Vavilovskii Zhurnal Genetiki i Selektsii
Vavilov Journal of Genetics and Breeding. 2025;29(3):358-369

doi 10.18699/vjgb-25-39

Modern methods in peach (*Prunus persica*) genome research

I.V. Rozanova [1, 2], E.A. Vodiasova [2]

[1] Federal Research Center the N.I. Vavilov All-Russian Institute of Plant Genetic Resources (VIR), St. Petersburg, Russia
[2] The Nikitsky Botanical Gardens – National Scientific Centre of RAS, Nikita, Yalta, Republic of Crimea, Russia
i.rozanova@vir.nw.ru

Abstract. Peach (*Prunus persica* (L.) Batsch) is one of the main agricultural stone fruit crops of the family Rosaceae. Modern breeding is aimed at improving the quality of the fruit, extending the period of its production, increasing its resistance to unfavorable environmental conditions and reducing the total cost of production of cultivated varieties. However, peach breeding is an extremely long process: it takes 10–15 years from hybridization of the parental forms to obtaining fruit-bearing trees. Research into peach varieties as donors of desirable traits began in the 1980s. The first version of the peach genome was presented in 2013, and its appearance contributed to the identification and localization of loci, followed by the identification of candidate genes that control the desired trait. The development of NGS has accelerated the development of methods based on the use of diagnostic DNA markers. Approaches that allow accelerating classical breeding processes include marker-oriented selection (MOS) and genomic selection. In order to develop DNA markers associated with the traits under investigation, it is necessary to carry out preliminary mapping of loci controlling economically desirable traits and to develop linkage maps. SNP-chip approaches and genotyping by sequencing (GBS) methods are being developed. In recent years, genome-wide association analysis (GWAS) has been actively used to identify genomic loci associated with economically important traits, which requires screening of large samples of varieties for hundreds and thousands of SNPs. Study on the pangenome has shown the need to analyze a larger number of samples, since there is still not enough data to identify polymorphic regions of the genome. The aim of this review was to systematize and summarize the major advances in peach genomic research over the last 40 years: linkage and physical map construction, development of different molecular markers, full genome sequencing for peach, and existing methods for genome-wide association studies with high-density SNP markers. This review provides a theoretical basis for future GWAS analysis in order to identify high-performance markers of economically valuable traits for peach and to develop genomic selection of this crop.

Key words: *Prunus persica*; GWAS; selection; genotyping; SNP

For citation: Rozanova I.V., Vodiasova E.A. Modern methods in peach (*Prunus persica*) genome research. *Vavilovskii Zhurnal Genetiki i Selektsii = Vavilov J Genet Breed*. 2025;29(3):358-369. doi 10.18699/vjgb-25-399

Funding. The study is supported by the Kurchatov Genomic Centre of the NBG–NSC (No. 075-15-2019-1670).

Современные методы в исследованиях генома персика (*Prunus persica*)

И.В. Розанова [1, 2], Е.А. Водясова [2]

[1] Федеральный исследовательский центр Всероссийский институт генетических ресурсов растений им. Н.И. Вавилова (ВИР), Санкт-Петербург, Россия
[2] Никитский ботанический сад – Национальный научный центр РАН, п.г.т. Никита, Ялта, Республика Крым, Россия
i.rozanova@vir.nw.ru

Аннотация. Персик (*Prunus p*ersica (L.) Batsch) – одна из основных сельскохозяйственных плодовых косточковых культур семейства розоцветных. Современная селекция направлена на улучшение качества плодов, расширение сроков их получения, создание сортов с устойчивостью к неблагоприятным условиям среды и снижение общих затрат на производство культивируемых сортов. При этом селекция персика – долгий процесс: от гибридизации родительских форм до получения плодоносящих деревьев проходит 10–15 лет. Исследования сортов персика как доноров желаемых признаков начались с 1980-х годов. Первый вариант генома персика был представлен в 2013 г., и его появление способствовало определению и локализации локусов, с последующим обнаружением генов-кандидатов, под чьим контролем находится хозяйственно ценный признак. Развитие методов NGS ускорило продвижение подходов, основанных на применении диагностических ДНК-маркеров. К таким подходам, позволяющим ускорять процессы классической селекции, относятся маркер-ориентированная селекция и геномная селекция. Для того чтобы разработать ДНК-маркеры, ассоциированные с изучаемыми свойствами, необходимо провести предварительное картирование локусов, контролирующих хозяйственно ценные признаки, создать карты сцепления. Работы по пангеному показали необходимость анализировать

© Rozanova I.V., Vodiasova E.A., 2025
This work is licensed under a Creative Commons Attribution 4.0 License

большее количество образцов, так как до сих пор не хватает данных для нахождения полиморфных областей генома. Развиваются подходы использования SNP-чипов и методов генотипирования через секвенирование (GBS, genotyping-by-sequencing). В последние годы для обнаружения локусов генома, ассоциированных с хозяйственно ценными признаками, активно применяется метод полногеномного анализа ассоциаций (GWAS, genome-wide association studies), для использования которого требуется скрининг больших выборок сортов по сотням и тысячам SNP. Цель настоящего обзора – систематизация и обобщение основных достижений в области геномных исследований персика за последние 40 лет: построение карт сцепления и физических карт, получение различных молекулярных маркеров, полногеномное секвенирование для персика, а также описание существующих работ полногеномных исследований ассоциаций с маркерами SNP высокой плотности. Этот обзор обеспечивает теоретическую основу для проведения GWAS с целью выявления высокоэффективных маркеров хозяйственно ценных признаков для персика и развития геномной селекции этой культуры.
Ключевые слова: *Prunus persica*; GWAS; селекция; генотипирование; SNP

Introduction

Peach (*Prunus persica* L.) is one of the main agricultural stone crops of the temperate zone, consumed fresh and processed, contains high amounts of vitamins, minerals, fiber and antioxidant compounds, while being low in calories and therefore excellent for dietary menus. As a species, it originated about 2.5 million years ago in the southwestern Tibetan Plateau region of China, from where its domestication began 4,000–5,000 years ago (Yu Y. et al., 2018) (see the Figure).

The current peach gene pool is divided into three groups that are characterized by differing climatic growing conditions (see the Figure). These groups originated in China in different geographical regions. The Southern group originated in climates with mild winters and hot, humid summers. These peaches have a flat shape and are characterized by a slightly acidic "honey" flavor. The Northern group included peach genotypes found in regions with cold winters and hot, dry summers. These peaches are generally resistant to drought and cold, but are not adapted for growth in southern areas. The third group is found in the arid northwest of China. It includes nectarines and peaches with yellow flesh, in contrast to the white-flesh peaches typical of the rest of China (Scorza, 1991).

The crop spread to Europe more than 2,000 years ago, along ancient trade routes through Persia (Hesse, 1975; Byrne et al., 2012). Peach was introduced to the Americas by Spanish and Portuguese settlers 400 years ago (Hesse, 1975; Scorza, Okie, 1990; Faust, Timon, 1995).

By now, peach is the second most important temperate fruit crop after apple. More than 1,000 varieties of *P. persica* with different phenotypic variations for various traits such as shape, fruit size, flavor, flower type, etc. have been produced worldwide. According to FAO data (https://www.fao.org/faostat/en/#data/QCL), peach is now grown almost everywhere on all continents except Greenland, northern regions of Europe and some regions in central Africa (see the Figure).

There is little information in the literature about existing peach collections in the world. In the United States, the first



Distribution and origin of *P. persica*.

The arrows show the distribution routes of the peach. Countries where the peach currently grows and its production level are shown according to the color scheme. Origin of peach in China: the Southern group spans the Yangtze River; the Northern group is along the Yellow River; the third group originated in northwest China.

breeding programs appeared in the late 18th century (Hesse, 1975; Scorza, Okie, 1990; Faust, Timon, 1995). Since the founding of the USDA (United States Department of Agriculture) in 1889, more than 2,100 clones or seeds of peaches and nectarines were imported into America from China and other parts of the world. The collection housed at the USDA Plant Introduction Garden at Chico, in California, had about 700 unique peach accessions. It was used by most breeders but focused primarily on genotypes derived from crosses with a member of the southern group of Chinese cultivars, 'Chinese Cling' (Scorza et al., 1985). This variety was widely used for crop improvement. During the 1950–1960s, the collection gradually declined and was almost eliminated with the closure of the station in the late 1960s. Only 60 varieties survived and were transferred to farms in Byron, Ga., and Beltsville, Md. farms. These varieties have unique traits not found in the gene pool descending from 'Chinese Cling' (Werner, Okie, 1998). Currently, the gene pool of the peach population in the U.S. is considered to be the most impoverished.

In order to establish a peach cultivar collection in China, peach germplasm collection was initiated at the Zhengzhou Fruit Research Institute (ZFRI) in the 1960s. In 1986, the National Peach Collection was established, which consisted of more than 600 accessions by 2000 (Wang et al., 2001). To date, more than 1,200 peach accessions have been collected from around the world, including wild species, ancient cultivar populations, and modern cultivars (Lirong et al., 2020).

Two peach collections are located in northern Spain: the National Peach Collection Gene Pool "Centro de Investigación y Tecnología Agroalimentaria de Aragón" (CITA) and "Estación Experimental de Aula Dei" (EEAD-CSIC) (https://cita-aragon.es/en/history-mission-vision-and-aims/). The quantitative characteristics of the collections are not presented on the website.

