SOCIETY FOR EXPERIMENTAL BIOLOGY

Journal of Experimental Botany
https://doi.org/10.1093/jxb/erae034 Advance Access Publication 1 February 2024

REVIEW PAPER

Methods for assembling complex mitochondrial genomes in land plants

Helena Štorchová*, and Manuela Krüger

Institute of Experimental Botany, Czech Academy of Sciences, Rozvojová 313, CZ-16502 Prague, Czech Republic

* Correspondence: storchova@ueb.cas.cz

Received 22 November 2023; Editorial decision 25 January 2024; Accepted 29 January 2024

Editor: Martin Janda, University of South Bohemia in České Budějovice, Czech Republic

Abstract

The large size and complex structural rearrangements inherent in the mitochondrial genomes of land plants pose challenges for their sequencing. Originally, the assembly of these genomes required the cloning of mitochondrial DNA fragments followed by Sanger sequencing. Subsequently, the advent of next-generation sequencing significantly expedited the process. This review highlights examples of plant mitochondrial genome assembly employing various technologies, including 454 sequencing, Illumina short sequencing reads, and Pacific Biosciences or Oxford Nanopore Technology long sequencing reads. The combination of short and long reads in hybrid assembly has proven to be the most efficient approach for achieving reliable assemblies of land plant mitochondrial genomes.

Keywords: Hybrid assembly, land plants, mitochondrial genome, next-generation sequencing, rearrangement, recombination, Unicycler.

Introduction

Unlike the uniform mitochondrial genomes (mitogenomes) of animals, which have standard sizes of 15–20 kb and conserved genomic structure and gene order (Taanman, 1999), plant mitogenomes are considerably variable in terms of size, structure, and gene ordering (Chevigny *et al.*, 2020). This diversity is more pronounced in land plants than in algae (Archaeplastida), and reaches its maximum in angiosperms (Christensen, 2021). A typical mitogenome of flowering plants is ~400–600 kb in length, mitogenomes as large as 11.7 Mbp (*Larix sibirica*; Putintseva *et al.*, 2020) and 11.3 Mbp (*Silene conica*; Sloan *et al.*, 2012a) have been reported. This increase in mitogenome size is attributed to the enlargement of intergenic regions, whereas the number of genes remains relatively stable, ranging between 40 and 70 (Palmer and Herbon, 1988; Mower 2020).

The estimation of the complete sequence (assembly) of land plant mitogenomes presents a challenging task, not only due to their large size but also because of structural rearrangements mediated by recombination across numerous repeats. The rearrangements in plant mitogenomes are likely associated with double-strand breaks, which disrupt DNA integrity (Abdelnoor *et al.*, 2003; Shedge *et al.*, 2007). DNA repair of double-strand breaks relies on homologous recombination or non-homologous end joining, which generate crossovers, chimeric genes, new repeats, and structural changes (Christensen, 2013). Mitogenomes exist in several configurations, which often consist of individual DNA molecules (subgenomes) that may be mutually joint or remain autonomous, forming a multipartite genome genome (Sloan *et al.*, 2012b;). This complexity

is further amplified by frequent transfers of DNA from plastids and the nucleus to plant mitochondria (Ellis, 1982; Goremykin *et al.*, 2012; Gandini and Sanchez-Puerta, 2017).

The fluidity of mitogenomes (particularly the variability of intergenic regions and frequent genomic rearrangements) has slowed down their sequencing compared with plastid genomes, which are conserved in size and structure (Ruhlman and Jansen, 2014). Complete mitogenome sequences of ~600 species were available in NCBI GenBank as of November 2023, compared with more than 13 000 complete plastid genome sequences. Only a few species, primarily crops, are represented by multiple completely sequenced mitogenomes (*Pisum sativum*, *Zea mays*, *Oryza sativa*, *Solanum tuberosum*, *Arabidopsis thaliana*, and *Silene vulgaris*). This contrasts with the dramatic increase of complete plastid genomic sequences, which are now often used to decipher phylogenetic relationships at both interspecific and intraspecific levels (Skuza *et al.*, 2023; Zhang *et al.*, 2023). The recent surge in next-generation sequencing (NGS) techniques, particularly those capable of generating long reads such as Pacific Biosciences (PacBio) and Oxford Nanopore Technologies (ONT), has facilitated the assembly of complex mitogenomes. However, no universal pipeline has yet been developed to reliably assemble any plant mitogenome. Although comprehensive reviews dealing with the structure and evolution of plant mitogenomes have been published (Gualberto and Newton, 2017; Mower, 2020; Petersen *et al.*, 2020), a summary of current progress in the methods used to assemble complex mitogenomes in land plants is still missing. This review aims to fill this gap.

