Proposal of a new nomenclature for introns in protein-coding genes in fungal mitogenomes

Shu Zhang¹ and Yong-Jie Zhang²*

Abstract

Fungal mitochondrial genes are often invaded by group I or II introns, which represent an ideal marker for understanding fungal evolution. A standard nomenclature of mitochondrial introns is needed to avoid confusion when comparing different fungal mitogenomes. Currently, there has been a standard nomenclature for introns present in rRNA genes, but there is a lack of a standard nomenclature for introns present in protein-coding genes. In this study, we propose a new nomenclature system for introns in fungal mitochondrial protein-coding genes based on (1) three-letter abbreviation of host scientific name, (2) host gene name, (3), one capital letter P (for group I introns), S (for group II introns), or U (for introns with unknown types), and (4) intron insertion site in the host gene according to the cyclosporin-producing fungus Tolypocladium inflatum. The suggested nomenclature was proved feasible by naming introns present in mitogenomes of 16 fungi of different phyla, including both basal and higher fungal lineages although minor adjustment of the nomenclature is needed to fit certain special conditions. The nomenclature also had the potential to name plant/protist/animal mitochondrial introns. We hope future studies follow the proposed nomenclature to ensure direct comparison across different studies.

Keywords: Nomenclature, Fungi, Mitogenome, Intron, Protein-coding gene

INTRODUCTION

Fungi constitute a huge group of highly diverse organisms, with 2.2–3.8 million estimated species and 144,000 currently known species on Earth (Hawksworth and Lücking 2017; Cannon et al. 2018). They were traditionally divided into four groups: chytridiomycetes, zygomycetes, ascomycetes, and basidiomycetes according to morphological traits associated with reproduction. Molecular phylogenetics and more recently phylogenomics recognized eight phyla in Fungi, namely Microsporidia, Cryptomycota, Blastocladiomycota, Chytridiomycota, Zoopagomycota, Mucoromycota, Ascomycota, and Basidiomycota (Spatafora et al. 2017). Aside from a few early divergent lineages and anaerobic organisms, almost all fungi contain mitochondria and mitogenomes in their cells (Bullerwell and Lang 2005; van der Giezen et al. 2005). Over recent years, mitogenomes of an increasing number of fungal species are sequenced. As of July 2019, mitogenomes from at least 300 fungal species are available with representatives from all major fungal groups. Fungal mitogenomes typically contain 15 standard protein-coding genes, two rRNA genes and a variable number of tRNA genes. These protein-coding genes are \textit{atp6}, \textit{atp8}, \textit{atp9}, \textit{cob}, \textit{cox1}, \textit{cox2}, \textit{cox3}, \textit{nad1}, \textit{nad2}, \textit{nad3}, \textit{nad4}, \textit{nad4L}, \textit{nad5}, \textit{nad6}, and \textit{rps3} (Lang 2018), and some of them may be absent from certain fungal mitogenomes (Koszul et al. 2003).

Introns as mobile elements are frequently observed in mitochondrial protein-coding and/or rRNA genes of fungi. One gene may also be simultaneously invaded by multiple introns (e.g., four introns in \textit{cob} and seven introns in \textit{cox1} in \textit{Isaria cicadae}) (Fan et al. 2019). Mitochondrial introns are divided into two groups (I and II) based on their secondary structure and splicing mechanism (Saldanha et al. 1993), with group I introns being abundant in fungal mitogenomes. Different fungal species or even different individuals of a particular fungus may show diversity in number and insertion position of mitochondrial introns (Kosa et al. 2006; Zhang et al. 2015; Zhang et al. 2017a;
Wang et al. 2018; Fan et al. 2019; Nie et al. 2019). Introns contribute to fungal mitogenome expansion/variability and represent an ideal marker for understanding fungal evolution (Zhang et al. 2015).

Currently, there has been a nomenclature for introns present in rRNA genes (Johansen and Haugen 2001). According to the nomenclature, introns are often found at a limited number of insertion sites in highly conserved regions of rRNA genes from nuclei, mitochondria, and chloroplasts, and therefore, a given rRNA sequence can be aligned with the chosen standard rRNA sequences of Escherichia coli to locate and name potential introns. For mitochondrial protein-coding genes, however, it is difficult to align their sequences with corresponding E. coli sequences due to high sequence divergence. In most literatures, introns in protein-coding genes are generally named serially according to their appearance in a particular host gene (e.g., cox1-i1, cox1-i2, and cox1-i3) (Deng et al. 2016; Zhang et al. 2017b; Zhang et al. 2017c). This naming strategy is not convenient for scientific communication and comparison of introns across different mitogenomes. A standard nomenclature of mitochondrial introns is needed to avoid confusion when comparing different fungal mitogenomes.

