

RESEARCH ARTICLE

# Novel Introner-Like Elements in fungi Are Involved in Parallel Gains of Spliceosomal Introns

Jérôme Collemare<sup>1,2</sup>\*, Henriek G. Beenen<sup>1,3</sup>, Pedro W. Crous<sup>4</sup>, Pierre J. G. M. de Wit<sup>1</sup>, Ate van der Burgt<sup>1,3</sup>

**1** Laboratory of Phytopathology, Wageningen University, Wageningen, The Netherlands, **2** Present address: UMR1345 IRHS-INRA, Beaucouzé, France, **3** Present address: Dyadic, Wageningen, The Netherlands, **4** Evolutionary Phytopathology, CBS-KNAW Fungal Biodiversity Centre, Utrecht, The Netherlands

☞ These authors contributed equally to this work.

\* [j.collemare@gmail.com](mailto:j.collemare@gmail.com)



## Abstract

Spliceosomal introns are key components of the eukaryotic gene structure. Although they contributed to the emergence of eukaryotes, their origin remains elusive. In fungi, they might originate from the multiplication of invasive introns named Introner-Like Elements (ILEs). However, so far ILEs have been observed in six fungal species only, including *Fulvia fulva* and *Dothistroma septosporum* (*Dothideomycetes*), arguing against ILE insertion as a general mechanism for intron gain. Here, we identified novel ILEs in eight additional fungal species that are phylogenetically related to *F. fulva* and *D. septosporum* using PCR amplification with primers derived from previously identified ILEs. The ILE content appeared unique to each species, suggesting independent multiplication events. Interestingly, we identified four genes each containing two gained ILEs. By analysing intron positions in orthologues of these four genes in Ascomycota, we found that three ILEs had inserted within a 15 bp window that contains regular spliceosomal introns in other fungal species. These three positions are not the result of intron sliding because ILEs are newly gained introns. Furthermore, the alternative hypothesis of an inferred ancestral gain followed by independent losses contradicts the observed degeneration of ILEs. These observations clearly indicate three parallel intron gains in four genes that were randomly identified. Our findings suggest that parallel intron gain is a phenomenon that has been highly underestimated in ILE-containing fungi, and likely in the whole fungal kingdom.

## OPEN ACCESS

**Citation:** Collemare J, Beenen HG, Crous PW, de Wit PJGM, van der Burgt A (2015) Novel Introner-Like Elements in fungi Are Involved in Parallel Gains of Spliceosomal Introns. PLoS ONE 10(6): e0129302. doi:10.1371/journal.pone.0129302

**Academic Editor:** Jean-François Pombert, Illinois Institute of Technology, UNITED STATES

**Received:** February 10, 2015

**Accepted:** May 8, 2015

**Published:** June 5, 2015

**Copyright:** © 2015 Collemare et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Data Availability Statement:** All relevant data are within the paper and its Supporting Information files.

**Funding:** These authors have no support or funding to report.

**Competing Interests:** The authors have declared that no competing interests exist.

## Introduction

Spliceosomal introns are hallmarks of the eukaryotic gene structure that are involved in regulation of gene expression and diversification of the protein repertoire [1–3]. Despite their predicted key role in emergence and evolution of eukaryotes, their origin is still enigmatic. Studies

that attempted to trace the evolutionary history of introns since the last eukaryotic common ancestor (LECA) provided only a few hints. Indeed, modeling of intron dynamics, as well as intron gain and loss analyses across eukaryotes, indicate that intron losses prevail, suggesting that LECA was intron-rich [4–6]. However, these studies also suggest bursts of intron gains in certain lineages like in the ancestor of metazoans [6]. This indicates that intron gain might have prevailed in some lineages at particular times during evolution. For example, recent population genomics studies showed large-scale intron gains in the crustacean *Daphnia pulex* [7,8] and in the fungus *Zyoseptoria tritici* [9]. However, these events occur infrequently and so far reports on intron gains are scarce compared to those on intron losses. In this regard, complex presence-absence patterns of introns in orthologous genes are usually considered to be an indication of evolutionary conservation [10,11]. Accordingly, only recurrent losses explained complex presence-absence patterns in Angiosperms, and no recurrent gains could be inferred [12]. Therefore, independent gains of introns at the same position, also known as parallel gains, are expected to occur at an extremely low rate.

The discussion on the origin of spliceosomal introns has recently received new perspectives with the discovery of Introner Elements (IEs) in the alga *Micromonas pusilla* [13,14] and of Introner-Like Elements (ILEs) in six fungal species belonging to the *Dothideomycetes* [15]. Both IEs and ILEs are spliceosomal introns that have invaded algal and fungal genomes, where they are present in hundreds of near-identical copies in unrelated genes. Consistent with this observation, ILEs were shown to represent up to 90% of the most recent single intron gains in the fungus *Zyoseptoria tritici* [16], whereas other types of intron duplication, including tandem duplication within the same gene, contributed to less than 1% of single intron gains [16]. Similarly, more than 500 novel IE insertions were observed in marine metagenomes [14].

Identification of ILEs can be challenging because they rapidly become indistinguishable from regular spliceosomal introns (RSIs), which supports the hypothesis that ILEs are their predecessors [15]. ILE multiplication is likely still ongoing in *Fulvia fulva* (formerly *Cladosporium fulvum*) and *Z. tritici* because their genomes contain ILEs that are 99% and 98% identical, respectively. In contrast, ILEs in the genome of *Dothistroma septosporum* might have lost their ability to multiply because their pairwise identity is much lower [15]. In each fungus, ILEs could be grouped in up to eight different families based on sequence identity [15]. In a given species, some ILE families still comprise active elements with high pairwise identity while others do not. For example in *Z. tritici*, ILEs from the *mg01* and *mg02* families are likely active, while the *mg05* and *mg06* families seem to have lost their ability to multiply [15]. Thus, some ILEs appear still mobile, resulting in contemporary intron gains in several fungal species.

Here, we provide evidence for the presence of novel ILEs in eight non-sequenced fungal species using PCR amplification with primers derived from previously identified ILEs. Strikingly, we found parallel gains of ILEs at positions occupied by RSIs in distant fungal species. Our study shows that ILEs occur in many fungal species and can be responsible for parallel intron gains.

## Materials and Methods

### Phylogeny of fungal species

Internal transcribed spacer (ITS) and partial large-subunit (LSU) ribosomal DNA sequences of *F. fulva*, *D. septosporum*, *Pseudocercospora fijiensis*, *Z. tritici* and *Aspergillus niger* were retrieved from the Joint Genome Institute MycoCosm portal [17] (<http://genome.jgi.doe.gov>; August 29, 2013) with BLASTN using sequences from [S1 Table](#) as query.