Early assemblies of plant mitogenomes relied on cloned mitochondrial fragments

The first completely sequenced mitogenome of land plants was that of *Marchantia polymorpha* (Oda *et al.*, 1992). It was described as a single circular molecule of approximately 184 400 bp. The authors used Sanger sequencing of genomic fragments cloned into cosmid vectors, which, despite being the best available method at the time, was slow and laborious. The circular structure of the mitogenome was confirmed by electron microscopy (Oda *et al.*, 1992). However, a detailed examination of the *M. polymorpha* mitogenome using pulsed-field gel electrophoresis revealed that circular molecules represented only a small portion of mitochondrial DNA. The majority of the *M. polymorpha* mitogenome existed as linear subgenomic or concatemeric molecules, sometimes forming branched structures (Oldenburg and Bendich, 2001). Although it was known that the depiction of plant mitogenomes as circular molecules was not entirely accurate, this convention persisted due to technical difficulties in recognizing their multiple conformations. The 'master circle' proved to be useful for comparing gene content, open reading frames, and specific intergenic sequences among individual mitogenome assemblies (Allen *et al.*, 2007; Goremykin *et al.*, 2009; Cuenca *et al.*, 2013). Plant mitogenomes were sequenced, using the Sanger method, from fragments of mitochondrial DNA cloned into cosmids or bacterial artificial chromosomes in the first decade of the 21st century (Unseld *et al.*, 1997; Kubo *et al.*, 2000; Notsu *et al.*, 2002).

NGS methods made it possible to assemble plant mitogenomes without previous cloning of mitochondrial DNA

The first NGS method, 454 sequencing (or pyrosequencing), generated hundreds of thousands of sequencing reads of ~600 bp in size, an amount sufficient for the assembly of highly rearranged angiosperm mitogenomes. The GS *De Novo* Assembler (Newbler) developed by Roche appeared to be an ideal tool to decipher multiple configurations of land-plant mitogenomes. It assembles individual contigs, that is, sets of overlapping DNA fragments representing contiguous genomic regions. The contigs generated by Newbler correspond to repeats and unique regions between them. The output file 454ContigGraph.txt provided information about contig connections necessary to reconstruct the rearrangements of the plant mitogenome produced by recombination across the repeats. Newbler did not include a tool for the visualization of the graph, but Iorizzo *et al.* (2012) created the Perl program bb.454con-tignet (http://www.vcru.wisc.edu/simonlab/sdata/software/), which is capable of depicting connections among the contigs. Unfortunately, Newbler 3.0 cannot be currently downloaded from a public link, because Roche stopped supporting it.

The 454 sequencing method was used to assemble, for example, the mitogenome of rice (Fujii *et al.*, 2010), *Raphanus sativus* (Tanaka *et al.*, 2012), *Liriodendron tulipifera* (Richardson *et al.*, 2013), and the highly rearranged mitogenome of *S. vulgaris* (Sloan *et al.*, 2012b; Štorchová *et al.*, 2018). The assembly was facilitated by using purified or enriched mitochondria for DNA extraction, but no cloning procedure was necessary with this method. The contigs corresponding to plastid inserts in mitogenomes could have been distinguished from plastid DNA based on their coverages.

Short reads generated by Illumina sequencing cannot resolve large repeats

Illumina sequencing generates millions of highly accurate reads, which decreases the cost of analyses (Satam *et al.*, 2023). However, the reads are short (50–300 bp) and the size of insert from which paired-end reads are derived is typically 500 bp or less. Unlike 454 reads, which are somewhat longer, the short Illumina reads are less suitable for assembling multiple configurations of land-plant mitogenomes; this is because they cannot resolve large repeats, as they do not span them.

Several programs, such as Velvet (Zerbino, 2010), SPAdes (Bankevich *et al.*, 2012), and NOVOplasty (Dierckxsens *et al.*, 2016), were developed for the assembly of organellar genomes from whole-genome short-read sequencing data. The SPAdes assembler cuts the reads into shorter strings, called *k*-mers, whose sizes depend on the read length. The software connects individual *k*-mers based on their overlaps and constructs a de Bruijn graph by traversing through the edges connecting the *k*-mers (Pevzner *et al.*, 2004). The final genomic assembly is generated as a path through all the edges of the graph, revisiting the repeats two or more times according to their copy number. If the repeats are not properly identified, the assembly is not accurate and simple paths are not found.

