# Application of Genomic Technologies to the Breeding of Trees

*Maria L. Badenes[1]\*, Angel Fernández i Martí[2,3], Gabino Ríos[1] and María J. Rubio-Cabetas[2]*

[1] Instituto Valenciano de Investigaciones Agrarias, Valencia, Spain, [2] Hortofruticulture Department, Agrifood Research and Technology Centre of Aragon, Zaragoza, Spain, [3] Genome Center, University of California, Davis, Davis, CA, USA

The recent introduction of next generation sequencing (NGS) technologies represents a major revolution in providing new tools for identifying the genes and/or genomic intervals controlling important traits for selection in breeding programs. In perennial fruit trees with long generation times and large sizes of adult plants, the impact of these techniques is even more important. High-throughput DNA sequencing technologies have provided complete annotated sequences in many important tree species. Most of the high-throughput genotyping platforms described are being used for studies of genetic diversity and population structure. Dissection of complex traits became possible through the availability of genome sequences along with phenotypic variation data, which allow to elucidate the causative genetic differences that give rise to observed phenotypic variation. Association mapping facilitates the association between genetic markers and phenotype in unstructured and complex populations, identifying molecular markers for assisted selection and breeding. Also, genomic data provide *in silico* identification and characterization of genes and gene families related to important traits, enabling new tools for molecular marker assisted selection in tree breeding. Deep sequencing of transcriptomes is also a powerful tool for the analysis of precise expression levels of each gene in a sample. It consists in quantifying short cDNA reads, obtained by NGS technologies, in order to compare the entire transcriptomes between genotypes and environmental conditions. The miRNAs are non-coding short RNAs involved in the regulation of different physiological processes, which can be identified by high-throughput sequencing of RNA libraries obtained by reverse transcription of purified short RNAs, and by *in silico* comparison with known miRNAs from other species. All together, NGS techniques and their applications have increased the resources for plant breeding in tree species, closing the former gap of genetic tools between trees and annual species.

Keywords: next generation sequencing, genome sequences, RNA-seq, association mapping, insertional mutant populations

## INTRODUCTION

Breeding of tree species is a longer and time consuming process compared to annual species. The long generation time impairs the development of crosses until the species overcome the juvenile period, being between 3 and 10 years depending on the species. In addition, the large size of trees makes it difficult to obtain large numbers of individuals for segregation families, and

precludes the management of plants under greenhouse-controlled conditions. Some otherwise useful techniques aimed at identifying gene variants such as 'tilling' and 'ecotilling' can hardly be applied to trees. Consequently, breeding by molecular marker assisted selection ("molecular breeding") has been limited to a reduced number of traits, mainly related to disease resistance and traits with oligogenic control (Gessler and Pertot, 2012; Zuriaga et al., 2013). The identification and functional analysis of gene orthologs from model plants presents a bottleneck for efficient transformation and regeneration protocols in many of these species.

However, the availability of genome sequences along with next generation sequencing (NGS) techniques may provide new tools to overcome most of the problems faced by tree breeding. Dissection of complex traits in many important tree species has become possible through the availability of genome sequences obtained by high-throughput DNA sequencing technologies along with phenotypic variation data. Such techniques offer shortcuts for the discovery of genes linked to selected traits and simplify the analysis of diversity in a population. Association mapping (AM) and genome wide association studies (GWASs) facilitate the association between genetic markers and phenotype in unstructured and complex populations, allowing the identification of molecular markers for assisted selection and breeding. Genotyping by sequencing (GBS) (Elshire et al., 2011; Poland et al., 2012) procedures provide 1000s of markers in a population, allowing the identification of genome regions involved in traits of interest. Concerning forest tree species, characterized by large genomes such as loblolly pine, eucalyptus and oak, NGS approaches have been pivotal in providing new releases of genome sequences and improved assemblage of former ones. All together, the NGS techniques provide new tools for overcoming the long breeding process of trees and increase the breeding efficiency. In this paper we aim to review the main molecular techniques applied recently to the discovery of genes in tree species and how they can benefit the breeding programs.

## NEXT GENERATION SEQUENCING

Future progress in fruit tree breeding will increasingly rely on understanding the links between specific genotypes and their influence on phenotypes. In this regard, increased knowledge and availability of DNA sequences for individual cultivars will provide key information for achieving specific breeding objectives. The recent introduction of NGS technologies (Shendure and Ji, 2008; Metzker, 2010) represents a major revolution in providing new tools for marking and identifying the genes and/or genomic intervals controlling important traits for selection in breeding programs. With the ability to cost effectively produce millions of DNA sequence reads in a single run, these technologies pave the way for detailed expression analyses, mutation mapping, polymorphism discovery, studies on non-coding regions, and other uses potentially impacting the speed at which questions can be addressed and results can be accomplished in plant breeding (Mardis, 2008).

Among the different NGS platforms, Roche 454® was the first commercially successful next generation system in Margulies et al. (2005). One year later, SOLiD® was purchased by Applied Biosystems. In 2007, Solexa released the genome analyzer (GA), and the company was purchased by Illumina®. In early 2010, Illumina launched HiSeq 2000, and in the same year Ion Torrent was bought by Life Technologies®. These platforms employ methodologies described in the literature (Bentley et al., 2008; McKernan et al., 2009; Perkel, 2011).

Currently, third generation technologies are being introduced that streamline sequencing protocols. Helicos Heliscope® (Thompson and Steinmann, 2010), Complete Genomics® (Drmanac et al., 2010), Nanopore® (Greninger et al., 2015) and Pacific Biosciences SMRT® (Eid et al., 2009) have incorporated new modifications. First, PCR is not needed before sequencing, and secondly, the signal is captured in real time, which means that the signal is a fluorescent (Pacbio) or electric current (Nanopore) monitored during the enzymatic reaction of adding nucleotides to the complementary strand. Additionally, all of them process millions of sequence reads in parallel with very long reads, in some cases up to 10 kb long (English et al., 2012). These high-throughput DNA sequencing technologies require the development of new bioinformatic tools (algorithms and software), for storage, retrieval and analyses of huge amounts of genome-wide sequence data. Recently, the development of the newest Oxford™ nanopore technology has provided novel improvements in molecular sensing such as real-time data streaming, improved simplicity, efficiency and scalability of workflows as well as direct analysis of the molecule of interest. These platforms along with the new bioinformatic tools have provided complete annotated sequences in many important horticultural crops and related wild species in a relatively short time. These sequences are stored in databases such as "*Sequence Read Archive*" from NCBI[1], PLantGDB[2], Phytozome[3] and EnsemblPlants[4].

Genome sequencing before 2010 was based on Sanger technology, which needed large DNA fragments cloned in BACs (Bacterial Artificial Chromosome). Using this technique, *Arabidopsis thaliana* was the first plant to have its genome sequenced, in 2000 by the *Arabidopsis* genome initiative. Lately, important cereal species were sequenced such as rice (*Oryza sativa* L., 389 Mb) in 2002 by IRGSP, corn (*Zea mays* L., 2300 Mb) by Schnable et al. (2009). However, bigger genomes such as oat (*Hordeum vulgare* L., 5100 Mb) and wheat (*Triticum aestivum* L., 17000 Mb), with 80–90% of repetitive sequences, were improved after NGS could be used. The genome of important species such as tomato was completed using NGS after preliminary approaches employing Sanger sequencing (Tomato Genome Consortium, 2012). Similarly, many tree species had few sequences published in databases before the appearance of second generation sequencing approaches.

---

[1]http://www.ncbi.nlm.nih.gov/sra

[2]http://www.plantgdb.org

[3]http://www.phytozome.net

[4]http://plants.ensembl.org/index.html

Fruit tree genome assemblies have often involved a combination of NGS and Sanger technologies. Chagne (2015) reviewed the status of genome sequencing in temperate, tropical and nut tree species, with high quality genome assemblies developed mainly for diploid species. Polyploidy is a widespread feature in fruit tree species that represents a real challenge to obtain whole-genome assemblies. However, new molecular and bioinformatics methods have been developed in order to improve current genome assemblies. For instance, the TruSeq Synthetic Long Read Sequencing technology from Illumina (McCoy et al., 2014) has been utilized to assemble the polyploid sugarcane genome sequence (approximately 1 Gbp) using synthetic long reads (Riaño-Pachón et al., 2016). For this purpose, reads longer than 1.5 kbp and BLAST hits against sequences from Viridiplantae were used for genome assembly using the OLC approach (McCoy et al., 2014).

The black cottonwood (*Populus trichocarpa*) was the first perennial tree species with an available genome sequence, obtained using the Sanger technology (Tuskan et al., 2006). The first peach genome assembled (v1.0; Verde et al., 2013) was also obtained using this technology. After the introduction of NGS, the genome sequence of more species became available, and some perennial tree species were included. **Table 1** summarizes the genomes available both in fruit and in forest tree species. Former assembly of the peach genome has been improved and recently released v2.0 using NGS platforms[5]. In addition, partial sequences from cherry and almond have been made available at the *Rosaceae* Genome Database (Koepke et al., 2013[6]). Concerning forest tree species, characterized by large genomes such as loblolly pine, eucalyptus and oak, the genomes have been released recently using only NGS approaches. The sequences available are indicated in **Table 1**.

Another interesting tool that combines the use of genome sequence with identification of markers for genotyping and identification of causal mutations is the "*Genotyping by Sequencing*" technique (Elshire et al., 2011; Poland et al., 2012). GBS utilizes barcodes for the multiplex sequencing of genomic representations of different genotypes, allowing the genetic mapping in a segregant population and the identification of many SNPs. These SNPs, after a first filtering based on quality criteria, may be utilized for SNP genotyping using different technologies (Perkel, 2008; Edwards et al., 2014). GBS provides a rapid and low-cost tool to genotype breeding populations (including the two parents), allowing plant breeders to implement GWAS, genetic linkage analysis, molecular marker discovery, and genomic selection (GS) under a large scale of plant breeding programs. Additionally, GBS has been shown to be a valid tool for genomic diversity studies (Fu and Peterson, 2011). Recently this approach was reported for quantitative trait loci (QTL) analyses in peach and other species (Ward et al., 2013; Bielenberg et al., 2015). Several microarrays using the Infinium[R] platform for SNP genotyping are available in different Rosaceae

[5]https://www.rosaceae.org/species/prunus_persica/genome_v2.0.a1
[6]http://www.rosaceae.org

**TABLE 1 | Reference genome sequences of fruit and forest tree species.**

| Fruit species | Common names | Reference |
|---|---|---|
| *Carica papaya* | Papaya | Ming et al., 2008 |
| *Malus domestica* | Apple | Velasco et al., 2010 |
| *Theobroma cacao* | Cacao | Argout et al., 2011 |
| *Fragaria vesca* | Strawberry | Shulaev et al., 2011 |
| *Prunus mume* | Japanese apricot | Zhang et al., 2012 |
| *Musa acuminata* | Banana | D'Hont et al., 2012 |
| *Phoenix dactylifera* | Date palm | Al-Mssallem et al., 2013 |
| *Pyrus bretschneideri* | Pear | Wu et al., 2013 |
| *Prunus persica* | Peach | Verde et al., 2013 |
| *Citrus sinensis* | Sweet orange | Xu et al., 2013 |
| *Morus notabilis* | Mulberry | He et al., 2013 |
| *Coffea canephora* | Coffee | Denoeud et al., 2014 |
| *Elais guineensis* and *Elais oleifera* | South American oil palm | Singh et al., 2013 |
| *Pyrus communis* | Pear | Chagné et al., 2014 |
| *Citrus genus* | Orange | Wu G. A.et al., 2014 |

| Forest tree species | Common names | Reference |
|---|---|---|
| *Populus trichocarpa* | Balck cottonwood | Tuskan et al., 2006 |
| *Betula nana* | Dwarf birch | Wang et al., 2013 |
| *Betula pendula* | Silver birch | Birchgenome.org |
| *Hevea brasiliensis* | Rubber tree | Rahman et al., 2013 |
| *Picea glauca* | White spruce | Birol et al., 2013 |
| *Picea abies* | Norway spruce | Nystedt et al., 2013 |
| *Pinus taeda* | Loblolly pine | Neale et al., 2014 |
| *Eucalyptus grandis* | Eucalytus | Myburg et al., 2014 |
| *Quercus ruber* | Oak | Plomion et al., 2016 |

species such as peach, apple, and cherry (Chagné et al., 2012; Verde et al., 2012).