All DNA sequences ([S1 File](#)) were aligned using MUSCLE [18] and poorly aligned regions were removed with Gblocks [19], allowing smaller final blocks, gap positions within the final

blocks and less strict flanking positions. The edited alignment was used to construct a maximum-likelihood phylogenetic tree using all sites in MEGA5 [20]. The Hasegawa-Kishino-Yano substitution model (uniform rate) was used and a hundred bootstrap replications were performed.

## Genomic DNA isolation and PCR

Fungal strains obtained from CBS-KNAW were grown in 50 mL PDB (BD) for 7 to 9 days at 22°C, shaking at 150 rpm. Mycelium was retrieved by filtering through miracloth, rinsed with water and snap frozen in liquid nitrogen. Mycelium was then ground with a tissue lyser (Retsch mix miller). About 100 mg of powder were incubated in 750  $\mu$ L of CTAB buffer (2% CTAB, 100 mM Tris pH 8, 20 mM EDTA and 1.4 M NaCl) for 1h at 57°C. Then 750  $\mu$ L of phenol/chloroform were added, and tubes were centrifuged at 10,000 x g for 10 min. The aqueous phase was transferred to a new tube and a second extraction was performed using 500  $\mu$ L of phenol/chloroform. Tubes were centrifuged at 10,000 x g for 10 min and the aqueous phase was transferred into a new tube. Genomic DNA was precipitated with 750  $\mu$ L of isopropanol, and tubes were centrifuged at 4,500 x g for 10 min. Supernatant was discarded and the pellet was washed with 70% ethanol. After centrifugation at 4,500 x g for 5 min, pellet was air-dried and dissolved in 100  $\mu$ L of TE overnight at 4°C. RNA was removed by adding 1  $\mu$ L of RNase A (10 mg/mL), incubating at 37°C for 2h. Genomic DNA was purified adding 200  $\mu$ L of phenol/chloroform. Tubes were centrifuged for 10 min at 10,000 x g and the aqueous phase was transferred into a new tube. DNA was precipitated by adding 200  $\mu$ L of isopropanol and centrifugation at 4,500 x g for 10 min. Pellet was washed with 70% ethanol, centrifuged at 4,500 x g for 5 min and air-dried. The pellet was dissolved in 100  $\mu$ L of water. Concentration was evaluated using a Nanodrop 1000 spectrophotometer (Thermo Scientific).

To design PCR primers for each family, ILEs exhibiting more than 80% pairwise identity (according to van der Burgt *et al.*, 2012) were aligned with MUSCLE [18]. Conserved nucleotide stretches that might serve as primers were identified. Primers could be designed to specifically amplify ILEs from *cf01* and *cf02* families (S2 Table). For ILE families that are shared by *F. fulva* and *D. septosporum*, degenerated primers had to be designed, using sequence stretches with at most four variable nucleotides (S2 Table).

About 50 ng of genomic DNA were used as template for PCR using GoTaq Flexi (Promega), 1 mM MgCl<sub>2</sub> and 5 pmol of each primer (S2 Table). After 2 min at 95°C, 35 cycles of the following program were used: 1 min at 95°C, 1 min at 55°C and 1 min at 72°C, followed by a final step of 5 min at 72°C.

## PCR fragment purification and sequencing

PCR products were either run on 2% agarose gels or 20% acrylamide gels (run for 7–8h at 50 V). Acrylamide gels were stained with 5,000 times diluted GelRed for 30 min shaking. Gels were washed with water for 15 min shaking prior to imaging.

PCR fragments were purified from agarose gels using the Wizard Gel and PCR clean-up system kit (Promega) according to the manufacturer's recommendations. Fragments were cloned into PGEM-T vector (Promega) and transformed into DH5 $\alpha$  *Escherichia coli* cells [21]. Plasmids were recovered from eight white colonies and insert size was checked by restriction digestion. Plasmids with inserts of different sizes were sent for sequencing (Macrogen).

PCR fragments were also recovered from acrylamide gels. Cut gel slices were crushed in tubes and two volumes of elution buffer were added (3.85 g ammonium acetate, 0.215 g magnesium acetate, 200  $\mu$ L 0.5 M EDTA, 1 mL 10% SDS in 100 mL water). Tubes were incubated at 37°C for 4h on a rotating wheel and then centrifuged at 10,000 x g for 1 min at 4°C.

Supernatant was transferred into a new tube and 0.5 volume of elution buffer was added to the pellet. After brief vortex, tubes were centrifuged at 10,000 x g for 1 min at 4°C. Both supernatants were combined and two volumes of cold 100% ethanol were added. After incubation on ice for 30 min, DNA was pelleted by centrifugation at 10,000 x g for 10 min at 4°C. Supernatant was discarded and pellet dissolved in 200 µL of TE. DNA was purified by adding 25 µL of 3 M sodium acetate pH 5.2 and two volumes of cold 100% ethanol. After incubation on ice for 30 min, DNA was recovered by centrifugation at 10,000 x g for 10 min at 4°C. Supernatant was discarded and pellet washed with 70% ethanol and centrifuged at 10,000 x g for 1 min at 4°C. The pellet was air-dried and dissolved in 10 µL of TE. 1 µL of this purified DNA fragment was used for a PCR reaction as described above, which was directly sent for sequencing (CBS; fragments and PCR primers mixed together with a BigDye Terminator Cycle Sequencing Kit v. 3.1 (Applied Biosystems) and analysed on an ABI Prism 3100 DNA Sequencer (Perkin-Elmer) [22]).

Sequences obtained from the PCR fragments were trimmed (S2 File) and aligned to the consensus sequences of ILE families using Muscle [18]. Obtained alignments were edited in GeneDoc [23].

