



Plant Archives

Journal homepage: <http://www.plantarchives.org>
DOI Url : <https://doi.org/10.51470/PLANTARCHIVES.2023.v23.no1.028>

DISCLOSING THE APPLICATIONS AND PROSPECTS OF WHOLE GENOME SEQUENCING IN ORNAMENTAL CROPS: A REVIEW

Sarankumar Chandran^{1*}, Valarmathi Muthu², Vaseka Subramani³, Sinthiga Susan Kirubakaran³, Vaishnavi Prabakaran³, Sarmila Subramanian³, Siva Nandhini Karthikeyan³, Soundarya Ravichandran³, Sindhuja Chokkalingam³, Sowmiya Jayakumar³, Prasanna Devi Devendiran³, Dhivya Durairaja³, Pavithra Natrajan³, Sumithra Palanivel³, Kalpana Devi Murugan³ and Nivetha Kesavan³.

¹Department of Plant Breeding and Genetics, Adhiparasakthi Horticultural College, G.B. Nagar, Kalavai, Tamil Nadu, India.

²Department of Biotechnology, Adhiparasakthi Agricultural College, G.B. Nagar, Kalavai, Tamil Nadu, India.

³Department of Horticulture, Adhiparasakthi Horticultural College, G.B. Nagar, Kalavai, Tamil Nadu, India.

Corresponding Author: saran32388@gmail.com

(Date of Receiving : 18-11-2022; Date of Acceptance : 23-01-2023)

ABSTRACT

Plants with ornamental value, processing typical value of lovely flowers, and uniformity in their structure are designated as 'ornamental plants'. China is noted as 'mother of garden' and has copious resources for ornamental plants. With the use of rigorous whole genome sequencing technology, it is possible to study the evolutionary process, variations in the genome level, gene regulation, and compelling biological systems. The advancement in genomic techniques has been very slow for a long time because of the nature of enormous genetic diversity and the high cost of sequencing. Based on the new genome sequencing technologies and reduced cost of sequencing, more than sixty-nine ornamental plants have been sequenced since the first sequenced ornamental crop '*Prunus mume*' was in 2012. In this review, we have covered whole genome sequencing and resequencing techniques of ornamental crops. Also discussed the essential findings of whole genome sequencing information and its potential applications.

Keywords : Whole Genome Sequencing, Ornamental Crops.

Introduction

Genomics refers to a branch of genetics that includes genome mapping, sequencing, and whole-genome functional analysis. The whole genome is used as the research object, to evaluate all of the genetic information in organisms' full genomes. The primary goal of genomics research is to interpret the whole-genome sequence, including genomic variations and gene regulation, through mining and expression to gain a deeper understanding of biological mechanisms, formulate more effective breeding strategies, broaden the mining breadth and depth of excellent alleles in germplasm resources, and increase operability for improving complex traits and the efficiency of breeding new varieties. Ornamental plants, an important component of agriculture and horticulture, play an important role in beautifying and improving human living conditions, fostering human sentiment, and driving structural changes in the agricultural business. *Arabidopsis thaliana* genome was the first to be published in 2000. With the advent of next-generation and high-throughput sequencing, sequencing technologies have evolved as their costs have reduced, making whole-genome sequencing of many plants possible. According to preliminary data, 400 plants have had their entire genome

sequenced (Fan, 2019). This advancement provides more abundant genetic data for plant variety studies, allowing breeders to conduct complete multidimensional research in genetics, genomics, and molecular breeding. This opens up new development options and provides the impetus for the breeding of more plants which leads to a new revolution in breeding technology. Since the genome sequencing of the first ornamental plant (*Prunus mume*) was achieved in 2012, more than 65 ornamental plants' genomes have been sequenced in less than ten years (Zhang *et al.*, 2012). The whole-genome sequencing results from these ornamental plant species have created a massive resource platform for molecular biology research in ornamental horticulture, which not only contributes to the understanding of genome structure and function in ornamental horticulture but also has significant guiding significance for exploring the origin and evolution of ornamental plants, mapping and cloning the functional genes of important traits, and accelerating the course of research.

Whole Genome Sequencing of Ornamental Plants

As of 30 October 2020, the whole-genome sequences and draft genome sequences of 69 ornamental plants had

