diversity

MDPI

Review

Nanopore Sequencing Technology as an Emerging Tool for Diversity Studies of Plant Organellar Genomes

Jakub Sawicki *, Katarzyna Krawczyk, Łukasz Paukszto, Mateusz Maździarz, Mateusz Kurzyński, Joanna Szablińska-Piernik and Monika Szczecińska

Department of Botany and Evolutionary Ecology, University of Warmia and Mazury in Olsztyn, Plac Łódzki 1, 10-1719 Olsztyn, Poland; katarzyna.krawczyk@uwm.edu.pl (K.K.); lukasz.paukszto@uwm.edu.pl (Ł.P.); mateusz.mazdziarz@uwm.edu.pl (M.M.); wodajo@pm.com (M.K.); joanna.szablinska@uwm.edu.pl (J.S.-P.); monika.szczecinska@uwm.edu.pl (M.S.)

* Correspondence: jakub.sawicki@uwm.edu.pl

Abstract: In this comprehensive review, we explore the significant role that nanopore sequencing technology plays in the study of plant organellar genomes, particularly mitochondrial and chloroplast DNA. To date, the application of nanopore sequencing has led to the successful sequencing of over 100 plant mitochondrial genomes and around 80 chloroplast genomes. These figures not only demonstrate the technology's robustness but also mark a substantial advancement in the field, highlighting its efficacy in decoding the complex and dynamic nature of these genomes. Nanopore sequencing, known for its long-read capabilities, significantly surpasses traditional sequencing techniques, especially in addressing challenges like structural complexity and sequence repetitiveness in organellar DNA. This review delves into the nuances of nanopore sequencing, elaborating on its benefits compared to conventional methods and the groundbreaking applications it has fostered in plant organellar genomics. While its transformative impact is clear, the technology's limitations, including error rates and computational requirements, are discussed, alongside potential solutions and prospects for technological refinement.

Keywords: plastome; plant mitogenome; nanopore sequencing; target enrichment; adaptive sampling

Citation: Sawicki, J.; Krawczyk, K.; Paukszto, Ł.; Maździarz, M.; Kurzyński, M.; Szablińska-Piernik, J.; Szczecińska, M. Nanopore Sequencing Technology as an Emerging Tool for Diversity Studies of Plant Organellar Genomes. *Diversity* **2024**, *16*, 173. https://doi.org/10.3390/d16030173

Academic Editor: Michael Wink

Received: 6 February 2024
Revised: 5 March 2024
Accepted: 6 March 2024
Published: 7 March 2024

1. Introduction

Plant organellar genomes, encompassing mitochondrial DNA (mtDNA) and chloroplast DNA (cpDNA), are fundamental to understanding plant cellular functions and evolutionary history. Mitochondrial genomes are central to energy production, while chloroplast genomes are essential for photosynthesis and other biosynthetic activities. Both organelles are hypothesized to have originated from ancient endosymbiotic events [1], a concept further substantiated by similarities in their genome structures and gene content compared to their bacterial ancestors [2].

Plant mitochondrial genomes are characterized by significant variability in size, ranging from 109 kb to over 10 Mb, and structural complexity [3]. This variation is attributed to frequent recombination, horizontal gene transfer, and gene loss [4]. Unlike animal mtDNA, plant mtDNA includes a broader range of genes, encoding elements of the electron transport chain, ribosomal proteins, tRNAs, and unique functional elements [5].

Chloroplast genomes, generally more conserved than mitochondrial genomes, range from 120 to 160 kb and contain genes crucial for photosynthesis, ribosomal proteins, and RNA components [6]. Despite this conservation, cpDNA exhibits structural and content variations among plant lineages, indicative of a complex evolutionary trajectory involving gene duplication and transfer [7].

The exploration of these genomes has been propelled by advancements in sequencing technologies, which has provided deeper insights into their structure and evolution [8].

These studies are vital for understanding plant biology, including cellular processes, phylogeny, evolutionary adaptation, and responses to environmental stresses [9,10], thereby offering avenues for agricultural innovation and conservation strategies.

Recent advancements in molecular biology and genomics have revolutionized the study of plant evolution, barcoding, and phylogeny, particularly through the exploration of chloroplast and mitochondrial genomes. This review focuses on the application of cutting-edge nanopore sequencing technology in these realms, emphasizing its impact on our understanding of plant genetic diversity and evolutionary history.

Nanopore sequencing, characterized by its ability to produce long reads of native DNA, allows for comprehensive genome mapping, including the resolution of repetitive sequences and complex genomic regions. This technology significantly enhances our ability to assemble complete chloroplast and mitochondrial genomes, facilitating detailed comparative genomic studies across diverse plant taxa [11,12]. Moreover, nanopore sequencing has improved the accuracy of plant barcoding, enabling precise species identification and tracing of plant lineage relationships, vital for biodiversity conservation and ecological studies [13].

By integrating recent scientific findings and technological breakthroughs, this review aims to shed light on the growing importance of nanopore sequencing in plant evolutionary studies, providing a roadmap for future research in this dynamic field.

2. Challenges, Limitations, and Recent Improvements in Nanopore Sequencing Technology

Despite the revolutionary impact of nanopore sequencing in plant organellar genomics, the technology is not without its challenges and limitations. A primary concern is the relatively high error rate associated with nanopore sequencing compared to more established techniques like Illumina sequencing. While error rates have been steadily decreasing with technological advancements, they still pose a challenge, particularly for precise applications like single-nucleotide polymorphism (SNP) detection or quantitative gene expression analysis [14]. This necessitates the use of sophisticated bioinformatic tools and error-correction algorithms to ensure data accuracy [15].

Technical challenges also persist, including the need for high-quality, high-molecular-weight DNA samples and the complexity of data management and analysis due to the large volume of data generated. The long-read nature of nanopore sequencing produces significantly larger datasets, requiring substantial computational resources for data processing and storage [16]. Single flow cell of a PromethION device can produce up to 2 Tb raw data, which should be kept in mind during the design of an experiment and the planning of data storage resources. The recently introduced Dorado basecaller for optimal performance and high-quality models with enabled modified base identification requires a top-tier GPU to keep up with real-time data processing [17]. This can be a limiting factor, particularly in low-resource settings or for researchers without access to adequate computational infrastructure. Moreover, while nanopore sequencing is adept at handling larger and more complex genomes, its efficiency can vary depending on the genomic context and the specific organism under study. For example, regions with significantly high GC content or repetitive sequences may still pose challenges, necessitating the integration of other sequencing methods or specialized sample preparation techniques [18,19].

In the last two years, Oxford Nanopore has made significant advancements across multiple areas to improve nanopore sequencing's quality and output. These include introducing Kit14 chemistry that enables modal raw read accuracies above Q20 (99%+) in simplex mode and approaching Q30 (99.9%) in duplex mode [20]. The new R10.4.1 pore further boosts accuracy by incorporating a modified motor protein and optimized run conditions, achieving up to 99.3% raw read accuracy at high yields with Kit14.

Moreover, the percentage of duplex reads constantly increase from 1% in Kit12 (with 10.4 pores) to 10–20% in Kit14 [21,22] and even further using recently developed High-Duplex flow cells (10.4.1 HD)

Basecalling methods have been also advanced through machine learning optimizations, now reaching Q50 for human consensus accuracy and 99% indel accuracy. Key factors powering these improvements are enhanced enzyme processivity, tuned pore dimensions, better current resolution (increased sampling rate from 4 kHz to 5 kHz), and larger training datasets for basecalling algorithms. Altogether, progress in sequencing accuracy achieved in the last two years enabled the assembly of high-quality genomes based exclusively on nanopore technology (Figure 1).

