

# Plant Genome Sequencing Technology

Subjects: **Biochemistry & Molecular Biology**

Contributor: Robert Henry

The development of technology that can capture large volumes of sequence data at low costs and with high accuracy has driven the acceleration of plant genome sequencing advancements. The sequencing of all plant species is a long-term goal that may become key to effectively supporting life on Earth through the improved management of plants in wild populations and their selection and genetic enhancement for use in agriculture and food production.

DNA sequencing

plant genome

long read sequencing

chromosome assembly

## 1. DNA Isolation

The starting point for the sequencing of plant genomes is obtaining a sample of DNA to sequence. The ease of obtaining a DNA sample of suitable quality for sequencing varies greatly between species. Plants contain many secondary metabolites, proteins and polysaccharides that may interfere with DNA extraction and become a source of contaminants that often reduce the efficiency of DNA extraction. Current technologies require a minimum amount of DNA and the DNA must be pure (free from contaminants that may inhibit sequencing) to facilitate efficient sequencing and the generation of large volumes of data. The amount of DNA required for long read sequencing has been greater (usually  $\mu\text{g}$  quantities) than that needed for short read sequencing (<https://dnatech.genomecenter.ucdavis.edu/pacbio-library-prep-sequencing>; accessed on 6 June 2022). For long read sequencing, the DNA must be intact (not degraded) so that long sequences are present in the sample and can be sequenced to extract long reads. Simple methods that were suitable for DNA extraction in the past for purposes such as PCR analysis <sup>[1]</sup> are no longer adequate, so the development of species- <sup>[2]</sup> or tissue-specific methods that can support next generation sequencing are often required <sup>[3]</sup>. Some species are especially difficult and require the isolation of the nucleus <sup>[4]</sup> first as a source of DNA that may be free from contaminants from other parts of the plant cell. The isolation of nuclei in plants is challenging for some of the same reasons that DNA isolation is difficult. The disruption of the plant cell wall requires forces that may damage organelles, such as nuclei and shear DNA.

## 2. Sort Read Sequences

The first set of next generation sequencing technologies provided larger volumes of short DNA sequences. The accuracy of these short sequences and the volume of data have since increased dramatically. The length of these sequences started at around 30 bp and has rapidly advanced to 100–150 bp. Paired-end sequencing has extended this technology to allow for the production of sequences of around 400 bp, but most applications currently deliver

sequences of around 150 bp. Illumina sequencing platforms are the dominant technology used for short read sequencing. This technology conducts sequencing by synthesis in a very large number of parallel reactions. The incorporation of nucleotides is monitored as the DNA is copied. Other techniques (e.g., solid [\[5\]](#) and 454 sequencing [\[6\]](#)) have been replaced by new technologies because they generally offered lower accuracy or data volumes, resulting in relatively higher costs. Ion Torrent sequencing is used for the rapid determination of a sequence of large numbers of small sequences, such as amplicon sequencing and 16 S metagenomic sequencing. In plants, this has been used for chloroplast sequencing [\[7\]](#).

## **3. Long Read Sequences**

The assembly of plant genomes with large numbers of repetitive sequences is not possible with only short read sequences. Therefore, technology that allows much longer sequences to be generated has been key to simplifying genome assembly. The length of these sequences and their accuracy have improved greatly since the technology was first introduced.

### **3.1. PacBio**

Pacific Biosciences (PacBio) has developed a long-read sequencing platform that provides accurate long read sequencing. The single-molecule real-time (SMRT) sequencing involves monitoring the incorporation of fluorescent-labeled nucleotides [\[8\]](#). Recently, single long reads (also known as continuous long reads (CLR)) have largely been replaced by HiFi reads, which provide a consensus sequence based on sequencing a long fragment of DNA (approximately 15,000 bp) multiple times by first circularizing the DNA and reading around the circle many times [\[9\]](#). The repeated sequencing of the same molecule allows a highly accurate sequence to be generated as the circular consensus sequence (CCS) is read. The quality of the genomes that are generated by the assembly of these reads into contigs has been improved by the application of optimized assembly tools, such as those provided by hifiasm.

### **3.2. ONT**

Oxford Nanopore Technologies (ONT) provides a long read sequencing technology that delivers accurate sequence data quickly. The sequence is determined by measuring the changes in electrical currents as the DNA is passed through a pore. The ONT platform generates very long reads and has the advantage of very low instrumentation costs. This platform has continued to improve and deliver very long read sequences with increasing accuracy [\[10\]](#). ONT sequencing has been widely applied to very rapid sequencing, such as that required for diagnostics [\[11\]](#), due to the advantage of having portable instruments. The chromosome-level assemblies of plant genomes can be achieved in combination with methods, such as optical mapping [\[12\]](#).

