SEQUENCE ANALYSIS OF PARTIAL TEF1α AND RPB2 GENES OF DIFFERENT PLEUROTUS ERYNGII ISOLATES BY MEANS OF PCR-RFLP

JÓZSEF SZARVAS1, KÁROLY PÁL2 AND ANDRÁS GEÖSEL3
1Mushroom Spawn Plant and Strain Research Laboratory, Biokékes Nonprofit Ltd., H-3395 Demjen, Top. Nr. 0287/8. Hungary
2Department of Microbiology and Food Technology, Eszterházy Károly College, H-3300 Eger, Leányka str. 6-8. Hungary
3Department of Vegetable and Mushroom Growing, Faculty of Horticultural Science, Corvinus University of Budapest, H-1118 Budapest, Villányi út 29-43. Hungary
szarvasjosef@hotmail.com, karoly.pal@ektf.hu, andras.geosel@uni-corvinus.hu

ABSTRACT

The Pleurotus eryngii species complex includes several varieties and certain groups with ambiguous taxonomic position. At the moment the following taxa are recognized: var. eryngii; var. ferulae /syn.: P. fuscus var. ferulae; var. elaeoselini; var. nebrodensis; var. tingitanus; var. tuoliensis; P. hadamardii; P. fossulatus. In the course of our studies sequence analysis of the tef1α and rpb2 genes was performed, in order to distinguish varieties and reveal variability between isolates. Specific regions of the tef1α and rpb2 genes were amplified in PCR experiments. The fragments were cleaned, sequenced, aligned then BLAST searched against the NCBI GenBank nucleotide database. Some point mutations were detected in the sequences, which were used for selection of differentiating restriction enzymes for subsequent PCR-RFLP experiments.

The tef1α gene sequences showed 100% identity in each strain. In contrast to that, point mutations were detected in the 21, 372 and 957 positions of the rpb2 sequences. In silico digestion was performed on the sequences and two restriction endonucleases, the BsmA I and TspDI I were selected for PCR-RFLP experiments. As a result of the digestions, the isolates could be grouped into two groups with both enzymes. At the same time, BLAST search with both amplified sequences did not give reliable information neither on varietas, nor on species level. The rpb2 locus, with its higher level of polymorphism, is a potential candidate for differentiation between varieties or identification on varietas level.

Based on the results we plan to investigate which loci and molecular methods may be suitable for differentiation of isolates.

Keywords: Pleurotus eryngii, tef1α, rpb2, PCR-RFLP

INTRODUCTION

As it was described in former studies, the Pleurotus eryngii species complex consists of numerous varieties and taxa with uncertain taxonomic positions. Nevertheless, each species share a common feature: they are associated with certain species of the Apiaceae (Umbelliferae) and Asteraceae (Compositae) families, as facultative biotrophs or white rot fungi in their natural habitat [1–7].

The term ‘species complex’ is used in general for fungi species that are closely related, but partially or completely incompatible [8, 9, 4, 10].

Identification of the species might be very difficult, because the classical, mostly micro- and macro morphological characters might be masked by environmental effects. Hence, taxonomic research is often based on nucleic acids nowadays. The ITS (Internal Transcribed Spacer) region of the ribosomal gene cluster is one of the most frequently investigated DNA fragment in fungal taxonomy. In former studies, phylogenetic relations amongst members of the P. eryngii species complex were thought to be revealed on basis of the ITS region and partial α-tubulin gene sequences. Unfortunately, none of these regions showed such variability that could be used for differentiation of varieties [11–14].

Additional candidates for molecular phylogenetics and identification on varietas level are the genes coding certain regions of the translation elongation factor (EF1α) and the second largest subunit of the RNA polymerase II (rpb2) [11, 12, 14-19]. The EF1α is a binding protein used in ribosomal protein synthesis in eukaryotic cells. Marongiu et al. [20] found
nucleotide substitutions in the tef1ά gene (coding EF1ά) that might be useful in differentiation of *eryngii* and *ferulae* varieties.