In Russia, the largest peach collection is located in the Nikita Botanical Garden in the Crimea and has 790 peaches and 85 nectarines (Smykov et al., 2021).

Peach is diploid ($2n = 16$), self-pollinating, with a base chromosome number of eight and belongs to the Rosaceae family, subfamily Prunoideae (Bassi, Monet, 2008). It has a lower level of genetic variability compared to other *Prunus* cultivars. On the one hand, the ability to self-pollinate is a limiting factor in breeding programs, on the other hand, in combination with such biological features as small genome size (265 Mb) (Arumuganathan, Earle, 1991) and diploid set of chromosomes, and as a result of its economic value, peach is an excellent model for genomic studies of stone fruits of the Rosaceae family (Monet et al., 1996; Abbott et al., 2002). The genomes of different *Prunus* species are highly conserved (Dirlewanger et al., 2004), allowing many of the major genes and loci of peach and other *Prunus* species to be placed on the same genetic map (Abbott et al., 2008).

Traditional peach seedling breeding is a labor-intensive process that takes 10–15 years from the initial crossing to the emergence of a new cultivar (Bliss, 2010; Ru et al., 2015). In addition, peach breeding programs require significant acreage due to the large size of the trees, as well as financial resources to cover the ongoing costs for technical treatments such as spraying herbicides, insecticides and fungicides, planting, pruning, thinning and watering. With the development of genetics, studies on the genetic diversity of the crop began (Herrero et al., 1964), the use of diagnostic DNA markers was developed (Callahan et al., 1991; Lambert et al., 2016; Demirel et al., 2024), and the advent of NGS sequencing techniques (Micali et al., 2015; Kim et al., 2021) allowed the generation of high-quality whole genome sequences for genomic breeding approaches.

The aim of this review is to summarize the results of genetic and breeding works for the *P. persica* culture based on the application of NGS methods.

Methods before NGS: isoenzymes, DNA markers, first linkage maps

Despite the significant progress made by peach breeders over the past hundred years, traditional seedling breeding is a labor-intensive process since, in temperate climates, peach trees require at least three years to reach fruiting maturity before progeny fruit quality can be evaluated (Bliss, 2010). In the late 1980s, it was recognized that markers, the alleles of which have distinct differences at the phenotypic level, could be useful in the analysis of complex traits (Monet, 1988). And markers, which are closely linked to traits that appear late in development, can be valuable for early tree selection (Chaparro et al., 1994).

Research on peach varieties as donors of desirable traits began in the 1980s (see the Table). Protein markers, or isozymes (isoenzymes), were the first to be used as potential markers to identify particularly valuable hybrids. Isoenzymes are different variants of an enzyme (different amino acid sequence isoforms of the same enzyme) that differ in electrophoretic mobility. Isoenzyme analysis could be used to distinguish hybrids between plum and peach from plum offspring (Parfitt et al., 1985), peach and almond hybrids (Arulsekar et al., 1986a; Chaparro et al., 1987). G.E. Jr. Carter with colleagues showed that differences in protein structure were sufficient to distinguish each peach cultivar (Carter, Brock, 1980) or to reveal their similarities (Arulsekar et al., 1986b).

Over time, a sufficient number of morphological markers, the localization of which on the chromosome is known, have become available. In this case, by analyzing F_2 populations from crosses, it is possible to determine the chromosomal positions of isozyme loci relative to morphological markers. R.E. Durham with colleagues identified the presence of separate, independently inherited loci by examining isoenzymes such as diaphorase, malate dehydrogenase and peroxidase (Durham et al., 1987). A total of four phenotypic trait linkage groups with isoenzymes were known in the early 1990s (Bailey, French, 1949; Monet et al., 1985; Monet, 1988; Monet, Gibault, 1991). However, low frequency and other drawbacks have prevented the widespread use of protein markers in breeding programs.

With the development of sequencing methods, a technological breakthrough occurred, and new methods of DNA polymorphism analysis appeared, which led to the emergence of molecular markers characterized by high frequency of occurrence in the genome. A molecular marker (DNA marker) corresponds to a gene or genomic region having different variants (alleles) and associated with different phenotypic manifestations (Khlestkina, 2014). Based on their use, approaches have emerged that complement classical breeding methods by

History of the genetic and genomic *P. persica* research

Year	Significant studies	Reference
1980–1990	Morphological traits localization on peach chromosomes using isoenzyme analysis techniques	Bailey, French, 1949; Carter, Brock, 1980; Monet et al., 1985; Monet, 1988; Monet, Gibault, 1991
1992	Generation of the first map showing linkage of markers to peach traits (35 RFLP markers for five bonding groups)	Belthoff et al.,1993
1995	Creation of a reference genetic map (T × E)	Foolad et al., 1995
2012	The DNA chip IPSC 9K SNP v1 creation	Verde et al., 2012
2013	Creation of the first version of the peach reference genome Peach v 1.0 (dihaploid 'Lovell' variety)	International Peach Genome Initiative et al., 2013
2015	Application of the GBS method on the peach	Bielenberg et al., 2015
2017	Updated peach reference genome Peach v 2.0 (dihaploid 'Lovell' variety)	Verde et al., 2017
2019	The 18K SNP array v2 DNA chip creation	Gasic et al., 2019
2020	Peach pangenome establishment (100 accessions)	Cao et al., 2020
2021	The 'Chinese Cling' variety sequencing (the variety central to the development of cultivated peaches)	Cao et al., 2021

searching for DNA markers associated with valuable traits in order to accelerate the breeding process.

Economic traits such as productivity, quality, maturity, tolerance to biotic and abiotic stresses are quantitative traits and are polygenic. The search and labeling of polymorphic loci associated with quantitative traits, or, in other words, Quantitative Trait Loci (QTL), is an agronomically important task. By accumulating information on molecular markers, it is possible to create genetic maps, the purpose of which is to identify neutrally inherited markers in close proximity to genetic determinants (loci or genes) that control the manifestation of certain traits, including quantitative traits (Chesnokov, Artem'eva, 2011).

QTLs for traits such as ripening period, fruit weight, size and texture, pH, titratable acidity and soluble solids have been found in peach (Quarta et al., 2002). Fruit quality QTLs tend to cluster in several genomic regions, especially in linkage groups 4, 5 and 6 (Dirlewanger et al., 2009). Similarly, most QTL and disease resistance genes are also clustered (Pflieger et al., 1999). This observation suggests that (1) a small number of Mendelian factors can explain most of the genetic variability in fruit quality traits and (2) traits of different characteristics often share common QTLs (Dirlewanger et al., 1999). Consequently, common QTLs usually correspond to different closely related genes or to a single gene with pleiotropic effects on many traits influenced by the same physiological process (Quilot et al., 2004).

Linkage maps approximate the genomic position and genetic distances between markers using linkage analysis of genetic data (Paterson, 1996; Jones et al., 1997; Collard et al., 2005). The construction of a genetic linkage map is based on meiotic events, where genetic recombination occurs, leading to the development of recombinant genotypes. The lower the recombination frequency between molecular markers, the more likely they are to be linked and in the same linkage group (Paterson, 1996; Jones et al., 1997; Collard et al., 2005). Markers are called unlinked if their recombination frequency is greater than 50 % and, thus, they are located in different linkage groups. Recently, such calculations have been done using software, e. g. MapChart 2.2 (Voorrips, 2002), Mapmaker (Lander et al., 1987), TMAP (Cartwright et al., 2007), MapQTL (Van Ooijen, 2009), Joinmap (Van Ooijen, 2006).

The first maps showing the association between peach phenotypic traits and markers appeared in 1992 (Belthoff et al., 1992). L.E. Belthoff and colleagues (1992) developed a genetic map containing five linkage groups using 35 RFLP markers. J.X. Chaparro and colleagues developed a linkage map for peach using 83 RAPD markers and two isozymes (Chaparro et al., 1994). Then, by investigating the F_2 generation obtained from crosses between almond (cultivar 'Texas') and peach (cultivar 'Early Gold'), a genetic map (T × E) was created, which was later used as a *Prunus* reference map (Foolad et al., 1995; Joobeur et al., 1998; Pozzi, Vecchietti, 2009). It included all eight clutch groups and covered a total distance of 491 cM. Microsatellite markers (Simple Sequences Repeats, SSR) have been widely used as diagnostic DNA markers in peach research since the 2000s (Sosinski et al., 2000; Dirlewanger et al., 2002; Hong et al., 2013).

The *Prunus* 'T × E' reference map contains 1,947 anchor markers (i. e., evenly distributed throughout the *Prunus* genome) with known map locations (Howad et al., 2005; Dirlewanger et al., 2007; Pozzi, Vecchietti, 2009), which allowed comparisons between the peach genomes and the rest of the *Prunus* species. This facilitated the subsequent development of intraspecific maps for peach and other maps of interspecific relationships in *Prunus* (Howad et al., 2005; Dirlewanger et al., 2007; Pozzi, Vecchietti, 2009).

Currently, 70 linkage maps have been generated for peach and related interspecific hybrids, which can be found in the Rosaceae Genome Database (GDR; http://www.rosaceae.org/) (Jung et al., 2008, 2014) as well as in J.A. Salazar et al. (2013).

These genetic linkage maps continue to serve as effective tools for comparison with the *Prunus* 'T × E' reference map.

Peach genomic studies

To identify and localize loci, with the following identification of candidate genes under the control of an economically valuable qualitative or quantitative peach trait, it is necessary to obtain the complete genome of *P. persica* (Tanksley et al., 1989; Winter, Kahl, 1995; Paterson, 1996; Jones et al., 1997; Collard et al., 2005).

