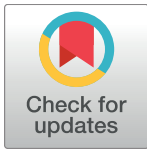


RESEARCH ARTICLE

The mitochondrial genome of the plant-pathogenic fungus *Stemphylium lycopersici* uncovers a dynamic structure due to repetitive and mobile elements

Mario Emilio Ernesto Franco¹, Silvina Marianela Yanil López¹, Rocio Medina¹, César Gustavo Lucentini¹, Maria Inés Troncozo², Graciela Noemí Pastorino², Mario Carlos Nazareno Saparrat^{2,3,4}, Pedro Alberto Balatti^{1,2*}



1 Centro de Investigaciones de Fitopatología, Facultad de Ciencias Agrarias y Forestales, Universidad Nacional de La Plata - Comisión de Investigaciones Científicas de la Provincia de Buenos Aires, La Plata, Buenos Aires, Argentina, **2** Cátedra de Microbiología Agrícola, Facultad de Ciencias Agrarias y Forestales, Universidad Nacional de La Plata, La Plata, Buenos Aires, Argentina, **3** Instituto de Botánica Carlos Spegazzini, Facultad de Ciencias Naturales y Museo, Universidad Nacional de La Plata, La Plata, Buenos Aires, Argentina, **4** Instituto de Fisiología Vegetal, Facultad de Ciencias Naturales y Museo-Facultad de Ciencias Agrarias y Forestales, Universidad Nacional de La Plata - Consejo Nacional de Investigaciones Científicas y Técnicas, La Plata, Buenos Aires, Argentina

* pbalatti@gmail.com

OPEN ACCESS

Citation: Franco MEE, López SMY, Medina R, Lucentini CG, Troncozo MI, Pastorino GN, et al. (2017) The mitochondrial genome of the plant-pathogenic fungus *Stemphylium lycopersici* uncovers a dynamic structure due to repetitive and mobile elements. PLoS ONE 12(10): e0185545. <https://doi.org/10.1371/journal.pone.0185545>

Editor: Marie-Joelle Virolle, Université Paris-Sud, FRANCE

Received: July 19, 2017

Accepted: September 14, 2017

Published: October 3, 2017

Copyright: © 2017 Franco 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: The mitochondrial genome of *Stemphylium lycopersici* CIDEFI-216 has been deposited at DDBJ/EMBL/GenBank under the accession no. KX453765.

Funding: This research was partially supported by the Agencia Nacional de Promoción Científica y Tecnológica (ANPCyT) of the Ministro de Ciencia, Tecnología e Innovación Productiva through the projects PICT 2012-2760 (PAB) and PICT 2015-1620 (MCNS). The funders had no role in study

Abstract

Stemphylium lycopersici (Pleosporales) is a plant-pathogenic fungus that has been associated with a broad range of plant-hosts worldwide. It is one of the causative agents of gray leaf spot disease in tomato and pepper. The aim of this work was to characterize the mitochondrial genome of *S. lycopersici* CIDEFI-216, to use it to trace taxonomic relationships with other fungal taxa and to get insights into the evolutionary history of this phytopathogen. The complete mitochondrial genome was assembled into a circular double-stranded DNA molecule of 75,911 bp that harbors a set of 37 protein-coding genes, 2 rRNA genes (*rns* and *rnI*) and 28 tRNA genes, which are transcribed from both sense and antisense strands. Remarkably, its gene repertoire lacks both *atp8* and *atp9*, contains a free-standing gene for the ribosomal protein S3 (*rps3*) and includes 13 genes with homing endonuclease domains that are mostly located within its 15 group I introns. Strikingly, subunits 1 and 2 of cytochrome oxidase are encoded by a single continuous open reading frame (ORF). A comparative mitogenomic analysis revealed the large extent of structural rearrangements among representatives of Pleosporales, showing the plasticity of their mitochondrial genomes. Finally, an exhaustive phylogenetic analysis of the subphylum Pezizomycotina based on mitochondrial data reconstructed their relationships in concordance with several studies based on nuclear data. This is the first report of a mitochondrial genome belonging to a representative of the family Pleosporaceae.

design, data collection and analysis, decision to publish, or preparation of the manuscript.

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

Introduction

Stemphylium (teleomorph: *Pleospora*) is a dematiaceous fungal genus belonging to the subphylum Pezizomycotina (Ascomycota). Species within this genus are widely distributed in nature, where they are found establishing pathogenic, saprotrophic or endophytic relationships with a broad range of plant hosts. Pathogenic representatives of this genus cause severe yield reductions and economic losses on several crops worldwide [1]. Among these phytopathogens, *Stemphylium lycopersici* has been associated with 36 plant hosts [2], it affects horticultural crops such as tomato [3] and pepper [4] as well as ornamental plants such as chrysanthemum [5] and kalanchoe [6]. Several studies mainly focused on morphological, physiological, pathogenic and genetic variation in different populations of *S. lycopersici* have been performed [7–10], but much research remains to be done concerning the molecular mechanisms underlying all these aspects. Regarding this, the whole-genome sequencing project of *S. lycopersici* CIDEFI-216 [11] might shed light on many of these issues and might also contribute to a better understanding of the evolutionary history of *Stemphylium*.

Mitochondria are semi-autonomous organelles present in most eukaryotic cells and are the main responsible of a variety of crucial cellular processes such as energy production, cell growth and apoptosis [12, 13]. They are thought to be descendants of an endosymbiotic α -proteobacterium that was engulfed by a proto-eukaryotic host cell more than one billion years ago in an event that gave rise to one of the most extreme and effective symbiotic relationship in nature. During this co-evolution, the genome of the ancient endosymbiont has undergone a remarkable size reduction since much of its genetic material have been transferred to the host or lost [14, 15].

Fungal mitochondrial DNAs are typically small, double-stranded DNA molecules with a highly compact gene organization that are present in multiple copies in each cell. The actual topology these molecules adopt *in vivo* is certainly under debate. While some authors assume a circular monomeric conformation, other authors suggest a linear concatemeric structure that, once it has been assembled, would also lead to a circular-mapping molecule. As a matter of fact, given the probability that the mitochondrial DNA of most fungal species replicates by the rolling-circle mechanism, it was also proposed that the actual genome architecture is a combination of both circular and linearly repeated. Fungal mitochondrial genomes usually harbor 14 core-genes encoding proteins involved in electron transport and oxidative phosphorylation, including the apocytochrome b (*cob*), three subunits of the cytochrome c oxidase (*cox1*, *cox2*, *cox3*), seven subunits of the reduced nicotinamide adenine dinucleotide ubiquinone oxidoreductase (*nad1*, *nad2*, *nad3*, *nad4*, *nad4L*, *nad5* and *nad6*) and three subunits of the ATP synthase (*atp6*, *atp8*, *atp9*), as well as untranslated genes of the small and large ribosomal RNA (rRNA) subunits (*rns* and *rnl*, respectively) and a set of transfer RNA (tRNA) genes. Other genes that are occasionally found within mitochondrial genomes of fungi are those that encode the ribosomal protein S3 (RPS3, coded by *rps3*) and the RNA subunit of the mitochondrial RNase P (*rnpB*). Last, but not least, fungal mitochondrial genomes also are characterized by a variable number of group I and group II introns that may carry homing endonuclease genes (HEGs) with LAGLIDADG or GIY-YIG motifs [14–18]. HEGs are selfish genetic mobile elements that encode site-specific-sequence-tolerant DNA endonucleases whose catalytic activity promotes their own propagation by the introduction DNA double-strand breaks (DSBs) into alleles lacking the endonuclease-coding sequence and by the subsequent repair of these DSBs via homologous recombination using as template the endonuclease-containing allele [19–21]. This mechanism of propagation can result in the insertion, deletion or mutation of DNA sequences [22]. Indeed, mobile introns and HEGs represent one of the major sources of variability within fungal mitochondrial genomes [16, 23, 24].

Mitochondrial genomes provide a valuable source of information to study evolutionary biology and systematics in eukaryotes. The reason for this is that they harbor conserved genes that code for proteins involved in the electron transport-oxidative phosphorylation system and that the mitochondrial DNA evolve faster than nuclear DNA, therefore, they allow the development of a robust phylogenetic analysis [25–27]. They have been successfully used in the solving of many unclear phylogenies [28, 29]. Besides, the relatively small size, circular-mapping topology and multi-copy nature of mitochondrial genomes facilitate their whole-sequencing and assembly and, thus, their study as an entity. This latter aspect is of major importance, because a complete mitochondrial genome sequence reveals the actual gene content of the organelle. Thus, it also provides additional information regarding the mitochondrial genome organization and enables the evaluation of structural rearrangements by means of comparative studies [14, 30, 31]. In addition, the high polymorphism frequently found within introns or intergenic regions of well-conserved mitochondrial genes makes these sequences useful for the study of genetic diversity among or within populations [32–34]. Mitochondrial genomes also are important in phytopathogens control, since they are the target of strobilurins, fungicides that prevent the synthesis of ATP by blocking the electron transfer at the quinol oxidation site by the cytochrome *bc₁* complex. Therefore, resistance of several fungal plant pathogens to strobilurins was associated with mutations in the mitochondrial gene *cob* [35, 36].

In this study, we assembled and annotated the mitochondrial genome of *S. lycopersici* CIDEFI-216, a fungus isolated from a tomato plant with symptoms of gray spot disease. We have thoroughly characterized it in terms of its gene content and organization, codon usage and by the occurrence of repetitive elements. We have also explored the evolutionary dynamics of the mitochondrial genomes of Dothideomycetes by a comparative mitogenomic approach. Finally, we carried out an extensive phylogenetic analysis aimed to infer the evolutionary relationships among 82 Pezizomycotina species, as revealed by the concatenated amino-acid sequences of 12 conserved mitochondrial-encoded proteins. This work describes mitochondrial genome of *S. lycopersici* CIDEFI-216, which, to our knowledge, is the first report for a member of the family Pleosporaceae (Pleosporales, Ascomycota).

Materials and methods

Sequence data and mitochondrial genome de novo assembly

The mitochondrial genome of *S. lycopersici* CIDEFI-216 was sequenced on an Illumina HiSeq 2000 platform (Illumina, San Diego, CA) using a 100 bp paired-end approach (insert size of 300 bp) in the course of the *S. lycopersici* CIDEFI-216 whole genome shotgun project (BioProject: PRJNA274742, BioSample: SAMN03332054) [11]. Even though the data were originated from total genomic DNA, the mitochondrial fraction of the sequence reads was expected to be over-represented as most eukaryotic cells contain multiple mitochondria, each with several copies of the genome. Hence, only a subset of 4,000,000 sequence reads (Sequence Read Archive: SRX872418) were used for the mitochondrial genome assembly. Low-quality bases at the ends of the sequence reads were trimmed off and the mitochondrial genome was assembled with the Geneious 9.1.2 *de novo* assembler [37] using low sensitivity, enabling the option to circularize contigs with matching ends. This procedure was repeated using different sets of 4,000,000 sequence reads in order to verify the recovery of the same mitochondrial DNA molecule.

Mitochondrial genome annotation

Gene annotation was performed with MFannot (<http://megasun.bch.umontreal.ca/cgi-bin/mfannot/mfannotInterface.pl>) using the NCBI translation Table 4 (The Mold, Protozoan, and

Coelenterate Mitochondrial Code and the Mycoplasma/Spiroplasma Code). MFannot predictions were individually checked to confirm genes boundaries as well as intron-exon boundaries by aligning the predicted genes with their orthologous in closely-related fungal species. Transfer-RNA annotations were further confirmed by tRNAscan-SE 1.21 [38]. Functional annotation of the predicted open reading frames (ORFs) was complemented with Blast2GO Basic [39]. A physical map of the mitochondrial genome was created with OrganellarGenomeDRAW (OGDRAW) v 1.2 [40] and modified with Inkscape 0.91 (<http://inkscape.org/>). The mitochondrial genome of *S. lycopersici* CIDEFI-216 has been deposited at DDBJ/EMBL/GenBank under the accession number KX453765.

Identification of repetitive sequences

Repetitive sequences were identified using the Vmatch software (<http://www.vmatch.de/>). Both direct and inverted repeats (palindromes) were computed by extending a seed of 30 nucleotides in length in both directions allowing for matches, mismatches, insertions and deletions with an X-drop value for edit distance extension of 3. The screening was restricted to repetitive sequences of at least 30 nucleotides in length and with a minimum match identity of 80%. Following the analysis, overlapped repeats were manually deleted. A physical map of the repetitive sequences in the mitochondrial genome of *S. lycopersici* CIDEFI-216 was created with Circos [41] and modified with Inkscape 0.91 (<http://inkscape.org/>).

Comparative mitogenomics

The extent of structural rearrangements in the mitochondrial genomes of *S. lycopersici* CIDEFI-216 and the Pleosporales species whose mitochondrial genomes were publicly available in the NCBI Genome database (<https://www.ncbi.nlm.nih.gov/genome/browse/?report=5>) by the end of December 2016, namely *Didymella pinodes* (GenBank accession number KT946597), *Parastagonospora nodorum* (GenBank accession number EU053989; [42]) and *Shiraia bambusicola* (GenBank accession number KM382246; [43]), was assessed by means of a multiple alignment using Mauve [44]. For comparative purposes, the first residue of the large ribosomal subunit was arbitrary set as the origin of all mitochondrial sequences. Mauve was run using the default settings and the backbone output was modified with Inkscape 0.91 (<http://inkscape.org/>).

