


## RESEARCH ARTICLE

# Mitochondrial genome sequence of *Phytophthora sansomeana* and comparative analysis of *Phytophthora* mitochondrial genomes

Guohong Cai <sup>\*</sup>, Steven R. Scofield

Crop Production and Pest Control Research Unit, Agricultural Research Service, USDA, and College of Agriculture, Purdue University, West Lafayette, Indiana, United States of America

\* [Guohong.cai@usda.gov](mailto:Guohong.cai@usda.gov)

## Abstract

*Phytophthora sansomeana* infects soybean and causes root rot. It was recently separated from the species complex *P. megasperma sensu lato*. In this study, we sequenced and annotated its complete mitochondrial genome and compared it to that of nine other *Phytophthora* species. The genome was assembled into a circular molecule of 39,618 bp with a 22.03% G+C content. Forty-two protein coding genes, 25 tRNA genes and two rRNA genes were annotated in this genome. The protein coding genes include 14 genes in the respiratory complexes, four ATP synthase genes, 16 ribosomal proteins genes, a *tatC* translocase gene, six conserved ORFs and a unique *orf402*. The tRNA genes encode tRNAs for 19 amino acids. Comparison among mitochondrial genomes of 10 *Phytophthora* species revealed three inversions, each covering multiple genes. These genomes were conserved in gene content with few exceptions. A 3' truncated *atp9* gene was found in *P. nicotianae*. All 10 *Phytophthora* species, as well as other oomycetes and stramenopiles, lacked tRNA genes for threonine in their mitochondria. Phylogenomic analysis using the mitochondrial genomes supported or enhanced previous findings of the phylogeny of *Phytophthora* spp.

## OPEN ACCESS

**Citation:** Cai G, Scofield SR (2020) Mitochondrial genome sequence of *Phytophthora sansomeana* and comparative analysis of *Phytophthora* mitochondrial genomes. PLoS ONE 15(5): e0231296. <https://doi.org/10.1371/journal.pone.0231296>

**Editor:** Hector Escriva, Laboratoire Arago, FRANCE

**Received:** March 20, 2020

**Accepted:** April 28, 2020

**Published:** May 14, 2020

**Copyright:** This is an open access article, free of all copyright, and may be freely reproduced, distributed, transmitted, modified, built upon, or otherwise used by anyone for any lawful purpose. The work is made available under the [Creative Commons CC0](https://creativecommons.org/licenses/by/4.0/) public domain dedication.

**Data Availability Statement:** The mitochondrial assembly was submitted to Genbank under accession: MH936679.

**Funding:** This study was supported by the United States Department of Agriculture. The authors, as USDA employees, performed study design, data collection and analysis, and manuscript preparation.

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

## Introduction

The genus *Phytophthora* includes many devastating pathogens infecting economically important crops [1]. Perhaps the best known and most historically significant among those is *P. infestans*, the pathogen that causes potato late blight and the culprit behind the Irish potato famine [2]. During the past 20 years, the number of species described in the *Phytophthora* genus has expanded significantly. With approximately 55 species described in 1999, it expanded to 105 species by 2007 [3] and 117 species by 2012 [4]. Currently, at least 185 formally described species and provisional species have been reported [5–7]. *Phytophthora* is a genus in the *Peronosporales* of oomycetes. Oomycetes produce hyphae and are morphologically similar to fungi. However, they are phylogenetically distant. Oomycetes belong to the major group Stramenopila, which also include diatoms and brown algae [8–10].

Phylogenetic analysis grouped *Phytophthora* species into 10 clades [6, 11, 12]. *P. sansomeana* is one of the recently described species in clade 8 [13]. Previously included in the species complex *P. megasperma sensu lato*, it is morphologically similar to but phylogenetically distinct from *P. megasperma sensu strictu*. *P. sansomeana* was isolated from soybean in Indiana in 1990 [14] and has now been reported in China [15], Canada and multiple Midwest states in the USA [16, 17]. It lives in the soil, infects soybean roots, and causes discoloration and rotting of lateral root and internal discoloration and rotting of the taproot. Above ground, it can cause yellowing and stunting, or whole plant wilting, but the chocolate-colored stem discoloration which is typically associated with late-season infection by another soybean root rot pathogen, *P. sojae*, is usually absent in *P. sansomeana*-infected soybean plants. This pathogen overwinters as oospore in the soil and in soybean debris. Also included in this species are isolates from Douglas-fir and alfalfa [13].

Interest in the genus *Phytophthora* has been increasing due to its economic impact. Molecular data serves as the foundation of many studies. Genome sequences have been available for many *Phytophthora* species. Complete mitochondrial genomes from eight species have been published. These include *P. infestans* [18, 19], *P. andina*, *P. ipomoeae*, *P. mirabilis*, *P. phaseoli* [20], *P. nicotianae* [21], *P. sojae* and *P. ramorum* [22]. Another species, *P. polonica*, also has its complete mitochondrial genome publicly available in GenBank (Table 1). Inconsistencies in annotation have been observed (see “Results and discussion” for details). Mitochondrial genomes have been used to study the population of *P. infestans* [23–26] as well as resolving the phylogenetic relationship of *Phytophthora* species [12, 19]. Molecular information from *P. sansomeana* is limited. In this study, we sequenced and annotated its complete mitochondrial genome, compared it to that of other *Phytophthora* species currently publicly available, and examined phylogenetic relationship among these species using their mitochondrial genomes.

**Table 1. Mitochondrial genomes of 10 *Phytophthora* species included in the analysis in this study.**

Species	Accession #	Length (bp)	GC (%)	Clade <sup>a</sup>	Reference
<i>P. infestans</i>					
Haplotype Ia	AY894835	37,922	22.32	1c	[19]
Ib	U17009	37,957	22.29	1c	[18]
IIa	AY898627	39,870	22.29	1c	[19]
IIb	AY898628	39,840	22.30	1c	[19]
<i>P. andina</i>					
Haplotype Ia	KJ408269 <sup>b</sup>	37,883	22.29	1c	[20]
Ic	HM590419	37,874	22.14	1c	[20]
<i>P. ipomoeae</i>	HM590420	37,872	22.39	1c	[20]
<i>P. mirabilis</i>	HM590421	37,779	22.38	1c	[20]
<i>P. phaseoli</i>	HM590418	37,914	22.14	1c	[20]
<i>P. nicotianae</i>	KY851301	37,561	21.88	1	[21]
<i>P. sojae</i>	DQ832717	42,977	21.70	7	[22]
<i>P. polonica</i>	KT946598	40,467	21.57	9	N/A
<i>P. ramorum</i>	DQ832718	39,314	21.98	8c	[22]
<i>P. sansomeana</i>	MH936679	39,618	22.03	8a	This study

<sup>a</sup> The designation of phylogenetic clade is based on [11, 12].

