Genomic Tools and Prospects for New Breeding Techniques in Flower Bulb Crops

F.A. Krens\textsuperscript{a}  K. Kamo\textsuperscript{b}
Wageningen UR Plant Breeding  Floral & Nursery Plants Research Unit
PO Box 16, 6700 AA Wageningen  Bldg. 10A Room 126 BARC West, USDA
The Netherlands  Beltsville, MD 20705-2350

\textbf{Keywords:} genome sequencing, promoters, genes, genetic modification, disease resistance, flower traits, marketing, bulbs

\textbf{Abstract}

For many of the new breeding techniques sequence information is of the utmost importance. In addition to current breeding techniques, such as marker-assisted selection (MAS) and genetic modification (GM), new breeding techniques such as zinc finger nucleases, oligonucleotide-mediated mutagenesis, RNAi and cisgenesis are totally dependent on knowing gene or allele sequences. Many plant species have been fully sequenced, but none of them represents an ornamental crop, let alone a member of the monocotyledonous, bulbiferous geophyte order, such as the \textit{Liliales}. For a limited number of these flower bulb species, some EST-libraries have been generated and a similarly low number of promoters and genes have been isolated, characterized, or used for transformation. Most of the functional analyses and applications in genetically modified crops were ectopic, i.e., not in flower bulb crops themselves. Genes and promoters originating from other organisms have been used in GM of flower bulbs aiming for the introduction of disease resistance or the modification of flower traits. In order to enhance consumer acceptance of transgenic GM flower bulbs and to facilitate obtaining EU approval for cultivation and market introduction, the marker-free technology is being tested.

\textbf{INTRODUCTION}

New breeding techniques encompass not only molecular marker-assisted selection and genetic modification but also among others zinc-finger nucleases-mediated genome editing, oligonucleotide mutagenesis, RNAi technology and the concept of cisgenesis (Lusser et al., 2011, 2012). Sequence information on genes or alleles, and expression regulating sequences such as promoters and terminators is a prerequisite for all of them. Sequence information on simple sequence repeats (SSRs or microsatellites), on expressed parts of the genome such as expressed sequence tags (ESTs) or preferably on the entire genome greatly facilitates the identification or development of molecular markers, that can subsequently be applied in marker-assisted selection. It also allows the identification and isolation of genes and even alleles and their promoters and terminators. Gene functions can be deduced from homology with known genes present in international databases. In bulbous ornamental crops large sequencing efforts are scarce and sequence information is lacking for most crops or very limited. Here, we present an overview of the state-of-the-art on present day sequence information on flower bulb crops and how this affects the use of genetic modification for crop improvement and the development of cisgenic or marker-free approaches.

\textbf{GENOME SEQUENCING}

Whole genome sequencing has become easier, faster and cheaper over the years with the technological advances that were made in the sequencing hardware. Table 1 shows the list of fully sequenced plant genomes (June 2012) which is rapidly expanding.

\textsuperscript{a} Frans.Krens@wur.nl
\textsuperscript{b} Kathryn.Kamo@ars.usda.gov
Efficient sequencing requires the crop-of-interest to be diploid, homozygous and with a small genome size. This allows not just determining the order of the base pairs, but also the assembly into contigs and linkage groups or chromosomes. The sequencing of a whole genome of a plant is completed when gene functions are annotated and available to the scientific community by entering it into a public database. It is clear from Table 1 that whole genome sequencing of ornamental crops is lacking, with the possible exception of columbine, *Aquilegia formosa*. Flower bulb crops are certainly not present.

One of the main reasons for this are the huge genome sizes of most of the representatives of the Liliaceae, Iridaceae, Asparagaceae, Melanthiaceae, and other flower bulb ornamentals as shown in Table 2. With the number of genes being rather similar among organisms, i.e., approximately between 25,000 and 40,000, the differences in genome size are due to repetitive DNA sequences. Such sequences can consist of simple sequence repeats or microsatellites, gene or domain (specific cluster of genes) duplications, chromosome duplications, retrotransposons, and microRNA (miRNA) family clusters. Ultimately, the entire genome can be duplicated leading to polyploidy. This repetitiveness can lead to serious problems in contig assembly, e.g., after sequencing bacterial artificial chromosome (BAC) libraries or after whole genome shotgun sequencing. When the repetitive sequence is larger than the read obtained from the sequencing equipment, one could start assembling e.g., chromosome 1 and without being aware through the repetitive sequences cross over to another chromosome only to finish ultimately in a third one. Thus, one produces chimeric virtual chromosomes and false genome organisation. A potential solution for this could be to wait for the development of sequencing techniques yielding larger reads or the use of highly saturated, high density molecular maps. However, such maps are not available in bulbs.

