

Review

Recent progress in whole genome sequencing, high-density linkage maps, and genomic databases of ornamental plants

Masafumi Yagi*

Institute of Vegetable and Floriculture Science, NARO, 2-1 Fujimoto, Tsukuba, Ibaraki 305-0852, Japan

Genome information is useful for functional analysis of genes, comparative genomic analysis, breeding of new varieties by marker-assisted selection, and map-based gene isolation. Genome-related research in ornamental plants has been relatively slow to develop because of their heterozygosity or polyploidy. Advances in analytical instruments, such as next-generation sequencers and information processing technologies have revolutionized biology, and have been applied in a large number and variety of species, including ornamental plants. Recently, high-quality whole genome sequences have been reported in plant genetics and physiology studies of model ornamentals, such as those in genus *Petunia* and Japanese morning glory (*Ipomoea nil*). In this review, whole genome sequencing and construction of high-density genetic linkage maps based on SNP markers of ornamentals will be discussed. The databases that store this information for ornamentals are also described.

Key Words: genetic linkage map, marker, simple sequence repeat (SSR), single nucleotide polymorphism (SNP), whole genome, next-generation sequencing (NGS).

Introduction

Ornamental plants include a wide range of plants. Consumers demand new types of floricultural crops and are eager for new cultivars with ornamental value, such as flowers with new shapes or colors. Therefore, the lifespan of ornamental plant cultivars is generally much shorter than that of other crops, and breeders or breeding companies continuously strive to develop and release new cultivars (Shibata 2008). For major ornamentals such as chrysanthemum, rose, and carnation, new ornamental cultivars are commonly produced by hybridization between elite cultivars and propagated asexually. The genetic background of most ornamentals is highly heterozygous, with polyploidy also being observed in some species. This situation complicates detailed genetic analyses using crossing populations, and, as a consequence, the development of sophisticated breeding strategies in ornamentals has lagged behind those for most agricultural crops (Yagi 2015). Genomics analysis tools include expressed sequence tags (ESTs), bacterial artificial chromosome (BAC) libraries, physical and genetic linkage maps, and molecular markers, which are useful for genetic

analyses of useful traits and for isolating the responsible genes (Han *et al.* 2007). Previously, I described the developments in genomic analysis resources and next-generation sequencing (NGS) applications to major ornamentals (Yagi 2015). At that time, comprehensive whole transcriptome sequences had been obtained for most of the major ornamentals, but the whole genome sequences were obtained only for carnation. Until now, the genome sequences of major ornamentals such as chrysanthemum, rose, and lily have not been reported, but high-quality genome sequences of parental wild species of petunia and Japanese morning glory (*Ipomoea nil*), which have been used as model plants in plant genetics and physiology studies, have been reported recently (Bombarely *et al.* 2016, Hoshino *et al.* 2016). In this review, I describe the latest reports of whole genome sequences, high-density linkage maps, and genomic databases for ornamental plants.

Whole genome sequences

After the release of the *Arabidopsis* genome in 2000 (The Arabidopsis Genome Initiative 2000) and the advent of NGS technology in 2005, the number of sequenced plant genomes increased rapidly to more than 100 (Michael and VanBuren 2015). High-throughput and low-cost genome sequencing technologies have enabled the determination of whole genome sequences in many non-model plant species.

Communicated by Akemi Ohmiya

Received July 31, 2017. Accepted October 19, 2017.

First Published Online in J-STAGE on February 17, 2018.

*Corresponding author (e-mail: myagi@affrc.go.jp)

Most of the massive transcriptome datasets generated from major ornamentals, including chrysanthemum (Wang *et al.* 2013), rose (Dubois *et al.* 2012), and carnation (Tanase *et al.* 2012), have been obtained using NGS technology (Yagi 2015). Whole genome sequences in ornamentals have been reported for carnation (Yagi *et al.* 2014a), *Phalaenopsis equestris* (Cai *et al.* 2015b), *Dendrobium officinale* (Yan *et al.* 2015), *Primula veris* (Nowak *et al.* 2015), *Dendrobium catenatum* (Zhang *et al.* 2016b), *Phalaenopsis* (Huang *et al.* 2016), *Petunia* (Bombarley *et al.* 2016), *Ipomoea nil* (Hoshino *et al.* 2016), *Hibiscus syriacus* (Kim *et al.* 2017), and *Helianthus annuus* (Badouin *et al.* 2017) (Table 1). In addition, snapdragon (*Antirrhinum majus*) genomic sequences are available for restricted users on the Snapdragon Genome Database website (<http://snapdragon.genomics.org.cn/page/species/index.jsp>). The shotgun genome sequence of Madagascar periwinkle, *Catharanthus roseus* (L.) G. Don (Kellner *et al.* 2015) and *Rosa roxburghii* Tratt (Lu *et al.* 2016), both of which have medical value, have been surveyed using only single-length libraries.

Carnation

In terms of use in actual breeding, carnation (*Dianthus caryophyllus* L.) genomics has been widely applied. When breeding for resistance to carnation bacterial wilt (CBW), Onozaki *et al.* (2004) developed a tightly linked marker for CBW resistance derived from *Dianthus capitatus* Balbis ex DC. ssp. *andrzejowskianus* Zapal. by bulked segregant analysis. Using this marker in an actual breeding program, we produced the new resistant cultivar ‘Karen Rouge’, which is resistant to CBW (Yagi *et al.* 2010). In parallel with the

breeding, we updated the carnation linkage maps to conduct QTL analysis and identify the linked markers for important traits (Yagi *et al.* 2006, 2012, 2013, 2017). The breeding process of ‘Karen Rouge’ and related genomic research in carnation have been well reviewed by Yagi (2013, 2015). Using the developed linkage maps, we identified linked markers for flower-type (Onozaki *et al.* 2006, Yagi *et al.* 2014b) and CBW resistance from line 85-11, which has a different resistance from *D. capitatus*, and identified QTLs for flower anthocyanin content (Yagi *et al.* 2012, 2013). Carnation, a heterologous diploid flower crop, was the first ornamental flower to be sequenced (Yagi *et al.* 2014a). The genome of ‘Francesco’, which is the leading cultivar in Japan, was sequenced (Yagi *et al.* 2014a). HiSeq and GS FLX+ sequencers together produced 45,088 scaffolds that spanned 568.9 Mb (91% of the carnation genome). A total of 43,266 protein-coding genes were deduced. Mapping of 248 core eukaryotic genes using the CEGMA program indicated that 96% of the core genes were completely covered in the genome assembly. In addition, genes related to flower color, flower longevity, flower scent, and disease resistance were annotated.