Phylogenetic analysis

The phylogeny of the subphylum Pezizomycotina was inferred by the concatenated amino acid sequences of 12 proteins encoded by the conserved mitochondrial genes involved in the oxidative phosphorylation and electron transport, namely: *atp6*, *cob*, *cox1*, *cox2*, *cox3*, *nad1*, *nad2*, *nad3*, *nad4*, *nad4L*, *nad5* and *nad6*. This data set excluded the products of the genes *atp8* and *atp9* as they are not shared across the entire subphylum. The analysis comprised all Pezizomycotina species whose mitochondrial genome was publicly available in the NCBI Genome database (<https://www.ncbi.nlm.nih.gov/genome/browse/?report=5>) by the end of December 2016 and contained the 12 conserved mitochondrial genes mentioned above. Following this criterion, a representative of the subphylum Taphrinomycotina (*Pneumocystis carinii*) and two representatives of the subphylum Saccharomycotina (*Yarrowia lipolytica* and *Komagataella pastoris*) were selected as outgroups. Hence, the analysis involved a total of 85 species representing 54 genera. Amino acid sequences were concatenated and aligned with MUSCLE [45] in Geneious 9.1.2 [37]. The resultant alignment was automatically curated with Gblocks 0.91b [46] using default settings except for the minimum length of a block which was set to 2. The selection of the best-fit model of protein evolution was carried out with ProtTest 3.2 [47] using

the Akaike Information Criterion [48]. The phylogeny was reconstructed using maximum-likelihood (best-fit model: LG+I+G+F) in PhyML 3.0 [49] and Bayesian inference (best-fit available model: LG+I+G) in MrBayes 3.2 [50]. The support of the maximum-likelihood tree was estimated through 1000 bootstrap replicates [51]. For the Bayesian analysis, two independent Metropolis-coupled Markov chain Monte Carlo (MCMCMC) with four chains (one cold and three heat) were performed through one million of generations. Trees in each chain were sampled every 250th generation, the first 25% of trees were removed as burn-in and posterior probabilities were determined from the remaining trees. Phylogenetic trees were edited in Fig-Tree v1.4.2 (<http://tree.bio.ed.ac.uk/software/figtree/>) and Inkscape 0.91 (<http://inkscape.org/>).

Results and discussion

General features

The mitochondrial genome of *S. lycopersici* CIDEFI-216 was *de novo* assembled into a single, circular, double-stranded DNA molecule of 75,911 bp, with mean coverage of 117.74 (minimum coverage of 24; maximum coverage of 200). It contains a set of 37 protein-coding genes, 2 rRNA genes (*rns* and *rnl*) and 28 tRNA genes. Genes are transcribed from both DNA strands, whose 62.9% correspond to non-coding DNA. With a GC content of 29.6%, it reaches the upper limit of the normal values for fungal mitochondria ($24.4 \pm 7.3\%$, estimated from all fungal mitochondrial genomes deposited in the Genome database of the NCBI by January 2016). The GC content of the protein-coding genes is quite similar to the overall GC content (29.1%), however considerable deviations are found in the GC content of rRNAs and tRNA genes, reaching 35% and 39%, respectively (Fig 1, Tables 1 and 2).

Protein-coding genes

The structural and functional annotation of the mitochondrial genome of *S. lycopersici* CIDEFI-216 led to the identification of 37 protein-coding genes. Twelve of them were related to the electron transport and oxidative phosphorylation system, and included *cob*, *cox1*, *cox2*, *cox3*, *nad1*, *nad2*, *nad3*, *nad4*, *nad4L*, *nad5*, *nad6* and *atp6*, which are considered core protein-coding genes. We also found a gene homologous to *rps3*, 13 genes homologues to GIY-YIG or LAGLIDADG homing endonucleases, a gene with homology to an N-acetyltransferase and 10 protein-coding genes with unknown function (Table 3 and S1 Table). All genes are transcribed in sense and anti-sense directions. Protein-coding genes account for about 57.81% of the mitochondrial genome, whilst its coding sequences represent the 44.69%. The observed difference is attributed to the presence of 15 group I introns, many of them harboring ORFs that mostly encode GIY-YIG or LAGLIDADG homing endonucleases. The core protein-coding genes exhibited some notable features. Undoubtedly, the most striking feature was the fusion between *cox1* and *cox2*, as suggested by its gene architecture. This putative *cox1/2* gene fusion was riddled with introns, accounting for nearly half of the introns found in the mitochondrial genome. Another remarkable feature was the absence of genes encoding the subunits 8 and 9 of the ATP synthase (*atp8* and *atp9*, respectively), since both genes are conserved among fungal mitochondria [16, 31, 52]. The aforementioned characteristics are further discussed below.

ATP synthase subunits 8 and 9 are encoded by the nuclear genome

In billion of years of co-evolution with its host-cell and specialization into an energy-producing organelle, the ancestor of mitochondria lost most of its genetic material. This might have led to the elimination of genes whose function either become obsolete or were relieved by the corresponding nuclear-gene counterpart. Indeed, it has been shown that most of its genetic material

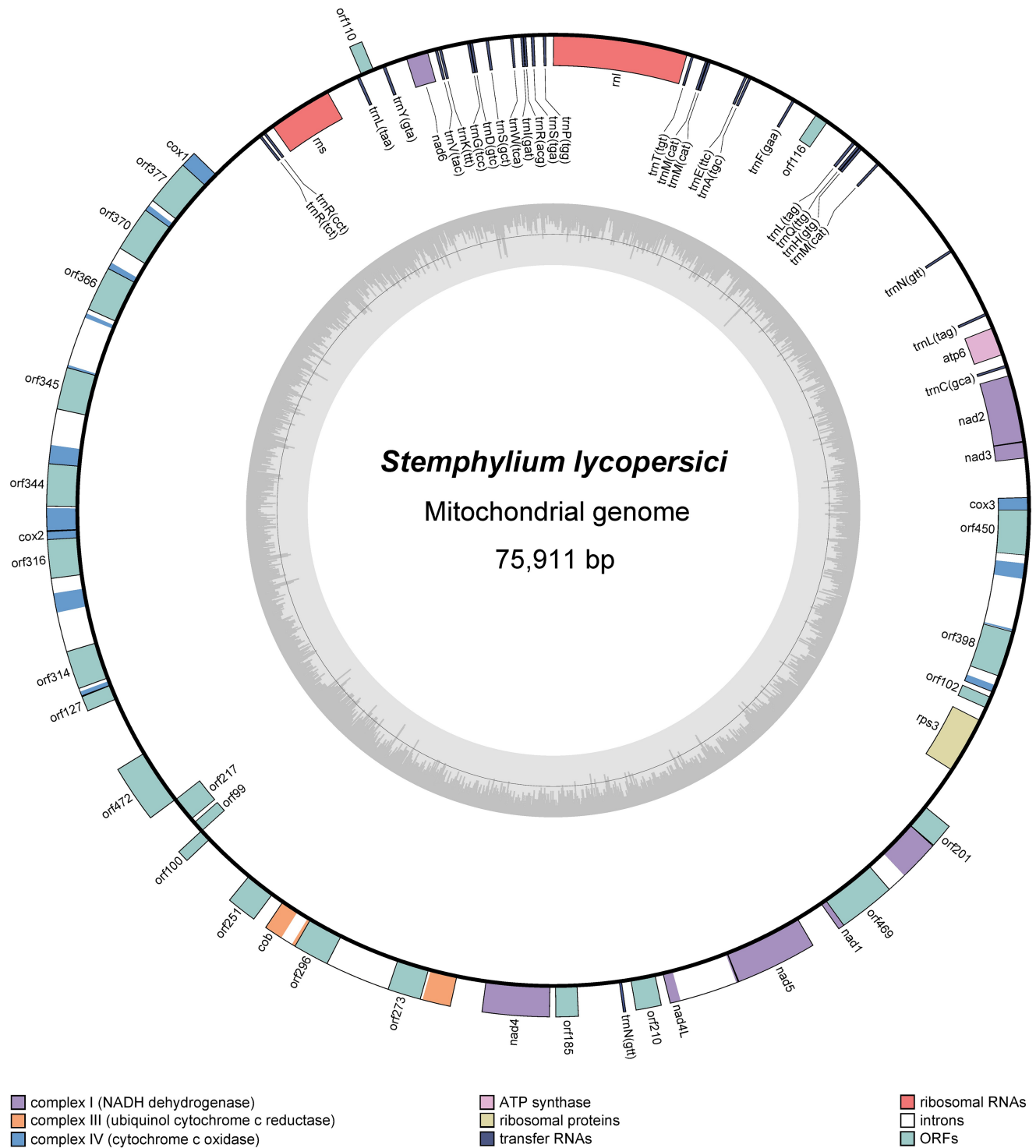


Fig 1. Circular map of the mitochondrial genome of *S. lycopersici* CIDEFI-216. Genes are represented by boxes, which are colored on the basis of their function. The inner boxes represent genes that are transcribed in a clockwise direction and the outer boxes show genes transcribed in a counterclockwise direction. The internal plot shows the GC content, with the black line marking 50%. The exact positions of the genes are indicated in Table 2. The physical map was created with OGDRAW.

<https://doi.org/10.1371/journal.pone.0185545.g001>

Table 1. General features of the mitochondrial genome of *S. lycopersici* CIDEFI-216.

Feature	Value
Genome size (bp)	75,911
GC content (%)	29.6
No. of protein-coding genes	37
GC content of protein-coding genes (%)	29.1
No. of rRNAs	2
GC content of rRNAs genes (%)	35.0
No. of tRNAs	28
GC content of tRNAs genes (%)	39.6
No. of introns	15
Intragenic regions (%)	30.2
Intergenic regions (%)	32.7
Repetitive regions (%)	5.6

<https://doi.org/10.1371/journal.pone.0185545.t001>

has been migrating to the nucleus as part of a process that is still ongoing. This is such that almost the entire mitochondrial proteome is actually encoded by nuclear-genes, synthesized in the cytosol and imported into the mitochondria by various different mechanisms, among which is the recognition of a mitochondrial-targeting peptide attached to the N-terminus of the corresponding pre-protein [15, 53].

The fungal mitochondrial genomes sequenced so far generally has a standard set of 14 protein-coding-genes that includes *atp6*, *atp8*, *atp9*, *cob*, *cox1*, *cox2*, *cox3*, *nad1*, *nad2*, *nad3*, *nad4*, *nad4L*, *nad5* and *nad6* [14, 31, 52]. There are, however, exceptions, as is attested by the absence of the *atp9* gene in *Podospora anserina* (Sordariales) [54]; of *atp8* and *atp9* genes in the Pleosporales *D. pinodes* (KT946597), *P. nodorum* [42] and *S. bambusicola* [43]; of *cox3* and *nad6* genes in *Acremonium implicatum* (Hypocreales) [55]; and the lack of *nad1*, *nad2*, *nad3*, *nad4*, *nad4L*, *nad5* and *nad6* genes in the mitochondrial genome of several yeasts [56–60]. We found that the mitochondrial genome of *S. lycopersici* CIDEFI-216 lacks both *atp8* and *atp9* genes, which is in agreement with observations made in other representatives of Pleosporales. Interestingly, these two genes are present in the mitochondrial genome of the Capnodiales *Zasmidium cellare* [61] and *Zymoseptoria tritici* [62], which suggests that these genes might have been lost at some point after the divergence between Dothideomycetidae and Pleosporomycetidae (Dothideomycetes). We searched for the presence of these genes in the nuclear of genome *S. lycopersici* CIDEFI-216 by doing BLAST [63] searches using the *atp8* and *atp9* nucleotide and amino-acid sequences of the closest related *Z. cellare* and *Z. tritici* as queries in a database composed by the 214 scaffolds of the draft genome of *S. lycopersici* CIDEFI-216 and its predicted proteome (GenBank accession number LGLR00000000.1). We found a gene highly homologous to *atp8* (Locus tag TW65_04673) in the scaffold 66 (GenBank accession number LGLR01000223) and another one highly homologous to *atp9* (Locus tag TW65_03276) in the scaffold 40 (GenBank accession number LGLR01000197). Both nuclear-genes were larger than their mitochondrial homologues, leaving a flanking sequence at their 3'-ends (N-terminus of each protein). The refinement of the alignment (ClustalW, default settings) to the homologous regions yielded a pairwise identity of 66.2% and 70.7% at the nucleotide level and 64.7% and 70.2% at the amino-acid level for *atp8* and *atp9*, respectively (S1 Fig). Furthermore, TargetP 1.1 [64] predicted the presence of N-terminal mitochondrial-targeting peptides in the putative products of *atp8* and *atp9* nuclear homologues (S2 Table).

Table 2. Gene organization of the mitochondrial genome of *S. lycopersici* CIDEFI-216.