### **3.3. Other**

Several technology providers have developed pseudo-long reads that are created by linking short reads. These techniques may produce long reads at lower costs, but the long reads that are generated often do not match the accuracy of the current long read methods [13]. These technologies have been developed by Universal Sequencing Technology [14], MGI [15] and 10× genomics [16]. Despite the great contribution that long read sequencing technology has made to the efficient production of high-quality plant genomes, the emergence of further advances in long read sequencing technologies remains one of the key areas that may contribute to future advances.

### 3.4. Advances

Genome sequencing and assembly requires an adequate depth of sequencing. The size of contigs that can be assembled as long read sequence data has been shown to increase in an almost linear way [17]. The size of the assembled genomes reduces slightly with more contiguous assemblies, probably due to the joining of homologous contig ends, as does the completeness [17]. Improved software has also enabled improvements in the assembly of long read sequences [18]. The use of hifiasm has been shown to allow the haplotype-resolved assembly of the large (30 Gb) genome of the Californian redwood (*Sequoia sempervirens*) [19]. These advances are illustrated by the quality of the early plant genomes relative to those that are being generated by the latest technology. The first rice genome, which was reported in 2002, was highly fragmented while current technology delivers sequence contigs that are often full-length chromosomes [18].

## 4. Chromosome-Level Assembly

The ultimate aim of genome sequencing is to obtain a complete genome sequence of each chromosome, from telomere to telomere. This relies on evidence from beyond the DNA sequence data. Physical and genetic mapping methods have been used to achieve the chromosome-level assembly of contigs that were generated from sequencing data [20]. Recently, the advances in sequencing technology have made it possible to generate many full-length chromosomes from the sequence data alone [18]. The complete assembly of sequence contigs into whole chromosomes has been widely achieved using genetic mapping data, chromatin mapping (Hi-C) or optical mapping. Hi-C [21] involves the mapping of chromatin by crosslinking the DNA in the intact chromatin, digesting the DNA and then sequencing (short reads) the DNA fragments at the ends of the crosslinks. These are used to position the sequence contigs along the chromosome. Optical mapping (Bionano) can also be used to locate sequences along the DNA sequence and to scaffold the sequence contigs [22]. Many projects have combined these technologies to support the generation of high-quality genomes. Recent advances in long read sequencing have enabled the generation of long contigs of highly accurate sequences, reducing reliance on these techniques for high-level assembly. They remain essential for the de novo assembly of most chromosome-level genomes. High-quality sequence contigs in combination with genetic mapping data, Bionano optical data or Hi-C chromatin mapping have generally succeeded in achieving chromosome-level assemblies of plant genomes. A report on more than 100 chromosome-level assemblies in 2021 [23] found only a 73% coverage of the pseudomolecules that represent the chromosomes. The combination of long read sequencing and the use of these tools has resulted in the recent reporting of many high-quality chromosome-level genome sequences (**Table 1**), with the quality improving greatly along with the most recent technology.

**Table 1.** Some recent chromosome-level assemblies of plant genomes.

Species		Technique *	Reference
Amorphophallus	<i>Amorphophallas konjac</i>	Hi-C	[24]
Apple	<i>Malus domestica</i>	Genetic Map	[25]
Avocado	<i>Persea americana</i>	Genetic Map	[26]
Banana	<i>Musa balbisiana</i>	Hi-C	[27]
Camphor	<i>Cinnamomum camphora</i>	Hi-C	[28]
Carrot	<i>Daucus carota</i>	Genetic Map	[29]
Chinese Skullcap	<i>Scutellaria baicalensis</i>	Hi-C	[30]
Crucihimalaya	<i>Crucihimalaya lasiocarpa</i>	Hi-C	[31]
Cucumber	<i>Cucumis metuliferus</i>	Hi-C	[32]
Eucalypt	<i>Corymbia citriodora</i>	Genetic Map	[20]
Field Pennycresss	<i>Thlaspi arvense</i>	Genetic Map/Hi-C/Bionano	[33]
Ginger	<i>Zingiber officinale</i>	Hi-C	[34]
Ginkgo	<i>Ginkgo biloba</i>	Hi-C	[35]
Joboba	<i>Simmondsia chinensis</i>	Hi-C	[36]
Macadamia	<i>Macadamia janseni</i>	Hi-C	[37]
	<i>Macadamia integrifolia</i>	Genetic Map	[38]
Paper Mulberry	<i>Broussonetis papyrifera</i>	Hi-C	[39]
Peach	<i>Prunus persica</i>	Hi-C	[40]
Peanut	<i>Arachis hypogaea</i>	Hi-C	[41]
Speranskia	<i>Speranskia yunnanensis</i>	Hi-C	[42]
Taxus	<i>Taxus chinensis</i>	Hi-C	[43]
Tea	<i>Camellia sinensis</i>	Hi-C	[44]
Water Caltrop	<i>Trapa spp.</i>	Hi-C	[45]

## 5. Haplotype-Resolved Genomes

\* The technique used to achieve the chromosome-level assembly of the sequence contigs.