Orchid

Orchidaceae constitutes the largest family of flowering plants, with the number of species possibly exceeding 25,000 (Hsiao *et al.* 2011). Among them, *Phalaenopsis* species are popular ornamental plants worldwide because of their elegant appearance and extended longevity, which makes them of great economic importance for the floral industry. *P. equestris* is an important breeding parent because

Table 1. Whole genome sequences of ornamental plant species

Species	Cultivar/strain name	Estimated genome size	Chromosome number	Total length of assembled genome sequence	Number of scaffolds	Scaffold N50	Numbers of predicted protein coding gene	Sequencer	Reference	Country
<i>Dianthus caryophyllus</i>	Francesco	622 Mb	2n = 2x = 30	568.9 Mb	45,088	61 Kb	43,266	HiSeq1000, GS FLX+	Yagi <i>et al.</i> 2014a	Japan
<i>Phalaenopsis equestris</i>	Unnamed inbred line	1.2 Gb	2n = 2x = 38	1.1 Gb	236,185	359 Kb	29,431	HiSeq 2000	Cai <i>et al.</i> 2015b	China, Taiwan, Belgium
<i>Dendrobium officinale</i>	GREEN	1.3 Gb	2n = 2x = 38	1.4 Gb	751,466	76 Kb	34,699	HiSeq2000, PacBio	Yan <i>et al.</i> 2015	China
<i>Primula veris</i>	Unnamed cultivated line	479 Mb	2n = 2x = 22	310.1 Mb	8,764	164 Kb	19,507	Hiseq2000, Miseq, Ion Proton, PacBio	Nowak <i>et al.</i> 2015	Switzerland, Norway, Germany
<i>Dendrobium catenatum</i>	Unnamed wild plants	1.1 Gb	2n = 2x = 38	1.0 Gb	72,903	391 Kb	28,910	HiSeq2000	Zhang <i>et al.</i> 2016b	China, Taiwan, Belgium
<i>Phalaenopsis</i>	KHM190	3.5 Gb	2n = 2x = 38	3.1 Gb	149,151	101 Kb	41,153	HiSeq2000	Huang <i>et al.</i> 2016	Taiwan
<i>Petunia axillaris</i>	N	1.4 Gb	2n = 2x = 14	1.3 Gb	83,639	1236 Kb	32,928	HiSeq2500, PacBio	Bombarley <i>et al.</i> 2016	USA, Switzerland, China, Netherlands, Germany, Italy, New Zealand, Spain, UK, France
<i>Petunia inflata</i>	S6	1.4 Gb	2n = 2x = 14	1.3 Gb	136,283	884 Kb	36,697			
<i>Ipomoea nil</i>	Tokyo Kokei Standard	750 Mb	2n = 2x = 30	734.8 Mb	3,416	2880 Kb	42,783	PacBio, HiSeq2500	Hoshino <i>et al.</i> 2016	Japan
<i>Hibiscus syriacus</i>	Serial number 520	1.9 Gb	2n = 4x = 80	1.8 Gb	77,492	140 Kb	87,603	HiSeq2000	Kim <i>et al.</i> 2017	Korea
<i>Helianthus annuus</i>	XRQ	3.6 Gb	2n = 2x = 34	2.9 Gb	12,318	524 Kb	52,232	PacBio RS II	Badouin <i>et al.</i> 2017	Canada, France, USA, Israel, UK

of its many colorful flowers in a single inflorescence (Cai *et al.* 2015b). The *P. equestris* genome is relatively small (1.2 Gb) compared with the genomes of other species in the same genus or even other genera (Leitch *et al.* 2009). *P. equestris* is not only a model plant of orchids, but also is the first plant with crassulacean acid metabolism (CAM) to have its genome sequenced (Cai *et al.* 2015b). Many orchids use the CAM pathway for photosynthesis rather than the C3 pathway, and CAM is considered to be an adaptation to arid environments. The CAM pathway evolved convergently in many different plant lineages, and it has been estimated that components of the CAM pathway are encoded in the genomes of about 6% of all flowering plant species (Silvera *et al.* 2010). The total *P. equestris* genome assembly amounted to 1.1 Gb, which is 93% of the estimated total genome size. A total of 29,431 protein-coding genes were predicted. Phylogenetic analysis using the number of synonymous substitutions per synonymous site (K_S) found evidence for an orchid-specific paleopolyploidy event that preceded the radiation of most orchid clades. A phylogenetic analysis of MADS-box genes known for their roles in flower development detected expanded and diversified families of MADS-box genes that might contribute to the highly specialized morphology of orchid flowers. Huang *et al.* (2016) reported a 3.1-Gb draft genome assembly of a *Phalaenopsis* cultivar and identified 41,153 protein-coding genes, which is 11,722 more than the number estimated from the *P. equestris* genome (Cai *et al.* 2015b). This difference in gene numbers may be due to the different approaches used and the different genetic backgrounds of the *Phalaenopsis* cultivar ‘KHM190’ and *P. equestris* inbred line.

Dendrobium is one of the largest genus of Orchidaceae and contains approximately 1200 species, which are distributed widely in the Asia-Pacific region and show great diversification of morphological characters (Takamiya *et al.* 2011). *Dendrobium officinale* Kimura & Migo is a traditional Chinese orchid herb that has both ornamental value and a broad range of therapeutic effects, and its genome has been sequenced (Yan *et al.* 2015). The assembled genome size was 1.4 Gb, which is bigger than the estimated genome size of 1.3 Gb. Zhang *et al.* (2016b) reported a high-quality genome assembly of *Dendrobium catenatum* using a combination of second and third generation sequencers. The assembled genome was 1.0 Gb, which was similar to the estimated genome size of 1.1 Gb and the scaffold N50 was 391 kb, which is longer than the 76 Kb for the *D. officinale* genome. In the *K*-mer analyses, multiple peaks were detected in the *D. officinale* genome, which indicated the genome assembly contained the artificial hybrid sequence and was of low quality, whereas in *D. catenatum* the peaks were normalized. A total of 789 Mb of repetitive elements occupying more than 78.1% of the *D. catenatum* genome were annotated, suggesting that *Dendrobium* genomes may contain high numbers of repeat sequences. Recently, a high-quality genome assembly of *Apostasia*, a group of primitive orchids,

was obtained, together with improved *Phalaenopsis* and *Dendrobium* genomes using PacBio and 10X Genomics linked-reads (Zhang *et al.* 2017). The N50 lengths of the scaffolds in these two genomes improved significantly: for *P. equestris*, the scaffold N50 increased from 359 Kb to 1.2 Mb, and for *D. catenatum* the scaffold N50 increased from 391 Kb to 1.1 Mb.

Primula

The PacBio third generation sequencer was first used in ornamentals to sequence the genome of *Primula veris* (Nowak *et al.* 2015). The assembled genome had a scaffold N50 of 164 kb, covering 64.7% of the estimated 479 Mb genome. The *P. veris* genome was the first genome to be assembled from a heterostylous species. Within the *Primula* genus, *P. vulgaris* (primrose) and *P. elatior* (oxlip) have emerged as model plants for evolutionary, ecological, and conservation studies, and as the paradigm for exploring the genetic control of distyly.