In our previous studies, we have tried to designate introns based on their insertion positions, but a mitogenome is arbitrarily selected from species under investigation (Fan et al. 2019; Zhang et al. 2019). In this study, we aim to propose a standard nomenclature for introns in protein-coding genes in fungal mitogenomes and test its applicability using fungal species from a broad range of taxonomic classification. To know if the suggested nomenclature can apply to “cross-kingdom” mitochondrial introns, some plant/protist/animal introns are also examined.

METHODS

In order to establish a standard nomenclature for introns in protein-coding genes across the kingdom Fungi, it is necessary to find an appropriate reference mitogenome. By looking at fungal species with available mitogenomes, we choose the mitogenome of the cyclosporin-producing fungus Tolypocladium inflatum ARSEF 3280 (accession number NC_036382) as the reference mitogenome. The 25,328-bp mitogenome of T. inflatum contains all the 15 protein-coding genes typically found in fungal mitogenomes, and there is no intron in any of these protein-coding genes (Zhang et al. 2017d). We did not choose the best-understood model fungi: baker’s yeast Saccharomyces cerevisiae, the fission yeast Schizosaccharomyces pombe, the opportunistic fungal pathogen Candida albicans, the filamentous euascomycete Neurospora crassa, etc. This is because the yeasts Sa. cerevisiae and Sc. pombe both lack genes coding for NADH dehydrogenases in their mitogenomes (Foury et al. 1995), and C. albicans and N. crassa contain introns in many different protein-coding genes (Borkovich et al. 2004; Bartelli et al. 2013). We also did not choose the human mitochondrial genome, which was selected as the reference to name introns found in nad5 and cox1 in certain metazoans (Emblem et al. 2011). This is because the human mitogenome contains only 13 standard protein-coding genes without atp9 and rps3. The latter two genes are known to harbor introns in fungal mitogenomes.

Both basal and higher fungi may contain introns in their mitogenomes. We randomly selected representative species in each fungal phylum to locate and name possible introns (Table 1). Determination of the insertion position of an intron relies on alignment between sequences of its host gene and corresponding gene sequences of T. inflatum (Additional file 1). Although there are many sequence alignment programs available, we recommend using MAFFT (https://mafft.cbrc.jp/alignment/software/), which is fast when aligning long sequences containing many introns and can always generate satisfactory alignment according to our experience. The default setting of MAFFT works well in most cases. If exon-intron boundaries are not correctly identified (probably due to the interference of intron sequences or presence of short exons) under the default settings, one may consider adjusting the alignment parameters (e.g., try ‘Unalignlevel > 0’ and possibly ‘Leave gappy regions’ by selecting the G-INS-1 or G-INS-i alignment strategy) and/or importing additional sequences to align from a species closely related the test species. In addition, it is always advisable to refer to known annotation results and/or characteristic nucleotides at splice sites of group I/II introns (Cech 1988) to ensure correct alignment and identification of exon-intron boundaries.