## Identification of orthologues and intron landscape analysis

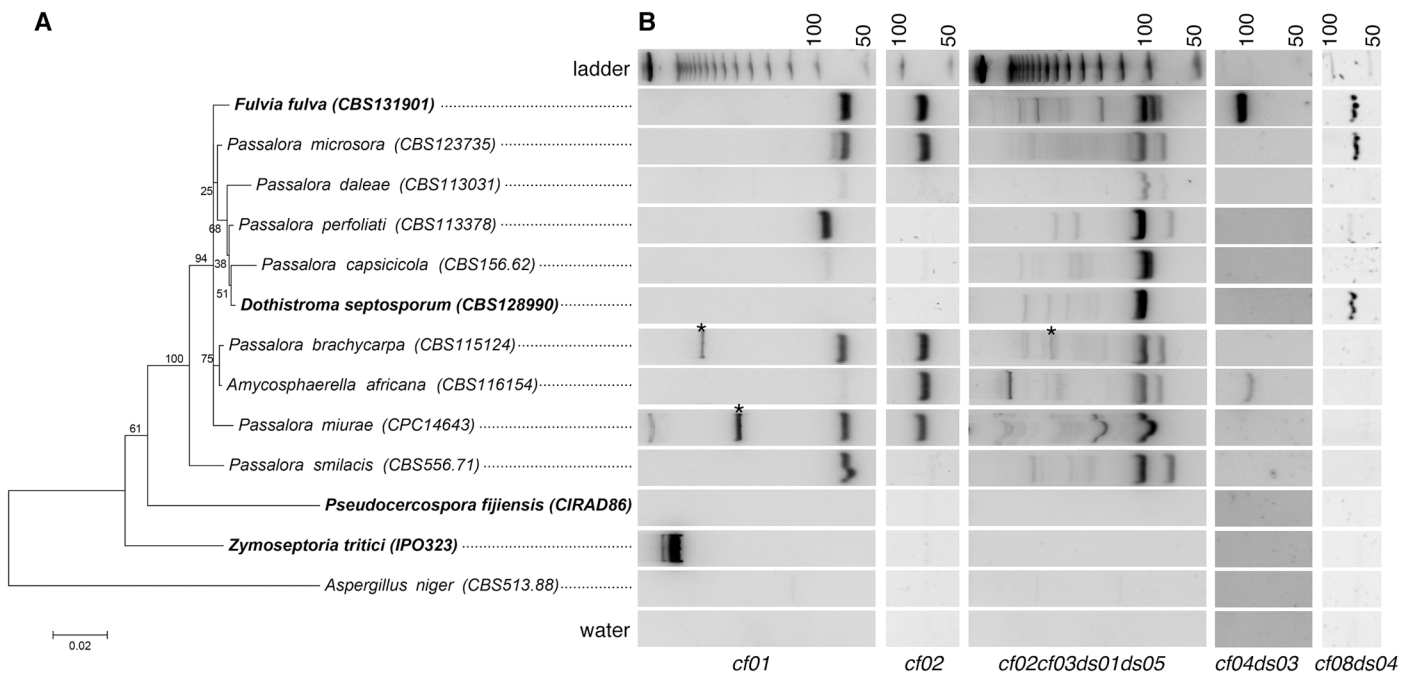
When two different ILEs were amplified in a PCR fragment, the exonic sequence in between was used as a query for a BLASTN search (word size 4, no filtering for low complexity regions and performing gapped alignment) in the masked assemblies of *F. fulva* and *D. septosporum* (<http://genome.jgi.doe.gov>) [24] and in the nr database at NCBI ([www.ncbi.nlm.nih.gov](http://www.ncbi.nlm.nih.gov)). The predicted protein sequence of Fulfu186212, Fulfu193200, Dotse59237 and Zymtr43851, in which new ILE insertions were detected, were used to search for homologues in all Dikarya fungal genomes available at the JGI MycoCosm portal using BLASTP (October 4, 2013) [17]. All hits were aligned using Muscle [18]. The obtained alignments were used to manually remove predicted proteins that contained large deletions or insertions. This step has been iteratively performed until alignments of predicted proteins of similar length were obtained. These alignments were then used to construct minimum evolution phylogenetic trees in MEGA5 [20], using default parameters with the Jones-Taylor-Thornton (JTT) amino acid substitution model, 100 bootstraps replications and gap site complete deletion. Phylogenetic trees were used to discriminate between orthologues and paralogues. Distant homologues were removed and new alignments and phylogenetic trees were constructed in order to confidently identify orthologues. Putative orthologues from Basidiomycota were used as outgroup. The final alignments and phylogenetic trees were performed as described above, with the exception that a maximum likelihood tree was built, using all sites.

For each selected protein, the genomic DNA sequence was retrieved at JGI, from the predicted start to stop codons. DNA sequences were aligned using MUSCLE [18] and manually edited in GeneDoc [23] in order to correct misaligned exon-intron-exon boundaries. This step is particularly needed when two introns are separated by only a few nucleotides in different fungal species. Intron positions were determined using DNA alignments and mapped onto protein alignments. Conserved domains were sought in the NCBI Conserved Domain Database [25] using the predicted protein sequence of Fulfu186212, Fulfu193200, Dotse59237 and Zymtr43851.

## Results and Discussion

### Specific amplification of Introner-Like Elements in non-sequenced fungal species

*F. fulva* and *D. septosporum* are closely related fungal species with different ILE activities, as suggested by the higher pairwise identity of certain ILE families in *F. fulva* [15]. However, both species also share some ILE families [15]. We hypothesized that ILEs with significant sequence similarity to elements from described families might be present in other related fungal species. Thus, we assessed by PCR on genomic DNA the presence of ILEs in eight non-sequenced fungal species that are phylogenetically related to *F. fulva* and *D. septosporum* (Fig 1A). Genomic DNA of the latter two species was used as positive controls and DNA of the more distantly related species *Pseudocercospora fijiensis* (formerly *Mycosphaerella fijiensis*), *Z. tritici* and *Aspergillus niger* as negative controls. Specific oligonucleotides were designed to amplify ILEs related to *F. fulva* cf01 and cf02 families; degenerated oligonucleotides were designed to amplify ILEs related to families shared by *F. fulva* and *D. septosporum* (cf02cf03ds01ds05, cf04ds03 and cf08ds04) (S2 Table). As expected, all oligonucleotide pairs amplified fragments in *F. fulva*, but not in the negative controls, proving their specificity (Fig 1B). The large fragment amplified in *Z. tritici* with the cf01 oligonucleotide pair corresponds to the 5'UTR and start of the gene containing the mg01020 ILE as revealed by sequencing (S2 File). The cf01 and cf02 pairs did not amplify fragments in *D. septosporum*, again proving their specificity. Degenerated oligonucleotide pairs amplified fragments in both *F. fulva* and *D. septosporum*, except for the cf04ds03 pair that is more specific to *F. fulva* ILEs (Fig 1B). PCR results obtained for other fungal species



**Fig 1. Detection of Introner-Like Elements (ILEs) in closely related dothideomycetous fungal species.** (A) Maximum-likelihood phylogenetic tree using ITS and LSU sequences. Fungal species in which ILEs have been previously identified are highlighted in bold. *Aspergillus niger* belongs to the *Eurotiomycetes* and serves to root the tree. Accession numbers of fungal species from the CBS-KNAW collection are indicated in between brackets. The scale bar indicates the number of substitutions per site. (B) PCR with primers specific to single (cf01 and cf02) or shared (cf02cf03ds01ds05, cf04ds03 and cf08ds04) ILE families between *F. fulva* and *Dothistroma septosporum* was performed using genomic DNA. PCR products were run on 20% acrylamide gels. The first row shows a 50 bp-step DNA ladder and the last row shows the water control. Asterisks indicate fragments that correspond to two different ILEs as revealed by sequencing.