Most published plant genomes are collapsed representations of the diploid genome as a single sequence, with a random inclusion of one of the two alleles at each heterozygous position. Only recently has it become possible to assemble each haplotype separately [46]. This has been the result of advances in both sequencing technology and sequence assembly tools. Current technology suggests that most genomes can now be sequenced at the haplotype level, thereby replacing the reporting of collapsed genomes with the sequences of the two haplotypes.

## 6. Pan-Genomes

The sequencing of plant genomes has shown that significant differences may be found within a plant species, which means that more than one reference genome is required to represent the species. The sequencing of plant genomes has also demonstrated that many genes are variably present in different individuals within a species. These presence/absence differences have led to the construction of pan-genomes, which represent the complete set of genes found within a population. A genome that includes all of the variations within a group of plants is known as a pan-genome. The pan-genome concept is a powerful tool for plant breeders for the analysis of gene pools [47]. Pan-genomes can be generated at different levels to represent the diversity that is found within, for example, domesticated gene pools, species or genera.

## 7. Transcriptomes

Transcriptome sequencing is an important tool for the analysis of the expressed regions of a genome. This is key to understanding gene functions and the determination of the genetic basis of important plant traits [48][49][50][51][52][53]. Transcriptome sequencing complements genome sequencing in genome characterization. Transcripts provide physical evidence that the sequence is formed of the expressed and complementary predictions of genes, based on the sequence alone. The comparison of the transcriptomes of different genotypes from different tissues or cell types at different stages of development and under different environmental conditions allows for the discovery of the genes that control plant traits and has become a key approach in plant biology and the discovery of genes for selection in plant breeding. Single-cell transcriptomics has become a powerful tool for understanding gene expression at the cell and tissue level but has had limited application in plants [54], partly due to the difficulty in isolating specific plant cells without disrupting expression.

### 7.1. RNAseq

The quantitative analysis of the levels of expression of genes in any specific cell, tissue, organ, genotype or development stage is widely determined by RNA sequencing (RNAseq) [55][56]. RNAseq has largely replaced earlier array-based or gene by gene analysis tools as it provided a more unbiased analysis of the whole transcriptome.

An analysis of the gene expression in the highly polyploid sugarcane genome revealed that while the different alleles of most genes are expressed in direct proportion to their abundance in the genome, some genes show highly biased patterns of expression [57]. In hexaploid wheat, subgenome-specific responses to diseases have also been reported [58].

## 7.2. Long Read Transcriptomes

Long read sequencing is a method that has been applied to the analysis of plant transcriptomes, which reveals the diversity of full-length transcripts and defines the variations in splicing and intron retention in gene expression [59]. The long read sequencing of transcriptomes avoids the challenge of the assembly of many closely related transcripts from short reads. Unique 3' and 5' sequences may be separated by common intervening sequences, which creates the risk of incorrectly combining the ends of the transcripts when using short reads.

Some examples of the application of long read sequencing to the analysis of plant transcriptomes of increasing complexity can be found for polyploid species in **Table 2**.

**Table 2.** The long read sequencing of polyploid transcriptomes.

Species	Ploidy	Reference
Coffee	4X	[59]
Wheat	6X	[60]
Strawberry	8X	[61]
Sugarcane	12X	[62]

# 8. Organelle Genome Sequencing

Plant cells usually contain a single nucleus and many organelles, probably hundreds of mitochondria and thousands of chloroplasts. Sequencing the organelle genomes is complicated by the transfer of genes between these genomes. The nuclear genome often contains many insertions of large and small sequences of organellar genomes. Many early methods struggled to distinguish organellar gene sequences from those of copies that were inserted into the nuclear genome because they relied on PCR amplification or organelle separation [63]. Nuclear inserts may represent versions of organellar genomes that were transferred in the past and that have diverged since insertion.

## 8.1. Chloroplast Genomes

The chloroplast genomes of plants are highly conserved sequences of 100–150 Kb, containing around 100 genes [64]. The structure of most chloroplasts is similar, with four components including inverted repeats that separate large and small single-copy regions. Chloroplasts have been widely used in plant identification due to their presence in all green plants and the high copy numbers in the cell that simplify the detection of chloroplast sequences. Early approaches that relied on chloroplast isolation or PCR amplification were plagued by confusion due to the copies of chloroplast sequences in the nuclear and mitochondrial genomes. Recent approaches [65] rely on the higher abundance of chloroplast genome sequence reads in short read sequence data to clearly distinguish

the correct sequence of the relevant chloroplast [\[66\]](#). The development of software tools now allows for the efficient extraction of accurate whole chloroplast genome sequences from even low (nuclear) coverage sequencing datasets. The annotation of chloroplast genomes that were generated in this way has resulted in the identification of around 100 genes with increasingly well-defined functions [\[67\]](#).