<sup>b</sup> This genome is incomplete.

<https://doi.org/10.1371/journal.pone.0231296.t001>

## Materials and methods

### Nucleic acids extraction and sequencing

The type strain of *P. sansomeana*, 1819b, was obtained from American Type Culture Collection (ATCC #: MYA-4455). It was maintained on lima bean agar. For nucleic acids extraction, it was grown in half-strength lima bean broth on a lab bench for 8 days. DNA was extracted from harvested mycelium using the Gentra Puregene Yeast/Bact kit (Qiagen).

DNA was sequenced using both PacBio and Illumina technologies. For PacBio sequencing, a 20-kb insert library was constructed and then run on two SMRT cells on the RS II sequencing platform using the P6-C4 chemistry. For Illumina sequencing, a TruSeq DNA PCR-free library with mean insert size of approximately 440 bp was constructed according to manufacturer's instructions. It was sequenced on the HighSeq2500 platform.

### Assembly, annotation and comparative genomics

The PacBio reads were assembled using HGAP (RS\_HGAP\_Assembly.3 protocol implemented in PacBio smrtanalysis 2.3.0 software package) and polished using Quiver [27]. Illumina read-pairs were mapped onto the assembly using Bowtie2 [28], and Pilon [29] was used to further improve the assembly by utilizing the aligned Illumina reads.

The mitochondrial genome of *P. sansomeana* was first annotated using MFannot (<http://megasun.bch.umontreal.ca/cgi-bin/mfannot/mfannotInterface.pl>) using the standard genetic code. Like plant mitochondria, but different from those of animal and fungi, oomycete mitochondria use the standard code [30]. The automatic annotation was then verified manually. Mitochondrial genomes of nine other *Phytophthora* species were downloaded from GenBank (Table 1). Protein coding genes among the species were compared. Genes identified in one species but not in another were manually verified or corrected. For this purpose, ORFfinder (<https://www.ncbi.nlm.nih.gov/orffinder/>) was used to find open reading frames (ORFs) and BLAST searches were used to identify similarity. ORFfinder was also used to find ORFs encoding at least 100 aa in the intergenic regions of all *Phytophthora* mitochondrial genomes. Ribosomal RNA (rRNA) gene sequences were used to search NCBI non-redundant database and SILVA rRNA database [31]. Transfer RNA (tRNA) genes were analyzed using tRNAscan-SE 2.0 [32] in the organellar mode. A circular map of the mitochondrial genome of *P. sansomeana* was generated using OGDRAW [33]. Its genome assembly and annotation were submitted to GenBank under accession number MH936679.

### Phylogenomic analysis

To construct phylogenetic tree based on the mitochondrial genomes, individual protein coding genes and rRNA genes were aligned with Muscle [34]. The alignments were visually inspected to remove gaps and regions deemed ambiguously aligned (usually between closely adjacent gaps). The alignments were concatenated and a maximum-likelihood tree based on the Tamura-Nei nucleotide model [35] was constructed using MEGA7 [36]. One thousand bootstraps were used to test the support of individual branches. *Pythium ultimum* [37] and *Pythium insidiosum* [38] were used as outgroups.

## Results and discussion

### Assembly and annotation of *P. sansomeana* mitochondrial genome

PacBio sequencing generated 383,775 subreads totaling 2.57 Gb after filtering and adapter removal. The N50 of subread length was 9,846 bp. The mitochondrial genome of *P.*

*sansomeana* was assembled into a single contig of 55.5 kb using PacBio reads, with a 15.8 kb direct duplicated region at both ends, indicating a circular topology. To verify this, the contig was broken at a random position in the non-duplicated region and the two segments were re-connected by merging the duplicated regions. PacBio reads were then mapped to this re-arranged contig. Visual inspection of read alignments and coverage confirmed its circular topology. After polishing, the PacBio assembly resulted in a molecule of 39,546 bp. Illumina sequencing generated 43.3 million read-pairs at 151 bp x 2 in length. A total of 2,249,965 read-pairs were mapped to the mitochondrial genome. Based on the alignments of Illumina reads, 72 single nucleotide deletions were corrected.

The mitochondrial genome of *P. sansomeana* was assembled into a circular molecule of 39,618 bp with a 22.03% G+C content. It was predicted to encode 42 protein coding genes, two ribosomal RNA genes, and 25 tRNA genes. The genes were encoded by both strands. The protein coding genes included 10 NADH dehydrogenase genes in respiratory complex I (*nad1*, *nad2*, *nad3*, *nad4*, *nad4L*, *nad5*, *nad6*, *nad7*, *nad9* and *nad11*), one cytochrome C reductase gene (*cob*) in complex III, and three cytochrome c oxidase genes in complex IV (*cox1*, *cox2* and *cox3*). Also included were four ATP synthase genes (*atp1*, *atp6*, *atp8* and *atp9*), five large subunit ribosomal protein genes (*rpl2*, *rpl5*, *rpl6*, *rpl14* and *rpl16*), and 11 small subunit ribosomal protein genes (*rps2*, *rps3*, *rps4*, *rps7*, *rps8*, *rps10*, *rps11*, *rps12*, *rps13*, *rps14* and *rps19*) (Fig 1). We annotated a sec-independent translocase gene (*tatC*) in the twin-arginine translocation system. The annotation of this gene was inconsistent in other *Phytophthora* mitochondrial genomes (see analysis below). The six conserved ORFs with no annotated function found in other *Phytophthora* mitochondrial genomes (Table 1) were also found in *P. sansomeana* (*orf32*, *ymf96* (= *orf79*), *ymf98* (= *orf142*), *ymf99* (= *orf217*), *ymf100* (= *orf100*), and *ymf101* (= *orf64*)). *P. sansomeana* mitochondria also included a unique ORF predicted to encode a protein of 402 aa (*orf402*) (Fig 1). No detectable similarity was found between *orf402* and any known genes in GenBank. All protein coding genes and ORFs had the TAA stop codon except the *nad11* gene, which had the TGA stop codon.