The rpb2 gene has a strongly conserved domain, so it is a good candidate for PCR primers. Formerly, the rpb2 was used (with additional genomic regions) for identification of *Cortinarius* and *Inocybe* species [18, 21].

In addition to the beta-tubulin, tef1 and rpb2 genes, there are numerous other loci available for fungal taxonomists, including nuclear and mitochondrial ribosomal rDNA (18S, 28S, ITS, IGS, mtSSU, mtLSU) and protein coding genes, e.g. rpb1, gamma-actin or ATP synthase. In previous experiments, those protein coding genes that are present in the genome in a single copy, were successful candidates in phylogenetic experiments [22–25]. Nevertheless, those genes that have multiple copies in the genome, might be potential sources of misleading results [26]. The type of gene one would like to use for phylogenetic analysis depends on how long evolutionary history is intended to be revealed. Quickly evolving genes offer good support for detecting recent evolutionary events, whereas genes with lower mutation rate are more suitable for long term evolutionary research. Ribosomal genes, such as ITS, IGS and mtSSU rDNA are subjects of high and moderate rate of mutation; hence, these loci are good choice for discrimination of isolates at genus and species level. Unfortunately, non-coding regions of these genes tend to undergo length variation, which might cause serious problems during alignment. Nowadays sequencing is popular method for investigation of organisms, but sequencing of entire fungal genomes or even multiple loci is not an option for many taxa, because fungal material convenient for molecular experiments is often limited and culturing of numerous species is not possible [27].

**MATERIALS AND METHODS**

**Strains**

King oyster mushroom strains were isolated from grassy fields of Novaj, Eger Felnémet-Pásztorvölgy, Bogács, Tószeg, Kecskemét and Heves (Hungary) in autumns between 2006 and 2008. Furthermore, some strains were made available to us from strain collection of the Species Research Laboratory of Korona Mushroom Union. Codes and place of origin of the strains used in this study are the following: Hungarian strains: PEA – Tószeg; PEP – Eger, Pásztorvölgy; PEG – Tószeg; PEF-i – Demjén, Vas-tanya; PEF – Kecskemét; PEC – Eger, Pásztorvölgy; Ple1V – Novaj; Ple2V – Novaj; Ple3V – Bogács; Ple4V – Hevesi Füves Puszta; Ple5V – Novaj; Ple6V – Novaj. Strains from abroad: PES – the Netherlands; PESZM – Malaysia; PEL – North-Italy; PEK – China.

**DNA extraction and PCR conditions**

DNA extraction was done according to the method of Szarvas et al. (2011) [28]. Amplification and sequence analysis was performed on the tef1ά and rpb2 regions, in order to differentiate *P. eryngii* strains and varieties.

Fig.1 shows the expected approx. 550 bp long fragment of the tef1ά locus, amplified in the PCR reactions. Primers annealed to and amplified the 4-6 exon regions (green) that included the 4-5 introns (I.4 and I.5), as well. Sequences of the primers are the following: EF595F (5’ CGT GAC TTC ATC AAG AAC ATG 3’) and EF1160R (5’ CCG ATC TTG TAG ACG TCC TG 3’).

**Figure 1.** Annealing sites of primers (empty arrows) in the tef1ά gene (coding translation elongation factor EF1ά). Yellow: introns (I.1-I.7); green: exons (based on data of Wendland & Kothe 1997 [29]).
An almost 1100 bp long fragment of the *rpb2* (RNA-polymerase II) locus was amplified by PCR, too. The primers amplified the 4-5 exon (domain 7-11) region (Fig. 2). Primers used in this reaction were bRPB2-6.9F (=b6.9F) (5’ TGG ACN CAY TGY GAR ATY CAY CC 3’) and bRPB2-11R1 (=b11R1) (5’ TGG ATY TTG TCR TCC ACC AT 3’).