The most recent assembler, GetOrganelle (Jin *et al.*, 2020), is a complex toolkit that uses a number of programs and scripts, including SPAdes, BLAST, or Bowtie2 (Langmead and Salzberg, 2012). It successfully assembles plastid genomes that are conserved in size (~150 kb) and structure with two single copy regions and a large inverted repeat. However, its application to the assembly of land-plant mitogenomes is limited by the short length of repeats. For example, GetOrganelle was used to assemble the mitogenome of the carnivorous plant *Genlisea tuberosa* from Illumina MiSeq with paired ends of 2 × 300 bp, which was ~730 kb in size, but the repeats were shorter than 300 bp (Matos *et al.*, 2022).

Long-read sequencing and hybrid assembly

Long-read sequencing technologies such as single-molecule real-time sequencing from PacBio and ONT have improved the accuracy in the assembly of large nuclear genomes, particularly in repetitive regions (Satam *et al.* 2023). Their disadvantage, compared with the more accurate Illumina sequencing, is a high error rate of ~10%. The assembler Canu (Koren *et al.*, 2017) performs sequence correction of long reads as the first step using the alignment of homologous reads. The recently developed PacBio high fidelity (HiFi) technology produces highly accurate long reads by finding consensus from circular sequencing. The pipeline MitoHiFi (Uliano-Silva *et al.*, 2023) makes it possible to assemble eukaryotic mitogenomes from HiFi sequencing reads, but it was optimized for small, compact animal mitogenomes. Canu could be used to decipher plant mitogenome sequences from long PacBio HiFi sequencing reads (Lu *et al.*, 2022).

However, frequent rearrangements of plant mitogenomes caused by recombinations across large repeats present a challenge even for long-read-based assembly methods. The presence of reads derived from various genomic configurations prevents the assembly of long continuous sequences, resulting in highly fragmented output. The Flye assembler (Kolmogorov *et al.*, 2019) takes repeats into consideration and generates concatenates derived from disjointed genomic regions. The final assembly graph is created with the help of individual reads bridging the repeats. Flye is therefore more suitable for plant mitogenome assembly than Canu, which generates a more fragmented assembly. The performance of Flye, and the resulting assembly, depends on the correct identification of the repeats. This is more challenging in highly recombinogenic mitogenomes with complex repeats consisting of several shorter repeat units (subrepeats). As an illustration, the complex repeat could be depicted as ABCDE, whereas the repeats BCE, ACD, BE, or A, B, D are located somewhere else in the mitogenome.

The Flye assembly derived from ONT whole-genome sequencing data visualized by using Bandage is shown in Fig. 1 (Wick *et al.*, 2015). Although the plastid genome was assembled accurately, the mitogenome is represented by a spider-like graph through which an unequivocal path cannot be found. Some repeats were assembled as short individual edges and at the same time as a part of long edges, which obscures mutual connections.

The most promising way to achieve a reliable assembly of land-plant mitogenome became hybrid assembly, which combines long- and short-read sequencing data. The Unicycler pipeline (Wick *et al.*, 2017) was originally developed to assemble short circular bacterial genomes, using numerous software tools,

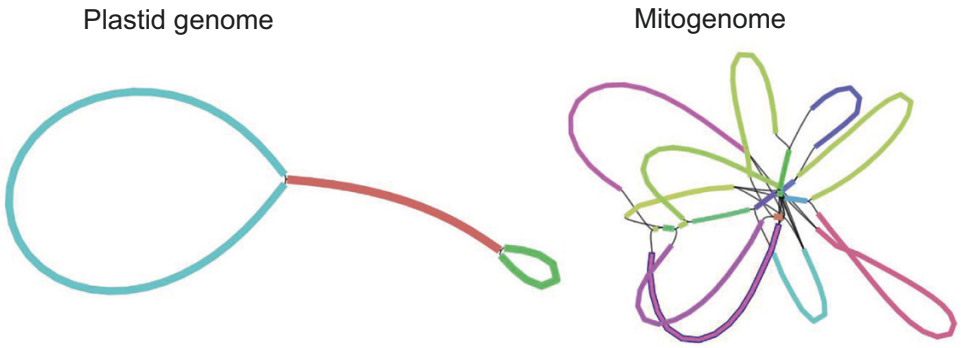

Fig. 1. Flye output graph made of ONT whole-genome sequencing data of *Silene fabaria* visualized by Bandage. The plastid genome is displayed on the left. The large and small loops correspond to the single-copy regions; the stem depicts inverted repeats. The mitochondrial genome is represented by the complex graph on the right. Unlike the plastid genome, a simple path through this mitochondrial graph cannot be found.