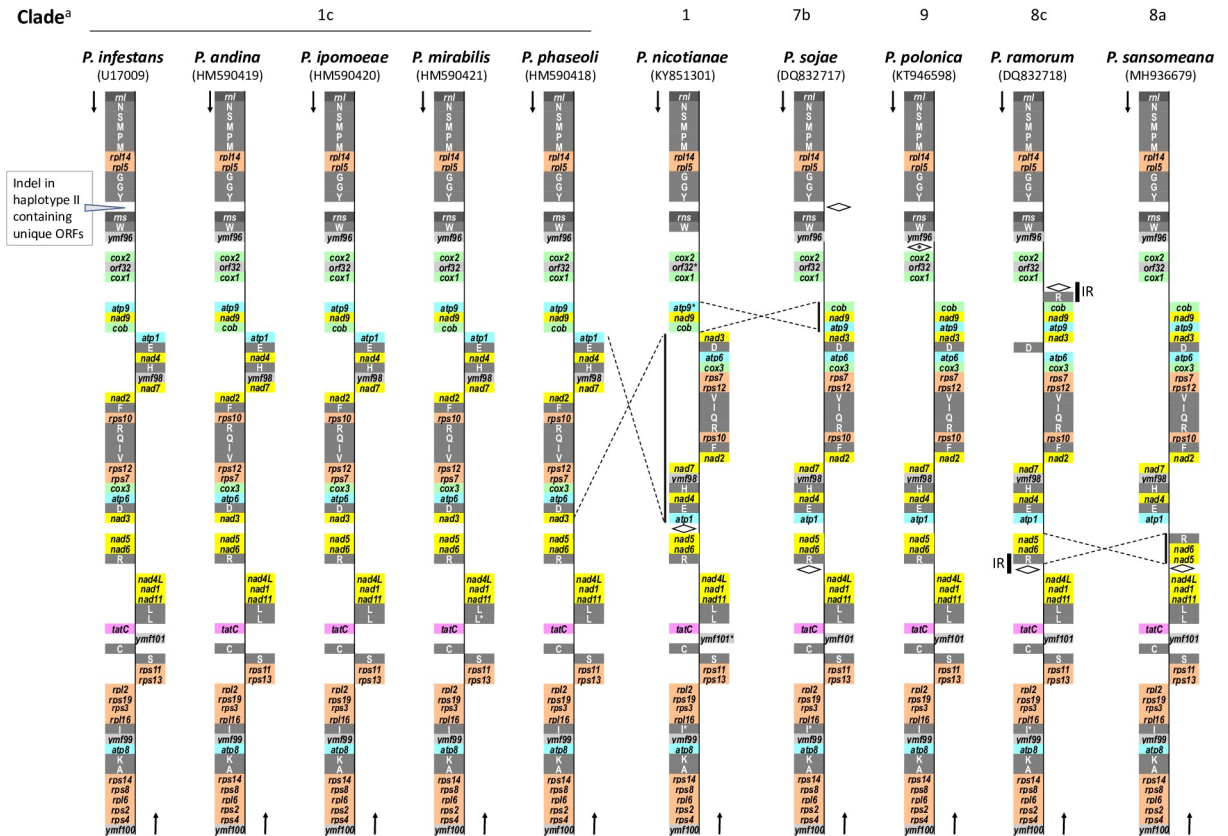
There was one large subunit rRNA gene (*rnl*) and one small subunit rRNA gene (*rns*). The 25 tRNA genes encoded tRNAs for 19 amino acids. There were two tRNA genes for arginine, glycine, leucine, methionine and serine, and one each for the other 14 amino acids except threonine (Fig 1). There was a third tRNA gene located between *rpl16* and *ymf99* (Fig 1) that had a CAU anticodon, but we interpreted it as *trnI*<sub>CAU</sub> rather than *trnM*<sub>CAU</sub> based on its similarity to a homologous tRNA gene at the same location in *P. infestans*. Experimentally confirmed in *Escherichia coli* [39] and inferred in *P. infestans* [18], it was assumed that the cytosine in *trnI*<sub>CAU</sub> anticodon was post-transcriptionally modified to lysidine that would enable it to recognize the AUA codon for isoleucine.

The genome was compact. Coding regions approximately comprised 90.1% of the genome. There was no intron in any of the genes. Of the 69 intergenic regions, 53 were 50 bp or less in length and 41 were less than 20 bp. There were three overlaps: between *rps7* and *rps12* (26 bp), between *nad1* and *nad11* (4bp), and between *tatC* and *ymf101* (4bp).

## Protein coding gene contents in the mitochondrial genomes of 10 *Phytophthora* species

Mitochondrial genomes of nine other *Phytophthora* species were publicly available and they were downloaded from GenBank (Table 1). Throughout the analysis below, please refer to Table 1 for reference and GenBank accessions. These 10 genomes are similar in size, ranging from 37,561 bp in *P. nicotianae* to 42,977 bp in *P. sojae*. G+C contents range from 21.57% in *P. polonica* to 22.39% in *P. ipomoeae*. Based on seven nuclear loci and four mitochondrial loci,





**Fig 2. Linear representation of gene contents and organization of mitochondrial genomes of 10 *Phytophthora* species.** RNA genes are in white letters with rRNA genes in dark grey background and tRNA genes in grey background. Protein coding genes are in dark letters with the following backgrounds: cytochrome C reductase and oxidases, green; ATP synthases, blue; NADH dehydrogenases, yellow; ribosomal protein genes, brown; twin-arginine translocase, purple; and conserved ORFs with unknown function, light grey. The diamond symbols represent regions with one or more ORFs unique to that particular species. In species other than *P. sansomeana*, protein coding genes with an asterisk are those identified in this study not presented in GenBank annotations. tRNA genes with an asterisk are those whose annotations are modified in this study, either in coding strand or in identity, comparing to GenBank annotations. *tatC* gene was previously given the gene symbol *ymf16* (*ymf1* is used to designate conserved ORF in mitochondria without annotated function) or *SecY* in species other than *P. sansomeana*. Our analysis showed that it should be named *tatC* (see text for detail). IR, inverted repeat. <sup>a</sup> Clade designation is based on [11, 12].