RESULTS AND DISCUSSION

We propose a new nomenclature system for introns in fungal mitochondrial protein-coding genes based on (1) three-letter abbreviation of host scientific name, (2) host gene name, (3) one capital letter P (for group I introns, meaning position or primary for easy memorization), S (for group II introns, meaning site or secondary), or U (for introns with unknown types), and (4) intron insertion site in the host gene according to T. inflatum (Additional file 1). When there is no ambiguity (e.g., when just talking about introns in a particular species or in a particular host gene of a species), host scientific name and/or host gene name may be omitted. In any case, however, the letter P/S/U and insertion site of an intron should never be omitted. Using the nomenclature, previously reported introns could be renamed. Examples of renaming are the group II intron Sce.cox1S169 (former aI1) from Saccharomyces cerevisiae cox1 at site 169, and the group 1 intron Cgl.cox1P240 (former CgCox1.1) from Candida glabrata cox1 at position...
| Fungal taxa            | Accession | Length (nt) | Phylum               | Class             | Order             | Family             | Code |
|-----------------------|-----------|-------------|----------------------|-------------------|-------------------|---------------------|------|
| **Basal fungi**       |           |             |                      |                   |                   |                     |      |
| Rozella allomycis     | NC_021611 | 12,055      | Cryptomycota         |                   |                   |                     | 4    |
| Rhizopus oryzae       | NC_006836 | 54,178      | Mucoromycota         | Mucorales         | Rhizopodaceae     |                     | 1    |
| Conidiobolus heterosporus | MK049352 | 53,364      | Zoopagomycota        | Entomophthoromyctes | Entomophthorales | Ancylistaceae      | 4    |
| Alomyces macrogynus   | NC_001715 | 57,473      | Blastocladiomycota   | Blastocladiomyctes | Blastocladiales   | Blastocladiales    | 4    |
| Hyaloraphidium curvatum | NC_003048 | 29,593      | Chytridiomycota      | Monoblepharomyctes | Monoblepharidales |                     | 4    |
| **Higher fungi**      |           |             |                      |                   |                   |                     |      |
| Candida albicans      | NC_002653 | 40,420      | Ascomycota           | Saccharomyctes    | Saccharomyctales  | Debaryomyctales     | 4    |
| Grosmannia piceipera a | FJ717837  | 2928        | Ascomycota           | Sordariomyctes    | Ophiostomatales   | Ophiostomatales     | 4    |
| Isaria cicadae        | MH922223  | 56,581      | Ascomycota           | Sordariomyctes    | Hypocreales       | Cordycipitaceae     | 4    |
| Neurospora crassa     | NC_026614 | 64,840      | Ascomycota           | Sordariomyctes    | Sordariales       | Sordariales         | 4    |
| Saccharomyces cerevisiae | NC_001224 | 85,779      | Ascomycota           | Saccharomyctes    | Saccharomyctales  | Saccharomyctales    | 3    |
| Schizosaccharomyces pombe | NC_001326 | 19,431      | Ascomycota           | Schizosaccharomyctes | Schizosaccharomyctales | Schizosaccharomyctales | 4    |
| Tolypocladium inflatum | NC_036382 | 25,328      | Ascomycota           | Sordariomyctes    | Hypocreales       | Ophiocordycipitaceae | 4    |
| Cryptococcus neoformans | NC_004336 | 24,874      | Basidiomycota        | Tremellomyctes    | Tremellales       | Tremellales         | 4    |
| Puccinia striiformis  | NC_039655 | 101,521     | Basidiomycota        | Pucciniomyctes    | Pucciniiales      | Pucciniaceae       | 4    |
| Tilletia indica       | NC_009880 | 65,147      | Basidiomycota        | Exobasidiomyctes  | Tilletiales       | Tilletiaceae       | 4    |
| Tricholoma matsutake  | NC_028135 | 76,037      | Basidiomycota        | Agaricomycetes    | Agaricales        | Tricholomataceae    | 4    |

*For Grosmannia piceipera, only sequences of the mI gene are known. In this species, an mI group IA intron (mI2449) encodes an rps3 gene which is further fragmented by the insertion of a group IC2 intron (Rudski and Hausner 2012)*
The suggested nomenclature is flexible to fit some special conditions. Firstly, although we suggest three-letter abbreviation of host scientific name, four-or-more-letter abbreviation may be used in cases where the three-letter abbreviation cannot discriminate among all species under investigation. An example is introns at position 717 in nad5 in *Candida pseudojiufengensis* (Cpse.nad5U717) and *Candida psychrophila* (Cpsy.nad5P717) (Table 2, lines 11–12). Secondly, twintrons (twin introns) have been described from some fungal mitogenomes with various combinations of group I or II introns nested inside each other or situated next to each other (Hafez and Hausner 2015; Deng et al. 2016). The internal/external or upstream/downstream members of a twintron could be named alphabetically. An example is the side-by-side twintron in cox3 in *Hypomyces aurantius*, where two group IA introns are arranged in tandem (Deng et al. 2016). The upstream intron of the twintron can be named as Hau.cox3P640a and the downstream one as Hau.cox3P640b (Table 2, lines 13–14). Finally, although introns present at an identical insertion site among different strains of a particular species are generally conserved, distantly related introns are sometimes detected among different strains. Introns of this kind can be named numerically. For example, Hth.cobP429 in different strains of *Hirsutella thompsonii* showed length variations (e.g., 2.7 kb in ARSEF 9457 and 4.8 kb in ARSEF 1947) (Wang et al. 2018), and the two variants may be named as Hth.cobP429–1 in ARSEF 9457 and Hth.cobP429–2 in ARSEF 1947 (Table 2, lines 15–16).