The sequencing of the maternal (e.g., chloroplast) and nuclear genomes of plants has frequently revealed discordant phylogenies [\[68\]\[69\]\[70\]](#), suggesting widespread reticulate evolution in plant populations (**Table 3**). Chloroplast transfers between species during rare events results in “chloroplast capture” by closely related species.

**Table 3.** Discordant phylogenies for chloroplast and nuclear genome sequences.

Species	Reference
Apple Tribe	<a href="#">[71]</a>
Eucalypts	<a href="#">[70]</a>
Osmorhiza spp.	<a href="#">[72]</a>
<i>Pedicularis</i> spp.	<a href="#">[73]</a>
Sorghum	<a href="#">[68]</a>
Rice ( <i>Oryza</i> )	<a href="#">[69]</a>

## 8.2. Plant Mitochondrial Genomes

The mitochondrial genomes of plants [\[74\]](#) are much larger and less conserved than the chloroplast genomes and as a result, they have been much less studied than chloroplasts. The mitochondrial genome, as with the nuclear genome, may include sequences that were derived from the chloroplast that have been inserted into the genome at various times throughout its evolutionary history. Due to the relatively higher number of mitochondrial genomes in cells, these sequences are even more likely to be confused with chloroplast genome sequences than chloroplast sequences that were inserted into nuclear genomes.

## References

1. Graham, G.C.; Mayers, P.; Henry, R.J. A simplified method for the preparation of fungal genomic DNA for PCR and RAPD analysis. *BioTechniques* 1994, 16, 48–50.
2. Nath, O.; Fletcher, S.J.; Hayward, A.; Shaw, L.M.; Agarwal, R.; Furtado, A.; Henry, R.J.; Mitter, N. A Comprehensive High-Quality DNA and RNA Extraction Protocol for a Range of Cultivars and Tissue Types of the Woody Crop Avocado. *Plants* 2022, 11, 242.

3. Healey, A.; Furtado, A.; Cooper, T.; Henry, R.J. Protocol: A simple method for extracting next-generation sequencing quality genomic DNA from recalcitrant plant species. *Plant Methods* 2014, 10, 21.
4. Guilfoyle, T.J. Chapter 8 Isolation and Characterization of Plant Nuclei. *Method Cell Biol.* 1995, 50, 101–112.
5. Garrido-Cardenas, J.A.; Garcia-Maroto, F.; Alvarez-Bermejo, J.A.; Manzano-Agugliaro, F. DNA Sequencing Sensors: An Overview. *Sensors* 2017, 17, 588.
6. Bundock, P.C.; Elliott, F.G.; Ablett, G.; Benson, A.D.; Casu, R.E.; Aitken, K.S.; Henry, R.J. Targeted single nucleotide polymorphism (SNP) discovery in a highly polyploid plant species using 454 sequencing. *Plant Biotechnol. J.* 2009, 7, 347–354.
7. Brozynska, M.; Furtado, A.; Henry, R.J. Direct Chloroplast Sequencing: Comparison of Sequencing Platforms and Analysis Tools for Whole Chloroplast Barcoding. *PLoS ONE* 2014, 9, e110387.
8. Rhoads, A.; Au, K.F. PacBio Sequencing and Its Applications. *Genom. Proteom. Bioinf.* 2015, 13, 278–289.
9. Hon, T.; Mars, K.; Young, G.; Tsai, Y.-C.; Karalius, J.W.; Landolin, J.M.; Maurer, N.; Kudrna, D.; Hardigan, M.A.; Steiner, C.C.; et al. Highly accurate long-read HiFi sequencing data for five complex genomes. *Sci. Data* 2020, 7, 399.
10. Wang, Y.H.; Zhao, Y.; Bollas, A.; Wang, Y.R.; Au, K.F. Nanopore Sequencing Technology, Bioinformatics and Applications. *Nat. Biotechnol.* 2021, 39, 1348–1365.
11. Wang, J.; Bhakta, N.; Miller, V.A.; Revsine, M.; Litzow, M.; Paietta, E.; Roberts, K.; Gu, Z.; Mullighan, C.; Jones, C.; et al. Nanopore RNA Sequencing as A Diagnostic Tool for Acute Leukemia in Low Resource Settings. *Pediatr. Blood Cancer* 2021, 68, S107.
12. Belser, C.; Istace, B.; Denis, E.; Dubarry, M.; Baurens, F.-C.; Falentin, C.; Genete, M.; Berrabah, W.; Chèvre, A.-M.; Delourme, R.; et al. Chromosome-scale assemblies of plant genomes using nanopore long reads and optical maps. *Nat. Plants* 2018, 4, 879–887.
13. Murigneux, V.; Rai, S.K.; Furtado, A.; Bruxner, T.J.C.; Tian, W.; Harliwong, I.; Wei, H.; Yang, B.; Ye, Q.; Anderson, E.; et al. Comparison of long-read methods for sequencing and assembly of a plant genome. *GigaScience* 2020, 9, giaa146.
14. Chen, Z.; Pham, L.; Wu, T.-C.; Mo, G.; Xia, Y.; Chang, P.L.; Porter, D.; Phan, T.; Che, H.; Tran, H.; et al. Ultralow-input single-tube linked-read library method enables short-read second-generation sequencing systems to routinely generate highly accurate and economical long-range sequencing information. *Genome Res.* 2020, 30, 898–909.