as shown in Fig. 2. In the first step, SPAdes creates a graph from short reads with minimum dead ends, then the coverage depth is assigned to the contigs, which helps to recognize the repeats. The graphs are simplified by resolving repeats by bridging with short and long reads. The final assembly joins adjacent contigs and repeats, and is corrected by Pilon (Walker *et al.*, 2014).

Hybrid assembly of plant mitogenomes does not require isolated or enriched mitochondria as the starting material for DNA extraction, and so whole-genome sequencing data are often used. However, more accurate assembly is achieved when the reads derived from mitochondrial DNA are selected before inputting them to Unicycler. Minimap2 (Li, 2018) or BLAST can retrieve long reads based on their similarity with related completely sequenced plant mitogenomes (Fig. 3). Flye or Canu create mitochondrial contigs, which are used as a reference for a second round of similarity searching and assembly. Short reads and long reads are then mapped against the mitochondrial contigs and uploaded to Unicycler. This approach was applied, for example, on the mitogenomic assembly of *Avena longiglumis* (Liu *et al.*, 2023). Alternatively, short reads may be assembled by GetOrganelle (Jin *et al.* 2020) and the resulting output used to retrieve mitochondrial long reads (Ni *et al.*, 2023;Yu *et al.*, 2023).

It is advantageous to try several analytical pipelines and to compare the assemblies obtained by various methods. Mitochondrial short and long reads can be uploaded to Newbler, regardless of whether they were produced by 454 sequencing. The output 454ContigGraph.txt provides independent identification of repetitive and unique regions, which helps to refine and confirm the final assembly (Fischer *et al.*, 2022).

Mitochondrial heteroplasmy, defined as the coexistence of two or more distinct mitogenomes in the same cell, is common in plants (Woloszynska, 2010). The final mitogenomic assembly represents the prevalent genomic structural variant, supported by the majority of reads and omitting less frequent rearrangements. The heteroplasmy in single nucleotide variants is concealed in the consensus sequences obtained by the last polishing steps. It can be revealed in the read alignments produced by using various variant call programs such as samtools/pileup and Genome Analysis Toolkit (GATK), as summarized by Yao *et al.* (2020).

Unicycler enables the distinction between the mitochondrial DNA of the mitogenome and the mitochondrial DNA inserted into the nuclear genome (NUMT), by considering their coverages. Mitochondrial DNA is more abundant than nuclear DNA, especially in organs rich in mitochondria, such as flower buds (Sloan *et al.*, 2012b). The 454ContigGraph displays NUMTs as contigs with low coverage and without any connection to other contigs. The validation of the final assembly by an independent method, such as Southern hybridization, is highly recommended (Sloan *et al.*, 2010; Iorizzo *et al.*, 2012).

Conclusions

Recent advances in NGS and the development of new bioinformatic programs have provided sufficient tools to assemble highly recombinogenic and rearranged mitogenomes of land plants. Using a combination of several methods, rather than a single assembler, is helpful. The sequencing of complete plant