The suggested nomenclature has been successfully applied to name introns in 16 fungi from different phyla, including both basal and higher fungal lineages (Table 3). These fungi contain introns in all protein-coding genes except atp8, nad2, and nad6, and cob and cox1 are most frequently invaded by introns. These introns are mostly group I introns, but we also find fewer group II introns as well as few introns with undetermined types. There are a total of 149 introns at 74 insertion sites in these fungi. Using the suggested nomenclature, intron positions in a particular gene can be directly observed and compared across different species. We find some points frequently inserted by introns in different species (e.g., cobP490, cox1P386, cox1P720, cox1P1107). From the intron insertion site numbers, one can also easily understand the phase of insertion sites in these fungi. Using the suggested nomenclature, intron positions in a particular gene can be directly observed and compared across different species. We find some points frequently inserted by introns in different species (e.g., cobP490, cox1P386, cox1P720, cox1P1107). From the intron insertion site numbers, one can also easily understand the phase of.

### Table 2: Representative examples of the new nomenclature of introns in protein-coding genes

| Line | New name | Old name | Fungal taxa | Host gene | Accession | Note | Reference |
|------|----------|----------|-------------|-----------|-----------|------|-----------|
| 1    | Sce.cox1S169 | al1      | *Saccharomyces cerevisiae* | cox1      | NC_001224 | Group II intron | Fourny et al. 1998 |
| 2    | Sce.cox1P771 | al5a     | *Saccharomyces cerevisiae* | cox1      | NC_001224 | Group I intron | Fourny et al. 1998 |
| 3    | Sce.cox1P1107 | alSβ   | *Saccharomyces cerevisiae* | cox1      | NC_001224 | Group I intron | Fourny et al. 1998 |
| 4    | Sce.cox1S1132 | al5γ    | *Saccharomyces cerevisiae* | cox1      | NC_001224 | Group I intron | Fourny et al. 1998 |
| 5    | Cgl.cox1P240 | Cgcox1.1 | *Candida glabrata* | cox1      | NC_004691 | Group I intron | Koszul et al. 2003 |
| 6    | Cgl.cox1P386 | Cgcox1.2 | *Candida glabrata* | cox1      | NC_004691 | Group I intron | Koszul et al. 2003 |
| 7    | Cgl.cox1P971 | Cgcox1.3 | *Candida glabrata* | cox1      | NC_004691 | Group I intron | Koszul et al. 2003 |
| 8    | Cme.cobP393  | b1       | *Candida metapsilosis* | cob       | NC_006971 | Group I intron | Kosa et al. 2005 |
| 9    | Hth.nad1P636 | nad1-i1  | *Hirsutella thompsonii* | nad1      | NC_040165 | Group I intron | Wang et al. 2018 |
| 10   | Ici.atp9P181 | atp9-i1  | *Isaria cicadae* | atp9      | MH922223 | Group I intron | Fan et al. 2019 |
| 11   | Cpse.nad5U717 |      | *Candida pseudojiufengensis* | nad5      | NC_022156 | Unknown intron type | Unpublished |
| 12   | Cpsy.nad5P717 |      | *Candida psychrophila* | nad5      | NC_036103 | Group I intron | Unpublished |
| 13   | Hau.cox3P640a | cox3-2  | *Hypomyces aurantius* | cox3      | NC_030206 | 1st one in twintron | Deng et al. 2016 |
| 14   | Hau.cox3P640b | cox3-2  | *Hypomyces aurantius* | cox3      | NC_030206 | 2nd one in twintron | Deng et al. 2016 |
| 15   | Hth.cobP429–1 | cox-i2 | *Hirsutella thompsonii* | cox3      | NC_040165 | Strain: ARSEF 9457 | Wang et al. 2018 |
| 16   | Hth.cobP429–2 | cox-i2 | *Hirsutella thompsonii* | cox3      | NC_040165 | Strain: ARSEF 9457 | Wang et al. 2018 |
| 17   | Zsa.nad5P717 | NDS–717 | *Zoanthus sansibaricus* | nad5      | KY888672 | Coral: Group I intron | Chi and Johansen 2017 |
| 18   | Zsa.cox1P867 | COI-867 | *Zoanthus sansibaricus* | cox1      | KY888672 | Coral: Group I intron | Chi and Johansen 2017 |
| 19   | Mbr.nad5P717 |      | *Monosiga brevicollis* | nad5      | AF538053 | Protist: Group I intron | Burger et al. 2003 |
| 20   | Ddi.cox2P357 | Ddi.cox2 | *Dictyostelium discoideum* | cox2      | NC_000895 | Protist: Group I intron | Ogawa et al. 2000 |
| 21   | Mpo.nad5P717 |      | *Marchantia polymorpha* | nad5      | M68929  | Plant: Group I intron | Oda et al. 1992 |
| 22   | Ath.cox2S691 |      | *Arabidopsis thaliana* | cox2      | NC_037304 | Plant: Group II intron | Sloan et al. 2018 |