The first Peach v1.0 genome variant was submitted in 2013 (International Peach Genome Initiative et al., 2013). The target for the full genome sequencing was a dihaploid genotype of the cultivar 'Lovell' (Plov2_2N), which was read by the Sanger method with 8.5-fold coverage. Eight pseudomolecules, reflecting the eight peach chromosomes, were assembled. The resulting genome assembly (Peach v1.0), spanning 227.3 Mb, of which 218.4 Mb (96 %) was decoded, contained 27,852 annotated genes, with an average density of 1.22 genes per 10 Kb. In 2017, an updated Peach v2.0 reference map was published (Verde et al., 2017), constructed by repeated NGS sequencing of the 'Lovell' dihaploid sample on the Illumina platform. The sequence totaled 227.4 Mb, which is only slightly longer than the first variant, but deeper resequencing resulted in 225.7 Mb (99.2 %) of the sequence being decoded. The approximate positions of the centromeric regions of the chromosomes were determined based on repetitive regions with low gene concentration and low recombination frequency. Based on the reference genome sequence, researchers studied the evolutionary history of the peach fruit (Cao et al., 2014, 2020; Yu Y. et al., 2018), identified domestication regions (Cao et al., 2014; Akagi et al., 2016; Li Y. et al., 2019), and genes controlling economically valuable traits (Cao et al., 2014, 2016, 2019).

The Peach v2.0 physical map was correlated with four genetic maps: (1) 67 forms from a mapping population of an interspecific cross between almond and peach 'Texas' × 'Earligold' F_2 (T × E (Joobeur et al., 1998)); (2) 242 forms from a mapping population derived from IF7310828 × Ferganensis BC1 (P × F (Dettori et al., 2001)); (3) 305 seedlings from the mapping population 'Contender' × 'Ambra' F_2 (C × A (Eduardo et al., 2011)); (4) 62 hybrids from the cross 'Maria Dolce' × SD81 F_1 (MD × SD). The mapping strategy involved an approach using SSR and SNP markers targeting specific regions of the peach genome and a full-genome approach using the IPSC 9K SNP v1 chip (Verde et al., 2012). The loci derived from the genetic maps were mapped to physical positions using the MareyMap package (Siberchicot et al., 2017). For each linkage map used in this study (T × E, C × A, P × F and MD × SD), recombination rates were estimated as the ratio between genetic (cM) and physical (million base pairs) distances.

However, a single reference assembly does not reflect intraspecific variability, so there is a need to investigate the genetic variation of different peach cultivars and their wild relatives using pangenomic analysis. Similar work on pangenome construction has been carried out for many crops such as soybean (Li Y.H. et al., 2014; Liu Y. et al., 2020), rice (Zhao et al., 2018), sunflower (Hübner et al., 2019), tomato (Gao et al., 2019), barley (Jayakodi et al., 2020). In 2014, several accessions' genomes were comparatively analyzed. To assess the process of peach domestication, 11 peach accessions (including the dihaploid 'Lovell' PLov2-2N used for reference assembly as a control) and one each of *P. ferganensis, P. kansuensis, P. davidiana* and *P. mira* were re-sequenced. *P. ferganensis* is considered a wild undomesticated peach or, more likely, represents an intermediate variant in the peach domestication. Using a set of 953,357 high-quality SNPs identified in *P. persica* and *P. ferganensis* samples, nucleotide sequence diversity was assessed for eight collected chromosomes (International Peach Genome Initiative et al., 2013).

In the first chromosome, the number of genetic variants at polymorphic loci was minimal. The greatest diversity among SNPs was observed in the distal region of the short arm of chromosome 2 and in the distal region of the long arm of chromosome 4. The density of genes encoding receptor proteins from the family of conserved nucleotide-binding leucine-rich proteins (R-proteins), which are involved in immunity, was 5-fold higher on chromosome 2 than in the rest of the genome (Dodds, Rathjen, 2010). Immunity-related regions are rapidly evolving, so the diversity detected is natural. It is known from the literature that genes associated with fruit ripening are located on chromosome 4 (Eduardo et al., 2011; Dirlewanger et al., 2012).

Since the study included samples with different ripening times, there is a high level of variability in the region associated with ripening due to the given sampling parameters. However, genotypes of *P. kansuensis, P. davidiana* or *P. mira* mature at the same optimal time and no SNP diversity between regions was found in these species (International Peach Genome Initiative et al., 2013). Similarly, studies by sequencing of six related peach (*P. persica*) accessions were also conducted (Guan et al., 2019). The genomic variations identified showed that the comparison of different crop genotypes is effective for the development of DNA markers. These works support the need to analyze more accessions, as there are still insufficient data to identify the polymorphic regions of the genome.

The first peach pangenome, consisting of 100 sequenced samples of *P. persica*, was obtained in 2020 (Cao et al., 2020). Also, in this work, the *de novo* genomes of four wild peach relatives, *P. mira, P. davidiana, P. kansuensis* and *P. ferganensis*, were assembled. When the sequenced peach accessions were compared with the reference genome (Verde et al., 2017), an average of 3.4 % of reads in each accession failed to match the reference genome, and these reads were assembled *de novo* by the researchers. In total, an additional 2.52 Mb of new sequences containing 2,833 contigs (>500 bp) of potential significance were obtained. Additionally, 923 new genes were identified in the newly assembled sequences (Cao et al., 2020). The total number of genes in the pangenome was 27,796. Genes were divided into conserved genes which were common to all 100 samples (24,971, 89.9 %), and variable genes (2,803, 10.1 %), the presence of which was detected in less than 99 % of the samples (Cao et al., 2020).

Pangenomic analysis revealed the presence of resistance genes (R-genes) among the variable gene set. A similar situation has been observed in soybean (Li Y.H. et al., 2014) and rice (Zhao et al., 2018). It is hypothesized that variations in resistance (R) gene copy number may help explain differences in resistance between wild and cultivated accessions (Li Y.H.

et al., 2014). Also, using peach pangenome, we found that 63 % of ornamental, 88 % of local, and 91 % of improved cultivars had a set of "optional" four genes encoding geraniol-8-hydroxylases, which are involved in the biosynthesis of terpenes, which play an important role in plant life and have anticarcinogenic, antiseptic and antimicrobial effects. These genes may have been under positive selection pressure both during domestication and during the breeding process.

When comparing *P. persica* with four wild species of the genus *Prunus* collected *de novo* by K. Cao and colleagues, it was found that 34.7 % of all genes found based on homology for encoded proteins were represented in all five species. At the same time, species-specific genes were found in *P. mira* (543 specific genes), *P. davidiana* (485), *P. kansuensis* (194), *P. ferganensis* (197), and *P. persica* (320). Such studies allow the identification of genes that confer species-specific properties. For example, a nematode resistance gene was identified in *P. kansuensis* (Cao et al., 2020). Such work makes it possible to identify differences in the genomes of closely related species and varieties, which is necessary for the identification of genes responsible for valuable qualities and traits of plants.

In 2021, K. Cao and colleagues (Cao et al., 2021) sequenced the 'Chinese Cling' cultivar, which is very important historically and central to the cultivated peach development in Europe (Byrne et al., 2000), Japan (Yamamoto et al., 2003) and the USA (Aranzana et al., 2010). The assembled genome contained 247.33 million base pairs, representing 99.8 % of the putative genome. Its comparison with the 'Lovell' reference genome revealed 685,407 novel SNPs, 162,655 insertions and deletions, and 16,248 copy number variation (CNV) structural variants. Gene family analysis revealed a reduction in gene families involved in the biosynthesis of flavones, flavonols, flavonoids and monoterpenoids compared to the 'Lovell' variety genome.

Thus, the genomic approach allows the comparative analysis of varieties and identification of variable genes (or loci in the genome) that may be responsible for different varietal traits. Such studies remain essential for further development of genomic selection in peach.

New approaches in peach research with NGS

Polymorphism analysis methods have evolved from rather labor-intensive isoenzyme-related and RFLP methods to high-throughput sequencing methods. Comparative studies of genetic distances between peach accessions estimated using SNP and SSR markers have been conducted. In the early 2000s, methods using SSR markers became dominant. M.T. Hamblin and colleagues showed that 89 SSR markers did a better job of clustering the samples of the study sample of 259 maize inbred lines than a set of 847 SNP markers. The researchers concluded that a large number of polymorphic single nucleotide loci are needed for qualitative analysis using SNP markers (Hamblin et al., 2007). Currently, tens of thousands of SNP loci are being analyzed. J.M. Yu and colleagues (2009) calculated that the power of 1,000 SNPs is similar to that of 100 SSRs for estimating population structure and relatedness. At the same time, SSR markers remain a major option for screening plant genetic resource collections (Nybom, Lācis, 2021) and for passporting samples (Trifonova et al., 2021).

SNP markers have a higher distribution frequency in the genome compared to SSR markers, which makes them more functional when polymorphisms within specific genes are required for targeted studies. The first SNP detection technologies were *in silico* search for SNPs by analyzing EST databases followed by PCR-based validation (Batley et al., 2003), and SNP detection by resequencing transcripts using the Sanger method (Morozova, Marra, 2008). However, these methods were unable to detect SNPs in intergenic and non-coding regions. The advent of GBS approaches and the development of DNA chips have overcome the problems associated with the low throughput ability and high cost of SNP detection (Mardis, 2008) and now allow cost-effective and time-efficient detection of SNPs at significant loci. More and more diagnostic SNP markers are now being used in breeding programs. The use of insertions/deletions as markers is also common, but their reproducibility is lower than that of SNPs.

