Plant Biotechnology 28, 123–130 (2011)
DOI: 10.5511/plantbiotechnology.11.0106a

Original Paper

The new FioreDB database provides comprehensive information on plant transcription factors and phenotypes induced by CRES-T in ornamental and model plants

Nobutaka Mitsuda[1,*], Yuko Takiguchi[1], Masahito Shikata[2], Kimiyo Sage-Ono[3], Michiyuki Ono[3], Katsutomo Sasaki[2], Hiroyasu Yamaguchi[2], Takako Narumi[2,a], Yuri Tanaka[4], Masao Sugiyama[4], Tomomichi Yamamura[4], Teruhiko Terakawa[4], Keiko Gion[5], Ryota Suzuri[5], Yoshikazu Tanaka[5], Takashi Nakatsuka[6], Souichi Kimura[6], Masahiro Nishihara[6], Tomoyuki Sakai[7], Reiko Endo-Onodera[7], Kumi Saitoh[7], Kanji Isuzugawa[7], Yoshimi Oshima[1], Tomotsugu Koyama[1,b], Miho Ikeda[1], Megumi Narukawa[1,c], Kyoko Matsui[1], Masaru Nakata[1], Norihiro Ohtsubo[2], Masaru Ohme-Takagi[1]

[1] Bioproduction Research Institute, National Institute of Advanced Industrial Science and Technology, Tsukuba, Ibaraki 305-8562, Japan; [2] National Institute of Floricultural Science, National Agriculture and Food Research Organization, Tsukuba, Ibaraki 305-8519, Japan; [3] Gene Research Center, Graduate School of Life and Environmental Sciences, University of Tsukuba, Tsukuba, Ibaraki 305-8572, Japan; [4] Central Research Laboratories, Hokko Chemical Industry Co., Atsugi, Kanagawa 243-0023, Japan; [5] Institute for Plant Science, Suntory Holdings Ltd., Mishima, Osaka 618-8503, Japan; [6] Iwate Biotechnology Research Center, Kitakami, Iwate 024-0003, Japan; [7] Horticultural Experiment Station, Yamagata Integrated Agricultural Research Center, Sagae, Yamagata 991-0043, Japan
E-mail: nobutaka.mitsuda@aist.go.jp Tel: +81-29-861-2641 Fax: +81-29-861-3026

Received November 19, 2010; accepted January 6, 2011 (Edited by T. Mizoguchi)

Abstract FioreDB (http://www.cres-t.org/fiore/public_db/) is a database of phenotypes induced by Chimeric REpressor gene-Silencing Technology (CRES-T) in ornamental and model plants. CRES-T induces a loss-of-function phenotype of a transcription factor (TF) gene by expression of a chimeric repressor produced by fusion of a TF to the strong transcriptional repression domain (SRDX). The earlier version of FioreDB provided phenotypic information induced by many kinds of chimeric repressors in various plants including torenia, chrysanthemum, cyclamen, gentian, morning glory, lisianthus and *Arabidopsis*. Phenotypic information, however, was not linked with gene information. We report here the development of the new FioreDB that provides more than 300 phenotypic information of various plants, linked to more than 100 TFs. FioreDB also provides information about classification of TFs, putative repression motifs found in TFs and other proteins, and incorporates publicly available gene information such as sequences and microarray data for all *Arabidopsis* genes. The new FioreDB described here, will be a valuable resource for basic research of TFs and for the manipulation of traits of agronomically important plants by CRES-T, especially from the point of view of horticulture.

Key words: Chimeric repressor gene-silencing technology (CRES-T), database, FioreDB, flower, SRDX, transcription factor.

CRES-T is a recently developed powerful technology that induces a loss-of-function phenotype of plant TFs by expression of a chimeric repressor in which a TF is fused with the strong transcriptional repression domain, SRDX (Hiratsu et al. 2003). The chimeric repressor dominantly represses the expression of target genes even in the presence of endogenous and functionally redundant TFs (Hiratsu et al. 2003; Mitsuda and Ohme-Takagi 2009 and references therein). We have shown that a chimeric repressor, constructed with a TF derived from the model plant *Arabidopsis thaliana*, often worked efficiently in other plant species without any modification (Gion et al. 2011; Tanaka et al. 2011; Shikata et al. 2011). These advantages have prompted us to apply CRES-T to

---

Abbreviations: CRES-T, chimeric repressor gene-silencing technology; TF, transcription factor.

[a] Present address: Faculty of Agriculture, Kagawa University, Miki, Kagawa 761-0795, Japan

[b] Present address: Division of Integrated Life Science, Graduate School of Biostudies, Kyoto University, Sakyo, Kyoto 606-8502, Japan

[c] Present address: Integrated Genomics Research Team, Plant Science Center, RIKEN, Yokohama, Kanagawa 230-0045, Japan

This article can be found at http://www.jspcmb.jp/

agronomically important crops and ornamental plants with high chromosome ploidy and limited gene information that often hamper us from generating loss-of-function mutants. FioreDB is a database that was created to provide phenotypic information induced by CRES-T in various ornamental plants including torenia, chrysanthemum, cyclamen, gentian, morning glory, lisianthus and a model plant *Arabidopsis* (Mitsuda et al. 2008, http://www.cres-t.org/fiore/public_db/). The phenotypic information was accumulated from more than 100 kinds of transgenic plants that express chimeric repressors of different *Arabidopsis* TFs (Mitsuda et al. 2008). However, the link between gene information and phenotypic information was not provided in the earlier version of FioreDB. We have now developed a new FioreDB database that provides information on more than 300 phenotypes, which newly include that of rose and carnation, by application of CRES-T for more than 100 TFs. The new FioreDB is available through a simple web-based interface that provides easy access to information about TFs that induce phenotype of interest by applying CRES-T. FioreDB also provides information on TF classification based on our original criteria along with details of the transcriptional repression domain found in *Arabidopsis* proteins including proteins that are not TFs. In addition, FioreDB contains information on all *Arabidopsis* genes, including sequences and microarray data from the Arabidopsis Information Resource (TAIR) (http://www.arabidopsis.org/, Swarbreck et al. 2008) and other web sites. The new FioreDB will contribute to basic research of plant TFs and to the manipulation of traits of agronomically important plants using CRES-T, especially from the point of view of horticulture.

