Assembling high-quality plant genomes with Oxford Nanopore

High-quality plant genome assemblies are pivotal for advancing research and agricultural biotechnology, supporting the study of biodiversity to inform conservation strategies and enabling the detailed characterisation and improvement of crops for sustainable productivity. However, plant genomes can be large, complex, and highly repetitive, resulting in significant challenges when using legacy short-read sequencing technologies for their assembly. Furthermore, the use of PCR precludes the ability to capture sequences that are not amenable to amplification, while epigenetic modifications must be detected indirectly via chemical treatments such as bisulfite conversion.

Using Oxford Nanopore reads of unrestricted length — from short to ultra-long — it is now possible to generate high-quality plant genome assemblies with a simple, streamlined workflow. Long nanopore reads can span large repetitive or highly uniform sequences and structural variants, whilst sequencing of native DNA captures sequences that are inaccessible to PCR. Epigenetic modifications can also be detected alongside canonical base sequences within the same sequencing run, providing multiomic insights from a single dataset. The versatile range of high-output PromethION™ devices allows labs to scale sequencing capacity to fit projects of varying sizes, sample volumes, and budgets, offering tailored solutions for diverse sequencing needs.

This workflow introduces how to generate a high-quality genome assembly with nanopore sequencing on a PromethION, starting from a plant leaf sample.

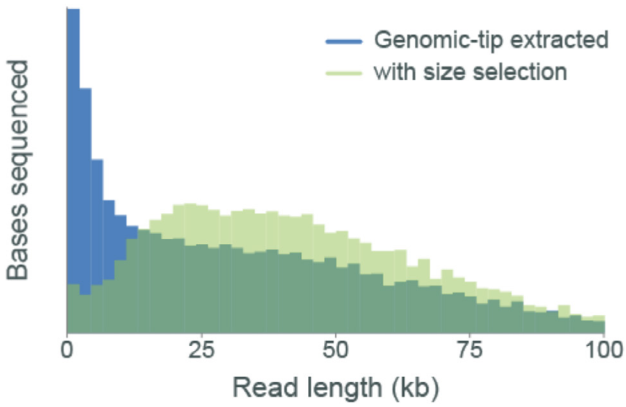
Extraction:
obtaining high molecular-weight DNA

View the extraction protocol:
nanoporetech.com/fever-tree-extraction

To maximise outputs of long reads in nanopore sequencing, it is important to extract high-quality, high molecular-weight DNA samples. We recommend extracting DNA from plant leaf samples using our protocol based on **Carlson lysis buffer**, followed by purification using the **QIAGEN Genomic-tip 500/G**.

Size selection is not a required step in sample preparation, but can be utilised to increase read length N50 in sequencing. If sufficient starting sample is available, we recommend using the optional size selection step featured in the extraction protocol, which will select for fragments over 10 kb in length.

When working with degraded samples composed of shorter fragment lengths, the size selection step can be left out, ensuring sufficient input material is available to proceed with library preparation.



Library preparation:
generating sequencing libraries without PCR

Find out more about preparing DNA libraries:
nanoporetech.com/prepare-dna

We recommend preparing libraries using the **Ligation Sequencing Kit**. This PCR-free approach retains long fragments of native DNA, ensuring regions that would be challenging to amplify are captured in sequencing, and preserves epigenetic modifications.

The Ligation Sequencing Kit is optimised for high accuracy and output, for good depth of coverage across the plant genome, and takes approximately one hour from extracted DNA to sequencing-ready library.

Sequencing:

scaling output to your needs with PromethION

Learn more about the PromethION range:
nanoporetech.com/promethion

The **PromethION** range of nanopore sequencing devices offers ideal solutions for on-demand, high-output whole-genome sequencing. The PromethION 24 provides sequencing on up to 24 individually addressable flow cells and features powerful onboard compute. For smaller projects, the PromethION 2 devices deliver PromethION-scale sequencing for any lab, on up to two flow cells run concurrently or independently.

For optimal assembly metrics, we recommend generating sufficient sequencing data for a minimum of 30x depth of coverage for haploid or diploid plant genomes. For genomes with higher ploidy, this should be increased by an additional 15x per ploidy. We recommend using the super accuracy (SUP) basecalling mode for the highest quality reads. Additionally, 5mC and 6mA modified base detection can be enabled during the basecalling stage to capture methylation information for downstream analyses. This will not impact the assembly workflow but will provide additional epigenetic insights.

Analysis:

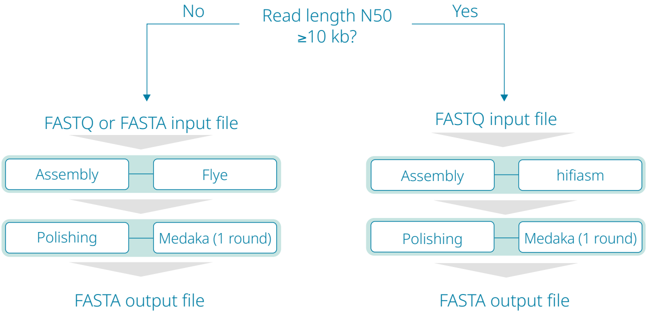
constructing a high-quality plant genome assembly

Find out more about analysing nanopore sequencing data:
nanoporetech.com/analyse

Prior to genome assembly, we recommend using **SeqKit**[1,2] to filter for reads with quality scores of over 10. For assembly, we recommend using the assembly tool **hifiasm**[3–5] (v0.24.0 or later) with settings optimised for simplex assembly using Oxford Nanopore reads (enabled via the '`--ont`' flag). From an input of FASTQ files, this assembler generates an n-ploid assembly where all alleles are represented, though individual contigs may include alleles from multiple haplotypes.

If working with degraded plant samples yielding read length N50s below 10 kb, the assembly tool **Flye**[6] is recommended due to its robustness across various species and datasets. Flye accepts FASTQ or FASTA files as input and produces haplotype-collapsed assemblies, though handling of ploidy can be limited in highly heterozygous genomes.

For polishing, we advise running **Medaka** as an optional step. A single round of polishing with Medaka has been shown to improve consensus accuracy in many cases, enhancing the overall assembly quality.



Read length N50 ≥10 kb?

No

FASTQ or FASTA input file

Assembly – Flye

Polishing – Medaka (1 round)

FASTA output file

Yes

FASTQ input file

Assembly – hifiasm

Polishing – Medaka (1 round)

FASTA output file

Find out more at: nanoporetech.com/plant-research

Oxford Nanopore Technologies

oxford nanopore technologies
@nanopore
@nanoporetech.com
www.nanoporetech.com

References:

1. Shen, W., Sipos, B., and Zhao, L. *iMeta* 3:e191 (2024). DOI: https://doi.org/10.1002/imt2.191
2. Shen, W. et al. *PLOS ONE* 11(10):e0163962 (2016). DOI: https://doi.org/10.1371/journal.pone.0163962
3. Cheng, H. et al. *Nat. Methods* 18:170–175 (2021). DOI: https://doi.org/10.1038/s41592-020-01056-5
4. Cheng, H. et al. *Nat. Biotechnol*. 40:1332–1335 (2022). DOI: https://doi.org/10.1038/s41587-022-01261-x
5. Cheng, H. et al. *Nat. Methods* 21:967–970 (2024). DOI: https://doi.org/10.1038/s41592-024-02269-8
6. Kolmogorov, M. et al. *Nat. Biotechnol*. 37:540–546 (2019). DOI: https://doi.org/10.1038/s41587-019-0072-8

WF_1281(EN)_V2_17Feb2025