As an alternative to sequencing the whole genome, one can sequence the parts that are expressed. cDNA sequencing, looking at Expressed Sequence Tags (ESTs), was done in some bulb crops, e.g., lily (Yu et al., 2003; Okada et al., 2006), iris (Tang et al., 2009) and saffron (D’Agostino et al., 2007). The data were used in functional analysis and in molecular marker development. Other examples of large scale sequencing of the transcriptome in flower bulb crops, e.g., for the identification of single nucleotide polymorphisms (SNPs) as molecular markers (as in onion) or for providing information on allele numbers, are not reported in literature, yet.

### PROMOTER SEQUENCES FROM BULBS

#### General

For the generation of genetically modified plants one needs genes or the coding regions of genes and regulatory sequences such as promoters and terminators. In order to drive expression efficiently in the tissue or at the developmental stage that is desired, the promoter needs to be carefully selected. Promoters used in plant expression originate from the T-DNA of *Agrobacterium tumefaciens*, e.g., the mas2 or nos promoters, from plant DNA viruses, e.g., the 35S or 19S promoters from the *Cauliflower mosaic virus* (CaMV), or from plant genes themselves. Table 3 shows the promoters isolated and characterized from bulbs. Eight out of ten are involved in tissue-specific expression in flowers or flower organs, however, also a constitutive promoter was identified and isolated from *Gladiolus* callus (Kamo et al., 1999, 2000). Because the bulbous monocotyledonous crops are notoriously recalcitrant in transformation and forming transgenic plants, the characterization was primarily done in dicotyledonous model crops, such as *Arabidopsis thaliana*, except for two cases. Promoters isolated from bulb species themselves or from related species could be used in a cisgenic or intragenic approach when combined with bulb genes or coding regions (Schouten et al., 2006; Rommens et al., 2007). At present, the promoters from bulbs as presented in Table 3 are not in use.

#### Gladiolus

In *Gladiolus* multiple promoters have been tested for performance in suspension
cells and in leaves. The promoters were placed in front of the gus reporter gene and GUS expression was checked 48 hours after bombardment. In suspension cells, monocot-derived promoters from cereal crops, such as rice and maize, expressed poorly in *Gladiolus* as compared to dicot-targeted promoters (*A. tumefaciens* or CaMV35S derived; Fig. 1). When tested in leaves, of the 13 different promoters (CaMV 35S, duplicated CaMV 35S, rice ACT1, *Arabidopsis* UBQ3, *Arabidopsis* UBQ10, rolD, mannoipine synthase, translation elongation factor 1 subunit α, maize UBII, maize ADH, potato UBI3, potato UBI7, phosphoenolpyruvate carboxylase) only the CaMV35S promoter showed satisfactory expression levels (Kamo and Blowers, 1999; Kamo et al., 2000). Expression in leaves is important when aiming for inducing resistance against viruses by genetic modification. As the CaMV35S promoter was covered by intellectual property rights (patented), highly efficient, constitutive promoters were looked for within *Gladiolus* itself. Genomic libraries were made and screened using the conserved sequence of the ubiquitin gene known from rice. Three candidate ubiquitin promoters were isolated and analysed (Joung and Kamo, 2006). One of them, the GUBQ1 promoter, showed high levels of expression in leaves, roots and callus similar to that of the CaMV35S promoter. The intron present in the promoter sequence that proved to be important for this expression was more similar to dicots than cereal monocots in AT content. The GUBQ1 promoter also gave high expression in dicots, such as tobacco and *Arabidopsis* (Joung and Kamo, 2006; Kamo, pers. commun.).