Materials and methods

The data and software implemented in FioreDB

All sequence data (coding sequences and genomic sequences) for *Arabidopsis thaliana* that is stored in FioreDB was downloaded from TAIR (TAIR9) (http://www.arabidopsis.org/, Swarbreck et al. 2008). Gene sequences and their flanking regions were excised from the genome sequence using an original Perl script (http://www.perl.org/). We redefined 1,906 genes and classified them into 48 different TF families based on our analysis of how they were represented in four different TF databases (RARTF, Iida et al. 2005; AtTFDB, Davuluri et al. 2003; PlnTFDB, Riano-Pachon et al. 2007; DATF, Guo et al. 2005) (Table S1). The conserved domain of each family was identified by HMMER3 (http://hmmer.org/) and manual curation. The amino-acid sequences were aligned using MAFFT (Katoh et al. 2002; Katoh and Toh 2008) with the default settings and modified for clarity using BOXSHADE software (http://www.ch.embnet.org/software/BOX_doc.html). Phylogenetic trees were constructed by the Neighbor-Joining method (Saitou and Nei 1987) and the reliability of clusters was measured by the bootstrap test (1000 replicates) (Felsenstein 1985) with ClustalW software (Chenna et al. 2003). Sequence data for rice (Itoh et al. 2007; Tanaka et al. 2008) was downloaded from the Rice Annotation Project Database (RAP-DB) (http://rapdb.dna.affrc.go.jp/), and the data for poplar (Tuskan et al. 2006), sorghum (Paterson et al. 2009), soybean (Schmutz et al. 2010) and *Physcomitrella* (Rensing et al. 2008) was obtained from Phytozome (http://www.phytozome.net/). Putative orthologs of the *Arabidopsis* genes in these plants were identified by a BLAST-based method. Using the *Arabidopsis* gene as the original query for the BLAST search against the data set of each plant, the top-hit gene for each plant that reversely top-hit the original query gene (mutual top-hit) was defined as a class I ortholog. A class I ortholog with a BLAST Score Ratio (BSR) (Rasko et al. 2005) >0.4 was defined as class Ia and the remaining class I orthologs (BSR <0.4) were defined as class Ib. The top-hit genes in each plant that did not reversely top-hit to the original query or non top-hit genes which top-hit back to the original query with BSRs >0.4 were defined as class II orthologs. Other genes with BSRs >0.4 were defined as class III orthologs. Putative repression domains represented by "(F/L/I/M)DL(D/E/N/Q/R)X$^{1\text{-}3}$P", "(non-hydrophobic)L(non-hydrophobic)L(non-hydrophobic)L", "(K/R)LFGV", "LLLFR" or "TLXLF(P/R)" were searched for using an original Perl script (Hiratsu et al. 2004; Ikeda et al. 2009; Ikeda and Ohme-Takagi 2009; Matsui et al. 2008; Mitsuda and Ohme-Takagi 2009). X$^{1\text{-}3}$ represents 1 to 3 residues of any amino acid and non-hydrophobic represents an A, R, N, D, C, Q, E, G, H, K, M, S or T residue. Development series (Schmid et al. 2005), abiotic-stress series (Kilian et al. 2007), light series, pathogen series and tiling array series (Laubinger et al. 2008; Zeller et al. 2009) microarray data were downloaded from the Weigel World web site (http://www.weigelworld.org/). Diurnal microarray data (Michael et al. 2008; Mockler et al. 2007) was downloaded from ArrayExpress (Accession Number: E-MEXP-1304, http://www.ebi.ac.uk/arrayexpress/, Parkinson et al. 2009). All other microarray data were downloaded from the AtGenExpressJPN web site (http://igrt0.psc.riken.jp/, Goda et al. 2008). All CGIs for FioreDB were written in Perl and run on an Apache web server (http://www.apache.org/) with the MySQL database system (http://www.mysql.com/).

Results and discussion

The new home page of FioreDB

The home page of the new FioreDB (http://www.cres-t.org/fiore/public_db/) has been updated in three main ways. First, a keyword search box has been provided (Figure 1A). Users can, for example, search for *Arabidopsis* genes using a keyword such as a common gene name (e.g. AGAMOUS or AG) or an AGI gene code (e.g. AT4G18960). The detailed gene information is displayed on a new page with phenotype if the data exists (see later section). Gene information were not provided by "Search by Phenotype" and "Browse Constructs" in the older version of FioreDB, but the new FioreDB provides actual information regarding the precise gene name that induces the phenotype of interest (see later section). Second, newly classified TF families are shown at the bottom of the home page (Figure 1A) and hyperlinked so that users can browse the phylogenetic tree of each of the TF families by clicking the TF family name (see later section). The new FioreDB also has a role of TF database. Third, the top banner shows a randomly selected photograph of a plant with a phenotype induced by CRES-T instead of the wild-type flowers of various plants shown in the previous version of FioreDB (Figure 1A). This banner appears on every page of the FioreDB and by clicking on it, the user can browse detailed information about the gene that induces the phenotype on the linked page.