# GENOME-WIDE GENETIC DIVERSITY STUDIES

One of the main challenges in breeding is to access and use the wide genetic variation present in germplasm collections and their wild relatives. Most of the high-throughput genotyping platforms described previously are being used for studies on diversity and population structure.

Genome-wide surveys of genetic diversity are useful to elucidate the causative genetic differences that give rise to observed phenotypic variation, providing a foundation for dissecting complex traits through genome-wide association studies. The use of genome and transcriptome sequencing for SNP discovery has resulted in large SNPs collections in most crops. In diversity studies, the availability of sequences allows the identification of polymorphism rates for individual SNP markers, assisting in the selection of those SNPs with biological meaning and those that are highly polymorphic between groups and populations. The detected polymorphisms facilitate the identification of mutations related to phenotypic variation enabling a better knowledge and wider use of the diversity in breeding.

Affymetrix chips have been applied to the identification of mutations related to phenotypic variation (Hill et al., 2013). Large collections of SNPs are being validated and applied for different purposes such as map construction, map saturation, genome-wide diversity studies, and AM analyses.

In the frame of the *International Peach SNP Consortium* (IPSC), an Illumina II 9K SNP chip was specifically designed (Verde et al., 2012) and used for genotyping of variety collections of peach, cherry and almond (Romeu et al., 2014; Sánchez et al., 2014). The *RosBREED SNP Consortium* (IRSC) developed a 9K chip based on Illumina technology aimed at studying the allelic diversity of *Malus* for breeding purposes (Chagné et al., 2012).

In fruit species, the most commonly used markers for diversity studies are still simple sequence repeats (SSRs) or microsatellites. The first SSR studies in Rosaceae focused on different species from the genus *Prunus* (Dirlewanger et al., 2002; Hormaza, 2002; Aranzana et al., 2003), *Malus* (Silverberg-Dilworth et al., 2006), and *Pyrus* (Yamamoto et al., 2002). Several local germplasm collections from temperate fruits have been studied (Bouhadida et al., 2011; Pina et al., 2014). However, in many crops progress in breeding relies on increasing the genetic variability available. Many studies of diversity are focused on the evaluation of genetic introgression from wild relatives. For instance, Richards et al. (2009) studied 20 populations of *Malus sieversii* prospected in China, concluding that diversity within each population was higher than among populations, which leads to the conclusion that this species represents a valuable source of diversity for apple breeding. In *Prunus* genus, introgressions from wild relatives resulted in gain for the commercial varieties, mainly concerning pathogen resistance (Quilot et al., 2004; Esmenjaud and Dirlewanger, 2007; Marandel et al., 2009). Other studies have focused on wild relatives from almond and apricot as a source of diversity (He et al., 2007; Zeinalabedini et al., 2008; Gross et al., 2012; Fernández i Martí et al., 2014). Increase in variability via wild relatives has been explored in other genera with relevant economic impacts such as *Citrus* (Barkley et al., 2006), and *Olea* (Belaj et al., 2010; Fernández i Martí et al., 2015a). Germplasm resources from underutilized fruits have been carried out in *Eriobotrya* (Gisbert et al., 2009) and *Diospyrus* (Naval et al., 2010).

SSRs from EST libraries were proven to be useful for genotyping diversity in fruit genera such as *Citrus* (Luro et al., 2008), *Malus* (Gasic et al., 2009), and *Prunus* (Lazzari et al., 2008). Additionally, partial sequences of chloroplast DNA have been used for establishing phylogenetic relationships among species of different genera of fruit crops (Bausher et al., 2006; Bielsa et al., 2014).

Although SSRs are still useful markers, the trend is to develop and use SNPs. As genome sequences become available, the collection of Affymetrix chips provides large numbers of SNPs as unique markers that allow accurate identification of differences among genotypes affording many applications in genetic diversity studies, such as measuring levels of introgression among populations and targeting of specific genes and their evolution.

## GENOME-WIDE SEQUENCES FOR GENE MINING

Among other applications, genomic data have been used for the *in silico* identification and characterization of genes and gene families related to important traits, thus enabling new tools for molecular marker assisted selection and plant breeding. Many gene families have been identified using this approach, among which are numerous transcription factor families. As an example, in *Prunus mume* MADS-box genes encoding transcription factors involved in crucial roles in plant development, especially in flower and fruit development were identified (Xu et al., 2014). In *Malus* x *domestica* the expansin gene family included 41 genes putatively related to cell-wall-loosening processes required for cell expansion and fruit softening (Zhang S.et al., 2014). The Golden Delicious apple genome was similarly mined for lipoxygenase *LOX* genes coding for enzymes that catalyze the dioxygenation of polyunsaturated fatty acids. In this study *MdLOX1a* and *MdLOX5e* were identified as candidate genes for fruit aroma volatile production in apple, based on expression and genetic evidences (Vogt et al., 2013). Mining the draft genome of papaya served to study the structure of different gene families in this species lacking recent genome-wide duplications, to accelerate the construction of physical maps of sex chromosomes and to identify a gene controlling fruit flesh color (Ming et al., 2012). The causative allele *br*, named *broomy* for pillar appearance phenotype, located previously in the linkage group 2 of the peach genetic map has been recently identified in the same region using NGS (Dardick et al., 2013). This gene, ppa010082, is an ortholog of the rice gene *TILLER ANGLE CONTROL1 (TAC1)* (Yu et al., 2007).

## RE-SEQUENCING GENOMES

With advancements in NGS technology, whole genome re-sequencing is currently the most rapid and effective method to unravel, at the genomic level, the underlying mechanisms of species origin, development, growth, and evolution. Material from wild relatives, ancestors, landraces held in germplasm collections and modern cultivars of crop species will offer a useful gene pool to cope with existing and new breeding challenges. The main applications of genome re-sequencing include detection of genetic differences between variants, transposon fingerprinting for assessing germplasm diversity and lineages, and mapping loci associated with specific traits, such as disease resistance, fruit quality or other important agronomical traits.

To date, re-sequencing initiatives have been launched for several species including 100s of cultivars, such as the "150 Tomato Genome ReSequencing project" or "Resequencing 302 wild and cultivated accessions identifies genes related to domestication and improvement in soybean." In the latter case, researchers were able to identify almost 10,000 new SNPs and 876,799 indels. In addition, the depth of sequencing data allowed them to identify a total of 1,614 copy number variations (CNVs) and 6,388 segmental deletions comprising 15.14 and 73.6 Mb sequences, respectively (Zhou et al., 2015). On the other hand,

the 150 Tomato Genome Re-sequencing project aims to explore the genetic variation available in old breeding tomato stocks and wild tomato relatives (Aflitos et al., 2014).

Whole genome re-sequencing has been applied to most of sequenced tree genomes. In *Populus*, linkage disequilibrium studies based on genome re-sequencing suggested the feasibility of genome-wide association studies in undomesticated populations (Slavov et al., 2012). Moreover it was used to elucidate the origin of cultivated mandarin and orange trees, formed by interspecific introgressions of *Citrus maxima* into ancestral mandarin species (Wu G. A.et al., 2014). Pooled re-sequencing of 240 *Eucalyptus* genomes served to develop a 60K SNP chip suitable for different *Eucalyptus* tree species (Silva-Junior et al., 2015).

Thus, re-sequencing the genomes from multiple individuals provides an important improvement over the SNP identification procedure. The development of more comprehensive SNP and molecular marker databases will offer an extensive number of genotype markers that can be used in whole-genome linkage disequilibrium (WGLD) and candidate gene association studies (Jackson et al., 2006).

## GENOME-WIDE TOOLS FOR FINE MAPPING AND DISSECTION OF TRAITS

In mapping populations, where major QTLs linked to target traits are described, resequencing individual genomes with extreme phenotypes followed by alignment of short reads against the reference genome facilitates the identification of polymorphisms associated with such traits. Using this procedure Zhebentyayeva et al. (2014) identified two dormancy associated genes *PpeDAM5* and *PpeDAM6* as strong candidates for chilling requirement and bud dormancy release in peach. Pirona et al. (2013) used two segregating populations and the Illumina 9K SNP array to redefine the map position of the fruit maturity date (MD) locus of peach. A sequence variant in the NAC gene ppa008301m co-segregated with the MD locus, suggesting this gene as a candidate for controlling fruit ripening time in peach. Vendramin et al. (2014) used a linkage map from a F2 peach progeny to map the nectarine locus in a 635 kb interval. The subsequent inspection of the genes annotated in the peach genome sequence (Peach v1.0) led to the identification of the MYB gene *PpeMYB25* as a candidate gene for trichome formation on fruit. In addition, three independent mutations in a single gene coding for a putative carotenoid cleavage dioxygenase (*PpCCD4*) were found associated with the yellow flesh trait in peach, after analyzing 37 varieties including ancestral relatives (Falchi et al., 2013).

In apple, the *Co* gene responsible for the columnar growth habit phenotype was fine mapped in a 200-kb region of the linkage group 10 (Bai et al., 2012; Moriya et al., 2012). Evidence from high throughput genomic sequencing of the 'Wijcik' mutant columnar habit and its wild-type 'McIntosh' with standard habit pointed to a 1956 bp insertion of a mobile DNA element as the likely origin of *Co* mutation (Wolters et al., 2013; Otto et al., 2014). The insertion was located in an intergenic region, but strongly affected the expression of a 2OG-Fe(II)

oxygenase gene named *MdCo31*. The involvement of *MdCo31* in the columnar growth habit was functionally confirmed by the replication of a *Co*-like phenotype in *Arabidopsis thaliana* expressing constitutively this gene (Wolters et al., 2013).