One common approach to SNP determination is genotyping using microarrays, or DNA chips. SNPs on the chip have been developed in such a way that it is possible to differentiate the samples under investigation in the pool of samples. DNA chips have been developed for many commercially important crops on two different platforms: the Illumina Infinium platform (6K for cherry (Peace et al., 2012), 8K for apple (Chagné et al., 2012), 18K for grape (Laucou et al, 2018) and 6K for avocado (Kuhn et al., 2019)) and the Axiom platform (480K SNPs for apple (Bianco et al., 2016), 68K for rose (Koning-Boucoiran et al., 2015), 700K for walnut (Marrano et al., 2019), and 70K and 200K for pear (Montanari et al., 2019; Li X. et al., 2019)).

In order to establish the medium-density Infinium SNP platform suitable for genotyping the peach gene pool, 56 breeding-significant peach accessions spanning the crop gene pool were selected. The samples selected were those used in international peach breeding programs, contributing to the breeding gene pool according to pedigree records, and based on parentage estimates from SSR studies showing genetic diversity. Over 1 million SNPs were obtained and tested, of which exactly 9,000 passed quality control, were genetically informative and formed the platform for genotyping, the first International Peach SNP Consortium (IPSC) peach 9K SNP array v1 chip (Verde et al., 2012). SNPs on the chip were distributed evenly across all eight chromosomes and the average spacing was 26.7 bp (Verde et al., 2012).

Platform validation was performed on 709 peach accessions comprising two independent evaluation samples: 232 accessions from the European Union and 479 accessions from the USA. The EU panel included 229 peach cultivars, and three wild species of the genus *Prunus* or their hybrids with peach. The US panel included pedigree varieties, breeding lines, and seedlings. Overall, the sampled material consisted of 45 % cultivars, 4 % improved breeding lines, and 51 % seedlings. Specimens clearly related to either peach or almond accounted for 82 and 2 %, respectively, while 16 % of the genotyped material was of interspecific origin (with almond).

In the next step, the peach 9K SNP array v1 platform was extended to 18K. The new chip included 9,000 SNPs from the previous version and 7,206 SNPs identified by sequencing 49 samples and uniformly distributed across all peach

chromosomes (Gasic et al., 2019). The uniform distribution of polymorphisms selected for the chip throughout the genome (the number of gaps smaller than 0.3 million base pairs reduced to 2 on the chromosomes 3 and 8) allows finding associations linked to the traits of interest.

Currently, genotyping by sequencing (GBS) has become the most common method of analyzing SNP markers for genome research. The term "GBS" is already used as an umbrella term for various NGS-based high-throughput genotyping methods under development (Rasheed et al., 2017). In plants, this method was first described by R.J. Elshire et al. in 2011 (2011).

Genotyping methods are used both for sequence determination and to identify associations between phenotype and genotype. Since the peach genome has now been sequenced, the identification of genomic regions associated with a trait can be performed immediately to search for candidate genes. The GBS method has been applied in peach research since 2015 (Bielenberg et al., 2015). Research in quantitative genetics is conducted equally using GBS (Cao et al., 2016, 2019; Guan et al., 2019; Li Y. et al., 2019; Meng et al., 2019; Guajardo et al., 2020; Thurow et al., 2020; Huang et al., 2021; Liu J. et al., 2021; Tan et al., 2021; Li X. et al., 2022, 2023), as well as using SNP chips (Micheletti et al., 2015; Akagi et al., 2016; Font i Forcada et al., 2019; Cirilli et al., 2021; da Silva Linge et al., 2021; Fu et al., 2021; Mas-Gómez et al., 2021, 2022).

Both GBS and SNP-chip genotyping have their advantages and disadvantages. For example, the diversity of biallelic SNPs collected at chip creation is limited, while the GBS method can cover and identify significant SNPs associated with a trait that are, however, not included in the chip set. Conversely, GBS often includes a large amount of missing data and coverage must be high enough to ensure reproducibility between the samples studied (Nybom, Lācis, 2021). GBS is currently used more frequently than SNP chips because this approach can be applied to crops for which the reference genome has not yet been sequenced. At least 96 samples are required for large-scale genotyping with GBS or with SNP chips (Zurn et al., 2020).

Analysis of associations between genomic loci and phenotypic traits

Today, modern technologies make it possible to perform genome-wide association studies (GWAS), the results of which are effectively used in breeding programs because they allow simultaneous genomic analysis of several hundred varieties for tens of thousands of loci, comparing the associations between different alleles and the trait of interest. By creating an appropriate sample, GWAS can identify loci for several economically valuable traits at once. This step expands the ability to select markers for agronomically important traits. In the future, the use of molecular markers will allow the selection of desired genotypes among breeding hybrids, which is actively used in marker-assisted breeding (MAB) programs (Khlestkina, 2014). The identification of significant associations facilitates the development of new markers, which can be used to set the required criteria for the variety to be developed.

In peach populations, due to the low level of genetic diversity, association mapping must consider linkage disequilibrium (LD), which is the non-random relationship between two alleles that causes certain allelic combinations to occur most frequently. The method is sensitive to the presence of a large number of related samples in the population structure, leading to spurious associations between phenotypes and marker loci (Mariette et al., 2010). Thus, if a particular combination of alleles confers an adaptive advantage, its frequency will increase relative to the frequency expected under random assignment. Several studies using SSR markers have been conducted in peach in varieties with different genetic backgrounds, and their results indicate that linkage disequilibrium is quite high in this crop.

Kinship between varieties and selection increase the level of linkage disequilibrium. It has been found to range from 6.01 to 20 cM (Aranzana et al., 2010; Cao et al., 2012; Font i Forcada et al., 2013). One strategy to deal with high linkage disequilibrium is to use SNPs that are not correlated with each other for analyses (e. g., taking $r^2 = 0.20$ as a measure of allelic association). Several algorithms exist to prune SNPs in this way or to reduce the degree of linkage disequilibrium between SNPs. Popular pruning strategies are implemented in PLINK 1.07/1.9 (Purcell et al., 2007), which sequentially scan the genome for correlated SNP pairs using only allele counting. As a result, only one representative SNP is retained for each region where highly correlated SNPs are present (Joiret et al., 2019).

The GWAS method has now identified genomic regions associated with many peach traits. Agronomic traits such as maturation, fruit pubescence, flesh colour, texture, flesh colour around the stone, fruit weight and soluble solids content are being studied (Micheletti et al., 2015; Cao et al., 2016; Elsadr, 2016; Font i Forcada et al., 2019; Li Y. et al., 2019; Liu H. et al., 2019; Thurow et al., 2020; Cirilli et al., 2021; da Silva Linge et al., 2021; Mas-Gómez et al., 2021, 2022; Tan et al., 2021; Li X. et al., 2023), as well as seed characteristics (kernel flavor) (Cao et al., 2016), pollen fertility traits (Huang et al., 2021), flower characteristics (Micheletti et al., 2015; Cao et al., 2016; Elsadr, 2016; Meng et al., 2019; Tan et al., 2021). There are works on peach resistance to various diseases (Fu et al., 2021; Li X. et al., 2022), cold and drought tolerance (Li Y. et al., 2019; Tan et al., 2021).

The above works demonstrate the potential value of the GWAS method for identifying new genomic regions associated with phenotypic traits of agricultural importance. This method can also be used to refine data on previously discovered QTLs (e. g., to more accurately determine the size of the locus under study) and facilitate the discovery of genes controlling the trait under investigation.

Conclusion

With the development of NGS approaches, several peach cultivars have been sequenced, providing a basis for whole-genome association studies. The large diversity of cultivars in existing collections allows not only to assess the diversity of the crop's gene pool, but also to search for marker-trait associations. Modern genotyping methods using GBS and SNP chips allow the identification of new markers that enrich the peach database. On the one hand, these new associations are of fundamental interest, contributing to the identification of peculiarities of genome evolution, individual development

of the peach tree and mechanisms of response to various environmental stimuli, and on the other hand, they are the basis for applied work aimed at developing effective markers and their use in obtaining new peach varieties with specified characteristics. This approach makes it possible to accelerate the breeding time of this stone fruit.

However, difficulties remain in the field of association mapping in peach breeding programs. This is mainly due to the fact that the number of samples in the collections studied should be at least 100 to reflect the degree of efficiency. In addition, the relatedness of the varieties and hybrids under study should be assessed beforehand when compiling the sample set. Thus, when working with peach collections, preliminary analysis of genetic diversity and relatedness is necessary, which is also better performed using SNPs.

