

BRIEF REPORT

Open Access



Plants acquired mitochondrial linear plasmids horizontally from fungi likely during the conquest of land

Yutong Wei¹, Zhen Gong¹ and Guan-Zhu Han^{1*}

Abstract

Mitochondrial linear plasmids have been sporadically reported in fungi and plants. Yet, much remains obscure about the diversity, distribution, and evolution of mitochondrial linear plasmids. Here, through phylogenomic analyses across 7,163 cellular organisms (including 991 plants), we find that mitochondrial linear plasmids are widely present in land plants and fungi. Phylogenetic analyses indicate that plants are likely to have acquired mitochondrial linear plasmids horizontally from fungi before or during the conquest of terrestrial environments by plants. Gene content analyses show that mitochondrial linear plasmids harbor a highly dynamic and promiscuous repertoire of genes. Our study refines the understanding of the origin and evolution of mitochondrial linear plasmids.

Keywords Genome evolution, Horizontal gene transfer, Phylogenetics, Plant terrestrialization, Mitochondrial linear plasmids

Main text

Transposable elements (TEs) comprise a major proportion of the plant genomes, contributing substantially to the evolution of their genome complexity. Among diverse TEs, Polintons (also known as Mavericks) are unusual large DNA transposons that have been thought to be present widely in the genomes of animals and protists [1–3]. Typically, the Polinton genomes are flanked by terminal inverted repeats (TIRs) and encode several conserved proteins, including protein-primed family-B DNA polymerase (pPolB), retroviral-like (RVE) integrase, DNA packaging ATPase, and maturation protease [3]. Moreover, many Polintons encode homologs of viral capsid proteins, indicating that they might produce virions and represent *bona fide* viruses [4]. Therefore, Polintons have

been proposed to possess dual lifestyles as both transposons and viruses [3]. Interestingly, Polintons share a set of conserved genes with a variety of mobile genetic elements that pertain to viruses, virophages, TEs, or plasmids, forming a highly complex evolutionary network [3, 5]. Polintons, adenoviruses, bidnaviruses, virophages, transpovirons, cytoplasmic linear plasmids, mitochondrial linear plasmids, and diverse phages (e.g., tectiviruses) encode a common hallmark protein, namely pPolB [3]. For convenience, we use Polinton-like mobile genetic elements (Polin-MEs) to refer to these pPolB-encoding mobile genetic elements.

Among Polin-MEs, mitochondrial linear plasmids have been sporadically reported in fungi and plants, but their origins remain largely mysterious [6, 7]. Here, through phylogenomic analyses across 7,163 eukaryotes, we found that plants are likely to have acquired mitochondrial linear plasmids horizontally from fungi at least before or during the conquest of terrestrial environments. Our phylogenomic analyses provide insights into

*Correspondence:

Guan-Zhu Han
guanzhu@njnu.edu.cn

¹ College of Life Sciences, Nanjing Normal University, Nanjing 210023, Jiangsu, China



the origin, evolution, and diversity of mitochondrial linear plasmids in plants.

Distribution of mitochondrial linear plasmids in plants

First, to explore the distribution of the mitochondrial linear plasmids in plants, we used a combined similarity search and phylogenetic analysis approach to identify mitochondrial linear plasmids in a total of 991 representative plant species (including nuclear genomes and mitochondrial genomes) (Table S1). These plant species cover the major diversity of plants, including 781 angiosperms, 15 gymnosperms, 4 hornworts, 8 liverworts, 19 mosses, 125 chlorophytes, 1 glaucophyte, and 19 rhodophytes (Table S1). Mitochondrial linear plasmids were identified in nuclear genomes and/or mitochondrial genomes of a limited range of green plants, including angiosperms, gymnosperms, ferns, and liverworts (Fig. 1B), indicating that mitochondrial linear plasmids are likely to be only present in land plants.