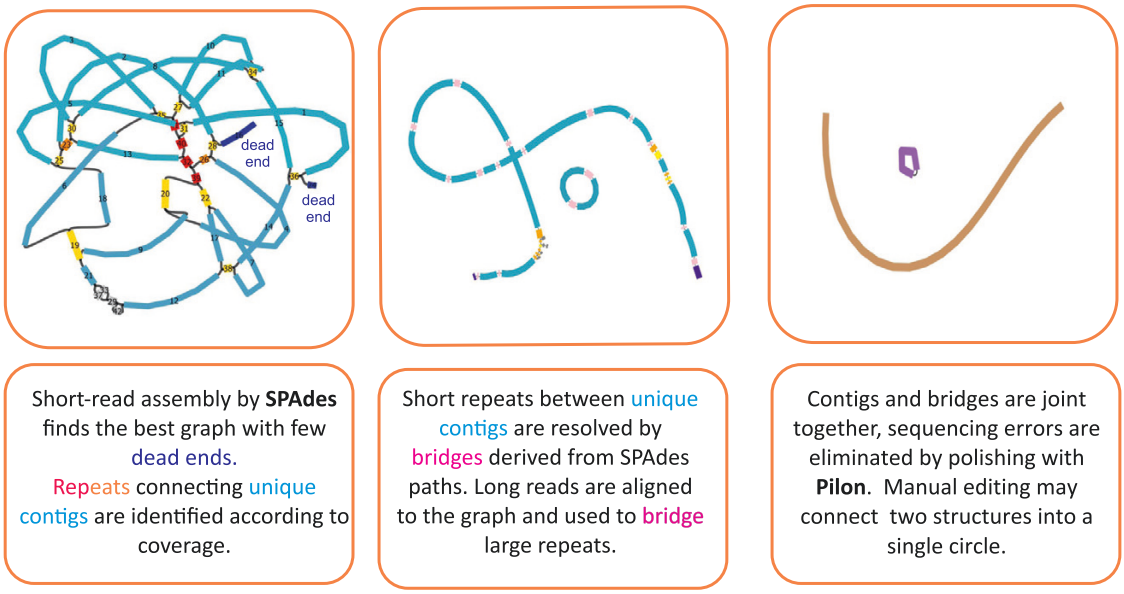

Fig. 2. Unicycler hybrid assembly of mitochondrial short- and long-read sequencing data of *Silene fabaria*. The process initiates with SPAdes assembly, followed by the resolution of repeats, and correction of the final sequence. Manual editing may be necessary to enhance the continuity of the assembly, such as connecting dead ends when supported by long reads. Unique contigs are marked in blue, while yellow and red bars represent repetitive sequences. The pink regions indicate bridges formed from long reads aligned to two adjacent unique contigs. Two final contigs (brown and violet) may be manually connected into a single circle.

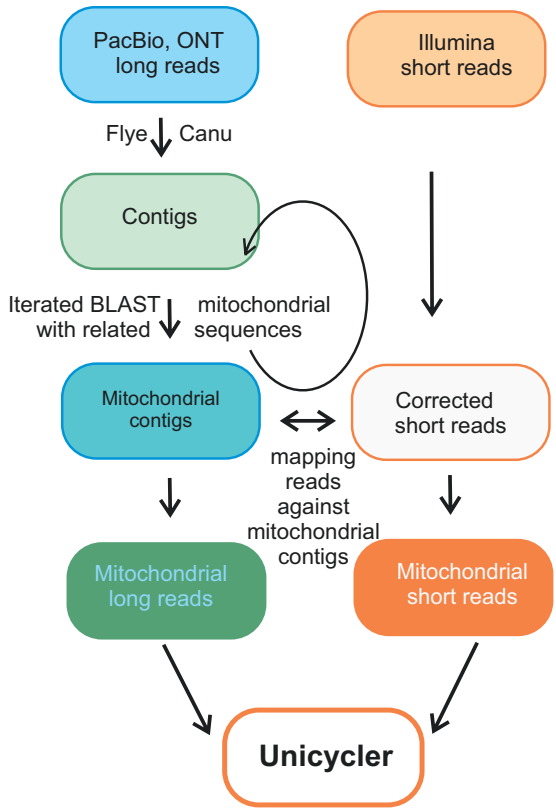

Fig. 3. Simplified view of the land plant mitogenome assembly pipeline. The pre-selection of mitochondrial reads is a recommended step in the pipeline.

mitogenomes will contribute to phylogenetic studies, because mitogenomes are larger than plastid genomes and may contain a greater total number of polymorphisms than plastid DNA despite a lower mutation rate (Duminil and Besnard, 2021). Dense sampling of completely sequenced mitogenomes in plant genera and families, known to have rapid acceleration of substitution rates (Mower *et al.*, 2007) or pervasively rearranged mitogenomes (Sloan *et al.*, 2012b), will shed light on the origin and evolution of these extreme traits. Complete sequences of plant mitogenomes will facilitate the search for chimeric mitochondrial genes, which are often responsible for cytoplasmic male sterility in crops. They will also contribute to better understanding of the relationship between the evolution of plant reproduction systems and mitochondrial DNA.

Acknowledgements

The authors thank Oushadee A. J. Abeywardana for helpful discussions.