References

Abbott A.G., Georgi L., Yvergniaux D., Wang Y., Blenda A., Reighard G., Inigo M., Sosinski B. Peach: the model genome for *Rosaceae*. *Acta Hortic*. 2002;575:145-155. doi 10.17660/ActaHortic.2002.575.14

Abbott A.G., Arús P., Scorza R. Genetic engineering and genomics. In: Layne D., Bassi D. (Eds) The Peach Botany, Production and Uses. London: CAB International, 2008;85-105. doi 10.1079/9781845933869.0085

Akagi T., Hanada T., Yaegaki H., Gradziel T.M., Tao R. Genome-wide view of genetic diversity reveals paths of selection and cultivar differentiation in peach domestication. *DNA Res*. 2016;23(3):271-282. doi 10.1093/dnares/dsw014

Aranzana M.J., Abbassi E.K., Howad W., Arús P. Genetic variation, population structure and linkage disequilibrium in peach commercial varieties. *BMC Genet*. 2010;11:69. doi 10.1186/1471-2156-11-69

Arulsekar S., Parfitt D.E., Kester D.E. Comparison of isozyme variability in peach and almond cultivars. *J Hered*. 1986a;77(4):272-274. doi 10.1093/oxfordjournals.jhered.a110235

Arulsekar S., Parfitt D.E., Beres W., Hansche P.E. Genetics of malate dehydrogenase isozymes in the peach. *J Hered*. 1986b;77(1):49-51. doi 10.1093/oxfordjournals.jhered.a110166

Arumuganathan K., Earle E.D. Nuclear DNA content of some important plant species. *Plant Mol Biol Rep*.1991;9:208-218. doi 10.1007/BF02672069

Bailey J.S., French A.P. The Inheritance of Certain Fruit and Foliage Characters in the Peach. Amherst, MA: University of Massachusetts Press, 1949

Bassi D., Monet R. Botany and taxonomy. In: Layne D.R., Bassi D. (Eds) The Peach: Botany, Production and Uses. Wallingford: CAB International, 2008;1-36. doi 10.1079/9781845933869.0001

Batley J., Barker G., O'Sullivan H., Edwards K.J., Edwards D. Mining for single nucleotide polymorphisms and insertions/deletions in maize expressed sequence tag data. *Plant Physiol*. 2003;132(1):84-91. doi 10.1104/pp.102.019422

Belthoff L.E., Ballard R., Abbott A., Morgens P., Callahan A., Scorza R., Baird W.V., Monet R. Development of a saturated linkage map of *Prunus persica* using molecular based marker systems. *Acta Hortic*. 1993;336:51-56. doi 10.17660/ActaHortic.1993.336.5

Bianco L., Cestaro A., Linsmith G., Muranty H., Denancé C., Théron A., Poncet C., … Davassi A., Laurens F., Velasco R., Durel C.E., Troggio M. Development and validation of the Axiom® Apple480K SNP genotyping array. *Plant J*. 2016;86(1):62-74. doi 10.1111/tpj.13145

Bielenberg D.G., Rauh B., Fan S., Gasic K., Abbott A.G., Reighard G.L., Okie W.R., Wells C.E. Genotyping by sequencing for SNP-based linkage map construction and QTL analysis of chilling requirement and bloom date in peach [*Prunus persica* (L.) Batsch]. *PloS One*. 2015;10(10):e0139406. doi 10.1371/journal.pone.0139406

Bliss F.A. Marker-assisted breeding in horticultural crops. *Acta Hort*. 2010;859:339-350. doi 10.17660/ActaHortic.2010.859.40

Byrne D.H., Sherman W.B., Bacon T.A. Stone fruit genetic pool and its exploitation for growing under warm winter conditions. In: Erez A. (Ed.) Temperate Fruit Crops in Warm Climates. Dordrecht: Springer, 2000;157-230. doi 10.1007/978-94-017-3215-4_8

Byrne D.H., Bassols M., Bassi D., Piagnani M., Gasic K., Reighard G., Moreno M., Pérez S. Peach. In: Badenes M.L., Byrne D.H. (Eds) Fruit Breeding. New York: Springer Science, 2012;505-570. doi 10.1007/978-1-4419-0763-9_14

Callahan A., Scorza R., Morgens P., Mante S., Cordts J., Cohen R. Breeding for cold hardiness: searching for genes to improve fruit quality in cold-hardy peach germplasm. *HortScience*. 1991;26(5):522-526. doi 10.21273/HORTSCI.26.5.522

Cao K., Wang L., Zhu G., Fang W., Chen C., Luo J. Genetic diversity, linkage disequilibrium, and association mapping analyses of peach (*Prunus persica*) landraces in China. *Tree Genet Genomes*. 2012;8(5):975-990. doi 10.1007/s11295-012-0477-8

Cao K., Zheng Z., Wang L., Liu X., Zhu G., Fang W., Cheng S., … Li Y., Li H., Guo J., Xu X., Wang J. Comparative population genomics reveals the domestication history of the peach, *Prunus persica*, and human influences on perennial fruit crops. *Genome Biol*. 2014;15:415. doi 10.1186/s13059-014-0415-1

Cao K.E., Zhou Z., Wang Q., Guo J., Zhao P., Zhu G., Fang W., Chen C., Wang X., Wang X., Tian Z., Wang L. Genome-wide association study of 12 agronomic traits in peach. *Nat Commun*. 2016;7(1):13246. doi 10.1038/ncomms13246

Cao K., Li Y., Deng C.H., Gardiner S.E., Zhu G., Fang W., Chen C., Wang X., Wang L. Comparative population genomics identified genomic regions and candidate genes associated with fruit domestication traits in peach. *Plant Biotechnol J*. 2019;17(10):1954-1970. doi 10.1111/pbi.13112

Cao K., Peng Z., Zhao X., Li Y., Liu K., Arus P., Zhu G., Deng S., Fang W., Chen C., Wang X., Wu J., Fei Z., Wang L. Pan-genome analyses of peach and its wild relatives provide insights into the genetics of disease resistance and species adaptation. *BioRxiv*. 2020. doi 10.1101/2020.07.13.200204

Cao K., Yang X., Li Y., Zhu G., Fang W., Chen C., Wang X., Wu J., Wang L. New high-quality peach (*Prunus persica* L. Batsch) genome assembly to analyze the molecular evolutionary mechanism of volatile compounds in peach fruits. *Plant J*. 2021;108(1):281-295. doi 10.1111/tpj.15439

Carter G.E. Jr., Brock M.M. Identification of peach cultivars through protein analysis. *HortScience*. 1980;15(3):292-293

Cartwright D.A., Troggio M., Velasco R., Gutin A. Genetic mapping in the presence of genotyping errors. *Genetics*. 2007;176(4):2521-2527. doi 10.1534/genetics.106.063982

Chagné D., Crowhurst R.N., Troggio M., Davey M.W., Gilmore B., Lawley C., Vanderzande S., … Wilhelm L., Van de Weg E., Gardiner S.E., Bassil N., Peace C. Genome-wide SNP detection, validation, and development of an 8K SNP array for apple. *PLoS One*. 2012;7(2):e31745. doi 10.1371/journal.pone.0031745

Chaparro J.X., Durham R.E., Moore G.A., Sherman W.B. Utilization of isozyme techniques to identify peach × 'Nonpareil' almond hybrids. *HortScience*. 1987;22(2):300-302. doi 10.21273/HORTSCI.22.2.300

Chaparro J.X., Werner D.J., O'Malley D., Sederoff R.R. Targeted mapping and linkage analysis of morphological isozyme, and RAPD markers in peach. *Theor Appl Genet*. 1994;87(7):805-815. doi 10.1007/BF00221132

Chesnokov Yu.V., Artem'eva A.M. Association mapping in plants (review). *Sel'skokhozyaystvennaya Biologiya = Agricultural Biology*. 2011;46(5):3-16 (in Russian)

Cirilli M., Baccichet I., Chiozzotto R., Silvestri C., Rossini L., Bassi D. Genetic and phenotypic analyses reveal major quantitative loci associated to fruit size and shape traits in a non-flat peach collection

(*P. persica* L. Batsch). *Hortic Res.* 2021;8:232. doi 10.1038/s41438-021-00661-5

Collard B.C.Y., Jahufer M.Z.Z., Brouwer J.B., Pang E.C.K. An introduction to markers, quantitative trait loci (QTL) mapping and marker-assisted selection for crop improvement: The basic concepts. *Euphytica*. 2005;142:169-196. doi 10.1007/s10681-005-1681-5

da Silva Linge C., Cai L., Fu W., Clark J., Worthington M., Rawandoozi Z., Byrne D.H., Gasic K. Multi-locus genome-wide association studies reveal fruit quality hotspots in peach genome. *Front Plant Sci.* 2021;12:644799. doi 10.3389/fpls.2021.644799

Demirel S., Pehluvan M., Aslantaş R. Evaluation of genetic diversity and population structure of peach (*Prunus persica* L.) genotypes using inter-simple sequence repeat (ISSR) markers. *Genet Resour Crop Evol*. 2024;71(3):1301-1312. doi 10.1007/s10722-023-01691-9

Dettori M.T., Quarta R., Verde I. A peach linkage map integrating RFLPs, SSRs, RAPDs, and morphological markers. *Genome.* 2001; 44(5):783-790. doi 10.1139/g01-065

Dirlewanger E., Moing A., Rothan C., Svanella L., Pronier V., Guye A., Plomion C., Monet R. Mapping QTLs controlling fruit quality in peach (*Prunus persica* (L.) Batsch). *Theor Appl Genet*. 1999;98: 18-31. doi 10.1007/s001220051035

Dirlewanger E., Cosson P., Tavaud M., Aranzana M., Poizat C., Zanetto A., Arús P., Laigret F. Development of microsatellite markers in peach [*Prunus persica* (L.) Batsch] and their use in genetic diversity analysis in peach and sweet cherry (*Prunus avium* L.). *Theor Appl Genet.* 2002;105(1):127-138. doi 10.1007/s00122-002-0867-7

Dirlewanger E., Graziano E., Joobeur T., Garriga-Calderé F., Cosson P., Howad W., Arús P. Comparative mapping and marker-assisted selection in Rosaceae fruit crops. *Proc Natl Acad Sci USA.* 2004;101(23): 9891-9896. doi 10.1073/pnas.0307937101