Horizontal transfer of mitochondrial linear plasmids from fungi to plants

To explore the potential source of plant mitochondrial linear plasmids, we used the combined similarity search and phylogenetic analysis approach to identify mitochondrial linear plasmids in the genomes of 6,172 eukaryotes, including 77 Excavata species, 5,792 Amorphea species, 3 Haptista species, 295 SAR (Stramenopila, Alveolata, and Rhizaria) species, and 5 Cryptista species (Fig. 1A) (Table S1). Mitochondrial linear plasmids were identified in nuclear genomes and/or mitochondrial genomes of a large number of fungi, suggesting that horizontal transfer(s) of mitochondrial linear plasmids might have taken place between fungi and plants (Fig. 1A). Different from fungal retroplasmids that replicate via reverse transcription and encode reverse transcriptases, mitochondrial linear plasmids we identified here encode pPolB proteins. To further explore the evolutionary relationships of mitochondrial linear plasmids from plants and fungi, we performed phylogenetic analyses based on pPolB, a hallmark protein for analyzing the relationship of Polin-MEs [3].

Phylogenetic analyses of pPolB proteins clearly show that mitochondrial linear plasmids from land plants cluster together (UBoot=100%) and nest within the diversity of the plasmids from fungi with robust support (Fig. 2 and Fig. S1). It should be noted that most of the mitochondrial linear plasmid sequences from liverworts and ferns were too short to be used for reliable phylogenetic analyses, and we used only one pPolB protein from the mitochondrial genome of *Marchantia paleacea* (Supplemental Data Set 1). Nevertheless, our phylogenomic analyses support that plants are likely to have acquired mitochondrial linear plasmids horizontally from fungi.

We then closely interrogated the distribution of mitochondrial linear plasmids in fungi, and found that they are widely distributed in the genomes of terrestrial fungi, including Zygomycota, Mucoromycota, Basidiomycota, and Ascomycota (Fig. 1C). Interestingly, fungal plasmids that are closely related to plant plasmids are from species of Mucoromycota (Mortierellales, Mucorales, and Endogonales) and Basidiomycota (Microbotryales, Agaricales, Polyporales, and Russulales) (Fig. 1C), many of which can establish symbiosis with diverse plants, such as species within Agaricales, Russulales, and Endogonales [8, 9]. Taken together, we hypothesize that plants might have acquired mitochondrial linear plasmids horizontally from fungi probably through symbioses such as mycorrhizal symbiosis.

Frequent host switching of mitochondrial linear plasmids in plants

We observed that the phylogeny of mitochondrial linear plasmids (based on the pPolB proteins) is generally incongruent with the phylogeny of their plant hosts. Moreover, we compared the phylogeny of mitochondrial linear plasmids with that of their plant hosts. Indeed, we found no statistically significant congruence in phylogenies between plant mitochondrial linear plasmids and their hosts (Fig. S2, Table S2). These results suggest that plant mitochondrial linear plasmids might have undergone frequent host switching.

(See figure on next page.)

Fig. 1 Distribution of the mitochondrial linear plasmids in eukaryotes. **A** Distribution of the mitochondrial linear plasmids in major eukaryote groups. Numbers indicate the numbers of species used. Closed and open green circles indicate the presence and absence of the mitochondrial linear plasmids in the corresponding eukaryote groups, respectively. **B** Distribution of mitochondrial linear plasmids in plants. Numbers near plants indicate the numbers of species used. Closed and open orange circles indicate the presence and absence of the mitochondrial linear plasmids in the corresponding plant groups, respectively. Numbers in the circles represent the number of species with mitochondrial linear plasmids. **C** Distribution of mitochondrial linear plasmids in fungi. Numbers near fungi indicate the numbers of species used. Closed and open blue circles indicate the presence and absence of the mitochondrial linear plasmids in the corresponding fungal groups, respectively. Numbers in the circles represent the number of species with mitochondrial linear plasmids