Finding a TF that induces a phenotype of interest

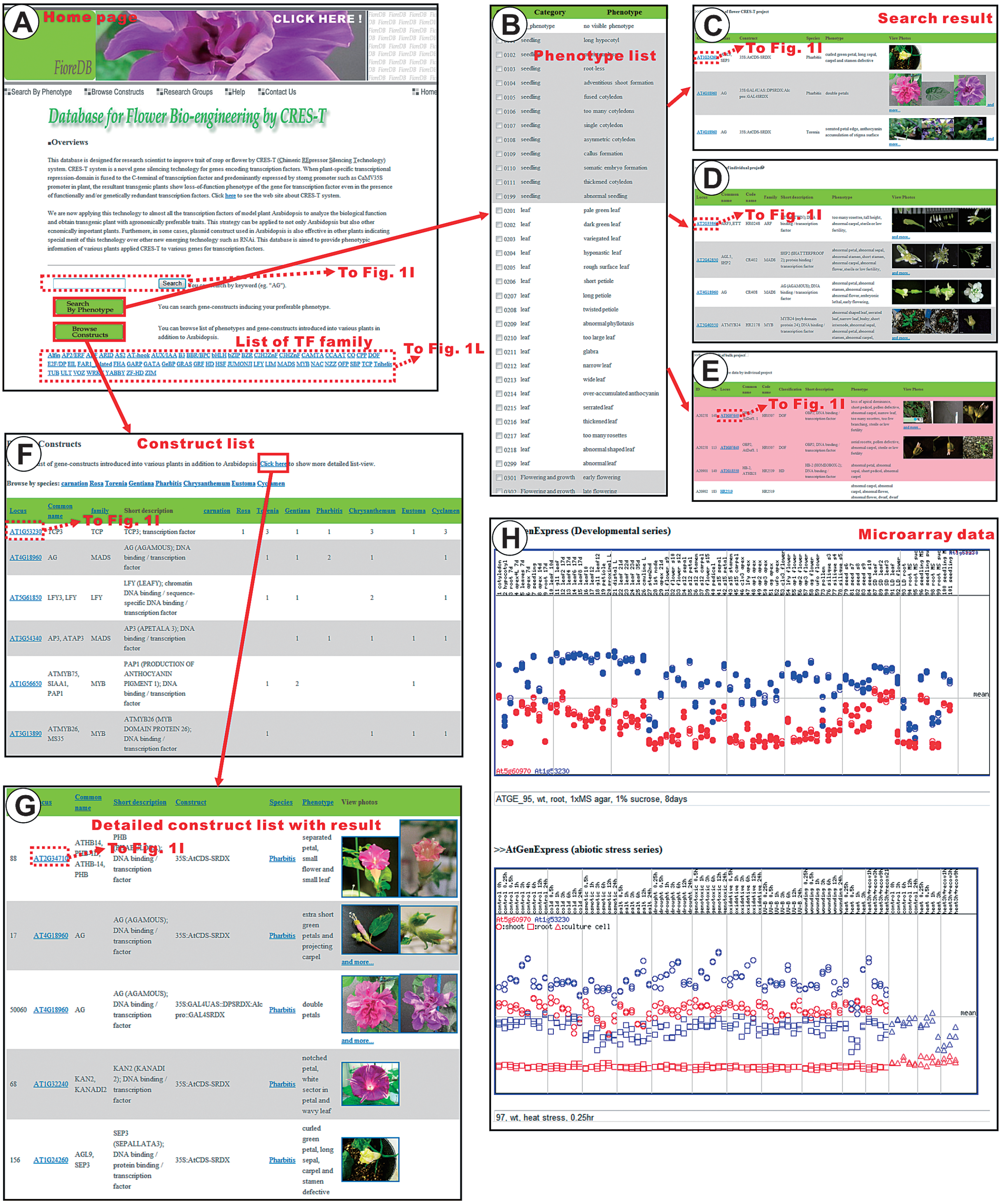
Users can search for the TF that induces a phenotype caused by CRES-T by clicking on the "Search by Phenotype" hyperlink in the home page. Users can select phenotypes of interest from the list containing 80 kinds of phenotypes (Figure 1B) (Mitsuda et al. 2008). We adopted two strategies to generate transgenic plants; the individual project and the bulk project. In the individual project, we made multiple transgenic plants by transforming each gene construct one by one. In the bulk project, mixtures of 40 to 60 gene constructs were transformed into *Arabidopsis* and torenia and several hundred transgenic plants were cultivated to identify, in an efficient manner, TFs that induce intriguing phenotypes (Mitsuda et al. 2008; Shikata et al. 2011). In *Arabidopsis*, mixtures of T1 transgenic seeds for the more than 1,500 TFs that were prepared one by one in the individual project, were grown on a large scale (more than 10,000 plants) (Mitsuda et al. 2008). In torenia, two different mixtures of gene constructs for around 50 TF genes that were expected to induce, with high efficiency, phenotypic changes in flower were transformed (Shikata et al. 2011). Search results contain data from both projects that is provided independently as three lists of TF genes: TFs that induce the selected phenotypes in ornamental plants in both projects; *Arabidopsis* TFs identified in the individual project; and TFs of *Arabidopsis* from the bulk project (Figures 1C-E). For ornamental plants, the *Arabidopsis* gene is listed even if the orthologous gene in the plant was used. If photographic data exists, it is shown in the rightmost column (Figures 1C-E). If the data are linked with gene information, users can browse the gene information and additional detailed phenotype information by clicking the hyperlinked gene name in the leftmost column (see later section). All the data on ornamental plants and part of the data (50 of the 350 TFs that were individually examined) on *Arabidopsis* is linked with gene information. All the results of our CRES-T projects are summarized in Table 1. In total, data for about 400 phenotypes with gene information is stored. More than half the phenotypes are from *Arabidopsis*. The efficiency of induction of abnormal phenotypes was high in *Arabidopsis* (257 phenotypes by 50 TFs) and morning glory (36 phenotypes by 11 TFs), while in gentian (4 phenotypes by 23 TFs) it was low, probably because of gene silencing of the transgene (Mishiba et al. 2005; Mishiba et al. 2010). In ornamental plants, flower phenotypes, especially in the petals, were frequently observed because the petals of ornamental plants are conspicuous and modification of the petal is the main focus of our project.

Browsing constructs

Users can browse all *Arabidopsis* TF genes used in our project as a list, viewed by clicking the "Browse Constructs" link on the home page. We made transgenic ornamental plants that express more than 100 individual TF genes. The list view provides a non-redundant gene list that summarizes the experiments performed in each plant (Figure 1F). A more detailed list view (Figure 1G) showing all of the experiments in individual projects for each or all of the plants can be obtained by clicking the hyperlinked plant name or "here" at the top of the page shown in Figure 1F. This detailed list view exhibits photographic data, if it exists, in the rightmost column (Figure 1G). Representative results are summarized in Table 2. The TEOSINTE BRANCHED1, CYCLOIDEA and PCF (TCP) TFs induced the expected phenotypes in most examined species even when the *Arabidopsis* gene was employed (Table 2) (Gion et al. 2011; Narumi et al. 2011; Tanaka et al. 2011), while the *Arabidopsis* MADS box TFs, such as AGAMOUS (AG) and APETALA3 (AP3), did not induce the expected phenotypes in plants other than *Arabidopsis* (Table 2). However, as shown for morning glory, when orthologs of the MADS box TFs from each species was employed, the expected phenotypes were induced (Table 2) (Sage-Ono et al. 2011). A number of induced phenotypes remain to be characterized in detail.