been published, including herbaceous plants such as carnation (*Dianthus caryophyllus*), phalaenopsis (*Phalaenopsis aphrodite*), orchid (*Apostasia odorata*), sacred lotus (*Nelumbo nucifera*), chrysanthemum (*Dendranthema morifolium*), and *Dionaea muscipula*, and woody plants, such as mei (*Prunus mume*), Yoshino cherry (*Prunus yedoensis*), sweet osmanthus (*Osmanthus fragrans*), peony (*Paeonia suffruticosa*), and Chinese rose (*Rosa chinensis*). The number of ornamental plant genomes sequenced per year grew considerably from one in 2012 to 17 in 2018. From 2016 to 2018, more than ten species were sequenced for three years. China has finished or led genome sequencing for 32 ornamental plants independently, followed by Japan and the United States, each of which has completed genome sequencing for more than ten species. Except for the double-haploid material with relatively strong homozygosity utilised for *R. chinensis*, all of the other plants were sequenced using wild diploids or cultivars with relatively uncertain genetic histories and low heterozygosity (Raymond *et al.*, 2018; Hibrand *et al.*, 2018). Long-read sequencers are used in conjunction with optical maps to create high-quality chromosome-level genomic assemblies (Schwartz *et al.*, 1993). The PacBio RS II system was originally used to generate the 1.27 Gb genome assembly of *Dendrobium officinale* for ornamental plants (Yan *et al.*, 2015). Long-range scaffolding approaches like high-throughput chromosomal conformation capture (Hi-C) make chromosome-scale contig assembly possible. In this regard, the recently assembled *Rosa chinensis* (515 Mb) genome assembly has a contig N50 of 24 Mb, making it one of the most comprehensive plant genomes (Raymond *et al.*, 2018). Given the extensive use of Illumina HiSeq, Nanopore, PacBio, and Hi-C technologies, the contig N50 values of *Gardenia jasminoides* and *Chimonanthus praecox* can now exceed 44 and 65.35 Mb, respectively, which was impossible five years ago. In general, next-generation sequencing on the Illumina platform (HiSeq 2000/2500/4000 and HiSeq X ten) is used in conjunction with third-generation sequencing (PacBio and Nanopore) and Hi-C technology. The assembled genome size of sequenced ornamental plants ranges from 237 Mb to 13.79 Gb, with a scaffold N50 of 13.8 Kb to 65.35 Mb. phylogenetic trees for all species having a published genome, which are classified into 21 orders and 35 families. The Rosaceae, Orchidaceae, and Asteraceae representative species for which high-quality sequencing has been completed were reported and discussed.

Applications

Gene annotation study

The practice of adding biological information to a species' genomic sequence using bioinformatics approaches is known as gene annotation. It recognizes noncoding RNA gene fragments, recognizes elements on genes (gene prediction), and adds biological information to the elements for sequence repeat identification, noncoding RNA prediction, gene structure prediction, and gene function annotation. This method allows for the identification of genes related to ornamental horticultural properties such as flowering regulation, flower color, floral smell, plant type, dormancy, cold resistance, and disease resistance. The dormancy-associated MADS-box transcription factor (DAM) family, which is involved in dormancy induction and release, is particularly important in ornamental plants (Sasaki *et al.*, 2011). Zhang *et al.*, 2018 found six DAM genes in the *P.*

mume genome tandem array and confirmed that the distribution pattern was comparable with prior investigations of the peach genome. Raymond *et al.*, 2018, found novel candidate genes likely implicated with recurrent blooming in *Rosa*, including TFL1, SPT, and DOG1.

Comparative Genomic Studies

Comparative genomics research examines known genes and genome structures to better understand the roles of related genes, their expression mechanisms, and the evolutionary links of species. The gathering of genomic information from numerous closely related species allows for a more complete and in-depth comparative genomics study. Furthermore, it is critical to conduct an in-depth comparative analysis of the collinear relationship between the genome sequences of two plants to analyze the origin and evolutionary relationship of plants, as well as to investigate important chromosome fragments or gene clusters that control major plant traits, as this information can be used to guide the discovery and cloning of important genes. By comparing Rosaceae genomes, Zhang *et al.*, 2018 created nine ancestral chromosomes of the Rosaceae family. These researchers discovered for the first time that ancestral chromosomes evolved into eight existing chromosomes in *P. mume* through 11 fusions, seven existing chromosomes in strawberry (*Fragaria ananassa*) through 15 fusions, and 17 existing chromosomes in apple (*Malus domestica*) through one whole genome duplication event plus five fusions. These findings provide an important foundation for future research into the origins and development of the Rosaceae.

Resequencing

Whole-genome resequencing entails sequencing genomes from different individuals of species with known genome sequences and analyzing differences between individuals or populations. Whole-genome resequencing technology can be used to undertake rapid resource screening, identify a huge number of genetic variations, and perform genetic evolution analysis and candidate gene prediction for essential features. These findings are important for discovering valuable genetic resources and for horticulture crop breeding, and so have great research and industry significance. Researchers studied the genetic architecture of floral features and the history of plant domestication in *P. mume* by resequencing 348 *P. mume* accessions and three additional *Prunus* species at an average sequencing depth of 19.3. In mei accessions, highly admixed population structure and introgression from *Prunus* species were found. Huang *et al.*, 2018 sequenced and examined the genomes of 19 lotus germplasms, providing a reliable and detailed understanding of the genome evolution of various lotus germplasms and pointing to important mutations responsible for rhizome growth.

Genome-Wide Association Studies

A genome-wide association study (GWAS) is a comparison or correlation analysis that uses millions of SNPs in the genome as biological genetic markers. It is a novel method for identifying genetic variants that influence complex behaviors through comparison. With the advancement of genomics research and DNA microarray technology, a GWAS may now provide an overall overview of multiple essential features at the same time, making it suited for the study of complex qualities. Association studies

between genes and phenotypes are carried out at the genome-wide level, using various centers, huge samples, and repeated verifications. This approach has been used to screen for and identify significant genes for important economic features in agriculture. Researchers identified significant quantitative trait loci (QTLs) and genomic areas in *P. mume* that contain many genes associated with petal color, stigma color, calyx color, bud color, stamina filament color, wood color, petal number, pistil character, bud aperture, and branching phenotype (Zhang *et al.*, 2018). Taken together, the identification of genetic loci related to floral and other features provides greater insight into the genetic mechanisms behind *P. mume* domestication and opens up options to construct genomic selection strategies to improve the performance of ornamental species. The GWAS approach was also used to identify the crucial ornamental attribute of flowering duration in sunflowers and roses (Badouin *et al.*, 2017).