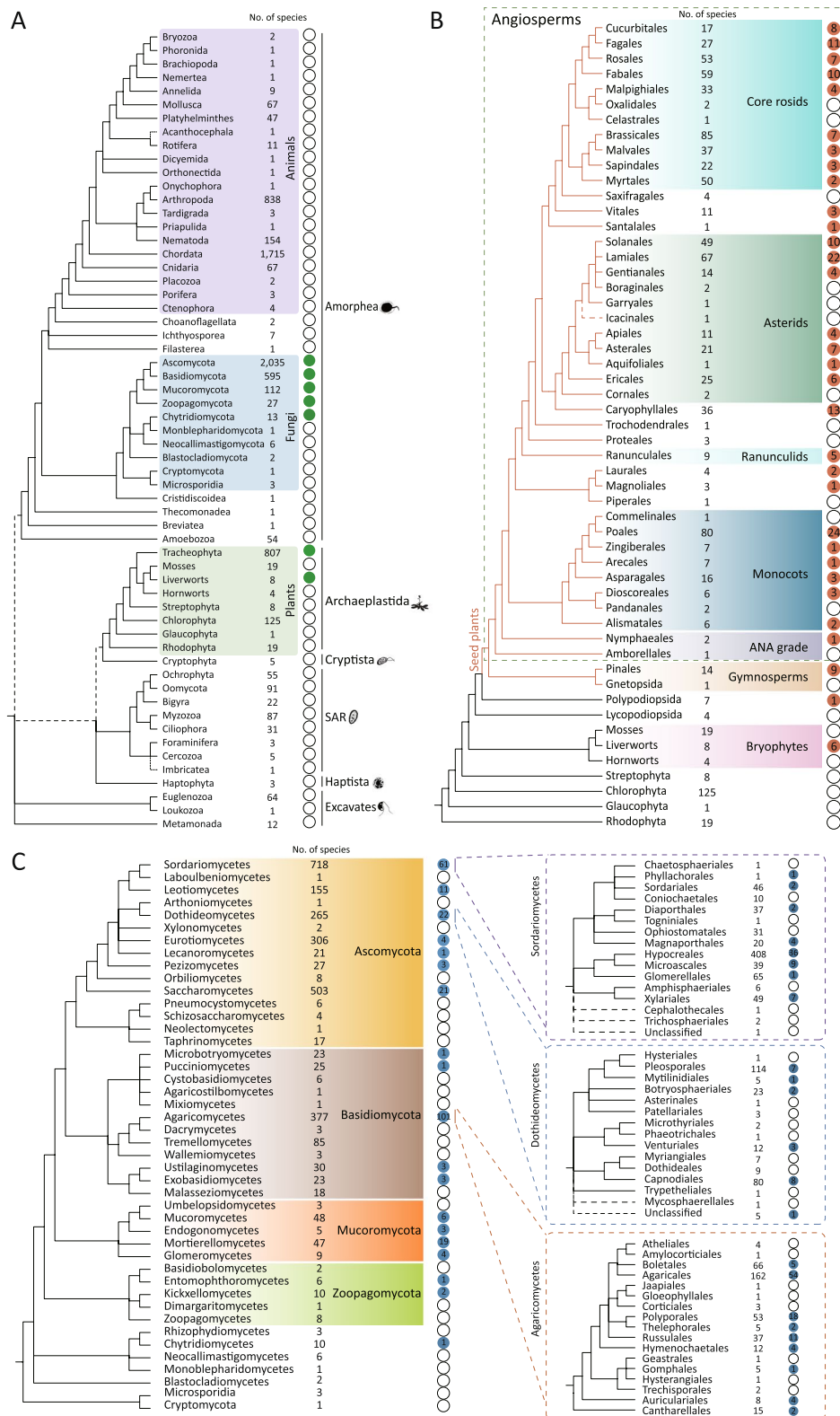


Fig. 1 (See legend on previous page.)