*a Examples from lines 1 to 16 are fungal species, and those from lines 17 to 22 are plant/protist/animal species as indicated in the column “Note”*
| Fungal taxa                        | atp6 | atp8 | atp9 | cob | cox1   | cox2 | cox3 | nad1 | nad2 | nad3 | nad4 | nad4L | nad5 | nad6 | rps3 | No. introns | No. genes with introns |
|-----------------------------------|------|------|------|-----|--------|------|------|------|------|------|------|-------|------|------|------|-------------|------------------------|
| Cryptomycota                      |      |      |      |     |        |      |      |      |      |      |      |       |      |      |     |            |                         |
| Rozella allomycis                 | 0    | 0    | 0    | 0   | P731   | 0    | 0    | 0    | 0    | 0    | 0    | 0     | 0    | 0    | 0    | 1           | 1                      |
| Mucoromycota                      | 0    | 0    | P157 | P393, P400 | P386, P615, P720 | P685 | P219 | 0    | 0    | 0    | 0    | 0     | 0    | 0    | 0    | 11          |                         |
| Zoopagomyctota                    | U374 | 0    | P69  | P201, P393, P429, P506, P600, P759, P823 | P212, P281, S313, P372, P386, P615, P678, P698, P709, P710, P720, P731, P807, P867, P1107, P1125, P1230, P1296 | P685 | P447 | U166, P636 | 0    | 0    | P915  | P426, U1059 | 0           | 0           | 24          | 5                        |
| Blastocladiomycota                |      |      |      |     |        |      |      |      |      |      |      |      |       |      |      |     |            |                         |
| Allomyces macrogynus              | 0    | 0    | 0    | P201, P417, P429, P490, P600, P759 | P221, P281, S313, P372, P386, P615, P678, P698, P709, P710, P720, P731, P807, P1107, P1230, P1296 | P685 | P447 | U166, P636 | 0    | 0    | P915  | P426, P717, P934 | 0           | 0           | 24          | 5                        |
| Chytridiomycota                   |      |      |      |     |        |      |      |      |      |      |      |      |       |      |      |     |            |                         |
| Hyaloraphidium curvatum           | 0    | 0    | 0    | P411 | 0      | 0    | 0    | 0    | 0    | 0    | 0    | 0     | 0     | 0    | 0    | 0           | 1                      |
| Ascomycota                        |      |      |      |     |        |      |      |      |      |      |      |      |       |      |      |     |            |                         |
| Candida albicans                  | 0    | 0    | 0    | U393, U429 | P386, P709, P720, P1107 | 0    | 0    | 0    | 0    | 0    | 0    | 0     | 0    | 0    | 0    | 1           | 1                      |
| Grosmannia piceiperda             | NA   | NA   | NA   | NA   | NA     | NA   | NA   | NA   | NA   | NA   | NA   | NA    | NA    | NA   | NA   | NA   | P159        | 1                       |
| Isaria cicadae                    | U572 | 0    | P181 | P393, P400, P506, P823 | P212, P281, P709, P720, P731, P1057, P1281 | P228, P357, P631 | P219, P631 | 0    | 0    | 0    | 0    | 0     | P417, P570 | 0           | 0           | 21          | 8                        |
| Neurospora crassa                 | P344 | 0    | 0    | P393, P400 | 0      | 0    | 0    | P636 | 0    | P90  | P505  | P263  | P324  | P717 | 0    | 0           | 9                       |
| Saccharomyces cerevisiae           | 0    | 0    | 0    | S415, P429, P506, P759, P823 | S169, S205, P240, P720, P971, P1107, S132 | 0    | 0    | 0    | 0    | 0    | 0    | 0     | 0     | 0           | 0                       |
| Schizosacharomyces pombe          | 0    | 0    | 0    | 5687 | P386, P731 | 0    | 0    | 0    | 0    | 0    | 0    | 0     | 0     | 0    | 0    | 0           | 0                       |
| Tolypocladium inflatum            | 0    | 0    | 0    | 0    | 0      | 0    | 0    | 0    | 0    | 0    | 0    | 0     | 0     | 0    | 0    | 0           | 0                       |
| Basidiomycota                     |      |      |      |     |        |      |      |      |      |      |      |      |       |      |      |     |            |                         |
| Cryptococcus neoformans           | 0    | 0    | 0    | P490 | 0      | 0    | 0    | P166 | 0    | 0    | 0    | 0     | 0     | 0    | 0    | 0           | 2                       |
| Puccinia striiformis              | 0    | 0    | 0    | U429, P809 | P386, U615, P709, P720, P807, U1107 | P314 | 0    | 0    | 0    | 0    | S495  | 0     | P417, P717 | 0           | 0           | 12          | 5                        |
| Tilletia indica                   | 0    | 0    | 0    | P490 | P212, P386, P709, P1107, P1305 | 0    | 0    | 0    | 0    | 0    | 0    | 0     | 0     | P417, P717 | 0           | 0           | 8           | 3                        |
Table 3 | Intron positions in mitochondrial protein-coding genes of selected fungal species  