Browsing gene information

The new FioreDB contains information, not only on TFs but on all *Arabidopsis* genes, that was not provided in previous version. The information from TAIR (Swarbreck et al. 2008) includes the gene alias and description, the gene ontology classification, the gene model, sequences of the coding regions, the amino acid sequences, and the genomic sequences around the genes. In addition, information from the Arabidopsis Small RNA Project (ASRP) (Backman et al. 2008) for miRNAs/tasiRNAs is provided if the gene is targeted by them. An example of the gene view is shown in Figure 1I. In the second panel on the gene view page, the categorization of genes, classified as TFs, in FioreDB and four other TF databases (RARTF, Iida et al. 2005; AtTFDB, Davuluri et al. 2003;



Database for Flower Bio-engineering by CRES-T

Overviews

This database is designed for research scientist to improve trait of crop or flower by CRES-T (Chimeric REpressor Silencing Technology) system. CRES-T system is a novel gene silencing technology for genes encoding transcription factors. When plant-specific transcriptional repression-domain is fused to the C-terminal of transcription factor and predominantly expressed by storng promoter such as CaMV35S promoter in plant, the resultant transgenic plants show loss-of-function phenotype of the gene for transcription factor even in the presence of functionally and/or genetically redundant transcription factors. Click here to see the web site about CRES-T system.

We are now applying this technology to almost all the transcription factors of model plant Arabidopsis to analyze the biological function and obtain transgenic plant with agronomically preferable traits. This strategy can be applied to not only Arabidopsis but also other ecnomically important plants. Furthemore, in some cases, plasmid construct used in Arabidopsis is also effective in other plants indicating special merit of this technology over other new emerging technology such as RNAi. This database is aimed to provide phenotypic information of various plants applied CRES-T to various genes for transcription factors.

You can search by keyword (eg. "AG").

You can search gene-constructs inducing your preferable phenotype.

You can browse list of phenotypes and gene-constructs introduced into various plants in addition to Arabidopsis.

Alfin AP2/ERF ARF ARID AS2 AT-hook AUX/IAA B3 BBR/BPC bHLH bZIP BZR C2H2ZnF C3HZnF CAMTA CCAAT CO CPP DOF E2F/DP EIL FAR1_related FHA GARP GATA GeBP GRAS GRF HD HSF JUMONJI LFY LIM MADS MYB NAC NZZ OFP SBP TCP Trihelix TUB ULT VOZ WRKY YABBY ZF-HD ZIM

	Category	Phenotype
	_phenotype	no visible phenotype
☐	seedling	long hypocotyl
☐ 0102	seedling	
☐ 0103	seedling	root-less
☐ 0104	seedling	adventitious shoot formation
☐ 0105	seedling	fused cotyledon
☐ 0106	seedling	too many cotyledons
☐ 0107	seedling	single cotyledon
☐ 0108	seedling	asymmetric cotyledon
☐ 0109	seedling	callus formation
☐ 0110	seedling	somatic embryo formation
☐ 0111	seedling	thickened cotyledon
☐ 0199	seedling	abnormal seedling
☐ 0201	leaf	pale green leaf
☐ 0202	leaf	dark green leaf
☐ 0203	leaf	variegated leaf
☐ 0204	leaf	hyponastic leaf
☐ 0205	leaf	rough surface leaf
☐ 0206	leaf	short petiole
☐ 0207	leaf	long petiole
☐ 0208	leaf	twisted petiole
☐ 0209	leaf	abnormal phyllotaxis
☐ 0210	leaf	too large leaf
☐ 0211	leaf	glabra
☐ 0212	leaf	narrow leaf
☐ 0213	leaf	wide leaf
☐ 0214	leaf	over-accumulated anthocyanin
☐ 0215	leaf	serrated leaf
☐ 0216	leaf	thickened leaf
☐ 0217	leaf	too many rosettes
☐ 0218	leaf	abnormal shaped leaf
☐ 0299	leaf	abnormal leaf
☐ 0301	Flowering and growth	early flowering
☐ 0302	Flowering and growth	late flowering

>> of flower CRES-T project

	lias	Construct	Species	Phenotype	View Photos
AT1G24260	AGL9, SEP3	35S:AtCDS-SRDX	Pharbitis	curled green petal, long sepal, carpel and stamen defective	
AT4G18960	AG	35S:GAL4UAS::DPSRDX:Alc pro::GAL4SRDX	Pharbitis	double petals	and more...
AT4G18960	AG	35S:AtCDS-SRDX	Torenia	serrated petal edge, anthocyanin accumulation of stigma surface	and more...

>> of individual project

Locus	Common name	Code name	Family	Short description	Phenotype	View Photos
AT2G33860	ARF3, ETT	HR0248	ARF	DNA binding / transcription factor	too many rosettes, tall height, abnormal carpel, sterile or low fertility,	and more...
AT2G42830	AGL5, SHP2	CR402	MADS	SHP2 (SHATTERPROOF 2); protein binding / transcription factor	abnormal petal, abnormal sepal, abnormal stamen, short stamen, abnormal carpel, abnormal flower, sterile or low fertility,	and more...
AT4G18960	AG	CR408	MADS	AG (AGAMOUS); DNA binding / transcription factor	abnormal petal, abnormal stamen, abnormal carpel, abnormal flower, embryonic lethal, early flowering,	
AT5G40350	ATMYB24	HR2178	MYB	MYB24 (myb domain protein 24); DNA binding / transcription factor	abnormal shaped leaf, serrated leaf, narrow leaf, bushy, short internode, abnormal sepal, abnormal petal, abnormal stamen, abnormal carpel,	and more...