Dirlewanger E., Cosson P., Boudehri K., Renaud C., Capdeville G., Tauzin Y., Laigret F., Moing A. Development of a second-generation genetic linkage map for peach [*Prunus persica* (L.) Batsch] and characterization of morphological traits affecting flower and fruit. *Tree Genet Genomes.* 2007;3:1-13. doi 10.1007/s11295-006-0053-1

Dirlewanger E., Claverie J., Iezzoni A.F., Wünsch A. Sweet and sour cherries: linkage maps, QTL detection and marker assisted selection. In: Folta K.M., Gardiner S.E. (Eds) Genetics and Genomics of Rosaceae. Plant Genetics and Genomics: Crops and Models. Vol. 6. New York, NY: Springer, 2009;291-313. doi 10.1007/978-0-387-77491-6_14

Dirlewanger E., Quero-García J., Le Dantec L., Lambert P., Ruiz D., Dondini L., Illa E., Quilot-Turion B., Audergon J.M., Tartarini S., Letourmy P., Arús P. Comparison of the genetic determinism of two key phenological traits, flowering and maturity dates, in three *Prunus* species: peach, apricot and sweet cherry. *Heredity.* 2012;109(5): 280-292. doi 10.1038/hdy.2012.38

Dodds P.N., Rathjen J.P. Plant immunity: towards an integrated view of plant pathogen interactions. *Nat Rev Genet.* 2010;11(8):539-548. doi 10.1038/nrg2812

Durham R.E., Moore G.A., Sherman W.B. Isozyme banding patterns and their usefulness as genetic markers in peach. *J Am Soc Hortic Sci.* 1987;112(6):1013-1018. doi 10.21273/JASHS.112.6.1013

Eduardo I., Pacheco I., Chietera G., Bassi D., Pozzi C., Vecchietti A., Rossini L. QTL analysis of fruit quality traits in two peach intraspecific populations and importance of maturity date pleiotropic effect. *Tree Genet Genomes*. 2011;7:323-335. doi 10.1007/s11295-010-0334-6

Elsadr H. A genome wide association study of flowering and fruit quality traits in peach [(*Prunus persica* (L.) Batsch]: Doctoral dissertation. University of Guelph, 2016

Elshire R.J., Glaubitz J.C., Sun Q., Poland J.A., Kawamoto K., Buckler E.S., Mitchell S.E. A robust, simple genotyping-by-sequencing (GBS) aproach for high diversity species. *PloS One.* 2011;6(5): e19379. doi 10.1371/journal.pone.0019379

Faust M., Timon B. Origin and dissemination of the peach. In: Janick J. (Ed.) Horticultural Reviews. John Wiley & Sons, Inc., 1995;331-379. doi 10.1002/9780470650585.ch10

Font i Forcada C., Oraguzie N., Igartua E., Moreno M.Á., Gogorcena Y. Population structure and marker-trait associations for pomological traits in peach and nectarine cultivars. *Tree Genet Genomes.* 2013;9:331-349. doi 10.1007/s11295-012-0553-0

Font i Forcada C., Guajardo V., Chin-Wo S.R., Moreno M.Á. Association mapping analysis for fruit quality traits in *Prunus persica* using SNP markers. *Front Plant Sci*. 2019;9:2005. doi 10.3389/fpls.2018.02005

Foolad M.R., Arulsekar S., Becerra V., Bliss F.A. A genetic map of *Prunus* based on an interspecific cross between peach and almond. *Theor Appl Genet*. 1995;91:262-269. doi 10.1007/BF00220887

Fu W., da Silva Linge C., Gasic K. Genome-wide association study of brown rot (*Monilinia* spp.) tolerance in peach. *Front Plant Sci.* 2021;12:635914. doi 10.3389/fpls.2021.635914

Gao L., Gonda I., Sun H., Ma Q., Bao K., Tieman D.M., Burzynski-Chang E.A., … van der Knaap E., Huang S., Klee H.J., Giovannoni J.J., Fei Z. The tomato pan-genome uncovers new genes and a rare allele regulating fruit flavor. *Nat Genet.* 2019;51(6):1044-1051. doi 10.1038/s41588-019-0410-2

Gasic K., Da Silva Linge C., Bianco L., Troggio M., Rossini L., Bassi D., Aranzana M.J., Arus P., Verde I., Peace C., Iezzoni A. Development and evaluation of a 9K SNP addition to the peach IPSC 9K SNP array v1. *HortScience*. 2019;54(9S):S188

Guajardo V., Solís S., Almada R., Saski C., Gasic K., Moreno M.Á. Genome-wide SNP identification in *Prunus* rootstocks germplasm collections using Genotyping-by-Sequencing: phylogenetic analysis, distribution of SNPs and prediction of their effect on gene function. *Sci Rep.* 2020;10(1):1467. doi 10.1038/s41598-020-58271-5

Guan L., Cao K., Li Y., Guo J., Xu Q., Wang L. Detection and application of genome-wide variations in peach for association and genetic relationship analysis. *BMC Genet.* 2019;20(1):101. doi 10.1186/s12863-019-0799-8

Hamblin M.T., Warburton M.L., Buckler E.S. Empirical comparison of simple sequence repeats and single nucleotide polymorphisms in assessment of maize diversity and relatedness. *PLoS One.* 2007;2(12): e1367. doi 10.1371/journal.pone.0001367

Herrero J., Cambra M., Tabuenca M.C. Cartografía de Frutales de Hueso y Pepita. Zaragoza: Estación Experimental de Aula Dei (EEAD-CSIC), 1964

Hesse C.O. Peaches. In: Janick J., Moore J.N. (Eds) Advances in Fruit Breeding. West Lafayette, Ind.: Purdue University Press, 1975; 285-335

Hong J.H., Yi S.I., Kwon Y.S., Kim Y., Choi K.J. Genetic diversity analysis of peach [*Prunus persica* (L.) Batsch] varieties using SSR markers. *Korean J Breed Sci.* 2013;45(3):201-211. doi 10.9787/KJBS.2013.45.3.201

Howad W., Yamamoto T., Dirlewanger E., Testolin R., Cosson P., Cipriani G., Monforte A.J., Georgi L., Abbott A.G., Arus P. Mapping with a few plants: using selective mapping for microsatellite saturation of the Prunus reference map. *Genetics.* 2005;171(3):1305-1309. doi 10.1534/genetics.105.043661

Huang Z., Shen F., Chen Y., Cao K., Wang L. Preliminary identification of key genes controlling peach pollen fertility using genome-wide association study. *Plants*. 2021;10(2):242. doi 10.3390/plants10020242

Hübner S., Bercovich N., Todesco M., Mandel J.R., Odenheimer J., Ziegler E., Lee J.S., ... Kubach T., Muños S., Langlade N.B., Burke J.M., Rieseberg L.H. Sunflower pan-genome analysis shows that hybridization altered gene content and disease resistance. *Nat Plants*. 2019;5(1):54-62. doi 10.1038/s41477-018-0329-0

International Peach Genome Initiative; Verde I., Abbott A.G., Scalabrin S., Jung S., Shu S., Marroni F., … Silva H., Salamini F., Schmutz J., Morgante M., Rokhsar D.S. The high-quality draft ge-

nome of peach (*Prunus persica*) identifies unique patterns of genetic diversity, domestication and genome evolution. *Nat Genet.* 2013; 45(5):487-494. doi 10.1038/ng.2586

Jayakodi M., Padmarasu S., Haberer G., Bonthala V.S., Gundlach H., Monat C., Lux T., ... Mayer K.F.X., Spannagl M., Li C., Mascher M., Stein N. The barley pan-genome reveals the hidden legacy of mutation breeding. *Nature*. 2020;588(7837):284-289. doi 10.1038/s41586-020-2947-8

Joiret M., Mahachie John J.M., Gusareva E.S., Van Steen K. Confounding of linkage disequilibrium patterns in large scale DNA based gene-gene interaction studies. *BioData Min*. 2019;12:11. doi 10.1186/s13040-019-0199-7

Jones N., Ougham H., Thomas H. Markers and mapping: we are all geneticists now. *New Phytol*. 1997;137(1):165-177. doi 10.1046/j.1469-8137.1997.00826.x

Joobeur T., Viruel M.A., de Vicente M.C., Jáuregui B., Ballester J., Dettori M.T., Verde I., Truco M.J., Messeguer R., Batlle I., Quarta R., Dirlewanger E., Arús P. Construction of a saturated linkage map for *Prunus* using an almond × peach F2 progeny. *Theor Appl Genet.* 1998;97:1034-1041. doi 10.1007/s001220050988

Jung S., Staton M., Lee T., Blenda A., Svancara R., Abbott A., Main D. GDR (Genome Database for Rosaceae): integrated web-database for Rosaceae genomics and genetics data. *Nucleic Acids Res*. 2008; 36:D1034-D1040. doi 10.1093/nar/gkm803

Jung S., Ficklin S.P., Lee T., Cheng C.-H., Blenda A., Zheng P., Yu J., Bombarely A., Cho I., Ru S., Evans K., Peace C., Abbott A.G., Mueller L.A., Olmstead M.A., Main D. The genome database for Rosaceae (GDR): year 10 update. *Nucleic Acids Res*. 2014;42: D1237-D1244. doi 10.1093/nar/gkt1012

Khlestkina E.K. Molecular markers in genetic studies and breeding. *Russ J Genet Appl Res.* 2014;4;236-244. https://link.springer.com/article/10.1134/S2079059714030022#citeas