In the Caucasian persimmon (*Diospyros lotus*), a conserved small RNA-dependent mechanism was found to determine sex in this dioecious plant, after high throughput genomic sequencing of sex-segregant pools of F1 plants (Akagi et al., 2014).

As an alternative to analyses in controlled crosses, AM in unstructured and complex populations is now being largely applied to many crops. Genetic resources consist of a large number of accessions with different histories, mutations, and recombination events and may represent a large reservoir of phenotypic and molecular diversity. The AM strategy has been proposed to identify polymorphisms involved in phenotypic variations and may be useful in identifying interesting alleles/traits for breeding purposes. This approach relies on the strength of association between genetic markers and phenotype. Thus, it detects and locates genes relative to an existing map of genetic markers (Mackay and Powell, 2007). AM has been successfully applied in mapping genes involved in several traits in different plant species (maize, sunflower, lettuce, potato, tomato, wheat, etc.), but only a few studies have been carried out in fruit tree crops, such as peach (Font i Forcada et al., 2013; Micheletti et al., 2015), apple (Cevik et al., 2010), pear (Oraguzie et al., 2010), almond (Font i Forcada et al., 2015a,b), and apricot (Mariette et al., 2016). In almond, these studies have been able to detect the genomic regions where candidate genes involved in the accumulation of important compounds for fruit quality such as tocopherol and phytosterol are located, using the peach reference genome. In apricot, the implementation of GWASs for Plum Pox virus (PPV) resistance utilizing NGS platforms to genotype a broad spectrum of the available apricot breeding germplasm verified a previously described single family QTL and specific candidate genes. Thus, demonstrating the utility of the combined approaches of both single family QTL and GWAS studies (Mariette et al., 2016).

## INSERTIONAL MUTANT POPULATIONS

Collections of insertional mutants by T-DNA and transposon tagging have become a powerful tool for functional genomics in model plant species. In *Arabidopsis thaliana*, high-throughput procedures for plant transformation, insertion-site recovery and DNA sequencing allowed the generation of libraries of indexed T-DNA insertions at the beginning of the century (Alonso et al., 2003). They have taken advantage of sequence tagging of target genes for functional gene characterization by direct and reverse genetics approaches, providing a vast amount of information in databases and the literature. The analyses of mutant populations provide direct functional links between genes and phenotypes, help to integrate *in silico* analysis of gene and protein expression, and facilitate the association studies of natural genetic polymorphism and the phenotypic analysis of adaptation to environment. However some mutagenesis approaches cannot be applied to trees because

of their outcrossing breeding systems, high heterozygosity, large body size, and long juvenile period. In contrast, gene-tagging approaches that use insertional mutagenesis to create dominant phenotypes are ideally suited for trees. The availability of poplar genome sequence allowed the first project of insertional mutations in a perennial tree species, using T-DNA activation tagging as a source of dominant and semidominant mutations (Busov et al., 2003, 2005). A dwarf poplar mutant was identified among a collection of 627 independent activation-tagged lines which contained a hyperactivated GA2-oxidase gene, leading to altered specific gibberellin contents (Busov et al., 2003). Tree stature is considered an interesting trait for breeders with potential impact on wood and fruit production and management costs. In perennial trees these studies are preliminary; however, in species with efficient transformation and regeneration protocols they may be combined with NGS technologies to constitute a valuable resource for functional genomics and breeding.

## DEEP SEQUENCING OF TRANSCRIPTOMES: RNA-SEQ

RNA sequencing (RNA-seq) is a powerful tool for the analysis of transcriptomes due to the precise measurement of the expression level of each gene in a sample by quantifying short cDNA reads obtained by NGS technologies, allowing the comparison of entire transcriptomes between genotypes and conditions (Wang et al., 2009; Martin and Wang, 2011). RNA reads may be mapped on a reference genome or simply aligned and assembled when a sequenced genome is not available, which makes this approach suitable for any plant species.

In fruit trees, the study of transcriptomes has been widely used for characterizing pathogen infection and resistance pathways. Rodamilans et al. (2014) analyzed the hypersensitive response of plum to PPV infection. Rubio et al. (2015) evaluated the transcriptomic changes on peach leaves infected by PPV aimed at studying the plant defense response. Socquet-Juglard et al. (2013) identified genes involved in the defense response to *Xanthomonas arboricola* in peach. In these and many other studies, genes involved in cell wall metabolism, photosynthesis, hormone signaling, and plant defense mechanisms were found differentially regulated following pathogen infection.

RNA-seq was also applied to physiological issues and developmental transitions specific to perennial plants, such as bud dormancy release. Zhu et al. (2015) performed the *de novo* transcriptome assembly and expression profiling of flower buds from Chinese cherry (*Prunus pseudocerasus*). Results identified dormancy-associated MADS-box, *AGAMOUS*-like, and *APETALA3*-like as genes related to bud dormancy release. Jia et al. (2014) analyzed by RNA-seq the continuous flowering trait in longan (*Dimocarpus longan*), an evergreen subtropical species. Results identified candidate genes providing new insight into the molecular process of regulating flowering time in woody plants. Genes homologous to *SHORT VEGETATIVE PHASE* (*SVP*),

*GIGANTEA* (*GI*), *F-BOX 1* (*FKF1*), and *EARLY FLOWERING 4* (*ELF4*) were found differentially expressed in cultivars flowering throughout the year and cultivars flowering only once in the season.

Other tree-specific processes approached by transcriptomic studies include wood quality traits, secondary cell wall formation and lignin synthesis, performed mostly in *Populus* (Chen et al., 2015), *Eucalyptus* (Thavamanikumar et al., 2014), *Acacia* (Wong et al., 2011), and *Tectona* (Galeano et al., 2015) genera.

Among many other studies employing RNA-seq in perennial tree species, those investigating the transcriptomic response to water deficit (Villar et al., 2011; Cossu et al., 2014; Dong et al., 2014; Behringer et al., 2015), oxidative stress (Lu et al., 2014), nutrient deficiency (Fan et al., 2014) and heat shock (Chen et al., 2014) are especially abundant and relevant under a perspective of climate change.

## CRISPR/Cas9 TECHNOLOGY

Clustered regularly interspaced short palindromic repeats (CRISPR)/CRISPR-associated (Cas) protein 9 (CRISPR/Cas9) has emerged as an alternative to classical plant breeding and transgenic methods to improve crop plants. Until 2013, the dominant genome editing tools were zinc finger nucleases (ZFNs; Kim et al., 1996) and transcription activator-like effector nucleases (TALENs, Christian et al., 2010). Recently, CRISPR/Cas9 offered an alternative to ZFNs and TALENs for genome editing. CRISPR/Cas9 depends on small RNA for sequence-specific cleavage. Because only programmable RNA is required to generate sequence specificity, CRISPR/Cas9 is easily applicable and has developed very fast over the past year. Thus, this technique is expected to revolutionize the field of genomics. Based on the acumen of CRISPR/Cas system, it can be utilized for introducing desired changes like targeted single and multiple gene knock-outs of detrimental genes in plants (Brouns et al., 2008) and introducing SNPs into a gene of interest (Voytas, 2013) for improvement of economic traits. Until now, it has been successfully applied in many plant species such as *Arabidopsis thaliana* (Mao et al., 2013), *Oryza sativa* (Shan et al., 2013), *Sorghum bicolor* (Jiang et al., 2013), *Solanum lycopersicum* (Ito et al., 2015), *Citrus sinensis* (Jia and Wang, 2014), *Zea mays* (Svitashev et al., 2015), or *Populus tomentosa* (Fan et al., 2015).

Clustered regularly interspaced short palindromic repeats (CRISPR)/CRISPR-associated technology uses an RNA to target the Cas9 nuclease to a particular site in the genome. Cas9 unwinds the DNA at the target site and restricts both strands of the DNA, creating a double-stranded break. Lately, the break can be repaired by non-homologous end joining (NHEJ), creating a high likelihood of introducing INDELS leading to frame shift mutations in a gene or locus of interest (Nekrasov et al., 2013; Ran et al., 2013). NGS may provide the sequence data required for the selection of gene targets, and to avoid single guide RNAs (sgRNA) with a high likelihood of off-targeting. The potential off-target sequences have to be identified by searching the particular plant genome database via BLASTN against the sgRNA target site within the

genes of study. Also several software tools[78] are available to predict potential off-target mutagenesis caused by Cas9/sgRNA effects for gene-editing tools. Only two studies in perennial species have applied the CRISPR/Cas9 technology, one in sweet orange (Jia and Wang, 2014), and other in *Populus tomentosa* (Fan et al., 2015); however, to date there are no studies linking NGS and CRISPR/Cas9 or showing if genetic changes induced by Cas9/sgRNA are inherited to subsequent generations in those species. Thus, the main applications of this technique are targeting food quality traits, susceptibility to pathogens or metabolomics engineering in diverting important regulatory pathways from valuable end-products (Rani et al., 2016).

## EPIGENOMES VS. TRANSCRIPTOMES

Modulation of developmental processes by epigenetic events involving gene expression regulation by DNA and chromatin modifications persistent through repeated cycles of cell division is an emerging issue in plant science. Recently, genome wide studies of methylated DNA immunoprecipitation followed by high-throughput sequencing (MeDIP-seq) were used to characterize the DNA methylome of *Populus trichocarpa* during *in vitro* culture and plant transformation methods, in micropropagated explants, calli, and regenerated plants (Vining et al., 2013). MeDIP-seq was also employed to find tissue-specific features in *Populus* (Vining et al., 2012; Lafon-Placette et al., 2013) and developmentally regulated regions in leaves of *Eucalyptus globulus* (Hasbún et al., 2016). On the other hand, numerous chromatin modifications involving histone acetylation, methylation, phosphorylation, and ubiquitination among others can be assayed by genome wide sequencing of immunoprecipitated chromatin (ChIP-seq). In perennial trees, ChIP-seq has been used to identify dormancy-dependent H3K27 trimethylated regions in peach (de la Fuente et al., 2015) and to map trimethylated H3K4 in developing xylem of *Eucalyptus grandis* (Hussey et al., 2015).

Most interestingly, these works allowed a double assessment of gene regulation by both epigenomic (MeDIP-seq, ChIP-seq) and transcriptomic methods (RNA-seq, microarray hybridization, *in silico* data), offering a multidimensional view of the regulatory details associated with a particular process, and confirming the conservation of chromatin effects on the regulation of gene expression in the species under study.

ChIP-seq has been also utilized for the identification of binding targets of a particular transcription factor involved in cambium development and differentiation in *Populus* (Liu et al., 2015), which opens an additional way for the molecular dissection of regulatory pathways in perennial trees.

On the other hand, a previous study carried out by Fernández i Martí et al. (2015b) showed that methylation DNA may be the

---

[7]http://chopchop.cbu.uib.no
[8]http://cbi.hzau.edu.cn/crispr/

reason and origin of the self-compatible trait in almond. Further studies are being undertaken in other fruit tree species aimed at confirming this hypothesis.