<https://doi.org/10.1371/journal.pone.0231296.g002>

this was caused by a deletion from the 3' end of the *atp9* gene that extended into the intergenic region between *atp9* gene and *nad9* gene. This deletion removed the last 13 amino acids in *atp9* protein and replaced it with five non-homologous amino acids (Fig 3).

Species	Location	Predicted protein sequence
<i>P. infestans</i>	la 10230..10457	MLLQASKFLGAGLATLGLIGAGIGIGNVFGSLIIGISRNPSLQQELMRTAILGFALTEAIALFCLMMAFLILFAF
	lb 10266..10493	MLLQASKFLGAGLATLGLIGAGIGIGNVFGSLIIGISRNPSLQQELMRTAILGFALTEAIALFCLMMAFLILFAF
	lla 12177..12404	MLLQASKFLGAGLATLGLIGAGIGIGNVFGSLIIGISRNPSLQQELMRTAILGFALTEAIALFCLMMAFLILFAF
	llb 12148..12375	MLLQASKFLGAGLATLGLIGAGIGIGNVFGSLIIGISRNPSLQQELMRTAILGFALTEAIALFCLMMAFLILFAF
<i>P. andina</i>	la 10222..10449	MLLQASKFLGAGLATLGLIGAGIGIGNVFGSLITGISRNPSLQQELMRTAILGFALTEAIALFCLMMAFLILFAF
lc 10215..10442	MLLQASKFLGAGLATLGLIGAGIGIGNVFGSLIIGISRNPSLQQELMRTAILGFALTEAIALFCLMMAFLILFAF	
<i>P. ipomoeae</i>	10230..10457	MLLQASKFLGAGLATLGLIGAGIGIGNVFGSLIIGISRNPSLQQELMRTAILGFALTEAIALFCLMMAFLILFAF
<i>P. mirabilis</i>	10198..10425	MLLQASKFLGAGLATLGLIGAGIGIGNVFGSLIIGISRNPSLQQELMRTAILGFALTEAIALFCLMMAFLILFAF
<i>P. phaseoli</i>	10248..10475	MLLQASKFLGAGLATLGLIGAGIGIGNVFGSLIIGISRNPSLQQELMRTAILGFALTEAIALFCLMMAFLILFAF
<i>P. nicotianae</i>	9379..9582	MLLQASKFLGAGLATLGLIGAGIGIGNVFGSLIIGISRNPSLQQELMRTAILGFALTEAIALYQIQI
<i>P. sojae</i>	13409..13182	MLLQASKFLGAGLATLGLIGAGIGIGNVFGSLIIGISRNPSLQQELMRTAILGFALTEAIALFCLMMAFLILFAF
<i>P. polonica</i>	7376..7149	MLLQASKFLGAGLATLGLAGAGIGIGNVFGSLVIGISRNPSLQQELMRTAILGFALTEAIALFCLMMAFLILFAF
<i>P. ramorum</i>	12818..12591	MLLQASKFLGAGLATLGLIGAGIGIGNVFGSLIIGISRNPSLQQELMRTAILGFALTEAIALFCLMMAFLILFAF
<i>P. sansomeana</i>	12823..12596	MLLQASKFLGAGLATLGLIGAGIGIGNVFGSLIIGISRNPSLQQELMRTAILGFALTEAIALFCLMMAFLILFAF

**Fig 3. Deduced protein sequences of *atp9* gene in the mitochondrial genomes of 10 *Phytophthora* species.** Amino acids differ from the consensus are shaded in grey. A deletion at the 3' end of this gene in *P. nicotianae* results in the replacement of the last 13 amino acids with five non-homologous ones (boxed in red).

<https://doi.org/10.1371/journal.pone.0231296.g003>

In addition to the 34 genes described above, another gene shared by these mitochondrial genomes was annotated with function, but its annotation in these genomes was inconsistent. *yfm16* (= *orf244*) is a conserved ORF in all 10 *Phytophthora* species that encodes a protein of approximately 244 amino acids. In the 13 mitochondrial genomes of nine *Phytophthora* species available in GenBank (Table 1), this gene was given the name *yfm16* in most cases and its product was annotated as SecY-independent transporter protein. However, in *P. ramorum* and *P. sojae*, this gene was named *SecY*, and in *P. nicotianae*, its product was named SecY. The name "SecY" is not compatible with the meaning of "SecY-independent".

There are three translocation systems in bacteria that transport distinct subset of proteins into or cross the inner membrane. SecYEG is the main translocase that transports unfolded peptides [40, 41], of which SecY is the main transmembrane subunit. YidC functions as an insertase and also assists SecYEG in the assembly of membrane proteins [42]. The twin-arginine translocation (Tat) system is capable of translocating folded or even multi-subunit protein complexes across inner membrane [43]. Proteins translocated by the Tat system have a twin-arginine motif in their signal peptides. The TatC protein is an essential component of the Tat system.