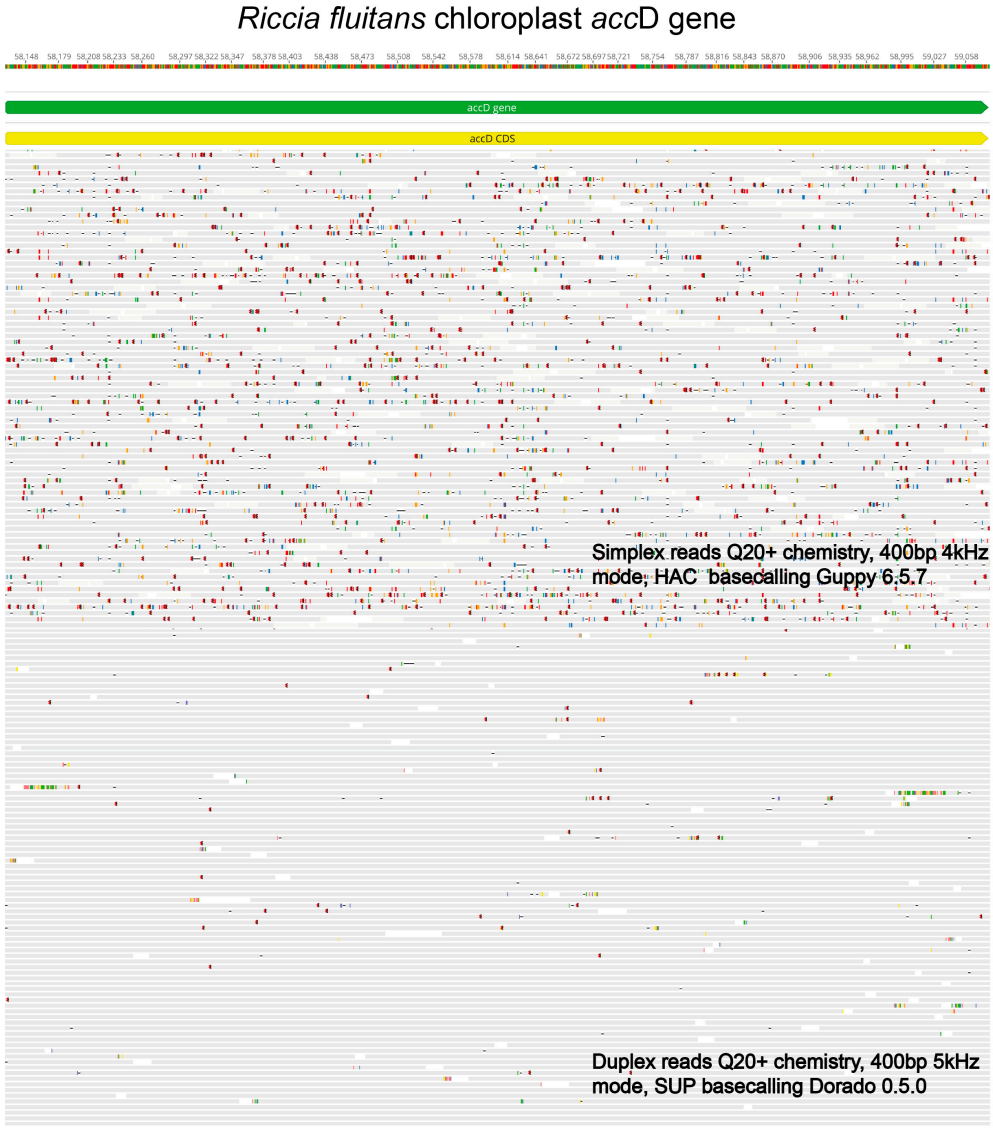

Figure 1. Comparison of mapping simplex and duplex reads onto *acc*D gene of *Riccia fluitans* plastome.

In comparison, Illumina sequencing platforms have long been the gold standard for high-throughput sequencing in terms of accuracy, achieving read accuracies above Q30 (99.9%) across their range of instruments, such as the NovaSeq, HiSeq, and MiSeq series. Illumina's technology is based on sequencing by synthesis (SBS) and is renowned for its high accuracy, especially for short-read sequencing applications. However, the recently introduced short-read Onso platform by PacBio can outperform NovaSeq by promising a delivery of 90% of reads with >Q40.

3. Organellar Genomes Sequenced Using Nanopore Technology

In the investigation of chloroplast genome sequencing employing nanopore technology, a progressive increase in the number of sequenced genomes was observed over the years,

as delineated by the collected data. Commencing in 2018 with a solitary genome, a gradual escalation was evident, with totals of 3, 7, 9, and 8 genomes sequenced in the years 2019, 2020, 2021, and 2022, respectively (Figure 2). Notably, the year 2023 marked a peak in this trend, with 12 chloroplast genomes being sequenced.

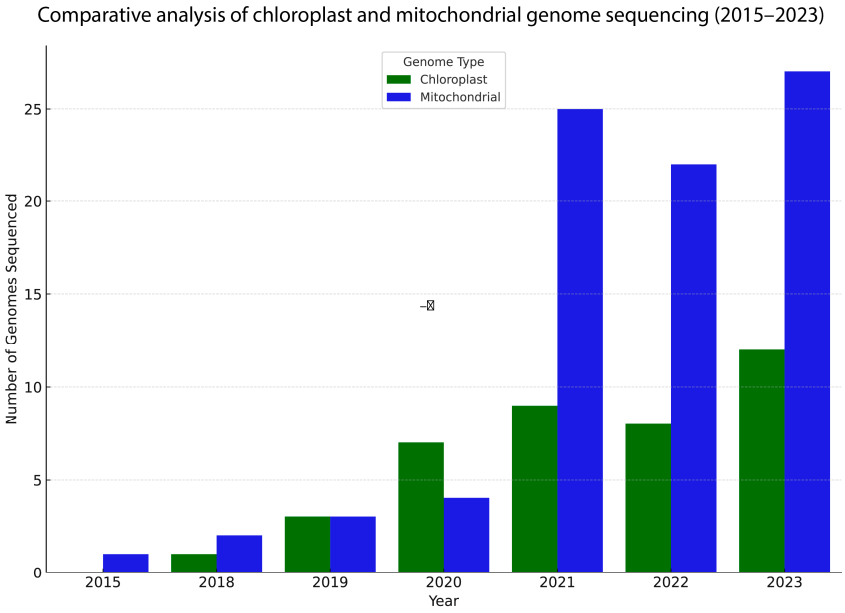

Figure 2. Plant organellar genomes in GenBank database sequenced using nanopore technology.

The temporal analysis of mitochondrial genome (mitogenome) sequencing from 2015 to 2023 reveals a notable escalation in the volume of sequencing endeavors. The inception of this period in 2015 is marked by the sequencing of a solitary mitogenome. A latent period follows until 2018, with a marginal increase of two sequences, and a steady yet modest growth is observed in subsequent years, with counts of three and four in 2019 and 2020, respectively. The year 2021 stands as a pivotal point, witnessing a significant surge to 25 sequenced mitogenomes, indicative of technological advancements and heightened research interest in this domain. Although a slight reduction to 22 sequences occurs in 2022, the upward trajectory is reinstated in 2023 with the sequencing of 27 mitogenomes. This dataset underscores the progressive intensification in mitochondrial genome research, reflecting broader trends in molecular biology and genomics, and highlighting the expanding scope of mitochondrial studies in diverse organisms.