**Figure 2.** Structure of *rpb2* locus of Basidiomycetes [30]. Primer positions and direction of synthesis are shown by arrows. Eukaryotic conserved regions are shown in blue.

In the first step conditions of the PCR reaction were optimized. Annealing temperature of primers was determined by temperature gradient PCR experiments, it was followed by the regular reactions. DNA of the 16 *P. eryngii* strains was used as template, after the concentration of each one was set to 10 ng/μl, in order to avoid false results due to unequal amount. The PCR reactions were based on the DreamTaq Kit (Fermentas) and set up as follows (for one reaction): 2.5 μl DreamTaq Buffer, 0.5 μl 25 mM MgCl₂, 0.5 μl 10 mM dNTP, 0.2 μl DreamTaq Polymerase (5U/μl), 1 μl forward primer (5 μM/μl), 1 μl reverse primer (5 μM/μl), 2 μl template DNA (10 ng/μl); 17.3 μl distilled water. The PCR program was optimized for both primer pairs and was the following: 1. 95 °C 3 min; 2. 95 °C 5 s, 57 °C 10 s, 72 °C 30 s; 3. 72 °C 5 min. The second step was repeated 35 times. PCR reactions were performed in a Corbett Research Thermocycler (Corbett Life Science, Australia) instrument.

The amplicons were visualized by gel electrophoresis on 1% (w/v) agarose gel (SeaKem LE Agarose, Lonza, USA) in 1× TBE (Tris-Borate-EDTA) for 30 min at 130 V and made visible by GelRed (Biotium, USA) staining and UV transillumination in a BioDocAnalyze (Biometra, Germany) instrument. A 100 bp BenchTop DNA ladder (Promega, USA) was used as molecular size marker.

**Sequencing and alignment of sequences.**

The fragments were cleaned by EZ-10 Spin Column Kit PCR Purification Kit (Biobasic Inc., Canada), according to the manufacturer’s instructions. Quality of the processed amplicons was checked again on 1% agarose gel as above and 10 μl of each sample was sent to be sequenced with both primers in the sequencing laboratory (LGC Genomics, Germany).

In the next step the fragments sequenced from the reverse primer were used to generate their reverse complement version, in order to be used for alignment. Reverse complement sequences were made by the EditSeq software of DNAStar package. It was followed by the alignment of reverse complement sequences with fragments sequenced by the forward primer, and then the complete fragments of each strain were aligned with each other. Both steps were performed by ClustalX (http://www.clustal.org/clustal2/), and the final alignment was visualized by BoxShade http://www.ch.embnet.org/software/BOX_form.html). Due to the alignment point mutations, identities and differences between strains became visible, which might help in identification of varieties in future experiments.

Identification of strains was performed by blasting the aligned sequences against the GenBank database of NCBI by BLAST algorithm.

**Analysis of fragments by RFLP (Restriction Fragment Length Polymorphism).**

Alignment of sequences revealed the existence of point mutations in some strains in the same position. These minor differences offer the possibility for differentiation of strains by RFLP, if proper restriction enzymes are available. Restriction enzymes
were selected by using the RestrictionMapper online tool (http://www.restrictionmapper.org/); \textit{in silico} digestion was performed on the sequences to find enzyme(s) capable of cleaving the DNA at the point mutations.

In total, two enzymes were found that showed ability to cut the DNA at the mutations. These enzymes were \textit{BsmAI} and \textit{TspDTI}, with recognition sites and digestion profile showed on Fig. 3 and Fig. 4.

\textbf{BsmAI} \hfill \textbf{TspDTI}

\[
\begin{align*}
5' & \text{GTCTC} (N) \downarrow 3' \\
3' & \text{CAGAG} (N) \uparrow 5' \\
5' & \text{ATGAANNNNNNNNNNN} \downarrow 3' \\
3' & \text{TACTTNNNNNNNNNNNNNNNN} \uparrow 5'
\end{align*}
\]

\textbf{Figure 3.} Cleavage site of \textit{BsmAI}. \hfill \textbf{Figure 4.} Cleavage site of \textit{TspDTI}.