Petunia

Petunia hybrida is a popular bedding plant that has a long history as a genetic model system. The genus *Petunia* is in the Solanaceae family and is native to South America. *Petunia* forms a separate and early branching clade rather than the typical $x = 12$ found for most Solanaceae crown-group species, including important crops such as tomato, potato, tobacco, pepper, and eggplant (Särkinen *et al.* 2013). Bombarely *et al.* (2016) reported the whole-genome sequencing and assembly of inbred derivatives of two wild parents, *P. axillaris* and *P. inflata*. The assemblies covered 1.26 Gb (90.2%) and 1.29 Gb (91.3%) of their diploid genomes (1.4 Gb; $2n = 14$), and contained 32,928 and 36,697 protein-coding genes, respectively. The estimated size of both genomes was 1.4 Gb, using *K*-mer size. The genomes revealed that the *Petunia* lineage had experienced at least two rounds of hexaploidization: the older gamma event, which is shared with most eudicots, and a more recent Solanaceae event that is shared with tomato and other Solanaceae species.

Ipomoea nil

I. nil is a traditional floricultural plant in Japan. It belongs to the family Convolvulaceae, and *Ipomoea* is the largest genus in the family. Hoshino *et al.* (2016) sequenced the genome of *I. nil*, which has been used as a model plant to study the genetic basis of floricultural traits using mainly third-generation PacBio sequencing platforms. The assembled genome of *I. nil* had a scaffold N50 of 2.88 Mb (contig N50 of 1.87 Mb), and covered 98% of the estimated 750-Mb genome. The average contig N50 length for all published genomes is 50 kb; thus, the average contig N50 length of the *I. nil* genome assembly is much longer (Hoshino *et al.* 2016). A major obstacle in utilizing whole-genome shotgun assemblies for important research applications such as gene isolation or comparative genomics has been the lack of

chromosomal positioning and contextualization of short sequence contigs (Mascher and Stein 2014). The ultimate goal of the anchoring process is to identify pseudomolecules, and to accurately order single sequence scaffolds on each chromosome with as few intervening gaps as possible (Tang *et al.* 2015). The scaffolds, which cover 91.4% of the assembly, have been anchored to 15 pseudo-chromosomes in *I. nil*.

Hibiscus

Hibiscus syriacus is a commonly grown ornamental species with attractive white, pink, red, lavender, and purple flowers displayed over a long blooming period, though individual flowers last only a day. The genome of tetraploid *H. syriacus* ($2n = 4x = 80$) was assembled by Kim *et al.* (2017). The assembled genome had a scaffold N50 of 140 kb, covering 92% of the estimated 1.9-Gb genome. Comparative genomic analysis of Malvaceae species, including *H. syriacus*, *Theobroma cacao* (diploid), and *Gossypium raimondii* (diploid cotton), provided clues about the recent polyploidization in *H. syriacus* by whole-genome duplications and unequal regulation of gene dosage by subsequent paleopolyploidy. This was the first report on whole genome sequence analysis of polyploidy woody plants and the effects of whole-genome duplication on their unique phenotypes (Kim *et al.* 2017).

Sunflower

Recently, the genome sequence of domesticated sunflower, *Helianthus annuus* L., which is a global oil crop ornamental, was reported (Badouin *et al.* 2017). *Helianthus* belongs to the Asteraceae family, which includes chrysanthemum. The sunflower genome is 3.6 Gb, but despite the high interest, assembling the genome has been extremely difficult because it consists mainly of long and highly similar repeats. This complexity has challenged leading-edge assembly protocols for close to a decade. To finally overcome this challenge, Badouin *et al.* (2017) generated $102\times$ coverage of the genome using a PacBio RS II platform. The assembled genome was equivalent to 2.9 Gb (scaffold N50 of 524 kb), which is 80% of the estimated genome size. The sunflower genome was predicted to encode 52,232 protein-coding genes. Using four high-density genetic maps and BAC end sequences, the genome sequences, which included 97% of the gene, were anchored to 17 pseudo-chromosomes.

Other major ornamentals

Whole genome sequences of some ornamentals have been reported recently, but the genome sequences of major ornamentals, including polyploids such as chrysanthemum and rose, have not been reported so far. Rose (*Rosa hybrida*) is the most advanced ornamental species in terms of genomic analysis (Yagi 2015). Ploidy levels of rose species range from $2x$ to $8x$, with the majority of wild species being diploid and most cultivars being tetraploid ($2n = 4x = 28$) (Debener and Linde 2009). Diploid *R. chinensis* cv. Old Blush was chosen as a model to overcome the complexity of

the polyploids (Dubois *et al.* 2012). Recently, a shotgun genome sequence of *R. roxburghii* Tratt was revealed, making it the *Rosa* genome to be published (Lu *et al.* 2016).

Chrysanthemum (*Chrysanthemum morifolium*) is also one of the most important ornamental crops worldwide. Most cultivated chrysanthemum varieties are hexaploid ($2n = 6x = 54$), with somatic chromosome numbers ranging from $2n = 47$ to 63 both between and within plants (Anderson 2006). The *C. morifolium* genome was estimated to be approximately 9.4 Gb (<http://data.kew.org/cvalues/>). We chose *C. seticuspe*, a diploid wild species, as a model for chrysanthemum and sequence its genome (estimated 3 Gb). Information from wild diploid species should help to resolve the genetic complexity of the hexaploidy and autopolyploidy chrysanthemum genomes. Among allopolyploid species, genome assemblies are available for octoploid strawberry (Hirakawa *et al.* 2014), tetraploid oilseed rape (Chalhoub *et al.* 2014), and tetraploid upland cotton (Li *et al.* 2015) (Ming and Wai 2015). To my knowledge, no de novo assembly of an autopolyploid plant genome has been reported in which each of the homeologous chromosomes was reconstructed separately (Jiao and Schneeberger 2017). Instead, assemblies of diploid or even haploid genomes have been used to assemble the genomes of autopolyploid species (The Potato Genome Sequencing Consortium 2011). Depending on the level of divergence between homeologous chromosomes, it may be possible to assemble autopolyploid genomes into a 'pseudo-haploid' sequence where polymorphic sites could be annotated in subsequent steps. Alternatively, third-generation technologies could be used to bridge gaps between neighboring polymorphisms, thereby distinguishing homeologous chromosomes. Obviously, obtaining longer reads or DNA molecules will help to improve the reconstruction of individual homeologs (Jiao and Schneeberger 2017).

In addition to recent progress in long-read DNA sequencing, novel technologies have emerged, which promise to improve scaffolding and eventually eliminate the need for genetic or physical mapping. These technologies include the Irys system by BioNano Genomics (www.bionanogenomics.com), the Hi-C protocol called Chicago, which has been provided as a service by Dovetail Genomics since 2014 (www.dovetailgenomics.com), and 10X Genomics, in 2015, integrated their proprietary GemCode technology into their latest system called Chromium (www.10xgenomics.com).

High-density linkage maps

High-density reference genetic linkage maps constructed with genome-wide molecular markers are important for many genetic and breeding applications, including marker-assisted selection, mapping of QTLs, identifying DNA markers for finger printing, and map-based gene cloning.

