PROGRESS IN SEQUENCING OF THE COMPLETE HAPLOTYPE-RESOLVED DIPLOID GENOMES OF PLANTS

[1]Chemeris D.A., [2]Kuluev B.R., [3]Patrushev M.V., [2]Garafutdinov R.R., [2]Chemeris A.V.

[1]GENVED LLC, Moscow, Russia, E-mail: dch@dch.ru.net
[2]Institute of Biochemistry and Genetics, Ufa Federal Research Centre,
Russian Academy of Sciences, Ufa, Russia
[3]The National Research Center «Kurchatov Institute», Moscow, Russia

Resume

Now most of the genomes of higher organisms are assembled for haploid set of chromosomes, in which DNA fragments from paired chromosomes from different parents alternate in the mosaic order. In this regard, such assemblies are more logical to consider as quasi-genomes, since they do not provide accurate information about the relationship between genotype and phenotype. A new level of genome assembly is haplotype-resolved diploid genomes of higher organisms, including plants and their polyploid forms that have undergone cytological and functional diploidization. Given the importance of such genome assemblies, a large number of corresponding computer programs have been developed that allow phased genome assembly and analysis. To date, haplotype-resolved diploid genomes have been assembled for more than 60 species of plants. Among them are various agricultural plants, ornamental plants, fruit trees and shrubs, forest trees, medicinal plants and others. For the most part, they are represented by diploid plants, but there are also triploids, auto- and allo-tetraploids, hexaploids and even octaploids. The sizes of the assembled genomes also vary greatly - from 135 million bp to 21.6 billion bp. In many articles it is note that phased genome assembly helped to more accurately determine individual important characteristics of the studied plants, including origin, evolution, domestication, yield, resistance to diseases and pests, as well as better understanding of metabolic pathways and mechanisms of heterosis.

Introduction

Exactly half a century ago, in 1973, two papers were published [Gilbert, Maxam, 1973; Sanger et al., 1973], in which the determination of DNA sequences of 24 nucleotides (the lactose-repressor binding site) and 50 nucleotides (a fragment of phage f1) was reported. Several years later two fast methods of DNA sequencing appeared [Maxam, Gilbert, 1977; Sanger et al., 1977]. But we will not dedicate in detail on the methods of DNA sequencing, as the history of their development in the last quarter of the 20th century is quite detailed by us earlier [Chemeris et al., 1999]. In another our work, we briefly reviewed the evolution of DNA sequencing methods of new generations [Zubov et al., 2021]. However, we can also recommend referring to another article, briefly describing the development of DNA sequencing methods of different generations over the past 40 years [Shendure et al., 2017].

Now the speed of sequencing in general compared to the same 1973 increased by trillions of times, and only in the last quarter of a century the cost of sequencing decreased by about a million times. This led to the possibility of sequencing complete genomes in a very short time and at an affordable price. To date, more than a thousand plant genomes belonging to more than 800 species and subspecies have been sequenced with varying degrees of completeness.

In the second half of the 1990s, even before the human genome was read, a catchy phrase appeared - "postgenome era" [Nowak, 1995]. For plants, this phrase appeared shortly after the completion of the sequencing Arabidopsis genome [Jander et al., 2002]. However, its use was premature, as the post-genomic era has not yet arrived, because modern genomes of higher organisms including plants are consist of consensus sequences from their parents in mosaic order and moreover from haploid set of chromosomes whereas phenotype depends on full (diploid) set of chromosomes. Thus, such genomes can be considered as quasihaploid genomes or quasi-genomes.

References genomes

If we do not take into account the sequencing of the yeast genome, then the first sequenced eukaryotic genomes in draft versions at the turn of the centuries were the genomes of model organisms - nematode *Caenorhabditis elegans* [The C. elegans Sequencing Consortium, 1998], fruit fly *Drosophila melanogaster* [Adams et al., 2000], *Arabidopsis thaliana* [The Arabidopsis Genome Initiative, 2000], as well as the human genome *Homo sapiens* [Lander et al., 2001; Venter et al., 2001]. Such genomes sequenced the first for any group of organisms have got designation as reference genomes for the species.

DNA polymorphism is much more than expected. In addition to the most massive Single-Nucleotide Polymorphism (SNP), much more contribution to the phenotypic manifestations is made by changes in the genome in the form of various Structural Variation (SV), including Insertion / Deletion (InDel), Copy Number Variations (CNV), Presence / Absence Variations (PAV), as well as translocations and inversions (TrIn) of genome regions. It was necessary to display these differences between the sequenced samples, for which the concept of "pangenome" was quite suitable, which should be considered as a kind of set of all genes, regulatory elements and non-coding regions present in different samples of the studied group.

Pangenomes

The term "pangenome" was proposed in 2005 to describe the genomes of several different isolates of the bacteria *Streptococcus agalactiae* [Tettelin et al., 2005]. Later, this term was extended to other organisms. The concept of plant pangenome was first proposed in 2007 when analyzing partially sequenced genomes of maize, revealing significant differences between the inbred lines Mo17 and B73, indicating that one reference genome of the species does not provide all the completeness of information [Morgante et al., 2007]. Subsequently, information on the diversity of genomes of one species continued to accumulate. Large-scale sequencing of plant pangenomes began in 2010, when the complete genomes of 6 inbred lines of maize [Lai et al., 2010], as well as 17 wild and 14 cultivated samples of soybean [Lam et al. 2010] were sequenced.

Only in the last couple of years, there have been many reviews on this topic, among which there are many that directly link pangenomics with the improvement of agricultural crops, as well as contain an analysis of relevant resources and specialized software tools [Hameed et al., 2022; Petereit et al., 2022; Tay Fernandez et al., 2022; Naithani et al., 2023; Shi et al., 2023; Wang et al., 2023a]. In them, one can find statements that agricultural science has entered the pangenomic era, and this is quite fair. In one of the articles, a timeline provides information on almost three dozen pangenomes of agricultural plants, with mention in some cases of haplotype assembly [Li et al., 2022]. At the same time, it was previously directly indicated that haplotype phasing (i.e. sequencing of diploid genomes) is a new frontier in the assembly of plant pangenomes [Michael, VanBuren, 2020], required for breeding work of a new quality.

Diploid genome sequencing

So, the top level of assembly of the nuclear genetic material of higher organisms can only be diploid, assembled on chromosome level. It must be read in T2T format, for which it is necessary to perform haplotype assembly of phased DNA regions exclusively *de novo*. This requires a significantly increased coverage of the genome during its sequencing and a combination of less accurate long/ultra-long reads with high-precision short reads. It also involves the use of other approaches: optical mapping, Hi-C sequencing, gamete binning, Strand-seq, linked-read sequencing and/or trio-sequencing, and the application of appropriate programs and algorithms for such phased assembly, of which more than fifty have already been created.

Despite the fact that phased assembly of diploid genomes of higher organisms remains rare, nevertheless, individual researchers have been conducting it for quite a long time. Many computer programs have been developed and now there is a whole arsenal of them, designed for using different source data. In this regard, one can pay attention to the review [Zhang et al., 2020], in which a stable trend in whole-genome sequencing for the restoration of phased diploid genomes is noted and confidence is expressed that in the coming years this will

become a routine procedure. Even more detailed, similar computer programs are considered in the review [Garg, 2021], where they were divided into groups by purpose; among others, it contains several programs that allow conducting such phased assembly *de novo* of polyploid genomes. In another recent review article [Guk et al., 2022], devoted to phased assembly of genomes, along with other programs, the assembler TrioCANU is mentioned, which implies operating with genomes of parental forms and their offspring. A lot of useful information is contained in a recent review, devoted to programs-assemblers of phased sequences [Duitama, 2023].

A significant part of the phased sequence assembly programs for the reconstruction of diploid genomes and their analysis appeared in the last couple of years – JTK [Masutani et al., 2023], Verkko [Rautiainen et al., 2023], GameteBinning_prac [Sun et al., 2023], mLinker [Sinha, Zhang, 2023], flopp [Shaw, Yu, 2022], HaploMaker [Fruzangohar et al., 2022], HaploSync [Minio et al., 2022], gcaPDA [Xie et al., 2022], LongPhase [Lin et al., 2022], PolyHaplotyper [Voorrips, Tumino, 2022], SubPhaser [Jia et al., 2022] and others. However, some of similar software products have a longer history, as can be seen from Table 1.

Given that many plants are characterized by polyploidy, programs for assembling such genomes in a phased manner in table 1 are especially distinguished. As for other programs, some of them allow further improvement of the assembly produced, as well as visualization of the results, including those oriented to polyploid organisms.

A special mention deserves the new assembly program Verkko, or rather a pipeline of several software products, which was used to assemble the most complete diploid human genome, based on the reference genome T2T-CHM13, which allowed 20 out of 46 chromosomes to be restored with 99.9997% accuracy [Rautiainen et al., 2023]. Previously developed assemblers were designed for hybrid assembly of long and short reads, while Verkko takes long and ultra-long reads into analysis, and this is the success in phased chromosome-level assembly in T2T format.

Despite of that some of the above-mentioned programs were originally developed for assembling the human genome, but since the general structure of eukaryotic genomes is similar, they can theoretically be suitable for plant diploid genomes, although plant genomes usually carry more repetitive DNA. Moreover, many plants are polyploids, which makes assembly more difficult.