Reaction setup for one reaction with the \textit{BsmAI} (New England Biolabs, USA) digestion was the following: 3 \text{ il NE buffer 4}, 1 \text{ il \textit{BsmAI}}, 16 \text{ il distilled water and 10 il PCR mix, containing the amplified fragments. Digestion was performed at 55 °C for 60 min.}

Reaction setup for one reaction with the \textit{TspDTI} (EURx, Poland) digestion was the following: 3 \text{ il 1 \times \textit{TspDTI} buffer, 1 il \textit{TspDTI}}, 16 \text{ il distilled water, 10 10 il PCR mix, containing the amplified fragments. Digestion was performed at 70 °C for 3 hours.}

The amplicons were visualized by gel electrophoresis on 2\% (w/v) agarose gel (SeaKem LE Agarose, Lonza, USA) in 1× TBE (Tris-Borate-EDTA) for 180 min at 80 V and made visible by GelRed (Biotium, USA) staining and UV transillumination in a BioDocAnalyze (Biometra, Germany) instrument. A 100 bp BenchTop DNA ladder (Promega, USA) was used as molecular size marker.

\textbf{RESULTS AND DISCUSSION}

\textbf{PCR experiments}

Amplification of the \textit{tefl\textalpha} gene resulted only one fragment, which size was the expected 550 bp and non-specific product was not found (Fig. 5). In contrast to that, non-specific fragments were visible below the main product (1000 bp) on agarose gel after the amplification of \textit{rpb2} gene, probably due to the fact that primers used in this reaction were degenerated. Nevertheless, these by-products could be easily removed by cleaning and only the main product was present and used for further processing (Fig. 6).

\textbf{Figure 5.} Amplification of the \textit{tefl\textalpha} gene of 16 \textit{P. eryngii} strains resulted only the expected fragment \hfill \textbf{Figure 6.} Main products of the \textit{rpb2} PCR with DNA of 16 \textit{P. eryngii} strains after removing non-specific products.

After sequencing and alignment of sequences, we had the data available for further analysis: 553 bp long fragment of the translation elongation factor (*tef1α*) and 1031 bp long fragment of the RNA polymerase II (*rpb2*).

**Results of sequencing and BLAST search**

Sequence analysis of the *tef1α* fragments showed that no polymorphism was present amongst the strains; homology of sequences was 100%. Rodriguez Estrada *et al.* (2010) [14] found minor differences between the *tef1* sequences of var. *nebrodensis* and var. *eryngii* isolates which were used for discrimination of these varieties. Based on our results we can state that there was no var. *nebrodensis* amongst our isolates.

In the same study, Rodriguez Estrada *et al.* (2010) [14] found that differences in the *rpb2* gene are sufficient for discrimination of *eryngii, elaeoselini* and *ferulae* varieties. In our sequences we found only a few nucleotide substitutions (Table 1).

<table>
<thead>
<tr>
<th>Position of nucleotide</th>
<th>Nucleotide</th>
<th>Isolate</th>
</tr>
</thead>
<tbody>
<tr>
<td>21</td>
<td>C</td>
<td>Ple-1V and Ple-2V</td>
</tr>
<tr>
<td></td>
<td>T</td>
<td>other strains</td>
</tr>
<tr>
<td>372</td>
<td>G</td>
<td>Ple-1V, Ple-2V, PEG, Ple-3V, Ple-4V</td>
</tr>
<tr>
<td></td>
<td>A</td>
<td>other strains</td>
</tr>
<tr>
<td>957</td>
<td>G</td>
<td>Ple-1V, Ple-2V, Ple-3V, Ple-4V, PEG, PES</td>
</tr>
<tr>
<td></td>
<td>A</td>
<td>other strains</td>
</tr>
</tbody>
</table>

Aligned and BoxShade processed sequences are shown on Fig. 7. Since nucleotide substitutions might mean differences between varieties, we tried to check our sequences in the GenBank database.