>>S lt of bulk project

e data by indivisual project

ID	No.	Locus	Common name	Code name	Classification	Short description	Phenotype	View Photos
A20258	148	AT1G07640	OBP2, AtDof1. 1	HR0507	DOF	OBP2; DNA binding / transcription factor	loss of apical dominance, short pedicel, pollen defective, abnormal carpel, narrow leaf, too many rosettes, too few branching, sterile or low fertility	and more...
A20258	113	AT1G07640	OBP2, AtDof1. 1	HR0507	DOF	OBP2; DNA binding / transcription factor	aerial rosette, pollen defective, abnormal carpel, sterile or low fertility	
A20901	140	AT2G18550	HB-2, ATHB21	HR2539	HD	HB-2 (HOMEOBOX-2); DNA binding / transcription factor	abnormal petal, abnormal sepal, short pedicel, abnormal carpel	
A20902	183	HR2519		HR2519			abnormal carpel, abnormal carpel, abnormal flower, abnormal flower, dwarf, dwarf	

Constructs

ist of gene-constructs introduced into various plants in addition to Arabidopsis Click here to show more detailed list-view.

Browse by species: carnation Rosa Torenia Gentiana Pharbitis Chrysanthemum Eustoma Cyclamen

Locus	Common name	family	Short description	carnation	Rosa	Torenia	Gentiana	Pharbitis	Chrysanthemum	Eustoma	Cyclamen
AT1G53230	TCP3	TCP	TCP3; transcription factor		1	3	1	1	3	1	3
AT4G18960	AG	MADS	AG (AGAMOUS); DNA binding / transcription factor			1	1	2	1		1
AT5G61850	LFY3, LFY	LFY	LFY (LEAFY); chromatin DNA binding / sequence-specific DNA binding / transcription factor			1	1		2		1
AT3G54340	AP3, ATAP3	MADS	AP3 (APETALA 3); DNA binding / transcription factor				1	1	1	1	1
AT1G56650	ATMYB75, SIAA1, PAP1	MYB	PAP1 (PRODUCTION OF ANTHOCYANIN PIGMENT 1); DNA binding / transcription factor			1	2			1	
AT3G13890	ATMYB26, MS35	MYB	ATMYB26 (MYB DOMAIN PROTEIN 26); DNA binding / transcription factor			1			1	1	1

	cus	Common name	Short description	Construct	Species	Phenotype	View photos
88	AT2G34710	ATHB14, PHB-1D, ATHB-14, PHB	PHB (PHABULOSA); DNA binding / transcription factor	35S:AtCDS-SRDX	Pharbitis	separated petal, small flower and small leaf	
17	AT4G18960	AG	AG (AGAMOUS); DNA binding / transcription factor	35S:AtCDS-SRDX	Pharbitis	extra short green petals and projecting carpel	and more...
50060	AT4G18960	AG	AG (AGAMOUS); DNA binding / transcription factor	35S:GAL4UAS::DPSRDX:Alc pro::GAL4SRDX	Pharbitis	double petals	and more...
68	AT1G32240	KAN2, KANADI2	KAN2 (KANADI 2); DNA binding / transcription factor	35S:AtCDS-SRDX	Pharbitis	notched petal, white sector in petal and wavy leaf	
156	AT1G24260	AGL9, SEP3	SEP3 (SEPALLATA3); DNA binding / protein binding / transcription factor	35S:AtCDS-SRDX	Pharbitis	curled green petal, long sepal, carpel and stamen defective	

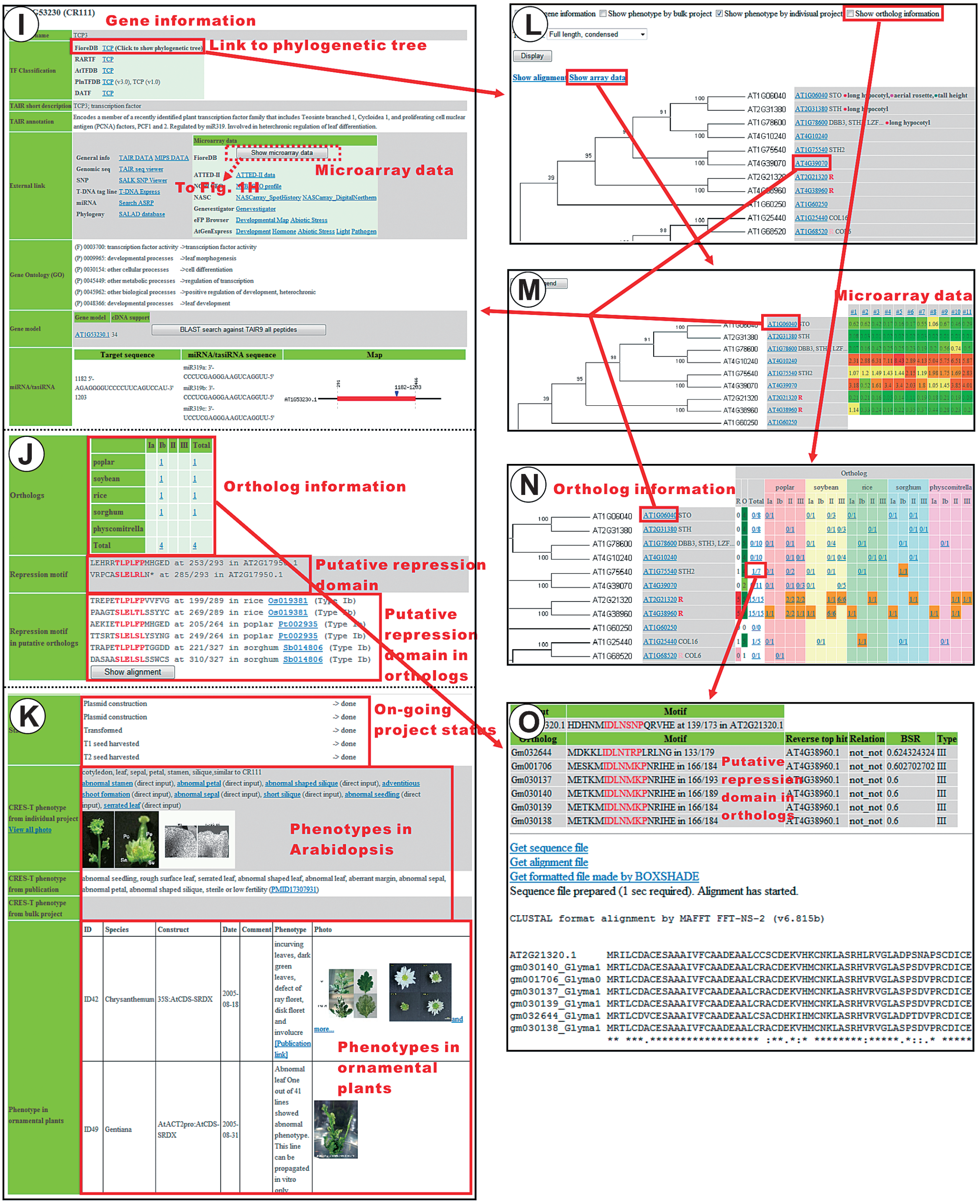
enExpress (Developmental series)