**GENES ISOLATED FROM BULBS**

More than 30 genes have been isolated and characterized from various bulbous crops (Table 4). Some are involved with plant defense against viruses, fungi or insects, others in determining nutritional value, e.g., carotenoid biosynthesis. The vast majority of the isolated genes play some role in determining flower color or flower development. A large number of MADS box genes was isolated from lily, narcissus, crocus and alstroemeria (Table 4). The isolation was primarily from cDNA libraries after identification of putative function by homology. The function was mostly verified by ectopic expression in *Arabidopsis*.

**GENETICALLY MODIFIED BULB CROPS**

**Gus**

Over the years protocols for gene transfer, either by bombardment or by *Agrobacterium*, have been developed based on monitoring GUS expression as a reporter (Jefferson et al., 1987). This also occurred in many bulb crops, not only in the major crops such as *Lilium*, *Tulipa*, *Iris* and *Gladiolus*, but including crops as *Agapanthus*, *Muscari*, *Narcissus*, *Alstroemeria*, *Hyacinthus*, *Ornithogalum* and *Hemerocallis*. It proved that the principle of gene transfer to bulb crops is feasible, however, the transformation efficiency is generally rather low.

**Morphological Traits**

Genetically modified bulbs or geophytes have been produced aimed at altering flower shape or color, at reducing pollen production in lily, changing plant habitus or leaf color (Table 5). Although transgenesis was confirmed by PCR or Southern analysis, not all traits could be observed in the plants generated. In *Caladium*, Li et al. (2005) successfully changed in the variegated leaves of the recipient the white parts into red-colored parts by introducing the anthocyanin biosynthesis regulatory genes Lc and C1. Fully-developed modified plants were obtained.

**Disease Resistance**

Several studies have been concerned with introducing resistance against viruses, bacteria or fungi (Table 6). Long-term disease resistance in the field still has to be proven in all cases. Here, we will report on the state-of-the-art introduction of virus-resistance in
Gladiolus and insect resistance in Lilium Oriental hybrids.

1. Virus Resistance. Several approaches can be chosen when aiming to introduce virus resistance. They encompass using coat protein genes of the targeted virus, replicase or movement protein genes, and the genes can be in either direct or antisense orientation. RNA interference (RNAi), in which parts of the target gene are present as inverted repeat sequences, also provide a potentially interesting approach to knock out required gene functions in viral infections. Finally, genes coding for antibodies raised against viral proteins can be used. In Gladiolus the introduction of the coat protein gene of Bean yellow mosaic virus (BYMV) in both sense or antisense orientation provided short-term resistance, however, no long-term resistance (Kamo et al., 2005). A variety of genes was tried for conferring resistance against the Cucumber mosaic virus in Gladiolus. Among them were the coat protein subgroup I and II genes, the defective replicase gene (alone or in different combinations with coat protein genes) and short-chain variable fragment, scFv, antibody genes against the subgroup I and II CMV coat proteins. The highest frequency (14-16%) of resistant plants were obtained with the replicase or coat protein II genes. A somewhat lower frequency of 5-7% was obtained with the antibody genes and no resistant plants with the coat protein I gene (Kamo et al., 2010, unpublished). Long-term performance in the field and the mechanisms involved will be studied in the near future.

2. Insect Resistance. In lilies, the problems generated by virus infections are not tackled by introducing genes by genetic modification against the viruses themselves, but by targeting the vector, responsible for spread of various viruses, i.e., aphids. Two genes are used in a dual approach in order to prevent build-up of an aphid population. One, the linalool synthase gene, will produce a volatile repellent, discouraging the aphids to land on the plants and starting to feed by sucking the phloem. The second gene, a proteinase inhibitor, will reduce fecundity after ingestion by aphids not affected by the repellent. These genes were tested before in other crops and proved to work in controlling aphids and/or thrips (Aharoni et al., 2003; Annadana et al., 2002). In our laboratory, a vector system is used aimed at producing marker-free genetically modified (GM) plants (Schaart et al., 2004; Krens et al., 2004). The latter is required if one wants to introduce GM plants onto the European market. Twenty-two lily cultivars, among which Lilium longiflorum, Lilium Oriental hybrid, Oriental-Trumpet (OT) hybrid, Asiatic hybrid and Longiflorum-Asiatic (LA) hybrid, were selected based on recommendations by Dutch lily breeding companies. They were mostly diploid, but also some triploids and a tetraploid were present. After establishing their regeneration and transformation potential (Krens et al., 2009), several transformation experiments finally resulted in 27 PCR-confirmed GM lily plants from seven cultivars. Expression was demonstrated by RT-PCR in 19 of them, representing three diploid Oriental hybrid cultivars, ‘White Express’, ‘Sheila’ and ‘Lake Carey’ and one OT hybrid, ‘Robina’ (Krens et al., unpublished). Research will continue with biochemical analyses for the presence of linalool and equistat (the proteinase inhibitor) and with aphid resistance assays in the greenhouse. Subsequently, individual GM lilies with shown resistance against aphids will be made marker-free by induced, recombination-based marker excision.