Blasting the *tef1α* and *rpb2* sequences against the GenBank database resulted in interesting findings. Results of *tef1α* blasts showed that all of our isolates should be *P. fossulatus* with 100% identity and 0 E-value, and only 94% identity was found with *P. eryngii* sequences. Blasting with *rpb2* resulted that our isolates are *P. ostreatus* with 100% identity and 0 E-value. Similarity with *P. eryngii* var. *eryngii* sequences was only 70%, but the same value was found with var. *elaeoselini* and var. *ferulae* sequences, as well. Since these results were unexpected, we investigated a couple of sequences downloaded from the GenBank. Our analysis showed that there is uncertainty in certain extent in the data; hence, more sequences should be uploaded after thorough analysis in order to increase confidence of BLAST search results.

Concluding the results above, only the *rpb2* locus contained nucleotide substitutions in our isolates. Investigation of the sequences of both loci revealed nucleotide polymorphism only in a low degree. Our analysis showed that deposited sequences of *tef1α* and *rpb2* loci in the GenBank should be revised, because these sequences are not suitable for strain identification. At the same time it became clear that investigation of the ribosomal region, *tef1α* and *rpb2* genes is not sufficient for discrimination of strains and varieties of *P. eryngii*, more loci must be involved in phylogenetic experiments. It is clear that molecular identification of varieties still has some serious drawbacks, which need more time to eliminate.
<table>
<thead>
<tr>
<th>isolate</th>
<th>position</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ple_5V</td>
<td>TCACAATCAGGATGTTCCGTTTTACGTCAATGCTCGCTCCAGCAGGTTTGT</td>
</tr>
<tr>
<td>Ple_6V</td>
<td>TCACAATCAGGATGTTCCGTTTTACGTCAATGCTCGCTCCAGCAGGTTTGT</td>
</tr>
<tr>
<td>PEFI</td>
<td>TCACAATCAGGATGTTCCGTTTTACGTCAATGCTCGCTCCAGCAGGTTTGT</td>
</tr>
<tr>
<td>PE-SZM</td>
<td>TCACAATCAGGATGTTCCGTTTTACGTCAATGCTCGCTCCAGCAGGTTTGT</td>
</tr>
<tr>
<td>PEA</td>
<td>TCACAATCAGGATGTTCCGTTTTACGTCAATGCTCGCTCCAGCAGGTTTGT</td>
</tr>
<tr>
<td>PEFA</td>
<td>TCACAATCAGGATGTTCCGTTTTACGTCAATGCTCGCTCCAGCAGGTTTGT</td>
</tr>
<tr>
<td>PEL</td>
<td>TCACAATCAGGATGTTCCGTTTTACGTCAATGCTCGCTCCAGCAGGTTTGT</td>
</tr>
<tr>
<td>PEK_4</td>
<td>TCACAATCAGGATGTTCCGTTTTACGTCAATGCTCGCTCCAGCAGGTTTGT</td>
</tr>
<tr>
<td>PEC</td>
<td>TCACAATCAGGATGTTCCGTTTTACGTCAATGCTCGCTCCAGCAGGTTTGT</td>
</tr>
<tr>
<td>PEP</td>
<td>TCACAATCAGGATGTTCCGTTTTACGTCAATGCTCGCTCCAGCAGGTTTGT</td>
</tr>
<tr>
<td>PES</td>
<td>TCACAATCAGGATGTTCCGTTTTACGTCAATGCTCGCTCCAGCAGGTTTGT</td>
</tr>
<tr>
<td>Ple_E_1V</td>
<td>TCACAATCAGGATGTTCCGTTTTACGTCAATGCTCGCTCCAGCAGGTTTGT</td>
</tr>
<tr>
<td>Ple_2V</td>
<td>TCACAATCAGGATGTTCCGTTTTACGTCAATGCTCGCTCCAGCAGGTTTGT</td>
</tr>
<tr>
<td>Peg</td>
<td>TCACAATCAGGATGTTCCGTTTTACGTCAATGCTCGCTCCAGCAGGTTTGT</td>
</tr>
<tr>
<td>Ple_3V</td>
<td>TCACAATCAGGATGTTCCGTTTTACGTCAATGCTCGCTCCAGCAGGTTTGT</td>
</tr>
<tr>
<td>Ple_4V</td>
<td>TCACAATCAGGATGTTCCGTTTTACGTCAATGCTCGCTCCAGCAGGTTTGT</td>
</tr>
</tbody>
</table>