Gene	Start position	Stop position	Length (nt)	Orientation
<i>rnl</i>	1	3,459	3,459	forward
<i>trnT(tgt)</i>	3,559	3,629	71	forward
<i>trnM(cat)</i>	3,937	4,007	71	forward
<i>trnM(cat)</i>	4,013	4,085	73	forward
<i>trnE(ttc)</i>	5,014	5,086	73	forward
<i>trnA(tgc)</i>	5,119	5,190	72	forward
<i>trnF(gaa)</i>	6,366	6,438	73	forward
<i>orf116</i>	7,077	7,427	351	forward
<i>trnL(tag)</i>	8,240	8,322	83	forward
<i>trnQ(ttg)</i>	8,393	8,464	72	forward
<i>trnH(gtg)</i>	8,470	8,543	74	forward
<i>trnM(cat)</i>	9,059	9,130	72	forward
<i>trnN(gtt)</i>	11,959	12,029	71	forward
<i>trnL(tag)</i>	13,797	13,869	73	forward
<i>atp6</i>	14,170	14,943	774	forward
<i>trnC(gca)</i>	15,216	15,287	72	forward
<i>nad2</i>	15,490	17,220	1,731	forward
<i>nad3</i>	17,221	17,643	423	forward
<i>cox3</i>	18,638	23,719	5,082	forward
<i>orf450</i>	18,971	20,323	1,353	forward
<i>orf398</i>	22,090	23,286	1,197	forward
<i>orf102</i>	23,847	24,155	309	forward
<i>orf490</i>	24,489	25,961	1,473	forward
<i>orf201</i>	27,104	27,709	606	reverse
<i>nad1</i>	27,709	30,731	3,023	reverse
<i>orf469</i>	29,178	30,587	1,410	reverse
<i>nad5</i>	31,440	33,431	1,992	reverse
<i>nad4L</i>	33,431	35,147	1,717	reverse
<i>orf210</i>	35,347	35,979	633	reverse
<i>trnN(gtt)</i>	36,210	36,280	71	reverse
<i>orf185</i>	37,345	37,902	558	reverse
<i>nad4</i>	38,034	39,671	1,638	reverse
<i>cob</i>	40,418	45,259	4,842	reverse
<i>orf273</i>	41,156	41,977	822	reverse
<i>orf296</i>	43,538	44,428	891	reverse
<i>orf251</i>	45,585	46,340	756	reverse
<i>orf100</i>	47,708	48,010	303	reverse
<i>orf99</i>	47,994	48,293	300	forward
<i>orf217</i>	48,394	49,047	654	forward
<i>orf472</i>	49,044	50,462	1,419	reverse
<i>orf127</i>	52,012	52,395	384	reverse
<i>cox2^a</i>	52,395	56,445	4,051	reverse
<i>orf314</i>	52,587	53,531	945	reverse
<i>orf316</i>	55,300	56,250	951	reverse
<i>cox1^a</i>	56,446	66,406	9,961	reverse
<i>orf344</i>	57,032	58,066	1,035	reverse
<i>orf345</i>	59,369	60,406	1,038	reverse

(Continued)

Table 2. (Continued)

Gene	Start position	Stop position	Length (nt)	Orientation
<i>orf366</i>	61,869	62,969	1,101	reverse
<i>orf370</i>	63,510	64,622	1,113	reverse
<i>orf377</i>	64,886	66,019	1,134	reverse
<i>trnR(tct)</i>	67,861	67,931	71	forward
<i>trnR(cct)</i>	67,995	68,065	71	forward
<i>rns</i>	68,284	69,953	1,670	forward
<i>orf110</i>	70,908	71,240	333	reverse
<i>trnY(gta)</i>	71,478	71,562	85	forward
<i>nad6</i>	72,116	72,691	576	forward
<i>trnV(tac)</i>	72,876	72,948	73	forward
<i>trnK(ttt)</i>	72,981	73,052	72	forward
<i>trnG(tcc)</i>	73,711	73,783	73	forward
<i>trnD(gtc)</i>	73,786	73,857	72	forward
<i>trnS(gct)</i>	74,188	74,267	80	forward
<i>trnW(tca)</i>	74,820	74,891	72	forward
<i>trnI(gat)</i>	75,077	75,148	72	forward
<i>trnR(acg)</i>	75,153	75,224	72	forward
<i>trnS(tga)</i>	75,345	75,429	85	forward
<i>trnP(tgg)</i>	75,655	75,727	73	forward

^a Both *cox1* and *cox2* are encoded by a single continuous open reading frame.

<https://doi.org/10.1371/journal.pone.0185545.t002>

A putative gene fusion between *cox1* and *cox2*

Cytochrome-c oxidase (EC 1.9.3.1) is a large oligomeric enzymatic complex that mediates the final step of the respiratory chain. In eukaryotes, this complex is composed by three large subunits, the cytochrome-c oxidase subunits I, II and III (COX1, COX2 and COX3, respectively), which comprise the catalytic center of the enzymatic complex together with a variable number of nuclear-encoded polypeptides [65]. Generally, COX1, COX2 and COX3 are encoded by three independent genes (*cox1*, *cox2* and *cox3*, respectively) that are located in the mitochondrial genome, with few exceptions in *cox2* and *cox3* [14, 66]. Given the important biological process that they fulfill, they are highly conserved even among distantly-related species, because of this, *cox1* has been used as a molecular barcode in several taxonomic and evolutionary studies [67, 68].

The predicted gene structure of the mitochondrial genome of *S. lycopersici* CIDEFI-216 presented an unusual gene organization of two subunits of the cytochrome-c oxidase. Specifically, COX1 and COX2 are encoded by a single continuous ORF that is 2,370 bp long and begin and terminate with the canonical start and stop codons ATG and TAA, respectively. This so-called *cox1/2* ORF contains the coding sequences of both COX1 and COX2 in the same frame lacking a canonical stop codon between them, which raises the possibility of a fusion transcript and/or protein. In this hypothetical transcript, COX1 sequence is extended to nucleotide 1,653, where the COX2 sequence begins.

This peculiar gene organization also was found in other representatives of the Pleosporales such as *D. pinodes* (KT946597) and *P. nodorum* [42]. In *S. bambusicola*, however, the aforementioned subunits are encoded by a pair of genes that are adjacent with no intergenic or overlapping nucleotides [43]. Similarly, a *cox1/2* ORF was found in the protozoans *Acanthamoeba castellanii* [69, 70], *Dictyostelium discoideum* [71] and *Polysphondylium pallidum*

Table 3. Characteristics of the protein-coding genes of the mitochondrial genome of *S. lycopersici* CIDEFI-216.

Gene	Introns	Coding sequence density (%)	Start Codon	Stop Codon	Product	Length (aa)
<i>atp6</i>	0	100	ATG	TAG	ATP synthase F0 subunit a	257
<i>cob</i>	2	23.9	ATG	TAG	Apocytochrome b	385
<i>cox1</i> ¹	6	16.6	ATG	TGT ²	Cytochrome c oxidase subunit 1	551
<i>cox2</i> ¹	2	17.7	GAT ³	TAA	Cytochrome c oxidase subunit 2	238
<i>cox3</i>	3	15.9	ATG	TAG	Cytochrome c oxidase subunit 3	269
<i>nad1</i>	1	36.9	ATG	TAA	NADH dehydrogenase subunit 1	371
<i>nad2</i>	0	100	ATG	TAA	NADH dehydrogenase subunit 2	576
<i>nad3</i>	0	100	ATG	TAG	NADH dehydrogenase subunit 3	140
<i>nad4</i>	0	100	ATG	TAA	NADH dehydrogenase subunit 4	545
<i>nad4L</i>	1	100	ATG	TAA	NADH dehydrogenase subunit 4L	89
<i>nad5</i>	0	100	ATG	TAA	NADH dehydrogenase subunit 5	663
<i>nad6</i>	0	100	ATG	TAA	NADH dehydrogenase subunit 6	191
<i>rps3</i>	0	100	ATG	TAA	Ribosomal protein S3	490
<i>orf99</i>	0	100	ATG	TAA	Hypothetical protein	99
<i>orf100</i>	0	100	ATG	TAG	Hypothetical protein	100
<i>orf102</i>	0	100	ATG	TAA	LAGLIDADG domain-containing protein	102
<i>orf110</i>	0	100	ATG	TAA	Hypothetical protein	110
<i>orf116</i>	0	100	ATG	TAA	Hypothetical protein	116
<i>orf127</i>	0	100	ATG	TAA	Hypothetical protein	127
<i>orf185</i>	0	100	ATG	TAA	Hypothetical protein	185
<i>orf201</i>	0	100	ATG	TAG	GIY-YIG domain-containing protein	201
<i>orf210</i>	0	100	ATG	TAA	Hypothetical protein	210
<i>orf217</i>	0	100	ATG	TAA	N-acetyltransferase domain-containing protein	217
<i>orf251</i>	0	100	ATG	TAA	Hypothetical protein	251
<i>orf273</i> ⁴	0	100	ATG	TAA	LAGLIDADG domain-containing protein	273
<i>orf296</i> ⁴	0	100	AAA ³	TAG	LAGLIDADG domain-containing protein	296
<i>orf314</i> ⁴	0	100	ATG	TAA	GIY-YIG domain-containing protein	314
<i>orf316</i> ⁴	0	100	AAA ³	TAA	GIY-YIG domain-containing protein	316
<i>orf344</i> ⁴	0	100	CAA ³	TAA	LAGLIDADG domain-containing protein	344
<i>orf345</i> ⁴	0	100	AAA ³	TAG	Hypothetical protein	345
<i>orf366</i> ⁴	0	100	AAA ³	TAA	LAGLIDADG domain-containing protein	366
<i>orf370</i> ⁴	0	100	GGA ³	TAA	GIY-YIG domain-containing protein	370
<i>orf377</i> ⁴	0	100	GAA ³	TAG	LAGLIDADG domain-containing protein	377
<i>orf398</i> ⁴	0	100	ATG	TAG	LAGLIDADG domain-containing protein	398
<i>orf450</i> ⁴	0	100	AAA ³	TAA	LAGLIDADG domain-containing protein	450
<i>orf469</i> ⁴	0	100	TGT ³	TAA	GIY-YIG domain-containing protein	469
<i>orf472</i>	0	100	ATG	TAA	Hypothetical protein	472

The conserved mitochondrial genes involved in the electron transport and oxidative phosphorylation are in bold type letter.

¹ Both *cox1* and *cox2* seems to be encoded by a single continuous open reading frame.

² Stop codon not determined. Possibly, the ORF is co-translated with the downstream exon(s) and then further processed to yield the corresponding mature proteins.

³ Start codon not determined. Possibly, the ORF is co-translated with the upstream exon(s) and then further processed to yield the corresponding mature proteins.

⁴ Gene located within an intron.

<https://doi.org/10.1371/journal.pone.0185545.t003>



Fig 2. Schematic sequence alignment of COX1 and COX2 with two different gene configurations. Amino-acid identities are highlighted with colors. COX1, COX2 and COX1/2 regions are indicated by green, yellow and blue boxes. Intron positions are represented by a discontinuity in the boxes. Any recognizable homing endonuclease domain, LAGLIDADG domain and/or GIY-YIG domain found within the hypothetical proteins encoded by the intronic ORFs are depicted by the letters H, L and G, respectively. The amino-acid sequences were aligned using MUSCLE with the default settings in Geneious 9.1.2. The detailed alignment is found in S2 Fig.

<https://doi.org/10.1371/journal.pone.0185545.g002>

(AY700145), and the cercozoan *Paracercomonas marina* (KP165385), which, in spite of the large evolutionary distance to fungi from Pleosporales, are placed near the base of radiation that leads to the animals and fungi [25, 72]. Indeed, conserved homologous blocks belonging to both subunits can be clearly identified within the multiple sequence alignment of the putative COX1/2 protein with homologous sequences from Pleosporales and the above-mentioned protists (Fig 2 and S2 Fig).

Such gene architecture was first reported by Burger et al. in 1995, describing the mitochondrial genome of *A. castellanii* [69]. A year later, Lonergan and Gray proved that, in *A. castellanii*, this ORF is transcribed into a large polycistronic mRNA that upon translation give rise to two products [70]. Next, Pellizzari et al found in the mitochondrial genome of *D. discoideum* evidences of a similar *cox1/2* transcript [71]. However, whether COX1 and COX2 are synthesized separately or as a fused protein in *S. lycopersici* remains unclear. These subunits might be encoded by the nuclear genome, whereas the mitochondrial *cox1/2* represents a dysfunctional copy. However, we were unable to find the coding sequences of both subunits in the nuclear genome of *S. lycopersici* CIDEFI-216 by doing BLAST searches against its predicted ORFeome and proteome. So, it appears that the mitochondrial *cox1/2* locus is unique and, thus, has to be functional. The *cox1/2* transcript may also contain an unusual stop codon immediately after the *cox1* CDS, but this is rather unlikely in view of the codon usage found in the conserved protein-coding genes of this mitochondrial genome, which will be discussed in a later section. Finally, we could not rule out that the *cox1/2* primary-transcript might be edited so as to include a termination codon in the mature mRNA, or be endonucleolytically processed to yield the two mature mRNAs. Future studies should be aimed at answering this question.