Author contributions

HŠ and MK conceptualized, wrote, and approved the final version of the manuscript.

Conflict of interest

The authors declare no conflicts of interests.

Funding

This work was supported by the Czech Science Foundation (GAČR) grant no. 19-01639S to HŠ.

References

Abdelnoor RV, Yule R, Elo A, Christensen AC, Meyer-Gauen G, Mackenzie SA. 2003. Substoichiometric shifting in the plant mitochondrial genome is influenced by a gene homologous to *MutS*. Proceedings of the National Academy of Sciences, USA **100**, 5968–5973.

Allen JO, Fauron CM, Minx P, *et al*. 2007. Comparisons among two fertile and three male-sterile mitochondrial genomes of maize. Genetics **177**, 1173–1192.

Bankevich A, Nurk S, Antipov D, *et al*. 2012. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. Journal of Computational Biology **19**, 455–477.

Chevigny N, Schatz-Daas D, Lofti F, Gualberto JM. 2020. DNA repair and the stability of the plant mitochondrial genome. International Journal of Molecular Sciences **21**, 328.

Christensen AC. 2013. Plant mitochondrial genome evolution can be explained by DNA repair mechanisms. Genome Biology and Evolution **5**, 1079–1086.

Christensen AC. 2021. Plant mitochondria are a riddle wrapped in a mystery inside an enigma. Journal of Molecular Evolution **89**, 151–156.

Cuenca A, Petersen G, Seberg O. 2013. The complete sequence of the mitochondrial genome of *Butomus umbellatus* – a member of an early branching lineage of monocotyledons. PLoS ONE **8**, e61552.

Dierckxsens N, Mardulyn P, Smits G. 2016. NOVOPlasty: *de novo* assembly of organelle genomes from whole genome data. Nucleic Acids Research **45**, e18.

Duminil J, Besnard G. 2021. Utility of the mitochondrial genome in plant taxonomic studies. In: Besse P, ed. Molecular plant taxonomy: methods and protocols. New York: Humana, 107–118.

Ellis J. 1982. Promiscuous DNA—chloroplast genes inside plant mitochondria. Nature **299**, 678–679.

Fischer A, Dotzek J, Walther D, Greiner S. 2022. Graph-based models of the *Oenothera* mitochondrial genome capture the enormous complexity of higher plant mitochondrial DNA organization. NAR Genomics and Bioinformatics **4**, lqac027.

Fujii S, Kazama T, Yamada M, Toriyama K. 2010. Discovery of global genomic re-organization based on comparison of two newly sequenced rice mitochondrial genomes with cytoplasmic male sterility-related genes. BMC Genomics **11**, 209.

Gandini CL, Sanchez-Puerta MV. 2017. Foreign plastid sequences in plant mitochondria are frequently acquired via mitochondrion-to-mitochondrion horizontal transfer. Scientific Reports **7**, 43402.

Goremykin VV, Lockhart PJ, Viola R, Velasco R. 2012. The mitochondrial genome of *Malus domestica* and the import-driven hypothesis of mitochondrial genome expansion in seed plants. The Plant Journal **71**, 615–626.

Goremykin VV, Salamini F, Velasco R, Viola R. 2009. Mitochondrial DNA of *Vitis vinifera* and the issue of rampant horizontal gene transfer. Molecular Biology and Evolution **26**, 99–110.

Gualberto JM, Newton KJ. 2017. Plant mitochondrial genomes: dynamics and mechanisms of mutation. Annual Review of Plant Biology **68**, 225–252.

Iorizzo M, Senalik D, Szklarczyk M, Grzebelus D, Spooner D, Simon P. 2012. *De novo* assembly of the carrot mitochondrial genome using next generation sequencing of whole genomic DNA provides first evidence

of DNA transfer into an angiosperm plastid genome. BMC Plant Biology **12**, 61.

Jin JJ, Yu WB, Yang JB, Song Y, dePamphilis CW, Yi TS, Li DZ. 2020. GetOrganelle: a fast and versatile toolkit for accurate de novo assembly of organelle genomes. Genome Biology **21**, 241.

Kolmogorov M, Yuan J, Lin Y, Pevzner PA. 2019. Assembly of long, error-prone reads using repeat graphs. Nature Biotechnology **37**, 540–546.

Koren S, Walenz BP, Berlin K, Miller JR, Bergman NH, Phillippy AM. 2017. Canu: scalable and accurate long-read assembly via adaptive *k*-mer weighting and repeat separation. Genome Research **27**, 722–736.