## GENOME-WIDE IDENTIFICATION OF microRNAs

MicroRNAs (*miRNAs*) are non-coding short RNAs involved in the regulation of different physiological processes through modification of the stability of complementary target transcripts. They can be identified by high-throughput sequencing of RNA libraries obtained by reverse transcription of purified short RNAs, and by *in silico* comparison with known miRNAs from other species. Numerous miRNA searches have been performed in perennial trees such as olive, persimmon, citrus, grapevine and peach, among others. In olive, a comprehensive study described miRNAs related to alternate bearing and fruit development. The differential accumulation of miRNAs under different developmental phases and tissues indicated that control of nutrition and hormone, together with flowering processes had a noteworthy impact on alternate bearing in olive (Yanik et al., 2013). The miRNA profiling at different stages of fruit development in pear allowed the identification of factors involved in fruit development and fruit quality through the regulation of lignin synthesis, sugar and acid metabolism and hormone signaling pathways (Wu J.et al., 2014). In persimmon, the identification of miRNAs involved in regulation of proanthocyanidin biosynthesis added valuable information on the mechanisms of natural astringency removal (Luo et al., 2015). The comparative profiling of miRNAs between red-flesh mutant and wild type sweet orange contributed to the identification of miRNA-mediated molecular processes regulating lycopene accumulation in sweet orange (Xu et al., 2010). In addition, novel cold responsive (Zhang X.-N.et al., 2014) and early flowering (Sun et al., 2012) microRNAs were identified in *Poncirus trifoliata*. Deep sequencing has been also used to identify miRNAs in grapevine (Pantaleo et al., 2010) and apple (Xia et al., 2012). In peach, the analysis of miRNA accumulation served to identify miRNA related to drought and chilling stresses (Barakat et al., 2012; Eldem et al., 2012).

In apple the columnar apple phenotype is connected to gene regulation by miRNAs, since transgenic apple expressing the *MdDRB1* (*Double Stranded RNA Binding Protein*) gene, associated with the biogenesis of miRNA, produces a phenotype similar to the columnar apple type of 'Wijcik' mutant (You et al., 2014).

## IMPACT OF NGS ON BREEDING

All the techniques described are providing new tools that overcome most of the problems faced in tree breeding. There are many examples of how these new tools benefit tree breeding. To describe the current and future impact in specific breeding programs we reviewed the impact in peach. We selected this species because among the temperate fruit crops, peach is the

third most important in the world after apples and pears. The breeding activity in this species is very active; more than 100 new cultivars are released per year (Strada and Fideghelli, 2003). Peach has a small genome and relatively short generation time, around 3 years, becoming the model for fruit tree species. This species summarizes both: high breeding and molecular activity. In this section we review some examples in which the NGS techniques have been applied in peach breeding and how in recent years they are generating results that will change the procedures of selection in the current breeding programs.

The peach genome sequence was available in Verde et al. (2013). The availability of a complete peach genome assembly combined with data from a long history of breeding, including several peach maps from families segregating for important traits and sets of phenotypic data, provide new opportunities for use of genomic and phenotypic data; all together will increase the power of the classical breeding.

Fruit quality traits showing quantitative inheritance are among the most important traits used in peach breeding. In the past, they have been difficult for phenotypic screening and needed large progenies. Martínez-García et al. (2013) built a high-density SNP map, based on whole-genome sequencing. They detected significant QTLs and candidate genes for quality traits related to chilling injury. Fruit ripening or MD is an important trait for lengthening the peach season. Fine mapping and bioinformatic analysis of the genome sequence associated with the MD locus allowed the identification of a sequence variant co-segregating with the MD trait. This variant provides potential marker-assisted breeding of new cultivars differing in MD. Another important fruit quality trait is resistance to cold storage, slow ripening (SR) is an interesting trait for selecting this resistance. Whole genome analysis allowed the identification of a deletion co-localizing with the SR trait (Nuñez-Lillo et al., 2015). Fruit acidity is a major determinant of fruit quality in peach. Wang et al. (2016) used genome-wide association studies by re-sequencing 129 varieties to identify a SNP linked to the acidity trait. Genotyping of a collection of 436 varieties served to verify this marker, thus enabling marker assisted selection of non-acid genotypes. The nectarine trait is due to a mutation resulting in a glabrous phenotype. Analysis of genomic re-sequencing data from peach/nectarine accessions pointed to an insertion of a LTR retroelement as the likely cause of the nectarine phenotype (Vendramin et al., 2014). Red color of flesh is due to high anthocyanin content leading to antioxidant properties that are attractive to consumers, and consequently has become an interesting trait in many breeding programs. Mapping projects combined with whole genome sequencing identified three genes of the dihydroflavonol-4-reductase family as good candidates for the control of this trait (Shen et al., 2013).

Similarly to fruit quality, high-throughput sequencing helped to dissect complex traits and to identify candidate genes for other traits such as adaptability to chilling (Bielenberg et al., 2015), and resistance to brown rot (Martínez-García et al., 2013), bacterial spot (Yang et al., 2013), and PPV (Zuriaga et al., 2013).

In all cited cases, NGS played an important role in the identification of gene candidates and allelic variants responsible for complex traits, allowing more efficient screening procedures of such traits by molecular markers developed from those allele variants and/or genes. The information emerging is changing breeding procedures very rapidly.

## CONCLUSION

The profusion of recent NGS-based studies performed in trees, briefly outlined in this review, firstly indicates that the old promise of genomics to make gene approaches affordable to many non-model species is already a matter of fact, driven by technological and important economic improvements. In fact, DNA sequencing cost based on calculations performed by the National Human Genome Research Institute has decreased from about $5,000 per raw megabase in 2001 to $0.015 in 2015[9]. While the task to identify genes related to a given biological process in any plant species has become easier and cheaper, publication requirements are increasingly exigent. For processes specific to tree species, a simple list of identified transcripts was often considered as a sufficient contribution to the basic knowledge of its regulation; however, currently, additional experimental evidences are expected to support genomic data. Thus, advanced resources are required in order to go one step further, or to tackle issues with wider interest. Such methodologies should optimally offer functional information of genes identified by means of biochemical analysis, genetic manipulation of gene expression and gene disruption or editing using the CRISPR/Cas9 or related systems. A sustained advance of tree science relies eventually on the use of these and other forthcoming methodologies capable of bridging the gap with standard model species.

## AUTHOR CONTRIBUTIONS

MB and MR-C revised the sections of NGS, Genome wide association, Genome-wide tools for fine mapping and dissection of traits. GR revised Insertional mutant populations, Deep sequencing of transcriptomes: RNA-seq. AF revised Genome-wide identification of microRNAs. All authors revised the final version of the MS and agree with the content.

## FUNDING

---

[9]https://www.genome.gov/sequencingcostsdata/

## ACKNOWLEDGMENTS

## REFERENCES

Aflitos, S., Schijlen, E., De Jong, H., De Ridder, D., Smit, S., Finkers, R., et al. (2014). Exploring genetic variation in the tomato (*Solanum* Section *Lycopersicon*) clade by whole-genome sequencing. *Plant J.* 80, 136–148. doi: 10.1111/tpj.12616

Akagi, T., Henry, I. M., Tao, R., and Comai, L. (2014). A Y-chromosome-encoded small RNA acts as a sex determinant in persimmons. *Science* 346, 646–650. doi: 10.1126/science.1257225

Al-Mssallem, I. S., Hu, S., Zhang, X., Lin, Q., Liu, W., Tan, J., et al. (2013). Genome sequence of the date palm *Phoenix dactylifera* L. *Nat. Commun.* 4, 2274. doi: 10.1038/ncomms3274

Alonso, J. M., Stepanova, A. N., Leisse, T. J., Kim, C. J., Chen, H., Shinn, P., et al. (2003). Genome-wide insertional mutagenesis of *Arabidopsis thaliana*. *Science* 301, 653–657. doi: 10.1126/science.1086391

Aranzana, M. J., Pineda, A., Cosson, P., Dirlewanger, E., Ascasibar, J., Cipriani, G., et al. (2003). A set of simple-sequence repeat (SSR) markers covering the *Prunus* genome. *Theor. Appl. Genet.* 106, 819–825. doi: 10.1007/s00122-002-1094-y

Argout, X., Salse, J., Aury, J. M., Guiltinan, M. J., Droc, G., Gouzy, J., et al. (2011). The genome of *Theobroma cacao*. *Nat. Genet.* 43, 101–108. doi: 10.1038/ng.736

Bai, T., Zhu, Y., Fernandez-Fernandez, F., Keulemans, J., Brown, S., and Xu, K. (2012). Fine genetic mapping of the Co locus controlling columnar growth habit in apple. *Mol. Genet. Genomics* 287, 437–450. doi: 10.1007/s00438-012-0689-5

Barakat, A., Sriram, A., Park, J., Zhebentyayeva, T., Main, D., and Abbott, A. (2012). Genome wide identification of chilling responsive microRNAs in *Prunus persica*. *BMC Genomics* 13:481. doi: 10.1186/1471-2164-13-481

Barkley, N. A., Roose, M. L., Krueger, R. R., and Federici, C. T. (2006). Assessing genetic diversity and population structure in a citrus germplasm collection utilizing simple sequence repeat markers (SSRs). *Theor. Appl. Genet.* 112, 1519–1531. doi: 10.1007/s00122-006-0255-9

Bausher, M. G., Singh, N. D., Lee, S. B., Jansen, R. K., and Daniell, H. (2006). The complete chloroplast genome sequence of *Citrus sinensis* (L.) Osbeck var 'Ridge Pineapple': organization and phylogenetic relationships to other angiosperms. *BMC Plant Biol.* 6:21. doi: 10.1186/1471-2229-6-21

Behringer, D., Zimmermann, H., Ziegenhagen, B., and Liepelt, S. (2015). Differential gene expression reveals candidate genes for drought stress response in *Abies alba* (pinaceae). *PLoS ONE* 10:e0124564. doi: 10.1371/journal.pone.0124564

Belaj, A., Muñoz-Diez, C., Baldoni, L., Satovic, Z., and Barranco, D. (2010). Genetic diversity and relationships of wild and cultivated olives at regional level in Spain. *Sci. Hortic.* 124, 323–330. doi: 10.1016/j.scienta.2010.01.010

Bentley, D. R., Balasubramanian, S., Swerdlow, H. P., Smith, G. P., Milton, J., Brown, C. G., et al. (2008). Accurate whole human genome sequencing using reversible terminator chemistry. *Nature* 456, 53–59. doi: 10.1038/nature07517

Bielenberg, D. G., Rauh, B., Fan, S., Gasic, K., Abbott, A. G., Reighard, G. L., et al. (2015). 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* 10:e0139406. doi: 10.1371/journal.pone.0139406

Bielsa, B., Jiwan, D., Fernandez i Marti, A., Dhingra, A., and Rubio-Cabetas, M. J. (2014). Detection of SNP and validation of a SFP InDel (deletion) in inverted repeat region of the *Prunus* species chloroplast genome. *Sci. Hortic.* 168, 108–112. doi: 10.1016/j.scienta.2014.01.028

Birol, I., Raymond, A., Jackman, S. D., Pleasance, S., Coope, R., Taylor, G. A., et al. (2013). Assembling the 20 Gb white spruce (*Picea glauca*) genome from whole-genome shotgun sequencing data. *Bioinformatics* 29, 1492–1497. doi: 10.1093/bioinformatics/btt178