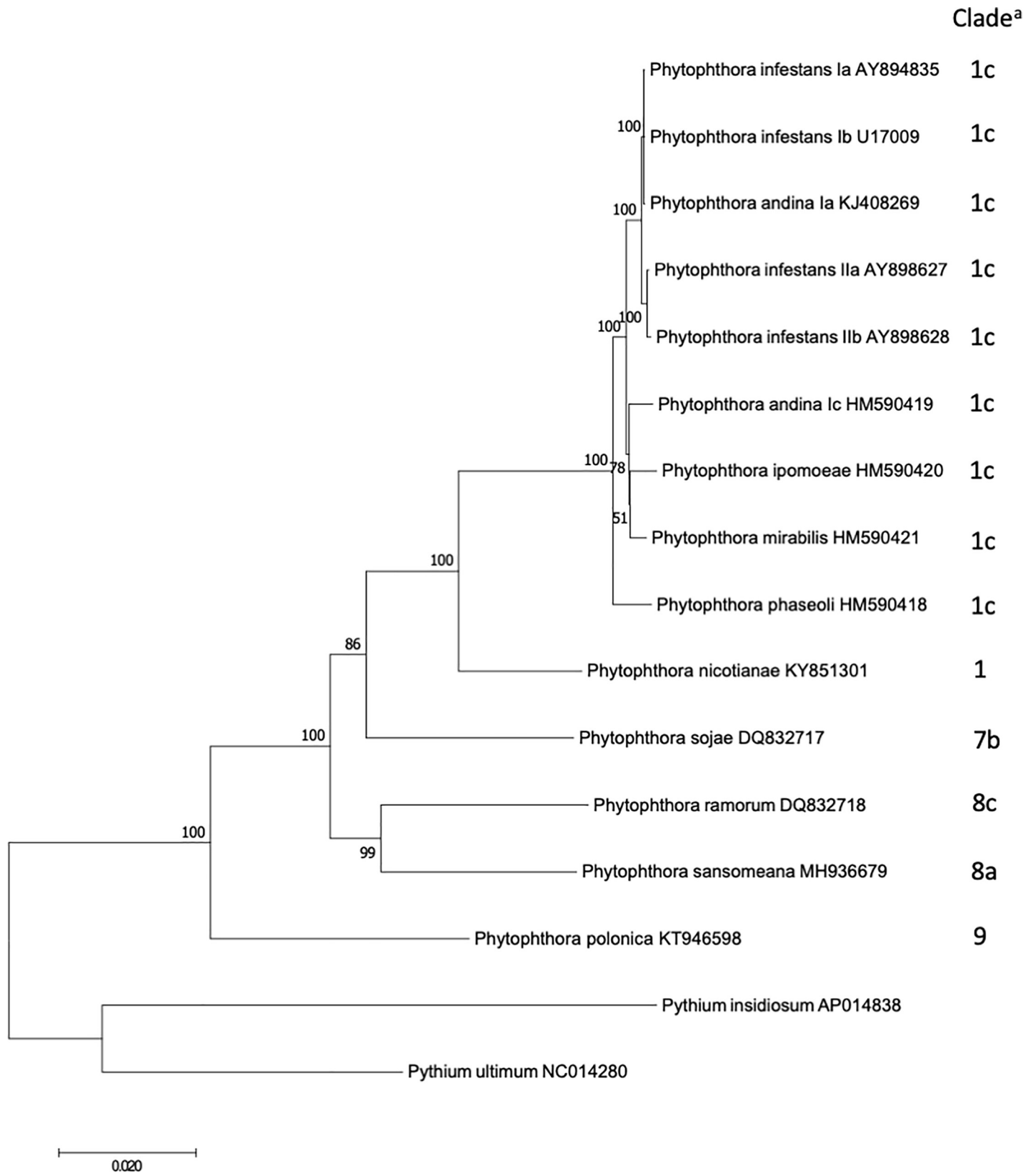
In a previous study, *TatC* gene was identified in the mitochondrial genomes of oomycetes [44]. To determine the function of *yfm16* gene, its predicted protein sequences from the 10 *Phytophthora* species were used to search Pfam database version 32 [45]. The only hit was TatC (sec-independent translocase protein, E values < 1e-20). As such, we annotated this gene as *tatC* (Fig 2).

Six conserved ORFs have been reported in the mitochondrial genomes of these 10 species. Of these, *orf32*, *yfm98*, *yfm99* and *yfm100* were found in all 10 species. *orf32* was not reported in *P. nicotianae* but our analysis found this ORF in the expected location (nucleotide 7530..7628) (Fig 2). *yfm96* was not found in *P. nicotianae*. Of the five species in subclade 1c, *yfm101*, a small ORF encoding approximately 64 aa, was found in *P. infestans* but missing in three species, *P. ipomoeae*, *P. mirabilis* and *P. phaseoli* (Fig 2). Interestingly, In *P. andina*, *yfm101* was missing in haplotype 1c but our analysis identified it in haplotype 1a (nucleotide 30270..30476). *yfm101* was not reported in *P. nicotianae*, but our analysis identified it at nucleotide 30206..30015.

Unique ORFs have been reported in several species. In *P. infestans* haplotypes IIa and IIb, an insertion of approximately 2 kb encodes multiple ORFs [19]. Also reported were *orf183* in *P. nicotianae*, two inverted copies of *orf176* in *P. ramorum*, and six ORFs at two location in *P. sojae*. Additionally, we identified *orf402* in *P. sansomeana* and also found an ORF with 173 codons (excluding the stop codon) in *P. polonica* at nucleotide 540..1061 (Fig 2).

## RNA gene contents in the mitochondrial genomes of 10 *Phytophthora* species

Each genome had two rRNA genes for the large and small subunit ribosomal RNAs (*rnl* and *rns*, respectively). All the genomes encoded 25 tRNA genes and these genes had the same anticodons (S1 Table). The only exception was *P. ramorum*, in which an inverted repeat of 1,150bp resulted in a second copy of *trnR<sub>UCU</sub>* and *orf176* (Fig 2 and S1 Table). None of the mitochondrial genomes encoded tRNA for threonine. A tRNA gene for threonine was not found in the mitochondrial genomes of other oomycetes including *Pythium insidiosum* [38], *Pythium ultimum* [37], *Peronospora effuse* [46], *Pseudoperonospora humuli* [47], *Pseudoperonospora cubensis* [48], *Saprolegnia ferax* [49], *Achlya hypogyna* and *Thraustotheca clavate* [50]. Two stramenopiles, *Thalassiosira pseudonana* [51] and *Phaeodactylum tricorutum* [52], also lacked tRNA genes for threonine in their mitochondrial genomes, suggesting that this is a shared trait of mitochondria in stramenopiles.



**Fig 4. A maximum-likelihood tree based on the mitochondrial genomes.** Branch support over 50% out of 1,000 bootstraps are shown. <sup>a</sup> Clade designation is based on [11, 12].

<https://doi.org/10.1371/journal.pone.0231296.g004>



## Organization of the mitochondrial genomes of 10 *Phytophthora* species

As discussed above, the mitochondrial genomes in these 10 species were conserved in gene content with few exceptions. These genomes were also conserved in gene order and coding strand except for three inversions. One inversion event occurred within clade 1. *P. nicotianae* belongs to clade 1 and is basal to subclades 1b and 1c. Compared to *P. nicotianae* and species in clades 7, 8 and 9, a large section between *atp1* gene and *nad3* gene was inverted in the five species in subclade 1c: *P. andina*, *P. infestans*, *P. ipomoeae*, *P. mirabilis* and *P. phaseoli*. This section contained 11 protein coding genes and eight tRNA genes. A second inversion occurred between species in clade 1 and species in clades 7, 8 and 9. This inversion covered three protein coding genes: *atp9*, *nad9* and *cob*. A third inversion occurred within clade 8. In *P. sansomeana* (subclade 8a), the section covering *nad5*, *nad6* and *trnR<sub>UCU</sub>* was inverted comparing to *P. ramorum* (subclade 8c) and species in other clades (Fig 2).