Comparative analysis with transcriptome data

RNA sequencing is a new method that employs next-generation sequencing to analyze transcriptomes. It can acquire sequence and expression information for almost all transcripts from specific cells or tissues in a specific state, including protein-coding mRNAs and various noncoding RNAs, as well as the expression abundance of different transcripts generated by alternative gene splicing, comprehensively and rapidly. The transcriptome is an unavoidable link that connects the genome's genetic information with the proteome's biological activity. Transcriptional regulation is currently the most well-studied and important regulatory technique in organisms. Transcriptome investigations are the foundation and beginning point for gene function structure studies, and they are the first topic to address after whole-genome sequencing is completed. Furthermore, transcriptome analysis yields a vast number of molecular markers, such as SSR and SNP. All of the sequence information, expression data, and molecular markers aid in the genetic mapping of QTLs for key ornamental traits in ornamental plants and contribute to the development of molecular markers in close linkage with excellent traits for use in molecular marker-assisted flower breeding. Based on the *P. mume* genome sequence, significant variations in gene expression were discovered between the bud and squaring stages, and 7,813 DEGs were identified, providing a unique perspective on floral fragrance production in *P. mume* (Zhao *et al.*, 2017). In a comparative transcriptome analysis, the water lily genome revealed varying genomic markers of ancient vascular cambium losses, and the expression patterns of floral ABCE genes, floral scent, and color genes were screened from the DEGs (Zhang *et al.*, 2020).

Development of SNP microarrays

SNPs can occur in coding regions, noncoding regions, or gene spacer regions, depending on their location in the gene. They are DNA molecular markers that have the greatest polymorphisms in the genome and are distinguished by their large number, uniform distribution, and ease of typing. SNPs can be used to identify genetic variation and genotype-associated traits. The use of SNPs as molecular markers to create distinct genetic maps of the genome has become a crucial aspect of research for analyzing genome diversity, getting domesticated selection regions, and

screening essential genes for important phenotypes. *P. mume's* genome sequence and resequencing revealed a total of 1,298,196 raw SNPs within coding regions of genes, 733,292 of which were nonsynonymous. Furthermore, by merging transcriptome data, 76 SNPs within DEGs linked with petal, stigma, calyx, and bud color were discovered (Zhang *et al.*, 2018). Based on SNP locations in resequenced species, sacred lotus, wild lotus, and Thai lotus showed greater divergence and genetic diversity than farmed lotus (Huang *et al.*, 2018).

Association of gene exploitation with various ornamental traits

A huge number of genes, in the region of 19,507-87,603, are annotated for each flowering species during whole-genome sequencing. Important genes involved in floral development, flower color production, and stress resistance can be uncovered through additional research. This promotes the development of distinct, high-quality, and resistant variants or kinds of a species, as well as significant references for improving ornamental and resistance properties in other flowering species.

Candidate genes governing flower development

Flower blooming is a process that involves the formation of inflorescence meristems and flower meristem tissues via floral induction and a variety of internal and external factors, followed by the formation of floral organ primordia and the eventual release of floral bud dormancy to form floral organs. *A. thaliana*, the flowering process is governed by a complex regulatory network, with at least seven flowering regulation pathways identified (Zhao *et al.*, 2017). Floral development genes can affect the production of inflorescence meristems as well as the direction of newly created floral primordia. These genes regulate plant blooming time by directing the production of inflorescence meristems or flower meristems, and mutations in these genes can result in mutants that flower earlier or later. The other class includes genes that control the production of floral organs, and mutations in these genes can result in homeoboxes (Zhao *et al.*, 2017). During long-term artificial domestication and cultivation, the form and number of floral organs in ornamental plants have undergone significant modifications, such as double petals, multiple sepals, and multiple pistils and stamens growing into independent flowers. These variations enhance the ornamental appeal of ornamental plants while also offering valuable materials for research into floral organ development in plants. Some key genes related to flowering transition and flower development have been analyzed using genomic data analysis as an important scientific issue, such as those in *Tarenaya hassleriana* (Cheng *et al.*, 2013), *Dendrobium officinale* (Yan *et al.*, 2015), *Primula veris* (Nowak *et al.*, 2015), *Dendrobium catenatum* (Zhang *et al.*, 2016), *Hibiscus syriacus* (Kim *et al.*, 2016), *Rosa spp* (Nakamura *et al.*, 2018), *Chrysanthemum spp* (Song *et al.*, 2018; Hirakawa *et al.*, 2019), and *Nymphaea colorata* (Zhang *et al.*, 2020).