A comparative analysis of the sequencing trends for mitochondrial and chloroplast genomes from 2015 to 2023 reveals distinct patterns and intensities in research trajectories. Mitochondrial genome sequencing exhibits a latent initial phase with minimal activity up to 2018, followed by a gradual increase and a striking surge in 2021, culminating in 27 sequenced mitogenomes by 2023. This pattern suggests a significant technological or research impetus around 2021, possibly reflecting a maturation of nanopore sequencing technology in both wet lab and bioinformatics fields. In contrast, chloroplast genome sequencing, tracked from 2018, shows a more uniform and steady increase year-by-year, starting with a single genome and consistently rising to 12 by 2023.

In plant organellar genomics, nanopore sequencing is often the method of choice for assembling mitochondrial genomes (mitogenomes) due to its unique advantages in addressing the complexities of these genomes, which are generally larger and more intricate than chloroplast genomes, which in most cases can be successfully assembled using a usually low DNA demand and well establishes short-read technologies. Plant mitogenomes are characterized by a high degree of repetitive DNA and significant intergenic regions, alongside frequent recombination, horizontal gene transfers, and structural variations [23]. The long-read capability of nanopore sequencing is particularly beneficial in spanning

across these repetitive sequences, which can be up to 40 kbp long [24,25]. This technology also adeptly resolves structural variations and recombination events, providing a more comprehensive understanding of plant mitochondrial DNA. A recent studies highlights the efficacy of long-read sequencing in unravelling the intricacies of plant mitochondrial genomes, reflecting its growing importance in plant genomics research [26,27].

Most of the studies on plant organellar genomes, performed using older pores or chemistries, also used short-read sequencing to improve the quality of an assembly [25,28,29]. The integration of data from both Oxford Nanopore and Illumina technologies in organellar genome studies leverages the long reads of nanopore for contiguous assembly and the high accuracy of Illumina for base-level detail. This hybrid sequencing strategy has been widely adopted for its synergistic benefits, effectively addressing the unique assembly challenges posed by organellar genomes due to their structure and size variability [25,30,31]. The combined approach not only enhances the quality of genome assemblies but also exemplifies the evolving strategies in genomic research to overcome specific technological limitations [18].

Recent improvement in nanopore sequencing technology enable the elimination of short-read technologies from laboratory pipelines and open new possibilities of assembling high-quality organellar genomes based exclusively on nanopore technology [22,32–35]. This trajectory underscores the growing adoption and efficacy of nanopore technology in the realm of organellar genome sequencing, reflecting its expanding role in molecular biology and genomics research.

4. Plastid and Mitochondrial DNA Extraction and Enrichment

Extracting high-molecular-weight (HMW) DNA from plant tissues presents a suite of challenges that are both unique and demanding, largely due to the complex nature of plant cell walls and the abundance of secondary metabolites. Plant cell walls, composed of cellulose, hemicellulose, and lignin, create a robust barrier that is difficult to penetrate [36]. This necessitates the use of harsh physical or enzymatic methods to disrupt the cell wall processes, which can often result in the shearing of DNA and the reduction in its molecular weight. Furthermore, the high content of polysaccharides and polyphenolic compounds in plant tissues complicates the extraction process [37]. These compounds can co-precipitate with nucleic acids, leading to impure DNA samples that are unsuitable for downstream applications such as long-read sequencing or genome assembly. The presence of mentioned contaminants and undigested RNA can significantly hinder the sequencing process. These substances can clog the nanopores, leading to interruptions in sequencing runs, reduced efficiency, and increased sequencing error rates. Polysaccharides and polyphenolic compounds, often co-extracted with DNA from biological samples, can adhere to DNA molecules and interfere with their passage through the nanopore, while RNA, if not adequately removed during the DNA purification process, can similarly block nanopores and compete with DNA for sequencing, thereby diluting the focus on target DNA sequences [38]. This pore clogging not only affects the throughput of the sequencing process but also compromises the quality of the sequencing data, making it challenging to achieve a high quality.

The presence of these secondary metabolites not only interferes with the purity and integrity of the extracted DNA but also with the efficacy of enzymatic reactions in library preparation process, as they can inhibit enzymes critical to DNA repair process and ligation steps. Moreover, the wide diversity of plant species, each with its unique set of secondary metabolites and cell wall compositions, means that a one-size-fits-all approach to HMW DNA extraction is often impractical. Developing species-specific protocols or adapting existing ones to mitigate these issues requires extensive empirical optimization, further complicating the extraction process.

Another challenge in extracting HMW DNA from plant tissues is the requirement to maintain the integrity of the DNA during extraction. High-molecular-weight DNA is particularly susceptible to mechanical shearing forces introduced during tissue homoge-

nization and liquid handling. Thus, gentle and carefully optimized protocols are essential to preserve the long fragments necessary for certain types of genomic analyses. The combination of these challenges necessitates innovative approaches and methodologies for HMW DNA extraction from plant tissues, emphasizing the need for continued research and development in this area to facilitate advancements in plant genomics and breeding programs. Researchers are continually seeking to refine and develop methods that can overcome these obstacles, ensuring the extraction of high-quality, high-molecular-weight DNA that is essential for the comprehensive analysis and understanding of plant genomes [39].

The enrichment of plastid and mitochondrial DNA prior to sequencing is an extremely beneficial step in plant genomics, addressing several pivotal challenges inherent to the sequencing process. In plant cells, organelle DNA represents a relatively minor fraction of the total cellular DNA, with chloroplast DNA typically constituting about 0.5% to 1% and mitochondrial DNA ranging from 0.1% to 1% of the total cellular DNA content, depending on the tissue type and developmental stage [40]. This disparity underscores the importance of reducing the proportion of nuclear DNA in sequencing samples, as unenriched samples can lead to the inadequate representation of organelle genomes. The challenge is further compounded by the presence of nuclear sequences resembling organelle DNA, known as nuclear mitochondrial DNA segments (NUMTs) and nuclear plastid DNA (NUPTs), which can result in misassemblies and incorrect genomic interpretations if not adequately filtered out [41].

From a cost-effectiveness perspective, the financial implications of high-throughput sequencing necessitate the maximization of useful data output. Enrichment enables a higher proportion of sequencing reads to be directly relevant to the targeted organelle genomes, thus optimizing the sequencing efforts and ensuring more comprehensive coverage without the need for prohibitively deep sequencing of the entire sample. This approach is especially crucial in large-scale or multi-sample studies, where the cost savings can be substantial [42].

The enhancement of data quality and accuracy is another crucial benefit of organelle DNA enrichment. In the context of organelle genomics, the integrity of the sequence data is paramount, particularly given the complex evolutionary history and genomic mosaicism characteristic of plant organelles. Enrichment minimizes the inclusion of non-target DNA, thus reducing the risk of chimeric sequence assembly and misinterpretation. This is critical in research focused on understanding the evolutionary dynamics, gene content, and functional roles of organelle genomes. The accuracy of organelle genome assembly is essential for deciphering their evolutionary trajectories and functional contributions to plant physiology and development [43].

Enrichment of organellar genome DNA for nanopore sequencing can be performed using PCR-based methods or by capturing native plastid or mitochondrial DNA using bead-based probes [44,45].

Amplifying the entire chloroplast genome using Long-Range PCR (LR-PCR) often involves the use of universal primers, which are designed to bind to conserved regions across a wide range of plant species [46]. This approach, exemplified by studies like Goremykin et al. [47] and Raubeson et al. [48], is advantageous for its broad applicability. Universal primers facilitate the amplification of large sections of the chloroplast genome in diverse plant groups, from angiosperms to gymnosperms. The use of these primers simplifies the process of amplifying the entire chloroplast genome by reducing the need for species-specific primer design, making LR-PCR more accessible and efficient. However, the reliance on universal primers also brings challenges. While they are designed to target conserved regions, variations in these regions across different species can lead to inconsistent amplification results. Furthermore, the optimization of LR-PCR conditions remains complex, as it requires careful calibration to ensure efficient and accurate amplification of large DNA fragments. The subsequent assembly of these overlapping fragments, especially in genomes with high repetitive content, adds another layer of complexity. This approach can be improved by tuning the primers toward being taxon-specific, narrowing the tax-

onomic groups, but increasing the successful amplification rate and completeness of the assembly [31].