## Phylogeny of *Phytophthora* species based on mitochondrial genomes

Other than *atp9*, 34 protein coding genes with annotated functions and the two rRNA genes, *rnl* and *rns*, were used to construct the phylogenetic tree (Fig 4). Previously, *Phytophthora* species were grouped into 10 clades [11, 12]. Our analysis included species from four clades and the result was in agreement with previous studies. Species from individual clades were grouped together with bootstrap support 86% or higher. Within clade 1, previous studies showed moderate support that *P. nicotianae* was basal to subclades 1b and 1c based on seven nuclear genes [11] and four mitochondrial genes [12]. Our analysis supported the conclusion that *P. nicotianae* was basal to subclade 1c (100% bootstrap support). Species in subclade 1b was not included in this study. Within clade 1c, *P. phaseoli* diverged first and was basal to other species. *P. andina* haplotype Ia was grouped with haplotype I of *P. infestans*, while haplotype Ic was grouped with *P. ipomoeae* and *P. mirabilis*. These findings were in agreement with previous reports [20, 53]. The phylogenetic analysis was dominated by clade 1 species. Decreasing the number of species in clade 1 did not change the topology of the tree (not shown).

In summary, we sequenced and annotated the complete mitochondrial genome of the soybean pathogen *P. sansomeana* and compared it to that of nine other *Phytophthora* species. Inconsistencies in annotation among these mitochondrial genomes were corrected. These genomes were found to be conserved in gene content with few exceptions. Three inversion events, each covering multiple genes, were observed among these genomes. Phylogenomic analysis using the mitochondrial genomes supported or enhanced previous findings.

Mention of trade names or commercial products in this publication is solely for the purpose of providing specific information and does not imply recommendation or endorsement by the U.S. Department of Agriculture. USDA is an equal opportunity provider and employer.

## Supporting information

**S1 Table. tRNA genes in the mitochondrial genomes of 10 *Phytophthora* species.**  
(XLSX)

## Author Contributions

**Conceptualization:** Guohong Cai.

**Data curation:** Guohong Cai.

**Formal analysis:** Guohong Cai.

**Funding acquisition:** Guohong Cai.

**Investigation:** Guohong Cai.

**Methodology:** Guohong Cai.

**Writing – original draft:** Guohong Cai.

**Writing – review & editing:** Guohong Cai, Steven R. Scofield.

## References

1. Erwin DC, Ribeiro OK. *Phytophthora* diseases worldwide. St. Paul, Minnesota: American Phytopathological Society (APS Press); 1996. xii + 562 p.
2. Bourke A. 'The visitation of god?' The potato and the great Irish famine. Dublin: Lilliput Press Ltd; 1993. x + 230 p.
3. Braiser CM. *Phytophthora* biodiversity: how many *Phytophthora* species are there? In: Goheen EM, Frankel S.J., editor. *Phytophthoras in Forests and Natural Ecosystems*. Albany, CA, USA: USDA Forest Service: General Technical Report PSW-GTR-221; 2007. p. 101–15.
4. Martin FN, Abad ZG, Balci Y, Ivors K. Identification and detection of *Phytophthora*: reviewing our progress, identifying our needs. *Plant Dis*. 2012; 96(8):1080–103. <https://doi.org/10.1094/PDIS-12-11-1036-FE> PMID: 30727075
5. Park J, Park B, Veeraraghavan N, Jung K, Lee Y-H, Blair JE, et al. *Phytophthora* Database: a forensic database supporting the identification and monitoring of *Phytophthora*. *Plant Dis*. 2008; 92(6):966–72. <https://doi.org/10.1094/PDIS-92-6-0966> PMID: 30769728
6. Yang X, Tyler BM, Hong C. An expanded phylogeny for the genus *Phytophthora*. *IMA Fungus*. 2017; 8(2):355–84. Epub 2017/11/21. <https://doi.org/10.5598/ima fungus.2017.08.02.09> PMID: 29242780.
7. IDphy: Molecular and morphological identification of *Phytophthora* based on the types [Internet]. 2019. Available from: <https://idtools.org/id/phytophthora/>.
8. Baldauf SL, Roger AJ, Wenk-Siefert I, Doolittle WF. A kingdom-level phylogeny of eukaryotes based on combined protein data. *Science*. 2000; 290(5493):972–7. Epub 2000/11/04. <https://doi.org/10.1126/science.290.5493.972> PMID: 11062127.
9. Cavalier-Smith T. Membrane heredity and early chloroplast evolution. *Trends Plant Sci*. 2000; 5(4):174–82. Epub 2000/03/31. [https://doi.org/10.1016/s1360-1385\(00\)01598-3](https://doi.org/10.1016/s1360-1385(00)01598-3) PMID: 10740299.
10. Sogin ML, Silberman JD. Evolution of the protists and protistan parasites from the perspective of molecular systematics. *Int J Parasitol*. 1998; 28(1):11–20. [https://doi.org/10.1016/s0020-7519\(97\)00181-1](https://doi.org/10.1016/s0020-7519(97)00181-1) PMID: 9504331
11. Blair JE, Coffey MD, Park S-Y, Geiser DM, Kang S. A multi-locus phylogeny for *Phytophthora* utilizing markers derived from complete genome sequences. *Fungal Genetics and Biology*. 2008; 45(3):266–77. <https://doi.org/10.1016/j.fgb.2007.10.010> PMID: 18039586
12. Martin FN, Blair JE, Coffey MD. A combined mitochondrial and nuclear multilocus phylogeny of the genus *Phytophthora*. *Fungal Genetics and Biology*. 2014; 66:19–32. <https://doi.org/10.1016/j.fgb.2014.02.006> PMID: 24603058
13. Hansen EM, Wilcox WF, Reeser PW, Sutton W. *Phytophthora rosacearum* and *P. sansomeana*, new species segregated from the *Phytophthora megasperma* “complex”. *Mycologia*. 2009; 101(1):129–35. <https://doi.org/10.3852/07-203> PMID: 19271675
14. Reeser PW, Scott DH, Ruhl DE. Recovery of race nonclassifiable *Phytophthora megasperma* f. sp. *glycinea* from soybean roots in Indiana in 1990. *Phytopathology*. 1991; 81:1201.
15. Tang QH, Gao F, Li GY, Wang H, Zheng XB, Wang YC. First report of root rot caused by *Phytophthora sansomeana* on soybean in China. *Plant Dis*. 2010; 94(3):378. Epub 2010/03/01. <https://doi.org/10.1094/pdis-94-3-0378a> PMID: 30754211.
16. Alejandro Rojas J, Jacobs JL, Napieralski S, Karaj B, Bradley CA, Chase T, et al. Oomycete species associated with soybean seedlings in North America—Part I: Identification and pathogenicity characterization. *Phytopathology*. 2016; 107(3):280–92. <https://doi.org/10.1094/PHTO-04-16-0177-R> PMID: 27801078
17. Bienapfl JC, Malvick DK, Percich JA. Specific molecular detection of *Phytophthora sojae* using conventional and real-time PCR. *Fungal Biol*. 2011; 115(8):733–40. Epub 2011/08/02. <https://doi.org/10.1016/j.funbio.2011.05.007> PMID: 21802053.