ATGE_95, wt, root, 1xMS agar, 1% sucrose, 8days

>>AtGenExpress (abiotic stress series)

97, wt, heat stress, 0.25hr

Figure 1. The new FioreDB. (A) Home page of the new FioreDB. A photograph of a plant with an impressive phenotype induced by CRES-T is randomly selected from the database and is used as the top banner on each page. Detailed gene information can be obtained by clicking on the photograph. (B) Phenotype list of "Search by Phenotype". Users can select a phenotype of interest from 80 classified phenotypes. (C) Result of a phenotype search in ornamental plants. The gene name is shown at the leftmost column. (D) Result of a phenotype search in the *Arabidopsis* individual project. (E) Result of a phenotype search in the *Arabidopsis* bulk project. (F) Non-redundant list of genes introduced into various ornamental plants shown by clicking "Browse Construct" on home page. (G) List of all experiments performed in our project. (H) Expression profile from various microarray data. Two genes are comparatively shown.



```
LEHRRTLPLFPMHGED at 253/293 in AT2G17950.1
VRPCASLELRLN* at 285/293 in AT2G17950.1
```

```
TREPETLPLFPVVFVG at 199/289 in rice Os019381 (Type Ib)
PAAGTSLELTLSSYYC at 269/289 in rice Os019381 (Type Ib)
AEKIETLPLFPMHGED at 205/264 in poplar Pt002935 (Type Ib)
TTSRTSLELSLYSYNG at 249/264 in poplar Pt002935 (Type Ib)
TRAPETLPLFPTGGDD at 221/327 in sorghum Sb014806 (Type Ib)
DASAASLELSLSSWCS at 310/327 in sorghum Sb014806 (Type Ib)
```

Get sequence file
Get alignment file
Get formatted file made by BOXSHADE
Sequence file prepared (1 sec required). Alignment has started.

```
CLUSTAL format alignment by MAFFT FFT-NS-2 (v6.815b)


AT2G21320.1      MRILCDACESAAAIVFCAADEAALCCSCDEKVHKCNKLASRHLRVGLADPSNAPSCDICE
gm030140_Glyma1  MRTLCDACESAAAIVFCAADEAALCRACDEKVHMCNKLASRHVRVGLASPSDVPRCDICE
gm001706_Glyma0  MRTLCDACESAAAIVFCAADEAALCRACDEKVHMCNKLASRHVRVGLASPSDVPRCDICE
gm030137_Glyma1  MRTLCDACESAAAIVFCAADEAALCRACDEKVHMCNKLASRHVRVGLASPSDVPRCDICE
gm030139_Glyma1  MRTLCDACESAAAIVFCAADEAALCRACDEKVHMCNKLASRHVRVGLASPSDVPRCDICE
gm032644_Glyma1  MRTLCDVCESAAAIVFCAADEAALCSACDHKIHMCNKLASRHVRVGLADPTDVPRCDICE
gm030138_Glyma1  MRTLCDACESAAAIVFCAADEAALCRACDEKVHMCNKLASRHVRVGLASPSDVPRCDICE
                 ** ***.******************  :**.*:* ********:*****.*::.* *****
```

Figure 1. (I) (J) (K) Gene information displayed in FloreDB. (I) Detailed gene information sourced mainly from TAIR. (J) Information about the putative orthologs and repression domain among them. (K) Information about the status of on-going CRES-T projects. Phenotypic information for *Arabidopsis* and ornamental plants. (L) Phylogenetic tree of TF family. The phenotypes of CRES-T in *Arabidopsis* can be shown. The red and pink "R"s represents genes with a putative repression domain in the coding region. (M) Phylogenetic tree of TFs displayed with microarray data. (N) Phylogenetic tree of TFs with ortholog information from five plant species. The "R" column indicates the number of species with a putative repression domain in the ortholog. The "O" column indicates the number of species with a putative ortholog of the gene. The "Total" column indicates number of orthologs with a putative repression domain compared with the total number of orthologs found in the five displayed species. The "Ia", "Ib", "II" and "III" columns indicate the number of each type of ortholog (see Materials and methods) with a putative repression domain compared with the number of each type of the ortholog found in each of the species. (O) Ortholog information displayed by clicking the hyperlinks in the "Ia", "Ib", "II" and "III" columns in (N) Relationship of hyperlink is shown by red arrows. Long arrows are shown by dashed lines and abbreviated for clarity.

Table 1. Summary of flower CRES-T project by phenotype

Phenotype	At	Tf	Cm	Pn	Gt	Cp	Rh	Dc	Eg	Total
abnormal seedling	24 (+93)									24 (+93)
hyponastic leaf	(+24)	4	3			4				11 (+24)
rough surface leaf	10 (+27)			2			2			14 (+27)
serrated leaf	3 (+24)	1	2	1		4	1			12 (+24)
other leaf abnormality	42 (+212)	5	6	8	2		1			64 (+212)
early flowering	11 (+12)		2		1				1	15 (+12)
late flowering	2 (+11)		2							2 (+11)
dwarf	6 (+63)	3	1							10 (+63)
other growth abnormality	14 (+81)									14 (+81)
abnormal stem	16 (+21)									16 (+21)
abnormal branching	14 (+51)									14 (+51)
abnormal petal	17 (+37)	24		6	1	4	3	1	2	58 (+37)
abnormal stamen	10 (+13)			1						11 (+13)
abnormal carpel	5 (+3)	1		3					1	10 (+3)
other flower abnormality	38 (+117)	18	4	8			1		1	70 (+117)
low fertility	22 (+103)			7						29 (+103)
other fruit abnormality	20 (+118)									20 (+118)
abnormal root	1 (+28)									1 (+28)
Total	255 (+1038)	56	20	36	4	12	8	1	5	397 (+1038)
Number of constructs	50 (+299)	102	36	11	23	13	4	1	7	

Except for the *Arabidopsis* bulk project, the results of all of our projects are summarized. The top heading represents the abbreviations used for the transformed plants. The numbers in the table indicate the number of constructs with gene information and the numbers in parentheses indicate the number of constructs without gene information that induce the indicated phenotype in each plant. At, *Arabidopsis thaliana*; Tf, *Torenia fournieri*; Cm, *Chrysanthemum x morifolium*; Pn, *Pharbitis nil* (Japanese morning glory); Gt, *Gentiana triflora x Gentiana scabra*; Cp, *Cyclamen persicum*; Rh, *Rosa x hybrida*; Dc, *Dianthus caryophyllus* (carnation); Eg, *Eustoma grandiflorum*.