Among various library preparation protocols dealing mainly with native DNA, options for low-input samples are also included, comprising whole genome amplification and low input by PCR protocols, and offering two alternative approaches to sample for amplification. Since organellar DNA is usually obtained in high yields in both PCR and rolling circle amplification (RCA) methods, the increase in cpDNA and mtDNA in the final sequencing result is often 2–3 fold higher in comparison to native DNA sequencing. The downside of this approach, besides the loss of modification signals, is inconsistent coverage of assembled genomes (Figure 3). Both RCA and EXP-PCA001 kits do not deliver coverage smoothness as provided by native DNA sequencing. Another issue of PCR-based enrichment methods is the possibility of random introduction of errors during the DNA synthesis and generation of false polymorphism [49]. This phenomenon occurs because the DNA polymerases used in PCR, despite their high fidelity, are not infallible and can incorporate incorrect nucleotides during DNA synthesis. Over multiple cycles of amplification, these errors can accumulate, leading to mutations that were not present in the original DNA template. This aspect of PCR-based methods can result in the generation of false polymorphisms, where variant sequences appear to exist but are in fact artifacts of the amplification process rather than genuine genetic variations present in the sample [50].

For many plant samples or herbarium specimens, the availability of tissue is limited; therefore, the application of these protocols may enable complete organellar genomes sequences even from small amounts of input DNA like 1–10 ng. PCR-enriched libraries, despite the above issues, usually generate higher yield from flow cells than native DNA-based libraries, even without mid-run flushing.

Enriching organellar DNA using magnetic beads is a refined and effective method in plant genomics, characterized by its specificity and the saving of the epigenetic signal of native nucleotides. In this process, magnetic beads coated with specific ligands or antibodies selectively bind to total DNA. After extracting total DNA from plant tissues, the DNA is incubated with these prepared beads, allowing for the selective binding of cpDNA and/or mtDNA. A magnetic field is then applied to separate the bead-bound cpDNA/mtDNA from nuclear DNA, followed by the elution of the organellar DNA. This method offers the rapid processing of samples with a high level of purity and specificity. However, it requires a significant investment in both the cost of specialized beads and technical expertise [12,44].

Although the custom enrichment kits are relatively expensive, the standard microbial DNA enrichment kits work well with both plastid and mitochondrial DNA [22,51].

This kit achieves enrichment by selectively binding and removing the CpG-methylated nuclear DNA using MBD2-Fc protein prebound to magnetic beads. Using this approach enabled an increase in the organellar DNA fraction by up to 50% in some liverworts species [22]. The use of bead-based methods for organellar DNA enrichment, while effective in certain contexts, has notable drawbacks, particularly concerning the large amount of input DNA required and the low recovery of targeted DNA. Firstly, the necessity for a substantial quantity of input DNA can be a significant limitation, especially when working with samples that are small, rare, or difficult to collect. This requirement often renders the method impractical for studies involving endangered species and small plant samples or herbarium specimens, where the DNA quantity is inherently limited.

In addition to the experimental enrichment of plastid and mitochondrial DNA, the recently developed nanopore adaptive sampling (NAS) technology offers a method to enhance specific preselected sequences during the sequencing process. Adaptive sampling, also known as “Read Until”, is a feature that enables real-time, selective sequencing of DNA molecules as they pass through the nanopore. This technology allows the sequencing device to actively decide whether to continue reading a particular DNA strand or eject it in favor of a different fragment [52,53].

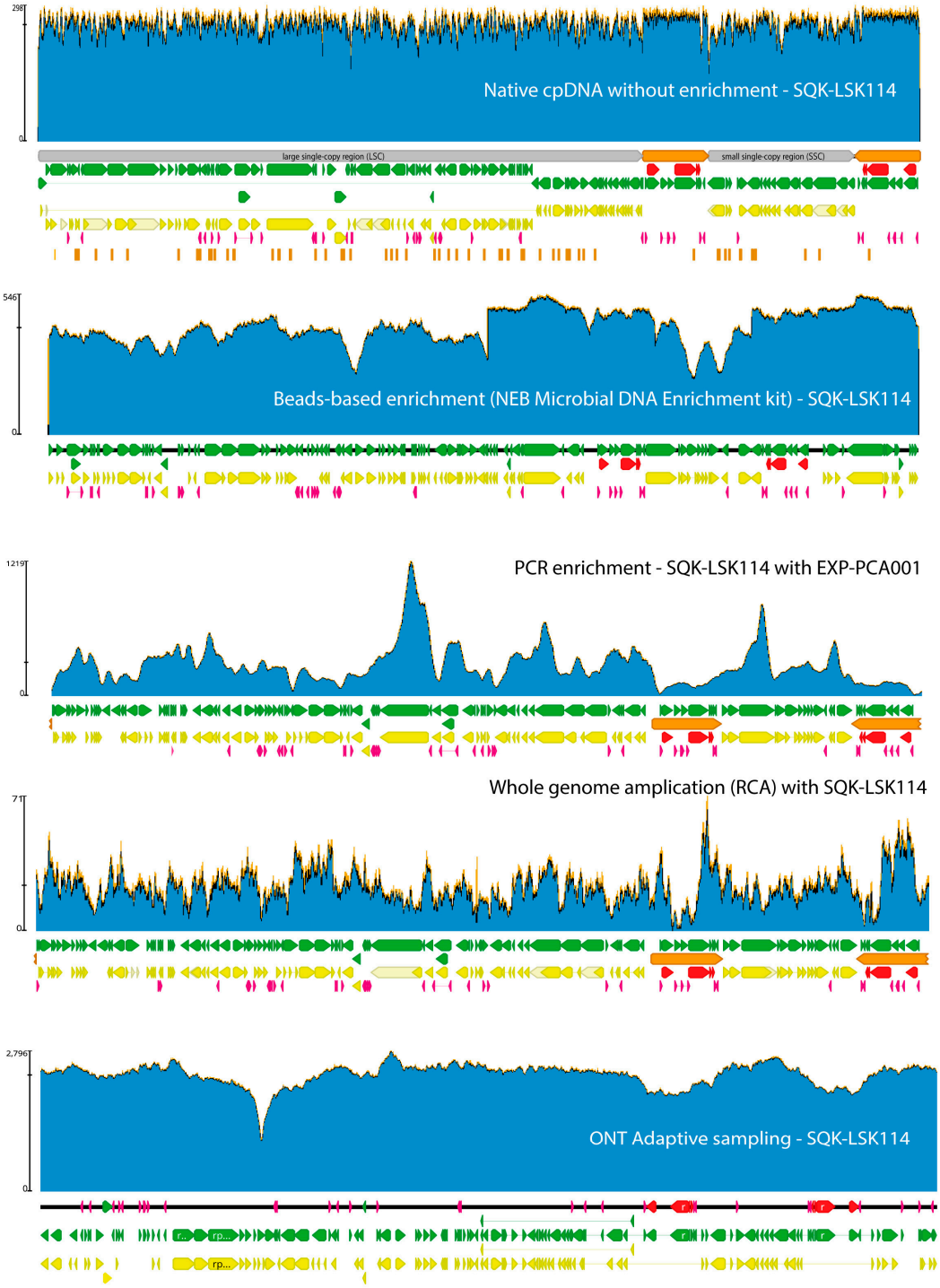

Figure 3. Coverage fluctuation of plastid genome assemblies using native and enriched DNA. The various bryophyte plastid genomes of similar lengths (ca 120 kbp) were sequenced using different approaches. From top to bottom: *Apopellia endiviifolia* (without enrichment), *Scapania undulata* (with organellar DNA enriched using bead-based method), *Orthotrichum cupulatum* (total DNA enrichment via standard PCR), *Riccia fluitans* (DNA enrichment using rolling circle amplification), and *Riccia glauca* (adaptative sampling using *Riccia fluitans* plastome as a reference).