The *cox1/2* gene sequence has a length of 14,012 nucleotides, contains 8 group I introns, 5 of which harbor ORFs that might encode homing endonucleases, as it could be expected by their predicted LAGLIDADG or GIY-YIG domains. Interestingly, the location of some of these introns is well-conserved, even among distantly-related species (Fig 2 and S2 Fig). So, it appears that *cox1* is a hotspot for insertion of group I introns [73, 74]. Whether group I introns encode homing endonucleases functionally active or are just footprints of previous insertional events, they provide valuable information to trace evolutionary history. Interestingly, there is substantial evidence that group I intron-encoded homing endonucleases are involved in processes such as genome rearrangements and horizontal gene transfer [75–77], suggesting these two possible scenarios that could explain the observed gene fusions between *cox1* and *cox2* in quite distantly-related taxa. The gene fusion might have arisen independently in different lineages due to the exon shuffling promoted by the homing endonuclease-mediated mobility of group I

introns. Alternatively, the fusion might be the result of a horizontal gene transfer event that might have occurred at some point during the evolution of Dothideomycetes and Protozoans.

A free-standing *rps3* and an intronless *rnl*

The mitochondrial genome of *S. lycopersici* CIDEFI-216 contains a free-standing gene that is 1,473 bp long that code for a 490 amino-acids protein with partial homology to several fungal RPS3, a key protein component of the ribosome, essential for protein translation. *Rps3* is widely distributed throughout all life kingdoms and is certainly diverse in sequence, structure and location [78, 79]. The first mitochondrial ribosomal protein described in fungi, specifically in yeast, was VAR1 that is coded by the free-standing gene *var1* [80]. Later, Burke and RajBhandary found an intronic ORF for the putative ribosomal protein S3 within the *rnl* U11 group I intron (also known as mL2449 group I intron) in *Neurospora crassa* (Sordariales) [81]. Currently, the general consensus is that RPS3, VAR1 and S5 are homologs [78, 79]. Although the longest protein orthologs usually contain partial sequences with no significant homology to any known protein, sometimes they contain partial sequences homologous to known proteins such as in some *Ophiostoma* spp. (Ophiostomatales), where *rps3* is fused to homing endonuclease domains [82]. Furthermore, while in *P. nodorum*, *rps3* also contains an internal sequence highly homologous to a partial sequence of *cox1* [42, 79], in *Sphaeronomella fimicola* (Microascales), *rps3* includes an in-frame micro-satellite insertion that results in the presence of 16 consecutive alanines [79]. In light of the findings described above, it is evident that *rps3* has a rather complex evolutionary history, which makes more interesting their sequence analysis among representatives of Ascomycota. Therefore, we used all the available mitochondrial genomes in the GenBank (Accessed in April 2017) and, when necessary, annotated them as described previously. We found that, in Saccharomycotina, the ribosomal protein S3 is coded by a free-standing gene and, on the other hand, the *rnl* usually contains an intron that in some cases harbor an ORF homologous to homing endonucleases. In Taphrinomycotina, *rps3* exists as a free-standing gene, through no intron was found in *rnl*. Within Pezizomycotina, *rps3* was found within the mL2449 group I intron of the *rnl* in members of Eurotiomycetes and Sordariomycetes. However, in Leotiomycetes and Lecanoromycetes, *rps3* exists either as a free-standing gene or within the mL2449 group I intron of *rnl*. Surprisingly, Dothideomycetes contains an intronless *rnl* and, while the mitochondrial genomes of *Z. cellare* and *Z. tritici* (Capnodiales) lack the *rps3*, the ones of *D. pinodes*, *S. bambusicola*, *S. lycopersici* and *P. nodorum* (Pleosporales) contain a free-standing *rps3*, that only in the latter is fused with a partial sequence of *cox1* (Fig 3 and S3 Fig). Probably, these gain-loss events that led to the configurations described might have been driven by the presence of an essential gene such as *rps3* within the mL2449 group I intron that ensured the maintenance of this neutral element, which otherwise might have eventually been lost, but also by the homing endonuclease-mediated mobility that might be responsible for the relocation of *rps3* from its original position (whichever it was) to the mL2449 group I intron and/or its settlement as a free-standing gene. This is further supported

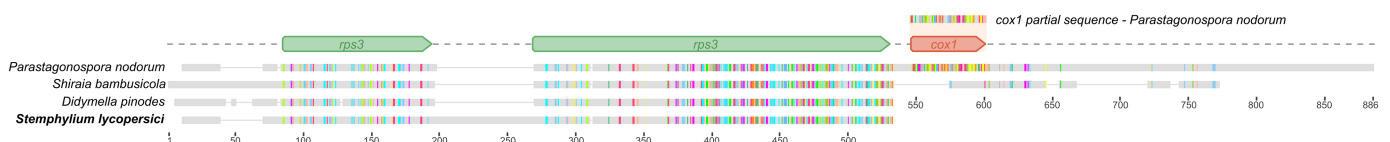


Fig 3. Schematic sequence alignment of RPS3 from Pleosporales species. It can be clearly seen conserved blocks of homology to RPS3 shared between the analyzed Pleosporales species. Regions of RPS3 that are conserved among Pleosporales are indicated with green boxes. The region homologous to COX1 particularly found in *P. nodorum* is indicated with a red box. The amino-acid sequences were aligned using MUSCLE with the default settings in Geneious 9.1.2. The detailed alignment is found in S3 Fig.

<https://doi.org/10.1371/journal.pone.0185545.g003>

by the presence of LAGLIDADG homing endonuclease-*rps3* fusions within the *rnl* Group I Intron in some *Ophiostoma* spp. [82] as well as by the fact that the *rps3* of *P. nodorum* presents a footprint of a previous insertion of the *cox1* [32, 79], a gene that is usually invaded by these mobile genetic elements [75–77]. It is worth mentioning that we did not find a *rps3* counterpart in the nuclear genome of any representative of Dothideomycetes.

Fertile ground for introns and homing endonucleases

Group I and group II introns are ribozymes that catalyze their splicing from a precursor RNA, restoring the translational reading frame and generating, in this way, a functional product. This selfish DNA is frequently present in mitochondrial genomes, though while group I introns are abundant in fungal mitochondrial genomes, group II introns are predominant in those of plants [17]. Both types of introns propagate themselves as mobile genetic elements, through different mechanisms. Mobility of group I intron is due to the activity of homing of endonucleases, whose coding sequences reside in non-critical sequences of the intron. On the other hand, mobility of group II introns relies on a mechanism known as retro-homing, which is catalyzed by an intron-encoded protein with reverse-transcriptase and maturase domains, and sometimes an endonuclease domain [83]. In fact, homing endonucleases themselves can be mobile elements, moving independently from their host intron, like the free-standing homing endonucleases, which are frequent within phage genomes [21]. It is intriguing how such genes settled within unspoiled introns to yield these composite mobile elements. Whatever the reason, this co-evolution left several footprints along genomes over eons and, hence, these genetic fossils are a valuable source of information for evolutionary studies.

In the mitochondrial genome of *S. lycopersici* CIDEFI-216 we found 15 introns, all corresponding to group I, many of them harboring ORFs that mostly encode GIY-YIG or LAGLIDADG domain-containing proteins. All introns were located within conserved genes involved in the electron transport and the oxidative phosphorylation system, which, ordered by frequency, are: *cox1/2* (8), *cox3* (3), *cob* (2), *nad1* (1) and *nad4L* (1) (Table 3). As it was mentioned before, *cox1* is a common target for the insertion of introns in fungi and other eukaryotes [75–77]. The mitochondrial genome of *S. lycopersici* CIDEFI-216 also contains 13 genes, whose hypothetical proteins present domains characteristic of homing endonucleases (S1 Table). While 8 of the hypothetical proteins carried the LAGLIDADG domain, 5 of them carried the GIY-YIG domain. Most of the homing endonuclease genes in the mitochondrial genome of *S. lycopersici* CIDEFI-216 are placed within group I introns, however two of them exist as free-standing genes (*orf102* for the LAGLIDADG domain-containing protein and *orf201* for the GIY-YIG domain-containing protein). Interestingly, they account for more than one-third of the total protein-coding genes of the mitochondrial genome.

It is important to compare intron positions within the *cox1/2* (*cox1* and *cox2*) sequence from the mentioned Dothideomycetes, slim molds and cercozoan representatives (Fig 2). Among the known sequences of *cox1/2* in Dothideomycetes, *S. lycopersici* has the highest number of introns (8 introns harboring 5 homing endonucleases), followed by *D. pinodes* (7 introns but 6 homing endonucleases) (GenBank accession number KT946597). Furthermore, while *S. bambusicola* has a *cox1* that contains a single intron (1 homing endonuclease), *cox2* lacks introns [43]. On the other hand, *cox1/2* in *P. nodorum* and *cox1* and *cox2* in *Z. cellare* and *Z. tritici* (Capnodiales) lack introns [42, 61, 62]. In the distantly-related slime mold *D. discoideum*, *cox1/2* contains 4 introns with 4 homing endonucleases (two in the same intron) [71], while the protozoans *A. castellanii* [69] and *P. pallidum* [70], and the cercozoan *P. marina* (GenBank accession number KP165385) lack introns. The presence of an intron in the same position in a distantly-related organism suggests that such structural feature could be the result

of an ancestral event of insertion. Considering this, the first intron of *D. pinodes* and the first intron of *D. discoideum* might have a common ancestral origin, though the latter lost their intron-mobility enzyme. However, all the other studied organisms lost this intron at some point of their evolutionary history. Hence, we hypothesized that the remaining introns might have arisen from more recent insertional events. Interestingly, the position of the third intron of *D. pinodes* and the single intron within *cox1* of *S. bambusicola* correlates with a slight shift, regarding the position of the second intron of *D. discoideum*. Similarly, the sixth intron of *S. lycopersici* and the fourth intron of *D. pinodes* with the third intron of *D. discoideum* are also correlated. We might have detected cases of lost-and-gain events, not only because of the aforementioned slight variation in introns position between these distantly-related organisms, but also because of the low level of amino-acid identity shared by their mobility proteins. We might speculate that the earliest insertional event gave rise to the seventh intron in *S. lycopersici* and the sixth intron in *D. pinodes*, since both of them contains LAGLIDADG domain-containing homing endonucleases with a high level of pairwise identity at the amino-acid level (Pairwise% Identity: 87.9%; Pairwise% Positive (BLSM62): 94.8%).

tRNAs genes and codon bias

The mitochondrial genome of *S. lycopersici* CIDEFI-216 has a set of 28 tRNA genes, including all those needed to decode the codons of each of the 20 standard amino-acids required for the synthesis of the predicted mitochondrial-encoded proteins (Table 4). As a result of this, mitochondria might not need to import any nuclear-encoded tRNA from the cytosol to synthesize

Table 4. tRNAs within the mitochondrial genome of *S. lycopersici* CIDEFI-216.

Amino Acid	Anticodon	Number
Ala	UGC	1
Arg	ACG	1
Arg	CCU	1
Arg	UCU	1
Asn	GUU	2
Asp	GUC	1
Cys	GCA	1
Gln	UUG	1
Glu	UUC	1
Gly	UCC	1
His	GUG	1
Ile	GAU	1
Leu	UAA	1
Leu	UAG	2
Lys	UUU	1
Met	CAU	3
Phe	GAA	1
Pro	UGG	1
Ser	GCU	1
Ser	UGA	1
Thr	UGU	1
Trp	UCA	1
Tyr	GUA	1
Val	UAC	1

<https://doi.org/10.1371/journal.pone.0185545.t004>

their proteins, as it occurs in many fungal, plant and animal mitochondria [14–16]. The tRNA gene repertoire of *S. lycopersici* CIDEFI-216 includes 3 copies of *trnM(cat)*, 2 of *trnN(gtt)* and 2 of *trnL(tag)*. It also has tRNA genes with different anticodons that incorporate the same amino-acid, namely: *trnR(acg)*, *trnR(cct)* and *trnR(tct)* for Arginine; *trnE(ttc)* and *trnE(ttc)* for Glutamine; and *trnS(tga)* and *trnS(gct)* for Serine. Most tRNA genes are transcribed in a clockwise direction from a region spanning upstream of *rns* and downstream of *atp6*. The sole exception is one of the two copies of *trnN(gtt)*, which is located considerably farther downstream from the tRNA-rich region and is transcribed in counterclockwise direction (Fig 1).

While we were trying to identify the translational start and stop codons of the mitochondrial genes of *S. lycopersici* CIDEFI-216, a number of rather unusual translational signals drew our attention. Specifically, the identification of a putative termination codon TGT of COX1, which was followed by a hypothetical initiation codon TGT of COX2, in addition to the presumptive initiation codons AAA, CAA, GGA, GAA and TGT used for many homing endonucleases (Table 3). The existence of such signals in the mitochondrial genome of *S. lycopersici* CIDEFI-216 is very unlikely considering the genetic code that is used in closely-related species [42, 43, 61, 62], but also because the only difference within conserved genes occurred with COX1 and COX2, which has already been discussed in detail. A possible explanation for this observation might be the co-translation of these ORFs carrying suspicious initiation codons alongside with their upstream exon(s), which later should be processed in order to obtain the corresponding mature protein. Thereby, these triplets may not function as initiation or termination codons, instead they may be translated into amino-acids. Even if these observations led us to think that all protein-coding genes start with the canonical initiation codon ATG and terminate with the ochre TAA or the amber TAG stop codons (Table 3), future functional studies should be conducted to validate these results. We also provide the detailed codon usage information in Table 5.