PlnTFDB, Riano-Pachon et al. 2007; DATF, Guo et al. 2005) is displayed. From this page, FioreDB provides a number of useful direct link to external databases such as TAIR, ASRP, T-DNA express (Alonso et al. 2003), ATTED-II (Obayashi et al. 2009) and the SALAD database (Mihara et al. 2010). If the gene has a potential repression domain, its conservation among putative orthologs of other plants is also displayed (Figure 1J). For genes categorized as TFs by our original criteria, additional information that includes the progress of the projects for transgenic *Arabidopsis* CRES-T lines and their phenotypes both in *Arabidopsis* and ornamental plants, when it exists, is provided (Figure 1K). If the information regarding phenotype of CRES-T is published elsewhere, direct link to the publication is also provided (Figure 1K). Users can also browse microarray data for, for example, developmental and environmental stress sourced from a number of web sites and can perform comparative analyses on multiple genes of interest (Figure 1H).

Using FioreDB as a TF database

In addition to provide phenotypic information of CRES-T, we have developed the new FioreDB as a TF database. We compared four representative TF database described above (Mitsuda and Ohme-Takagi 2009) and redefined 1,906 *Arabidopsis* genes, classifying them into 48 TF families based on our analysis of their representation in the four databases (Table S1). The 48 families are listed on the home page of the new FioreDB (Figure 1A). Users can browse the phylogenetic tree of each TF family, prepared by aligning the full-length or conserved-domain (putative DNA-binding domain) amino acid sequences of members of the family, by clicking the hyperlinked family name on the home page (Figure 1L). If a gene has a possible repression domain in the coding region it is marked with a pink "R" in the tree. If, in addition, the domain is conserved in at least two other species, then the gene is marked with a red "R" in the tree (Figure 1L). Users can browse phenotypes that were induced in our *Arabidopsis* project (Figure 1L) or microarray data beside the tree like the SALAD on ARRAY database (Figure 1M) (Mihara et al. 2010). The new FioreDB also provides ortholog information on the trees for rice, poplar, sorghum, soybean and moss, *Physcomitrella* (Figure 1N). Information regarding sequence conservation, especially of possible repression domains, among these species can easily be obtained (Figure 1O).

Conclusion

Here we report the development of FioreDB which, compared with the previous version (Mitsuda et al. 2008), has two new features; (i) most of the phenotypic

Table 2. Summary of flower CRES-T project by gene

Alias	Locus	Expected phenotype	At	Tf		Cm		Pn		Gt		Cp		Rh		Dc		Eg	
			At	At	Tf	At	Cm	At	Pn	At	Gt	At	Cp	At	Rh	At	Dc	At	Eg
TCP3	AT1G53230	serrated leaf and petal	E	E		E				X		E	E	E					
AG	AT4G18960	double flower	E	U		U		U	E	X	U	X							
LFY	AT5G61850	no flower	E	U		X				X		X							
AtRL2	AT2G21650	change of symmetry	U	U						X				U					
FIL	AT2G45190	filamentous flower	U			X		U		X									
SEP3	AT1G24260	leaf-like flower	X	U		U		U		X									
KAN1	AT5G16560	adaxialization	U					U		X									
TCP5	AT5G60970	serrated leaf and petal	E			E		E		E		E						X	
ANAC038	AT2G24430	cotyledon fusion	E									X							
AP3	AT3G54340	loss of petal and stamen	E			X		X		X		X						X	
AP1	AT1G69120	formation of secondary flower	E	U		U				X									
MYB26	AT3G13890	indehiscent anther	E	U		X						X						X	
PHB	AT2G34710	abaxialization	U	U				U											
AP2	AT4G36920	perianth organs to carpels	U	U								X		U		U			
TT8	AT4G09820	suppression of pigment	E	X						X									
PAP1	AT1G56650	suppression of pigment	E	U						X	E							E	
PAN	AT1G68640	increased petals	E			X				X								U	
STK	AT4G09960	abnormal ovule	U	U		X				X									
WUS	AT2G17950	loss of meristem	U	U		X				X									
EIN3	AT3G20770	ethylene insensitive	E			E		X		E									
MYB23	AT5G40330	glabra	E	X		X						X							
CUC1	AT3G15170	cotyledon fusion	E	X		X													

The results for 22 representative genes are summarized. The top row represents the abbreviations used for the transformed plants and the second row indicates the abbreviations for the plant from which the used gene was derived. At, *Arabidopsis thaliana*; Tf, *Torenia fournieri*; Cm, *Chrysanthemum*×*morifolium*; Pn, *Pharbitis nil* (Japanese morning glory); Gt, *Gentiana triflora*×*Gentiana scabra*; Cp, *Cyclamen persicum*; Rh, *Rosa*×*hybrid*a; Dc, *Dianthus caryophyllus* (carnation); Eg, *Eustoma grandiflorum*; E, expected phenotype; U, unexpected or uncharacterized phenotype; X, no aberrant phenotype.

data are linked with gene information, and (ii) FioreDB can now be used as a TF database. The systematic application of CRES-T in other plants is now in progress and the data will be added to FioreDB in the future. The new FioreDB will play a central role in basic research of plant TFs and in the genetic manipulation, using CRES-T, of traits of agronomically important plants, especially from the point of view of horticulture.