This advanced technique has been effectively utilized to acquire complete mitochondrial genomes from endangered species by enriching the genomic DNA in fecal samples of the host animals. Interestingly, this method can work with bait sequences that have as

low as 70% sequence identity to the target, indicating its potential utility in sequencing unknown species using reference sequences from distantly related organisms [54].

To understand its effectiveness, it is crucial to compare its performance with traditional nanopore sequencing methods. The research conducted in this field demonstrates that NAS is quite efficient at enriching specific reference sequences [55–58]. The enrichment levels observed in these studies vary, showing an increase from as little as 0.96 times to as much as 5.4 times when compared to standard sequencing experiments. This targeted approach is particularly advantageous for organellar genome sequencing, where specific regions of interest can be enriched in the sequencing data, thereby enhancing the efficiency and depth of coverage for mitochondrial and chloroplast genomes.

The application of adaptive sampling in organellar genomics creates exciting possibilities. For instance, it can facilitate the selective enrichment of mitochondrial and chloroplast DNA in mixed samples, effectively bypassing the predominance of nuclear DNA. This selective sequencing is especially beneficial in studies where organellar genomes are present in low copy numbers relative to nuclear DNA, a common scenario in plant cells. By focusing sequencing efforts on these organelles, researchers can achieve higher coverage and more detailed insights into their structure and function with reduced sequencing effort and cost. However, the availability and quality of input material still play a main role in sample processing workflows (Figure 4).

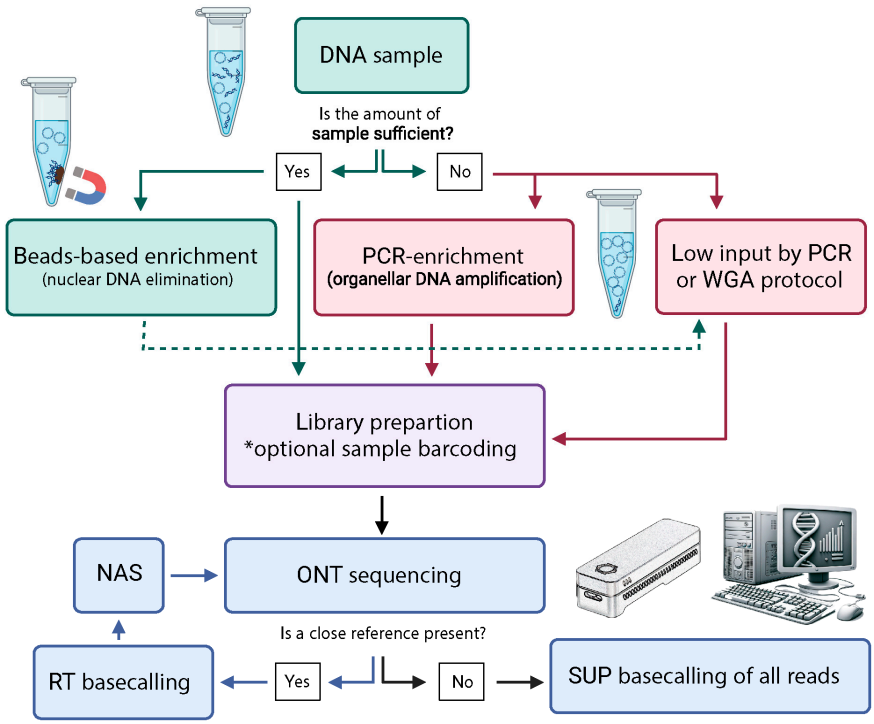

Figure 4. Possible workflows in sequencing plant organellar genomes using nanopore technology. Asterisk indicates optional samples multiplexing.

5. Dedicated Long-Reads Assemblers for Organellar Genomes

Currently, there are several tools available for assembling organelle genomes using long-read and hybrid data, such as organelle_pba [59], canu [60], unicycler [61], and flye [62]. However, these tools have specific limitations. organelle_pba, designed for PacBio data, is constrained as its Sprai [63] and Celera [64] assemblers are no longer updated, hindering its use with hybrid datasets. Syme et al.'s [65] method necessitates an additional step of manually filtering raw reads for the plastome, and it can result in multiple contigs. Canu's outcomes vary with different read coverages [30], and the unicycler, tailored for hybrid data, becomes significantly slower with larger input. Alternatively, a "fishing

approach" using either shasta for nanopore data [66] or hifiasm for PacBio data [67] could be employed to first assemble raw reads, followed by extracting plastid contigs using a reference genome. However, this method is time-consuming and resource-intensive for large datasets, and the accuracy of the assembly might be compromised by redundant sequences [68].

Recently published ptGAUL [69] stands out in the field of plastome assembly, offering accurate results in under 10 min using approximately 16 Gb of memory for datasets smaller than 10 Gb. This efficiency positions it as a notably faster option compared to previously published tools; however, for an optimal performance, it requires a relatively high N50. As a result, ptGAUL is particularly useful for phylogenetic and molecular evolutionary studies involving plastomes with extensively long repeat regions. However, it is important for researchers to remain vigilant when dealing with species that have multiple plastome types, such as *Eleocharis* [70] and *Monsonia* [71], even when using ptGAUL for plastome assembly.

6. Conclusions

In this review, we have explored the emerging role of nanopore sequencing technology in the study of plant organellar genomes, delving into its advantages, challenges, and future potential. Nanopore sequencing, with its capability to generate long reads, has shown immense promise in deciphering the complex and diverse structures of mitochondrial and chloroplast DNA in plants. This technology has addressed some of the longstanding limitations of short-read sequencing, allowing for a more comprehensive understanding of organellar genome architecture, evolutionary dynamics, and functional roles. Despite its transformative impact, challenges such as error rates and computational demands persist, necessitating ongoing technological and bioinformatic advancements.

The introduction of adaptive sampling technology by Oxford Nanopore marks a pivotal advancement in targeted sequencing, offering the potential to revolutionize organellar genomics further. By enabling the selective sequencing of specific genomic regions, this approach can significantly enhance the efficiency and depth of organellar genome analysis. As the technology continues to evolve, its application in plant organellar genomics is poised to provide critical insights into plant biology, contributing to advancements in crop improvement, conservation, and our understanding of plant evolution and diversity.

Overall, nanopore sequencing technology stands as a cornerstone in modern plant genomics, offering a powerful tool for exploring the intricate world of plant organelles. Its ongoing development is not just a testimony to technological innovation but also a beacon for future discoveries in plant science and beyond. As we continue to harness and refine this technology, we edge closer to unraveling the full complexity and potential of plant organellar genomes, paving the way for groundbreaking research in plant biology and biotechnology.

Author Contributions: Conceptualization, J.S. and M.S.; methodology, J.S.; software, J.S.; writing—original draft preparation, J.S., K.K., Ł.P., M.M., M.K., J.S.-P. and M.S.; writing—review and editing, J.S., K.K., Ł.P., M.M., M.K., J.S.-P. and M.S.; visualization, J.S. and K.K.; supervision, J.S.; project administration, J.S.; funding acquisition, J.S. All authors have read and agreed to the published version of the manuscript.