DEVELOPMENTS IN EU MARKET INTRODUCTION

Of the GM crops grown globally, only two species represent the ornamental sector, rose and carnation. Those crops have been altered in flower color having a blue or purple hue. Two GM rose lines are approved for commercial use, both of them in Japan and one of them in Australia. Recently, both acquired the deregulated status in the USA. The rose lines carry the nptII gene as selectable marker and did not enter the EU approval process. Two carnation GM cultivars are approved for commercial use in the European Union (EU); both carry a plant-derived sulfonylurea herbicide resistance gene as selectable marker. Cultivation and large-scale production is done in South America, Colombia and Ecuador or in Africa, Kenya or Ethiopia. In Directive 2001/18/EC of the European Parliament it is stated that the use of antibiotic resistance genes as selectable
markers should be phased out by 2008. Despite this statement, a GM potato with altered starch composition was authorized for commercial release in the EU and it carried an nptII selectable marker gene. Still, Dutch authorities emphasize that when applying for permission to prepare field trials with GM crops (environmental release), it is best to make sure that no antibiotic resistance genes are present anymore within the GM crop. This is why we use the excision-based marker removal system developed earlier in our laboratory (Schaart et al., 2004; Krens et al., 2004) in the production of insect-resistant lilies.

Until recently, it was generally believed that procedures for environmental release of GM ornamental crops would be easier, more rapid and cheaper because no large scale toxicity studies would be required as the crops were not eaten. In fact, within the report of the European Food Safety Authority (EFSA) on the Florigene carnation Moonaqum™ it is explicitly mentioned “Given that carnation Moonaqum 123.8.12 is not intended for human or animal consumption as food or feed but for ornamental use only, the GMO panel does not consider it necessary to perform a comprehensive food/feed safety assessment on the whole GM plant”. However, the report then continues with a thorough toxicity and allergenicity screen. (http://www.gmocompass.org/pdf/regulation/Flowers/Carnation_ moonaqum_efsa_opinion.pdf).

For the application for approval for market introduction in the EU it appeared that there is no existing form giving clear instructions on what information should be delivered.

One is referred to the carnation applications because those were approved. The minimum information required and what is superfluous in those files consisting of 238 pages is not defined. Extensive reference is made in the carnation files to years of experience with cultivation, handling and environmental risk issues obtained outside the EU, of course, in the case of a newly generated GM ornamental crop one cannot refer to earlier experience but one has still to acquire the information. For this, extensive field trials will be required once one knows what information is necessary. New dossiers are said to be obligatory for each cultivar, even with the same genes, for each derived cultivar, e.g., after crossing using the GM line as parent, even for every application with a given cultivar when this is different, e.g., for cut flowers or for potted plants. The estimated time-frame is 5.5 years from first applying for field trials to the decision by the EU on market introduction; costs per application seem to be at least € 200,000. This means that GM in the EU is not feasible for ornamental crops or for small and medium enterprises that are the predominant business forms in the ornamental sector. New technologies such as genome editing with zinc finger nucleases, mutation induction by oligonucleotide mediated mutation or cisgenesis when exempted or considered not a GMO technique might present alternatives depending on the outcome of ongoing discussions within the EU on these techniques and how to deal with them (http://www.efsa.europa.eu/en/efsajournal/pub/2561.htm; Lusser et al., 2012).