18. 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. *Curr Genet.* 1997; 31(5):380–95. Epub 1997/05/01. <https://doi.org/10.1007/s002940050220> PMID: 9162109.
19. Avila-Adame C, Gómez-Alpizar L, Zismann V, Jones KM, Buell CR, Ristaino JB. Mitochondrial genome sequences and molecular evolution of the Irish potato famine pathogen, *Phytophthora infestans*. *Curr Genet.* 2006; 49(1):39–46. <https://doi.org/10.1007/s00294-005-0016-3> PMID: 16328503
20. Lassiter ES, Russ C, Nusbaum C, Zeng Q, Saville AC, Olarte RA, et al. Mitochondrial genome sequences reveal evolutionary relationships of the *Phytophthora* 1c clade species. *Curr Genet.* 2015; 61(4):567–77. <https://doi.org/10.1007/s00294-015-0480-3> PMID: 25754775
21. Yuan X, Feng C, Zhang Z, Zhang C. Complete mitochondrial genome of *Phytophthora nicotianae* and identification of molecular markers for the oomycetes. *Frontiers in Microbiology.* 2017; 8(1484). <https://doi.org/10.3389/fmicb.2017.01484> PMID: 28848506
22. Martin FN, Bensasson D, Tyler BM, Boore JL. Mitochondrial genome sequences and comparative genomics of *Phytophthora ramorum* and *P. sojae*. *Curr Genet.* 2007; 51(5):285–96. <https://doi.org/10.1007/s00294-007-0121-6> PMID: 17310332
23. May KJ, Ristaino JB. Identity of the mtDNA haplotype(s) of *Phytophthora infestans* in historical specimens from the Irish Potato Famine. *Mycol Res.* 2004; 108(5):471–9. <https://doi.org/10.1017/S0953756204009876>.
24. Griffith GW, Shaw DS. Polymorphisms in *Phytophthora infestans*: Four mitochondrial haplotypes are detected after PCR amplification of DNA from pure cultures or from host lesions. *Applied and Environmental Microbiology.* 1998; 64(10):4007. PMID: 9758833
25. Carter DA, Archer SA, Buck KW, Shaw DS, Shattock RC. Restriction fragment length polymorphisms of mitochondrial DNA of *Phytophthora infestans*. *Mycol Res.* 1990; 94(8):1123–8. [https://doi.org/10.1016/S0953-7562\(09\)81344-0](https://doi.org/10.1016/S0953-7562(09)81344-0).
26. Martin MD, Ho SYW, Wales N, Ristaino JB, Gilbert MTP. Persistence of the mitochondrial lineage responsible for the Irish potato famine in extant New World *Phytophthora infestans*. *Molecular Biology and Evolution.* 2014; 31(6):1414–20. <https://doi.org/10.1093/molbev/msu086> PMID: 24577840
27. Chin C-S, Alexander DH, Marks P, Klammer AA, Drake J, Heiner C, et al. Nonhybrid, finished microbial genome assemblies from long-read SMRT sequencing data. *Nat Methods.* 2013; 10:563. <https://doi.org/10.1038/nmeth.2474> <https://www.nature.com/articles/nmeth.2474#supplementary-information>. PMID: 23644548
28. Langmead B, Salzberg SL. Fast gapped-read alignment with Bowtie 2. *Nat Meth.* 2012; 9(4):357–9. <https://doi.org/10.1038/nmeth.1923> <http://www.nature.com/nmeth/journal/v9/n4/abs/nmeth.1923.html#supplementary-information>. PMID: 22388286
29. Walker BJ, Abeel T, Shea T, Priest M, Abouelliel A, Sakthikumar S, et al. Pilon: An integrated tool for comprehensive microbial variant detection and genome assembly improvement. *PLoS One.* 2014; 9(11):e112963. <https://doi.org/10.1371/journal.pone.0112963> PMID: 25409509
30. Karlovsky P, Fartmann B. Genetic code and phylogenetic origin of oomycetous mitochondria. *J Mol Evol.* 1992; 34(3):254–8. <https://doi.org/10.1007/BF00162974> PMID: 1588598
31. Quast C, Pruesse E, Yilmaz P, Gerken J, Schweer T, Yarza P, et al. The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. *Nucleic Acids Res.* 2013; 41(Database issue):D590–6. Epub 2012/11/30. <https://doi.org/10.1093/nar/gks1219> PMID: 23193283; PubMed Central PMCID: PMC3531112.
32. Lowe TM, Chan PP. tRNAscan-SE On-line: integrating search and context for analysis of transfer RNA genes. *Nucleic Acids Res.* 2016; 44(W1):W54–7. Epub 2016/05/14. <https://doi.org/10.1093/nar/gkw413> PMID: 27174935; PubMed Central PMCID: PMC4987944.
33. Lohse M, Drechsel O, Kahlau S, Bock R. OrganellarGenomeDRAW—a suite of tools for generating physical maps of plastid and mitochondrial genomes and visualizing expression data sets. *Nucleic Acids Res.* 2013; 41(Web Server issue):W575–W81. <https://doi.org/10.1093/nar/gkt289> PMC3692101. PMID: 23609545
34. Edgar RC. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.* 2004; 32(5):1792–7. <https://doi.org/10.1093/nar/gkh340> PMID: 15034147.
35. Tamura K, Nei M. Estimation of the number of nucleotide substitutions in the control region of mitochondrial DNA in humans and chimpanzees. *Mol Biol Evol.* 1993; 10(3):512–26. Epub 1993/05/01. <https://doi.org/10.1093/oxfordjournals.molbev.a040023> PMID: 8336541.
36. Kumar S, Stecher G, Tamura K. MEGA7: Molecular evolutionary genetics analysis version 7.0 for bigger datasets. *Mol Biol Evol.* 2016; 33(7):1870–4. Epub 2016/03/24. <https://doi.org/10.1093/molbev/msw054> PMID: 27004904.