15. Wang, O.; Chin, R.; Cheng, X.; Wu, M.K.Y.; Mao, Q.; Tang, J.; Sun, Y.; Anderson, E.; Lam, H.K.; Chen, D.; et al. Efficient and unique cobarcoding of second-generation sequencing reads from long DNA molecules enabling cost-effective and accurate sequencing, haplotyping, and de novo assembly. *Genome Res.* 2019, 29, 798–808.
16. Singh, M.; Al-Eryani, G.; Carswell, S.; Ferguson, J.M.; Blackburn, J.; Barton, K.; Roden, D.; Luciani, F.; Phan, T.G.; Junankar, S.; et al. High-throughput targeted long-read single cell sequencing reveals the clonal and transcriptional landscape of lymphocytes. *Nat. Commun.* 2019, 10, 3120.
17. Sharma, P.A.O.; Alsubaie, B.; Al-Mssallem, I.; Nath, O.; Mitter, N.; Margarido, G.R.A.; Topp, B.; Murigneux, V.; Masouleh, A.K.; Furtado, A.; et al. Improvements in The Sequencing and Assembly of Plant Genomes. *Gigabyte* 2021.
18. Sharma, P.; Masouleh, A.K.; Topp, B.; Furtado, A.; Henry, R.J. De novo chromosome level assembly of a plant genome from long read sequence data. *Plant J.* 2021, 109, 727–736.
19. 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.
20. Healey, A.L.; Shepherd, M.; King, G.J.; Butler, J.B.; Freeman, J.S.; Lee, D.J.; Potts, B.M.; Silva-Junior, O.B.; Baten, A.; Jenkins, J.; et al. Pests, diseases, and aridity have shaped the genome of *Corymbia citriodora*. *Commun. Biol.* 2021, 4, 537.
21. Harewood, L.; Kishore, K.; Eldridge, M.; Wingett, S.; Pearson, D.; Schoenfelder, S.; Collins, V.P.; Fraser, P. Hi-C as a tool for precise detection and characterisation of chromosomal rearrangements and copy number variation in human tumours. *Genome Biol.* 2017, 18, 125.
22. Chen, P.; Jing, X.; Ren, J.; Cao, H.; Hao, P.; Li, X. Modelling BioNano optical data and simulation study of genome map assembly. *Bioinformatics* 2018, 34, 3966–3974.
23. Shirasawa, K.; Harada, D.; Hirakawa, H.; Isobe, S.; Kole, C. Chromosome-level de novo genome assemblies of over 100 plant species. *Breed. Sci.* 2021, 71, 117–124.
24. Gao, Y.; Zhang, Y.; Feng, C.; Chu, H.; Feng, C.; Wang, H.; Wu, L.; Yin, S.; Liu, C.; Chen, H.; et al. A chromosome-level genome assembly of *Amorphophallus konjac* provides insights into konjac glucomannan biosynthesis. *Comput. Struct. Biotechnol. J.* 2022, 20, 1002–1011.
25. Sun, X.; Jiao, C.; Schwaninger, H.; Chao, C.T.; Ma, Y.; Duan, N.; Khan, A.; Ban, S.; Xu, K.; Cheng, L.; et al. Phased diploid genome assemblies and pan-genomes provide insights into the genetic history of apple domestication. *Nat. Genet.* 2020, 52, 1423–1432.
26. Rendón-Anaya, M.; Ibarra-Laclette, E.; Méndez-Bravo, A.; Lan, T.; Zheng, C.; Carretero-Paulet, L.; Perez-Torres, C.A.; Chacón-López, A.; Hernandez-Guzmán, G.; Chang, T.-H.; et al. The avocado genome informs deep angiosperm phylogeny, highlights introgressive hybridization, and