Funding: Figures 2 and 3 were prepared from sequencing experiments financially supported by The National Science Center Kraków, Poland (Grant No. 2020/39/B/NZ8/02504).

Institutional Review Board Statement: Not applicable.

Data Availability Statement: Not applicable.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Gray, M.W. The Endosymbiont Hypothesis Revisited. *Int. Rev. Cytol.* **1992**, *141*, 233–357. [CrossRef] [PubMed]
2. Reconstructing Evolution: Gene Transfer from Plastids to the Nucleus—Bock—2008—BioEssays—Wiley Online Library. Available online: https://onlinelibrary.wiley.com/doi/10.1002/bies.20761 (accessed on 1 January 2024).
3. Palmer, J.D.; Herbon, L.A. Plant Mitochondrial DNA Evolves Rapidly in Structure, but Slowly in Sequence. *J. Mol. Evol.* **1988**, *28*, 87–97. [CrossRef]
4. Mower, J.P.; Stefanović, S.; Hao, W.; Gummow, J.S.; Jain, K.; Ahmed, D.; Palmer, J.D. Horizontal Acquisition of Multiple Mitochondrial Genes from a Parasitic Plant Followed by Gene Conversion with Host Mitochondrial Genes. *BMC Biol.* **2010**, *8*, 150. [CrossRef] [PubMed]
5. Knoop, V. When You Can't Trust the DNA: RNA Editing Changes Transcript Sequences. *Cell. Mol. Life Sci.* **2011**, *68*, 567–586. [CrossRef] [PubMed]
6. Green, B.R. Chloroplast Genomes of Photosynthetic Eukaryotes. *Plant J.* **2011**, *66*, 34–44. [CrossRef] [PubMed]
7. Jansen, R.K.; Cai, Z.; Raubeson, L.A.; Daniell, H.; dePamphilis, C.W.; Leebens-Mack, J.; Müller, K.F.; Guisinger-Bellian, M.; Haberle, R.C.; Hansen, A.K.; et al. Analysis of 81 Genes from 64 Plastid Genomes Resolves Relationships in Angiosperms and Identifies Genome-Scale Evolutionary Patterns. *Proc. Natl. Acad. Sci. USA* **2007**, *104*, 19369–19374. [CrossRef] [PubMed]
8. The First 50 Plant Genomes—Michael—2013—The Plant Genome—Wiley Online Library. Available online: https://acsess.onlinelibrary.wiley.com/doi/10.3835/plantgenome2013.03.0001in (accessed on 14 January 2024).
9. Sugiura, M. The Chloroplast Genome. *Plant Mol. Biol.* **1992**, *19*, 149–168. [CrossRef]
10. Smith, D.R. Plastid Genomes Hit the Big Time. *New Phytol.* **2018**, *219*, 491–495. [CrossRef]
11. Wang, W.; Lanfear, R. Long-Reads Reveal That the Chloroplast Genome Exists in Two Distinct Versions in Most Plants. *Genome Biol. Evol.* **2019**, *11*, 3372–3381. [CrossRef]
12. Wang, Y.; Zhao, Y.; Bollas, A.; Wang, Y.; Au, K.F. Nanopore Sequencing Technology, Bioinformatics and Applications. *Nat. Biotechnol.* **2021**, *39*, 1348–1365. [CrossRef]
13. Prudnikow, L.; Pannicke, B.; Wünschiers, R. A Primer on Pollen Assignment by Nanopore-Based DNA Sequencing. *Front. Ecol. Evol.* **2023**, *11*, 1112929. [CrossRef]
14. Jain, M.; Koren, S.; Miga, K.H.; Quick, J.; Rand, A.C.; Sasani, T.A.; Tyson, J.R.; Beggs, A.D.; Dilthey, A.T.; Fiddes, I.T.; et al. Nanopore Sequencing and Assembly of a Human Genome with Ultra-Long Reads. *Nat. Biotechnol.* **2018**, *36*, 338–345. [CrossRef] [PubMed]
15. Wick, R.R.; Judd, L.M.; Holt, K.E. Performance of Neural Network Basecalling Tools for Oxford Nanopore Sequencing. *Genome Biol.* **2019**, *20*, 129. [CrossRef] [PubMed]
16. Amarasinghe, S.L.; Su, S.; Dong, X.; Zappia, L.; Ritchie, M.E.; Gouil, Q. Opportunities and Challenges in Long-Read Sequencing Data Analysis. *Genome Biol.* **2020**, *21*, 30. [CrossRef]
17. Samarakoon, H.; Ferguson, J.M.; Gamaarachchi, H.; Deveson, I.W. Accelerated Nanopore Basecalling with SLOW5 Data Format. *Bioinformatics* **2023**, *39*, btad352. [CrossRef] [PubMed]
18. Henry, R.J. Progress in Plant Genome Sequencing. *Appl. Biosci.* **2022**, *1*, 113–128. [CrossRef]
19. Pucker, B.; Irisarri, I.; de Vries, J.; Xu, B. Plant Genome Sequence Assembly in the Era of Long Reads: Progress, Challenges and Future Directions. *Quant. Plant Biol.* **2022**, *3*, e5. [CrossRef]
20. Sanderson, N.D.; Kapel, N.; Rodger, G.; Webster, H.; Lipworth, S.; Street, T.L.; Peto, T.; Crook, D.; Stoesser, N. Comparison of R9.4.1/Kit10 and R10/Kit12 Oxford Nanopore Flowcells and Chemistries in Bacterial Genome Reconstruction. *Microb. Genom.* **2023**, *9*, 000910. [CrossRef]
21. Ni, Y.; Liu, X.; Simeneh, Z.M.; Yang, M.; Li, R. Benchmarking of Nanopore R10.4 and R9.4.1 Flow Cells in Single-Cell Whole-Genome Amplification and Whole-Genome Shotgun Sequencing. *Comput. Struct. Biotechnol. J.* **2023**, *21*, 2352–2364. [CrossRef]
22. Sawicki, J.; Krawczyk, K.; Kurzyński, M.; Maździarz, M.; Paukszto, Ł.; Sulima, P.; Szczecińska, M. Nanopore Sequencing of Organellar Genomes Revealed Heteroplasmy in Simple Thalloid and Leafy Liverworts. *Acta Soc. Bot. Pol.* **2023**, *92*, 172516. [CrossRef]
23. Kozik, A.; Rowan, B.A.; Lavelle, D.; Berke, L.; Schranz, M.E.; Michelmore, R.W.; Christensen, A.C. The Alternative Reality of Plant Mitochondrial DNA: One Ring Does Not Rule Them All. *PLoS Genet.* **2019**, *15*, e1008373. [CrossRef] [PubMed]
24. Wynn, E.L.; Christensen, A.C. Repeats of Unusual Size in Plant Mitochondrial Genomes: Identification, Incidence and Evolution. *G3 GenesGenomesGenetics* **2019**, *9*, 549–559. [CrossRef] [PubMed]
25. Szandar, K.; Krawczyk, K.; Myszczyński, K.; Ślipiko, M.; Sawicki, J.; Szczecińska, M. Breaking the Limits—Multichromosomal Structure of an Early Eudicot Pulsatilla Patens Mitogenome Reveals Extensive RNA-Editing, Longest Repeats and Chloroplast Derived Regions among Sequenced Land Plant Mitogenomes. *BMC Plant Biol.* **2022**, *22*, 109. [CrossRef] [PubMed]
26. Zou, Y.; Zhu, W.; Sloan, D.B.; Wu, Z. Long-Read Sequencing Characterizes Mitochondrial and Plastid Genome Variants in Arabidopsis Msh1 Mutants. *Plant J. Cell Mol. Biol.* **2022**, *112*, 738–755. [CrossRef]
27. Masutani, B.; Arimura, S.; Morishita, S. Investigating the Mitochondrial Genomic Landscape of Arabidopsis Thaliana by Long-Read Sequencing. *PLoS Comput. Biol.* **2021**, *17*, e1008597. [CrossRef] [PubMed]