Kubo T, Nishizawa S, Sugawara A, Itchoda N, Estiati A, Mikami T. 2000. The complete nucleotide sequence of the mitochondrial genome of sugar beet (*Beta vulgaris* L.) reveals a novel gene for $tRNA^{Cys}$(GCA). Nucleic Acids Research **28**, 2571–2576.

Langmead B, Salzberg SL. 2012. Fast gapped-read alignment with Bowtie 2. Nature Methods **9**, 357–359.

Li H. 2018. Minimap2: pairwise alignment for nucleotide sequences. Bioinformatics **34**, 3094–3100.

Liu Q, Yuan H, Xu J, Cui D, Xiong G, Schwarzacher T, Heslop-Harrison JS. 2023. The mitochondrial genome of the diploid oat *Avena longiglumis*. BMC Plant Biology **23**, 218.

Lu C, Gao LZ, Zhang QJ. 2022. A high-quality genome assembly of the mitochondrial genome of the oil-tea tree *Camellia gigantocarpa* (Theaceae). Diversity **14**, 850.

Matos RG, Silva SR, Płachno BJ, Adamec L, Michael TP, Varani AM, Miranda VFO. 2022. The complete mitochondrial genome of carnivorous *Genlisea tuberosa* (Lentibulariaceae): structure and evolutionary aspects. Gene **824**, 146391.

Mower JP. 2020. Variation in protein gene and intron content among land plant mitogenomes. Mitochondrion **53**, 203–213.

Mower JP, Touzet P, Gummow JS, Delph LF, Palmer JD. 2007. Extensive variation in synonymous substitution rates in mitochondrial genes of seed plants. BMC Evolutionary Biology **7**, 135.

Ni Y, Zhang X, Li J, Lu Q, Chen H, Ma B, Liu C. 2023. Genetic diversity of *Coffea arabica* L. mitochondrial genomes caused by repeat-mediated recombination and RNA editing. Frontiers in Plant Science **14**, 1261012.

Notsu Y, Masood S, Nishikawa T, Kubo N, Akiduki G, Nakazono M, Hirai A, Kadowaki K. 2002. The complete sequence of the rice (*Oryza sativa* L.) mitochondrial genome: frequent DNA sequence acquisition and loss during the evolution of flowering plants. Molecular Genetics and Genomics **268**, 434–445.

Oda K, Yamato K, Ohta E, Nakamura Y, Takemura M, Nozato N, Akashi K, Kanegae T, Ogura Y, Kohchi T. 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.

Oldenburg DJ, Bendich AJ. 2001. Mitochondrial DNA from the liverwort *Marchantia polymorpha*: circularly permuted linear molecules, head-to-tail concatemers, and a 5′ protein. Journal of Molecular Biology **310**, 549–562.

Palmer JD, Herbon LA. 1988. Plant mitochondrial DNA evolves rapidly in structure, but slowly in sequence. Journal of Molecular Evolution **28**, 87–97.

Petersen G, Anderson B, Braun HP, Meyer EH, Møller IM. 2020. Mitochondria in parasitic plants. Mitochondrion **52**, 173–182.

Pevzner PA, Tang H, Tesler G. 2004. *De novo* repeat classification and fragment assembly. In: RECOMB '04: Proceedings of the eighth annual international conference on research in computational molecular biology. New York: ACM Press, 213–222.

Putintseva YA, Bondar EI, Simonov EP, *et al*. 2020. Siberian larch (*Larix sibirica* Ledeb.) mitochondrial genome assembled using both short and long nucleotide sequence reads is currently the largest known mitogenome. BMC Genomics **21**, 654.

Richardson AO, Rice DW, Young GJ, Alverson AJ, Palmer JD. 2013. The 'fossilized' mitochondrial genome of *Liriodendron tulipifera*: ancestral gene content and order, ancestral editing sites, and extraordinarily low mutation rate. BMC Biology **11**, 29.

Ruhlman TA, Jansen RK. 2014. The plastid genomes of flowering plants. In: Maliga P, ed. Chloroplast biotechnology. Methods in molecular biology. Vol. **1132**. Totowa: Humana Press, 3–38.

Satam H, Joshi K, Mangrolia U, *et al*. 2023. Next-generation sequencing technology: current trends and advancements. Biology **12**, 997.