Kim J.S., Ku Y.S., Park S.G., Kim S.H., Park H.W., Won S.Y. Anticipated polymorphic SSRs and their application based on next generation sequencing of *Prunus persica*. *Korean J Breed Sci*. 2021;53(4): 350-360. doi 10.9787/KJBS.2021.53.4.350

Koning-Boucoiran C.F., Esselink G.D., Vukosavljev M., van't Westende W.P., Gitonga V.W., Krens F.A., Voorrips R.E., van de Weg W.E., Schulz D., Debener T., Maliepaard C., Arens P., Smulders M.J. Using RNA-Seq to assemble a rose transcriptome with more than 13,000 full-length expressed genes and to develop the WagRhSNP 68k Axiom SNP array for rose (*Rosa* L.). *Front Plant Sci*. 2015;6:249. doi 10.3389/fpls.2015.00249

Kuhn D.N., Livingstone D.S., Richards J.H., Manosalva P., Van den Berg N., Chambers A.H. Application of genomic tools to avocado (*Persea americana*) breeding: SNP discovery for genotyping and germplasm characterization. *Sci Hortic*. 2019;246:1-11. doi 10.1016/j.scienta.2018.10.011

Lambert P., Campoy J.A., Pacheco I., Mauroux J.B., Da Silva Linge C., Micheletti D., Bassi D., ... Pascal T., Troggio M., Aranzana M.J., Patocchi A., Arús P. Identifying SNP markers tightly associated with six major genes in peach [*Prunus persica* (L.) Batsch] using a high-density SNP array with an objective of marker-assisted selection (MAS). *Tree Genet Genomes.* 2016;12:121. doi 10.1007/s11295-016-1080-1

Lander E.S., Green P., Abrahamson J., Barlow A., Daly M.J., Lincoln S.E., Newburg L. Mapmaker: an interactive computer package for constructing primary genetic linkage maps of experimental and natural populations. *Genomics.* 1987;1(2):174-181. doi 10.1016/0888-7543(87)90010-3

Laucou V., Launay A., Bacilieri R., Lacombe T., Adam-Blondon A.-F., Bérard A., Chauveau A., … Maul E., Ponnaiah M., Töpfer R., Péros J.P., Boursiquot J.M. Extended diversity analysis of cultivated grapevine *Vitis vinifera* with 10K genome-wide SNPs. *PLoS One.* 2018;13(2):e0192540. doi 10.1371/journal.pone.0192540

Li X., Singh J., Qin M., Li S., Zhang X., Zhang M., Khan A., Zhang S., Wu J. Development of an integrated 200K SNP genotyping array and application for genetic mapping, genome assembly improvement and genome wide association studies in pear (*Pyrus*). *Plant Biotechnol J*. 2019;17(8):1582-1594. doi 10.1111/pbi.13085

Li X., Wang J., Su M., Zhou J., Zhang M., Du J., Zhou H., ... Fang W., Wang L., Jia H., Gao Z., Ye Z. Single nucleotide polymorphism detection for peach gummosis disease resistance by genome-wide association study. *Front Plant Sci.* 2022;12:763618. doi 10.3389/fpls.2021.763618

Li X., Wang J., Su M., Zhang M., Hu Y., Du J., Zhou H., Yang X., Zhang X., Jia H., Gao Z., Ye Z. Multiple-statistical genome-wide association analysis and genomic prediction of fruit aroma and agronomic traits in peaches. *Hortic Res.* 2023;10(7):uhad117. doi 10.1093/hr/uhad117

Li Y.H., Zhou G., Ma J., Jiang W., Jin L.G., Zhang Z., Guo Y., ... Chang R.Z., Jiang Z., Jackson S.A., Li R., Qiu L.J. *De novo* assembly of soybean wild relatives for pan-genome analysis of diversity and agronomic traits. *Nat Biotechnol.* 2014;32(10):1045-1052. doi 10.1038/nbt.2979

Li Y., Cao K.E., Zhu G., Fang W., Chen C., Wang X., Zhao P., Guo J., Ding T., Guan L., Zhang Q., Guo W., Fei Z., Wang L. Genomic analyses of an extensive collection of wild and cultivated accessions provide new insights into peach breeding history. *Genome Biol.* 2019;20(1):36. doi 10.1186/s13059-019-1648-9

Lirong W., Yong L., Gengrui Z., Weichao F., Changwen C., Ke C., Xinwei W. Peach genomics and breeding programs at Zhengzhou Fruit Research Institute, CAAS. *Acta Hortic*. 2020;1282:1-6. doi 10.17660/ActaHortic.2020.1282.1

Liu H., Cao K., Zhu G., Fang W., Chen C., Wang X., Wang L. Genome-wide association analysis of red flesh character based on resequencing approach in peach. *J Am Soc Hortic Sci.* 2019;144(3):209-216. doi 10.21273/JASHS04622-18

Liu J., Bao Y., Zhong Y., Wang Q., Liu H. Genome-wide association study and transcriptome of olecranon-type traits in peach (*Prunus persica* L.) germplasm. *BMC Genomics.* 2021;22(1):702. doi 10.1186/s12864-021-08017-y

Liu Y., Du H., Li P., Shen Y., Peng H., Liu S., Zhou G.A., … Wang Z., Zhu B., Han B., Liang C., Tian Z. Pan-genome of wild and cultivated soybeans. *Cell.* 2020;182(1):162-176. doi 10.1016/j.cell.2020.05.023

Mardis E.R. Next-generation DNA sequencing methods. *Annu Rev Genomics Hum Genet.* 2008;9(1):387-402. doi 10.1146/annurev.genom.9.081307

Mariette S., Tavaud M., Arunyawat U., Capdeville G., Millan M., Salin F. Population structure and genetic bottleneck in sweet cherry estimated with SSRs and the gametophytic self-incompatibility locus. *BMC Genet.* 2010;11:77. doi 10.1186/1471-2156-11-77

Marrano A., Martínez-García P.J., Bianco L., Sideli G.M., Di Pierro E.A., Leslie C.A., Stevens K.A., Crepeau M.W., Troggio M., Langley C.H., Neale D.B. A new genomic tool for walnut (*Juglans regia* L.): development and validation of the high-density AxiomTM *J. regia* 700K SNP genotyping array. *Plant Biotechnol J.* 2019; 17(6):1027-1036. doi 10.1111/pbi.13034

Mas-Gómez J., Cantín C.M., Moreno M.Á., Prudencio Á.S., Gómez-Abajo M., Bianco L., Troggio M., Martínez-Gómez P., Rubio M., Martínez-García P.J. Exploring genome-wide diversity in the national peach (*Prunus persica*) germplasm collection at CITA (Zaragoza, Spain). *Agronomy*. 2021;11(3):481. doi 10.3390/agronomy11030481

Mas-Gómez J., Cantín C.M., Moreno M.Á., Martínez-García P.J. Genetic diversity and genome-wide association study of morphological and quality traits in peach using two Spanish peach germplasm collections. *Front Plant Sci.* 2022;13:854770. doi 10.3389/fpls.2022.854770

Meng G., Zhu G., Fang W., Chen C., Wang X., Wang L., Cao K. Identification of loci for single/double flower trait by combining genome-wide association analysis and bulked segregant analysis in peach (*Prunus persica*). *Plant Breed.* 2019;138(3):360-367. doi 10.1111/pbr.12673

Micali S., Vendramin E., Dettori M.T., Verde I. Genetics and genomics of stone fruits. In: Agricultural and Food Biotechnologies of *Olea europaea* and Stone Fruits. Bentham, 2015;243-307. doi 10.2174/9781608059935115010008

Micheletti D., Dettori M.T., Micali S., Aramini V., Pacheco I., Da Silva Linge C., Foschi S., ... Rossini L., Verde I., Laurens F., Arús P., Aranzana M.J. Whole-genome analysis of diversity and SNP-major gene association in peach germplasm. *PloS One*. 2015;10(9):e0136803. doi 10.1371/journal.pone.0136803

Monet R. Peach genetics: past present and future. *Acta Hortic.* 1988;254:49-58. doi 10.17660/ActaHortic.1989.254.8

Monet R., Gibault B. Polymorphisme de l'alpha-amylase chez le pecher. Etude genetique. *Agronomie* (*France*). 1991;11(5):353-358

Monet R., Bastard Y., Gibault B. Genetic studies on the breeding of flat peaches. *Agronomie* (*France*). 1985;5(8):727-731

Monet R., Guye A., Roy M., Dachary N. Peach mendelian genetics: a short review and new results. *Agronomie*. 1996;16(5):321-329. doi 10.1051/agro:19960505

Montanari S., Bianco L., Allen B.J., Martínez-García P.J., Bassil N.V., Postman J., Knäbel M., … Langley C.H., Evans K., Dhingra A., Troggio M., Neale D.B. Development of a highly efficient Axiom™ 70 K SNP array for *Pyrus* and evaluation for high-density mapping and germplasm characterization. *BMC Genomics*. 2019;20(1):331. doi 10.1186/s12864-019-5712-3

Morozova O., Marra M.A. Aplications of next-generation sequencing technologies in functional genomics. *Genomics.* 2008;92(5):255-264. doi 10.1016/j.ygeno.2008.07.001

Nybom H., Lācis G. Recent large-scale genotyping and phenotyping of plant genetic resources of vegetatively propagated crops. *Plants.* 2021;10(2):415. doi 10.3390/plants10020415