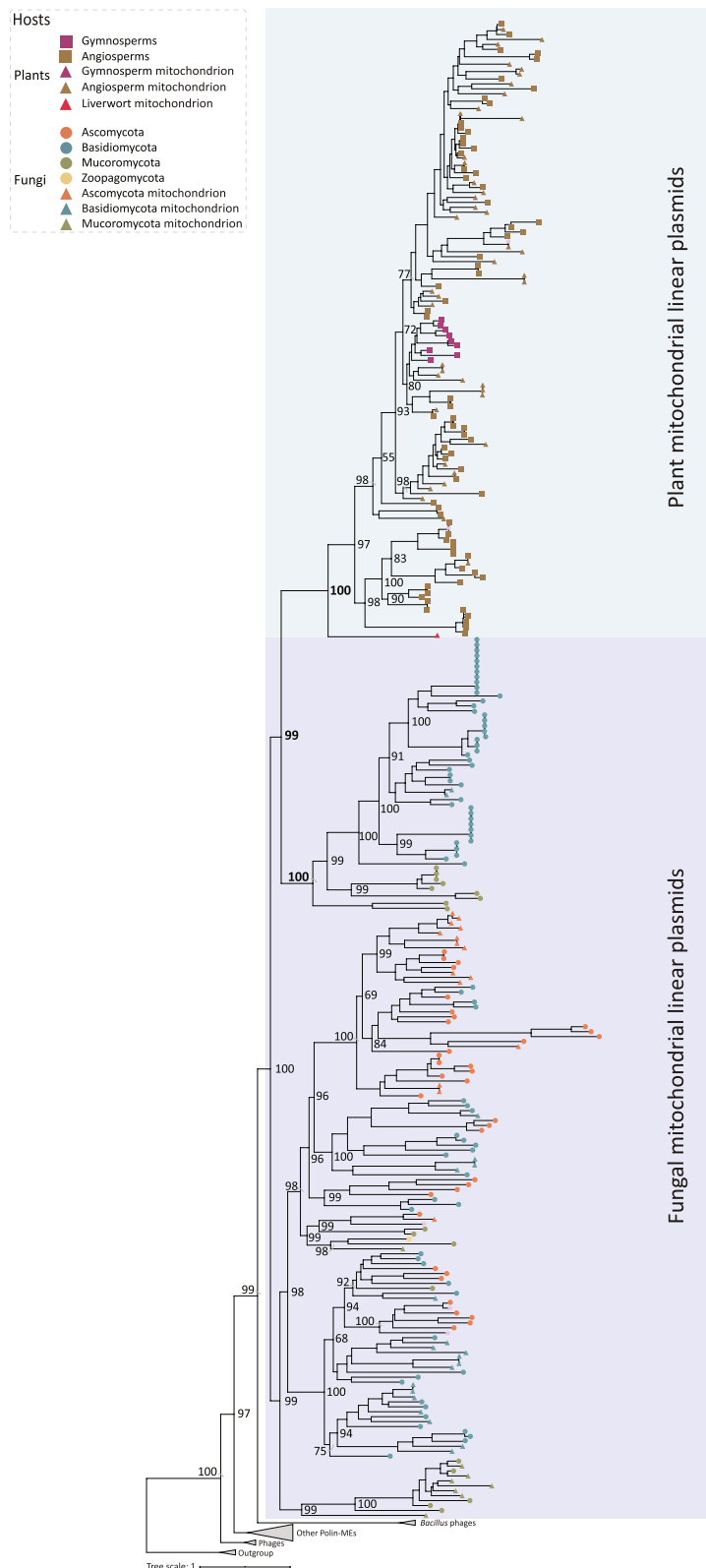


Fig. 2 Phylogenetic relationship among representative mitochondrial linear plasmids from plants and fungi. Mitochondrial linear plasmids from fungal and plant nuclear genomes are labeled with circles and rectangles, and are highlighted in purple and blue, respectively. Mitochondrial linear plasmids were identified in mitochondrial genomes are labeled with triangles. Representative pPolB proteins are labeled in pink stars. Branch support values are given near the selected nodes

Dynamic evolution of mitochondrial linear plasmid genome architectures

To explore the diversity and evolution of gene contents within mitochondrial linear plasmids, we identified a total of 95 complete elements characterized by the presence of two flanking TIRs from plants and

fungi, and annotated their protein domain architectures. Unlike canonical Polintons, mitochondrial linear plasmids of fungi or plants do not encode viral capsid proteins (MCP or Penton), which is consistent with their nature as plasmids (Fig. 3D). As expected, the hallmark protein domain, pPolB, was most frequently