Shedge V, Arrieta-Montiel M, Christensen AC, Mackenzie SA. 2007. Plant mitochondrial recombination surveillance requires unusual *RecA* and *MutS* homologs. Plant Cell **19**, 1251–1264.

Skuza L, Androsiuk P, Gastineau R, Paukszto L, Jastrzębski JP, Cembrowska-Lech D. 2023. Molecular structure, comparative and phylogenetic analysis of the complete chloroplast genome sequences of weedy rye *Secale cereale* ssp. *segetale*. Scientific Reports **13**, 5412.

Sloan DB, Alverson AJ, Chuckalovcak JP, Wu M, McCauley DE, Palmer JD, Taylor DR. 2012a. Rapid evolution of enormous, multichromosomal genomes in flowering plant mitochondria with exceptionally high mutation rates. PLoS Biology **10**, e1001241.

Sloan DB, Alverson AJ, Štorchová H, Palmer JD, Taylor DR. 2010. Extensive loss of translational genes in the structurally dynamic mitochondrial genome of the angiosperm *Silene latifolia*. BMC Evolutionary Biology **10**, 274.

Sloan DB, Müller K, Mccauley DE, Taylor DR, Štorchová H. 2012b. Intraspecific variation in mitochondrial genome sequence, structure, and gene content in *Silene vulgaris*, an angiosperm with pervasive cytoplasmic male sterility. New Phytologist **196**, 1228–1239.

Štorchová H, Stone JD, Sloan DB, Abeyawardana OAJ, Müller K, Walterová J, Pažoutová M. 2018. Homologous recombination changes the context of *Cytochrome b* transcription in the mitochondrial genome of *Silene vulgaris* KRA. BMC Genomics **19**, 874.

Taanman J-W. 1999. The mitochondrial genome: structure, transcription, translation and replication. Biochimica et Biophysica Acta **1410**, 103–123.

Tanaka Y, Tsuda M, Yasumoto K, Yamagishi H, Terachi T. 2012. A complete mitochondrial genome sequence of Ogura-type male-sterile cytoplasm and its comparative analysis with that of normal cytoplasm in radish (*Raphanus sativus* L.). BMC Genomics **13**, 352.

Uliano-Silva M, Ferreira JGRN, Krasheninnikova K, *et al*. 2023. MitoHiFi: a python pipeline for mitochondrial genome assembly from PacBio high fidelity reads. BMC Bioinformatics **24**, 288.

Unseld M, Marienfeld J, Brandt P, Brennicke A. 1997. The mitochondrial genome of *Arabidopsis thaliana* contains 57 genes in 366,924 nucleotides. Nature Genetics **15**, 57–61.

Walker BJ, Abeel T, Shea T, *et al*. 2014. Pilon: an integrated tool for comprehensive microbial variant detection and genome assembly improvement. PLoS One **9**, e112963.

Wick RR, Judd LM, Gorrie CL, Holt KE. 2017. Unicycler: resolving bacterial genome assemblies from short and long sequencing reads. PLoS Computational Biology **13**, e1005595.

Wick RR, Schultz MB, Zobel J, Holt KE. 2015. Bandage: interactive visualization of *de novo* genome assemblies. Bioinformatics **31**, 3350–3352.

Woloszynska M. 2010. Heteroplasmy and stoichiometric complexity of plant mitochondrial genomes—though this be madness, yet there's method in't. Journal of Experimental Botany **61**, 657–671.

Yao Z, You FM, N'Diaye A, Knox RE, McCartney C, Hiebert CW, Pozniak C, Xu W. 2020. Evaluation of variant calling tools for large plant genome re-sequencing. BMC Bioinformatics **21**, 360.

Yu R, Chen X, Long L, Jost M, Zhao R, Liu L, Mower JP, DePamphilis CW, Wanke S, Jiao Y. 2023. *De novo* assembly and comparative analyses of mitochondrial genomes in Piperales. Genome Biology and Evolution **15**, evad041.

Zerbino DR. 2010. Using the Velvet *de novo* assembler for short-read sequencing technologies. Current Protocols in Bioinformatics **31**, 11.5.1–11.5.12.

Zhang Q, Li J, Yu Y, Xu H. 2023. A new map of the chloroplast genome of *Hippophae* based on inter- and intraspecific variation analyses of 13 accessions belonging to eight *Hippophae* species. Brazilian Journal of Botany **46**, 367–382.