Repetitive elements accounted for about 5.6% of the mitochondrial genome

One of the most notable features of the mitochondrial genome of *S. lycopersici* CIDEFI-216 is the abundance of both direct and palindromic repetitive sequences. In this regard, we found 84 repetitive elements larger than 30 nucleotides, which accounted for about 5.6% of the mitochondrial genome and exhibited an overall GC content of 25.1%. These repetitive sequences were unevenly distributed mostly throughout intergenic regions of the mitochondrial genome and showed a notable preference for some regions, among the most prominent ones were those located upstream and downstream the *rns* and *rnl*, between *orf185* and *orf210* and downstream *rps3* genes (Fig 4). We also observed the occurrence of a large duplicated region that contains a partial sequence of *nad2* (Fig 4, green ribbon). This duplication is 1332 nucleotides long and their first 558 nucleotides have a 99.6% pairwise identity with the first third of *nad2*, however there is a nonsense mutation in the 17th codon which results in the conversion of a Serine into a stop codon (S17Stop). Within this partial sequence, there is also one more single nucleotide polymorphism that may result in a silent mutation (I219I). The sequence downstream nucleotide 558 had no homology to any known protein (Fig 5). Probably, selective pressure might have generated a premature stop-codon within the gene arisen from the partial duplication of *nad2*, which might avoid expression of such a large non-functional protein, a process that may be rather expensive and yet unaffordable in terms of cellular energy.

Although they have no obvious function, repetitive sequences may play a role in the dynamics and evolution of mitochondrial genomes. Indeed, it is accepted that repetitive sequences are frequently involved in recombinations events that may not only lead to changes in gene

Table 5. Codon usage of the protein-coding genes in the mitochondrial genome of *S. lycopersici* CIDEFI-216.

Codon	Amino acid	Percentage	Frequency	Codon	Amino acid	Percentage	Frequency
GCA	Ala	33.7	179	CTT	Leu	15.1	215
GCC	Ala	8.5	45	TTA	Leu	60.6	865
GCG	Ala	5.6	30	TTG	Leu	8.0	114
GCT	Ala	52.2	277	AAA	Lys	85.4	653
AGA	Arg	68.7	266	AAG	Lys	14.6	112
AGG	Arg	9.8	38	ATG	Met	100	215
CGA	Arg	3.6	14	TTC	Phe	26.4	204
CGC	Arg	1.8	7	TTT	Phe	73.6	568
CGG	Arg	2.6	10	CCA	Pro	21.8	81
CGT	Arg	13.4	52	CCC	Pro	8.4	31
AAC	Asn	24.9	175	CCG	Pro	11.9	44
AAT	Asn	75.1	527	CCT	Pro	58.0	215
GAC	Asp	18.6	77	AGC	Ser	7.3	70
GAT	Asp	81.4	336	AGT	Ser	29.3	282
TGC	Cys	13.0	18	TCA	Ser	23.0	222
TGT	Cys	87.0	120	TCC	Ser	4.9	47
CAA	Gln	79.9	226	TCG	Ser	2.8	27
CAG	Gln	20.1	57	TCT	Ser	32.8	316
GAA	Glu	74.0	333	ACA	Thr	39.2	235
GAG	Glu	26.0	117	ACC	Thr	7.2	43
GGA	Gly	33.7	207	ACG	Thr	5.8	35
GGC	Gly	4.2	26	ACT	Thr	47.8	287
GGG	Gly	11.9	73	TGA	Trp	72.8	99
GGT	Gly	50.2	309	TGG	Trp	27.2	37
CAC	His	27.4	65	TAC	Tyr	27.1	155
CAT	His	72.6	172	TAT	Tyr	72.9	416
ATA	Ile	47.8	506	GTA	Val	43.2	275
ATC	Ile	9.8	104	GTC	Val	3.9	25
ATT	Ile	42.3	448	GTG	Val	10.2	65
CTA	Leu	12.4	177	GTT	Val	42.7	272
CTC	Leu	1.3	19	TAA	*	72.2	26
CTG	Leu	2.7	38	TAG	*	27.8	10

* Stop codón

<https://doi.org/10.1371/journal.pone.0185545.t005>

order and orientation, but also to the generation of novel genes [84, 85]. Besides their intriguing origin and function (if any), mitochondrial repetitive sequences represent attractive molecular markers to study populations, considering that they can be easily amplified because of their high copy number.

Lack of synteny among mitochondrial genomes of four representatives of Pleosporales

The alignment of the mitochondrial genomes of *S. lycopersici* (KX453765), *D. pinodes* (KT946597), *P. nodorum* [42] and *S. bambusicola* [43] revealed a remarkable variation in gene order and orientation (Fig 6). However, there are some conserved arrangements, like *nad2-nad3* and *cox1-cox2* gene pairs, which are next to each other, or even fused. Also, *nad4L* is next to *nad5*, though in *P. nodorum* [42] a non-conserved ORF between them was found. Similarly,

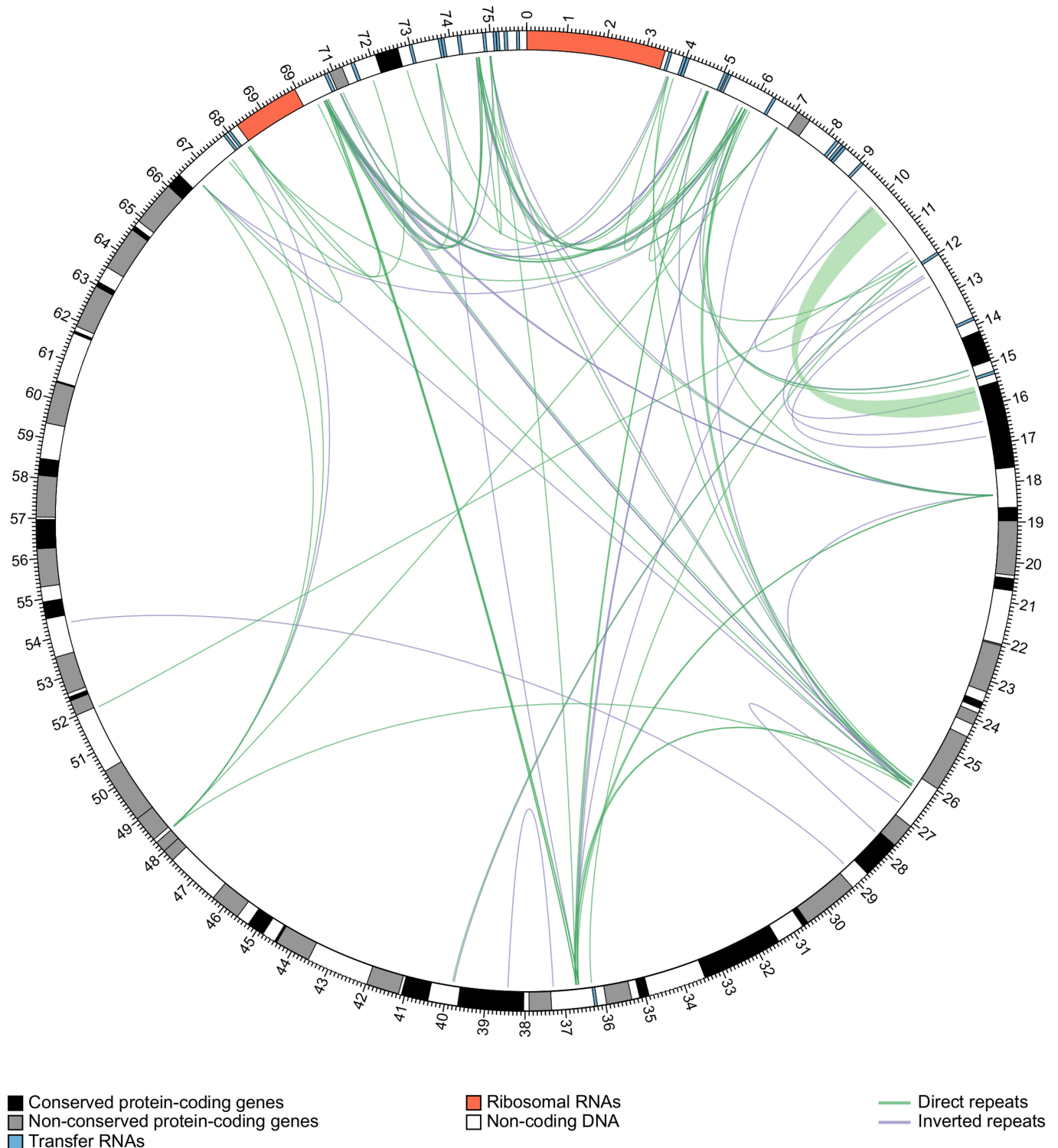


Fig 4. Map of the repetitive sequences in the mitochondrial genome of *S. lycopersici* CIDEFI-216. Lines connect repetitive sequences of at least 30 nucleotides in length and with a minimum match identity of 80%. A green ribbon shows the large duplication. The detailed references are given in the figure. Repetitive sequences were identified using the Vmatch software. The physical map was created with Circos.

<https://doi.org/10.1371/journal.pone.0185545.g004>

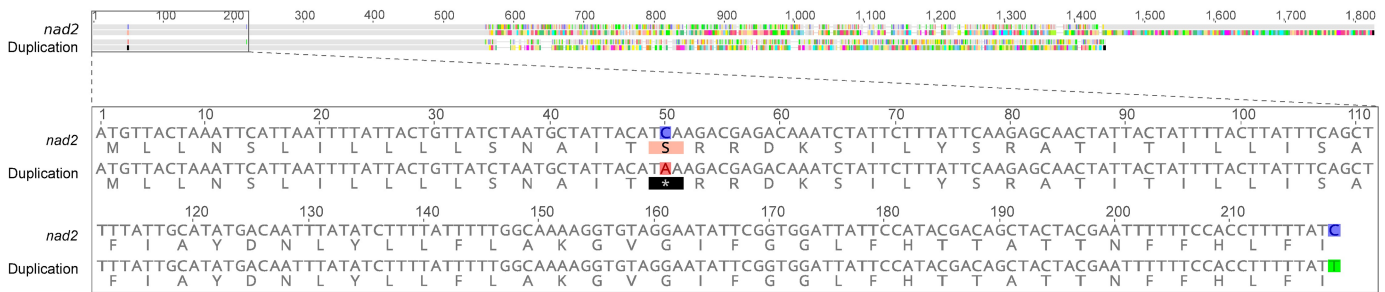


Fig 5. Sequence alignment between *nad2* and the large duplicated region found in the mitochondrial genome of *S. lycopersici*. The duplicated region has a length of 630 nucleotides, encompassing the first 590 nucleotides of *nad2*, albeit two SNPs (S17Stop and I219I) (See Fig 4). Nucleotide and amino-acid discrepancies are highlighted with colors. The nucleotide sequences were aligned using ClustalW with the default settings and were automatically translated using the translation Table 4 (The Mold, Protozoan, and Coelenterate Mitochondrial Code and the Mycoplasma/Spiroplasma Code) in Geneious 9.1.2.