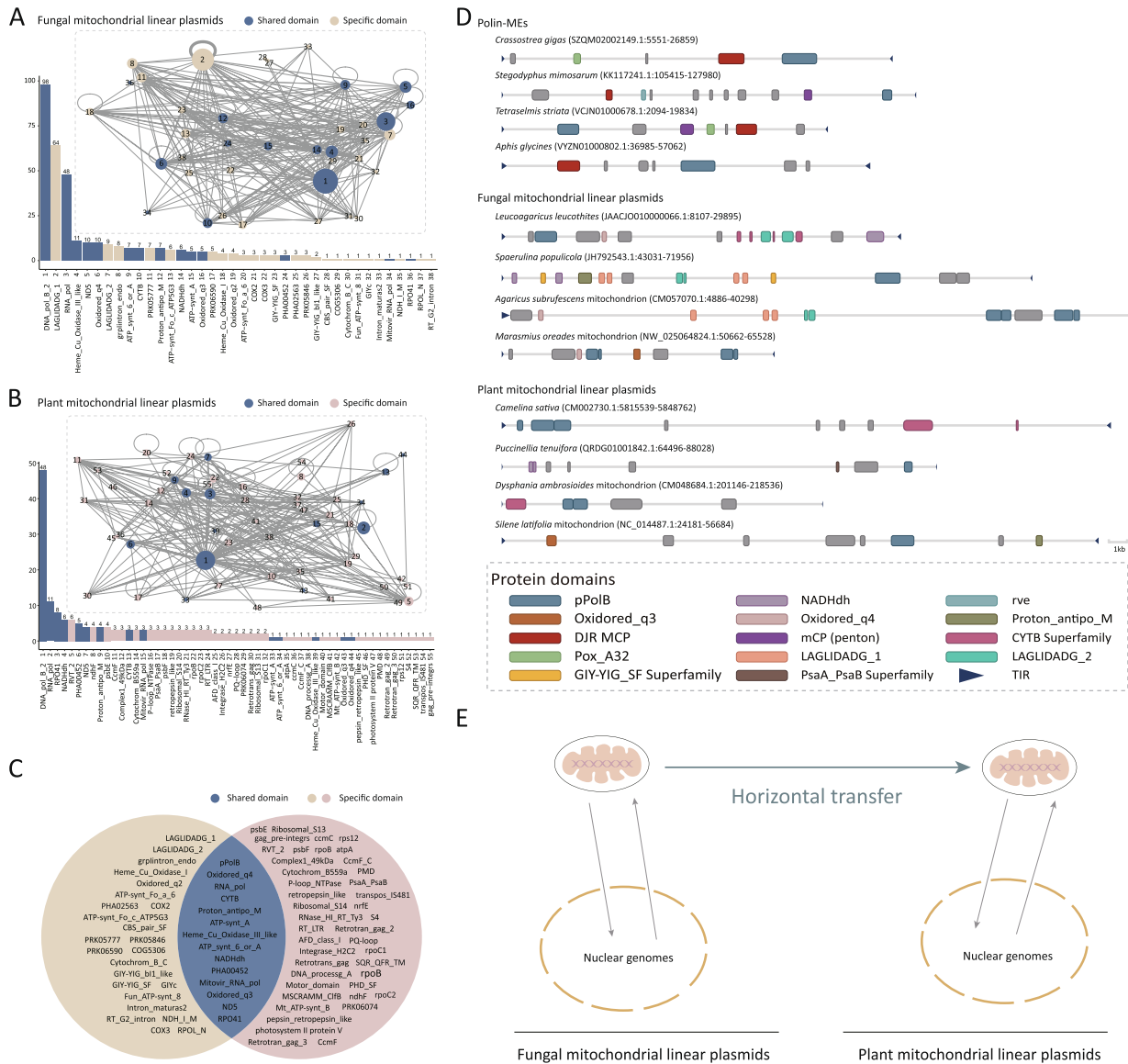


Fig. 3 Origin of the mitochondrial linear plasmids in seed plants. **A** Annotated protein domain architectures of complete mitochondrial linear plasmids from fungi. **B** Annotated protein domain architectures of complete mitochondrial linear plasmids from plants. Protein domains shared by fungal and plant mitochondrial linear plasmids are labeled in dark blue. Fungal and plant specific protein domains are labeled in beige and pink, respectively. Charts indicate the occurrence frequency of different protein domains in the corresponding groups. The circle size is proportional to the occurrence frequency of protein domains. Numbers near the circles represent domain names (important domains are specified) and are consistent with the numbers near the protein domain names in the charts. The thickness of lines linking two circles is proportional to the frequency at which the two protein domains appear in a complete element. **C** Venn diagram of annotated protein domains in fungal and plant mitochondrial linear plasmids. Protein domains shared by fungal and plant mitochondrial linear plasmids are highlighted in dark blue. Fungal and plant specific protein domains are highlighted in beige and pink, respectively. **D** Genome maps of representative Polin-MEs, fungal mitochondrial linear plasmids and plant mitochondrial linear plasmids. **E** Hypothetical scenarios for the origins and evolution of plant mitochondrial linear plasmids