Candidate gene governing anthocyanin synthesis

Flower color is one of the most important characteristics of ornamental plants. Anthocyanin is a necessary pigment for floral color, and its manufacture is catalyzed by several enzymes (Tanaka and Ohmia, 2008). Different anthocyanins are generated as a result of changes in

substituent groups at distinct places on the basic skeleton, resulting in different plant organ colors like as red, purple, blue-purple, and blue. Anthocyanins are flavonoid secondary metabolites in plants and the most abundant water-soluble pigments in nature, where they serve an important role in color production and antioxidation in plant flowers and fruits. The R2R3-MYB genes are involved in the synthesis of anthocyanin (Xu *et al.*, 2015). *P. mume* has 96 R2R3-MYB genes that have been classified into 35 subfamilies. Finally, the over-expression of PmMYB1 and PmMYBa1 in tobacco revealed their functions, which greatly increased anthocyanin accumulation in transgenic tobacco. The flower colors of PmMYB1-over expressing transgenic plants were substantially deeper, and the anthocyanin content of transgenic plants' corollas was significantly higher than that of the control (Zhang *et al.*, 2017). The major blue anthocyanidin pigment in water lily was identified as delphinidin 3'-O-, and some genes for an anthocyanidin synthase and a delphinidin-modification enzyme were screened by comparing the expression profiles of two *N. colorata* cultivars with white and blue petals (Zhang *et al.*, 2020). Blue blooms developed after the butterfly pea UDP (uridine diphosphate)-glucose: anthocyanin 3', the 5'-O-glucosyltransferase gene was transferred into chrysanthemums (Noda *et al.*, 2017). Two MYB transcription factors have been shown to impact flower color in *Rosa rugosa* via regulating flavonoid production in response to wounding and oxidation (Shen *et al.*, 2019). A chalcone synthase (PhCHS) associated in flavonoid biosynthesis and two anthocyanin O-methyltransferase (AOMT) genes were found in *Paeonia*, which was consistent with anthocyanin accumulation in petals (Du *et al.*, 2015; Gu *et al.*, 2019).

Candidate genes governing flower scent biosynthesis

The floral fragrance has tremendous aesthetic, economic, and application value as one of the quality features of ornamental plants. Secondary metabolites found in petals include esters, alcohols, ketones, aldehydes, terpenes, and volatile phenols, which are predominantly generated by terpene metabolism, phenylpropane metabolism, and the lipoxygenase pathway (Dudareva and Pichersky, 2011). Diverse petals contain different sorts of fragrance components, resulting in varied scents among different flower species. Zhang *et al.*, 2018 reported that the benzyl alcohol acetyltransferase (BEAT) gene can directly catalyze the synthesis of benzyl acetate, a critical component of the floral aroma in *P. mume*, in a study on the molecular mechanism responsible for the flowery scent in *P. mume*. Furthermore, 44 unique PmBEATs were discovered in *P. mume* based on genomic data from *P. mume* and *P. persica*, significantly more than the 16 reported in apple, 14 in strawberry, and four in grape. These PmBEAT genes arose from gene duplication events in *P. mume* species evolution, with retroduplication and tandem duplication being the two prevalent duplication patterns. Overexpression of the PmBEAT36 or PmBEAT37 genes increased benzyl acetate production in *P. mume* petal protoplasts, while interfering with the expression of these genes decreased benzyl acetate concentration significantly (Bao *et al.*, 2019). Zhao *et al.*, 2017 performed a comparative transcriptome analysis of flower genes linked with floral features in different developmental stages and tissues and identified 12 novel genes involved in floral fragrance production in *P. mume*.

Furthermore, five TFs from *Phalaenopsis bellina* (bHLH4, bHLH6, bZIP4, ERF1, and NAC1) are implicated in orchid floral monoterpenes. PrCYP79D73 is involved in floral volatile organic compounds and other nitrogen-containing volatiles in *Plumeria rubra* (Dhandapani *et al.*, 2019).

Candidate gene governing plant architecture

Plant architecture is rich and diversified as a result of long-term evolution, natural selection, and a complex regulatory process involving genetics and the environment. Various plant architecture traits are not only advantageous for the production of rich and diverse horticultural landscapes, but they are also advantageous for plant adaptability to complicated contexts and competition, as well as the consumption of light and nutrients. The findings, along with the completion of whole-genome sequencing for many ornamental plants of the genus *Prunus*, provide an important data foundation for further research into the fundamental genetic underpinnings of pendulous characteristics (Verde *et al.*, 2013). Zhang *et al.*, 2018, developed a high-density genetic map based on the eight scaffolds of the *P. mume* genome and mapped QTLs for important features such as plant type, flower color, petals, and leaves in *P. mume* using specific-length amplified fragment sequencing (SLAF). They discovered ten SLAF markers that were strongly associated to *P. mume* pendulous characteristics. The pendulous features of *P. mume* were finely mapped to a 1.14 cM area on chromosome 7, and 36 candidate genes that might be related to the pendulous traits of *P. mume* were suggested using these markers (Zhang *et al.*, 2015). Using genome and bulked segregant analyses, researchers made breakthroughs in the mining and tagging of genes for weeping and dwarf characteristics in peach (*P. persica*) (Hollender *et al.*, 2018).