doi:10.1371/journal.pone.0129302.g001



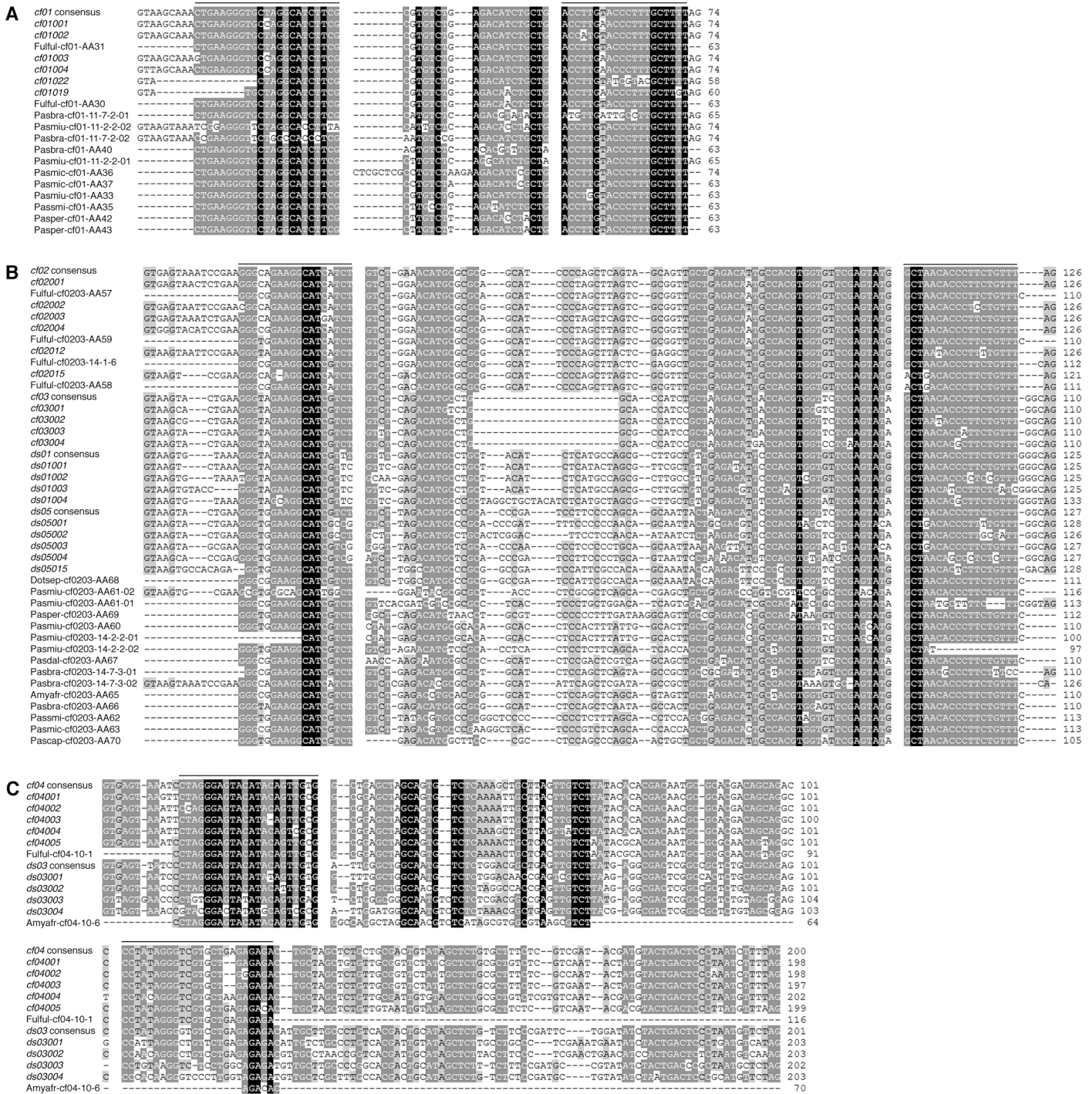
revealed diverse amplification profiles, suggesting the presence of ILEs in these non-sequenced genomes (Fig 1B). Sequencing of nearly all PCR fragments shown in Fig 1B confirmed that they are true ILEs (Fig 2 and S2 File). It is noteworthy that the acrylamide gel resolution did not allow the separation of ILEs that differ in only a few nucleotides in length, as shown by the cloning of different *F. fulva* ILEs from the same PCR fragment (Fig 2A). In addition to fragments of the expected sizes corresponding to the different ILE families [15], larger fragments were also amplified with the *cf01* and *cf02cf03ds01ds05* oligonucleotide pairs. These likely correspond to related ILEs of different size or to different ILEs inserted in the same locus. Remarkably, the amplification profile of each fungal species is unique (Fig 1B) and all amplified ILEs differed in sequence (Fig 2). Moreover, for a given family, signal intensity of the PCR fragments varied, suggesting either degenerated ILEs or fewer ILEs in species like *Passalora daleae*, *Passalora capsicicola*, *Amycosphaerella africana* and *Passalora smilacis* (Fig 1B). Altogether, these results show that ILEs are highly dynamic and that many likely represent unique intron gains.

### Identification of genes with recent multiple ILE insertions

Sequencing of four larger *cf01* and *cf02cf03ds01ds05* fragments from *Passalora brachycarpa* and *Passalora miurae* revealed in each case the presence of two ILEs located next to each other (between 145 and 626 bp apart; Fig 1B, Fig 3A–3D and S2 File). The exonic sequences between the two ILEs correspond to four different genes that putatively encode a nucleoside transporter, a peroxidase, a hydroxylase/oxidoreductase and a fungal transcription factor, respectively. All four genes belong to multigene families because many paralogues could be identified, even in distant species belonging to the Basidiomycota. However, only a few orthologues could be assigned confidently, mainly within the *Dothideomycetes* (Fig 3A–3D and S3 File). An intron landscape was determined for each of the four genes by mapping all intron positions identified in Ascomycota. The analysis revealed a majority of single presence intron positions (16 out of 39 analyzed intron positions, and up to 25 positions when single presences in a monophyletic clade were included) (Fig 3E). This finding is similar to previous reports in other eukaryotes that have experienced high rates of intron gains such as *Oikopleura dioica* and *Thalassiosira pseudonana*, in which more than 75% of intron positions are unique [26,27]. Thirteen positions show complex presence-absence patterns and a single position represents a single absence in a given species (Fig 3E). These results suggest a balanced number of putative intron gains and losses, consistent with previous intron gain and loss analyses in fungi [6,9,15,28,29].