37. Levesque CA, Brouwer H, Cano L, Hamilton J, Holt C, Huitema E, et al. Genome sequence of the necrotrophic plant pathogen *Pythium ultimum* reveals original pathogenicity mechanisms and effector repertoire. *Genome Biol.* 2010; 11(7):R73. <https://doi.org/10.1186/gb-2010-11-7-r73> PMID: 20626842
38. Tangphatsornruang S, Ruang-areerate P, Sangsrakru D, Rujirawat T, Lohnoo T, Kittichotirat W, et al. Comparative mitochondrial genome analysis of *Pythium insidiosum* and related oomycete species provides new insights into genetic variation and phylogenetic relationships. *Gene.* 2016; 575(1):34–41. <https://doi.org/10.1016/j.gene.2015.08.036> PMID: 26299654
39. Muramatsu T, Nishikawa K, Nemoto F, Kuchino Y, Nishimura S, Miyazawa T, et al. Codon and amino acid specificities of a transfer RNA are both converted by a single post-transcriptional modification. *Nature.* 1988; 336(6195):179–81. Epub 1988/11/10. <https://doi.org/10.1038/336179a0> PMID: 3054566
40. Powers T, Walter P. Co-translational protein targeting catalyzed by the *Escherichia coli* signal recognition particle and its receptor. *EMBO J.* 1997; 16(16):4880–6. Epub 1997/08/15. <https://doi.org/10.1093/emboj/16.16.4880> PMID: 9305630; PubMed Central PMCID: PMC1170123.
41. Hartl FU, Lecker S, Schiebel E, Hendrick JP, Wickner W. The binding cascade of SecB to SecA to SecY/E mediates preprotein targeting to the *E. coli* plasma membrane. *Cell.* 1990; 63(2):269–79. Epub 1990/10/19. [https://doi.org/10.1016/0092-8674\(90\)90160-g](https://doi.org/10.1016/0092-8674(90)90160-g) PMID: 2170023.
42. Wang P, Dalbey RE. Inserting membrane proteins: the YidC/Oxa1/Alb3 machinery in bacteria, mitochondria, and chloroplasts. *Biochim Biophys Acta.* 2011; 1808(3):866–75. Epub 2010/08/31. <https://doi.org/10.1016/j.bbamem.2010.08.014> PMID: 20800571.
43. Palmer T, Berks BC. The twin-arginine translocation (Tat) protein export pathway. *Nat Rev Microbiol.* 2012; 10(7):483–96. Epub 2012/06/12. <https://doi.org/10.1038/nrmicro2814> PMID: 22683878.
44. Petru M, Wideman J, Moore K, Alcock F, Palmer T, Dolezal P. Evolution of mitochondrial TAT translocases illustrates the loss of bacterial protein transport machines in mitochondria. *BMC Biol.* 2018; 16(1):141–. <https://doi.org/10.1186/s12915-018-0607-3> PMID: 30466434.
45. El-Gebali S, Mistry J, Bateman A, Eddy SR, Luciani A, Potter SC, et al. The Pfam protein families database in 2019. *Nucleic Acids Res.* 2018; 47(D1):D427–D32. <https://doi.org/10.1093/nar/gky995> PMID: 30357350
46. Fletcher K, Klosterman SJ, Derevnina L, Martin F, Bertier LD, Koike S, et al. Comparative genomics of downy mildews reveals potential adaptations to biotrophy. *BMC Genomics.* 2018; 19(1):851. Epub 2018/11/30. <https://doi.org/10.1186/s12864-018-5214-8> PMID: 30486780; PubMed Central PMCID: PMC6264045.
47. Rahman A, Gongora-Castillo E, Bowman MJ, Childs KL, Gent DH, Martin FN, et al. Genome sequencing and transcriptome analysis of the hop downy mildew pathogen *Pseudoperonospora humuli* reveal species-specific genes for molecular detection. *Phytopathology.* 2019; 109(8):1354–66. Epub 2019/04/03. <https://doi.org/10.1094/PHYTO-11-18-0431-R> PMID: 30939079.
48. Lu WJ, Hu WG, Wang GP. Complete mitochondrial genome of *Pseudoperonospora cubensis*. *Mitochondrial DNA Part A.* 2016; 27(5):3487–8. Epub 2015/07/18. <https://doi.org/10.3109/19401736.2015.1066361> PMID: 26186306.
49. Grayburn WS, Hudspeth DS, Gane MK, Hudspeth ME. The mitochondrial genome of *Saprolegnia ferax*: organization, gene content and nucleotide sequence. *Mycologia.* 2004; 96(5):981–9. Epub 2004/09/01. PMID: 21148919.
50. O'Brien MA, Misner I, Lane CE. Mitochondrial genome sequences and comparative genomics of *Achlya hypogyna* and *Thraustotheca clavata*. *J Eukaryot Microbiol.* 2014; 61(2):146–54. Epub 2013/11/21. <https://doi.org/10.1111/jeu.12092> PMID: 24252096.
51. Armbrust EV, Berges JA, Bowler C, Green BR, Martinez D, Putnam NH, et al. The genome of the diatom *Thalassiosira pseudonana*: ecology, evolution, and metabolism. *Science.* 2004; 306(5693):79–86. Epub 2004/10/02. <https://doi.org/10.1126/science.1101156> PMID: 15459382.
52. Oudot-Le Secq MP, Green BR. Complex repeat structures and novel features in the mitochondrial genomes of the diatoms *Phaeodactylum tricornutum* and *Thalassiosira pseudonana*. *Gene.* 2011; 476(1–2):20–6. Epub 2011/02/16. <https://doi.org/10.1016/j.gene.2011.02.001> PMID: 21320580.
53. Blair JE, Coffey MD, Martin FN. Species tree estimation for the late blight pathogen, *Phytophthora infestans*, and close relatives. *PLoS One.* 2012; 7(5):e37003. <https://doi.org/10.1371/journal.pone.0037003> PMID: 22615869