Candidate genes governing dormancy release

Prunus mume and *P. yedoensis* flowers are the first to bloom in the spring. Zhang *et al.*, 2015 investigated the molecular pathways behind *P. mume* dormancy rupture and flowering at low temperatures. The researchers discovered six dormancy-associated MADS-box (DAM) genes in the genome with a tandem repeat distribution. PmDAM1, PmDAM3, PmDAM2, PmDAM5, PmDAM4, and PmDAM6 are formed from a sequence of duplication events in the following order: PmDAM1, PmDAM3, PmDAM2, PmDAM5, PmDAM4, and PmDAM6. Tandem genes have not been detected in *M. domestica* or *F. ananassa*, although the molecular evolution pattern of DAM genes is unique to *Prunus* plants. This phenomenon could be attributed to *Prunus* plants blooming earlier, including *P. persica*, *P. mume*, apricot (*Armeniaca vulgaris*), and sweet cherry (*Prunus avium*), than most other flowering species (Zhang *et al.*, 2012). C-repeat-binding transcription factors control DAM genes (CBFs). In *P. persica* and plum, a conserved CBF site was discovered 1000 bp upstream of the DAM4-DAM6 transcription start site (*Prunus salicina*). According to the most recent research findings, a sense-response link between PmCBFs and PmDAMs exists in cold-induced dormancy and is concurrently regulated by six PmCBFs and PmDAM4-6 (Zhao *et al.*, 2018).

Candidate gene governing biotic stress resistance

Disease resistance is an important characteristic that has sparked interest in the study across all flowering plants. As a result, the whole-genome study concentrates on genes related

to disease resistance. R genes, which encode proteins with extremely high structural similarities, such as leucine zippers, nucleotide-binding sites, transmembrane domains, leucine-rich repeats, and similar extracellular regions of drosophilid toll protein and mammalian toll and interleukin-1 receptor, are primarily involved in plant disease resistance (TIR). The gene family with the greatest distribution and number of plant R genes is the nucleotide-binding site 'leucine-rich repeat genes'. The nucleotide-binding site is near the N-terminus of their encoded proteins, while the leucine-rich repeat is near the C-terminus. The N-terminus of proteins produced by various genes may also contain one or more of the two conserved motifs described below: the coiled-coil motif and the TIR motif. The genome of *P. mume* had 253 leucine-rich repeats receptor-like kinase (LRR-RLK) genes, and most pathogenesis-related (PR) gene families were significantly increased and ordered in tandem, particularly PR103 (Zhang *et al.*, 2021). Resistance (R) genes account for 0.53% of the total predicted genes in *Hibiscus syriacus*, which is lower than in other plants studied in genomic research (0.63 to 1.35%) (Kim *et al.*, 2016). The genome of *Asparagus setaceus* contained 76 R genes containing nucleotide-binding sites (NBSs), which were classified into five groups: TIR-NBS, CC-NBS-LRR, NBS-LRR, NBS, and CC-NBS. The NBS-LRR group was the largest, with 29 genes (Li *et al.*, 2020).

Candidate gene governing abiotic stress resistance

Low temperature, humidity, heat, drought, and saline-alkali environments all severely impede the growth and development of attractive plants. These circumstances can alter plant physiology, biochemistry, and morphology, and can even result in death. Because of this difficulty, decorative plant cultivation facilities are cumbersome and cannot be widely promoted, negatively impacting their quality and benefits. Low temperature is a significant element that limits plant growth, development, and geographical dissemination. Low-temperature stress is classified as chilling stress ($>0^{\circ}\text{C}$) and freezing stress (0°C). Plants from the tropics and subtropics are more vulnerable to cold; plants from temperate regions, on the other hand, have evolved intricate mechanisms to withstand and adapt to chilling (freezing) stress, protecting the plants from harm. Cold acclimation is a protective mechanism for plant adaptation and tolerance to low-temperature stress that is controlled by a complicated network (Shi *et al.*, 2018). The CBF route, in particular, is regarded as the most significant and well-studied (Liu *et al.*, 2018). 30 LEA genes were found based on *P. mume* genome data, and heterologous expression of PmLEA enhanced the cold resistance of *Escherichia coli* and tobacco (*Nicotiana tabacum*) (Bao *et al.*, 2017; Du *et al.*, 2013). Based on yeast two-hybrid and bimolecular fluorescence complementation investigations, a molecular regulation model of the PmDAM and PmCBF genes in response to dormancy and dormancy release of flower buds triggered by low-temperature signals⁹⁴ was developed (Zhao *et al.*, 2018).

Prospects of whole genome sequencing data for ornamental crops

The Earth BioGenome Effort (EBP) is a 10-year biology project that intends to sequence, classify, and characterize the genomes of all Earth's eukaryotic species. The primary scientific challenges for plants are increasing

agricultural yields and other agronomically essential features, producing biofuel, gene editing, and conserving endangered species. The 10,000 Plant Genome Sequencing Project (10KP), launched by the Beijing Genomics Institute in Shenzhen (BGI, Shenzhen), is a historic attempt to catalog plant genetic variation and represents a significant step forward in comprehending the tree of life. The National Engineering Research Center for Floriculture in China has proposed a preliminary proposal for the 100 Flowers Genome Sequencing Project. Many ornamentals have high ploidy levels and homologous polyploids (chrysanthemum and alfalfa) or exceptionally large genome sizes (lily and tulip), limiting the development and use of genome sequencing technology in ornamental plants. With the advancement of sequencing and bioinformatics analysis tools, as well as the ongoing introduction of new biological technologies, genomics research on ornamental plants have progressed faster and more efficiently. Despite significant problems in genome sequencing and assembly of flowering plants, the quality of genome assembly results is relatively excellent in terms of the analytical data from 69 flower species that underwent genome sequencing, and four of them were resequenced using updated sequence technology (Clouse *et al.*, 2016; Hibrand *et al.*, 2018; Gui *et al.*, 2018). As far as we know, at least a dozen ornamental plants are undergoing genome quality upgrading. As more ornamental plant genomes are sequenced, bioinformatics analysis could disclose important basic information about species origins and the genes that influence flower features. The advancement of genomics will undoubtedly fill knowledge gaps in traditional breeding strategies. The ultimate goal is to obtain the most effective and rapid approach for obtaining the ideal type of flower variation with fixed-point improvement and the aggregation of several elite features. Higher (flowering) plants numbered 30,000 in China, and some aesthetic blooming plants arrived in Europe relatively early (Wilson, 1929). Chinese people adore flowers and grow a variety of them, including mei, peony, chrysanthemum, rose, lily, lotus, and orchid. Because of the rapid growth of genome sequencing technology around the world, enormous amounts of whole-genome sequencing data are in desperate need of deep mining. A long-term strategic genomics research plan that is not limited to cultivated species but also considers the thorough development of the sequencing of important wild relatives of ornamental species in China as well as promoting the mining, protection, and utilization of important genetic resources should be developed. It is critical to abandon reliance on apparent phenotypes, shift investigations to genotype-dependent research, and shift single gene studies to GWAS.