We will not consider in more detail all these programs for assembling and analyzing phased diploid genomes, since this is the topic of a separate article.

Table 1.

Some programs for assembly and analysis of phased diploid genomes of higher organisms

Approach	Program	Reference
Reference-based phasing	RefEditor	Yuan et al., 2015
	FALCON-Unzip	Chin et al., 2016
	HapCUT2	Edge et al., 2017
	WhatsHap	Garg et al., 2018
	Aquila	Zhou et al., 2021
	HaploMaker	Fruzangohar et al., 2022
	HaploSync	Minio et al., 2022
	LongPhase	Lin et al., 2022
	mLinker	Sinha, Zhang, 2023
	Verkko	Rautiainen et al., 2023
	JTK	Masutani et al., 2023
Polyploid genomes		
	HapTree-X	Berger et al., 2015
	Hap10	Majidian et al., 2020
	Ranbow	Moeinzadeh et al., 2020
	nPhase	Abou Saada et al., 2021
De novo assembly	Supernova	Weisenfeld et al., 2017
	SGVar	Tian et al., 2018
	Platanus-allee	Kajitani et al., 2019
	phasebook	Luo et al., 2021
	DipAsm	Garg et al., 2021
	hifiasm	Cheng et al., 2021
	SpecHap	Yu et al., 2021
	FALCON-Phase	Kronenberg et al., 2021
	WHdenovo	Garg et al., 2020
Polyploid genomes		
	POLYTE	Baaijens, Schonhuth, 2019
	ALLHiC	Zhang et al., 2019
	SDip	Heller et al., 2020
	flopp	Shaw, Yu, 2022
	SubPhaser	Jia et al., 2022
Trio-binning	TrioCanu	Koren et al., 2018
	HAPDeNovo	Zhou et al., 2018
	trioPhaser	Miller, Piccolo, 2021
Gamete-binning	gcaPDA	Xie et al., 2022
	GameteBinning_prac	Sun et al., 2023
Other programs	Odintifier	Samaniego Castruita et al., 2015
	HaploMerger2	Huang et al., 2017
	Purge Haplotigs	Roach et al., 2018
	Samovar	Darby et al., 2019
	purge_dups	Guan et al., 2020
	SVIM-asm	Heller, Vingron, 2020
	VISOR	Bolognini et al., 2020
	PRINCESS	Mahmoud et al., 2021
	Hapo-G	Aury, Istace, 2021
	HapSolo	Solares et al., 2021
Polyploid genomes		
	AltHap	Hashemi et al., 2018
	CompHapDet	Sankararaman et al., 2020
	PolyCluster	Mazrouee et al., 2020
	PolyHaplotyper	Voorrips, Tumino, 2022

Genomes and diploid genomes of plants

More than a thousand complete quasi-genomes of plants with varying degrees of completeness are already known, and approximately 40% of them are made on the chromosome-level assembly [Sun et al., 2022]. T2T-genomes have been fully or partially sequenced in recent years for several plant species - maize *Zea mays* [Liu et al., 2020], banana *Musa acuminata* [Belser et al., 2021], watermelon *Citrullus lanatus* [Deng et al., 2022], rice *Oryza sativa* [Zhang et al., 2022], Arabidopsis [Wang et al., 2022a], potato *Solanum tuberosum* [Yang et al., 2023], bitter melon *Momordica charantia* [Fu et al., 2023]. There is no doubt that this list will expand soon. Recently, it was described in detail how to sequence plant genomes in such a way that they end with the T2T assembly format [Naito, 2023]. Also, in another recent review, it is emphasized that the era of T2T plant genomes assembled without gaps from telomere to telomere has arrived [Gladman et al., 2023]. It is hard to disagree with this, but knowledge of diploid genomes and pan-genomes is also needed. Previously, in a review article dedicated to plant genomes with known sizes from 135 Mb to 16.9 Gb, assembled at the chromosome level, which were known at that time for 114 species, special attention was paid to the issue of prospective assembly of phased genomes [Shirasawa et al., 2021].

Compared to a thousand quasi-haploid plant genomes, relatively few genomes have been assembled in haplotype-resolved format so far. Nevertheless, almost hundred diploid phased genomes belonging to more than 60 species of plants have been assembled with varying degrees of completeness. Among them are various agricultural plants, ornamental plants, fruit trees and shrubs, forest trees, medicinal plants and others. These plants are mostly diploids, but there are also triploids, tetraploids (auto- and allo-), hexaploids and even octaploids. The sizes of the assembled genomes also vary considerably - from 135 Mb in arabidopsis to 21.6 Gb in artichoke. Moreover, in many studies it is noted that phased genome assembly helped to better understand some important traits characteristic of the studied plants, including their origin, evolution, domestication, yield, disease and pest resistance, metabolic pathways, and heterosis mechanisms.

It is not surprising that one of the first plants with a phased genome was the model plant with a small genome - *Arabidopsis thaliana*. For this, the technology of trio-sequencing of the hybrid and its parental forms was used, and for comparative analysis the reference genome TAIR10 [Chin et al., 2016] was used. The same paper also presents data on sequencing the diploid genome of grape *Vitis vinifera* variety Cabernet Sauvignon, characterizing by a high level of heterozygosity, which facilitated the assembly of phased sequences. And grapes need to be given special attention, since for no other species so many

phased, haplotype-resolved genomes have been assembled.

Grapevine became the fourth plant species for which their quasi-genomes were sequenced [The French-Italian Public Consortium for Grapevine Genome Characterization, 2007]. The recently published book The Grape Genome describes the strategies for sequencing and assembling grapevine genomes, including haplotyped ones [Figuerosa-Balderas et al., 2019]. After the first paper on phasing the diploid genome of Cabernet Sauvignon grape variety, work on its improvement was continued [Minio et al., 2017]. Later, diploid genomes of a whole range of varieties were sequenced by different authors [Roach et al., 2018; Vondras et al., 2019; Zhou et al., 2019; Minio et al., 2019; 2022; Solares et al., 2021; Maestri et al., 2022; Smit et al., 2020]. Due to the assembly of phased genomes for a number of grape varieties, further studies of genome polymorphism in many cases were carried out taking into account the obtained data. Thus, using the improved chromosome-scale Cabernet Savignon genome sequence and phased assembly of nine wild and cultivars of grape, it was found that dioecy, characteristic of wild species, was lost during domestication, and a putative region in the genome responsible for it was identified [Massonnet et al., 2020]. An approach for identifying grape clones based on amplicon sequencing was developed [Urra et al., 2023]. Comparing the phased genomes of Merlot grape variety and its parental forms, a characteristic chimerism for it was detected [Sichel et al., 2023]. In addition to *V.vinifera*, phased genomes have been obtained for other species of this genus - *V.riparia*, *V.rupestris*, *V.berlandieri* [Minio et al., 2022], for the hybrid *V.labruscana* × *V.vinifera* [Shirasawa et al., 2022] and for the resistant to powdery mildew species *Muscadinia rotundifoila* [Cochetel et al., 2021; Massonnet et al., 2022], for the sample Gf.99-03, originating from the also resistant to downy mildew disease grape species *V.amurensis*, having the variety "Michurinets" as one of the predecessors [Frommer et al., 2023]. Recently, based on the phased genomes of nine wild North American grape species *V.acerifolia, V.aestivalis, V.arizonica, V.berlandieri, V.girdiana, V.monticola, V.mustangensis, V.riparia, V.rupestris* their super-pangenome was created [Cochetel et al., 2023].

Another woody climbing shrub, for which a phased genome assembly was performed, became kiwi fruit *Actinidia chinensis*, as well as *A.latifolia*, and their diploid genomes were assembled in T2T format with high accuracy [Han et al., 2023]. Also, phased genomes have been assembled for a whole range of fruit crops: mango *Mangifera indica* [Mango Genome Consortium, 2021], persimmon *Diospyros kaki* [Li et al., 2023], lychee *Litchi chinensis* [Hu et al., 2022], sweet orange *Citrus sinensis* [Wu et al., 2022], *C.australis* [Nakandala et al., 2023], various cultivars of apple tree *Malus domestica* [Sun et al., 2020; Khan et al., 2022, Švara et al., 2023], hybrid cherry trees *Cerasus × yedoensis* [Shirasawa et

al., 2019] and *Cerasus × kanzakura* [Shirasawa et al., 2021a], pear *Pyrus bretschneideri* [Shi et al., 2019], apricot *Prunus armeniaca* [Campoy et al., 2020].