| Fungal taxa               | atp6 | atp8 | atp9 | cob | cox1 | cox2 | cox3 | nad1 | nad2 | nad3 | nad4 | nad4L | nad5 | nad6 | rps3 | No. introns | No. genes with introns | No. intron insertion points |
|--------------------------|------|------|------|-----|------|------|------|------|------|------|------|-------|------|------|------|-------------|--------------------------|-----------------------------|
| Tricholoma matsutake     | 0    | 0    | 0    | P823| U281, P372, P386, P720, P900 | P318, P357 | P276 | 0    | 0    | 0    | 0    | 0    | P417 | 0    | 0    | 10            | 5                        |
| No. introns              | 3    | 0    | 3    | 35  | 65   | 10   | 4    | 8    | 0    | 2    | 3    | 1    | 14   | 0    | 1    | 149          |                          |
| No. intron-containing species | 3   | 0    | 3    | 13  | 11   | 6    | 3    | 6    | 0    | 2    | 3    | 1    | 7    | 0    | 1    |                          |
| No. intron insertion points | 3   | 0    | 3    | 13  | 29   | 7    | 3    | 3    | 0    | 2    | 3    | 0    | 7    | 0    | 1    | 74            |                          |

\( ^{a} \) Re-annotation was performed if online or published annotations failed to correctly identify introns. Intron types were determined by RNAweasel (http://megasun.bch.umontreal.ca/RNAweasel/)

\( ^{b} \) "-" indicates the absence of corresponding genes in a particular mitogenome

\( ^{c} \) NA, not available
an intron, which is phase 0 when an intron inserts between two codons (e.g., cobP393), and phase 1 or 2 when an intron inserts within a codon (e.g., cox1S205, cox1P386). These introns are often found at highly conserved regions (Additional file 2).

In addition to fungi, plants and protists (but rarely in animals) also contain group I or II introns in their mitochondrial genes (Oda et al. 1992; Ogawa et al. 2000; Burger et al. 2003; Chi and Johansen 2017). The nomenclature suggested in this study could potentially apply to plant/protist/animal mitochondrial introns (Table 2, lines 17–22; Additional file 2). Plant mitogenomes, however, are also known to encode several intron-containing protein genes (e.g., nad7, ccmC, rps10, rpl2) that are absent in fungal mitogenomes (Zhang et al. 2011; Sloan et al. 2018). Introns are even found in tRNA-coding genes in plant mitogenomes (Smith et al. 2011). An additional plant reference is necessary to name introns unique to plant mitogenomes.

CONCLUSIONS

A standard nomenclature was suggested for introns in protein-coding genes in fungal mitogenomes. It was proved feasible by naming introns present in mitogenomes of 16 fungi from a broad range of taxonomic classification, and it also had the potential to name introns in plant/protist/animal mitogenomes. Future studies should follow the proposed nomenclature to ensure direct comparison across different studies.

Additional files

Additional file 1: Sequences of protein-coding genes of Tolypocladium inflatum ARS52 3280 (accession number NC_036382). Insertion site of group I introns are shown in red, group II introns in green, and introns with undetermined intron types in shade. (DOCX 21 kb)

Additional file 2: Intron insertion sites for 22 common introns. Exon sequences of cob, cox1, cox2, nad1, and nad5 of different fungal taxa plus few non-fungal taxa were aligned by MAFFT, and visualization of the aligned sequences was performed using ESPript 3.0 (Robert and Gouet 2014) under default settings. Refer to Tables 1 and 2 for organisms represented by accession numbers, and the accession numbers of non-fungal taxa are marked in red boxes. Insertion sites of introns are shown using upward arrows. For phase 0 introns, conserved amino acids before and after insertion sites are listed. The amino acid glycine (G) is frequently seen before insertion sites of phase 0 introns. For phase 1 or 2 introns, conserved amino acids at insertion sites are given, and corresponding triplet codons are marked by a horizontal line. (PPTX 2235 kb)

Acknowledgements

Authors are thankful to the editor and two anonymous reviewers for their suggestions that helped us improve the manuscript.