identified. Most mitochondrial linear plasmids encode pPolB and RNA_pol (accession: cl44477), which indicates that they are capable of autonomous replication and transcription. Some domains are related to various transposable elements, and these domains might be derived from transposable element insertions rather than integral components of mitochondrial linear plasmids. Many host-derived domains also appear in mitochondrial linear plasmids but with relatively low frequency, such as mitochondrion-encoding genes and PsaA_PsaB-encoding genes (accession: cl08224) (Fig. 3A-C). Given the frequent movement between nuclear genomes and mitochondrial genomes observed for the plasmids (Fig. 1A), we suspect that mitochondrial linear plasmids might act as gene transfer agents which can transfer genes between nuclear genomes and mitochondrial genomes (Fig. 3E).

Domains related to homing endonucleases, such as GIY-YIG (accession: PF01541), PDDExK (accession: cl40440), LAGLIDADG (accession: PF00961), and HNH (accession: PF01844), were also frequently identified in mitochondrial linear plasmids (Fig. 3A-C). Homing endonuclease genes are selfish genetic elements that mobilize themselves through generating double strand breaks at specific target sites and getting copied across to the broken sites [10]. Therefore, we suspect that homing endonuclease domains might help drive the rapid spread of mitochondrial linear plasmids into related target sites within the host genome.

On the origin of mitochondrial linear plasmids in plants

In this study, we found that mitochondrial linear plasmids are widely present in the nuclear and/or mitochondrial genomes of land plants, and the distribution indicates that plants are likely to have acquired mitochondrial linear plasmids before or during the conquest of terrestrial environments by plants. Mitochondrial linear plasmids were identified in the genomes of terrestrial fungi (including Zygomycota, Mucoromycota, Basidiomycota, and Ascomycota), suggesting that mitochondrial linear plasmids might have emerged in fungi during their early transition to land in the Cryogenian period (~720 million years ago) [11]. However, these timescales should be taken with caution, because plants or fungi might have acquired mitochondrial plasmids later than estimated above (but spread to earlier-branching lineages through host switching) or earlier (but lost in earlier-branching lineages). Nevertheless, our analyses do show that plants might have acquired mitochondrial linear plasmids horizontally from fungi through the ancient symbiosis between plants and fungi.

Material and methods

Identification of mitochondrial linear plasmids in plants

We used a combined similarity search and phylogenetic analysis approach to identify mitochondrial linear plasmids based on the hallmark protein pPolB within the genomes of 991 plants [12, 13]. All the genomes of plants were retrieved from NCBI genome resources (Table S1). First, we used the tBLASTn algorithm to search plant genomes using various complete and truncated pPolb proteins (Supplemental Data Set 2) as queries with an *e* cut-off value of 10^{-5} . Then, we performed initial large-scale phylogenetic analyses of pPolB homologs from plants and representative pPolB proteins from Polin-MEs. Protein sequences were aligned using MAFFT 7.450 [14]. Large-scale phylogenetic analyses were performed using an approximate maximum likelihood method implemented in FastTree 2.1.10 [15].

Identification of mitochondrial linear plasmids in eukaryotes

To explore the potential source of the mitochondrial linear plasmids in plants, we used the aforementioned combined similarity search and phylogenetic analysis approach to screen the mitochondrial linear plasmids that are closely related to plant mitochondrial linear plasmids in a total of 6,172 eukaryotes. All the genomes of eukaryotes were retrieved from NCBI genome resources (Table S1).