Other crops that are consumed in one form or another and for which phased genomes have been assembled are: sugarcane *Saccharum spontaneum* with 32 and 40 chromosomes [Zhang et al., 2018; 2022b], pineapple *Ananas comosus* [Nashima et al., 2022; Yow et al., 2022], triploid banana *Musa acuminata* [Huang et al., 2023; Liu et al., 2023a; Xie et al., 2023], tetraploid blueberry *Vaccinum corymbosum* [Colle et al., 2019; Mengist et al., 2023], octaploid strawberry *Fragaria × ananasa* [Cheng et al., 2021; Hardigan et al., 2021; Jin et al., 2023; Mao et al., 2023], tea *Camellia sinensis* [Wang et al., 2021; Zhang et al., 2021; 2021a], hop *Humulus lupulus* [Padgitt-Cobb et al., 2020], *Vanilla planifolia* [Piet et al., 2022], diploid potato *Solanum tuberosum* [Zhou et al., 2020a], tetraploid potato *S.tuberosum* [Bao et al., 2022; Hoopes et al., 2022; Shaw, Yu, 2022; Sun et al., 2022], hexaploid sweet potato *Ipomoea batatas* [Yang et al., 2017; Moeinzadeh et al., 2020; Yan et al., 2024], maize *Zea mays* [Cheng et al., 2021]; cassava *Manihot esculenta* [Mansfeld et al., 2021; Long et al., 2022; Qi et al., 2022; Landi et al., 2023], pepper *Capsicum annuum* [Hulse-Kemp et al., 2018], hybrid pepper *C.annuum* [Delorean et al., 2023]; artichoke *Helianthus tuberosus* [Wang et al., 2023b]. A phased diploid genome of the agricultural plant alfalfa *Medicago sativa*, used as animal feed, was also assembled [Chen et al., 2020; Shen et al., 2020; Long et al., 2022a]. Another herbaceous plants whose phased genomes were assembled are lawn grass *Zoysia japonica* [Zhou et al., 2020] and tetraploid Shepherd's Purse *Capsella bursa-pastoris* [Moeinzadeh et al., 2020].

From medicinal plants, diploid genomes were assembled for several species: ginger *Zingiber officinale* [Cheng et al., 2021a; Li et al., 2021], *Suaeda glauca* [Yi et al., 2022; Cheng et al., 2023], *Bupleurum chinense* [Zhang et al., 2022a], baiji *Bletilla striata* [Jiang et al., 2022], patchouli *Pogostemon cablin* [Shen et al., 2022], mugwort *Artemisia annua* [Liao et al., 2022], rhubarb *Rheum officinale* [Zhang et al., 2023], wheel wingnut *Cyclocarya paliurus* [Yu et al., 2023]; buzhaye *Microcos paniculata* [Liu et al., 2023].

Among ornamental plants, diploid genomes have been assembled for the following species: hydrangea *Hydrangea macophylla* [Nashima et al., 2021; Wu et al., 2023], threefork morning glory *Ipomoea trifida* [Zhou et al., 2020], hexaploid *Chrysanthemum* [Voorrips, Tumino, 2022], bermudagrass *Cynodon dactylon* [Zhang et al., 2022].

Among woody plants, diploid genomes have been determined for triploid poplar *Populus tomentosa* [Tong et al., 2022], hexaploid bamboo *Dendrocalamus latiflorus* [Zheng et al., 2022], interspecific hybrid *Eucalyptus grandis × E.urophylla* [Lötter et al., 2022], California

redwood *Sequoia sempervirens* [Cheng et al., 2021]; oak *Quercus glauca* [Luo et al., 2023], cork oak *Q.variabilis* [Wang et al., 2023], honey plant - paperbark tree *Melaleuca quinquenervia* [Chen et al., 2023].

Special attention should be paid to some species. In order to construct phased genome pear *P.bretschneideri*, it was decided to sequence DNA from 12 pollen grains (actually gametes), which also allowed to obtain information about the features of recombination occurring in meiosis [Shi et al., 2019]. The comparison of phased haplotypes with the reference genome confirmed its mosaic structure, which, however, is not surprising. At the same time, the authors noted previously impossible analysis of allelic effects in pear gene expression with the genome of a regular quasi-haploid assembly. Sequencing 445 pollen grains, other authors reconstructed the phased diploid genome of apricot *P.armeniaca* at the level of pseudochromosomes with 99% accuracy [Campoy et al., 2020]. By performing a phased assembly of 135 single sperm cells of tea, it was possible to establish some places of crossing over [Zhang et al., 2021].

No less interest is aroused by sequencing in a phased haplotype format of the genomes of a dioecious plant *Asparagus kiusianus*, belonging to male and female individuals, which, as expected, can help better understand the mechanisms that determine the sex of such plants [Shirasawa et al., 2022a].

There is no doubt that the number of plants whose diploid genomes will be sequenced will soon increase, as there is a certain trend in whole-genome sequencing to restore phased diploid genomes, for which a large number of specialized programs have already been created for their assembly, which has already been noted above.

We are also interested in sequencing the diploid genome of the wild diploid wheat *Triticum sinskajae*, whose draft quasi-genome of about 5 billion bp we have recently assembled. Earlier, we sequenced the chloroplast genome of this wheat species [Kuluev et al., 2020], as well as similar genome of the einkorn *T.monococcum*, in the collection plantings of which the Sinskaya wheat was found. The data we obtained indicate that these are different species of wheat [Kuluev et al., 2024]. In this regard, it should be noted that the full quasi-genome of *T.monococcum* was recently sequenced [Ahmed et al., 2023].

Conclusion

There are already many examples of diploid genomes for different higher organisms, including plants and this gives confidence that in the future haplotype-resolved sequencing will become massive and quite affordable although one is much more difficult than sequencing quasi-genomes. In a recent review [Kong et al., 2023], it is said that accurate, gap-less, telomere-to-telomere, and fully phased complex plant genomes could soon become routine. But this requires a fundamentally

different DNA sequencing. We have repeatedly drawn attention to the need for sequencing diploid genomes [Kuluev et al, 2020; Baymiev et al., 2022; Chemeris et al., 2023] and we are working on it.

Sequencing the entire DNA of the full set of chromosomes of higher organisms, collected in the format from telomere to telomere - T2T haplotype-resolved genome - is undoubtedly a new level of genomic knowledge that will provide very important information, missing now for most eukaryotic species. After the emergence of available technology for sequencing and assembling diploid genomes, all known quasi-genomes will have to be sequenced again and assembled in a phased T2T format.

And until this level is massively achieved, there can be no talk of a "post-genomic era" or "post-genomic technologies". At the same time, there is a firm confidence that such an era will come, and, as far as we can judge, quite soon.

References

1. Abou Saada O, Tsouris A, Eberlein C, Friedrich A, Schacherer J. nPhase: an accurate and contiguous phasing method for polyploids. *Genome Biol*. 2021. V.22(1). 126. doi: 10.1186/s13059-021-02342-x

2. Adams MD, Celniker SE, Holt RA. et al. The genome sequence of *Drosophila melanogaster. Science*. 2000. V.287. P.2185-2195. doi: 10.1126/science.287.5461.2185

3. Ahmed HI, Heuberger M, Schoen A et al. Einkorn genomics sheds light on history of the oldest domesticated wheat. *Nature*. 2023. V.620(7975). P.830-838. doi: 10.1038/s41586-023-06389-7

4. Aury JM, Istace B. Hapo-G, haplotype-aware polishing of genome assemblies with accurate reads. *NAR Genom Bioinform*. 2021. V.3(2). lqab034. doi: 10.1093/nargab/lqab034

5. Baaijens JA, Schönhuth A. Overlap graph-based generation of haplotigs for diploids and polyploids. *Bioinformatics*. 2019. V.35(21). P.4281-4289. doi: 10.1093/bioinformatics/btz255

6. Bao Z, Li C, Li G et al. Genome architecture and tetrasomic inheritance of autotetraploid potato. *Mol Plant*. 2023. V.16(11). 1866. doi: 10.1016/j.molp.2023.10.005

7. Baymiev AlKh, Kuluev AR, Matniyazov RT, Garafutdinov RR, Bayimiev AnKh, Gimalov FR, Chemeris DA, Kuluev BR, Chemeris AV. The variety of quantitative estimates of the DNA content in plant nuclei and their dispersion, some terms and concepts (genome, C-value, pangenome). *Biomics*. 2022. V.14(1). P. 79-100. doi: 10.31301/2221-6197.bmcs.2022-6 (In Russian)

8. Belser C, Baurens FC, Noel B. et al. Telomere-to-telomere gapless chromosomes of banana using nanopore sequencing. *Commun Biol*. 2021. V.4(1). 1047. doi: 10.1038/s42003-021-02559-3

9. Berger E, Yorukoglu D, Berger B. HapTree-X: An Integrative Bayesian Framework for Haplotype Reconstruction from Transcriptome and Genome Sequencing Data. *Res Comput Mol Biol*. 2015. V.9029. P.28-29. doi: 10.1007/978-3-319-16706-0_4

10. Bolognini D, Sanders A, Korbel JO. et al. VISOR: a versatile haplotype-aware structural variant simulator for short- and long-read sequencing. *Bioinformatics*. 2020. V.36(4). P.1267-1269. doi: 10.1093/bioinformatics/btz719

11. Campoy JA, Sun H, Goel M. et al. Gamete binning: chromosome-level and haplotype-resolved genome assembly enabled by high-throughput single-cell sequencing of gamete genomes. *Genome Biol*. 2020. V.21(1). 306. doi: 10.1186/s13059-020-02235-5