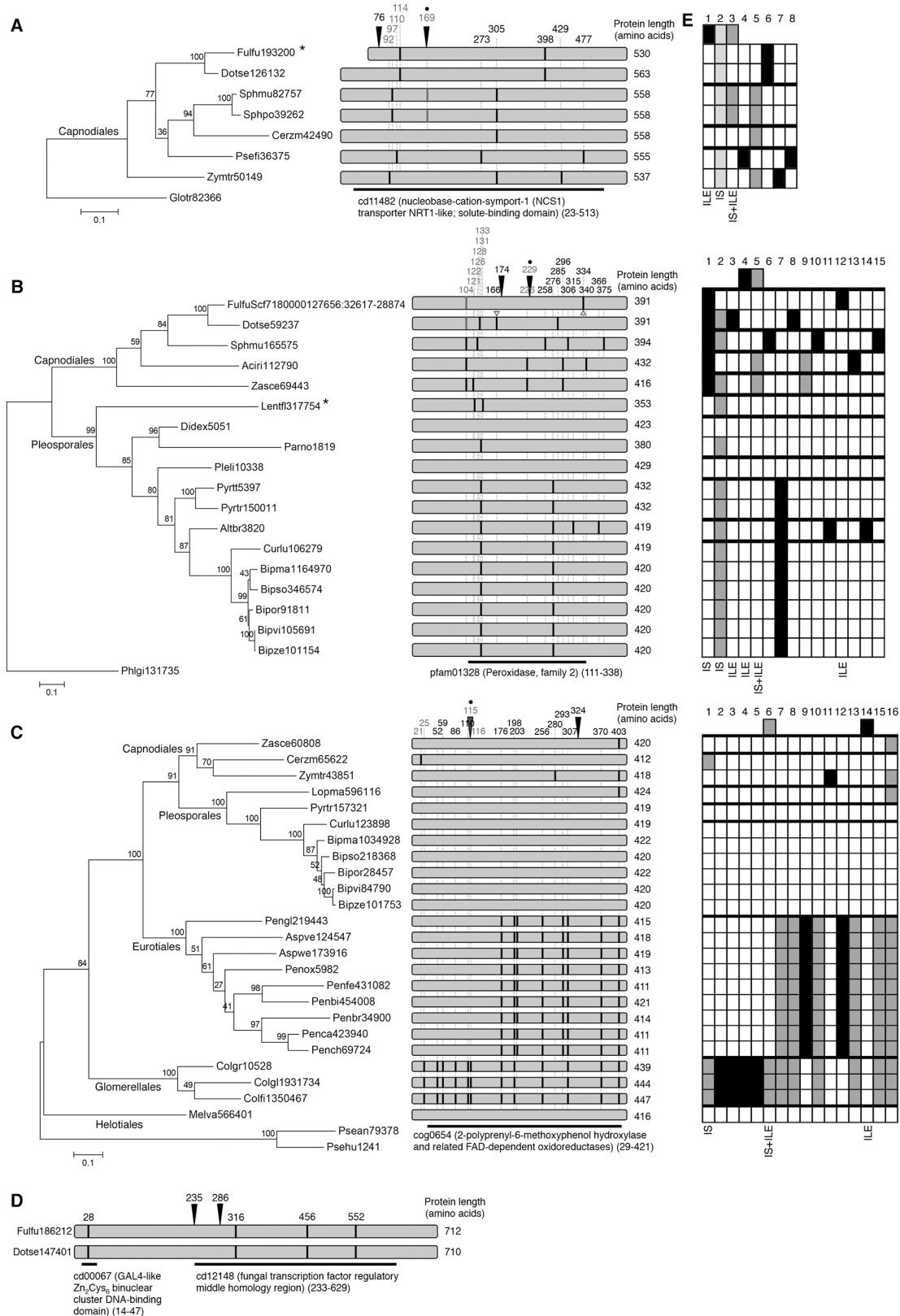
The two ILEs that inserted in the transporter gene in *P. brachycarpa* are located on each side of a conserved intron in *F. fulva* and *D. septosporum* (Fig 3A). This intron is conserved in *P. brachycarpa* (S2 and S3 Files), suggesting that ILEs recently inserted in this gene. Orthologues of the peroxidase gene in *F. fulva* and *D. septosporum* also contain ILEs, each at different positions (Fig 3B) [14]. The occurrence of four distinct ILEs in an orthologous gene of three different fungal species suggests that this gene contains unknown features that make it very attractive for ILE insertion. All observed ILE insertions occurred at positions lacking introns in all other species (Fig 3). Clearly these ILEs are recent intron gains in *P. brachycarpa* and *P. miurae*.

Intron landscape analysis revealed six positions where intron sliding might have occurred (Fig 3). Intron sliding is a poorly understood phenomenon that refers to intron positions separated by less than 15 nucleotides [30]. It has been proposed that these positions correspond to one and the same intron that would have slightly shifted due to splice site mutations or compensatory insertion and deletion, or would have been lost and immediately reinserted [31]. Three ILEs (positions 169, 229 and 115 in Fig 3A, 3B and 3C, respectively) are separated by one, eight and three nucleotides, respectively, from a position occupied by an RSI in other



**Fig 2. Alignments of newly discovered Introner-Like Elements (ILEs).** DNA sequences obtained from the PCR fragments were aligned, as references the consensus sequence and the four most conserved ILE sequences of the corresponding families in *F. fulva* and *Dothistroma septosporum*. The complete sequences of known ILEs amplified by PCR in this study are also included. Alignments are shown for ILEs related to (A) *cf01*, (B) *cf02cf03ds01ds05* and (C) *cf08ds04* families. Bars above the alignments indicate the oligonucleotide sequences. *Fulful*: *Fulvia fulva*; *Dotsep*: *Dothistroma septosporum*; *Amyafr*: *Amycosphaerella africana*; *Pasbra*: *Passalora brachycarpa*; *Paspcap*: *Passalora capsicicola*; *Pasdal*: *Passalora daleae*; *Pasmic*: *Passalora microsora*; *Pasmiu*: *Passalora miurae*; *Pasper*: *Passalora perfoliati*; *Pasmi*: *Passalora smilacis*.