CONCLUSIONS
Great challenges lie ahead in applying genetic modification or new breeding techniques in flower bulb crops. Sequence information on genes and promoters or on alleles is required for this. Transformation or macromolecule delivery systems combined with regeneration protocols still need developing or further optimization, and finally, market introductions should become economically viable.

ACKNOWLEDGEMENTS
The research on GM lilies by F.A. Krens et al. is funded by the Dutch Commodity Board for Horticulture.

Literature Cited
Terpenoid metabolism in wild-type and transgenic Arabidopsis plants. The Plant Cell 15:2866-2884.


Tables

Table 1. The list of fully sequenced plant genomes*, June 2012.

<table>
<thead>
<tr>
<th>Columbine</th>
<th>Poplar</th>
<th>Soybean</th>
<th>Date palm</th>
<th>Foxtail millet</th>
</tr>
</thead>
<tbody>
<tr>
<td>Potato</td>
<td>Strawberry</td>
<td>Cocoa</td>
<td>Rice</td>
<td>Sugarbeet</td>
</tr>
<tr>
<td>Grape</td>
<td>Castor bean</td>
<td>Papaya</td>
<td>Brachy</td>
<td>Flax</td>
</tr>
<tr>
<td>Cucumber</td>
<td>Apple</td>
<td>Thale cress</td>
<td>Maize</td>
<td>Cannabis</td>
</tr>
<tr>
<td>Chinese cabbage</td>
<td>Peach</td>
<td>Salt cress</td>
<td>Sorghum</td>
<td>Lotus japonicus</td>
</tr>
<tr>
<td>Medicago</td>
<td>Pigeon pea</td>
<td>Common bean</td>
<td>Cotton</td>
<td>Capsella rubella</td>
</tr>
</tbody>
</table>

Table 2. Some examples of genome sizes, including those of some representatives of flower bulb genera. Note that within a genus the species can show very large differences in genome size. The exemplary numbers are chosen by the authors in order to demonstrate the size and variation within flower bulb crops and to allow comparison with other organisms.

<table>
<thead>
<tr>
<th>Reference species</th>
<th>Haploid genome size in Mbp</th>
<th>Bulb species (family)</th>
<th>Haploid genome size in Mbp</th>
</tr>
</thead>
<tbody>
<tr>
<td>Humans</td>
<td>3.300</td>
<td>Gladiolus (Iridaceae)</td>
<td>1.100</td>
</tr>
<tr>
<td>Thale cress</td>
<td>125</td>
<td>Muscari (Asparagaceae)</td>
<td>4.200</td>
</tr>
<tr>
<td>Rice</td>
<td>430</td>
<td>Iris (Iridaceae)</td>
<td>12.000</td>
</tr>
<tr>
<td>Apple</td>
<td>750</td>
<td>Tulip (Liliaceae)</td>
<td>24.000</td>
</tr>
<tr>
<td>Maize</td>
<td>2.500</td>
<td>Hyacinth (Asparagaceae)</td>
<td>40.000</td>
</tr>
<tr>
<td>Barley</td>
<td>5.300</td>
<td>Lily (Liliaceae)</td>
<td>40.000</td>
</tr>
<tr>
<td>Wheat</td>
<td>16.000</td>
<td>Fritillaria (Liliaceae)</td>
<td>130.000</td>
</tr>
<tr>
<td></td>
<td></td>
<td><em>Paris japonica</em> (Melianthiaceae)</td>
<td>150.000</td>
</tr>
</tbody>
</table>

* Present world record holder.

Table 3. Promoters isolated and characterized from flower bulb crops.

<table>
<thead>
<tr>
<th>Genus</th>
<th>Number</th>
<th>Origin</th>
<th>Goal</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lilium</td>
<td>4</td>
<td>Flower (organ)</td>
<td>Tissue specific expression</td>
</tr>
<tr>
<td>Hyacinth</td>
<td>1</td>
<td>Shoot</td>
<td>Cold induced expression</td>
</tr>
<tr>
<td>Crocus</td>
<td>1</td>
<td>Flower (organ)</td>
<td>Tissue specific expression</td>
</tr>
<tr>
<td>Crocus</td>
<td>2</td>
<td>Flower (color)</td>
<td>Tissue specific expression</td>
</tr>
<tr>
<td>Gladiolus</td>
<td>1</td>
<td>Callus</td>
<td>Constitutively high</td>
</tr>
<tr>
<td>Gladiolus</td>
<td>1</td>
<td>Flower (stage)</td>
<td>Dev. stage specific expression</td>
</tr>
</tbody>
</table>

Table 4. Genes isolated and characterized from flower bulb crops.