<https://doi.org/10.1371/journal.pone.0185545.g005>

nad6 is nested between *rns* and *rnl* in a region rich in tRNA genes, but a non-conserved ORF is placed between *rns* and *nad6* in *D. pinodes* (KT946597) and *P. nodorum* [42]. In *S. bambusicola*, apart from *nad6*, *cox1* and *cox2* are also nested between *rns* and *rnl* [43]. It is also interesting the large difference in mitochondrial genomes sizes within the studied Pleosporales. In view of the multiple alignment, it seems that the mitochondrial genome expansion is related to

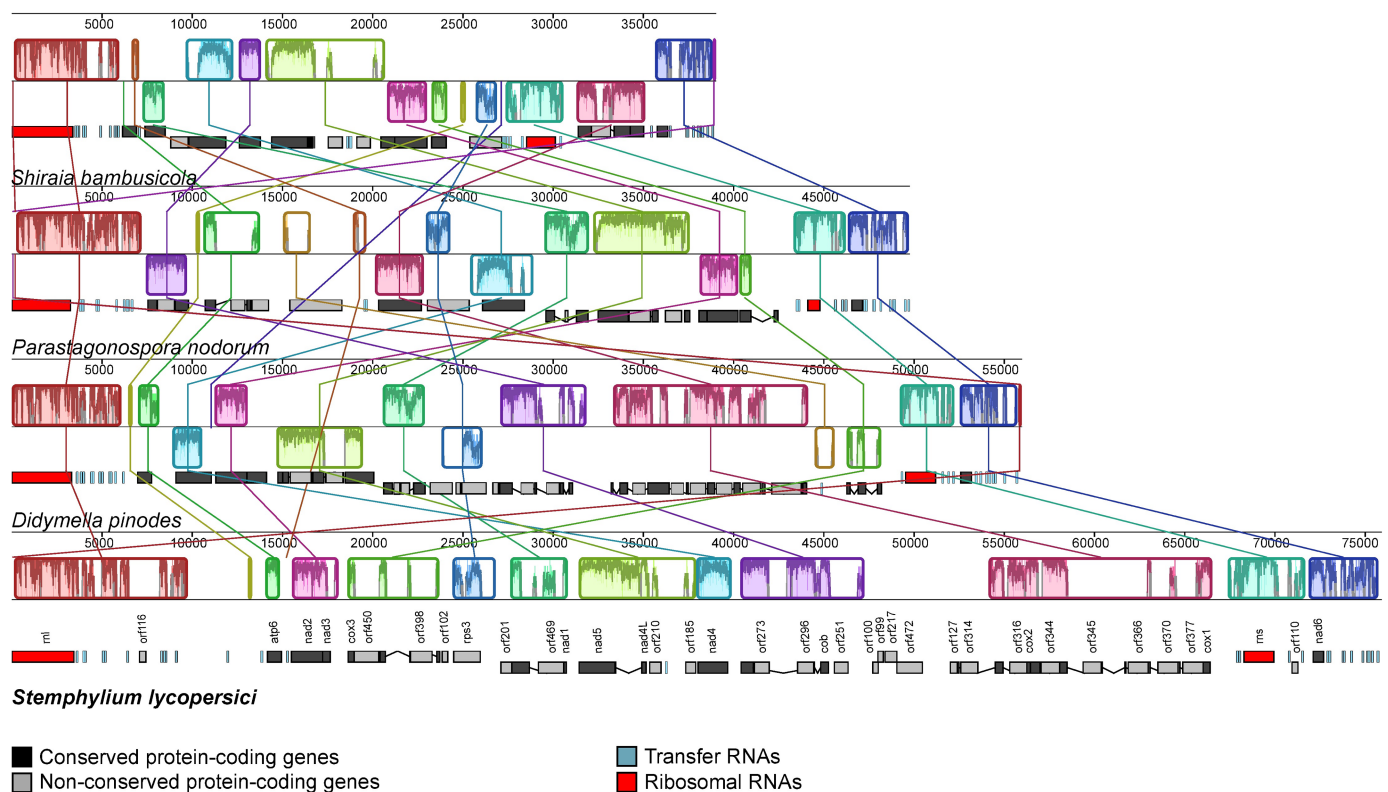


Fig 6. Whole mitochondrial alignments of four Pleosporales species. Gene order and orientation among mitochondrial genomes of the Pleosporales species *S. lycopersici* (KX453765), *D. pinodes* (KT946597), *P. nodorum* (EU053989) and *S. bambusicola* (KM382246). Locally collinear blocks are depicted by the same colors. The plot inside of them indicates the level of sequence similarity. The ruler above each genome represents the nucleotide positions. The detailed references are given in the figure. The mitochondrial genomes were arbitrary linearized starting at *ml* and were aligned using Mauve with the default settings. The backbone output was then modified with Inkscape 0.91.

<https://doi.org/10.1371/journal.pone.0185545.g006>

the presence of non-conserved ORFs, introns and repetitive elements, which, in turn, are particularly common within intergenic regions, as it was shown in the mitochondrial genome of *S. lycopersici* CIDEFI-216 (Fig 4).

Phylogenetic analysis of Pezizomycotina

Pezizomycotina is the largest subphylum of Ascomycota that includes remarkably diverse species in terms of their nutritional strategy [86]. Therefore, we further studied the systematic of Pezizomycotina and assessed the relationship of *S. lycopersici* with other taxa by means of a phylogenetic analysis using 12 conserved genes of 82 representatives of this subphylum. The analysis also included as outgroups a representative of Taphrinomycotina and two representatives of Saccharomycotina. Both maximum likelihood approach and Bayesian inference led to similar tree-topologies, except for the lack of congruence of a few internal nodes (Fig 7).

The phylogenetic tree strongly supports (bootstrap $\geq 99\%$ and posterior probabilities = 1) that Pezizomycotina is a monophyletic subphylum that include the clade Leotiomyceta, Pezizomycetes (specifically its order Pezizales), Eurotiomycetes and its orders (Chaetothiales, Onygenales and Eurotiales), Lecanoromycetes and its order (Peltigerales) and Sordariomycetes and its orders (Ophiostomatales, Diaporthales, Sordariales, Xylariales, Glomerellales and Microascales). Also, the class Dothideomycetes was resolved as monophyletic, albeit with a moderate support (bootstrap = 88%) in the maximum likelihood approach but with a strong support (posterior probabilities = 0.99) in the Bayesian inference. Nevertheless, both Capnodiales and Pleosporales were clearly shown to be monophyletic (bootstrap = 1 and posterior probabilities = 1). Similarly, the class Lecanoromycetes was resolved as monophyletic, but with lower support (bootstrap = 83% and posterior probabilities = 0.95), however members of the orders Helotiales and Leotiomycetes *inc. sed.* remained unsolved. The results described above are in accordance with the findings of Spatafora et. al, who reconstructed the phylogeny of Pezizomycotina based on five nuclear loci widely used in fungal taxonomy, namely: small subunit ribosomal DNA (SSU rDNA), long subunit ribosomal DNA (LSU rDNA), DNA-directed RNA polymerase II largest subunit (RPB1), DNA-directed RNA polymerase II second largest subunit (RPB2) and elongation factor 1-alpha (EF-1a) [86]. However, in our analysis members of Pezizomycotina are grouped together with the representative of Taphrinomycotina in a well-supported clade (bootstrap = 100% and posterior probabilities = 1), unlike most published data that support that Pezizomycotina and Saccharomycotina form a monophyletic group [86–89]. We argued that the observed difference might be due to some artifact of our phylogenetic reconstruction, particularly, by the Long Branch Attraction (LBA) effect, which could be explained by the clustering of fast-evolving lineages irrespective of their evolutionary relationships. As a matter of fact, mitochondrial data of some Taphrinomycotina and Saccharomycotina species are known to be subject to LBA artifacts [90, 91]. Also, it is worth mentioning that *K. pastoris* and *Y. lipolytica* (both Saccharomycotina) and *P. carinii* (Taphrinomycotina) have a particular characteristic that distinguish them from some of other taxa of their subphyla, which is the atypical presence of *atp6*, *cob*, *cox1*, *cox2*, *cox3*, *nad1*, *nad2*, *nad3*, *nad4*, *nad4L*, *nad5* and *nad6* genes within their mitochondrial genomes [91–93], a feature that was, in fact, used as a criterion for taxon sampling.

Regarding the relationships among the taxa belonging to Pleosporales, we found that *S. lycopersici* (Pleosporaceae) was clustered together with *P. nodorum* (Phaeosphaeriaceae) in a moderate-supported group (bootstrap = 75% and posterior probabilities = 1) with *S. bambusicola* (Shiraiaceae) as sister taxon, which were also in a clade that was separated from *D. pinodes* (Didymellaceae), though with low support values. However, it is not our intention to discuss the taxonomical relationship of Shiraiaceae within Pleosporales due to the lack of exhaustive

Numbers at the nodes represents maximum likelihood bootstrap support values (as a percentage of 1000 replicates) and Bayesian posterior probabilities, respectively. A hyphen (-) is used to indicate when a particular node was not recovered in the Bayesian inference. Branches with a dot represent branches whose lengths are not to scale and their actual lengths are indicated by the values placed above them. The scale bar represents the number of changes per sites.

<https://doi.org/10.1371/journal.pone.0185545.g007>

molecular phylogenetic studies including this family. Finally, it is important to point out that no correlation was found between the phylogenetic grouping of the Pleosporales representatives and the length of their mitochondrial genomes (Figs 6 and 7).

Conclusions

The mitochondrial genome of *S. lycopersici* CIDEFI-216 harbors 12 genes encoding proteins involved in the oxidative phosphorylation and electron transport systems (*atp6*, *cob*, *cox1*, *cox2*, *cox3*, *nad1*, *nad2*, *nad3*, *nad4*, *nad4L*, *nad5* and *nad6*). *Atp8* and *atp9*, two other genes widely conserved among mitochondrial genomes of filamentous ascomycetes, were found in the nuclear genome. The mitochondrial genome of *S. lycopersici* CIDEFI-216 displays an unusual gene structure around cytochrome-c oxidase subunits 1 and 2, which led to a single continuous ORF encoding both subunits, as occurs in other Pleosporales species and distantly-related protozoan taxa. The mitochondrial genome of *S. lycopersici* CIDEFI-216 also has a freestanding *rps3* gene and 13 genes whose hypothetical proteins possess at least one recognizable homing endonuclease domain. The analysis of such a limited number of samples raises the question if such features could be considered hallmarks of Pleosporales, which will be solved soon, in view of the increased rate of sequenced fungal genomes. The comparison of gene order and orientation in *S. lycopersici* and its close relatives in the order Pleosporales revealed the plasticity of their mitochondrial genomes, which seems to be mediated in large part by the action of repetitive elements and homing endonucleases. Finally, it was further proved the utility of fungal mitochondrial genomes in comparative and phylogenetic studies and highlighted their potential for population studies.

Supporting information

S1 Fig. Sequence alignments of the nuclear-encoded *atp8* and *atp9* of *S. lycopersici* CIDEFI-216 with their mitochondrial-encoded homologues in *Z. cellare* and *Z. tritici*. (a) Sequence alignment of *atp8*. (b) Sequence alignment of *atp9*. Nucleotide and amino-acid identities are highlighted with colors. The nucleotide sequences were aligned using ClustalW with the default settings and were automatically translated using the translation Table 4 (The Mold, Protozoan, and Coelenterate Mitochondrial Code and the Mycoplasma/Spiroplasma Code) in Geneious 9.1.2. In addition, for the particular case of *S. lycopersici*, the automated translation was also carried out using the translation Table 1 (Standard), achieving the same results. (EPS)

S2 Fig. Detailed sequence alignment of COX1 and COX2 with two different gene configurations. Detailed sequence alignment of the schematic sequence alignment presented in Fig 2. (EPS)

S3 Fig. Schematic sequence alignment of RPS3 from Pleosporales species. Detailed sequence alignment of the schematic sequence alignment presented in Fig 3. (EPS)

S1 Table. Functional annotation of the non-conserved protein-coding genes of the mitochondrial genome of *Stemphylium lycopersici* CIDEFI-216.

(XLSX)

S2 Table. TargetP 1.1 predictions.

(XLSX)

Author Contributions**Conceptualization:** Mario Emilio Ernesto Franco, Mario Carlos Nazareno Saparrat, Pedro Alberto Balatti.**Data curation:** Mario Emilio Ernesto Franco, Silvina Marianela Yanil López, Rocio Medina, César Gustavo Lucentini, Maria Inés Troncozo, Graciela Noemí Pastorino.**Formal analysis:** Mario Emilio Ernesto Franco, Silvina Marianela Yanil López, Rocio Medina, César Gustavo Lucentini, Maria Inés Troncozo, Graciela Noemí Pastorino.**Funding acquisition:** Mario Carlos Nazareno Saparrat, Pedro Alberto Balatti.**Investigation:** Mario Emilio Ernesto Franco, Mario Carlos Nazareno Saparrat, Pedro Alberto Balatti.**Methodology:** Mario Emilio Ernesto Franco, Mario Carlos Nazareno Saparrat, Pedro Alberto Balatti.**Project administration:** Mario Carlos Nazareno Saparrat, Pedro Alberto Balatti.**Resources:** Pedro Alberto Balatti.**Software:** Mario Emilio Ernesto Franco.**Supervision:** Mario Emilio Ernesto Franco, Mario Carlos Nazareno Saparrat, Pedro Alberto Balatti.**Validation:** Mario Emilio Ernesto Franco, Mario Carlos Nazareno Saparrat, Pedro Alberto Balatti.**Visualization:** Mario Emilio Ernesto Franco.**Writing – original draft:** Mario Emilio Ernesto Franco, Mario Carlos Nazareno Saparrat, Pedro Alberto Balatti.**Writing – review & editing:** Mario Emilio Ernesto Franco, Silvina Marianela Yanil López, Rocio Medina, César Gustavo Lucentini, Maria Inés Troncozo, Graciela Noemí Pastorino, Mario Carlos Nazareno Saparrat, Pedro Alberto Balatti.**References**

1. Câmara MP, O'Neill NR, Van Berkum P. Phylogeny of *Stemphylium* spp. based on ITS and glyceraldehyde-3-phosphate dehydrogenase gene sequences. *Mycologia*. 2002; 94(4), 660–672. <https://doi.org/10.1080/15572536.2003.11833194> PMID: 21156539
2. Farr DF, Rossman AY. Fungal databases, systematic mycology and microbiology laboratory, ARS, USDA. 2016; <http://nt.ars-grin.gov/fungalDATABASES/>. Accessed 2016.
3. Jones JB, Jones JP, Stall RE, Zitter TA. *Compendium of Tomato Diseases and Pests*. St. Paul, MN: American Phytopathological Society Press; 2014.
4. Kim BS, Yu SH, Cho HJ, Hwang HS. Gray leaf spot in peppers caused by *Stemphylium solani* and *S. lycopersici*. *The Plant Pathology Journal*. 2004; 20(2), 85–91. <https://doi.org/10.5423/PPJ.2004.20.2.085>