doi:10.1371/journal.pone.0129302.g002



**Fig 3. Analysis of intron positions in genes with recent insertions of Introner-Like Elements (ILE).** ILEs from *Passalora brachycarpa* inserted in genes encoding (A) a transporter and (B) a peroxidase. ILEs from *Passalora miurae* inserted in genes encoding (C) a hydroxylase/oxidoreductase and (D) a fungal



transcription factor. For each gene, a maximum likelihood phylogenetic tree was constructed with the predicted protein sequence of orthologues. The trees were rooted with the closest homologue found in Basidiomycota. Bootstrap values of 100 repeats are shown. Scale bar represents the number of substitutions per site. The numbers correspond to the protein ID from the Joint Genome Institute mycocosm portal, except for one gene that was not predicted in the *Fulvia fulva* genome and for which genomic coordinates are given (B). Orders in fungal classification are mentioned in the trees. On the right, diagrams depict aligned protein sequences and intron positions are indicated as black bars. Their positions in the protein alignment that served to build the phylogenetic trees are indicated above. Positions that are shown in grey highlight putative intron sliding. The black arrows indicate the positions where ILEs inserted in the genes of *P. miuriae* or *P. brachycarpa*. The open triangles indicate previously identified ILEs. Dots indicate positions where parallel intron gains have occurred. The black bar below each protein representation indicates conserved domains (positions in the protein alignments are indicated between brackets). Asterisks behind species names indicate genes that are likely pseudogenes because of an in frame stop codon. (E) Schematic overview of intron positions (numbered on top) in three genes. The first row shows intron positions in genes of *P. miuriae* or *P. brachycarpa*. Thick lines indicate monophyletic clades according to the phylogenetic trees. Black, dark grey and light grey squares indicate single presence of an intron position in a monophyletic clade, presence-absence polymorphism and single absence of an intron position in a monophyletic clade, respectively. White squares indicate absence of intron. The presence of ILE and occurrence of putative intron splicing (IS) are indicated below each scheme. *Aciri*: *Acidomyces richmondensis*; *Altbr*: *Alternaria brassicicola*; *Cerzm*: *Cercospora zaeae-maydis*; *Fulfu*: *Fulvia fulva*; *Bipze*: *Bipolaris zeicola*; *Bipma*: *Bipolaris maydis*; *Curlu*: *Curvularia lunata*; *Bipor*: *Bipolaris oryzae*; *Bipso*: *Bipolaris sorokiniana*; *Bipvi*: *Bipolaris victoriae*; *Didex*: *Didymella exigua*; *Dotse*: *Dothistroma septosporum*; *Lentfl*: *Lentithecium fluviatile*; *Pleli*: *Plenodomus lingam*; *Lopma*: *Lophiostoma macrostomum*; *Psefi*: *Pseudocercospora fijiensis*; *Pyrtr*: *Pyrenophora tritici-repentis*; *Pyrtrt*: *Pyrenophora teres f. teres*; *Sphmu*: *Sphaerulina musiva*; *Sphpo*: *Sphaerulina populicola*; *Parno*: *Parastagonospora nodorum*; *Zasce*: *Zasmidium cellare*; *Zymtr*: *Zymoseptoria tritici*; *Aspve*: *Aspergillus versicolor*; *Aspwe*: *Aspergillus wentii*; *Penbi*: *Penicillium bilaiae*; *Penbr*: *Penicillium brevicompactum*; *Penca*: *Penicillium canescens*; *Pench*: *Penicillium chrysogenum*; *Penfe*: *Penicillium fellutanum*; *Pengl*: *Penicillium glabrum*; *Penox*: *Penicillium oxalicum*; *Colgr*: *Colletotrichum graminicola*; *Colfi*: *Colletotrichum fiorinae*; *Colgl*: *Colletotrichum gloeosporioides*; *Melva*: *Meliniomyces variabilis*; *Glotr*: *Gloeophyllum trabeum*; *Phlgi*: *Phlebiopsis gigantea*; *Psean*: *Pseudozyma antarctica*; *Psehu*: *Pseudozyma hubeiensis*.

doi:10.1371/journal.pone.0129302.g003

fungal species (S3 File). According to the hypothesis of intron sliding, these three ILEs would be wrongly assigned as conserved ancestral introns. For each position, this hypothesis implies the ancestral insertion of an ILE that would have degenerated and, most often, been lost in all fungal species, but would not have degenerated in *P. brachycarpa* or *P. miuriae*. This hypothesis strongly contradicts the observed fast degeneration of ILEs [15] and the observation that lost introns tend to be short in fungi while ILEs are long [16]. Therefore, we propose that parallel gains of two different introns are the most parsimonious explanation for these three positions. In the transporter gene, such parallel gain might have occurred very recently (Fig 3A). Indeed, although introns in *Sphaerulina* species do not share significant sequence similarity with ILEs, they are longer (81 and 87 bp; S3 File) than RSIs (50–55 bp) in fungi, and thus might represent degenerated ILEs that had inserted in the common ancestor of both *Sphaerulina* species [14].

Depending on the evolutionary model used to infer rates of intron gains, parallel gains were predicted to account for 5–10% to most of shared intron positions [10,32,33]. However, studies on genes that had been transferred from mitochondria to the nucleus revealed a very limited number of shared positions, indicating a frequency of parallel gains as low as 2.3% [34,35]. Similarly, parallel gains were estimated to contribute to only 4.2% of intron positions in *T. pseudonana* [26]. Our results are in striking conflict with the low number of shared positions in species that experienced high rates of intron gains as the estimated parallel gain frequency in our small dataset is 23%. However, this discrepancy likely reflects that all three ILEs inserted in positions where intron sliding might have occurred. These observations suggest that actually parallel gains might have occurred at positions where intron sliding has been proposed in Fig 3. This is particularly true for the first intron position in the oxidoreductase gene, which is occupied by an intron in *Cercospora zaeae-maydis* and in the distant species of the order *Glomerellales* (Fig 3C). In addition, the second intron position in the peroxidase gene is considered as conserved because of the occurrence of many introns within 15 nucleotides (Fig 3B). However, this position is occupied by two introns in *Lentithecium fluviatile*, suggesting that they have resulted from at least two independent gain events during the evolutionary history of *Dothideomycetes*.