<table>
<thead>
<tr>
<th>Function</th>
<th>Genus</th>
<th>Number</th>
</tr>
</thead>
<tbody>
<tr>
<td>Resistance, defense</td>
<td>Galanthus, Tulipa, Iris, Crocus</td>
<td>4</td>
</tr>
<tr>
<td>Nutritional value</td>
<td>Crocus, Narcissus</td>
<td>4</td>
</tr>
<tr>
<td>Flower color</td>
<td>Narcissus, Muscari, Iris, Tulipa</td>
<td>7</td>
</tr>
<tr>
<td>Flower development</td>
<td>Lilium, Narcissus, Crocus, Alstroemeria</td>
<td>Numerous (MADS box genes)</td>
</tr>
<tr>
<td>Abiotic stress</td>
<td>Hyacinth, Lilium</td>
<td>2</td>
</tr>
<tr>
<td>Miscellaneous</td>
<td>Crocus, Gladiolus</td>
<td>2</td>
</tr>
</tbody>
</table>
Table 5. Morphological traits as a target for genetic modification in geophytes.

<table>
<thead>
<tr>
<th>Crop</th>
<th>Gene</th>
<th>Trait</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Caladium</td>
<td>Maize Lc+C1</td>
<td>Leaf color</td>
<td>Li et al., 2005</td>
</tr>
<tr>
<td>Lilium</td>
<td>Rol A,B,C</td>
<td>Dwarf</td>
<td>Mercuri et al., 2003</td>
</tr>
<tr>
<td>Lilium</td>
<td>Maize Zm401</td>
<td>Pollen-less</td>
<td>Li et al., 2008</td>
</tr>
<tr>
<td>Lilium</td>
<td>Carotenoids</td>
<td>Flower, leaf color</td>
<td>Azadi et al., 2010</td>
</tr>
<tr>
<td>Narcissus</td>
<td>Phytoene synthase</td>
<td>Flower color</td>
<td>Lu et al., 2007</td>
</tr>
<tr>
<td>Tricyrtis</td>
<td>Agapanthus MADS</td>
<td>Flower shape, color</td>
<td>Nakano et al., 2007</td>
</tr>
</tbody>
</table>

Table 6. Disease resistance as a target for genetic modification in geophytes.

<table>
<thead>
<tr>
<th>Crop</th>
<th>Gene</th>
<th>Target</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gladiolus</td>
<td>coat protein/replicase</td>
<td>Bean yellow mosaic virus (BYMV), Cucumber mosaic virus (CMV)</td>
<td>Kamo et al., 2005, 2010</td>
</tr>
<tr>
<td>Ornithogalum</td>
<td>coat protein/replicase</td>
<td>CMV</td>
<td>Cohen et al., 2005;</td>
</tr>
<tr>
<td></td>
<td>tachyplesin1</td>
<td>Erwinia</td>
<td>Van Emmenes et al., 2008</td>
</tr>
<tr>
<td>Zantedeschia</td>
<td>ferredoxin-like, pfpl</td>
<td>Erwinia</td>
<td>Yip et al., 2011</td>
</tr>
<tr>
<td>Hyacinthus</td>
<td>thaumatinII</td>
<td>Botrytis, Fusarium</td>
<td>Popowich et al., 2007</td>
</tr>
</tbody>
</table>

Figures

Fig. 1. A comparison of gene promoter activities in suspension cells of Gladiolus by measuring GUS expression 48 h after bombardment. Values represent relative GUS expression where 35S-GUS is set at 1.0. CaMV 35S= Cauliflower mosaic virus; mas2 = mannopine synthase; rolD = root loci D gene of Agrobacterium rhizogenes; Act1 = actin1 gene; Ubi1 = ubiquitin1 gene.