5. Nishi N, Muta T, Ito Y, Nakamura M, Tsukiboshi T. Ray speck of chrysanthemum caused by *Stemphylium lycopersici* in Japan. *Journal of General Plant Pathology*. 2009; 75(1), 80–82. <https://doi.org/10.1007/s10327-008-0133-1>
6. Tomioka K, Sato T, Sasaya T, Koganezawa H. Leaf spot of kalanchoe caused by *Stemphylium lycopersici*. *Japanese Journal of Phytopathology*. 1997; 63(4), 337–340. <https://doi.org/10.3186/jjphytopath.63.337>
7. Nasehi A, Kadir JB, Nasr-Esfahani M, Abed-Ashtiani F, Wong MY, Rambe SK, et al. Analysis of genetic and virulence variability of *Stemphylium lycopersici* associated with leaf spot of vegetable crops. *European journal of plant pathology*. 2014; 140(2), 261–273. <https://doi.org/10.1007/s10658-014-0460-3>
8. Nasehi A, Kadir JB, Esfahani MN, Mahmodi F, Golkhandan E, Akter S, et al. Cultural and physiological characteristics of *Stemphylium lycopersici* causing leaf blight disease on vegetable crops. *Archives of Phytopathology and Plant Protection*. 2014; 47(14), 1658–1665. <https://doi.org/10.1080/03235408.2013.853905>
9. Li F, Xue F, Yu X. GC–MS, FTIR and Raman Analysis of Antioxidant Components of Red Pigments from *Stemphylium lycopersici*. *Current microbiology*. 2017; 74(4), 532–539. <https://doi.org/10.1007/s00284-017-1220-3> PMID: 28255784
10. Franco MEE, Troncozo MI, López SMY, Lucentini G, Medina R, Saparrat, MC, et al. A survey on tomato leaf grey spot in the two main production areas of Argentina led to the isolation of *Stemphylium lycopersici* representatives which were genetically diverse and differed in their virulence. *European Journal of Plant Pathology*. 2017; 1–18. <https://doi.org/10.1007/s10658-017-1248-z>
11. Franco ME, López S, Medina R, Saparrat MC, Balatti P. Draft genome sequence and gene annotation of *Stemphylium lycopersici* strain CIDEFI-216. *Genome announcements*. 2015; 3(5), e01069–15. <https://doi.org/10.1128/genomeA.01069-15> PMID: 26404600
12. Chan DC. Mitochondria: dynamic organelles in disease, aging, and development. *Cell*. 2006; 125(7), 1241–1252. <https://doi.org/10.1016/j.cell.2006.06.010> PMID: 16814712
13. Basse CW. Mitochondrial inheritance in fungi. *Current opinion in microbiology*. 2010; 13(6), 712–719. <https://doi.org/10.1016/j.mib.2010.09.003> PMID: 20884279
14. Lang BF, Gray MW, Burger G. Mitochondrial genome evolution and the origin of eukaryotes. *Annual review of genetics*. 1999; 33(1), 351–397. <https://doi.org/10.1146/annurev.genet.33.1.351> PMID: 10690412
15. Bullerwell CE, Lang BF. Fungal evolution: the case of the vanishing mitochondrion. *Current opinion in microbiology*. 2005; 8(4), 362–369. <https://doi.org/10.1016/j.mib.2005.06.009> PMID: 15993645
16. Lang BF. Mitochondrial Genomes in Fungi. *Molecular Life Sciences: An Encyclopedic Reference*. 2014; 1–7. https://doi.org/10.1007/978-1-4614-6436-5_113-2
17. Lang BF, Laforest MJ, Burger G. Mitochondrial introns: a critical view. *Trends in Genetics*. 2007; 23(3), 119–125. <https://doi.org/10.1016/j.tig.2007.01.006> PMID: 17280737
18. Lambowitz AM, Belfort M. Mobile bacterial group II introns at the crux of eukaryotic evolution. *Microbiology spectrum*. 2015; 3(1). <https://doi.org/10.1128/microbiolspec.MDNA3-0050-2014> PMID: 26104554
19. Burt A, Koufopanou V. Homing endonuclease genes: the rise and fall and rise again of a selfish element. *Current opinion in genetics & development*. 2004; 14(6), 609–615. <https://doi.org/10.1016/j.gde.2004.09.010> PMID: 15531154
20. Stoddard BL. Homing endonuclease structure and function. *Quarterly reviews of biophysics* 2005; 38(1), 49–95. <https://doi.org/10.1017/S0033583505004063> PMID: 16336743
21. Edgell DR. Selfish DNA: homing endonucleases find a home. *Current Biology*. 2009; 19(3), R115–R117. <https://doi.org/10.1016/j.cub.2008.12.019> PMID: 19211047
22. Stoddard BL. Homing endonucleases: from microbial genetic invaders to reagents for targeted DNA modification. *Structure*. 2011; 19(1), 7–15. <https://doi.org/10.1016/j.str.2010.12.003> PMID: 21220111
23. Repar J, Warnecke T. Mobile introns shape the genetic diversity of their host genes. *Genetics*. 2017; 205(4), 1641–1648. <https://doi.org/10.1534/genetics.116.199059> PMID: 28193728
24. Wu B, Hao W. Horizontal transfer and gene conversion as an important driving force in shaping the landscape of mitochondrial introns. *G3: Genes, Genomes, Genetics*. 2014; 4(4), 605–612. <https://doi.org/10.1534/g3.113.009910> PMID: 24515269
25. Burger G, Gray MW, Lang BF. Mitochondrial genomes: anything goes. *Trends in genetics*. 2003; 19(12), 709–716. <https://doi.org/10.1016/j.tig.2003.10.012> PMID: 14642752
26. Pantou MP, Kouvelis VN, Typas MA. The complete mitochondrial genome of the vascular wilt fungus *Verticillium dahliae*: a novel gene order for *Verticillium* and a diagnostic tool for species identification. *Current genetics*. 2006; 50(2), 125–136. <https://doi.org/10.1007/s00294-006-0079-9> PMID: 16733756

27. Kouvelis VN, Sialakouma A, Typas MA. Mitochondrial gene sequences alone or combined with ITS region sequences provide firm molecular criteria for the classification of *Lecanicillium* species. *Mycological research*. 2008; 112(7), 829–844. <https://doi.org/10.1016/j.mycres.2008.01.016> PMID: 18501578
28. Bullerwell CE, Forget L, Lang BF. Evolution of monoblepharidalean fungi based on complete mitochondrial genome sequences. *Nucleic Acids Research*. 2003; 31(6), 1614–1623. <https://doi.org/10.1093/nar/gkg264> PMID: 12626702
29. van de Sande WW. Phylogenetic analysis of the complete mitochondrial genome of *Madurella mycetomatis* confirms its taxonomic position within the order Sordariales. *PLoS One*. 2012; 7(6), e38654. <https://doi.org/10.1371/journal.pone.0038654> PMID: 22701687
30. Paquin B, Laforest MJ, Forget L, Roewer I, Wang Z, Longcore J, et al. The fungal mitochondrial genome project: evolution of fungal mitochondrial genomes and their gene expression. *Current genetics*. 1997; 31(5), 380–395. <https://doi.org/10.1007/s002940050220> PMID: 9162109
31. Aguilera G, de Vienne DM, Ross ON, Hood ME, Giraud T, Petit E, et al. High variability of mitochondrial gene order among fungi. *Genome biology and evolution*. 2014; 6(2), 451–465. <https://doi.org/10.1093/gbe/evu028> PMID: 24504088
32. Formey D, Molès M, Haouy A, Savelli B, Bouchez O, Bécard G, et al. Comparative analysis of mitochondrial genomes of *Rhizophagus irregularis*–syn. *Glomus irregulare*–reveals a polymorphism induced by variability generating elements. *New phytologist*. 2012; 196(4), 1217–1227. <https://doi.org/10.1111/j.1469-8137.2012.04283.x> PMID: 22967288
33. Joardar V, Abrams NF, Hostetler J, Paukstelis PJ, Pakala S, Pakala SB, et al. Sequencing of mitochondrial genomes of nine *Aspergillus* and *Penicillium* species identifies mobile introns and accessory genes as main sources of genome size variability. *BMC genomics*. 2012; 13(1), 698. <https://doi.org/10.1186/1471-2164-13-698> PMID: 23234273
34. Zhang Y, Zhang S, Zhang G, Liu X, Wang C, Xu J. Comparison of mitochondrial genomes provides insights into intron dynamics and evolution in the caterpillar fungus *Cordyceps militaris*. 2015; *Fungal genetics and biology*, 77, 95–107. <https://doi.org/10.1016/j.fgb.2015.04.009> PMID: 25896956
35. Kim YS, Dixon EW, Vincelli P, Farman ML. Field resistance to strobilurin (QoI) fungicides in *Pyricularia grisea* caused by mutations in the mitochondrial cytochrome b gene. *Phytopathology*. 2003; 93(7), 891–900. <https://doi.org/10.1094/PHYTO.2003.93.7.891> PMID: 18943171
36. Ma Z, Michailides TJ. Advances in understanding molecular mechanisms of fungicide resistance and molecular detection of resistant genotypes in phytopathogenic fungi. *Crop Protection*. 2005; 24(10), 853–863. <https://doi.org/10.1016/j.cropro.2005.01.011>
37. Kearse M, Moir R, Wilson A, Stones-Havas S, Cheung M, Sturrock S, et al. Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics*. 2012; 28(12), 1647–1649. <https://doi.org/10.1093/bioinformatics/bts199> PMID: 22543367
38. Lowe TM, Eddy SR. tRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence. *Nucleic acids research*. 1997; 25(5), 955–964. <https://doi.org/10.1093/nar/25.5.0955> PMID: 9023104
39. Conesa A, Götz S, García-Gómez JM, Terol J, Talón M, Robles M. Blast2GO: a universal tool for annotation, visualization and analysis in functional genomics research. *Bioinformatics*. 2005; 21(18), 3674–3676. <https://doi.org/10.1093/bioinformatics/bti610> PMID: 16081474
40. Lohse M, Drechsel O, Bock R. OrganellarGenomeDRAW (OGDRAW): a tool for the easy generation of high-quality custom graphical maps of plastid and mitochondrial genomes. *Current genetics*. 2007; 52(5–6), 267–274. <https://doi.org/10.1007/s00294-007-0161-y> PMID: 17957369
41. Krzywinski M, Schein J, Birol I, Connors J, Gascoyne R, Horsman D, et al. Circos: an information aesthetic for comparative genomics. *Genome research*. 2009; 19(9), 1639–1645. <https://doi.org/10.1101/gr.092759.109> PMID: 19541911
42. Hane JK, Lowe RGT, Solomon PS, Tan K-C, Schoch CL, Spatafora JW, et al. Dothideomycete–plant interactions illuminated by genome sequencing and EST analysis of the wheat pathogen *Stagonospora nodorum*. *The Plant Cell*. 2007; 19(11), 3347–3368. <https://doi.org/10.1105/tpc.107.052829> PMID: 18024570
43. Shen XY, Li T, Chen S, Fan L, Gao J, Hou CL. Characterization and phylogenetic analysis of the mitochondrial genome of *Shiraia bambusicola* reveals special features in the order of Pleosporales. *PloS one*. 2015; 10(3), e0116466. <https://doi.org/10.1371/journal.pone.0116466> PMID: 25790308
44. Darling AC, Mau B, Blattner FR, Perna NT. Mauve: multiple alignment of conserved genomic sequence with rearrangements. *Genome research*. 2004; 14(7), 1394–1403. <https://doi.org/10.1101/gr.2289704> PMID: 15231754
45. Edgar RC. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic acids research*. 2004; 32(5), 1792–1797. <https://doi.org/10.1093/nar/gkh340> PMID: 15034147