12. Chemeris AV, Akhunov ED, Vakhitov VA. DNA sequencing. Moscow. Nauka. 1999. 429 p. (In Russian)

13. Chemeris DA, Kuluev BR, Garafutdinov RR, Gerashchenkov GA, Baymiev AnKh, Baymiev AlKh, Chemeris AV. The study of pangenomes and super-pangenomes is a necessary requirement for the creation of promising intragenic and cisgenic plants, as well as the disclosure of the nature of heterosis. *Biomics*. 2023. V.15(3). P.167-203. DOI: 10.31301/2221-6197.bmcs.2023-17 (In Russian)

14. Chen H, Zeng Y, Yang Y et al. Allele-aware chromosome-level genome assembly and efficient transgene-free genome editing for the autotetraploid cultivated alfalfa. *Nat Commun*. 2020. V.11(1). 2494. doi: 10.1038/s41467-020-16338-x

15. Chen SH, Martino AM, Luo Z, Schwessinger B, Jones A, Tolessa T, Bragg JG, Tobias PA, Edwards RJ. A high-quality pseudo-phased genome for *Melaleuca quinquenervia* shows allelic diversity of NLR-type resistance genes. *Gigascience*. 2023. V.12. giad102. doi: 10.1093/gigascience/giad102

16. Cheng H, Concepcion GT, Feng X, Zhang H, Li H. Haplotype-resolved de novo assembly using phased assembly graphs with hifiasm. *Nat Methods*. 2021. V.18(2). P.170-175. doi: 10.1038/s41592-020-01056-5

17. Cheng SP, Jia KH, Liu H et al. Haplotype-resolved genome assembly and allele-specific gene expression in cultivated ginger. *Hortic Res*. 2021a. V.8(1). 188. doi: 10.1038/s41438-021-00599-8

18. Cheng Y, Sun J, Jiang M et al. Chromosome-scale genome sequence of *Suaeda glauca* sheds light on salt stress tolerance in halophytes. *Hortic Res*. 2023. V.10(9). uhad161. doi: 10.1093/hr/uhad161

19. Chin CS, Peluso P, Sedlazeck FJ. et al. Phased diploid genome assembly with single-molecule real-time sequencing. *Nat. Methods*. 2016. V.13(12). P.1050-1054. doi: 10.1038/nmeth.4035

20. Cochetel N, Minio A, Guarracino A, Garcia JF, Figueroa-Balderas R, Massonnet M, Kasuga T, Londo JP, Garrison E, Gaut BS, Cantu D. A super-pangenome of the

North American wild grape species. *Genome Biol*. 2023. V.24(1). 290. doi: 10.1186/s13059-023-03133-2

21. Cochetel N, Minio A, Massonnet M. et al. Diploid chromosome-scale assembly of the *Muscadinia rotundifolia* genome supports chromosome fusion and disease resistance gene expansion during *Vitis* and *Muscadinia* divergence. *G3 (Bethesda)*. 2021. V.11(4). jkab033. doi: 10.1093/g3journal/jkab033

22. Colle M, Leisner C, Wai CM. et al. Haplotype-phased genome and evolution of phytonutrient pathways of tetraploid blueberry. *Gigascience*. 2019. V.8(3). giz012. doi: 10.1093/gigascience/giz012

23. Darby CA, Fitch JR, Brennan PJ. et al. Samovar: Single-Sample Mosaic Single-Nucleotide Variant Calling with Linked Reads. *iScience*. 2019. V.18. P.1-10. doi: 10.1016/j.isci.2019.05.037

24. Delorean EE, Youngblood RC, Simpson SA. et al. Representing true plant genomes: haplotype-resolved hybrid pepper genome with trio-binning. *Front Plant Sci*. 2023. V.14. 184112. doi: 10.3389/fpls.2023.1184112

25. Deng Y, Liu S, Zhang Y. et al. A telomere-to-telomere gap-free reference genome of watermelon and its mutation library provide important resources for gene discovery and breeding. *Mol Plant*. 2022. V.15(8). P.1268-1284. doi: 10.1016/j.molp.2022.06.010

26. Ebler J, Haukness M, Pesout T, Marschall T, Paten B. Haplotype-aware diplotyping from noisy long reads. *Genome Biol*. 2019. V.20(1). 116. doi: 10.1186/s13059-019-1709-0

27. Edge P, Bafna V, Bansal V. HapCUT2: robust and accurate haplotype assembly for diverse sequencing technologies. *Genome Res*. 2017. V.27(5). P.801-812. doi: 10.1101/gr.213462.116

28. Figueroa-Balderas R, Minio A, Morales-Cruz A. et al. Strategies for Sequencing and Assembling Grapevine Genomes. In: Cantu, D., Walker, M. (eds) The Grape Genome. 2019. P.77-88. doi: 10.1007/978-3-030-18601-2_5

29. French-Italian Public Consortium for Grapevine Genome Characterization. The grapevine genome sequence suggests ancestral hexaploidization in major angiosperm phyla. *Nature*. 2007. V.449(7161). P.463-467. doi: 10.1038/nature06148

30. Frommer B, Müllner S, Holtgräwe D. et al. Phased grapevine genome sequence of an Rpv12 carrier for biotechnological exploration of resistance to *Plasmopara viticola*. *Front Plant Sci*. 2023. V.14. 1180982. doi: 10.3389/fpls.2023.1180982

31. Fruzangohar M, Timmins WA, Kravchuk O, Taylor J. HaploMaker: An improved algorithm for rapid haplotype assembly of genomic sequences. *Gigascience*. 2022. V.11. giac038. doi: 10.1093/gigascience/giac038

32. Fu A, Zheng Y, Guo J. et al. Telomere-to-telomere genome assembly of bitter melon (*Momordica charantia* L. var. abbreviata Ser.) reveals fruit development, composition and ripening genetic characteristics. *Hortic Res*. 2022. V.10(1). uhac228. doi: 10.1093/hr/uhac228

33. Garg S. Computational methods for chromosome-scale haplotype reconstruction. *Genome Biol*. 2021. V.22(1). 101. doi: 10.1186/s13059-021-02328-9

34. Garg S, Aach J, Li H, Sebenius I, Durbin R, Church G. A haplotype-aware de novo assembly of related individuals using pedigree sequence graph. *Bioinformatics*. 2020. V.36(8). P.2385-2392. doi: 10.1093/bioinformatics/btz942

35. Garg S, Fungtammasan A, Carroll A. et al. Chromosome-scale, haplotype-resolved assembly of human genomes. *Nat Biotechnol*. 2021. V.39(3). P.309-312. doi: 10.1038/s41587-020-0711-0

36. Garg S, Rautiainen M, Novak AM, Garrison E, Durbin R, Marschall T. A graph-based approach to diploid genome assembly. *Bioinformatics*. 2018. V.34(13). P.i105-i114. doi: 10.1093/bioinformatics/bty279

37. Gilbert W, Maxam A. The nucleotide sequence of the lac operator. *Proc Natl Acad Sci USA*. 1973. V.70(12). P.3581-3584. doi: 10.1073/pnas.70.12.3581

38. 48. Gladman N, Goodwin S, Chougule K, McCombie WR, Ware D. Era of gapless plant genomes: innovations in sequencing and mapping technologies revolutionize genomics and breeding. *Curr Opin Biotechnol*. 2023. V.79. 102886. doi: 10.1016/j.copbio.2022.102886

39. Guan D, McCarthy SA, Wood J. et al. Identifying and removing haplotypic duplication in primary genome assemblies. *Bioinformatics*. 2020. V.36(9). P.2896-2898. doi: 10.1093/bioinformatics/btaa025

40. Guk JY, Jung M-J, Choi J-W. et al. De novo phasing resolves haplotype sequences in complex plant genomes. *Plant Biotechnol J*. 2022. V.20(6). P.1031-1041. doi: 10.1111/pbi.13815

41. Hameed A, Poznanski P, Nadolska-Orczyk A. et al. Graph Pangenomes Track Genetic Variants for Crop Improvement. *Int J Mol Sci*. 2022. V.23(21). 13420. doi: 10.3390/ijms232113420

42. Han X, Zhang Y, Zhang Q. et al. Two haplotype-resolved, gap-free genome assemblies for *Actinidia latifolia* and *Actinidia chinensis* shed light on the regulatory mechanisms of vitamin C and sucrose metabolism in kiwifruit. *Mol Plant*. 2023. V.16(2). P.452-470. doi: 10.1016/j.molp.2022.12.022

43. Hardigan MA, Feldmann MJ, Picot DDA. et al. Blueprint for Phasing and Assembling the Genomes of Heterozygous Polyploids: Application to the Octoploid Genome of Strawberry. *bioRxiv*. 2021.11.03.467115. doi: 10.1101/2021.11.03.467115

44. Hashemi A, Zhu B, Vikalo H. Sparse Tensor Decomposition for Haplotype Assembly of Diploids and Polyploids. *BMC Genomics*. 2018. V.19(Suppl 4). 191. doi: 10.1186/s12864-018-4551-y

45. Heller D, Vingron M. SVIM-asm: Structural variant detection from haploid and diploid genome assemblies. *Bioinformatics*. 2020. 36(22-23). P.5519–5521. doi: 10.1093/bioinformatics/btaa1034

46. Heller D, Vingron M, Church G, Li H, Garg S. SDip: A novel graph-based approach to haplotype-aware assembly based structural variant calling in targeted segmental duplications sequencing. *bioRxiv* 2020.02.25.964445. doi: 10.1101/2020.02.25.964445