SSR markers have many advantages over other molecular markers, such as genetic co-dominance. They are multi-allelic, relatively abundant, widely dispersed across the

genome, and easily and automatically scored (Powell *et al.* 1996). Over the past few years, SSR markers have been used in genetic diversity analysis, parentage assessment, species identification, and mapping genetic linkage (Feng *et al.* 2016). The rapid evolution of NGS technologies has made it easier to develop EST-SSR or genomic SSR markers in ornamental plants. In minor ornamentals such as hydrangea (Waki *et al.* 2017), gentian (Nakatsuka *et al.* 2012), and colored calla lily (*Zantedeschia rehmannii* Engl.) (Wei *et al.* 2016), SSR markers have been developed by NGS. However, linkage maps containing a large number of SSR markers that can be used in different cultivars or across species are scarce, with the exception of the SSR markers for carnation and rose. This implies that typing of SSR markers is still a costly and time-consuming process for ornamentals.

Among the ornamentals, the most number of linkage maps have been produced for rose (Yagi 2015). Until now, linkage maps for *Rosa* species have been produced for several diploid (Crespel *et al.* 2002, Debener and Mattiesch 1999, Dugo *et al.* 2005, Hosseini Moghaddam *et al.* 2012, Spiller *et al.* 2010, Yan *et al.* 2005) and tetraploid populations (Hibrand-Saint Oyant *et al.* 2008, Kawamura *et al.* 2011, Rajapakse *et al.* 2001, Zhang *et al.* 2006) using SSR, RAPD, AFLP, RFLP, and SCAR markers. NGS has made it possible to rapidly and cheaply identify large numbers of single nucleotide polymorphisms (SNPs) in genomes. SNPs are useful molecular markers because they are an abundant and frequent type of genetic variation (Ganal *et al.* 2009). SNPs were identified using de novo assembled RNA-Seq data generated from tetraploid cut and garden roses and from diploid *Rosa multiflora* (Koning-Boucoiran *et al.* 2015), and a total of 68,893 rose SNPs was included on the WagRhSNP Axiom array that they developed. As far as I know, this is the only SNP array that is currently available for ornamentals. In tetraploids such as rose, five alternative genotypes (aaaa, baaa, bbaa, bbba, and bbbb; nulliplex to quadruplex) can be present in a single locus. SNP markers are biallelic, so can only distinguish two alleles. To estimate the allele dosage, several methodologies and softwares have been proposed for SNP calling in polyploids (Clevenger *et al.* 2015). Voorrips *et al.* (2011) developed the R package fitTetra, which can efficiently assign five possible genotypes from bi-allelic marker data. TetraploidMap software for mapping tetraploids was publicly available (Hackett *et al.* 2007), but was restricted to markers that segregate as simplex \times nulliplex or duplex \times nulliplex. Preedy and Hackett (2016) proposed a robust method for rapid construction of high-density linkage maps for tetraploid species with large datasets based on a multidimensional scaling approach. By combining such SNP array systems and mapping methodologies, Vukosavljev *et al.* (2016) developed the first high-density linkage map for tetraploid rose. The largest linkage map had 1929 SNP markers and covered 1765.5 cM with an average marker distance of 0.9 cM. Compared with earlier tetraploid rose maps with average marker distances

from 2.4 cM (homologs integrated) to 5.3 cM (map per homolog), homolog coverage and marker density were much improved. This result shows that SNP array systems can be adopted successfully for tetraploid mapping and need not be confined to diploid populations. These technologies and methodologies that were developed for rose will be useful for mapping polyploids in the other ornamentals. Bourke *et al.* (2017) reported an ultra-high-density map using the 68-K WagRhSNP array. A total of 25,695 SNP markers were assigned to seven integrated consensus map (ICM) groups with the largest gap being 4.3 cM on ICM1. These maps are the highest-density linkage maps for *Rosa* published to date. Bourke *et al.* (2017) also mapped the woodland strawberry genome to the *Rosa* map to try to understand segmental allopolyploidy in polyploid species.

In early NGS analyses, whole-genome resequencing was used for SNP identification and genetic mapping in only a few fully sequenced model organisms (Ganal *et al.* 2009). Consequently, Miller *et al.* (2007) developed a cost-effective method for SNP detection using restriction-site associated DNA (RAD) sequencing. In carnation, we constructed a high-density linkage map based on a combination of SSR and RAD markers developed by double-digest RAD sequencing (Peterson *et al.* 2012) (Yagi *et al.* 2017). A total of 2404 (285 SSR and 2119 RAD) markers were assigned to 15 linkage groups spanning 971.5 cM, with an average marker interval of 0.4 cM. Comparative analysis between a reference linkage map (Yagi *et al.* 2013) and the high-density linkage map demonstrated that the marker positions and alignments showed good collinearity. Marker density (marker/cM) was 2.5—an increase compared with the previous map (0.42; Yagi *et al.* 2013). In *I. nil*, a RAD marker-based genetic linkage map was constructed to anchor the genome sequences in order to eliminate the possibility of mis-assembled chimeric scaffolds and establish pseudomolecules (Hoshino *et al.* 2016). Illumina sequencing using the RAD sequencing procedure was employed to construct high-density genetic linkage maps that yielded 86.1 million reads for the parent samples and 562.2 million reads for the 176 progeny samples, which resulted in 3733 SNP markers. Fifteen linkage maps were constructed using the SNP markers and were assigned to 91.4% of the assembly. Specific-locus amplified fragment sequencing (SLAF-seq) is another method for large-scale de novo SNP discovery and genotyping that was first described by Sun *et al.* (2013). Using SLAF-seq, Cai *et al.* (2015a) constructed an integrated genetic map of tree peony (*Paeonia Sect. Moutan*) that contained 1189 markers on the five linkage groups and spanned 920.7 cM, with an average inter-marker distance of 0.774 cM. Tree peonies (*Paeonia Sect. Moutan*), native to China, have become an international ornamental plant and are widely cultivated in many countries in Asia, America, Europe, as well as in Australia (Cheng 2007).

Advanced high-throughput genotyping technology has enabled the development of large-scale sequence markers and high-density linkage maps in ornamentals, which will

serve as an important foundation not only for QTL mapping, map-based gene cloning, and molecular breeding, but also for orienting sequence scaffolds and assembling genome sequences.

Databases for genomic analysis in ornamentals

Now, with the low cost of sequencing technologies, it is expected that a large amount of sequence information for ornamental plants will be released in the near future. Besides, the transcriptome data and DNA marker information, polymorphic information between cultivars or lines will also become available. It is important that all these data are accessible in an easy-to-use form. I have listed the web pages that may be useful for researchers and breeders of ornamentals in **Table 2**.

Databases that contain genome sequence information for carnation and *I. nil* are as follows. Carnation DB (<http://carnation.kazusa.or.jp/>) contains the assembled scaffold sequences (Yagi *et al.* 2014a), and annotated known and novel repetitive sequences, genes for non-coding RNAs (tRNA, rRNA, snoRNA, and miRNA), and potential protein-encoding genes. The *Ipomoea nil* Genome Project website (<http://viewer.shigen.info/asagao/index.php>), released in February 2017, contains the Asagao assembly of high-quality genome sequences that were released in 2016 (Hoshino *et al.* 2016) and allows them to be viewed in a browser. Blast searches and keyword searches by Gene Symbol, Gene/Transcript ID, GO ID, Pfam ID, and Description can be conducted, and sequences and annotations can be downloaded from the website.