Phylogenetic analysis

To explore the evolutionary relationships of the mitochondrial linear plasmids, we performed phylogenetic analyses using pPolB proteins of mitochondrial linear plasmids, some phages and Polin-MEs. The pPolB proteins were aligned using MAFFT 7.450 [14]. Phylogenetic analysis was performed using a maximum likelihood (ML) method implemented in IQ-tree 2 [16]. ModelFinder implemented in IQ-tree 2 was used to determine the best-fitting substitution model [16]. The node supports were evaluated using an ultrafast bootstrap method with 1,000 replicates [17]. Phylogenetic trees were annotated using iTOL [18].

Phylogeny congruence analysis

The phylogeny of plant mitochondrial linear plasmids was compared with that of their hosts using Jane 4 [19]. Different sets of cost values for five types of events (for cospeciation, duplication, duplication with host switch, loss, and failure to diverge: 0, 1, 2, 1, 1; -1, 0, 0, 0, 0; and 0, 1, 1, 2, 0) were examined [20]. The statistical analyses

were performed using the method of random parasite tree with the sample size of 500.

Domain architecture annotation

To identify complete mitochondrial linear plasmids, we used *gt tirvish* implemented in *GenomeTools* to identify the flanking TIRs [21]. We identified a total of 95 complete elements characterized by the presence of two flanking TIRs (Table S3). Complete mitochondrial linear plasmids with flanking TIRs were retrieved and annotated using various domain search or similarity search tools, including Conserved Domain search [22], HMMER [23, 24], tBLASTn, and BLASTp.

Abbreviations

TEs	Transposable elements
TIRs	Terminal inverted repeats
pPolB	Protein-primed family-B DNA polymerase
RVE	Retroviral-like endonuclease
Polin-MEs	Polinton-like mobile genetic elements
ML	Maximum likelihood

Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s13100-023-00304-7>.

Additional file 1: Figure S1. Phylogenetic relationship among representative mitochondrial linear plasmids from plants and fungi with all taxon names. **Figure S2.** Comparison of the phylogenies between plant mitochondrial linear plasmids and their hosts. The left is the phylogeny of plant mitochondrial linear plasmids based on the pPolB proteins, whereas the right is the host tree based on the plant tree of life. **Dataset S1.** Alignment of pPolB encoded by liverworts and ferns. **Dataset S2.** Representative pPolB proteins used in this study. **Table S1.** Information on species genomes used in this study. **Table S2.** Phylogeny congruence test for plant mitochondrial linear plasmids. **Table S3.** Information on fungus and plant genomes with the presence of mitochondrial linear plasmids.

Acknowledgements

Not applicable.

Authors' contributions

G.-Z.H. and Z.G. conceived and designed the research. Y.W. performed the research. G.-Z.H., Y.W. and Z.G. wrote the manuscript. All authors read and approved the final manuscript.

Funding

This work was supported by National Natural Science Foundation of China (31922001).

Availability of data and materials

Data sharing is not applicable to this article as no datasets were generated during the current study.

Declarations

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interests

The authors declare no competing interests.