**Figure 7.** Aligned *rpb2* sequences of 16 *P. eryngii* isolates
Results of RFLP analysis.

*In silico* digestion showed that both restriction enzymes were able to divide the strains into two groups.

**BsmAI**

*Group 1*: 504 bp, 456 bp, 71 bp (Ple-5V, Ple-6V, PEF-i, PE-SZM, PEA, PEF, PEL, PEK, PEC, PEP, PES)

*Group 2*: 504 bp, 373 bp, 83 bp, 71 bp (Ple-1V, Ple-2V, Ple-3V, Ple-4V, PEG)

Results of the digestion by the real *BsmAI* enzyme were the same as the *in silico* digestion, but not each fragment was well visible on the agarose gel. In case of strains belonging to Group 2 (Ple-1V, Ple-2V, Ple-3V, Ple-4V and PEG), the 456 bp fragment is not present, though it is well visible at the other strains. So, this enzyme can be used for sorting the isolates into two groups (*Fig. 8*). It should be noted that fragment sizes in of *in silico* digestion are smaller than size of real fragments, because sequences were chopped prior to the theoretical digestion.

The *TspDTI* enzyme was not able to separate the strains (*Fig. 9*), though it was successfully performed during *in silico* digestion. During digestion a 9 bp long fragment is removed from the 313 bp fragment, but this difference is too small to visualize it agarose gel, so polyacrylamide electrophoresis or capillary electrophoresis is highly recommended in this case.

**TspDTI theoretical digestion**

*Group 1*: 586 bp, 304 bp, 132 bp, 9 bp (Ple-1V, Ple-2V)

*Group 2*: 586 bp, 313 bp, 132 bp (Ple-3V, Ple-4V, Ple-5V, Ple-6V, PEF-i, PE-SZM, PEA, PEF, PEL, PEK, PEC, PEP, PES, PEG)

**SUMMARY**

In this study we presented the sequence analysis of *tef1á* and *rpb2* genes of 16 *P. eryngii* isolates (of mostly Hungarian origin) in order to reveal variability between isolates and varieties.

Fragments from of the *tef1á* and *rpb2* regions were amplified by specific primer pairs (*tef1á*: EF595F/EF1160R; *rpb2*: bRPB2-6.9F/bRPB2-11R1). The amplicons were cleaned, sequenced and blasted against the GenBank database. After sequence analysis, theoretical *in silico* digestion (RFLP) was performed on the sequences, which was followed by real digestion with restriction enzymes. The isolates showed 100% homology in the *tef1á* region, and a few nucleotide substitutions were found amongst the *rpb2* sequences of the various strains (in position 21, 372 and 957). The real digestion experiments were done with *BsmAI* and *TspDTI* restriction endonucleases, because these two enzymes were able to divide the isolates.
into 2 groups in the *in silico* experiments. The real digestions were successful; though, fragments from the *TspDII* digestion are advised to separate on polyacrylamide gel or with capillary electrophoresis, because differences were not visible on agarose gel. Results of the BLAST search were ambiguous on varietas and species level, as well. Since the *rpb2* locus showed higher degree of polymorphism than the *tef1α*, it has the potential to be used in PCR-RFLP experiments for discrimination of strains or maybe even of varieties. Nonetheless, in the future it is necessary to involve more loci and molecular methods in phylogenetic experiments aiming the differentiation of *P. eryngii* strains on varietas level.

**REFERENCES**