47. Hoopes G, Meng X, Hamilton JP et al. Phased, chromosome-scale genome assemblies of tetraploid potato reveal a complex genome, transcriptome, and predicted proteome landscape underpinning genetic diversity. *Mol Plant*. 2022. V.15(3). P.520-536. doi: 10.1016/j.molp.2022.01.003

48. Hu G, Feng J, Xiang X. et al. Two divergent haplotypes from a highly heterozygous lychee genome suggest independent domestication events for early and late-maturing cultivars. *Nat Genet*. 2022. V.54(1). P.73-83. doi: 10.1038/s41588-021-00971-3

49. Huang HR, Liu X, Arshad R, Wang X, Li WM, Zhou Y, Ge XJ. Telomere-to-telomere haplotype-resolved reference genome reveals subgenome divergence and disease resistance in triploid Cavendish banana. *Hortic Res*. 2023. V.10(9). uhad153. doi: 10.1093/hr/uhad153

50. Huang S, Kang M, Xu A. HaploMerger2: rebuilding both haploid sub-assemblies from high-heterozygosity diploid genome assembly. *Bioinformatics*. 2017. V.33(16). P.2577-2579. doi: 10.1093/bioinformatics/btx220

51. Hulse-Kemp AM, Maheshwari S, Stoffel K. et al. Reference quality assembly of the 3.5-Gb genome of *Capsicum annuum* from a single linked-read library. *Hortic Res*. 2018. V.5. 4. doi: 10.1038/s41438-017-0011-0

52. Jander G, Norris SR, Rounsley SD, Bush DF, Levin IM, Last RL. Arabidopsis map-based cloning in the post-genome era. *Plant Physiol*. 2002. V.129(2). P.440-450. doi: 10.1104/pp.003533

53. Jia KH, Wang ZX, Wang L et al. SubPhaser: a robust allopolyploid subgenome phasing method based on subgenome-specific k-mers. *New Phytol*. 2022. V.235(2). P.801-809. doi: 10.1111/nph.18173

54. Jiang L, Lin M, Wang H et al. Haplotype-resolved genome assembly of *Bletilla striata* (Thunb.) Reichb.f. to elucidate medicinal value. *Plant J*. 2022. V.111(5). P.1340-1353. doi: 10.1111/tpj.15892

55. Jin X, Du H, Zhu C et al. Haplotype-resolved genomes of wild octoploid progenitors illuminate genomic diversifications from wild relatives to cultivated strawberry. *Nat Plants*. 2023. V.9(8). P.1252-1266. doi: 10.1038/s41477-023-01473-2

56. Kajitani R, Yoshimura D, Okuno M. et al. Platanus-allee is a de novo haplotype assembler enabling a comprehensive access to divergent .heterozygous regions. *Nat Commun*. 2019. V.10(1). P.1702. doi: 10.1038/s41467-019-09575-2

57. Khan A, Carey SB, Serrano A, Zhang H, Hargarten H, Hale H, Harkess A, Honaas L. A phased, chromosome-scale genome of 'Honeycrisp' apple (*Malus domestica*). *GigaByte*. 2022. gigabyte69. doi: 10.46471/gigabyte.69

58. Khan AW, Garg V, Roorkiwal M, Golicz AA, Edwards D, Varshney RK. Super-Pangenome by Integrating the Wild Side of a Species for Accelerated Crop Improvement. *Trends Plant Sci*. 2020. V.25(2). P.148-158. doi: 10.1016/j.tplants.2019.10.012

59. Kong W, Wang Y, Zhang S, Yu J, Zhang X. Recent Advances in Assembly of Plant Complex Genomes. *Genomics Proteomics Bioinformatics*. 2023. V.21(3). P. 427-439. doi: 10.1016/j.gpb.2023.04.004

60. Koren S, Rhie A, Walenz BP. et al. De novo assembly of haplotype-resolved genomes with trio binning. *Nat Biotechnol*. 2018. V. 22. 10.1038/nbt.4277. doi: 10.1038/nbt.4277

61. Kronenberg ZN, Rhie A, Koren S. et al. Extended haplotype-phasing of long-read de novo genome assemblies using Hi-C. *Nat Commun*. 2021. V.12(1). 1935. doi: 10.1038/s41467-020-20536-y

62. Kuluev A., Kuluev B., Chemeris A. Sequencing and analysis of complete chloroplast genomes of einkorn wheats *Triticum sinskajae* and *Triticum monococcum* accession k-20970. *Genetic Resources and Crop Evolution*. 2024. (In Press)

63. Kuluev BR, Baymiev AnKh, Gerashchenkov GA, Yunusbaev UB, Garafutdinov RR, Alekseev YaI, Baymiev AlKh, Chemeris AV. One hundred years of haploid genomes. Now time comes for diploid genomes. *Biomics*. 2020. V. 12(4). P. 411-434. DOI: 10.31301/2221-6197.bmcs.2020-33 (In Russian)

64. Lai J, Li R, Xu X et al. Genome-wide patterns of genetic variation among elite maize inbred lines. *Nat Genet*. 2010. V.42(11). P.1027-1030. doi: 10.1038/ng.684

65. Lam HM, Xu X, Liu X. et al. Resequencing of 31 wild and cultivated soybean genomes identifies patterns of genetic diversity and selection. *Nat Genet*. 2010. V.42. P.1053–1059. doi: 10.1038/ng.715

66. Landi M, Shah T, Falquet L, Niazi A, Stavolone L, Bongcam-Rudloff E, Gisel A. Haplotype-resolved genome of heterozygous African cassava cultivar TMEB117 (*Manihot esculenta*). *Sci Data*. 2023. V.10(1). 887. doi: 10.1038/s41597-023-02800-0

67. Li H, Sun P, Wang Y. et al. Allele-aware chromosome-level genome assembly of the autohexaploid *Diospyros kaki* Thunb. *Sci Data*. 2023. V.10(1). 270. doi: 10.1038/s41597-023-02175-2

68. Li HL, Wu L, Dong Z et al. Haplotype-resolved genome of diploid ginger (*Zingiber officinale*) and its unique gingerol biosynthetic pathway. *Hortic Res*. 2021. V.8(1). 189. doi: 10.1038/s41438-021-00627-7

69. Li W, Liu J, Zhang H. et al. Plant pan-genomics: recent advances, new challenges, and roads ahead. *J Genet*

Genomics. 2022. V.49(9). P.833-846. doi: 10.1016/j.jgg.2022.06.004

70. Liao B, Shen X, Xiang L. et al. Allele-aware chromosome-level genome assembly of *Artemisia annua* reveals the correlation between ADS expansion and artemisinin yield. *Mol Plant*. 2022. V.15(8). P.1310-1328. doi: 10.1016/j.molp.2022.05.013

71. Lin JH, Chen LC, Yu SC, Huang YT. LongPhase: an ultra-fast chromosome-scale phasing algorithm for small and large variants. *Bioinformatics*. 2022. V.38(7). P.1816-1822. doi: 10.1093/bioinformatics/btac058

72. Liu D, Tian X, Shao S, Ma Y, Zhang R. Haplotype-resolved chromosomal-level genome assembly of Buzhaye (*Microcos paniculata*). *Sci Data*. 2023. V.10(1). 901. doi: 10.1038/s41597-023-02821-9

73. Liu J, Seetharam AS, Chougule K. et al. Gapless assembly of maize chromosomes using long-read technologies. *Genome Biol*. 2020. V.21(1). 121. doi: 10.1186/s13059-020-02029-9

74. Liu X, Arshad R, Wang X. et al. The phased telomere-to-telomere reference genome of *Musa acuminata*, a main contributor to banana cultivars. *Sci Data*. 2023a. V.10(1). 631. doi: 10.1038/s41597-023-02546-9

75. Long EM, Bradbury PJ, Romay MC, Buckler ES, Robbins KR. Genome-wide imputation using the practical haplotype graph in the heterozygous crop cassava. *G3 (Bethesda)*. 2022. V.12(1). jkab383. doi: 10.1093/g3journal/jkab383

76. Long R, Zhang F, Zhang Z. et al. Genome Assembly of Alfalfa Cultivar Zhongmu-4 and Identification of SNPs Associated with Agronomic Traits. *Genomics Proteomics Bioinformatics*. 2022a. V.20(1). P.14-28. doi: 10.1016/j.gpb.2022.01.002

77. Lötter A, Duong TA, Candotti J, Mizrachi E, Wegrzyn JL, Myburg AA. Haplogenome assembly reveals structural variation in Eucalyptus interspecific hybrids. *Gigascience*. 2022. V.12:giad064. doi: 10.1093/gigascience/giad064

78. Luo CS, Li TT, Jiang XL. et al. High-quality haplotype-resolved genome assembly for ring-cup oak (*Quercus glauca*) provides insight into oaks demographic dynamics. *Mol Ecol Resour*. 2023. e13914. doi: 10.1111/1755-0998.13914

79. Luo X., Kang X., Schönhuth A. phasebook: haplotype-aware de novo assembly of diploid genomes from long reads. *Genome Biol*. 2021. V.22(1). 299. doi: 10.1186/s13059-021-02512-x