The Rosaceae family includes a large number of important crops such as pear, apple, peach, and cherry in addition to rose. The Genome Database for Rosaceae (GDR, <https://www.rosaceae.org/>) was established in 2003 to integrate publicly available genetic and genomic data and to provide genome analysis tools for the worldwide Rosaceae genomics research community (Jung *et al.* 2004, 2008). GDR is a comprehensive online database resource of curated genomic, genetic, and breeding data, and analysis tools. Initially, genomic data consisted mostly of ESTs, which were collected, assembled into genus-specific unigene sets, and func-

tionally annotated, where possible, by association with known protein homologs. Genetic data such as linkage maps and genetic markers have also been collected and integrated into the GDR (Jung *et al.* 2014). For roses, the GDR contains the following data: gene sequences and annotations, genetic maps, markers, publications, SNP arrays, trait loci, and transcripts.

The genus *Petunia* is a member of the Solanaceae family, and it forms a separate and early branching clade within the family (Bombarely *et al.* 2016). The Sol Genomics Network (SGN, <http://solgenomics.net>) is a web portal with genomic and phenotypic data, and analysis tools for the Solanaceae family and closely related species (Fernandez-Pozo *et al.* 2015). SGN hosts whole genome data for an increasing number of Solanaceae family members including tomato, potato, pepper, eggplant, tobacco, and *Nicotiana benthamiana*. The database also stores loci and phenotype data that researchers can upload and edit with user-friendly web interfaces. The draft genome sequences of two *Petunia* species (*P. axillaris* and *P. inflata*) (Bombarely *et al.* 2016) are available on the SGN website.

The Orchidstra database has been under active development since 2013 (Su *et al.* 2013) and Orchidstra 2.0 (<http://orchidstra2.abrc.sinica.edu.tw>) was released recently. Orchidstra 2.0 was built with a new database system to accommodate the increasing amount of orchid transcriptome data and to store the annotations of 510,947 protein-coding genes and 161,826 noncoding transcripts, covering 18 orchid species belonging to 12 genera in five subfamilies of Orchidaceae (Chao *et al.* 2017).

The Kazusa DNA Research Institute in Japan developed the Plant Genome DataBase Japan (PGDBj, <http://pgdbj.jp/?ln=en>) to integrate plant genome-related information from other databases and from the literature (Asamizu *et al.* 2014). The PGDBj has three component databases (The Ortholog DB, The Plant Resource DB, The DNA Marker DB) and a cross-search engine that provides a seamless search over the contents of all three databases. Regarding ornamentals, information for chrysanthemum, carnation, and *I. nil* are currently available in the PGDBj.

These databases and tools are useful resources for genomics in ornamentals; however, their scope is still limited.

Table 2. Useful web pages for genomics of ornamental plant species

Name	Object	Target/Species	URL
Carnation DB	Database for genome sequence for carnation	Carnation	http://carnation.kazusa.or.jp/
Genome Database for Rosaceae (GDR)	Database for genomics for rose	Rosaceae species	https://www.rosaceae.org/
<i>Ipomoea nil</i> Genome Project	Database for genome sequence of Japanee morning glory	<i>Ipomoea nil</i>	http://viewer.shigen.info/asagao/index.php
Orchidstra 2.0	Database for transcriptome for orchid family	Orchidaceae family	http://orchidstra2.abrc.sinica.edu.tw/orchidstra2/index.php
Plant Genome DataBase Japan (PGDBj)	Integrated genomic database for plants	Chrysanthemum, carnation, <i>Ipomoea nil</i>	http://pgdbj.jp/index.html?ln=en
Sol Genomics Network (SGN)	Database for genome sequence for petunia	Solanaceae species	https://solgenomics.net/

In the near future, not only genome and transcriptome data but also metabolome, proteome, and phenome data, as well as many other experimental resources are likely to be released through a number of databases. The integration of these omics datasets in web databases will provide researchers with a wealth of desirable information, but it is also important for such data and registered information to be updated regularly.

Future perspectives

In this review, I focused on recent progress in whole genome sequencing, construction of a high-density genetic linkage map, and databases of ornamentals. The genomic analysis tools will help researchers and breeders develop new cultivars with desirable traits by marker-assisted selection and new genomic-based strategies. The availability of a huge number of markers and high-throughput genotyping using NGS or SNP arrays will enable genome-wide association analysis in addition to conventional QTL analysis. In addition, genomic selection (Meuwissen *et al.* 2001), which is a method to predict the genetic value of selection candidates based on the genomic estimated breeding value predicted from genotypic data obtained from high-density markers positioned throughout the genome, will enable the more accurate selection of complicated traits.

Genome sequences and related high-density linkage maps will also help to reveal the genes responsible for important traits. Recently, the application of gene editing technology using the CRISPR/Cas9 system in ornamentals has been reported (Kishi-Kaboshi *et al.* 2017, Zhang *et al.* 2016a). It is anticipated that CRISPR/Cas9 technology will promote gene functional analysis and genetic improvement in ornamental plants. The integration of genomics and gene editing technology will be a powerful tool for constructing agronomically designed ornamental plants.