## Conclusions

Our study revealed the presence of novel ILEs in eight additional non-sequenced fungal species. It suggests that ILEs are still mobile in these species where they contributed to recent intron gains. It is remarkable that amplification of novel ILEs resulted in the identification of four genes with high spliceosomal intron dynamics. The variation in intron positions in *Dothideomycetes* is consistent with ILE multiplication in this class of fungi. It was previously suggested that intron positions shared between fungi, animals and plants reflect evolutionary conservation rather than parallel gains [10,11]. However, we report three clear examples of parallel intron gains in fungi by studying intron positions of four genes only. This finding is consistent with previous reports, although mostly reported in single genes, of independent intron gains in diverse eukaryotic lineages [7,36–41]. Parallel gains were also recently reported in *Daphnia pulex* populations, in which they contributed to nearly half of all the newly discovered introns [8]. Altogether, our observations and those by others suggest that parallel intron gains might occur more frequently than previously thought in ILE-containing fungi, but also in the whole fungal kingdom and distant eukaryotes.

## Supporting Information

**S1 File. Sequences used to build the fungal species phylogenetic tree.**  
(PDF)

**S2 File. Sequences amplified by PCR.**  
(PDF)

**S3 File. DNA and amino acid alignments built to determine intron landscapes.**  
(PDF)

**S1 Table. GenBank accession numbers of the ITS and LSU sequences used in this study.**  
(PDF)

**S2 Table. Primers used in this study.**  
(PDF)

## Acknowledgments

We acknowledge Dr. Ewald Groenewald (CBS-KNAW, Utrecht) for providing the ITS and LSU sequences of non-sequenced fungal species, and Dr. Ronald de Vries (CBS-KNAW, Utrecht) for HB support.

## Author Contributions

Conceived and designed the experiments: JC PC PJGMW AVDB. Performed the experiments: JC HB. Analyzed the data: JC HB AVDB. Contributed reagents/materials/analysis tools: PC PJGMW. Wrote the paper: JC PJGMW.

## References

1. Le Hir H, Nott A, Moore MJ. How introns influence and enhance eukaryotic gene expression. *Trends Biochem Sci.* 2003; 28: 215–220. PMID: [12713906](#)
2. Koonin EV. The origin of introns and their role in eukaryogenesis: a compromise solution to the introns-early versus introns-late debate? *Biol Direct.* 2006; 1: 22. PMID: [16907971](#)
3. Warf MB, Berglund JA. Role of RNA structure in regulating pre-mRNA splicing. *Trends Biochem. Sci.* 2010; 35: 169–178. doi: [10.1016/j.tibs.2009.10.004](#) PMID: [19959365](#)

4. Roy SW, Gilbert W. Rates of intron loss and gain: implications for early eukaryotic evolution. *Proc Natl Acad Sci USA*. 2005; 102: 5773–5778. PMID: [15827119](#)
5. Roy SW, Irimia M. Mystery of intron gain: new data and new models. *Trends Genet*. 2009; 25: 67–73. doi: [10.1016/j.tig.2008.11.004](#) PMID: [19070397](#)
6. Csuros M, Rogozin IB, Koonin EV. A detailed history of intron-rich eukaryotic ancestors inferred from a global survey of 100 complete genomes. *PLoS Comput Biol*. 2011; 7: e1002150. doi: [10.1371/journal.pcbi.1002150](#) PMID: [21935348](#)
7. Li W, Tucker AE, Sung W, Thomas WK, Lynch M. Extensive, recent intron gains in *Daphnia* populations. *Science*. 2009; 326: 1260–1262. doi: [10.1126/science.1179302](#) PMID: [19965475](#)
8. Li W, Kuzoff R, Wong K, Tucker A, Lynch M. Characterization of newly gained introns in *Daphnia* populations. *Genome Biol Evol*. 2014; 6: 2218–2234. doi: [10.1093/gbe/evu174](#) PMID: [25123113](#)
9. Torriani SF, Stukenbrock EH, Brunner PC, McDonald BA, Croll D. Evidence for extensive recent intron transposition in closely related fungi. *Curr Biol*. 2011; 21: 2017–2022. doi: [10.1016/j.cub.2011.10.041](#) PMID: [22100062](#)
10. Sverdlov AV, Rogozin IB, Babenko VN, Koonin EV. Conservation versus parallel gains in intron evolution. *Nucleic Acids Res*. 2005; 33: 1741–1748. PMID: [15788746](#)
11. Carmel L, Rogozin IB, Wolf YI, Koonin EV. Patterns of intron gain and conservation in eukaryotic genes. *BMC Evol Biol*. 2007; 7: 192. PMID: [17935625](#)
12. Wang H, Devos KM, Bennetzen JL. Recurrent loss of specific introns during angiosperm evolution. *PLoS Genet*. 2014; 10: e1004843. doi: [10.1371/journal.pgen.1004843](#) PMID: [25474210](#)
13. Worden AZ, Lee JH, Mock T, Rouzé P, Simmons MP, Aerts AL, et al. Green evolution and dynamic adaptations revealed by genomes of the marine picoeukaryotes *Micromonas*. *Science*. 2009; 324: 268–272. doi: [10.1126/science.1167222](#) PMID: [19359590](#)
14. Verhelst B, van de Peer Y, Rouzé P. The complex intron landscape and massive intron invasion in a picoeukaryote provides insights into intron evolution. *Genome Biol Evol*. 2013; 5: 2393–2401. doi: [10.1093/gbe/evt189](#) PMID: [24273312](#)
15. van der Burgt A, Severing E, de Wit PJGM, Collemare J. Birth of new spliceosomal introns in fungi by multiplication of introner-like elements. *Curr Biol*. 2012; 22: 1260–1265. doi: [10.1016/j.cub.2012.05.011](#) PMID: [22658596](#)
16. Collemare J, van der Burgt A, de Wit PJGM. At the origin of spliceosomal introns: is multiplication of introner-like elements the main mechanism of intron gain in fungi? *Commun Integr Biol*. 2013; 6: e23147. doi: [10.4161/cib.23147](#) PMID: [23750299](#)
17. Grigoriev IV, Nikitin R, Haridas S, Kuo A, Ohm R, Otilar R, et al. MycoCosm portal: gearing up for 1000 fungal genomes. *Nucleic Acids Res*. 2014; 42: D699–704. doi: [10.1093/nar/gkt1183](#) PMID: [24297253](#)
18. Edgar RC. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res*. 2004; 32: 1792–1797. PMID: [15034147](#)
19. Castresana J. Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis. *Mol Biol Evol*. 2000; 17: 540–552. PMID: [10742046](#)
20. Tamura K, Peterson D, Peterson N, Stecher G, Nei M, Kumar S. MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Mol Biol Evol*. 2011; 28: 2731–2739. doi: [10.1093/molbev/msr121](#) PMID: [21546353](#)
21. Sambrook J, Russell DW. *Molecular cloning: a laboratory manual*. 3rd ed. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press; 2001.
22. Cheewangkoon R, Crous PW, Hyde KD, Groenewald JZ, To-Anan C. Species of *Mycosphaerella* and related anamorphs on *Eucalyptus* leaves from Thailand. *Persoonia*. 2008; 21: 77–91. doi: [10.3767/003158508X370857](#) PMID: [20396579](#)
23. Nicholas KB, Nicholas HB Jr, Deerfield DW II. GeneDoc: analysis and visualization of genetic variation. *EMBNEW.NEWS*. 1997; 4: 14.
24. de Wit PJGM, van der Burgt A, Ökmen B, Stergiopoulos I, Abd-Elsalam KA, Aerts AL, et al. The genomes of the fungal plant pathogens *Cladosporium fulvum* and *Dothistroma septosporum* reveal adaptation to different hosts and lifestyles but also signatures of common ancestry. *PLoS Genet*. 2012; 8: e1003088. doi: [10.1371/journal.pgen.1003088](#) PMID: [23209441](#)
25. Marchler-Bauer A, Derbyshire MK, Gonzales NR, Lu S, Chitsaz F, Geer LY, et al. CDD: NCBI's conserved domain database. *Nucleic Acids Res*. 2015; 43: D222–6. doi: [10.1093/nar/gku1221](#) PMID: [25414356](#)
26. Roy SW, Penny D. A very high fraction of unique intron positions in the intron-rich diatom *Thalassiosira pseudonana* indicates widespread intron gain. *Mol Biol Evol*. 2007; 24: 1447–1457. PMID: [17350938](#)