Received: 10 August 2023 Accepted: 11 October 2023

Published online: 17 October 2023

References

- Feschotte C, Pritham EJ. Non-mammalian c-integrases are encoded by giant transposable elements. *Trends Genet.* 2005;21(10):551–2.
- Kapitonov VV, Jurka J. Self-synthesizing DNA transposons in eukaryotes. *Proc Natl Acad Sci U S A.* 2006;103(12):4540–5.
- Krupovic M, Koonin EV. Polintons: a hotbed of eukaryotic virus, transposon and plasmid evolution. *Nat Rev Microbiol.* 2015;13(2):105–15.
- Krupovic M, Bamford DH, Koonin EV. Conservation of major and minor jelly-roll capsid proteins in Polinton (Maverick) transposons suggests that they are bona fide viruses. *Biol Direct.* 2014;29(9):6.
- Koonin EV, Krupovic M. Polintons, virophages and transpovirons: a tangled web linking viruses, transposons and immunity. *Curr Opin Virol.* 2017;25:7–15.
- Warren JM, Simmons MP, Wu Z, Sloan DB. Linear Plasmids and the Rate of Sequence Evolution in Plant Mitochondrial Genomes. *Genome Biol Evol.* 2016;8(2):364–74.
- Handa H. Linear plasmids in plant mitochondria: peaceful coexistences or malicious invasions? *Mitochondrion.* 2008;8(1):15–25.
- Hibbett DS, Matheny PB. The relative ages of ectomycorrhizal mushrooms and their plant hosts estimated using Bayesian relaxed molecular clock analyses. *BMC Biol.* 2009;10(7):13.
- Chang Y, Desirò A, Na H, Sandor L, Lipzen A, Clum A, Barry K, Grigoriev IV, Martin FM, Stajich JE, Smith ME, Bonito G, Spatafora JW. Phylogenomics of Endogonaceae and evolution of mycorrhizas within Mucoromycota. *New Phytol.* 2019;222(1):511–25.
- Burt A, Koufopanou V. Homing endonuclease genes: the rise and fall and rise again of a selfish element. *Curr Opin Genet Dev.* 2004;14(6):609–15.
- Lutzoni F, Nowak MD, Alfaro ME, Reeb V, Miadlikowska J, Krug M, Arnold AE, Lewis LA, Swofford DL, Hibbett D, Hilu K, James TY, Quandt D, Magallón S. Contemporaneous radiations of fungi and plants linked to symbiosis. *Nat Commun.* 2018;9(1):5451.
- Wang J, Han GZ. A sister lineage of sampled retroviruses corroborates the complex evolution of retroviruses. *Mol Biol Evol.* 2021;38(3):1031–9.
- Gong Z, Han GZ. Euphyllphyte paleoviruses illuminate hidden diversity and macroevolutionary mode of Caulimoviridae. *J Virol.* 2018;92(10):e02043–e2117.
- Katoh K, Standley DM. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol Biol Evol.* 2013;30(4):772–80.
- Price MN, Dehal PS, Arkin AP. FastTree 2—approximately maximum-likelihood trees for large alignments. *PLoS One.* 2010;5(3):e9490.
- Minh BQ, Schmidt HA, Chernomor O, Schrempf D, Woodhams MD, von Haeseler A, Lanfear R. IQ-TREE 2: new models and efficient methods for phylogenetic inference in the genomic era. *Mol Biol Evol.* 2020;37(5):1530–4.
- Hoang DT, Chernomor O, von Haeseler A, Minh BQ, Vinh LS. UFBoot2: Improving the Ultrafast Bootstrap Approximation. *Mol Biol Evol.* 2018;35(2):518–22.
- Letunic I, Bork P. Interactive Tree Of Life (iTOL) v4: recent updates and new developments. *Nucleic Acids Res.* 2019;47(W1):W256–9.
- Conow C, Fielder D, Ovadia Y, Libeskind-Hadas R. Jane: a new tool for the copyphylogeny reconstruction problem. *Algorithms Mol Biol.* 2010;5:16.
- Xu X, Zhao H, Gong Z, Han GZ. Endogenous retroviruses of non-avian/mammalian vertebrates illuminate diversity and deep history of retroviruses. *PLoS Pathog.* 2018;14(6):e1007072.
- Gremme G, Steinbiss S, Kurtz S. GenomeTools: a comprehensive software library for efficient processing of structured genome annotations. *IEEE/ACM Trans Comput Biol Bioinform.* 2013;10(3):645–56.
- Lu S, Wang J, Chitsaz F, Derbyshire MK, Geer RC, Gonzales NR, Gwadz M, Hurwitz DI, Marchler GH, Song JS, Thanki N, Yamashita RA, Yang M, Zhang

- D, Zheng C, Lanczycki CJ, Marchler-Bauer A. CDD/SPARCLE: the conserved domain database in 2020. *Nucleic Acids Res.* 2020;48(D1):D265–8.
23. Eddy SR. Profile hidden Markov models. *Bioinformatics.* 1998;14(9):755–63.
24. Gao Y, Wang W, Zhang T, Gong Z, Zhao H, Han GZ. Out of water: the origin and early diversification of plant R-Genes. *Plant Physiol.* 2018;177(1):82–9.
25. Kalyanamoorthy S, Minh BQ, Wong TKF, von Haeseler A, Jermin LS. ModelFinder: fast model selection for accurate phylogenetic estimates. *Nat Methods.* 2017;14(6):587–9.

Publisher's Note

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

Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more biomedcentral.com/submissions