28. Sawicki, J.; Bączkiewicz, A.; Buczkowska, K.; Górski, P.; Krawczyk, K.; Mizia, P.; Myszczyński, K.; Ślipiko, M.; Szczecińska, M. The Increase of Simple Sequence Repeats during Diversification of Marchantiidae, An Early Land Plant Lineage, Leads to the First Known Expansion of Inverted Repeats in the Evolutionarily-Stable Structure of Liverwort Plastomes. *Genes* **2020**, *11*, 299. [CrossRef]
29. Liu, J.; Ni, Y.; Liu, C. Polymeric Structure of the Cannabis Sativa L. Mitochondrial Genome Identified with an Assembly Graph Model. *Gene* **2023**, *853*, 147081. [CrossRef]
30. Wang, W.; Schalamun, M.; Morales-Suarez, A.; Kainer, D.; Schwessinger, B.; Lanfear, R. Assembly of Chloroplast Genomes with Long- and Short-Read Data: A Comparison of Approaches Using Eucalyptus Pauciflora as a Test Case. *BMC Genom.* **2018**, *19*, 977. [CrossRef]
31. Scheunert, A.; Dorfner, M.; Lingl, T.; Oberprieler, C. Can We Use It? On the Utility of de Novo and Reference-Based Assembly of Nanopore Data for Plant Plastome Sequencing. *PLoS ONE* **2020**, *15*, e0226234. [CrossRef]
32. The First Genome for the Cape Primrose Streptocarpus Rexii (Gesneriaceae), a Model Plant for Studying Meristem-Driven Shoot Diversity—Nishii—2022—Plant Direct—Wiley Online Library. Available online: https://onlinelibrary.wiley.com/doi/10.1002/pld3.388 (accessed on 1 January 2024).
33. Plášek, V.; Sawicki, J.; Seppelt, R.D.; Cave, L.H. Orthotrichum Cupulatum Hoffm. Ex Brid. Var. Lithophilum, a New Variety of Epilithic Bristle Moss from Tasmania. *Acta Soc. Bot. Pol.* **2023**, *92*, 1–8. [CrossRef]
34. Ciborowski, K.; Skierkowski, B.; Żukowska, K.; Krawczyk, K.; Sawicki, J. Nanopore Sequencing of Chloroplast Genome of Scapania Undulata (L.) Dumort., 1835 (Scapaniaceae, Jungermanniales). *Mitochondrial DNA Part B Resour.* **2022**, *7*, 1424–1426. [CrossRef]
35. López, K.E.R.; Armijos, C.E.; Parra, M.; Torres, M.d.L. The First Complete Chloroplast Genome Sequence of Mortiño (*Vaccinium floribundum*) and Comparative Analyses with Other *Vaccinium* Species. *Horticulturae* **2023**, *9*, 302. Available online: https://www.mdpi.com/2311-7524/9/3/302 (accessed on 15 January 2024). [CrossRef]
36. Fulton, T.M.; Chunwongse, J.; Tanksley, S.D. Microprep Protocol for Extraction of DNA from Tomato and Other Herbaceous Plants. *Plant Mol. Biol. Rep.* **1995**, *13*, 207–209. [CrossRef]
37. Porebski, S.; Bailey, L.G.; Baum, B.R. Modification of a CTAB DNA Extraction Protocol for Plants Containing High Polysaccharide and Polyphenol Components. *Plant Mol. Biol. Rep.* **1997**, *15*, 8–15. [CrossRef]
38. Russo, A.; Mayjonade, B.; Frei, D.; Potente, G.; Kellenberger, R.T.; Frachon, L.; Copetti, D.; Studer, B.; Frey, J.E.; Grossniklaus, U.; et al. Low-Input High-Molecular-Weight DNA Extraction for Long-Read Sequencing From Plants of Diverse Families. *Front. Plant Sci.* **2022**, *13*, 1494. [CrossRef]
39. Kang, M.; Chanderbali, A.; Lee, S.; Soltis, D.E.; Soltis, P.S.; Kim, S. High-Molecular-Weight DNA Extraction for Long-Read Sequencing of Plant Genomes: An Optimization of Standard Methods. *Appl. Plant Sci.* **2023**, *11*, e11528. [CrossRef]
40. Bock, R. The Give-and-Take of DNA: Horizontal Gene Transfer in Plants. *Trends Plant Sci.* **2010**, *15*, 11–22. [CrossRef]
41. Smith, D.R.; Keeling, P.J. Mitochondrial and Plastid Genome Architecture: Reoccurring Themes, but Significant Differences at the Extremes. *Proc. Natl. Acad. Sci. USA* **2015**, *112*, 10177–10184. [CrossRef] [PubMed]
42. Bensasson, D.; Zhang, D.-X.; Hartl, D.L.; Hewitt, G.M. Mitochondrial Pseudogenes: Evolution's Misplaced Witnesses. *Trends Ecol. Evol.* **2001**, *16*, 314–321. [CrossRef] [PubMed]
43. Sloan, D.B.; Alverson, A.J.; Chuckalovcak, J.P.; Wu, M.; McCauley, D.E.; Palmer, J.D.; Taylor, D.R. Rapid Evolution of Enormous, Multichromosomal Genomes in Flowering Plant Mitochondria with Exceptionally High Mutation Rates. *PLoS Biol.* **2012**, *10*, e1001241. [CrossRef]
44. Mariac, C.; Scarcelli, N.; Pouzadou, J.; Barnaud, A.; Billot, C.; Faye, A.; Kougbeadjo, A.; Maillol, V.; Martin, G.; Sabot, F.; et al. Cost-effective Enrichment Hybridization Capture of Chloroplast Genomes at Deep Multiplexing Levels for Population Genetics and Phylogeography Studies. *Mol. Ecol. Resour.* **2014**, *14*, 1103–1113. [CrossRef]
45. Wang, J.; Mu, W.; Yang, T.; Song, Y.; Hou, Y.G.; Wang, Y.; Gao, Z.; Liu, X.; Liu, H.; Zhao, H. Targeted Enrichment of Novel Chloroplast-Based Probes Reveals a Large-Scale Phylogeny of 412 Bamboos. *BMC Plant Biol.* **2021**, *21*, 76. [CrossRef]
46. Uribe-Convers, S.; Duke, J.R.; Moore, M.J.; Tank, D.C. A Long PCR–Based Approach for DNA Enrichment Prior to next-Generation Sequencing for Systematic Studies. *Appl. Plant Sci.* **2014**, *2*, 1300063. [CrossRef]
47. Goremykin, V.V.; Viola, R.; Hellwig, F.H. Removal of Noisy Characters from Chloroplast Genome-Scale Data Suggests Revision of Phylogenetic Placements of Amborella and Ceratophyllum. *J. Mol. Evol.* **2009**, *68*, 197–204. [CrossRef]
48. Raubeson, L.A.; Peery, R.; Chumley, T.W.; Dziubek, C.; Fourcade, H.M.; Boore, J.L.; Jansen, R.K. Comparative Chloroplast Genomics: Analyses Including New Sequences from the Angiosperms Nuphar Advena and Ranunculus Macranthus. *BMC Genomics* **2007**, *8*, 174. [CrossRef]
49. Potapov, V.; Ong, L.J. Examining Sources of Error in PCR by Single-Molecule Sequencing. *PLoS ONE* **2017**, *12*, e0169774. [CrossRef] [PubMed]
50. Acinas, G.A.; Sarma-Rupavtarm, R.; Klepac-Ceraj, V.; Polz, M.F. PCR-Induced Sequence Artifacts and Bias: Insights from Comparison of Two 16S rRNA Clone Libraries Constructed from the Same Sample. *App. Environ. Microb.* **2005**, *71*, 8966–8969. [CrossRef]
51. Yigit, E.; Hernandez, D.I.; Trujillo, J.T.; Dimalanta, E.; Bailey, C.D. Genome and Metagenome Sequencing: Using the Human Methyl-Binding Domain to Partition Genomic DNA Derived from Plant Tissues. *Appl. Plant Sci.* **2014**, *2*, 1400064. [CrossRef] [PubMed]