Literature Cited

- Anderson, N.O. (2006) *Chrysanthemum. Dendranthema × grandiflora Tzvelv*. In: Anderson, N.O. (ed.) Flower Breeding and Genetics. Springer, Dordrecht, pp. 389–437.
- Asamizu, E., H. Ichihara, A. Nakaya, Y. Nakamura, H. Hirakawa, T. Ishii, T. Tamura, K. Fukami-Kobayashi, Y. Nakajima and S. Tabata (2014) Plant Genome DataBase Japan (PGDBJ): a portal website for the integration of plant genome-related databases. *Plant Cell Physiol.* 55: e8.
- Badouin, H., J. Gouzy, C.J. Grassa, F. Murat, S.E. Staton, L. Cottret, C. Lelandais-Brière, G.L. Owens, S. Carrère, B. Mayjonade *et al.* (2017) The sunflower genome provides insights into oil metabolism, flowering and Asterid evolution. *Nature* 546: 148–152.
- Bombarely, A., M. Moser, A. Amrad, M. Bao, L. Bapaume, C.S. Barry, M. Bliiek, M.R. Boersma, L. Borghi, R. Bruggmann *et al.* (2016) Insight into the evolution of the Solanaceae from the parental genomes of *Petunia hybrida*. *Nat. Plants* 2: 16074.
- Bourke, P.M., P. Arens, R.E. Voorrips, G.D. Esselink, C.F. Koning-Boucoiran, W.P. Van't Westende, T. Santos Leonardo, P. Wissink, C. Zheng, G. van Geest *et al.* (2017) Partial preferential chromosome pairing is genotype dependent in tetraploid rose. *Plant J.* 90: 330–343.
- Cai, C., F.-Y. Cheng, J. Wu, Y. Zhong and G. Liu (2015a) The first high-density genetic map construction in tree peony (*Paeonia Sect. Moutan*) using genotyping by specific-locus amplified fragment sequencing. *PLoS ONE* 10: e0128584.
- Cai, J., X. Liu, K. Vanneste, S. Proost, W.-C. Tsai, K.-W. Liu, L.-J. Chen, Y. He, Q. Xu, C. Bian *et al.* (2015b) The genome sequence of the orchid *Phalaenopsis equestris*. *Nat. Genet.* 47: 65–72.
- Chalhoub, B., F. Denoeud, S. Liu, I.A. Parkin, H. Tang, X. Wang, J. Chiquet, H. Belcram, C. Tong, B. Samans *et al.* (2014) Early allopolyploid evolution in the post-neolithic *Brassica napus* oilseed genome. *Science* 345: 950–953.
- Chao, Y.-T., S.-H. Yen, J.-H. Yeh, W.-C. Chen and M.-C. Shih (2017) Orchidstra 2.0—A transcriptomics resource for the orchid family. *Plant Cell Physiol.* 58: e9.
- Cheng, F. (2007) Advances in the breeding of tree peonies and a cultivar system for the cultivar group. *Int. J. Plant Breed.* 1: 89–104.
- Clevenger, J., C. Chavarro, S.A. Pearl, P. Ozias-Akins and S.A. Jackson (2015) Single nucleotide polymorphism identification in polyploids: a review, example, and recommendations. *Mol. Plant* 8: 831–846.
- Crespel, L., M. Chirollet, C.E. Durel, D. Zhang, J. Meynet and S. Gudin (2002) Mapping of qualitative and quantitative phenotypic traits in *Rosa* using AFLP markers. *Theor. Appl. Genet.* 105: 1207–1214.
- Debener, T. and L. Mattiesch (1999) Construction of a genetic linkage map for roses using RAPD and AFLP markers. *Theor. Appl. Genet.* 99: 891–899.
- Debener, T. and M. Linde (2009) Exploring complex ornamental genomes: The rose as a model plant. *Crit. Rev. Plant Sci.* 28: 267–280.
- Dubois, A., S. Carrere, O. Raymond, B. Pouvreau, L. Cottret, A. Rocchia, J. Onesto, S. Sakr, R. Atanassova, S. Baudino *et al.* (2012) Transcriptome database resource and gene expression atlas for the rose. *BMC Genomics* 13: 638.
- Dugo, M.L., Z. Satovic, T. Millan, J.I. Cubero, D. Rubiales, A. Cabrera and A.M. Torres (2005) Genetic mapping of QTLs controlling horticultural traits in diploid roses. *Theor. Appl. Genet.* 111: 511–520.
- Feng, S., R. He, J. Lu, M. Jiang, X. Shen, Y. Jiang, Z. Wang and H. Wang (2016) Development of SSR markers and assessment of genetic diversity in medicinal *Chrysanthemum morifolium* cultivars. *Front. Genet.* 7: 113.
- Fernandez-Pozo, N., N. Menda, J.D. Edwards, S. Saha, I.Y. Teclé, S.R. Strickler, A. Bombarely, T. Fisher-York, A. Pujar, H. Foerster *et al.* (2015) The Sol Genomics Network (SGN)—from genotype to phenotype to breeding. *Nucleic Acids Res.* 43: D1036–D1041.
- Ganal, M.W., T. Altmann and M.S. Röder (2009) SNP identification in crop plants. *Curr. Opin. Plant Biol.* 12: 211–217.
- Hackett, C.A., I. Milne, J.E. Bradshaw and Z. Luo (2007) TetraploidMap for Windows: linkage map construction and QTL mapping in auto-tetraploid species. *J. Hered.* 98: 727–729.
- Han, Y., K. Gasic, B. Marron, J.E. Beever and S.S. Korban (2007) A BAC-based physical map of the apple genome. *Genomics* 89: 630–637.
- Hibrand-Saint-Oyant, L., L. Crespel, S. Rajapakse, L. Zhang and F. Foucher (2008) Genetic linkage maps of rose constructed with new microsatellite markers and locating QTL controlling flowering traits. *Tree Genet. Genomes* 4: 11.
- Hirakawa, H., K. Shirasawa, S. Kosugi, K. Tashiro, S. Nakayama, M. Yamada, M. Kohara, A. Watanabe, Y. Kishida, T. Fujishiro *et al.*