46. Castresana J. Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis. *Molecular biology and evolution*. 2000; 17(4), 540–552. <https://doi.org/10.1093/oxfordjournals.molbev.a026334> PMID: 10742046
47. Darriba D, Taboada GL, Doallo R, Posada D. ProtTest 3: fast selection of best-fit models of protein evolution. *Bioinformatics*. 2011; 27(8), 1164–1165. <https://doi.org/10.1093/bioinformatics/btr088> PMID: 21335321
48. Akaike H. A new look at the statistical model identification. *IEEE transactions on automatic control*. 1974; 19(6), 716–723. <https://doi.org/10.1109/TAC.1974.1100705>
49. Guindon S, Gascuel O. A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. *Systematic biology*. 2003; 52(5), 696–704. <https://doi.org/10.1080/10635150390235520> PMID: 14530136
50. Ronquist F, Teslenko M, van der Mark P, Ayres D, Darling A, Höhna S, et al. MrBayes 3.2: Efficient Bayesian phylogenetic inference and model choice across a large model space. *Systematic Biology*. 2012; 61(3), 539–542. <https://doi.org/10.1093/sysbio/sys029> PMID: 22357727
51. Felsenstein J. Confidence limits on phylogenies: an approach using the bootstrap. *Evolution*. 1985; 39(3), 783–791. <https://doi.org/10.1111/j.1558-5646.1985.tb00420.x> PMID: 28561359
52. Haas BJ, Zeng Q, Pearson MD, Cuomo CA, Wortman JR. Approaches to fungal genome annotation. *Mycology*. 2011; 2(3), 118–141. <https://doi.org/10.1080/21501203.2011.606851> PMID: 22059117
53. Chacinska A, Koehler CM, Milenkovic D, Lithgow T, Pfanner N. Importing mitochondrial proteins: machineries and mechanisms. *Cell*. 2009; 138(4), 628–644. <https://doi.org/10.1016/j.cell.2009.08.005> PMID: 19703392
54. Cummings DJ, McNally KL, Domenico JM, Matsuura ET. The complete DNA sequence of the mitochondrial genome of *Podospira anserina*. *Current genetics*. 1990; 17(5), 375–402. <https://doi.org/10.1007/BF00334517> PMID: 2357736
55. Yao Y, Lin R, Tian X, Shen B, Mao Z, Xie B. The complete mitochondrial genome of the nematophagous fungus *Acremonium implicatum*. *Mitochondrial DNA Part A*. 2016; 27(5), 3246–3247. <https://doi.org/10.3109/19401736.2015.1007367> PMID: 25630733
56. Solieri L. Mitochondrial inheritance in budding yeasts: towards an integrated understanding. *Trends in microbiology*. 2010; 18(11), 521–530. <https://doi.org/10.1016/j.tim.2010.08.001> PMID: 20832322
57. Freel KC, Friedrich A, Schacherer J. Mitochondrial genome evolution in yeasts: an all-encompassing view. *FEMS yeast research*. 2015; 15(4), fov023. <https://doi.org/10.1093/femsyr/fov023> PMID: 25969454
58. Foury F, Roganti T, Lecrenier N, Purnelle B. The complete sequence of the mitochondrial genome of *Saccharomyces cerevisiae*. *FEBS letters*. 1998; 440(3), 325–331. [https://doi.org/10.1016/S0014-5793\(98\)01467-7](https://doi.org/10.1016/S0014-5793(98)01467-7) PMID: 9872396
59. Bullerwell CE, Leigh J, Forget L, Lang BF. A comparison of three fission yeast mitochondrial genomes. *Nucleic acids research*. 2003; 31(2), 759–768. <https://doi.org/10.1093/nar/gkg134> PMID: 12527786
60. Talla E, Anthouard V, Bouchier C, Frangeul L, Dujon B. The complete mitochondrial genome of the yeast *Kluyveromyces thermotolerans*. *FEBS letters*. 2005; 579(1), 30–40. <https://doi.org/10.1016/j.febslet.2004.10.106> PMID: 15620687
61. Goodwin SB, McCorison CB, Cavaletto JR, Culley DE, LaButti K Baker SE, et al. The mitochondrial genome of the ethanol-metabolizing, wine cellar mold *Zasmidium cellare* is the smallest for a filamentous ascomycete. *Fungal biology*. 2016; 120(8), 961–974. <https://doi.org/10.1016/j.funbio.2016.05.003> PMID: 27521628
62. Torriani SF, Goodwin SB, Kema GH, Pangilinan JL, McDonald BA. Intraspecific comparison and annotation of two complete mitochondrial genome sequences from the plant pathogenic fungus *Mycosphaerella graminicola*. *Fungal Genetics and Biology*. 2008; 45(5), 628–637. <https://doi.org/10.1016/j.fgb.2007.12.005> PMID: 18226935
63. Altschul SF, Madden TL, Schäffer AA, Zhang J, Zhang Z, Miller W, et al. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic acids research*. 1997; 25(17), 3389–3402. <https://doi.org/10.1093/nar/25.17.3389> PMID: 9254694
64. Emanuelsson O, Brunak S, Von Heijne G, Nielsen H. Locating proteins in the cell using TargetP, SignalP and related tools. *Nature protocols*. 2007; 2(4), 953–971. <https://doi.org/10.1038/nprot.2007.131> PMID: 17446895
65. Capaldi RA. Structure and function of cytochrome c oxidase. *Annual review of biochemistry*. 1990; 59(1), 569–596. <https://doi.org/10.1146/annurev.bi.59.070190.003033> PMID: 2165384
66. Gray MW, Lang BF, Cedergren R, Golding GB, Lemieux C, Sankoff D, et al. Genome structure and gene content in protist mitochondrial DNAs. *Nucleic acids research*. 1998; 26(4), 865–878. <https://doi.org/10.1093/nar/26.4.865> PMID: 9461442

67. Hebert PD, Cywinska A, Ball SL. Biological identifications through DNA barcodes. *Proceedings of the Royal Society of London B: Biological Sciences*. 2003; 270(1512), 313–321. <https://doi.org/10.1098/rspb.2002.2218> PMID: 12614582
68. Hajibabaei M, Singer GA, Hebert PD, Hickey DA. DNA barcoding: how it complements taxonomy, molecular phylogenetics and population genetics. *TRENDS in Genetics*. 2007; 23(4), 167–172. <https://doi.org/10.1016/j.tig.2007.02.001> PMID: 17316886
69. Burger G, Plante I, Lonergan KM, Gray MW. The mitochondrial DNA of the amoeboid protozoan, *Acanthamoeba castellanii*: complete sequence, gene content and genome organization. *Journal of molecular biology*. 1995; 245(5), 522–537. <https://doi.org/10.1006/jmbi.1994.0043> PMID: 7844823
70. Lonergan KM, Gray MW. Expression of a Continuous Open Reading Frame Encoding Subunits 1 and 2 of Cytochrome c Oxidase in the Mitochondrial DNA of *Acanthamoeba castellanii*. *Journal of molecular biology*. 1996; 257(5), 1019–1030. <https://doi.org/10.1006/jmbi.1996.0220> PMID: 8632465
71. Pellizzari R, Anjard C, Bisson R. Subunits I and II of *Dictyostelium* cytochrome c oxidase are specified by a single open reading frame transcribed into a large polycistronic RNA. *Biochimica et Biophysica Acta (BBA)-Bioenergetics*. 1997; 1320(1), 1–7. [https://doi.org/10.1016/S0005-2728\(97\)00010-8](https://doi.org/10.1016/S0005-2728(97)00010-8)
72. Gray MW, Burger G, Lang BF. Mitochondrial evolution. *Science*. 1999; 283(5407), 1476–1481. <https://doi.org/10.1126/science.283.5407.1476> PMID: 10066161
73. Santamaria M, Vicario S, Pappadà G, Scioscia G, Scazzocchio C, Saccone C. Towards barcode markers in Fungi: an intron map of Ascomycota mitochondria. *BMC bioinformatics*. 2009; 10(6), S15. <https://doi.org/10.1186/1471-2105-10-S6-S15> PMID: 19534740
74. Férandon C, Moukha S, Callac P, Benedetto JP, Castroviejo M, Barroso G. The *Agaricus bisporus* *cox1* gene: the longest mitochondrial gene and the largest reservoir of mitochondrial group I introns. *PLoS One*. 2010; 5(11), e14048. <https://doi.org/10.1371/journal.pone.0014048> PMID: 21124976
75. Vaughn JC, Mason MT, Sper-Whitis GL, Kuhlman P, Palmer JD. Fungal origin by horizontal transfer of a plant mitochondrial group I intron in the chimeric *cox1* gene of *Peperomia*. *Journal of Molecular Evolution*. 1995; 41(5), 563–572. <https://doi.org/10.1007/BF00175814> PMID: 7490770
76. Cho Y, Palmer JD. Multiple acquisitions via horizontal transfer of a group I intron in the mitochondrial *cox1* gene during evolution of the Araceae family. *Molecular Biology and Evolution*. 1999; 16(9), 1155–1165. <https://doi.org/10.1093/oxfordjournals.molbev.a026206> PMID: 10486971
77. Sanchez-Puerta MV, Abbona CC, Zhuo S, Tepe EJ, Bohs L, Olmstead RG, et al. Multiple recent horizontal transfers of the *cox1* intron in Solanaceae and extended co-conversion of flanking exons. *BMC evolutionary biology*. 2011; 11(1), 277. <https://doi.org/10.1186/1471-2148-11-277> PMID: 21943226
78. Bullerwell CE, Burger G, Lang BF. A novel motif for identifying rps3 homologs in fungal mitochondrial genomes. *Trends in biochemical sciences*. 2000; 25(8), 363–365. [https://doi.org/10.1016/S0968-0004\(00\)01612-1](https://doi.org/10.1016/S0968-0004(00)01612-1) PMID: 10916154
79. Sethuraman J, Majer A, Iranpour M, Hausner G. Molecular evolution of the mtDNA encoded rps3 gene among filamentous ascomycetes fungi with an emphasis on the ophiostomatoid fungi. *Journal of molecular evolution*. 2009; 69(4), 372–385. <https://doi.org/10.1007/s00239-009-9291-9> PMID: 19826748
80. Groot GS, Mason TL, Harten-Loosbroek N. Var 1 is associated with the small ribosomal subunit of mitochondrial ribosomes in yeast. *Molecular and General Genetics MGG*. 1979; 174(3), 339–342. <https://doi.org/10.1007/BF00267808> PMID: 384172
81. Burke JM, Rajbhandary UL. Intron within the large rRNA gene of *N. crassa* mitochondria: a long open reading frame and a consensus sequence possibly important in splicing. *Cell*. 1982; 31(3), 509–520. [https://doi.org/10.1016/0092-8674\(82\)90307-5](https://doi.org/10.1016/0092-8674(82)90307-5)
82. Sethuraman J, Majer A, Friedrich NC, Edgell DR, Hausner G. Genes within genes: multiple LAGLI-DADG homing endonucleases target the ribosomal protein S3 gene encoded within an rnl group I intron of *Ophiostoma* and related taxa. *Molecular biology and evolution*. 2009; 26(10), 2299–2315. <https://doi.org/10.1093/molbev/msp145> PMID: 19597163
83. Bonen L, Vogel J. The ins and outs of group II introns. *TRENDS in Genetics*. 2001; 17(6), 322–331. [https://doi.org/10.1016/S0168-9525\(01\)02324-1](https://doi.org/10.1016/S0168-9525(01)02324-1) PMID: 11377794
84. Bowen NJ, Jordan IK. Transposable elements and the evolution of eukaryotic complexity. *Current issues in molecular biology*. 2002; 4, 65–76. PMID: 12074196
85. Phadnis N, Sia RA, Sia EA. Analysis of repeat-mediated deletions in the mitochondrial genome of *Saccharomyces cerevisiae*. *Genetics*. 2005; 171(4), 1549–1559. <https://doi.org/10.1534/genetics.105.047092> PMID: 16157666
86. Spatafora JW, Sung GH, Johnson D, Hesse C, O'Rourke B, Serdani M, et al. A five-gene phylogeny of Pezizomycotina. *Mycologia*. 2006; 98(6), 1018–1028. <https://doi.org/10.1080/15572536.2006.11832630> PMID: 17486977

87. Berbee ML. The phylogeny of plant and animal pathogens in the Ascomycota. *Physiological and Molecular Plant Pathology*. 2001; 59(4), 165–187. <https://doi.org/10.1006/pmpp.2001.0355>
88. Reeb V, Lutzoni F, Roux C. Contribution of RPB2 to multilocus phylogenetic studies of the euascomycetes (Pezizomycotina, Fungi) with special emphasis on the lichen-forming Acarosporaceae and evolution of polyspory. *Molecular phylogenetics and evolution*. 2004; 32(3), 1036–1060. <https://doi.org/10.1016/j.ympev.2004.04.012> PMID: 15288074
89. Schoch CL, Sung GH, Lopez-Giraldez F, Townsend JP, Miadlikowska J, Hofstetter V, et al. The Ascomycota tree of life: a phylum-wide phylogeny clarifies the origin and evolution of fundamental reproductive and ecological traits. *Systematic biology*. 2009; 58(2), 224–239. <https://doi.org/10.1093/sysbio/syp020> PMID: 20525580
90. Liu Y, Leigh JW, Brinkmann H, Cushion MT, Rodriguez-Ezpeleta N, Philippe H, et al. Phylogenomic analyses support the monophyly of Taphrinomycotina, including Schizosaccharomyces fission yeasts. *Molecular biology and evolution*. 2008; 26(1), 27–34. <https://doi.org/10.1093/molbev/msn221> PMID: 18922765
91. Ma L, Huang DW, Cuomo CA, Sykes S, Fantoni G, Das B, et al. Sequencing and characterization of the complete mitochondrial genomes of three *Pneumocystis* species provide new insights into divergence between human and rodent *Pneumocystis*. *The FASEB Journal*. 2013; 27(5), 1962–1972. <https://doi.org/10.1096/fj.12-224444> PMID: 23392351
92. Küberl A, Schneider J, Thallinger GG, Anderl I, Wibberg D, Hajek T, et al. High-quality genome sequence of *Pichia pastoris* CBS7435. *Journal of biotechnology*. 2011; 154(4), 312–320. <https://doi.org/10.1016/j.jbiotec.2011.04.014> PMID: 21575661
93. Magnan C, Yu J, Chang I, Jahn E, Kanomata Y, Wu J, et al. Sequence assembly of *Yarrowia lipolytica* strain W29/CLIB89 shows transposable element diversity. *PloS one*. 2016; 11(9), e0162363. <https://doi.org/10.1371/journal.pone.0162363> PMID: 27603307