Plant genomes have always been a significant subject of genetics because of their complexity and uniqueness. Before the second generation of high-throughput sequencing, sequencing costs were high and throughput was poor. It was too difficult or expensive for researchers to collect whole-genome sequences of high repeat sequence species. Many species with significant economic and decorative significance have yet to be submitted for comprehensive genome sequencing. In short, because ornamental plants are unique and diverse, there are both problems and opportunities in a genomic study of these species. Challenge: (1) A complicated genome. The phrase complicated genome refers to a type of genome that cannot

be studied directly using traditional sequencing and assembly procedures. It usually refers to a genome with a high number of repetitive sequences, a high heterozygosity, a high GC content, and difficulty in removing foreign DNA contamination. (2) Polyploidy. Ornamental plants frequently exhibit autopolyploidy. It is often created by doubling two or more sets of genomes, and it is extremely useful in genetic breeding and agricultural output. It is simple to connect erroneous allele fragments using traditional procedures, resulting in the improper connecting of homologous chromosomes. (3) Megagenome. Mega genomes are species having genomes greater than 10 Gb. The sequencing and analysis of these species are quite complicated, especially for assembly analysis, which is a big difficulty. *Paris japonica* is an interesting plant. Scientists discovered that it contains the world's largest genome, with 150 Gb, which is 50 times larger than that of humans. But the genomes of certain decorative plants have been judged complete, the assembly quality of other species is low, and a limited number of "holes" have not yet been plugged due to technical restrictions, although scientists' interest in this area is controversial. The most recent study indicates that segments in the genome that were previously thought to be irrelevant or "junk" have their meaning. These missing sequences are crucial, and we now have the potential to mine them. Third-generation sequencing technology (PacBio and Nanopore) can fill gaps in some difficult-to-assemble genomic areas caused by sequencing errors, repeat regions, heterochromatin, genomic polymorphisms, and second-generation sequencing preferences. The following innovative methods can be tried using third-generation sequencing technology to overcome the difficulty of sequencing the genomes of ornamental plants. (1) The pangenome. The pangenome encompasses both the essential and non-essential genomes. The core genome refers to genes that occur in all individuals, whereas the nonessential genome refers to genes that exist in just certain persons. (2) Hi-C. The advantages of Hi-C sequencing technology are as follows: on the one hand, no large number of F1 populations are required because only individuals are required; on the other hand, the haplotype genome can be separated without parent purification, making this method suitable for the assembly of a highly heterozygous genome that is difficult to assemble.

The notions of challenging genome sequencing and assembly quality have evolved and changed in tandem with the advancement of sequencing technology. We can't sequence everything only to sequence the genome. The goal of sequencing must be to disclose the species' main scientific problems. Transcriptomics, metabolomics, proteomics, degradomics, and phenomics research should be prioritized. With the publication of more genomic data, it has become increasingly difficult to evaluate, store, and exchange vast amounts of genome sequencing data. A critical issue is how to address the time and expense constraints that researchers confront to fulfill the goals of decreasing repetitious research, enhancing the practicability of scientific research, mining research content, and improving the transparency of scientific research. Furthermore, bioinformatics education and application in practice must be improved. We predict that, with the ongoing advancement of sequencing technology, whole-genome sequencing of horticulture crops will enter a rapid development stage shortly, resulting in enormous contributions to the global horticultural business.

