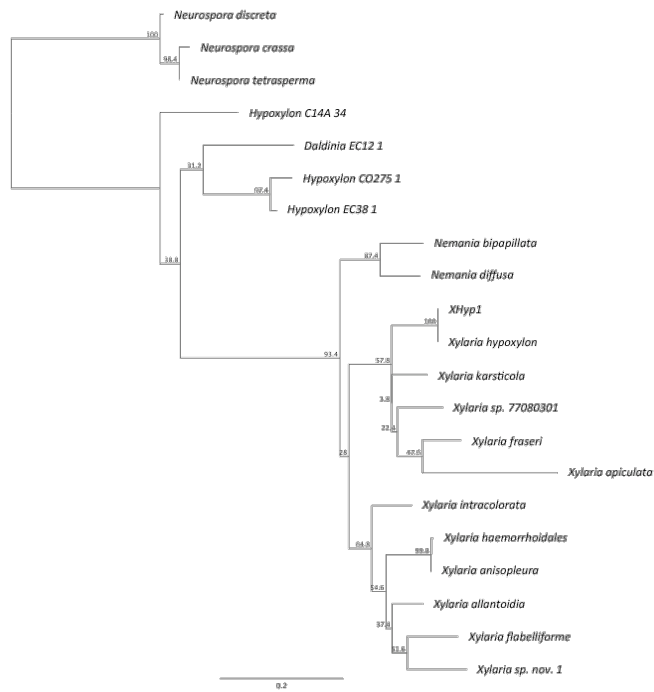
Figure 17. MCM7 F9/R595 Maximum Likelihood Tree
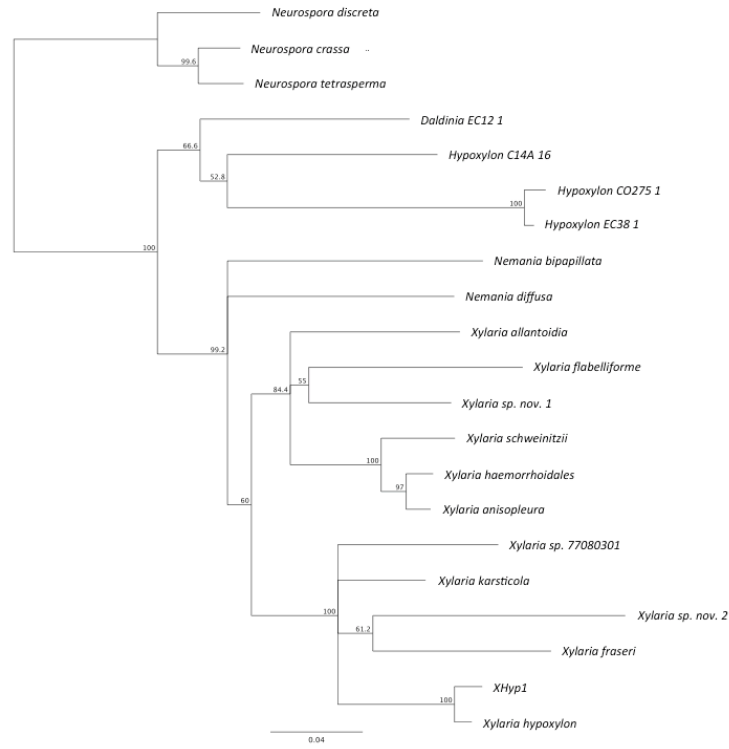
Figure 18. MCM5 F1188/1970 Neighbor-Joining Tree
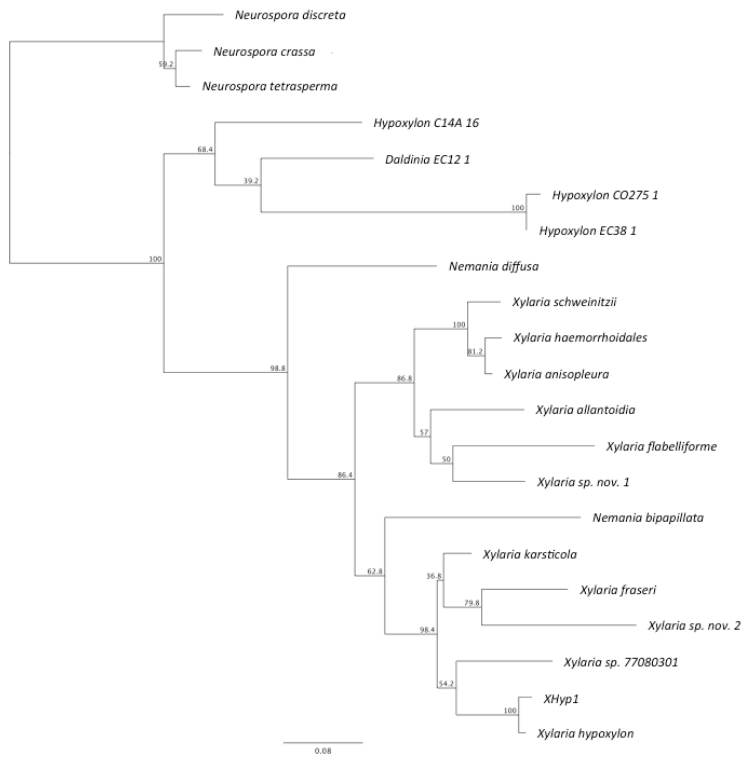
Figure 19. MCM5 F1188/1970 Maximum Likelihood Tree