Bouhadida, M., Moreno, M. A., Gonzalo, M. J., Alonso, J. M., and Gogorcena, Y. (2011). Genetic variability of introduced and local Spanish peach cultivars determined by SSRs markers. *Tree Genet. Genomes* 7, 257–270. doi: 10.1007/s11295-010-0329-3

Brouns, S. J., Jore, M. M., Lundgren, M., Westra, E. R., Slijkhuis, R. J., Snijders, A. P., et al. (2008). Small CRISPR RNAs guide antiviral defense in prokaryotes. *Science* 321, 960–964. doi: 10.1126/science.1159689

Busov, V. B., Brunner, A. M., Meilan, R., Filichkin, S., Ganio, L., Gandhi, S., et al. (2005). Genetic transformation: a powerful tool for dissection of adaptive traits in trees. *New Phytol.* 167, 9–18. doi: 10.1111/j.1469-8137.2005.01412.x

Busov, V. B., Meilan, R., Pearce, D. W., Ma, C., Rood, S. B., and Strauss, S. H. (2003). Activation tagging of a dominant gibberellin catabolism gene (GA 2-oxidase) from poplar that regulates tree stature. *Plant Physiol.* 132, 1283–1291. doi: 10.1104/pp.103.020354

Cevik, V., Ryder, C. D., Popovich, A., Manning, K., King, G. J., and Seymour, G. B. (2010). A FRUITFULL-like gene is associated with genetic variation for fruit flesh firmness in apple (*Malus domestica* Borkh.). *Tree Genet. Genomes* 6, 271–279. doi: 10.1007/s11295-009-0247-4

Chagne, D. (2015). Whole genome sequencing of fruit tree species. *Adv. Bot. Res.* 74, 1–37. doi: 10.1016/bs.abr.2015.04.004

Chagné, D., Crowhurst, R. N., Pindo, M., Thrimawithana, A., Deng, C., Ireland, H., et al. (2014). The draft genome sequence of European pear (*Pyrus communis* L. 'Bartlett'). *PLoS ONE* 9:e92644. doi: 10.1371/journal.pone.0092644

Chagné, D., Crowhurst, R. N., Troggio, M., Davey, M. W., Gilmore, B., Lawley, C., et al. (2012). Genome-wide SNP detection, validation, and development of an 8K SNP array for apple. *PLoS ONE* 7:e31745. doi: 10.1371/journal.pone.0031745

Chen, J., Chen, B., and Zhang, D. (2015). Transcript profiling of *Populus tomentosa* genes in normal, tension, and opposite wood by RNA-seq. *BMC Genomics* 16:164. doi: 10.1186/s12864-015-1390-y

Chen, J., Yin, W., and Xia, X. (2014). Transcriptome profiles of *Populus euphratica* upon heat shock stress. *Curr. Genomics* 15, 326–340. doi: 10.2174/1389202915054110610 1835

Christian, M., Cermak, T., Doyle, E. L., Schmidt, C., Zhang, F., Hummel, A., et al. (2010). Targeting DNA double-strand breaks with TAL effector nucleases. *Genetics* 186, 757–761. doi: 10.1534/genetics.110.120717

Cossu, R. M., Giordani, T., Cavallini, A., and Natali, L. (2014). High-throughput analysis of transcriptome variation during water deficit in a poplar hybrid: a general overview. *Tree Genet. Genomes* 10, 53–66. doi: 10.1007/s11295-013-0661-5

Dardick, C., Callahan, A., Horn, R., Ruiz, K. B., Zhebentyayeva, T., Hollender, C., et al. (2013). PpeTAC1 promotes the horizontal growth of branches in peach trees and is a member of a functionally conserved gene family found in diverse plant species. *Plant J.* 75, 618–630. doi: 10.1111/tpj.12234

de la Fuente, L., Conesa, A., Lloret, A., Badenes, M. L., and Ríos, G. (2015). Genome-wide changes in histone H3 lysine 27 trimethylation associated with bud dormancy release in peach. *Tree Genet. Genomes* 11, 45. doi: 10.1007/s11295-015-0869-7

Denoeud, F., Carretero-Paulet, L., Dereeper, A., Droc, G., Guyot, R., Pietrella, M., et al. (2014). The coffee genome provides insight into the convergent evolution of caffeine biosynthesis. *Science* 345, 1181–1184. doi: 10.1126/science.1255274

D'Hont, A., Denoeud, F., Aury, J.-M., Baurens, F.-C., Carreel, F., Garsmeur, O., et al. (2012). The banana (*Musa acuminata*) genome and the evolution of monocotyledonous plants. *Nature* 488, 213–217. doi: 10.1038/nature11241

Dirlewanger, E., Cosson, P., Tavaud, M., Aranzana, J., Poizat, C., Zanetto, A., et al. (2002). 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.* 105, 127–138. doi: 10.1007/s00122-002-0867-7

Dong, Y., Fan, G., Deng, M., Xu, E., and Zhao, Z. (2014). Genome-wide expression profiling of the transcriptomes of four *Paulownia tomentosa* accessions in response to drought. *Genomics* 104, 295–305. doi: 10.1016/j.ygeno.2014.08.008

Drmanac, R., Sparks, A. B., Callow, M. J., Halpern, A. L., Burns, N. L., Kermani, B. G., et al. (2010). Human genome sequencing using unchained base reads on self-assembling DNA nanoarrays. *Science* 327, 78–81. doi: 10.1126/science.1181498

Edwards, D., Zander, M., Dalton-Morgan, J., and Batley, J. (2014). New technologies for ultrahigh-throughput genotyping in plant taxonomy. *Methods Mol. Biol.* 1115, 151–175. doi: 10.1007/978-1-62703-767-9_8

Eid, J., Fehr, A., Gray, J., Luong, K., Lyle, J., Otto, G., et al. (2009). Real-time DNA sequencing from single polymerase molecules. *Science* 323, 133–138. doi: 10.1126/science.1162986

Eldem, V., Çelikkol Akçay, U., Ozhuner, E., Bakır, Y., Uranbey, S., and Unver, T. (2012). Genome-wide identification of miRNAs responsive to drought in peach (*Prunus persica*) by high-throughput deep sequencing. *PLoS ONE* 7:e50298. doi: 10.1371/journal.pone.0050298

Elshire, R. J., Glaubitz, J. C., Sun, Q., Poland, J. A., Kawamoto, K., Buckler, E. S., et al. (2011). A robust, simple genotyping-by-sequencing (GBS) approach for high diversity species. *PLoS ONE* 6:e19379. doi: 10.1371/journal.pone.0019379

English, A. C., Richards, S., Han, Y., Wang, M., Vee, V., Qu, J., et al. (2012). Mind the gap: upgrading genomes with Pacific Biosciences RS long-read sequencing technology. *PLoS ONE* 7:e47768. doi: 10.1371/journal.pone.0047768

Esmenjaud, D., and Dirlewanger, E. (2007). "Plum," in *Genome Mapping and Molecular Breeding. Fruits & Nuts*, ed. C. R. Kole (Heidelberg: Springer), 119–136.

Falchi, R., Vendramin, E., Zanon, L., Scalabrin, S., Cipriani, G., Verde, I., et al. (2013). Three distinct mutational mechanisms acting on a single gene underpin the origin of yellow flesh in peach. *Plant J.* 76, 175–187. doi: 10.1111/tpj.12283

Fan, D., Liu, T., Li, C., Jiao, B., Li, S., Hou, Y., et al. (2015). Efficient CRISPR/Cas9-mediated targeted mutagenesis in *Populus* in the first generation. *Sci. Rep.* 5, 12217. doi: 10.1038/srep12217

Fan, F., Cui, B., Zhang, T., Qiao, G., Ding, G., and Wen, X. (2014). The temporal transcriptomic response of *Pinus massoniana* seedlings to phosphorus deficiency. *PLoS ONE* 9:e105068. doi: 10.1371/journal.pone.0105068

Fernández i Martí, A., Font i Forcada, C., Rubio-Cabetas, M. J., Kamali, K., Wirthensohn, M., and Socias i Company, R. (2014). Molecular analyses of evolution and population structure in a worldwide almond [*Prunus dulcis* (Mill.) D.A. Webb syn. *P. amygdalus* Batsch] pool assessed by microsatellite markers. *Genet. Res. Crop Evol.* 62, 205–219. doi: 10.1007/s10722-014-0146-x

Fernández i Martí, A., Font i Forcada, C., Socias i Company, R., and Rubio-Cabetas, M. J. (2015a). Genetic relationships and population structure of local olive tree accessions from Northeastern Spain revealed by SSR markers. *Acta Physiol. Plant* 37, 1726–1735. doi: 10.1007/s11738-014-1726-2

Fernández i Martí, A., Gradziel, T. M., and Socias i Company, R. (2015b). Methylation of the Sf locus in almond is associated with S-RNase loss of function. *Plant Mol. Biol.* 86, 681–689. doi: 10.1007/s11103-014-0258-x

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

Font i Forcada, C., Oraguzie, N., Reyes-Chin-Wo, S., Espiau, M. T., Socias i Company, R., and Fernández i Martí, A. (2015a). Identification of genetic loci associated with quality traits in almond via association mapping. *PLoS ONE* 10:e0127656. doi: 10.1371/journal.pone.0127656

Font i Forcada, C., Velasco, L., Socias i Company, R., and Fernandez i Martí, A. (2015b). Association mapping for kernel phytosterol content in almond. *Front. Plant Sci.* 6:530. doi: 10.3389/fpls.2015.00530

Fu, Y. B., and Peterson, G. W. (2011). Genetic diversity analysis with 454 pyrosequencing and genomic reduction confirmed the eastern and western division in the cultivated barley gene pool. *Plant Genome* 4, 226–237. doi: 10.3835/plantgenome2011.08.0022

Galeano, E., Vasconcelos, T. S., Vidal, M., Mejia-Guerra, M. K., and Carrer, H. (2015). Large-scale transcriptional profiling of lignified tissues in *Tectona grandis*. *BMC Plant Biol.* 15:221. doi: 10.1186/s12870-015-0599-x

Gasic, K., Han, Y., Kertbundit, S., Shulaev, V., Iezzoni, A. F., Stover, E. W., et al. (2009). Characteristics and transferability of new apple EST-derived SSRs to other Rosaceae species. *Mol. Breed.* 23, 397–411. doi: 10.1007/s11032-008-9243-x

Gessler, C., and Pertot, I. (2012). Vf scab resistance of *Malus*. *Trees* 26, 95–108. doi: 10.1007/s00468-011-0618-y

Gisbert, A. D., Martínez-Calvo, J., Llácer, G., Romero, C., and Badenes, M. L. (2009). Genetic diversity evaluation of a loquat [*Eriobotrya japonica* (Thun.)