Adherence to national and international regulations

Not applicable.

Authors’ contributions

Y.J.Z designed the research and wrote the manuscript. S.Z performed the research. Both authors read and approved the final manuscript.

Funding

This study was funded by the National Natural Science Foundation of China (31872162), the Research Project Supported by Shanxi Scholarship Council of China (2017–015), Hundred Talents Program of Shanxi Province, and the Special Fund for Large Scientific Instruments and Equipment in Shanxi Province.

Availability of data and materials

All data used in this study are publicly available.

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Author details

1Institute of Applied Chemistry, Shanxi University, Taiyuan 030006, China.
2School of Life Science, Shanxi University, Taiyuan 030006, China.
3Institute of Applied Chemistry, Shanxi University, Taiyuan 030006, China.

Received: 8 June 2019 Accepted: 20 August 2019

Published online: 10 October 2019

References

Bartelli TF, Ferreira RC, Colombo AL, Briones MRS (2013) Intraspecific comparative genomics of Candida albicans mitochondria reveals non-coding regions under neutral evolution. Infection, Genetics and Evolution 14:302–312
Borkovich KA, Alex LA, Yarden O, Freitag M, Turner GE, Read ND et al (2004) Lessons from the genome sequence of Neurospora crassa: tracing the path from genomic blueprint to multicellular organism. Microbiology and Molecular Biology Reviews 68:1–108
Bullerwell CE, Lang BF (2005) Fungal evolution: the case of the vanishing mitochondrion. Current Opinion in Microbiology 8:362–369
Burger G, Forget L, Zhu Y, Gray MW, Lang BF (2003) Unique mitochondrial genome architecture in unicellular relatives of animals. Proceedings of the National Academy of Sciences 100:892–897
Cannon P, Aguire-Hudson B, Aime MC, Ainsworth AM, Bidartondo MI, Gaya E et al (2018) Definition and diversity. In: Willis KJ (ed) State of the World’s Fungi report. Royal Botanic Gardens, Kew, pp 4–11
Cech TR (1988) Conserved sequences and structures of group I introns. Annual Review of Biochemistry 57:955–985
Chi SI, Johansen SD (2017) Zoantharian mitochondrial genomes contain unique complex I introns and highly conserved intergenic regions. Gene 62824–31
Deng Y, Zhang Q, Ming R, Lin L, Lin X, Lin Y et al (2016) Analysis of the mitochondrial genome in Hypomyces aurantius reveals a novel trintron complex in fungi. International Journal of Molecular Sciences 17:1049
Emblem A, Karlsen BO, Evertsen J, Johansen SD (2011) Mitogenome rearrangement in the cold-water scleractinian coral Lophelia pertusa (Cnidaria, Anthozoa) involves a long-term evolving group I intron. Molecular Phylogenetics and Evolution 61:495–503
Fan W-W, Zhang S, Zhang Y-J (2019) The complete mitochondrial genome of the Chan-hua fungus Isaria cicadae: a tale of intron evolution in Cordycipitaceae. Environmental Microbiology 21:864–879
Fourny F, Roganti T, Lecrenier N, Purnelle B (1998) The complete sequence of the mitochondrial genome of Saccharomyces cerevisiae. FEBS Letters 440:325–331
Hafer M, Hausner G (2015) Convergent evolution of trintron-like configurations: one is never enough. RNA Biology 12:1275–1288
Hawksworth DL, Lücking R (2017) Fungal diversity revisited: 2.2 to 3.8 million species. Microbiol. Spectrum 5: FUNK-0052-2016
Johansen S, Haugen P (2001) A new nomenclature of group I introns in ribosomal DNA. RNA 7:935–936
Kosa P, Valach M, Tomaska L, Wolfe KH, Nosek J (2006) Complete DNA sequences of the mitochondrial genomes of the pathogenic yeasts Candida orthopsilosis and Candida metapsilosis: insight into the evolution of linear DNA genomes from mitochondrial telomere mutants. Nucleic Acids Research 34:2472–2481
Koszul R, Malpertuy A, Frangeul L, Bouchier C, Wincker P, Thierry A et al (2003) The complete mitochondrial genome sequence of the pathogenic yeast Candida (Torulopsis) glabrata. FEBS Letters 534:39–48