80. Maestri S, Gambino G, Lopatriello G. et al. 'Nebbiolo' genome assembly allows surveying the occurrence and functional implications of genomic structural variations in grapevines (*Vitis vinifera* L.). *BMC Genomics*. 2022. V.23(1). 159. doi: 10.1186/s12864-022-08389-9

81. Mahmoud M, Doddapaneni H, Timp W, Sedlazeck FJ. PRINCESS: comprehensive detection of haplotype resolved SNVs, SVs, and methylation. *Genome Biol*. 2021. V.22(1). 268. doi: 10.1186/s13059-021-02486-w

82. Majidian S, Kahaei MH, de Ridder D. Hap10: reconstructing accurate and long polyploid haplotypes using linked reads. *BMC Bioinformatics*. 2020. V.21(1). 253. doi: 10.1186/s12859-020-03584-5

83. Mango Genome Consortium. The 'Tommy Atkins' mango genome reveals candidate genes for fruit quality. *BMC Plant Biol*. 2021. V.21(1). 108. doi: 10.1186/s12870-021-02858-1

84. Mansfeld BN, Boyher A, Berry JC. et al. Large structural variations in the haplotype-resolved African cassava genome. *Plant J*. 2021. V.108(6). P.1830-1848. doi: 10.1111/tpj.15543

85. 105. Mao J, Wang Y, Wang B. et al. High-quality haplotype-resolved genome assembly of cultivated octoploid strawberry. *Hortic Res*. 2023. V.10(1). uhad002. doi: 10.1093/hr/uhad002

86. Massonnet M, Cochetel N, Minio A. et al. The genetic basis of sex determination in grapes. *Nat Commun*. 2020. V.11(1). 2902. doi: 10.1038/s41467-020-16700-z

87. Massonnet M, Vondras AM, Cochetel N. Et al. Haplotype-resolved powdery mildew resistance loci reveal the impact of heterozygous structural variation on NLR genes in *Muscadinia rotundifolia*. *G3 (Bethesda)*. 2022. V.12(8). jkac148. doi: 10.1093/g3journal/jkac148

88. Masutani B, Suzuki Y, Suzuki Y, Morishita S. JTK: targeted diploid genome assembler. *Bioinformatics*. 2023. V.39(7). btad398. doi: 10.1093/bioinformatics/btad398

89. Maxam AM, Gilbert W. A new method for sequencing DNA. *Proc Natl Acad Sci USA*. 1977. V.74(2). P.560-564. doi: 10.1073/pnas.74.2.560

90. Mazrouee S, Wang W. PolyCluster: Minimum Fragment Disagreement Clustering for Polyploid Phasing. *IEEE/ACM Trans Comput Biol Bioinform*. 2020. V.17(1). P.264-277. doi: 10.1109/TCBB.2018.2858803

91. Mengist MF, Bostan H, De Paola D. et al. Autopolyploid inheritance and a heterozygous reciprocal translocation shape chromosome genetic behavior in tetraploid blueberry (*Vaccinium corymbosum*). *New Phytol*. 2023. V.237(3). P.1024-1039. doi: 10.1111/nph.18428

92. Michael TP, VanBuren R. Building near-complete plant genomes. *Curr Opin Plant Biol*. 2020 V.54. P.26-33. doi: 10.1016/j.pbi.2019.12.009

93. Miller DB, Piccolo SR. trioPhaser: using Mendelian inheritance logic to improve genomic phasing of trios. *BMC Bioinformatics*. 2021. V.22(1). 559. doi: 10.1186/s12859-021-04470-4

94. Minio A, Cochetel N, Massonnet M. et al. HiFi chromosome-scale diploid assemblies of the grape rootstocks 110R, Kober 5BB, and 101-14 Mgt. *Sci Data*. 2022. V.9(1). 660. doi: 10.1038/s41597-022-01753-0

95.   Minio A, Cochetel N, Vondras AM, Massonnet M, Cantu D. Assembly of complete diploid-phased chromosomes from draft genome sequences. *G3 (Bethesda).* 2022a. V.12(8). jkac143. doi: 10.1093/g3journal/jkac143

96.   Minio A, Lin J, Gaut BS, Cantu D. How Single Molecule Real-Time Sequencing and Haplotype Phasing Have Enabled Reference-Grade Diploid Genome Assembly of Wine Grapes. *Front Plant Sci.* 2017. V.8. 826. doi: 10.3389/fpls.2017.00826

97.   Minio A, Massonnet M, Figuerosa-Balderas R, Castro A, Cantu D. Diploid Genome Assembly of the Wine Grape Carmenere. *G3 (Bethesda).* 2019. V.9(5). P.1331-1337. doi: 10.1534/g3.119.400030

98.   Moeinzadeh MH, Yang J, Muzychenko E. et al. Ranbow: A fast and accurate method for polyploid haplotype reconstruction. *PLoS Comput Biol.* 2020. V.16(5). e1007843. doi: 10.1371/journal.pcbi.1007843

99.   Morgante M, De Paoli E, Radovic S. Transposable elements and the plant pan-genomes. *Curr Opin Plant Biol.* 2007. V.10(2). P.149-155. doi: 10.1016/j.pbi.2007.02.001

100.   Naithani S, Deng CH, Sahu SK, Jaiswal P. Exploring Pan-Genomes: An Overview of Resources and Tools for Unraveling Structure, Function, and Evolution of Crop Genes and Genomes. *Biomolecules.* 2023. V.13(9). 1403. doi: 10.3390/biom13091403

101.   Naito K. How to Sequence and Assemble Plant Genomes. *Methods Mol Biol.* 2023. V.2632. P.57-77. doi: 10.1007/978-1-0716-2996-3_5

102.   Nakandala U, Masouleh AK, Smith MW. et al. Haplotype resolved chromosome level genome assembly of Citrus australis reveals disease resistance and other citrus specific genes. *Hortic Res.* 2023. V.10(5). uhad058. doi: 10.1093/hr/uhad058

103.   Nashima K, Shirasawa K, Ghelfi A. et al. Genome sequence of *Hydrangea macrophylla* and its application in analysis of the double flower phenotype. *DNA Res.* 2021. V.28(1). dsaa026. doi: 10.1093/dnares/dsaa026

104.   Nashima K., Shirasawa K., Isobe S. et al. Gene prediction for leaf margin phenotype and fruit flesh color in pineapple (*Ananas comosus*) using haplotype-resolved genome sequencing. *Plant J.* 2022. V.110(3). P.720-734. doi: 10.1111/tpj.15699

105.   Nowak R. Entering the postgenome era. *Science.* 1995. V.270. P.368-371.

106.   Padgitt-Cobb LK, Kingan SB, Wells J. et al. A draft phased assembly of the diploid Cascade hop (*Humulus lupulus*) genome. *Plant Genome.* 2021. V.14(1). e20072. doi: 10.1002/tpg2.20072

107.   Pan W, Zhao Y, Xu Y, Zhou F. WinHAP2: an extremely fast haplotype phasing program for long genotype sequences. *BMC Bioinformatics.* 2014. V.15. 164. doi: 10.1186/1471-2105-15-164

108.   Petereit J, Bayer PE, Thomas WJW, Tay Fernandez CG, Amas J, Zhang Y, Batley J, Edwards D. Pangenomics and Crop Genome Adaptation in a Changing Climate. *Plants (Basel).* 2022. V.11(15). 1949. doi: 10.3390/plants11151949

109.   Piet Q, Droc G, Marande W. et al. A chromosome-level, haplotype-phased *Vanilla planifolia* genome highlights the challenge of partial endoreplication for accurate whole-genome assembly. *Plant Commun.* 2022. V.3(5). 100330. doi: 10.1016/j.xplc.2022.100330

110.   Qi W, Lim Y-W, Patrignani A. et al. The haplotype-resolved chromosome pairs of a heterozygous diploid African cassava cultivar reveal novel pan-genome and allele-specific transcriptome features. *Gigascience.* 2022. V.11. giac028. doi: 10.1093/gigascience/giac028

111.   Rautiainen M, Nurk S, Walenz BP. et al. Telomere-to-telomere assembly of diploid chromosomes with Verkko. *Nat Biotechnol.* 2023. V.41(10). P.1474-1482. doi: 10.1038/s41587-023-01662-6

112.   Roach MJ, Johnson DL, Bohlmann J, van Vuuren HJJ, Jones SJM, Pretorius IS, Schmidt SA, Borneman AR. Population sequencing reveals clonal diversity and ancestral inbreeding in the grapevine cultivar Chardonnay. *PLoS Genet.* 2018. V.14(11). e1007807. doi: 10.1371/journal.pgen.1007807

113.   Roach MJ, Schmidt SA, Borneman AR. Purge Haplotigs: allelic contig reassignment for third-gen diploid genome assemblies. *BMC Bioinformatics.* 2018. V.19(1). 460. doi: 10.1186/s12859-018-2485-7

114.   Samaniego Castruita JA, Zepeda Mendoza ML, Barnett R. Et al. Odintifier--A computational method for identifying insertions of organellar origin from modern and ancient high-throughput sequencing data based on haplotype phasing. *BMC Bioinformatics.* 2015. V.16(1). 232. doi: 10.1186/s12859-015-0682-1