ACKNOWLEDGEMENTS

REFERENCES

Álvarez, I., & Wendel, J.F. 2003. Ribosomal ITS sequences and plant phylogenetic inference. Molecular Phylogenetics and Evolution 29(3):417-434.

Cai L, Jeewon R, Hyde KD. 2006. Phylogenetic investigations of Sordariaceae based on multiple gene sequences and morphology. Mycological Research 110(2):137-150.

Capella-Gutierrez S, Kauff F, Gabaldón T. 2014. A phylogenomics approach for selecting robust sets of phylogenetic markers. Nucleic Acids Research 42(7):54.

Compton T. 1990. Degenerate primers for DNA amplification. In: Innis M, Gelfand D, Sninsky J, White T, editors. PCR protocols: a guide to methods and applications. New York: Academic Press. p. 39-45.

Davis EC, Franklin JB, Shaw AJ, Vilgalys R. 2003. Endophytic *Xylaria* (Xylariaceae) among liverworts and angiosperms: phylogenetics, distribution, and symbiosis. American Journal of Botany 90(11):1661-1667.

Dennis R. 1956. Some Xylarias of tropical America. Kew Bull 11(3):401-444.

Dentinger B, Margaritescu S, Moncalvo J. 2009. Rapid and reliable high-throughput methods of DNA extraction for use in barcoding and molecular systematics of mushrooms. Molecular Ecology Resources 10(4):628-633.

Felsenstein J. 1985. Confidence limits on phylogenies: An approach using the bootstrap. Evolution 39(4):783-791.

Hsieh HM, Ju YM, Rogers JD. 2005. Molecular phylogeny of *Hypoxylon* and closely related genera. Mycologia 97(4):844-865.

Hsieh HM, Lin C, Fang M, Rogers JD, Fournier J, Lechat C, Ju Y. 2010. Phylogenetic status of *Xylaria* subgenus *Pseudoxylaria* among taxa of the subfamily Xylarioideae (Xylariaceae) and phylogeny of the taxa involved in the subfamily. Molecular Phylogenetics and Evolution 54(3):957-969.

Huelsenbeck JP. 1995. The robustness of two phylogenetic methods: Four-taxon simulations reveal a slight superiority of maximum likelihood over neighbor joining. Molecular Biology and Evolution 12(5):843-849.

Kiss L. 2012. Limits of nuclear ribosomal DNA internal transcribed spacer (ITS) sequences as species barcodes for fungi. Proceedings of the National Academy of Sciences USA 109:E1811.