Lang BF (2018) Mitochondrial genomes in Fungi. In: Wells RD, Bond JS, Klinman J, Masters BSS (eds) Molecular Life Sciences. Springer, New York, pp 722–728
Nie Y, Wang L, Cai Y, Tao W, Zhang Y-J, Huang B (2019) Mitochondrial genome of the entomopathoid fungus Conidiobolus heterosporus provides insights into evolution of basal fungi. Applied Microbiology and Biotechnology 103: 1379–1391
Oda K, Yamaoto K, Ohta E, Nakamura Y, Takemura M, Nozato N et al (1992) Gene organization deduced from the complete sequence of liverwort Marchantia polymorpha mitochondrial DNA. A primitive form of plant mitochondrial genome. Journal of Molecular Biology 223:1–7
Ogawa S, Yoshino R, Angata K, Iwamoto M, Pi M, Kuroe K et al (2000) The mitochondrial DNA of Oxystromella discoida em:em:em: complete sequence, gene content and genome organization. Molecular & General Genetics 263:514–519
Robert X, Gouet P (2014) Deciphering key features in protein structures with the new ENDest server. Nucleic Acids Research 42:W320–W324
Rudšek SM, Hauser M (2012) The mtDNA rps3 locus has been invaded by a group I intron in some species of Grosmannia. Mycoscience 53:471–475
Saldanha R, Mohr G, Belfort M, Lambowitz AM (1993) Group I and group II introns. The FASEB Journal 7:15–24
Sloan DB, Wu Z, Sharbrough J (2018) Correction of persistent errors in Arabidopsis reference mitochondrial genomes. Plant Cell 30:525–527
Smith DR, Burki F, Yamada T, Grimwood J, Grigoriev IV, Van Etten JL, Keeling PJ (2011) The GC-rich mitochondrial and plastid genomes of the green alga Coccorymna give insight into the evolution of organelle DNA nucleotide landscape. PLoS One 6:e23624
Spatafora JW, Arme MC, Grigoriev IV, Martin F, Stajich JE, Blackwell M (2017) The fungal tree of life: from molecular systematics to genome-scale phylogenies. Microbiology Spectrum 5: FUNK-0053-2016
van der Giezen M, Tovar J, Clark CG (2005) Mitochondrion-derived organelles in protists and fungi. Int Rev Cytol. 244:175–225
Wang L, Zhang S, Li J-H, Zhang Y-J (2018) Mitochondrial genome, comparative analysis and evolutionary insights into the entomopathogenic fungus Hirsutella thompsonii. Environmental Microbiology 20:3393–3405
Zhang S, Hao AJ, Zhao YY, Zhang XY, Zhang Y-J (2017a) Comparative mitochondrial genomics toward exploring molecular markers in the medicinal fungus Cordyceps militaris. Scientific Reports 7:40219
Zhang S, Wang XN, Zhang XL, Liu XZ, Zhang Y-J (2017b) Complete mitochondrial genome of the endophytic fungus Pestalotiopsis fici: features and evolution. Applied Microbiology and Biotechnology 101:1593–1604
Zhang S, Zhang Y-J, Li Z (2019) Complete mitogenome of the entomopathogenic fungus Sparothev insectorum RCEF 264 and comparative mitogenomics in Ophiostomatiales. Applied Microbiology and Biotechnology 13:5797–5809
Zhang X, Zhang R, Hou SY, Shi J, Guo SD (2011) Research progress on mitochondrial genome of higher plant. Journal of Agricultural Science and Technology 13:23–31
Zhang Y-J, Yang XQ, Zhang S, Humber RA, Xu J (2017d) Genomic analyses reveal low mitochondrial and high nuclear diversity in the cynipoid-producing fungus Tolypocladium inflatum. Applied Microbiology and Biotechnology 101:8517–8531
Zhang Y-J, Zhang H-Y, Liu XZ, Zhang S (2017c) Mitochondrial genome of the nematode endoparasitic fungus Hirsutella vernicola reveals a high level of synteny in the family Ophiocordycipitaceae. Applied Microbiology and Biotechnology 101:3295–3304
Zhang Y-J, Zhang S, Zhang G, Liu X, Wang C, Xu J (2015) Comparison of mitochondrial genomes provides insights into intron dynamics and evolution in the caterpillar fungus Cordyceps militaris. Fungal Genetics and Biology 77:95–107

Publisher’s Note
Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.