115.   Sanger F, Donelson JE, Coulson AR, Kössel H, Fischer D. Use of DNA polymerase I primed by a synthetic oligonucleotide to determine a nucleotide sequence in phage fl DNA. *Proc Natl Acad Sci USA.* 1973. V.70(4). P.1209-1213. doi: 10.1073/pnas.70.4.1209

116.   Sanger F, Nicklen S, Coulson AR. DNA sequencing with chain-terminating inhibitors. *Proc Natl Acad Sci USA.* 1977. V.74(12). P.5463-5467. doi: 10.1073/pnas.74.12.5463

117.   Sankararaman A, Vikalo H, Baccelli F. ComHapDet: a spatial community detection algorithm for haplotype assembly. *BMC Genomics.* 2020. V.21(Suppl 9). 586. doi: 10.1186/s12864-020-06935-x

118.   Shaw J, Yu YW. flopp: Extremely Fast Long-Read Polyploid Haplotype Phasing by Uniform Tree Partitioning. *J Comput Biol.* 2022. V.29(2). P.195-211. doi: 10.1089/cmb.2021.0436

119.   Shen C, Du H, Chen Z. et al. The Chromosome-Level Genome Sequence of the Autotetraploid Alfalfa and Resequencing of Core Germplasms Provide Genomic Resources for Alfalfa Research. *Mol Plant.* 2020. V.13(9). P.1250-1261. doi: 10.1016/j.molp.2020.07.003

120. Shen Y, Li W, Zeng Y. et al. Chromosome-level and haplotype-resolved genome provides insight into the tetraploid hybrid origin of patchouli. *Nat Commun*. 2022. V.13(1). 3511. doi: 10.1038/s41467-022-31121-w

121. Shendure J, Balasubramanian S, Church GM, Gilbert W, Rogers J, Schloss JA, Waterston RH. DNA sequencing at 40: past, present and future. *Nature*. 2017. V.550(7676). P.345-353. doi: 10.1038/nature24286

122. Shi D, Wu J, Tang H. et al. Single-pollen-cell sequencing for gamete-based phased diploid genome assembly in plants. *Genome Res*. 2019. V.29(11). P.1889-1899. doi: 10.1101/gr.251033.119

123. Shi J, Tian Z, Lai J, Huang X. Plant pan-genomics and its applications. *Mol Plant*. 2023. V.16(1). P.168-186. doi: 10.1016/j.molp.2022.12.009

124. Shirasawa K, Esumi T, Hirakawa H. et al. Phased genome sequence of an interspecific hybrid flowering cherry, 'Somei-Yoshino' (*Cerasus × yedoensis*). *DNA Res*. 2019. V.26(5). P.379-389. doi: 10.1093/dnares/dsz016

125. Shirasawa K., Harada D., Hirakawa H., Isobe S., Kole C. Chromosome-level de novo genome assemblies of over 100 plant species. *Breed Sci*. 2021. V.71(2). P.117-124. doi: 10.1270/jsbbs.20146

126. Shirasawa K., Hirakawa H., Azuma A. et al. De novo whole-genome assembly in an interspecific hybrid table grape, 'Shine Muscat'. *DNA Res*. 2022. V.29(6). dsac040. doi: 10.1093/dnares/dsac040

127. Shirasawa K, Itai A, Isobe S. Genome sequencing and analysis of two early-flowering cherry (*Cerasus × kanzakura*) varieties, 'Kawazu-zakura' and 'Atami-zakura'. *DNA Res*. 2021a. V.28(6). dsab026. doi: 10.1093/dnares/dsab026

128. Shirasawa K, Ueta S, Murakami K, Abdelrahman M, Kanno A, Isobe S. Chromosome-scale haplotype-phased genome assemblies of the male and female lines of wild asparagus (*Asparagus kiusianus*), a dioecious plant species. *DNA Res*. 2022a. V.29(1). dsac002. doi: 10.1093/dnares/dsac002

129. Sichel V, Sarah G, Girollet N, Laucou V, Roux C, Roques M, Mournet P, Cunff LL, Bert PF, This P, Lacombe T. Chimeras in Merlot grapevine revealed by phased assembly. *BMC Genomics*. 2023. V.24(1). 396. doi: 10.1186/s12864-023-09453-8

130. Sinha S, Zhang CZ. Determining Complete Chromosomal Haplotypes by mLinker. *Methods Mol Biol*. 2023. V.2590. P.149-159. doi: 10.1007/978-1-0716-2819-5_10

131. Smith HH. Effects of Genome Balance, Polyploidy, and Single Extra Chromosomes on Size in *Nicotiana*. *Genetics*. 1943. V.28(3). P.227-236.

132. Smit SJ, Vivier MA, Young PR. Comparative (Within Species) Genomics of the *Vitis vinifera* L. Terpene Synthase Family to Explore the Impact of Genotypic Variation Using Phased Diploid Genomes. *Front Genet*. 2020. V.11. 421. doi: 10.3389/fgene.2020.00421

133. Solares EA, Tao Y, Long AD, Gaut BS. HapSolo: an optimization approach for removing secondary haplotigs during diploid genome assembly and scaffolding. *BMC Bioinformatics*. 2021. V.22(1). 9. doi: 10.1186/s12859-020-03939-y

134. Švara A, Sun H, Fei Z, Khan A. Chromosome-level phased genome assembly of 'Antonovka' identified candidate apple scab resistance genes highly homologous to HcrVf2 and HcrVf1 on linkage group 1. *G3 (Bethesda)*. 2023. V.4. jkad253. doi: 10.1093/g3journal/jkad253

135. Sun H, Campoy JA, Schneeberger K. Gamete Binning to Achieve Haplotype-Resolved Genome Assembly. *Methods Mol Biol*. 2023. V.2590. P.201-218. doi: 10.1007/978-1-0716-2819-5_13

136. Sun H, Jiao WB, Krause K. et al. Chromosome-scale and haplotype-resolved genome assembly of a tetraploid potato cultivar. *Nat Genet*. 2022. V.54(3). P.342-348. doi: 10.1038/s41588-022-01015-0

137. Sun X, Jiao C, Schwaninger H. et al. Phased diploid genome assemblies and pan-genomes provide insights into the genetic history of apple domestication. *Nat Genet*. 2020. V.52(12). P.1423-1432. doi: 10.1038/s41588-020-00723-9

138. Sun Y, Shang L, Zhu QH, Fan L, Guo L. Twenty years of plant genome sequencing: achievements and challenges. *Trends Plant Sci*. 2022. V.27(4). P.391-401. doi: 10.1016/j.tplants.2021.10.006

139. Tay Fernandez CG, Nestor BJ, Danilevicz MF, Gill M, Petereit J, Bayer PE, Finnegan PM, Batley J, Edwards D. Pangenomes as a Resource to Accelerate Breeding of Under-Utilised Crop Species. *Int J Mol Sci*. 2022. V.23(5). 2671. doi: 10.3390/ijms23052671

140. Tettelin H, Masignani V, Cieslewicz MJ. et al. Genome analysis of multiple pathogenic isolates of Streptococcus agalactiae: implications for the microbial "pan-genome". *Proc Natl Acad Sci USA*. 2005. V.102(39). P.13950-13955. doi: 10.1073/pnas.0506758102

141. The Arabidopsis Genome Initiative. Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana*. *Nature*. 2000. V. 408. P.796–815. doi: 10.1038/35048692

142. The C.elegans Sequencing Consortium. Sequence and analysis of the genome of *C.elegans*. *Science*. 1998. V.282. P. 2012–2018.

143. Tian S, Yan H, Klee EW. et al. Comparative analysis of de novo assemblers for variation discovery in personal genomes. *Brief Bioinform*. 2018. V.19(5). P.893-904. doi: 10.1093/bib/bbx037

144. Tong S, Wang Y, Chen N. et al. PtoNF-YC9-SRMT-PtoRD26 module regulates the high saline tolerance of a triploid poplar. *Genome Biol*. 2022. V.23(1). 148. doi: 10.1186/s13059-022-02718-7

145. Urra C, Sanhueza D, Pavez C. et al. Identification of grapevine clones via high-throughput amplicon

sequencing: a proof-of-concept study. *G3 (Bethesda)*. 2023. V.13(9). jkad145. doi: 10.1093/g3journal/jkad145

146.   Vondras AM, Minio A, Blanco-Ulate B. et al. The genomic diversification of grapevine clones. *BMC Genomics*. 2019. V.20(1). 972. doi: 10.1186/s12864-019-6211-2

147.   Voorrips RE, Tumino G. PolyHaplotyper: haplotyping in polyploids based on bi-allelic marker dosage data. *BMC Bioinformatics*. 2022. V.23(1). 442. doi: 10.1186/s12859-022-04989-0

148.   Wang B, Yang X, Jia Y. et al. High-quality *Arabidopsis thaliana* Genome Assembly with Nanopore and HiFi Long Reads. *Genomics Proteomics Bioinformatics*. 2022. V.20(1). P.4-13. doi: 10.1016/j.gpb.2021.08.003

149.   Wang P, Yu J, Jin S. et al. Genetic basis of high aroma and stress tolerance in the oolong tea cultivar genome. *Hortic Res*. 2021. V.8(1). 107. doi: 10.1038/s41438-021-00542-x