References

- Badouin, H.; Gouzy, J.; Grassa, C.J.; Murat, F.; Staton, S.E.; Cottret, L.; Lelandais-Brière, C.; Owens, G.L.; Carrère, S.; Mayjonade, B. and Legrand, L. (2017). The sunflower genome provides insights into oil metabolism, flowering and Asterid evolution. *Nature*, 546 (7656):148-52.
- Bao, F.; Ding, A.; Zhang, T.; Luo, L.; Wang, J.; Cheng, T. and Zhang, Q. (2019). Expansion of PmBEAT genes in the *Prunus mume* genome induces characteristic floral scent production. *Horticulture Research*, 1: 6.
- Cheng, S.; van den Bergh, E.; Zeng, P.; Zhong, X.; Xu, J.; Liu, X.; Hofberger, J.; de Bruijn, S.; Bhide, A.S.; Kuelahoglu, C. and Bian, C. (2013). The *Tarenaya hassleriana* genome provides insight into reproductive trait and genome evolution of crucifers. *The Plant Cell*, 25(8): 2813-30.
- Clouse, J.W.; Adhikary, D.; Page, J.T.; Ramaraj, T.; Deyholos, M.K.; Udall, J.A.; Fairbanks, D.J.; Jellen, E.N. and Maughan, P.J. (2016). The amaranth genome: genome, transcriptome, and physical map assembly. *Plant Genome*.
- Dhandapani, S.; Jin, J.; Sridhar, V.; Chua, N.H. and Jang, I.C. (2019). CYP79D73 participates in biosynthesis of floral scent compound 2-phenylethanol in *Plumeria rubra*. *Plant Physiology*, 180(1):171-84.
- Du, H.; Wu, J.; Ji, K.X.; Zeng, Q.Y.; Bhuiya, M.W.; Su, S.; Shu, Q.Y.; Ren, H.X.; Liu, Z.A. and Wang, L.S. (2015). Methylation mediated by an anthocyanin, O-methyltransferase, is involved in purple flower coloration in *Paeonia*. *Journal of Experimental Botany*, 66(21): 6563-6577.
- Dudareva, N. and Pichersky, E. (2000). Biochemical and molecular genetic aspects of floral scents. *Plant Physiology*, 122(3): 627-634.
- Fan, J. *Plant Genomics*. (Science Press, 2019).
- Gu, Z.; Men, S.; Zhu, J.; Hao, Q.; Tong, N.; Liu, Z.A.; Zhang, H.; Shu, Q. and Wang, L. (2019). *Chalcone synthase* is ubiquitinated and degraded via interactions with a RING-H2 protein in petals of *Paeonia* 'He Xie'. *Journal of experimental botany*. 70(18): 4749-4762.
- Gui, S.; Peng, J.; Wang, X.; Wu, Z.; Cao, R.; Salse, J.; Zhang, H.; Zhu, Z.; Xia, Q.; Quan, Z.; Shu, L. (2018). Improving *Nelumbo nucifera* genome assemblies using high-resolution genetic maps and BioNano genome mapping reveals ancient chromosome rearrangements. *The Plant Journal*, 94(4): 721-34.
- Hibrand Saint-Oyant, L.; Ruttink, T.; Hamama, L.; Kirov, I.; Lakhwani, D.; Zhou, N.N.; Bourke, P.M.; Daccord, N.; Leus, L.; Schulz, D.; Van de Geest, H. (2018). A high-quality genome sequence of *Rosa chinensis* to elucidate ornamental traits. *Nature plants*, 4(7): 473-84.
- Hirakawa, H.; Sumitomo, K.; Hisamatsu, T.; Nagano, S.; Shirasawa, K.; Higuchi, Y.; Kusaba, M.; Koshioka, M.; Nakano, Y.; Yagi, M. and Yamaguchi, H. (2019). De novo whole-genome assembly in *Chrysanthemum seticospe*, a model species of *Chrysanthemums*, and its application to genetic and gene discovery analysis. *DNA research*. 26(3):195-203.
- Hollender, C.A.; Pascal, T.; Tabb, A.; Hadiarto, T.; Srinivasan, C.; Wang, W.; Liu, Z.; Scorza, R. and Dardick, C. (2018). Loss of a highly conserved sterile alpha motif domain gene (WEEP) results in pendulous