Lindl.] germplasm collection by SSR and S-allele fragments. *Euphytica* 168, 121–134. doi: 10.1007/s10681-009-9901-z

Greninger, A. L., Naccache, S. N., Federman, S., Yu, G., Mbala, P., Bres, V., et al. (2015). Rapid metagenomic identification of viral pathogens in clinical samples by real-time nanopore sequencing analysis. *Genome Med.* 7:99. doi: 10.1186/s13073-015-0220-9

Gross, B. L., Henk, A. D., Forsline, P. L., Richards, C. M., and Volk, G. M. (2012). Identification of interspecific hybrids among domesticated apple and its wild relatives. *Tree Genet. Genomes* 8, 1223–1235. doi: 10.1007/s11295-012-0509-4

Hasbún, R., Iturra, C., Bravo, S., Rebolledo-Jaramillo, B., and Valledor, L. (2016). Differential methylation of genomic regions associated with heteroblasty detected by M&M algorithm in the nonmodel species *Eucalyptus globulus* Labill. *Int. J. Genomics* 2016, 4395153. doi: 10.1155/2016/4395153

He, N., Zhang, C., Qi, X., Zhao, S., Tao, Y., Yang, G., et al. (2013). Draft genome sequence of the mulberry tree *Morus notabilis*. *Nat. Commun.* 4, 2445. doi: 10.1038/ncomms3445

He, T.-M., Chen, X.-S., Xu, Z., Gao, J.-S., Lin, P.-J., Liu, W., et al. (2007). Using SSR markers to determine the population genetic structure of wild apricot (*Prunus armeniaca* L.) in the Ily Valley of West China. *Genet. Resour. Crop Evol.* 54, 563–572. doi: 10.1007/s10722-006-0013-5

Hill, T. A., Ashrafi, H., Reyes-Chin-Wo, S., Yao, J., Stoffel, K., Truco, M.-J., et al. (2013). Characterization of *Capsicum annuum* genetic diversity and population structure based on parallel polymorphism discovery with a 30K unigene pepper genechip. *PLoS ONE* 8:e56200. doi: 10.1371/journal.pone.0056200

Hormaza, J. I. (2002). Molecular characterization and similarity relationships among apricot (*Prunus armeniaca* L.) genotypes using simple sequence repeats. *Theor. Appl. Genet.* 104, 321–328. doi: 10.1007/s001220100684

Hussey, S. G., Mizrachi, E., Groover, A., Berger, D. K., and Myburg, A. A. (2015). Genome-wide mapping of histone H3 lysine 4 trimethylation in *Eucalyptus grandis* developing xylem. *BMC Plant Biol.* 15:117. doi: 10.1186/s12870-015-0499-0

Ito, Y., Yokoi, A. N., Endo, M., Mikami, M., and Toki, S. (2015). CRISPR/ Cas9-mediated mutagenesis of the RIN locus that regulates tomato fruit ripening. *Biochem. Biophys. Res. Commun.* 467, 76–82. doi: 10.1016/j.bbrc.2015.09.117

Jackson, S., Rounsley, S., and Purugganan, M. (2006). Comparative sequencing of plant genomes: choices to make. *Plant Cell* 18, 1100–1104. doi: 10.1105/tpc.106.042192

Jia, H., and Wang, N. (2014). Targeted genome editing of sweet orange using Cas9/sgRNA. *PLoS ONE* 9:e93806. doi: 10.1371/journal.pone.0093806

Jia, T., Wei, D., Meng, S., Allan, A. C., and Zeng, L. (2014). Identification of regulatory genes implicated in continuous flowering of longan (*Dimocarpus longan* L.). *PLoS ONE* 9:e114568. doi: 10.1371/journal.pone.0114568

Jiang, W., Zhou, H., Bi, H., Fromm, M., Yang, B., and Weeks, D. P. (2013). Demonstration of CRISPR/Cas9/sgRNA-mediated tar- geted gene modification in *Arabidopsis*, tobacco, sorghum and rice. *Nucleic Acid Res.* 41, 188. doi: 10.1093/nar/gkt780

Koepke, T., Schaeffer, S., Harper, A., Dicenta, F., Edwards, M., Henry, R. J., et al. (2013). Comparative genomics analysis in Prunoideae to identify biologically relevant polymorphisms. *Plant Biotechnol. J.* 11, 883–893. doi: 10.1111/pbi.12081

Kim, Y. G., Cha, J., and Chandrasegaran, S. (1996). Hybrid restriction enzymes: zinc finger fusions to Fok I cleavage domain. *Proc. Natl. Acad. Sci. U.S.A.* 93, 1156–1160.

Lafon-Placette, C., Faivre-Rampant, P., Delaunay, A., Street, N., Brignolas, F., and Maury, S. (2013). Methylome of DNase I sensitive chromatin in *Populus trichocarpa* shoot apical meristematic cells: a simplified approach revealing characteristics of gene-body DNA methylation in open chromatin state. *New Phytol.* 197, 416–430. doi: 10.1111/nph.12026

Lazzari, B., Caprera, A., Vecchietti, A., Merelli, I., Barale, F., Milanesi, L., et al. (2008). Version VI of the ESTree db: an improved tool for peach transcriptome analysis. *BMC Bioinformatics* 9:S9. doi: 10.1186/1471-2105-9-S2-S9

Liu, L., Zinkgraf, M., Petzold, H. E., Beers, E. P., Filkov, V., and Groover, A. (2015). The *Populus* ARBORKNOX1 homeodomain transcription factor regulates woody growth through binding to evolutionarily conserved target genes of diverse function. *New Phytol.* 205, 682–694. doi: 10.1111/nph.13151

Lu, X., Kim, H., Zhong, S., Chen, H., Hu, Z., and Zhou, B. (2014). De novo transcriptome assembly for rudimentary leaves in *Litchi chinesis* Sonn. and

identification of differentially expressed genes in response to reactive oxygen species. *BMC Genomics* 15:805. doi: 10.1186/1471-2164-15-805

Luo, Y., Zhang, X., Luo, Z., Zhang, Q., and Liu, J. (2015). Identification and characterization of microRNAs from Chinese pollination constant non-astringent persimmon using high-throughput sequencing. *BMC Plant Biol.* 15:11. doi: 10.1186/s12870-014-0400-6

Luro, F. L., Costantino, G., Terol, J., Argout, X., Allario, T., Wincker, P., et al. (2008). Transferability of the EST-SSRs developed on Nules clementine (*Citrus clementina* Hort ex Tan) to other Citrus species and their effectiveness for genetic mapping. *BMC Genomics* 9:287. doi: 10.1186/1471-2164-9-287

Mackay, I., and Powell, W. (2007). Methods for linkage disequilibrium mapping in crops. *Trends Plant Sci.* 12, 57–63. doi: 10.1016/j.tplants.2006.12.001

Mao, Y., Zhang, H., Xu, N., Zhang, B., Gao, F., and Zhu, J. K. (2013). Application of the CRISPR/Cas system for efficient genome engineering in plants. *Mol. Plant* 6, 2008–2011. doi: 10.1093/mp/sst121

Marandel, G., Salava, J., Abbott, A., Candresse, T., and Decroocq, V. (2009). Quantitative trait loci metaanalysis of Plum pox virus resistance in apricot (*Prunus armeniaca* L.): new insights on the organization and the identification of genomic resistance factors. *Mol. Plant Pathol.* 10, 347–360. doi: 10.1111/j.1364-3703.2009.00535.x

Mardis, E. R. (2008). The impact of next-generation sequencing technology on genetics. *Trends Genet.* 24, 133–141. doi: 10.1016/j.tig.2007.12.007

Margulies, M., Egholm, M., Altman, W. E., Attiya, S., Bader, J. S., Bemben, L. A., et al. (2005). Genome sequencing in microfabricated high-density picolitre reactors. *Nature* 437, 376–380. doi: 10.1038/nature03959

Mariette, S., Wong Jun Tai, F., Roch, G., Barre, A., Chague, A., Decroocq, S., et al. (2016). Genome-wide association links candidate genes to resistance to Plum Pox Virus in apricot (*Prunus armeniaca*). *New Phytol.* 209, 773–784. doi: 10.1111/nph.13627

Martin, J. A., and Wang, Z. (2011). Next-generation transcriptome assembly. *Nat. Rev. Genet.* 12, 671–682. doi: 10.1038/nrg3068

Martínez-García, P. J., Parfitt, D. E., Bostock, R. M., Fresnedo-Ramírez, J., Vazquez-Lobo, A., Ogundiwin, E. A., et al. (2013). Application of genomic and quantitative genetic tools to identify candidate resistance genes for brown rot resistance in peach. *PLoS ONE* 8:e78634. doi: 10.1371/journal.pone.0078634

McCoy, R. C., Taylor, R. W., Blauwkamp, T. A., Kelley, J. L., Kertesz, M., Pushkarev, D., et al. (2014). Illumina TruSeq synthetic long-reads empower de novo assembly and resolve complex, highly-repetitive transposable elements. *PLoS ONE* 9:e106689. doi: 10.1371/journal.pone.0106689

McKernan, K. J., Peckham, H. E., Costa, G. L., McLaughlin, S. F., Fu, Y., Tsung, E. F., et al. (2009). Sequence and structural variation in a human genome uncovered by short-read, massively parallel ligation sequencing using two-base encoding. *Genome Res.* 19, 1527–1541. doi: 10.1101/gr.091868.109

Metzker, M. L. (2010). Sequencing technologies-the next generation. *Nat. Rev. Genet.* 11, 31–46. doi: 10.1038/nrg2626

Micheletti, D., Dettori, M. T., Micali, S., Aramini, V., Pacheco, I., Da Silva Linge, C., et al. (2015). Whole-genome analysis of diversity and SNP-major gene association in peach germplasm. *PLoS ONE* 10:e0136803. doi: 10.1371/journal.pone.0136803

Ming, R., Hou, S., Feng, Y., Yu, Q., Dionne-Laporte, A., Saw, J. H., et al. (2008). The draft genome of the transgenic tropical fruit tree papaya (*Carica papaya* Linnaeus). *Nature* 452, 991–996. doi: 10.1038/nature06856

Ming, R., Yu, Q., Moore, P. H., Chen, N. J., Wang, M. L., Shuler, M. A., et al. (2012). Genome of papaya, a fast growing tropical fruit tree. *Tree Genet. Genomes* 8, 445–462. doi: 10.1007/s11295-012-0490-y

Moriya, S., Okada, K., Haji, T., Yamamoto, T., and Abe, K. (2012). Fine mapping of Co, a gene controlling columnar growth habit located on apple (*Malus domestica* Borkh.) linkage group 10. *Plant Breed.* 131, 437–450. doi: 10.1111/j.1439-0523.2012.01985.x

Myburg, A., Grattapaglia, D., Tuskan, G., Hellsten, U., Hayes, R. D., Grimwood, J., et al. (2014). The genome of *Eucalyptus grandis*. *Nature* 501, 356–362. doi: 10.1038/nature13308

Naval, M. M., Zuriaga, E., Pecchioli, S., Llácer, G., Giordani, E., and Badenes, M. L. (2010). Analysis of genetic diversity among persimmon cultivars using microsatellite markers. *Tree Genet. Genomes* 6, 677–687. doi: 10.1007/s11295-010-0283-0