150.   Wang L, Li LL, Chen L. et al. Telomere-to-telomere and haplotype-resolved genome assembly of the Chinese cork oak (*Quercus variabilis*). *Front Plant Sci*. 2023. V.14. 1290913. doi: 10.3389/fpls.2023.1290913

151.   Wang S, Qian YQ, Zhao RP, Chen LL, Song JM. Graph-based pan-genomes: increased opportunities in plant genomics. *J Exp Bot*. 2023a. V.74(1). P.24-39. doi: 10.1093/jxb/erac412

152.   Wang S, Wang A, Chen R. et al. Haplotype-resolved chromosome-level genome of hexaploid Jerusalem artichoke provides insights into its origin, evolution, and inulin metabolism. *Plant Commun*. 2023b. 100767. doi: 10.1016/j.xplc.2023.100767

153.   Weisenfeld NI, Kumar V, Shah P, Church DM, Jaffe DB. Direct determination of diploid genome sequences. *Genome Res*. 2017. V.27(5). P.757-767. doi: 10.1101/gr.214874.116

154.   Wu B, Yu Q, Deng Z, Duan Y, Luo F, Gmitter FG Jr. A chromosome-level phased genome enabling allele-level studies in sweet orange: a case study on citrus Huanglongbing tolerance. *Hortic Res*. 2022. V.10(1). uhac247. doi: 10.1093/hr/uhac247

155.   Wu R, Kaiser AD. Structure and base sequence in the cohesive ends of bacteriophage lambda DNA. *J Mol Biol*. 1968. V.35(3). P.523-537. doi: 10.1016/s0022-2836(68)80012-9

156.   Wu X, Simpson SA, Youngblood RC, Liu XF, Scheffler BE, Rinehart TA, Alexander LW, Hulse-Kemp AM. Two haplotype-resolved genomes reveal important flower traits in bigleaf hydrangea (*Hydrangea macrophylla*) and insights into Asterid evolution. *Hortic Res*. 2023. V.10(12). uhad217. doi: 10.1093/hr/uhad217

157.   Xie M, Yang L, Jiang C. et al. gcaPDA: a haplotype-resolved diploid assembler. *BMC Bioinformatics*. 2022. V.23(1). 68. doi: 10.1186/s12859-022-04591-4

158.   Xie WZ, Zheng YY, He W. et al. Two haplotype-resolved genome assemblies for AAB allotriploid bananas provide insights into banana subgenome asymmetric evolution and *Fusarium* wilt control. *Plant Commun*. 2023. 100766. doi: 10.1016/j.xplc.2023.100766

159.   Yan M, Li M, Wang Y. et al. Haplotype-based phylogenetic analysis and population genomics uncover the origin and domestication of sweetpotato. *Mol. Plant*. 2024. doi: 10.1016/j.molp.2023.12.019

160.   Yang J, Moeinzadeh MH, Kuhl H. et al. Haplotype-resolved sweet potato genome traces back its hexaploidization history. *Nat Plants*. 2017. V.3(9). P.696-703. doi: 10.1038/s41477-017-0002-z

161.   Yang X, Zhang L, Guo X. et al. The gap-free potato genome assembly reveals large tandem gene clusters of agronomical importance in highly repeated genomic regions. *Mol Plant*. 2023. V.16(2). P.314-317. doi: 10.1016/j.molp.2022.12.010

162.   Yi L, Sa R, Zhao S. et al. Chromosome-Scale, Haplotype-Resolved Genome Assembly of *Suaeda Glauca*. *Front Genet*. 2022. V.13. 884081. doi: 10.3389/fgene.2022.884081

163.   Yow AG, Bostan H, Castanera R. et al. Improved High-Quality Genome Assembly and Annotation of Pineapple (*Ananas comosus*) Cultivar MD2 Revealed Extensive Haplotype Diversity and Diversified FRS/FRF Gene Family. *Genes (Basel)*. 2021. V.13(1). 52. doi: 10.3390/genes13010052

164.   Yu RM, Zhang N, Zhang BW. Et al. Genomic insights into biased allele loss and increased gene numbers after genome duplication in autotetraploid *Cyclocarya paliurus*. *BMC Biol*. 2023. V.21(1). 168. doi: 10.1186/s12915-023-01668-1

165.   Yu Y, Chen L, Miao X, Li SC. SpecHap: a diploid phasing algorithm based on spectral graph theory. *Nucleic Acids Res*. 2021. V.49(19). e114. doi: 10.1093/nar/gkab709

166.   Yuan S, Johnston HR, Zhang G, Li Y, Hu YJ, Qin ZS. One Size Doesn't Fit All - RefEditor: Building Personalized Diploid Reference Genome to Improve Read Mapping and Genotype Calling in Next Generation Sequencing Studies. *PLoS Comput Biol*. 2015. V.11(8). e1004448. doi: 10.1371/journal.pcbi.1004448

167.   Zhang B, Chen S, Liu J, Yan YB, Chen J, Li D, Liu JY. A High-Quality Haplotype-Resolved Genome of Common Bermudagrass (*Cynodon dactylon* L.) Provides Insights Into Polyploid Genome Stability and Prostrate Growth. *Front Plant Sci*. 2022. V.13. 890980. doi: 10.3389/fpls.2022.890980

168.   Zhang H, He Q, Xing L. et al. The haplotype-resolved genome assembly of autotetraploid rhubarb *Rheum officinale* provides insights into its genome evolution and massive accumulation of anthraquinones. *Plant Commun*. 2023. V.26. 100677. doi: 10.1016/j.xplc.2023.100677

169. Zhang Q, Li M, Chen X et al. Chromosome-Level Genome Assembly of *Bupleurum chinense* DC Provides Insights Into the Saikosaponin Biosynthesis. *Front Genet.* 2022a. V.13. 878431. doi: 10.3389/fgene.2022.878431

170. Zhang Q, Qi Y, Pan H. et al. Genomic insights into the recent chromosome reduction of autopolyploid sugarcane *Saccharum spontaneum*. *Nat Genet.* 2022b. V.54(6). P.885-896. doi: 10.1038/s41588-022-01084-1

171. Zhang W, Luo C, Scossa F, Zhang Q, Usadel B, Fernie AR, Mei H, Wen W. A phased genome based on single sperm sequencing reveals crossover pattern and complex relatedness in tea plants. *Plant J.* 2021. V.105(1). P.197-208. doi: 10.1111/tpj.15051

172. Zhang X, Chen S, Shi L. et al. Haplotype-resolved genome assembly provides insights into evolutionary history of the tea plant *Camellia sinensis*. *Nat Genet.* 2021a. V.53(8). P.1250-1259. doi: 10.1038/s41588-021-00895-y

173. Zhang X., Wu R., Wang Y., Yu J., Tang H. Unzipping haplotypes in diploid and polyploid genomes. *Comput Struct Biotechnol J.* 2020. V.18. P.66-72. doi: 10.1016/j.csbj.2019.11.011

174. Zhang X, Zhang S, Zhao Q, Ming R, Tang H. Assembly of allele-aware, chromosomal-scale autopolyploid genomes based on Hi-C data. *Nat Plants.* 2019. V.5(8). P.833-845. doi: 10.1038/s41477-019-0487-8

175. Zheng Y, Yang D, Rong J. et al. Allele-aware chromosome-scale assembly of the allopolyploid genome of hexaploid Ma bamboo (*Dendrocalamus latiflorus* Munro). *J Integr Plant Biol.* 2022. V.64(3). P.649-670. doi: 10.1111/jipb.13217

176. Zhou C, Olukolu B, Gemenet DC. et al. Assembly of whole-chromosome pseudomolecules for polyploid plant genomes using outbred mapping populations. *Nat Genet.* 2020. V.52(11). P.1256-1264. doi: 10.1038/s41588-020-00717-7

177. Zhou Q, Tang D, Huang W. et al. Haplotype-resolved genome analyses of a heterozygous diploid potato. *Nat Genet.* 2020a. V.52(10). P.1018-1023. doi: 10.1038/s41588-020-0699-x

178. Zhou X, Batzoglou S, Sidow A, Zhang L. HAPDeNovo: a haplotype-based approach for filtering and phasing de novo mutations in linked read sequencing data. *BMC Genomics.* 2018. V.19(1). 467. doi: 10.1186/s12864-018-4867-7

179. Zhou X, Zhang L, Weng Z, Dill DL, Sidow A. Aquila enables reference-assisted diploid personal genome assembly and comprehensive variant detection based on linked reads. *Nat Commun.* 2021. V.12(1). 1077. doi: 10.1038/s41467-021-21395-x

180. Zhou Y, Minio A, Massonnet M, Solares E, Lv Y, Beridze T, Cantu D, Gaut BS. The population genetics of structural variants in grapevine domestication. *Nat Plants.* 2019. V.5(9). P.965-979. doi: 10.1038/s41477-019-0507-8

181. Zubov VV, Chemeris DA, Vasilov RG, Kurochkin VE, Alekseev YaI. Brief history of high-throughput nucleic acid sequencing methods. *Biomics.* 2021. V.13(1). P. 27-46. DOI: 10.31301/2221-6197.bmcs.2021-4 (In Russian)