Parfitt D.E., Arulsekar S., Ramming D.W. Identification of plum× peach hybrids by isoenzyme analysis. *HortScience*. 1985;20(2):246-248

Paterson A.H. Making genetic maps. In: Paterson A.H. Genome Mapping in Plants. Academic Press, 1996;23-39

Peace C., Bassil N., Main D., Ficklin S., Rosyara U.R., Stegmeir T., Sebolt A., Gilmore B., Lawley C., Mockler T.C., Bryant D.W., Wilhelm L., Iezzoni A. Development and evaluation of a genome-wide 6K SNP array for diploid sweet cherry and tetraploid sour cherry. *PLoS One.* 2012;7(12):e48305. doi 10.1371/journal.pone.0048305

Pflieger S., Lefebvre V., Caranta C., Blattes A., Goffinet B., Palloix A. Disease resistance gene analogs as candidates for QTLs involved in pepper-pathogen interactions. *Genome.* 1999;42(6):1100-1110

Pozzi C., Vecchietti A. Peach structural genomics. In: Folta K.M., Gardiner S.E. (Eds) Genetics and Genomics of Rosaceae. Plant Genetics and Genomics: Crops and Models. Vol. 6. New York, NY: Springer, 2009;235-257. https://link.springer.com/book/10.1007/978-0-387-77491-6

Purcell S., Neale B., Todd-Brown K., Thomas L., Ferreira M.A.R., Bender D., Maller J., Sklar P., de Bakker P.I.W., Daly M.J., Sham P.C. PLINK: A tool set for whole-genome association and population-based linkage analysis. *Am J Hum Genet*. 2007;81(3):559-575. doi 10.1086/519795

Quarta R., Cedrola C., Dettori M.T., Verde I. QTL analysis of agronomic traits in a BC1 peach population. *Acta Hortic.* 2002;592:291-297. doi 10.17660/ActaHortic.2002.592.41

Quilot B., Wu B.H., Kervella J., Génard M., Foulongne M., Moreau K. QTL analysis of quality traits in an advanced backcross between *Prunus persica* cultivars and the wild relative species *P. davidiana*. *Theor Appl Genet.* 2004;109(4):884-897. doi 10.1007/s00122-004-1703-z

Rasheed A., Hao Y., Xia X., Khan A., Xu Y., Varshney R.K., He Z. Crop breeding chips and genotyping platforms: progress, challenges, and perspectives. *Mol Plant*. 2017;10(8):1047-1064. doi 10.1016/j.molp.2017.06.008

Ru S., Main D., Evans K., Peace C. Current applications, challenges, and perspectives of marker-assisted seedling selection in Rosaceae tree fruit breeding. *Tree Genet Genomes.* 2015;11:8. doi 10.1007/s11295-015-0834-5

Salazar J.A., Ruiz D., Campoy J.A., Sánchez-Pérez R., Crisosto C.H., Martínez-García P.J., Blenda A., Jung S., Main D., Martínez-Gómez P., Rubio M. Quantitative trait loci (QTL) and Mendelian trait loci (MTL) analysis in *Prunus*: a breeding perspective and beyond. *Plant Mol Biol Rep*. 2013;32:1-18. doi 10.1007/s11105-013-0643-7

Scorza R. Gene transfer for the genetic improvement of perennial fruit and nut crops. *HortScience*. 1991;26(8):1033-1035

Scorza R., Okie W.R. Peaches (*Prunus*). *Acta Hortic.* 1991;290:177-234. doi 10.17660/ActaHortic.1991.290.5

Scorza R., Mehlenbacher S.A., Lightner G.W. Inbreeding and coancestry of freestone peach cultivars of the eastern United States and implications for peach germplasm improvement. *J Am Soc Hortic Sci.* 1985;110(4):547-552. doi 10.21273/JASHS.110.4.547

Siberchicot A., Bessy A., Gueguen L., Marais G.A. Mareymap online: a user-friendly web application and database service for estimating recombination rates using physical and genetic maps. *Genome Biol Evol.* 2017;9(10):2506-2509. doi 10.1093/gbe/evx178

Smykov A., Shoferistov E., Korzin V., Mesyats N., Saplev N. Promising directions in the selection of peach, apricot and nectarine. *E3S Web Conf.* 2021;254:01010. doi 10.1051/e3sconf/202125401010

Sosinski B., Gannavarapu M., Hager L.D., Beck L.E., King G.J., Ryder C.D., Rajapakse S., Baird W.V., Ballard R.E., Abbott A.G. Characterization of microsatellite markers in peach [*Prunus persica* (L.) Batsch]. *Theor Appl Genet.* 2000;101:421-428. doi 10.1007/s001220051499

Tan Q., Li S., Zhang Y., Chen M., Wen B., Jiang S., Chen X., Fu X., Li D., Wu H., Wang Y., Xiao W., Li L. Chromosome-level genome assemblies of five *Prunus* species and genome-wide association studies for key agronomic traits in peach. *Hortic Res.* 2021;8(1):213. doi 10.1038/s41438-021-00648-2

Tanksley S.D., Young N.D., Paterson A.H., Bonierbale M.W. RFLP mapping in plant-breeding – new tools for an old science. *Nat Biotechnol.* 1989;7:257-264. doi 10.1038/nbt0389-257

Thurow L.B., Gasic K., Bassols Raseira M.C., Bonow S., Marques Castro C. Genome-wide SNP discovery through genotyping by sequencing, population structure, and linkage disequilibrium in Brazilian peach breeding germplasm. *Tree Genet Genomes.* 2020;16:10. doi 10.1007/s11295-019-1406-x

Trifonova A.A., Boris K.V., Mesyats N.V., Tsiupka V.A., Smykov A.V., Mitrofanova I.V. Genetic diversity of peach cultivars from the collection of the Nikita Botanical Garden based on SSR markers. *Plants.* 2021;10(12):2609. doi 10.3390/plants10122609

Van Ooijen J.W. Joinmap® 4. Software for the calculation of genetic linkage maps in experimental populations. ScienceOpen, Inc., 2006

Van Ooijen J.W. MapQTL® 6. Software for the mapping of quantitative trait loci in experimental populations of diploid species. ScienceOpen, Inc., 2009

Verde I., Bassil N., Scalabrin S., Gilmore B., Lawley C.T., Gasic K., Micheletti D, ... Aranzana M.J., Arús P., Iezzoni A., Morgante M., Peace C. Development and evaluation of a 9K SNP array for peach by internationally coordinated SNP detection and validation in breeding germplasm. *PloS One.* 2012;7(4):e35668. doi 10.1371/journal.pone.0035668

Verde I., Jenkins J., Dondini L., Micali S., Pagliarani G., Vendramin E., Paris R., ... Shu S., Grimwood J., Tartarini S., Dettori M.T., Schmutz J. The Peach v2. 0 release: high-resolution linkage map-

ping and deep resequencing improve chromosome-scale assembly and contiguity. *BMC Genomics.* 2017;18(1):225. doi 10.1186/s12864-017-3606-9

Voorrips R.E. MapChart: software for the graphical presentation of linkage maps and QTLs. *J Hered*. 2002;93(1):77-78. doi 10.1093/jhered/93.1.77

Wang L., Zhu G., Fang W. Peach germplasm and breeding programs at Zhengzhou in China. *Acta Hortic.* 2001;592:177-182. doi 10.17660/ActaHortic.2002.592.25

Werner D.J., Okie W.R. A history and description of the *Prunus persica* plant introduction collection. *HortScience.* 1998;33(5):787-793. doi 10.21273/HORTSCI.33.5.787

Winter P., Kahl G. Molecular marker technologies for plant improvement. *World J Microbiol Biotechnol*. 1995;11(4):438-448. doi 10.1007/BF00364619

Yamamoto T., Mochida K., Hayashi T. Shanhai Suimitsuto, one of the origins of Japanese peach cultivars. *J Japan Soc Hortic Sci*. 2003; 72(2):116-121

Yu J.M., Zhang Z.W., Zhu C.S., Tabanao D.A., Pressoir G., Tuinstra M.R., Kresovich S., Todhunter R.J., Buckler E.S. Simulation appraisal of the adequacy of number of background markers for relationship estimation in association mapping. *Plant Genome.* 2009; 2(1):63-77. doi 10.3835/plantgenome2008.09.0009

Yu Y., Fu J., Xu Y., Zhang J., Ren F., Zhao H., Tian S., … Wang G., Ma R., Jiang Q., Wei J., Xie H. Genome re-sequencing reveals the evolutionary history of peach fruit edibility. *Nat Commun.* 2018; 9(1):5404. doi 10.1038/s41467-018-07744-3

Zhao Q., Feng Q., Lu H., Li Y., Wang A., Tian Q., Zhan Q, ... Xu Q., Wang Z.X., Wei X., Han B., Huang X. Pan-genome analysis highlights the extent of genomic variation in cultivated and wild rice. *Nat Genet.* 2018;50(2):278-284. doi 10.1038/s41588-018-0041-z

Zurn J.D., Nyberg A., Montanari S., Postman J., Neale D., Bassil N. A new SSR fingerprinting set and its comparison to existing SSR- and SNP-based genotyping platforms to manage *Pyrus* germplasm resources. *Tree Genet Genomes.* 2020;16:72. doi 10.1007/s11295-020-01467-7

Conflict of interest. The authors declare no conflict of interest.

Received July 25, 2024. Revised January 24, 2025. Accepted January 27, 2025.