Neale, D. B., Wegrzyn, J. L., Stevens, K. A., Zimin, A. V., Puiu, D., Crepeau, M. W., et al. (2014). Decoding the massive genome of loblolly pine using haploid DNA

and novel assembly strategies. *Genome Biol.* 15, R59. doi: 10.1186/gb-2014-15-3-r59

Nekrasov, V., Staskawicz, B., Weigel, D., Jones, J. D., and Kamoun, S. (2013). Targeted mutagenesis in the model plant *Nicotiana benthamiana* using Cas9 RNA-guided endonuclease. *Nat. Biotechnol.* 31, 691–693. doi: 10.1038/nbt.2655

Nuñez-Lillo, G., Cifuentes-Esquivel, A., Troggio, M., Micheletti, D., Infante, R., Campos-Vargas, R., et al. (2015). Identification of candidate genes associated with mealiness and maturity date in peach [*Prunus persica* (L.) Batsch] using QTL analysis and deep sequencing. *Tree Genet. Genomes* 11, 86. doi: 10.1007/s11295-015-0911-9

Nystedt, B., Street, N. R., Wetterbom, A., Zuccolo, A., Lin, Y. C., Douglas, G., et al. (2013). The Norway spruce genome sequence and conifer genome evolution. *Nature* 497, 579–584. doi: 10.1038/nature12211

Oraguzie, N. C., Whitworth, C. J., Brewer, L., Hall, A., Volz, R. K., Bassett, H., et al. (2010). Relationships of PpACS1 and PpACS2 genotypes, internal ethylene concentration and fruit softening in European (*Pyrus communis*) and Japanese (*Pyrus pyrifolia*) pears during cold air storage. *Plant Breed.* 129, 219–226. doi: 10.1111/j.1439-0523.2009.01684.x

Otto, D., Peterson, R., Braukseipe, B., Braun, P., and Schmidt, E. R. (2014). The columnar mutation ("Co gene") of apple (*Malus domestica*) is associated with an integration of a Gypsy-like retrotransposon. *Mol. Breed.* 33, 863–880. doi: 10.1007/s11032-013-0001-3

Pantaleo, V., Szittya, G., Moxon, S., Miozzi, L., Moulton, V., Dalmay, T., et al. (2010). Identification of grapevine microRNAs and their targets using high-throughput sequencing and degradome analysis. *Plant J.* 62, 960–976. doi: 10.1111/j.0960-7412.2010.04208.x

Perkel, J. (2008). SNP genotyping: six technologies that keyed a revolution. *Nat. Methods* 5, 447–453. doi: 10.1038/nmeth0508-447

Perkel, J. (2011). Making contact with sequencing's fourth generation. *Biotechniques* 50, 93–95. doi: 10.2144/000113608

Pina, A., Urrestarazu, J., and Errea, P. (2014). Analysis of the genetic diversity of local apple cultivars from mountainous areas from Aragon (Northeastern Spain). *Sci. Hortic.* 174, 1–9. doi: 10.1016/j.scienta.2014.04.037

Pirona, R., Eduardo, I., Pacheco, I., Da Silva Linge, C., Miculan, M., Verde, I., et al. (2013). Fine mapping and identification of a candidate gene for a major locus controlling maturity date in peach. *BMC Plant Biol.* 13:166. doi: 10.1186/1471-2229-13-166

Plomion, C., Aury, J. M., Amselem, J., Alaeitabar, T., Barbe, V., Belser, V., et al. (2016). Decoding the oak genome: public release of sequence data, assembly, annotation and publication strategies. *Mol. Ecol. Resour.* 16, 254–265. doi: 10.1111/1755-0998.12425

Poland, J. A., Brown, P. J., Sorrells, M. E., and Jannink, J. L. (2012). Development of high-density genetic maps for barley and wheat using a novel two-enzyme genotyping-by-sequencing approach. *PLoS ONE* 7:e32253. doi: 10.1371/journal.pone.0032253

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

Rahman, A. Y. A., Usharraj, A. O., Misra, B. B., Thottathil, G. P., Jayasekaran, K., Feng, Y., et al. (2013). Draft genome sequence of the rubber tree *Hevea brasiliensis*. *BMC Genomics* 14:75. doi: 10.1186/1471-2164-14-75

Ran, F. A., Hsu, P. D., Lin, C. Y., Gootenberg, J. S., Konermann, S., Trevino, A. E., et al. (2013). Double nicking by RNA-guided CRISPR Cas9 for enhanced genome editing specificity. *Cell* 154, 1380–1389. doi: 10.1016/j.cell.2013.08.021

Rani, R., Yadav, P., Kalyani, M., Barbadikar, B., Baliyan, N., Malhotra, E. V., et al. (2016). CRISPR/Cas9: a promising way to exploit genetic variation in plants. *Biotechnol. Lett.* 38, 1991–2006. doi: 10.1007/s10529-016-2195-z

Riaño-Pachón, D. M., Mattiello, L., and Prado da Cruz, L. (2016). *Surveying the Complex Polyploid Sugarcane Genome Sequence Using Synthetic Long Reads*. Technical Report. Sao Paulo: Laboratório Nacional de Ciência e Tecnologia do Bioetanol.

Richards, C. M., Volk, G. M., Reilley, A. A., Henk, A. D., Lockwood, D. R., Reeves, P. A., et al. (2009). Genetic diversity and population structure in *Malus sieversii*, a wild progenitor species of domesticated apple. *Tree Genet. Genomes* 5, 339–347. doi: 10.1007/s11295-008-0190-9

Rodamilans, B., San León, D., Mühlberger, L., Candresse, T., Neumüller, M., Oliveros, J. C., et al. (2014). Transcriptomic analysis of *Prunus* domestica undergoing hypersensitive response to plum pox virus infection. *PLoS ONE* 9:e100477. doi: 10.1371/journal.pone.0100477

Romeu, J. F., Monforte, A. J., Sánchez, G., Granell, A., García-Brunton, J., Badenes, M. L., et al. (2014). Quantitative trait loci affecting reproductive phenology in peach. *BMC Plant Biol.* 14:52. doi: 10.1186/1471-2229-14-52

Rubio, M., Rodríguez-Moreno, L., Ballester, A. R., de Moura, M. C., Bonghi, C., Candresse, T., et al. (2015). Analysis of gene expression changes in peach leaves in response to Plum pox virus infection using RNA-Seq. *Mol. Plant Pathol.* 16, 164–176. doi: 10.1111/mpp.12169

Sánchez, G., Martínez, J., Romeu, J., García, J., Monforte, A. J., Badenes, M. L., et al. (2014). The peach volatilome modularity is reflected at the genetic and environmental response levels in a QTL mapping population. *BMC Plant Biol.* 14:137. doi: 10.1186/1471-2229-14-137

Schnable, P. S., Ware, D., Fulton, R. S., Stein, J. C., Wei, F., Pasternak, S., et al. (2009). The B73 maize genome: complexity, diversity, and dynamics. *Science* 326, 1112–1115. doi: 10.1126/science.1178534

Shan, Q., Wang, Y., Li, J., Zhang, Y., Chen, K., Liang, Z., et al. (2013). Targeted genome modification of crop plants using a CRISPR-Cas system. *Nat. Biotechnol.* 31, 686–688. doi: 10.1038/nbt.2650

Shen, Z., Confolent, C., Lambert, P., Poëssel, J. L., Quilot-Turion, B., Yu, M., et al. (2013). Characterization and genetic mapping of a new blood-flesh trait controlled by the single dominant locus DBF in peach. *Tree Genet. Genomes* 9, 1435–1446. doi: 10.1007/s11295-013-0649-1

Shendure, J., and Ji, H. (2008). Next-generation DNA sequencing. *Nat. Biotechnol.* 26, 1135–1145. doi: 10.1038/nbt1486

Shulaev, V., Sargent, D. J., Crowhurst, R. N., Mockler, T. C., Folkerts, O., Delcher, A. L., et al. (2011). The genome of woodland strawberry (*Fragaria vesca*). *Nat. Genet.* 43, 109–116. doi: 10.1038/ng.740

Silfverberg-Dilworth, E., Matasci, C. L., Van de Weg, W. E., Van Kaauwen, M. P. W., Walser, M., Kodde, L. P., et al. (2006). Microsatellite markers spanning the apple (*Malus x domestica* Borkh.) genome. *Tree Genet. Genomes* 2, 202–224. doi: 10.1007/s11295-006-0045-1

Silva-Junior, O. B., Faria, D. A., and Grattapaglia, D. (2015). A flexible multi-species genome-wide 60K SNP chip developed from pooled resequencing of 240 *Eucalyptus* tree genomes across 12 species. *New Phytol.* 206, 1527–1540. doi: 10.1111/nph.13322

Singh, R., Ong-Abdullah, M., Low, E.-T. L., Manaf, M. A. A., Rosli, R., Nookiah, R., et al. (2013). Oil palm genome sequence reveals divergence of interfertile species in Old and New worlds. *Nature* 500, 335–339. doi: 10.1038/nature 12309

Slavov, G. T., DiFazio, S. P., Martin, J., Schackwitz, W., Muchero, W., Rodgers-Melnick, E., et al. (2012). Genome resequencing reveals multiscale geographic structure and extensive linkage disequilibrium in the forest tree *Populus trichocarpa*. *New Phytol.* 196, 713–725. doi: 10.1111/j.1469-8137.2012. 04258.x

Socquet-Juglard, D., Kamber, T., Pothier, J. F., Christen, D., Gessler, C., Duffy, B., et al. (2013). Comparative RNA-seq analysis of early-infected peach leaves by the invasive phytopathogen *Xanthomonas arboricola* pv. pruni. *PLoS ONE* 8:e54196. doi: 10.1371/journal.pone.0054196

Strada, G. D., and Fideghelli, C. (2003). Le cultivar de drupacee introdotte del 1991 al 2001. *L'Informatore Agrario* 59, 65–70.