- reveals pathogen-influenced gene space adaptation. *Proc. Natl. Acad. Sci. USA* 2019, 116, 17081–17089.
27. Wang, Z.; Miao, H.; Liu, J.; Xu, B.; Yao, X.; Xu, C.; Zhao, S.; Fang, X.; Jia, C.; Wang, J.; et al. *Musa balbisiana* genome reveals subgenome evolution and functional divergence. *Nat. Plants* 2019, 5, 810–821.
  28. Shen, T.; Qi, H.; Luan, X.; Xu, W.; Yu, F.; Zhong, Y.; Xu, M. The chromosome-level genome sequence of the camphor tree provides insights into Lauraceae evolution and terpene biosynthesis. *Plant Biotechnol. J.* 2021, 20, 244–246.
  29. Iorizzo, M.; Ellison, S.; Senalik, D.; Zeng, P.; Satapoomin, P.; Huang, J.; Bowman, M.; Iovene, M.; Sanseverino, W.; Cavagnaro, P.; et al. A high-quality carrot genome assembly provides new insights into carotenoid accumulation and asterid genome evolution. *Nat. Genet.* 2016, 48, 657–666.
  30. Zhao, Q.; Yang, J.; Cui, M.-Y.; Liu, J.; Fang, Y.; Yan, M.; Qiu, W.; Shang, H.; Xu, Z.; Yidiresi, R.; et al. The Reference Genome Sequence of *Scutellaria baicalensis* Provides Insights into the Evolution of Wogonin Biosynthesis. *Mol. Plant* 2019, 12, 935–950.
  31. Feng, L.; Lin, H.; Kang, M.; Ren, Y.; Yu, X.; Xu, Z.; Wang, S.; Li, T.; Yang, W.; Hu, Q. A chromosome-level genome assembly of an alpine plant *Crucihimalaya lasiocarpa* provides insights into high-altitude adaptation. *DNA Res.* 2022, 29, dsac004.
  32. Ling, J.; Xie, X.X.; Gu, X.F.; Zhao, J.L.; Ping, X.X.; Li, Y.; Yang, Y.H.; Mao, Z.C.; Xie, B.Y. High-quality chromosome-level genomes of *Cucumis metuliferus* and *Cucumis melo* provide insight into *Cucumis* genome evolution. *Plant J.* 2021, 107, 136–148.
  33. Nunn, A.; Rodríguez-Arévalo, I.; Tandukar, Z.; Frels, K.; Contreras-Garrido, A.; Carbonell-Bejerano, P.; Zhang, P.; Cruz, D.R.; Jandrasits, K.; Lanz, C.; et al. Chromosome-level *Thlaspi arvense* genome provides new tools for translational research and for a newly domesticated cash cover crop of the cooler climates. *Plant Biotechnol. J.* 2022, 20, 944–963.
  34. Li, H.L.; Wu, L.; Dong, Z.M.; Jiang, Y.S.; Jiang, S.J.; Xing, H.T.; Li, Q.; Liu, G.C.; Tian, S.M.; Wu, Z.Y.; et al. Haplotype-Resolved Genome of Diploid Ginger (*Zingiber Officinale*) and Its Unique Gingerol Biosynthetic Pathway. *Hortic. Res.* 2021, 8, 189.
  35. Shan, H.; Kong, H. The genome of *Ginkgo biloba* refined. *Nat. Plants* 2021, 7, 714–715.
  36. Al-Dossary, O.; Alsubaie, B.; Kharabian-Masouleh, A.; Al-Mssallem, I.; Furtado, A.; Henry, R.J. The jojoba genome reveals wide divergence of the sex chromosomes in a dioecious plant. *Plant J.* 2021, 108, 1283–1294.
  37. Sharma, P.; Murigneux, V.; Haimovitz, J.; Nock, C.J.; Tian, W.; Masouleh, A.K.; Topp, B.; Alam, M.; Furtado, A.; Henry, R.J. The genome of the endangered *Macadamia janseni* displays little