- branch growth in peach trees. *Proceedings of the National Academy of Sciences*, 115(20): E4690-9.
- Huang, L.; Yang, M.; Li, L.; Li, H.; Yang, D.; Shi, T. and Yang, P. (2018). Whole genome re-sequencing reveals evolutionary patterns of sacred lotus (*Nelumbo nucifera*). *Journal of Integrative Plant Biology*, 60(1): 2-15.
- Kim, Y.M.; Kim, S.; Koo, N.; Shin, A.Y.; Yeom, S.I.; Seo, E.; Park, S.J.; Kang, W.H.; Kim, M.S.; Park, J. and Jang, I. (2017). Genome analysis of *Hibiscus syriacus* provides insights of polyploidization and indeterminate flowering in woody plants. *DNA Research*, 24(1): 71-80.
- Li, L.F.; Cushman, S.A.; He, Y.X. and Li, Y. (2020). Genome sequencing and population genomics modeling provide insights into the local adaptation of weeping forsythia. *Horticulture research*.
- Liu, J.; Shi, Y. and Yang, S. (2018). Insights into the regulation of C-repeat binding factors in plant cold signaling. *Journal of Integrative Plant Biology*, 60(9): 780-795.
- Nakamura, N.; Hirakawa, H.; Sato, S.; Otagaki, S.; Matsumoto, S.; Tabata, S. and Tanaka, Y. (2018). Genome structure of *Rosa multiflora*, a wild ancestor of cultivated roses. *DNA Research*, 25(2): 113-21.
- Noda, N.; Yoshioka, S.; Kishimoto, S.; Nakayama, M.; Douzono, M.; Tanaka, Y. and Aida, R. (2017). Generation of blue chrysanthemums by anthocyanin B-ring hydroxylation and glucosylation and its coloration mechanism. *Science Advances*, 3(7): e1602785.
- Nowak, M.D.; Russo, G.; Schlapbach, R.; Huu, C.N.; Lenhard, M. and Conti, E. (2015). The draft genome of *Primula veris* yields insights into the molecular basis of heterostyly. *Genome Biology*, 16(1): 1-7.
- Raymond, O.; Gouzy, J.; Just, J.; Badouin, H.; Verdenaud, M.; Lemainque, A.; Vergne, P.; Moja, S.; Choisne, N.; Pont, C. and Carrere, S. (2018). The *Rosa* genome provides new insights into the domestication of modern roses. *Nature Genetics*, 50(6): 772-777.
- Sasaki, R.; Yamane, H.; Ooka, T.; Jotatsu, H.; Kitamura, Y.; Akagi, T. and Tao, R. (2011). Functional and expressional analyses of PmDAM genes associated with endodormancy in Japanese apricot. *Plant Physiology*, 157(1): 485-497.
- Schwartz, D.C.; Li, X.; Hernandez, L.I.; Ramnarain, S.P.; Huff, E.J. and Wang, Y.K. (1993). Ordered restriction maps of *Saccharomyces cerevisiae* chromosomes constructed by optical mapping. *Science*, 262 (5130):110-114.
- Shen, Y.; Sun, T.; Pan, Q.; Anupol, N.; Chen, H.; Shi, J.; Liu, F.; Deqiang, D.; Wang, C.; Zhao, J.; Yang, S. (2019). Rr MYB 5-and Rr MYB 10-regulated flavonoid biosynthesis plays a pivotal role in feedback loop responding to wounding and oxidation in *Rosa rugosa*. *Plant biotechnology Journal*, 17(11): 2078-95.
- Shi, Y.; Ding, Y. and Yang, S. (2018). Molecular regulation of CBF signaling in cold acclimation. *Trends in plant science*. 23 (7): 623-37.
- Song, C.; Liu, Y.; Song, A.; Dong, G.; Zhao, H.; Sun, W.; Ramakrishnan, S.; Wang, Y.; Wang, S.; Li, T. and Niu, Y. (2018). The *Chrysanthemum nankingense* genome provides insights into the evolution and diversification of chrysanthemum flowers and medicinal traits. *Molecular Plant*. 11 (12): 1482-1491.
- Tanaka, H.; Hirakawa, H.; Kosugi, S.; Nakayama, S.; Ono, A.; Watanabe, A.; Hashiguchi, M.; Gondo, T.; Ishigaki, G.; Mugerza, M. and Shimizu, K. (2016). Sequencing and comparative analyses of the genomes of zoysiagrasses. *DNA Research*, 23(2): 171-180.
- Verde, I.; Abbott, A.G.; Scalabrin, S.; Jung, S.; Shu, S.; Marroni, F.; Zhebentyayeva, T.; Dettori, M.T.; Grimwood, J.; Cattonaro, F. and Zuccolo, A. (2013). The high-quality draft genome of peach (*Prunus persica*) identifies unique patterns of genetic diversity, domestication and genome evolution. *Nature genetics*, 45(5): 487-494.
- Wilson, E.H. China, Mother of Gardens. (The Stratford Company, 1929).
- Xu, W.; Dubos, C. and Lepiniec, L. (2015). Transcriptional control of flavonoid biosynthesis by MYB-bHLH-WDR complexes. *Trends in Plant Science*, 20(3): 176-185.
- Yan, L.; Wang, X.; Liu, H.; Tian, Y.; Lian, J.; Yang, R.; Hao, S.; Wang, X.; Yang, S.; Li, Q.; Qi, S. (2015). The genome of *Dendrobium officinale* illuminates the biology of the important traditional Chinese orchid herb. *Molecular Plant*, 8(6): 922-934.
- Zhang, G.Q.; Xu, Q.; Bian, C.; Tsai, W.C.; Yeh, C.M.; Liu, K.W.; Yoshida, K.; Zhang, L.S.; Chang, S.B.; Chen, F. and Shi, Y. (2016). The *Dendrobium catenatum* Lindl. genome sequence provides insights into polysaccharide synthase, floral development and adaptive evolution. *Scientific Reports*, 6(1):1-10.
- Zhang, J.; Zhang, Q.; Cheng, T.; Yang, W.; Pan, H.; Zhong, J.; Huang, L. and Liu, E. (2015). High-density genetic map construction and identification of a locus controlling weeping trait in an ornamental woody plant (*Prunus mume* Sieb. et Zucc). *DNA research*, 22(3): 183-191.
- Zhang, L.; Chen, F.; Zhang, X.; Li, Z.; Zhao, Y.; Lohaus, R.; Chang, X.; Dong, W.; Ho, S.Y.; Liu, X. and Song, A. (2020). The water lily genome and the early evolution of flowering plants. *Nature*, 577 (7788): 79-84.
- Zhang, Q.; Zhang, H.; Sun, L.; Fan, G.; Ye, M.; Jiang, L.; Liu, X.; Ma, K.; Shi, C.; Bao, F.; Guan, R. (2018). The genetic architecture of floral traits in the woody plant *Prunus mume*. *Nature communications*. 9 (1):1-2.
- Zhao, D.; Hamilton, J.P.; Pham, G.M.; Crisovan, E.; Wiegert-Rininger, K.; Vaillancourt, B.; DellaPenna, D. and Buell, C.R. (2017). De novo genome assembly of *Camptotheca acuminata*, a natural source of the anti-cancer compound camptothecin. *Giga Science*. 6(9): gix065.