27. Denoed F, Henriot S, Mungpakdee S, Aury JM, Da Silva C, Brinkmann H, et al. Plasticity of animal genome architecture unmasked by rapid evolution of a pelagic tunicate. *Science*. 2010; 330: 1381–1385. doi: [10.1126/science.1194167](https://doi.org/10.1126/science.1194167) PMID: [21097902](https://pubmed.ncbi.nlm.nih.gov/21097902/)
28. Nielsen CB, Friedman B, Birren B, Burge CB, Galagan JE. Patterns of intron gain and loss in fungi. *PLoS Biol*. 2004; 2: e422. PMID: [15562318](https://pubmed.ncbi.nlm.nih.gov/15562318/)
29. Sun Y, Whittle CA, Corcoran P, Johannesson H. Intron evolution in *Neurospora*: the role of mutational bias and selection. *Genome Res*. 2015; 25: 100–110. doi: [10.1101/gr.175653.114](https://doi.org/10.1101/gr.175653.114) PMID: [25342722](https://pubmed.ncbi.nlm.nih.gov/25342722/)
30. Rogozin IB, Lyons-Weiler J, Koonin EV. Intron sliding in conserved gene families. *Trends Genet*. 2000; 16: 430–432. PMID: [11050324](https://pubmed.ncbi.nlm.nih.gov/11050324/)
31. Stoltzfus A, Logsdon JM Jr, Palmer JD, Doolittle WF. Intron "sliding" and the diversity of intron positions. *Proc Natl Acad Sci U S A*. 1997; 94: 10739–10744. PMID: [9380704](https://pubmed.ncbi.nlm.nih.gov/9380704/)
32. Qiu WG, Schisler N, Stoltzfus A. The evolutionary gain of spliceosomal introns: sequence and phase preferences. *Mol Biol Evol*. 2004; 21: 1252–1263. PMID: [15014153](https://pubmed.ncbi.nlm.nih.gov/15014153/)
33. Nguyen HD, Yoshihama M, Kenmochi N. New maximum likelihood estimators for eukaryotic intron evolution. *PLoS Comput Biol*. 2005; 1: e79. PMID: [16389300](https://pubmed.ncbi.nlm.nih.gov/16389300/)
34. Yoshihama M, Nakao A, Nguyen HD, Kenmochi N. Analysis of ribosomal protein gene structures: implications for intron evolution. *PLoS Genet*. 2006; 2: e25. PMID: [16518464](https://pubmed.ncbi.nlm.nih.gov/16518464/)
35. Ahmadinejad N, Dagan T, Gruenheit N, Martin W, Gabaldón T. Evolution of spliceosomal introns following endosymbiotic gene transfer. *BMC Evol Biol*. 2010; 10: 57. doi: [10.1186/1471-2148-10-57](https://doi.org/10.1186/1471-2148-10-57) PMID: [20178587](https://pubmed.ncbi.nlm.nih.gov/20178587/)
36. Tarrio R, Rodríguez-Trelles F, Ayala FJ. A new *Drosophila* spliceosomal intron position is common in plants. *Proc Natl Acad Sci U S A*. 2003; 100: 6580–6583. PMID: [12750476](https://pubmed.ncbi.nlm.nih.gov/12750476/)
37. Krauss V, Pecyna M, Kurz K, Sass H. Phylogenetic mapping of intron positions: a case study of translation initiation factor eIF2gamma. *Mol Biol Evol*. 2005; 22: 74–84. PMID: [15356279](https://pubmed.ncbi.nlm.nih.gov/15356279/)
38. Flakowski J, Bolivar I, Fahmi J, Pawlowski J. Tempo and mode of spliceosomal intron evolution in actin of foraminifera. *J Mol Evol*. 2006; 63: 30–41. PMID: [16755352](https://pubmed.ncbi.nlm.nih.gov/16755352/)
39. Da Lage JL, Maczkowiak F, Cariou ML. Phylogenetic distribution of intron positions in alpha-amylase genes of bilateria suggests numerous gains and losses. *PLoS One*. 2011; 6: e19673. doi: [10.1371/journal.pone.0019673](https://doi.org/10.1371/journal.pone.0019673) PMID: [21611157](https://pubmed.ncbi.nlm.nih.gov/21611157/)
40. Oda T, Ohniwa RL, Suzuki Y, Denawa M, Kumeta M, Okamura H, et al. Evolutionary dynamics of spliceosomal intron revealed by *in silico* analyses of the P-Type ATPase superfamily genes. *Mol Biol Rep*. 2011; 38: 2285–2293. doi: [10.1007/s11033-010-0360-3](https://doi.org/10.1007/s11033-010-0360-3) PMID: [21046256](https://pubmed.ncbi.nlm.nih.gov/21046256/)
41. Da Lage JL, Binder M, Hua-Van A, Janeček S, Casane D. Gene make-up: rapid and massive intron gains after horizontal transfer of a bacterial  $\alpha$ -amylase gene to *Basidiomycetes*. *BMC Evol Biol*. 2013; 13: 40. doi: [10.1186/1471-2148-13-40](https://doi.org/10.1186/1471-2148-13-40) PMID: [23405862](https://pubmed.ncbi.nlm.nih.gov/23405862/)