- diversity but represents an important genetic resource for plant breeding. *Plant Direct* 2021, 5, e364.
38. Nock, C.J.; Baten, A.; Mauleon, R.; Langdon, K.S.; Topp, B.; Hardner, C.; Furtado, A.; Henry, R.J.; King, G.J. Chromosome-Scale Assembly and Annotation of the Macadamia Genome (*Macadamia integrifolia* HAES 741). *G3-Genes Genomes Genet.* 2020, 10, 3497–3504.
  39. Peng, X.; Liu, H.; Chen, P.; Tang, F.; Hu, Y.; Wang, F.; Pi, Z.; Zhao, M.; Chen, N.; Chen, H.; et al. A Chromosome-Scale Genome Assembly of Paper Mulberry (*Broussonetia papyrifera*) Provides New Insights into Its Forage and Papermaking Usage. *Mol. Plant* 2019, 12, 661–677.
  40. Lian, X.D.; Zhang, H.P.; Jiang, C.; Gao, F.; Yan, L.; Zheng, X.B.; Cheng, J.; Wang, W.; Wang, X.B.; Ye, X.; et al. De Novo Chromosome-Level Genome of A Semi-Dwarf Cultivar of *Prunus Persica* Identifies The Aquaporin Ptip2 as Responsible for Temperature-Sensitive Semi-Dwarf Trait And Ppb3-1 for Flower Type And Size. *Plant Biotechnol. J.* 2022, 20, 886–902.
  41. Zhuang, W.; Chen, H.; Yang, M.; Wang, J.; Pandey, M.K.; Zhang, C.; Chang, W.-C.; Zhang, L.; Zhang, X.; Tang, R.; et al. The genome of cultivated peanut provides insight into legume karyotypes, polyploid evolution and crop domestication. *Nat. Genet.* 2019, 51, 865–876.
  42. Yuan, G.; Tan, S.; Wang, D.; Yang, Y.; Tian, B. Chromosome-Level Genome Assembly of the Rare and Endangered Tropical Plant *Speranskia yunnanensis* (Euphorbiaceae). *Front. Genet.* 2022, 12, 755564.
  43. Xiong, X.; Gou, J.; Liao, Q.; Li, Y.; Zhou, Q.; Bi, G.; Li, C.; Du, R.; Wang, X.; Sun, T.; et al. The *Taxus* genome provides insights into paclitaxel biosynthesis. *Nat. Plants* 2021, 7, 1026–1036.
  44. Zhang, X.; Chen, S.; Shi, L.; Gong, D.; Zhang, S.; Zhao, Q.; Zhan, D.; Vasseur, L.; Wang, Y.; Yu, J.; et al. Haplotype-resolved genome assembly provides insights into evolutionary history of the tea plant *Camellia sinensis*. *Nat. Genet.* 2021, 53, 1250–1259.
  45. Lu, R.; Chen, Y.; Zhang, X.; Feng, Y.; Comes, H.P.; Li, Z.; Zheng, Z.; Yuan, Y.; Wang, L.; Huang, Z.; et al. Genome sequencing and transcriptome analyses provide insights into the origin and domestication of water caltrop (*Trapa* spp., Lythraceae). *Plant Biotechnol. J.* 2022, 20, 761–776.
  46. Low, W.Y.; Tearle, R.; Liu, R.; Koren, S.; Rhie, A.; Bickhart, D.; Rosen, B.D.; Kronenberg, Z.N.; Kingan, S.B.; Tseng, E.; et al. Haplotype-resolved genomes provide insights into structural variation and gene content in Angus and Brahman cattle. *Nat. Commun.* 2020, 11, 2071.
  47. Tao, Y.; Zhao, X.; Mace, E.; Henry, R.; Jordan, D. Exploring and Exploiting Pan-genomics for Crop Improvement. *Mol. Plant* 2018, 12, 156–169.
  48. Cheng, B.; Furtado, A.; Henry, R.J. The coffee bean transcriptome explains the accumulation of the major bean components through ripening. *Sci. Rep.* 2018, 8, 11414.

49. Gillies, S.A.; Futardo, A.; Henry, R.J. Gene expression in the developing aleurone and starchy endosperm of wheat. *Plant Biotechnol. J.* 2012, 10, 668–679.
50. Kasirajan, L.; Hoang, N.V.; Furtado, A.; Botha, F.C.; Henry, R.J. Transcriptome analysis highlights key differentially expressed genes involved in cellulose and lignin biosynthesis of sugarcane genotypes varying in fiber content. *Sci. Rep.* 2018, 8, 11612.
51. Nirmal, R.C.; Furtado, A.; Wrigley, C.; Henry, R.J. Influence of Gene Expression on Hardness in Wheat. *PLoS ONE* 2016, 11, e0164746.
52. Rangan, P.; Furtado, A.; Henry, R. Differential response of wheat genotypes to heat stress during grain filling. *Exp. Agric.* 2019, 55, 818–827.
53. Rangan, P.; Furtado, A.; Henry, R.J. The transcriptome of the developing grain: A resource for understanding seed development and the molecular control of the functional and nutritional properties of wheat. *BMC Genom.* 2017, 18, 766.
54. Cuperus, J.T. Single-cell genomics in plants: Current state, future directions, and hurdles to overcome. *Plant Physiol.* 2021, 188, 749–755.
55. Hoang, N.V.; Furtado, A.; Thirugnanasambandam, P.P.; Botha, F.C.; Henry, R.J. De novo assembly and characterizing of the culm-derived meta-transcriptome from the polyploid sugarcane genome based on coding transcripts. *Heliyon* 2018, 4, e00583.
56. Tirumala, R.K.; Subbaiyan, G.K.; Singh, A.K.; Furtado, A.; Henry, R.J. RNA-Seq to Understand Transcriptomes and Application in Improving Crop Quality. In *Comprehensive Foodomics*, 1st ed.; Cifuentes, A., Ed.; Elsevier: Amsterdam, The Netherlands, 2021; Volume 1, pp. 472–485.
57. Margarido, G.R.A.; Correr, F.H.; Furtado, A.; Botha, F.C.; Henry, R.J. Limited allele-specific gene expression in highly polyploid sugarcane. *Genome Res.* 2022, 32, 297–308.
58. Powell, J.J.; Fitzgerald, T.L.; Stiller, J.; Berkman, P.J.; Gardiner, D.M.; Manners, J.M.; Henry, R.J.; Kazan, K. The defence-associated transcriptome of hexaploid wheat displays homoeolog expression and induction bias. *Plant Biotechnol. J.* 2016, 15, 533–543.
59. Cheng, B.; Furtado, A.; Henry, R.J. Long-read sequencing of the coffee bean transcriptome reveals the diversity of full-length transcripts. *GigaScience* 2017, 6, gix086.
60. Zhou, S.; Zhang, J.; Han, H.; Zhang, J.; Ma, H.; Zhang, Z.; Lu, Y.; Liu, W.; Yang, X.; Li, X.; et al. Full-length transcriptome sequences of *Agropyron cristatum* facilitate the prediction of putative genes for thousand-grain weight in a wheat-*A. cristatum* translocation line. *BMC Genom.* 2019, 20, 1025.
61. Yuan, H.Z.; Yu, H.M.; Huang, T.; Shen, X.J.; Xia, J.; Pang, F.H.; Wang, J.; Zhao, M.Z. The Complexity of The *Fragaria X Ananassa* (Octoploid) Transcriptome by Single-Molecule Long-Read Sequencing. *Hortic. Res.* 2019, 6, 46.