Sun, L. M., Ai, X. Y., Li, W. Y., Guo, W. W., Deng, X. X., Hu, C. G., et al. (2012). Identification and comparative profiling of miRNAs in an early flowering mutant of trifoliate orange and its wild type by genome-wide deep sequencing. *PLoS ONE* 7:e43760. doi: 10.1371/journal.pone.004 3760

Svitashev, S., Young, J. K., Schwartz, C., Gao, H., Falco, S. C., and Cigan, A. M. (2015). Targeted mutagenesis, precise gene editing, and site-specific gene insertion in maize using Cas9 and guide RNA. *Plant Physiol.* 169, 931–945. doi: 10.1104/pp.15.00793

Thavamanikumar, S., Southerton, S., and Thumma, B. (2014). RNA-Seq using two populations reveals genes and alleles controlling wood traits and growth in *Eucalyptus nitens*. *PLoS ONE* 9:e101104. doi: 10.1371/journal.pone.010 1104

Thompson, J. F., and Steinmann, K. E. (2010). Single molecule sequencing with a HeliScope genetic analysis system. *Curr. Protoc. Mol. Biol.* 7:7.10. doi: 10.1002/0471142727.mb0710s92

Tomato Genome Consortium (2012). The tomato genome sequence provides insights into fleshy fruit evolution. *Nature* 485, 635–641. doi: 10.1038/nature11119

Tuskan, G. A., Difazio, S., Jansson, S., Bohlmann, J., Grigoriev, I., Hellsten, U., et al. (2006). The genome of black cottonwood, *Populus trichocarpa* (Torr. & Gray). *Science* 313, 1596–1604. doi: 10.1126/science.1128691

Velasco, R., Zharkikh, A., Affourtit, J., Dhingra, A., Cestaro, A., Kalyanaraman, A., et al. (2010). The genome of the domesticated apple (*Malus × domestica* Borkh.). *Nat. Genet.* 42, 833–839. doi: 10.1038/ng.654

Vendramin, E., Pea, G., Dondini, L., Pacheco, I., Dettori, M. T., Gazza, L., et al. (2014). A unique mutation in a MYB gene cosegregates with the nectarine phenotype in peach. *PLoS ONE* 9:e90574. doi: 10.1371/journal.pone.00 90574

Verde, I., Abbott, A. G., Scalabrin, S., Jung, S., Shu, S., Marroni, F., et al. (2013). The high-quality draft genome of peach (*Prunus persica*) identifies unique patterns of genetic diversity, domestication and genome evolution. *Nat. Genet.* 45, 487–494. doi: 10.1038/ng.2586

Verde, I., Bassil, N., Scalabrin, S., Gilmore, B., Lawley, C. T., Gasic, K., et al. (2012). Development and evaluation of a 9K SNP array for peach by internationally coordinated SNP detection and validation in breeding germplasm. *PLoS ONE* 7:e35668. doi: 10.1371/journal.pone.0035668

Voytas, D. F. (2013). Plant genome engineering with sequence-specific nucleases. *Annu. Rev. Plant Biol.* 64, 327–350. doi: 10.1146/annurev-arplant-042811-105552

Villar, E., Klopp, C., Noirot, C., Novaes, E., Kirst, M., Plomion, C., et al. (2011). RNA-Seq reveals genotype-specific molecular responses to water deficit in eucalyptus. *BMC Genomics* 12:538. doi: 10.1186/1471-2164-12-538

Vining, K., Pomraning, K. R., Wilhelm, L. J., Ma, C., Pellegrini, M., Di, Y., et al. (2013). Methylome reorganization during in vitro dedifferentiation and regeneration of *Populus trichocarpa*. *BMC Plant Biol.* 13:92. doi: 10.1186/1471-2229-13-92

Vining, K. J., Pomraning, K. R., Wilhelm, L. J., Priest, H. D., Pellegrini, M., Mockler, T. C., et al. (2012). Dynamic DNA cytosine methylation in the *Populus trichocarpa* genome: tissue-level variation and relationship to gene expression. *BMC Genomics* 13:27. doi: 10.1186/1471-2164-13-27

Vogt, J., Schiller, D., Ulrich, D., Schwab, W., and Dunemann, F. (2013). Identification of lipoxygenase (LOX) genes putatively involved in fruit flavour formation in apple (*Malus × domestica*). *Tree Genet. Genomes* 9, 1493–1511. doi: 10.1007/s11295-013-0653-5

Wang, N., Thomson, M., Bodles, W. J. A., Crawford, R. M. M., Hunt, H. V., Featherstone, A. W., et al. (2013). Genome sequence of dwarf birch (*Betula nana*) and cross-species RAD markers. *Mol. Ecol.* 22, 3098–3111. doi: 10.1111/mec.12131

Wang, Q., Wang, L., Zhu, G., Cao, K., Fang, W., Chen, C., et al. (2016). DNA marker-assisted evaluation of fruit acidity in diverse peach (*Prunus persica*) germplasm. *Euphytica* 210, 413–426. doi: 10.1007/s10681-016-1709-z

Wang, Z., Gerstein, M., and Snyder, M. (2009). RNA-Seq: a revolutionary tool for transcriptomics. *Nat. Rev. Genet.* 10, 57–63. doi: 10.1038/nrg 2484

Ward, J. A., Bhangoo, J., Fernández-Fernández, F., Moore, P., Swanson, J. D., Viola, R., et al. (2013). Saturated linkage map construction in *Rubus idaeus* using genotyping by sequencing and genome-independent imputation. *BMC Genomics* 14:2. doi: 10.1186/1471-2164-14-2

Wolters, P. J., Schouten, H. J., Velasco, R., Si-Ammour, A., and Baldi, P. (2013). Evidence for regulation of columnar habit in apple by a putative 2OG-Fe(II) oxygenase. *New Phytol.* 200, 993–999. doi: 10.1111/nph. 12580

Wong, M. M., Cannon, C. H., and Wickneswari, R. (2011). Identification of lignin genes and regulatory sequences involved in secondary cell wall formation in *Acacia auriculiformis* and *Acacia mangium* via de novo transcriptome sequencing. *BMC Genomics* 12:342. doi: 10.1186/1471-2164-12-342

Wu, G. A., Prochnik, S., Jenkins, J., Salse, J., Hellsten, U., Murat, F., et al. (2014). Sequencing of diverse mandarin, pummelo and orange genomes reveals complex history of admixture during citrus domestication. *Nat. Biotechnol.* 32, 656–662. doi: 10.1038/nbt.2906

Wu, J., Wang, D., Liu, Y., Wang, L., Qiao, X., and Zhang, S. (2014). Identification of miRNAs involved in pear fruit development and quality. *BMC Genomics* 15:953. doi: 10.1186/1471-2164-15-953

Wu, J., Wang, Z., Shi, Z., Zhang, S., Ming, R., Zhu, S., et al. (2013). The genome of the pear (*Pyrus bretschneideri* Rehd.). *Genome Res.* 23, 396–408. doi: 10.1101/gr.144311.112

Xia, R., Zhu, H., An, Y. Q., Beers, E. P., and Liu, Z. (2012). Apple miRNAs and tasiRNAs with novel regulatory networks. *Genome Biol.* 13, R47. doi: 10.1186/gb-2012-13-6-r47

Xu, Q., Chen, L. L., Ruan, X., Chen, D., Zhu, A., Chen, C., et al. (2013). The draft genome of sweet orange (*Citrus sinensis*). *Nat. Genet.* 45, 59–66. doi: 10.1038/ng.2472

Xu, Q., Liu, Y., Zhu, A., Wu, X., Ye, J., Yu, K., et al. (2010). Discovery and comparative profiling of microRNAs in a sweet orange red-flesh mutant and its wild type. *BMC Genomics* 11:246. doi: 10.1186/1471-2164-11-246

Xu, Z., Zhang, Q., Sun, L., Du, D., Cheng, T., Pan, H., et al. (2014). Genome-wide identification, characterisation and expression analysis of the MADS-box gene family in *Prunus* mume. *Mol. Genet. Genomics* 289, 903–920. doi: 10.1007/s00438-014-0863-z

Yamamoto, T., Kimura, T., Shoda, M., Imai, T., Saito, T., Sawamura, Y., et al. (2002). Genetic linkage maps constructed by using an interspecific cross between Japanese and European pears. *Theor. Appl. Genet.* 106, 9–18. doi: 10.1007/s00122-002-0966-5

Yang, N., Reighard, G., Ritchie, D., Okie, W., and Gasic, K. (2013). Mapping quantitative trait loci associated with resistance to bacterial spot (*Xanthomonas arboricola* pv. pruni) in peach. *Tree Genet. Genomes* 9, 573. doi: 10.1007/s11295-012-0580

Yanik, H., Turktas, M., Dundar, E., Hernandez, P., Dorado, G., and Unver, T. (2013). Genome-wide identification of alternate bearing-associated microRNAs (miRNAs) in olive (*Olea europaea* L.). *BMC Plant Biol.* 13:10. doi: 10.1186/1471-2229-13-10

You, C.-X., Zhao, Q., Wang, X.-F., Xie, X.-B., Feng, X.-M., Zhao, L.-L., et al. (2014). A dsRNA-binding protein MdDRB1 associated with miRNA biogenesis modifies adventitious rooting and tree architecture in apple. *Plant Biotechnol. J.* 12, 183–192. doi: 10.1111/pbi.12125

Yu, B., Lin, Z., Li, H., Li, X., Li, J., Wang, Y., et al. (2007). TAC1, a major quantitative trait locus controlling tiller angle in rice. *Plant J.* 52, 891–898. doi: 10.1111/j.1365-313X.2007.03284.x

Zeinalabedini, M., Majourhat, K., Khayam-Nekoui, M., Grigorian, V., Torchi, M., Dicenta, F., et al. (2008). Comparison of the use of morphological, protein and DNA markers in the genetic characterization of Iranian wild *Prunus* species. *Sci. Hortic.* 116, 80–88. doi: 10.1016/j.scienta.2007.10.022

Zhang, O., Chen, W., Sun, L., Zhao, F., Huang, B., Yang, W., et al. (2012). The genome of *Prunus* mume. *Nat. Commun.* 3, 1318. doi: 10.1038/ncomms2290

Zhang, S., Xu, R., Gao, Z., Chen, C., Jiang, Z., and Shu, H. (2014). A genome-wide analysis of the expansin genes in *Malus × domestica*. *Mol. Genet. Genomics* 289, 225–236. doi: 10.1007/s00438-013-0796-y

Zhang, X.-N., Li, X., and Liu, J.-H. (2014). Identification of conserved and novel cold-responsive microRNAs in trifoliate orange (*Poncirus trifoliata* (L.) Raf.) using high-throughput sequencing. *Plant Mol. Biol. Rep.* 32, 328–341. doi: 10.1007/s11105-013-0649-1

Zhebentyayeva, T. N., Fan, S., Chandra, A., Bielenberg, D. G., Reighard, G. L., Okie, W. R., et al. (2014). Dissection of chilling requirement and bloom date QTLs in peach using a whole genome sequencing of sibling trees from an F2 mapping population. *Tree Genet. Genomes* 10, 35–51. doi: 10.1007/s11295-013-0660-6

Zhu, Y., Li, Y., Xin, D., Chen, W., Shao, X., Wang, Y., et al. (2015). RNA-Seq-based transcriptome analysis of dormant flower buds of Chinese cherry (*Prunus pseudocerasus*). *Gene* 555, 362–376. doi: 10.1016/j.gene.2014.11.032

Zhou, Z., Jiang, Y., Wang, Z., Gou, Z., Lyu, J., Li, W., et al. (2015). Resequencing 302 wild and cultivated accessions identifies genes related to domestication and improvement in soybean. *Nat. Biotechnol.* 33, 408–414. doi: 10.1038/nbt.3096

Zuriaga, E., Soriano, J. M., Zhebentyayeva, T., Romero, C., Dardick, C., Cañizares, J., et al. (2013). Genomic analysis reveals MATH gene(s) as candidate(s) for *Plum pox virus* (PPV) resistance in apricot (*Prunus armeniaca* L.). *Mol. Plant Pathol.* 14, 663–677. doi: 10.1111/mpp.12037