- (2014) Dissection of the octoploid strawberry genome by deep sequencing of the genomes of *Fragaria* species. *DNA Res.* 21: 169–181.
- Hoshino, A., V. Jayakumar, E. Nitasaka, A. Toyoda, H. Noguchi, T. Itoh, T. Shin-I, Y. Minakuchi, Y. Koda, A.J. Nagano *et al.* (2016) Genome sequence and analysis of the Japanese morning glory *Ipomoea nil*. *Nat. Commun.* 7: 13295.
- Hosseini Moghaddam, H., L. Leus, J. De Riek, J. Van Huylbroeck and E. Van Bockstaele (2012) Construction of a genetic linkage map with SSR, AFLP and morphological markers to locate QTLs controlling pathotype-specific powdery mildew resistance in diploid roses. *Euphytica* 184: 413–427.
- Hsiao, Y.-Y., Y.-W. Chen, S.-C. Huang, Z.-J. Pan, C.-H. Fu, W.-H. Chen, W.-C. Tsai and H.-H. Chen (2011) Gene discovery using next-generation pyrosequencing to develop ESTs for *Phalaenopsis* orchids. *BMC Genomics* 12: 360.
- Huang, J.-Z., C.-P. Lin, T.-C. Cheng, Y.-W. Huang, Y.-J. Tsai, S.-Y. Cheng, Y.-W. Chen, C.-P. Lee, W.-C. Chung, B.C.-H. Chang *et al.* (2016) The genome and transcriptome of *Phalaenopsis* yield insights into floral organ development and flowering regulation. *PeerJ* 4: e2017.
- Jiao, W.B. and K. Schneeberger (2017) The impact of third generation genomic technologies on plant genome assembly. *Curr. Opin. Plant Biol.* 36: 64–70.
- Jung, S., C. Jesudurai, M. Staton, Z. Du, S. Ficklin, I. Cho, A. Abbott, J. Tomkins and D. Main (2004) GDR (Genome Database for Rosaceae): integrated web resources for Rosaceae genomics and genetics research. *BMC Bioinformatics* 5: 130.
- Jung, S., M.E. Staton, T. Lee, A. Blenda, R. Svancara, A.G. Abbott and D. Main (2008) GDR (Genome Database for Rosaceae): integrated web-database for Rosaceae genomics and genetics data. *Nucleic Acids Res.* 36: D1034–D1040.
- Jung, S., S.P. Ficklin, T. Lee, C.-H. Cheng, A. Blenda, P. Zheng, J. Yu, A. Bombarely, I. Cho, S. Ru *et al.* (2014) The Genome Database for Rosaceae (GDR): year 10 update. *Nucleic Acids Res.* 42: D1237–D1244.
- Kawamura, K., L. Hibrand-Saint Oyant, L. Crespel, T. Thouroude, D. Lalanne and F. Foucher (2011) Quantitative trait loci for flowering time and inflorescence architecture in rose. *Theor. Appl. Genet.* 122: 661–675.
- Kellner, F., J. Kim, B.J. Clavijo, J.P. Hamilton, K.L. Childs, B. Vaillancourt, J. Cepela, M. Habermann, B. Steuernagel, L. Clissold *et al.* (2015) Genome-guided investigation of plant natural product biosynthesis. *Plant J.* 82: 680–692.
- Kim, Y.-M., S. Kim, N. Koo, A.-Y. Shin, S.-I. Yeom, E. Seo, S.-J. Park, W.-H. Kang, M.-S. Kim, J. Park *et al.* (2017) Genome analysis of *Hibiscus syriacus* provides insights of polyploidization and indeterminate flowering in woody plants. *DNA Res.* 24: 71–80.
- Kishi-Kaboshi, M., R. Aida and K. Sasaki (2017) Generation of gene-edited *Chrysanthemum morifolium* using multicopy transgenes as targets and markers. *Plant Cell Physiol.* 58: 216–226.
- Koning-Boucoiran, C.F., G.D. Esselink, M. Vukosavljev, W.P. van't Westende, V.W. Gitonga, F.A. Krens, R.E. Voorrips, W.E. van de Weg, D. Schulz, T. Debener *et al.* (2015) Using RNA-Seq to assemble a rose transcriptome with more than 13,000 full-length expressed genes and to develop the WagRhSNP 68k Axiom SNP array for rose (*Rosa L.*). *Front. Plant Sci.* 6: 249.
- Leitch, I.J., I. Kahandawala, J. Suda, L. Hanson, M.J. Ingrouille, M.W. Chase and M.F. Fay (2009) Genome size diversity in orchids: consequences and evolution. *Ann. Bot.* 104: 469–481.
- Li, F., G. Fan, C. Lu, G. Xiao, C. Zou, R.J. Kohel, Z. Ma, H. Shang, X. Ma, J. Wu *et al.* (2015) Genome sequence of cultivated upland cotton (*Gossypium hirsutum* TM-1) provides insights into genome evolution. *Nat. Biotechnol.* 33: 524–530.
- Lu, M., H. An and L. Li (2016) Genome survey sequencing for the characterization of the genetic background of *Rosa roxburghii* Tratt and leaf ascorbate metabolism genes. *PLoS ONE* 11: e0147530.
- Mascher, M. and N. Stein (2014) Genetic anchoring of whole-genome shotgun assemblies. *Front. Genet.* 5: 208.
- Meuwissen, T.H., B.J. Hayes and M.E. Goddard (2001) Prediction of total genetic value using genome-wide dense marker maps. *Genetics* 157: 1819–1829.
- Michael, T.P. and R. VanBuren (2015) Progress, challenges and the future of crop genomes. *Curr. Opin. Plant Biol.* 24: 71–81.
- Miller, M.R., J.P. Dunham, A. Amores, W.A. Cresko and E.A. Johnson (2007) Rapid and cost-effective polymorphism identification and genotyping using restriction site associated DNA (RAD) markers. *Genome Res.* 17: 240–248.
- Ming, R. and C. Man Wai (2015) Assembling allopolyploid genomes: no longer formidable. *Genome Biol.* 16: 27.
- Nakatsuka, T., E. Yamada, M. Saito, T. Hikage, Y. Ushiku and M. Nishihara (2012) Construction of the first genetic linkage map of Japanese gentian (*Gentianaceae*). *BMC Genomics* 13: 672.
- Nowak, M.D., G. Russo, R. Schlapbach, C.N. Huu, M. Lenhard and E. Conti (2015) The draft genome of *Primula veris* yields insights into the molecular basis of heterostyly. *Genome Biol.* 16: 12.
- Onozaki, T., N. Tanikawa, M. Taneya, K. Kudo, T. Funayama, H. Ikeda and M. Shibata (2004) A RAPD-derived STS marker is linked to a bacterial wilt (*Burkholderia caryophylli*) resistance gene in carnation. *Euphytica* 138: 255–262.
- Onozaki, T., T. Yoshinari, T. Yoshimura, M. Yagi, S. Yoshioka, M. Taneya and M. Shibata (2006) DNA markers linked to a recessive gene controlling single flower type derived from a wild species, *Dianthus capitatus* ssp. *andrzejowskianus*. *Hort. Res. (Japan)* 5: 363–367.
- Peterson, B.K., J.N. Weber, E.H. Kay, H.S. Fisher and H.E. Hoekstra (2012) Double digest RADseq: an inexpensive method for de novo SNP discovery and genotyping in model and non-model species. *PLoS ONE* 7: e37135.
- Powell, W., M. Morgante, C. Andre, M. Hanafey, J. Vogel, S. Tingey and Antoni Rafalski (1996) The comparison of RFLP, RAPD, AFLP and SSR (microsatellite) markers for germplasm analysis. *Mol. Breed.* 2: 225–238.
- Preedy, K.F. and C.A. Hackett (2016) A rapid marker ordering approach for high-density genetic linkage maps in experimental auto-tetraploid populations using multidimensional scaling. *Theor. Appl. Genet.* 129: 2117–2132.
- Rajapakse, S., D.H. Byrne, L. Zhang, N. Anderson, K. Arumuganathan and R.E. Ballard (2001) Two genetic linkage maps of tetraploid roses. *Theor. Appl. Genet.* 103: 575–583.
- Särkinen, T., L. Bohs, R.G. Olmstead and S. Knapp (2013) A phylogenetic framework for evolutionary study of the nightshades (*Solanaceae*) a dated 1000-tip tree. *BMC Evol. Biol.* 13: 214.
- Shibata, M. (2008) Importance of genetic transformation in ornamental plant breeding. *Plant Biotechnol.* 25: 3–8.
- Silvera, K., K.M. Neubig, W.M. Whitten, N.H. Williams, K. Winter and J.C. Cushman (2010) Evolution along the crassulacean acid metabolism continuum. *Funct. Plant Biol.* 37: 995–1010.
- Spiller, M., R.G. Berger and T. Debener (2010) Genetic dissection of scent metabolic profiles in diploid rose populations. *Theor. Appl. Genet.* 120: 1461–1471.
- Su, C., Y.-T. Chao, S.-H. Yen, C.-Y. Chen, W.-C. Chen, Y.-C.A. Chang