52. Payne, A.; Holmes, N.; Clarke, T.; Munro, R.; Debebe, B.J.; Loose, M. Readfish Enables Targeted Nanopore Sequencing of Gigabase-Sized Genomes. *Nat. Biotechnol.* **2021**, *39*, 442–450. [CrossRef] [PubMed]
53. Loose, M.; Malla, S.; Stout, M. Real-Time Selective Sequencing Using Nanopore Technology. *Nat. Methods* **2016**, *13*, 751–754. [CrossRef]
54. Wanner, N.; Larsen, P.A.; McLain, A.; Faulk, C. The Mitochondrial Genome and Epigenome of the Golden Lion Tamarin from Fecal DNA Using Nanopore Adaptive Sequencing. *BMC Genom.* **2021**, *22*, 726. [CrossRef] [PubMed]
55. Gan, M.; Wu, B.; Yan, G.; Li, G.; Sun, L.; Lu, G.; Zhou, W. Combined Nanopore Adaptive Sequencing and Enzyme-Based Host Depletion Efficiently Enriched Microbial Sequences and Identified Missing Respiratory Pathogens. *BMC Genom.* **2021**, *22*, 732. [CrossRef] [PubMed]
56. Martin, S.; Heavens, D.; Lan, Y.; Horsfield, S.; Clark, M.D.; Leggett, R.M. Nanopore Adaptive Sampling: A Tool for Enrichment of Low Abundance Species in Metagenomic Samples. *Genome Biol.* **2022**, *23*, 11. [CrossRef] [PubMed]
57. Kipp, E.J.; Lindsey, L.L.; Milstein, M.S.; Blanco, C.M.; Baker, J.P.; Faulk, C.; Oliver, J.D.; Larsen, P.A. Nanopore Adaptive Sampling for Targeted Mitochondrial Genome Sequencing and Bloodmeal Identification in Hematophagous Insects. *Parasit. Vectors* **2023**, *16*, 68. [CrossRef] [PubMed]
58. Weilguny, L.; De Maio, N.; Munro, R.; Manser, C.; Birney, E.; Loose, M.; Goldman, N. Dynamic, Adaptive Sampling during Nanopore Sequencing Using Bayesian Experimental Design. *Nat. Biotechnol.* **2023**, *41*, 1018–1025. [CrossRef] [PubMed]
59. Soorni, A.; Haak, D.; Zaitlin, D.; Bombarely, A. Organelle_PBA, a Pipeline for Assembling Chloroplast and Mitochondrial Genomes from PacBio DNA Sequencing Data. *BMC Genom.* **2017**, *18*, 49. [CrossRef] [PubMed]
60. Koren, S.; Walenz, B.P.; Berlin, K.; Miller, J.R.; Bergman, N.H.; Phillippy, A.M. Canu: Scalable and Accurate Long-Read Assembly via Adaptive k-Mer Weighting and Repeat Separation. *Genome Res.* **2017**, *27*, 722–736. [CrossRef] [PubMed]
61. Wick, R.R.; Judd, L.M.; Gorrie, C.L.; Holt, K.E. Unicycler: Resolving Bacterial Genome Assemblies from Short and Long Sequencing Reads. *PLoS Comput. Biol.* **2017**, *13*, e1005595. [CrossRef]
62. Kolmogorov, M.; Yuan, J.; Lin, Y.; Pevzner, P.A. Assembly of Long, Error-Prone Reads Using Repeat Graphs. *Nat. Biotechnol.* **2019**, *37*, 540–546. [CrossRef]
63. Miyamoto, M.; Motooka, D.; Gotoh, K.; Imai, T.; Yoshitake, K.; Goto, N.; Iida, T.; Yasunaga, T.; Horii, T.; Arakawa, K.; et al. Performance Comparison of Second- and Third-Generation Sequencers Using a Bacterial Genome with Two Chromosomes. *BMC Genom.* **2014**, *15*, 699. [CrossRef]
64. Denisov, G.; Walenz, B.; Halpern, A.L.; Miller, J.; Axelrod, N.; Levy, S.; Sutton, G. Consensus Generation and Variant Detection by Celera Assembler. *Bioinformatics* **2008**, *24*, 1035–1040. [CrossRef]
65. Syme, A.E.; McLay, T.G.B.; Udovicic, F.; Cantrill, D.J.; Murphy, D.J.; McLay, T.G.B.; Udovicic, F.; Cantrill, D.J.; Murphy, D.J. Long-Read Assemblies Reveal Structural Diversity in Genomes of Organelles—An Example with Acacia Pycnantha. *Gigabyte* **2021**, *2021*, 1–23. [CrossRef] [PubMed]
66. Shafin, K.; Pesout, T.; Lorig-Roach, R.; Haukness, M.; Olsen, H.E.; Bosworth, C.; Armstrong, J.; Tigyi, K.; Maurer, N.; Koren, S.; et al. Nanopore Sequencing and the Shasta Toolkit Enable Efficient de Novo Assembly of Eleven Human Genomes. *Nat. Biotechnol.* **2020**, *38*, 1044–1053. [CrossRef] [PubMed]
67. Cheng, H.; Concepcion, G.T.; Feng, X.; Zhang, H.; Li, H. Haplotype-Resolved de Novo Assembly Using Phased Assembly Graphs with Hifiasm. *Nat. Methods* **2021**, *18*, 170–175. [CrossRef] [PubMed]
68. Zhang, Z.; Xie, P.; Guo, Y.; Zhou, W.; Liu, E.; Yu, Y. Easy353: A Tool to Get Angiosperms353 Genes for Phylogenomic Research. *Mol. Biol. Evol.* **2022**, *39*, msac261. [CrossRef]
69. Zhou, W.; Armijos, C.E.; Lee, C.; Lu, R.; Wang, J.; Ruhlman, T.A.; Jansen, R.K.; Jones, A.M.; Jones, C.D. Plastid Genome Assembly Using Long-Read Data. *Mol. Ecol. Resour.* **2023**, *23*, 1442–1457. [CrossRef]
70. Lee, C.; Ruhlman, T.A.; Jansen, R.K. Unprecedented Intraindividual Structural Heteroplasmy in Eleocharis (Cyperaceae, Poales) Plastomes. *Genome Biol. Evol.* **2020**, *12*, 641–655. [CrossRef]
71. Ruhlman, T.A.; Zhang, J.; Blazier, J.C.; Sabir, J.S.M.; Jansen, R.K. Recombination-Dependent Replication and Gene Conversion Homogenize Repeat Sequences and Diversify Plastid Genome Structure. *Am. J. Bot.* **2017**, *104*, 559–572. [CrossRef]