Larena I, Salazar O, González V, Julián MC, Rubio V. 1999. Design of a primer for ribosomal DNA internal transcribed spacer with enhanced specificity for Ascomycetes. Journal of Biotechnology 75(2):187-194.

Lee JS, Ko KS, Jung HS. 2000. Phylogenetic analysis of *Xylaria* based on nuclear ribosomal ITS1-5.8 S-ITS2 sequences. FEMS Microbiology Letters 187(1):89-93.

Lei M. 2005. The MCM complex: Its role in DNA replication and implications for cancer therapy. Current Cancer Drug Targets 5(5):365-80.

Lopez-Giraldez F and Townsend JP. 2011. PhyDesign: An online application for profiling phylogenetic informativeness. BMC Evolutionary Biology 11:152.

Peláez F, González V, Platas G, Sánchez Ballesteros J, Rubio V. 2008. Molecular phylogenetic studies within the Xylariaceae based on ribosomal DNA sequences. Fungal Diversity 31(1): 111-134.

Rogers JD. 1979. The Xylariaceae: systematic, biological and evolutionary aspects. Mycologia 71(1):1-42.

Rogers JD. 2000. Thoughts and musings on tropical Xylariaceae. Mycol Res 104(12):1412-1420.

Sánchez-Ballesteros, J., González, V., Salazar, O., Acero, J., Portal, M.A., Julián, M., ..., Platas, G. 2000. Phylogenetic study of *Hypoxylon* and related genera based on ribosomal ITS sequences. Mycologia 92(5):964-977.

Sanderson MJ. 1997. A nonparametric approach to estimating divergence times in the absence of rate constancy. Molecular Biology and Evolution 14(12):1218-1231.

Sanderson MJ. 2002. Estimating absolute rates of molecular evolution and divergence times: A penalized likelihood approach. Molecular Biology and Evolution 19(1):101-109.

Schoch CL, Seifert KA, Huhndorf S, Robert V, Spouge JL, Levesque CA, Chen W, Fungal Barcoding Consortium. 2012. Nuclear ribosomal internal transcribed spacer (ITS) region

as a universal DNA barcode marker for fungi. Proceedings of the National Academy of Sciences U S A 109(16):6241-6246.

Stadler M, Fournier J, Læssøe T, Decock C, Peršoh D, Rambold G. 2010. *Ruwenzoria*, a new genus of the Xylariaceae from Central Africa. Mycological Progress 9(2):169-179.

Stadler M, Kuhnert E, Peršoh D, Fournier J. 2013. The Xylariaceae as model example for a unified nomenclature following the "One Fungus-One Name" (1F1N) concept. Mycology 4(1):5-21.

Tang A, Jeewon R, Hyde K. 2009. A re-evaluation of the evolutionary relationships within the Xylariaceae based on ribosomal and protein-coding gene sequences. Fungal Diversity 34:127-155.

Triebel D, Peršoh D, Wollweber H, Stadler M. 2005. Phylogenetic relationships among *Daldinia*, *Entonaema*, and *Hypoxylon* as inferred from ITS nrDNA analyses of Xylariales. Nova Hedwigia 80(1-2):25-43.

Folmer O, Black M, Hoeh W, Lutz R, Vrijenhoek R. 1994. DNA primers for amplification of mitochondrial cytochrome c oxidase subunit I from diverse metazoan invertebrates. Molecular Marine Biology and Biotechnology 3(5):294-299.

Walker DM, Castlebury LA, Rossman AY, Sogonov MV, White JF. 2010. Systematics of genus *Gnomoniopsis* (Gnomoniaceae, Diaporthales) based on a three gene phylogeny, host associations and morphology. Mycologia 102(6):1479-1496.

Walker DM, Castlebury LA, Rossman AY, White Jr. JF. 2012. New molecular markers for fungal phylogenetics: two genes for species-level systematics in the Sordariomycetes (Ascomycota). Molecular Phylogenetics and Evolution 64(3):500-512.

Whalley A. 1996. The Xylariaceous way of life. Mycological Research 100(8):897-922.