- and M.-C. Shih (2013) Orchidstra: An integrated orchid functional genomics database. *Plant Cell Physiol.* 54: e11.
- Sun, X., D. Liu, X. Zhang, W. Li, H. Liu, W. Hong, C. Jiang, N. Guan, C. Ma, H. Zeng *et al.* (2013) SLAF-seq: an efficient method of large-scale de novo SNP discovery and genotyping using high-throughput sequencing. *PLoS ONE* 8: e58700.
- Takamiya, T., P. Wongsawad, N. Tajima, N. Shioda, J.F. Lu, C.L. Wen, J.B. Wu, T. Handa, H. Iijima, S. Kitanaka *et al.* (2011) Identification of *Dendrobium* species used for herbal medicines based on ribosomal DNA internal transcribed spacer sequence. *Biol. Pharm. Bull.* 34: 779–782.
- Tanase, K., C. Nishitani, H. Hirakawa, S. Isobe, S. Tabata, A. Ohmiya and T. Onozaki (2012) Transcriptome analysis of carnation (*Dianthus caryophyllus* L.) based on next-generation sequencing technology. *BMC Genomics* 13: 292.
- Tang, H., X. Zhang, C. Miao, J. Zhang, R. Ming, J.C. Schnable, P.S. Schnable, E. Lyons and J. Lu (2015) ALLMAPS: robust scaffold ordering based on multiple maps. *Genome Biol.* 16: 3.
- The Arabidopsis Genome Initiative (2000) Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana*. *Nature* 408: 796–815.
- The Potato Genome Sequencing Consortium (2011) Genome sequence and analysis of the tuber crop potato. *Nature* 475: 189–195.
- Voorrips, R.E., G. Gort and B. Vosman (2011) Genotype calling in tetraploid species from bi-allelic marker data using mixture models. *BMC Bioinformatics* 12: 172.
- Vukosavljev, M., P. Arens, R.E. Voorrips, W.P. van't Westende, G.D. Esselink, P.M. Bourke, P. Cox, W.E. van de Weg, R.G. Visser, C. Maliepaard *et al.* (2016) High-density SNP-based genetic maps for the parents of an outcrossed and a selfed tetraploid garden rose cross, inferred from admixed progeny using the 68k rose SNP array. *Hortic. Res.* 3: 16052.
- Waki, T., M. Kodama, M. Akutsu, K. Namai, M. Iigo, T. Kurokura, T. Yamamoto, K. Nashima, M. Nakayama and M. Yagi (2017) Development of DNA markers linked to double-flower and hortensia traits in *Hydrangea macrophylla* (Thunb.) Ser. *Hort. J.* doi: 10.2503/hortj.OKD-096.
- Wang, H., J. Jiang, S. Chen, X. Qi, H. Peng, P. Li, A. Song, Z. Guan, W. Fang, Y. Liao *et al.* (2013) Next-generation sequencing of the *Chrysanthemum nankingense* (Asteraceae) transcriptome permits large-scale unigene assembly and SSR marker discovery. *PLoS ONE* 8: e62293.
- Wei, Z., Z. Sun, B. Cui, Q. Zhang, M. Xiong, X. Wang and D. Zhou (2016) Transcriptome analysis of colored calla lily (*Zantedeschia rehmannii* Engl.) by Illumina sequencing: de novo assembly, annotation and EST-SSR marker development. *PeerJ* 4: e2378.
- Yagi, M., T. Onozaki, M. Taneya, H. Watanabe, T. Yoshimura, T. Yoshinari, Y. Ochiai and M. Shibata (2006) Construction of a genetic linkage map for the carnation by using RAPD and SSR markers and mapping quantitative trait loci (QTL) for resistance to bacterial wilt caused by *Burkholderia caryophylli*. *J. Japan Soc. Hort. Sci.* 75: 166–172.
- Yagi, M., T. Onozaki, H. Ikeda, N. Tanikawa, M. Shibata, T. Yamaguchi, K. Tanase, K. Sumitomo and M. Amano (2010) Breeding process and characteristics of carnation 'Karen Rouge' with resistance to bacterial wilt. *Bull. Natl. Inst. Flor. Sci.* 10: 1–10.
- Yagi, M., T. Kimura, T. Yamamoto, S. Isobe, S. Tabata and T. Onozaki (2012) QTL analysis for resistance to bacterial wilt (*Burkholderia caryophylli*) in carnation (*Dianthus caryophyllus*) using an SSR-based genetic linkage map. *Mol. Breed.* 30: 495–509.
- Yagi, M. (2013) Application of DNA markers for breeding carnations resistant to bacterial wilt. *JARQ* 47: 29–35.
- Yagi, M., T. Yamamoto, S. Isobe, H. Hirakawa, S. Tabata, K. Tanase, H. Yamaguchi and T. Onozaki (2013) Construction of a reference genetic linkage map for carnation (*Dianthus caryophyllus* L.). *BMC Genomics* 14: 734.
- Yagi, M., S. Kosugi, H. Hirakawa, A. Ohmiya, K. Tanase, T. Harada, K. Kishimoto, M. Nakayama, K. Ichimura, T. Onozaki *et al.* (2014a) Sequence analysis of the genome of carnation (*Dianthus caryophyllus* L.). *DNA Res.* 21: 231–241.
- Yagi, M., T. Yamamoto, S. Isobe, S. Tabata, H. Hirakawa, H. Yamaguchi, K. Tanase and T. Onozaki (2014b) Identification of tightly linked SSR markers for flower type in carnation (*Dianthus caryophyllus* L.). *Euphytica* 198: 175–183.
- Yagi, M. (2015) Recent progress in genomic analysis of ornamental plants, with a focus on carnation. *Hort. J.* 84: 3–13.
- Yagi, M., K. Shirasawa, T. Waki, T. Kume, S. Isobe, K. Tanase and H. Yamaguchi (2017) Construction of an SSR and RAD marker-based genetic linkage map for carnation (*Dianthus caryophyllus* L.). *Plant Mol. Biol. Rep.* 35: 110–117.
- Yan, L., X. Wang, H. Liu, Y. Tian, J. Lian, R. Yang, S. Hao, X. Wang, S. Yang, Q. Li *et al.* (2015) The genome of *Dendrobium officinale* illuminates the biology of the important traditional chinese orchid herb. *Mol. Plant* 8: 922–934.
- Yan, Z., C. Denneboom, A. Hattendorf, O. Dolstra, T. Debener, P. Stam and P.B. Visser (2005) Construction of an integrated map of rose with AFLP, SSR, PK, RGA, RFLP, SCAR and morphological markers. *Theor. Appl. Genet.* 110: 766–777.
- Zhang, B., X. Yang, C. Yang, M. Li and Y. Guo (2016a) Exploiting the CRISPR/Cas9 system for targeted genome mutagenesis in petunia. *Sci. Rep.* 6: 20315.
- Zhang, G.-Q., Q. Xu, C. Bian, W.-C. Tsai, C.-M. Yeh, K.-W. Liu, K. Yoshida, L.-S. Zhang, S.-B. Chang, F. Chen *et al.* (2016b) The *Dendrobium catenatum* Lindl. genome sequence provides insights into polysaccharide synthase, floral development and adaptive evolution. *Sci. Rep.* 12: 19029.
- Zhang, G.-Q., K.-W. Liu, Y.-Y. Hsiao, Z. Li, R. Lohaus, K. Yoshida, S.-C. Niu, S. Fujiwara, Y.-C. Lin, Y.-Q. Zhang *et al.* (2017) The *Apostasia* genome and the evolution of orchids. *Nature* 549: 379–383.
- Zhang, L.H., D.H. Byrne, R.E. Ballard and S. Rajapakse (2006) Microsatellite marker development in rose and its application in tetraploid mapping. *J. Am. Soc. Hort. Sci.* 131: 380–387.