62. Hoang, N.V.; Furtado, A.; Mason, P.J.; Marquardt, A.; Kasirajan, L.; Thirugnanasambandam, P.P.; Botha, F.C.; Henry, R.J. A survey of the complex transcriptome from the highly polyploid sugarcane genome using full-length isoform sequencing and de novo assembly from short read sequencing. *BMC Genom.* 2017, 18, 395.
63. Hoang, N.V.; Furtado, A.; McQualter, R.B.; Henry, R.J. Next generation sequencing of total DNA from sugarcane provides no evidence for chloroplast heteroplasmy. *New Negat. Plant Sci.* 2015, 1–2, 33–45.
64. Dobrogojski, J.; Adamiec, M.; Luciński, R. The chloroplast genome: A review. *Acta Physiol. Plant.* 2020, 42, 98.
65. Henry, R.; Rice, N.; Edwards, M.; Nock, C. Next-Generation Technologies to Determine Plastid Genome Sequences. In *Chloroplast Biotechnology*; Maliga, P., Ed.; Humana Press: Totova, NJ, USA, 2014; Volume 1132, pp. 39–46.
66. Nock, C.J.; Waters, D.L.; Edwards, M.A.; Bowen, S.G.; Rice, N.; Cordeiro, G.M.; Henry, R.J. Chloroplast genome sequences from total DNA for plant identification. *Plant Biotechnol. J.* 2010, 9, 328–333.
67. Guyeux, C.; Charr, J.-C.; Tran, H.T.M.; Furtado, A.; Henry, R.J.; Crouzillat, D.; Guyot, R.; Hamon, P. Evaluation of chloroplast genome annotation tools and application to analysis of the evolution of coffee species. *PLoS ONE* 2019, 14, e0216347.
68. Ananda, G.; Norton, S.; Blomstedt, C.; Furtado, A.; Møller, B.; Gleadow, R.; Henry, R. Phylogenetic relationships in the *Sorghum* genus based on sequencing of the chloroplast and nuclear genes. *Plant Genome* 2021, 14, e20123.
69. Brozynska, M.; Copetti, D.; Furtado, A.; Wing, R.A.; Crayn, D.; Fox, G.; Ishikawa, R.; Henry, R.J. Sequencing of Australian wild rice genomes reveals ancestral relationships with domesticated rice. *Plant Biotechnol. J.* 2016, 15, 765–774.
70. Healey, A.; Lee, D.J.; Furtado, A.; Henry, R.J. Evidence of inter-sectional chloroplast capture in *Corymbia* among sections *Torellianae* and *Maculatae*. *Aust. J. Bot.* 2018, 66, 369.
71. Hodel, R.G.J.; Zimmer, E.A.; Liu, B.-B.; Wen, J. Synthesis of Nuclear and Chloroplast Data Combined With Network Analyses Supports the Polyploid Origin of the Apple Tribe and the Hybrid Origin of the Maleae—Gillenieae Clade. *Front. Plant Sci.* 2022, 12, 820997.
72. Yoo, K.; Lowry, P.P.; Wen, J. Discordance of chloroplast and nuclear ribosomal DNA data in *Osmorhiza* (Apiaceae). *Am. J. Bot.* 2002, 89, 966–971.
73. Yu, W.-B.; Huang, P.-H.; Li, D.-Z.; Wang, H. Incongruence between Nuclear and Chloroplast DNA Phylogenies in *Pedicularis* Section *Cyathophora* (Orobanchaceae). *PLoS ONE* 2013, 8, e74828.

74. Gualberto, J.M.; Milesina, D.; Wallet, C.; Niazi, A.K.; Weber-Lotfi, F.; Dietrich, A. The plant mitochondrial genome: Dynamics and maintenance. *Biochimie* 2014, 100, 107–120.
- 

Retrieved from <https://encyclopedia.pub/entry/history/show/60981>