Acknowledgements

The authors appreciate the contribution of all developers of the software used in FioreDB and all contributors who provided data to FioreDB. This work was supported by the Program for Promotion of Basic and Applied Researches for Innovations in Bio-oriented Industry from the Bio-oriented Technology Research Advancement Institution.

References

Alonso JM, Stepanova AN, Leisse TJ, Kim CJ, Chen H, Shinn P, Stevenson DK, Zimmerman J, Barajas P, Cheuk R, et al. (2003) Genome-wide insertional mutagenesis of *Arabidopsis thaliana*. *Science* 301: 653–657

Backman TW, Sullivan CM, Cumbie JS, Miller ZA, Chapman EJ, Fahlgren N, Givan SA, Carrington JC, Kasschau KD (2008) Update of ASRP: the *Arabidopsis* small RNA project database. *Nucleic Acids Res* 36: D982–985

Chenna R, Sugawara H, Koike T, Lopez R, Gibson TJ, Higgins DG, Thompson JD (2003) Multiple sequence alignment with the clustal series of programs. *Nucleic Acids Res* 31: 3497–3500

Davuluri RV, Sun H, Palaniswamy SK, Matthews N, Molina C, Kurtz M, Grotewold E (2003) AGRIS: *Arabidopsis* gene regulatory information server, an information resource of *Arabidopsis cis*-regulatory elements and transcription factors. *BMC Bioinformatics* 4: 25–25

Felsenstein J (1985) Confidence limits on phylogenies: an approach using the bootstrap. *Evolution* 39: 783–791

Gion K, Suzuri R, Shikata M, Mitsuda N, Oshima Y, Koyama T, Ohme-Takagi M, Ohtsubo N, Tanaka Y (2011) Morphological changes of *Rosa*×*hybrida* by a chimeric repressor of *Arabidopsis* TCP3. *Plant Biotechnol* 28: 149–152

Goda H, Sasaki E, Akiyama K, Maruyama-Nakashita A, Nakabayashi K, Li W, Ogawa M, Yamauchi Y, Preston J, Aoki K, et al. (2008) The AtGenExpress hormone and chemical treatment data set: experimental design, data evaluation, model data analysis and data access. *Plant J* 55: 526–542

Guo A, He K, Liu D, Bai S, Gu X, Wei L, Luo J (2005) DATF: a database of *Arabidopsis* transcription factors. *Bioinformatics* 21: 2568–2569

Hiratsu K, Matsui K, Koyama T, Ohme–Takagi M (2003) Dominant repression of target genes by chimeric repressors that include the EAR motif, a repression domain, in *Arabidopsis*. *Plant J* 34: 733–739

Hiratsu K, Mitsuda N, Matsui K, Ohme-Takagi M (2004) Identification of the minimal repression domain of SUPERMAN shows that the DLELRL hexapeptide is both necessary and sufficient for repression of transcription in *Arabidopsis*. *Biochem Biophys Res Comm* 321: 172–178

Iida K, Seki M, Sakurai T, Satou M, Akiyama K, Toyoda T, Konagaya A, Shinozaki K (2005) RARTF: database and tools for complete sets of *Arabidopsis* transcription factors. *DNA Res* 12: 247–256

Ikeda M, Mitsuda N, Ohme-Takagi M (2009) *Arabidopsis*

WUSCHEL is a bifunctional transcription factor that acts as a repressor in stem cell regulation and as an activator in floral patterning. *Plant Cell* 21: 3493–3505

Ikeda M, Ohme-Takagi M (2009) A novel group of transcriptional repressors in *Arabidopsis*. *Plant Cell Physiol* 50: 970–975

Itoh T, Tanaka T, Barrero RA, Yamasaki C, Fujii Y, Hilton PB, Antonio BA, Aono H, Apweiler R, Bruskiewich R, et al. (2007) Curated genome annotation of Oryza sativa ssp. japonica and comparative genome analysis with *Arabidopsis thaliana*. *Genome Res* 17: 175–183

Katoh K, Misawa K, Kuma K, Miyata T (2002) MAFFT: a novel method for rapid multiple sequence alignment based on fast fourier transform. *Nucleic Acids Res* 30: 3059–3066

Katoh K, Toh H (2008) Recent developments in the MAFFT multiple sequence alignment program. *Brief Bioinform* 9: 286–298

Kilian J, Whitehead D, Horak J, Wanke D, Weinl S, Batistic O D'Angelo C, Bornberg-Bauer E, Kudla J, Harter K (2007) The atGenExpress global stress expression data set: protocols, evaluation and model data analysis of UV-B light, drought and cold stress responses. *Plant J* 50: 347–363

Laubinger S, Zeller G, Henz SR, Sachsenberg T, Widmer CK, Naouar N, Vuylsteke M, Scholkopf B, Ratsch G, Weigel D (2008) At-TAX: a whole genome tiling array resource for developmental expression analysis and transcript identification in *Arabidopsis thaliana*. *Genome Biol* 9: R112–R112

Matsui K, Umemura Y, Ohme-Takagi M (2008) AtMYBL2, a protein with a single MYB domain, acts as a negative regulator of anthocyanin biosynthesis in *Arabidopsis*. *Plant J* 55: 954–967

Michael TP, Mockler TC, Breton G, McEntee C, Byer A, Trout JD, Hazen SP, Shen R, Priest HD, Sullivan CM, et al. (2008) Network discovery pipeline elucidates conserved time-of-day-specific *cis*-regulatory modules. *PLoS Genet* 4: e14

Mihara M, Itoh T, Izawa T (2010) SALAD database: a motif-based database of protein annotations for plant comparative genomics. *Nucleic Acids Res* 38: D835–842

Mishiba K, Nishihara M, Nakatsuka T, Abe Y, Hirano H, Yokoi T, Kikuchi A, Yamamura S (2005) Consistent transcriptional silencing of 35S-driven transgenes in gentian. *Plant J* 44: 541–556

Mishiba K, Yamasaki S, Nakatsuka T, Abe Y, Daimon H, Oda M, Nishihara M (2010) Strict *De Novo* Methylation of the 35S Enhancer Sequence in Gentian. *PLoS ONE* 5: e9670

Mitsuda N, Ohme-Takagi M (2009) Functional analysis of transcription factors in *Arabidopsis*. *Plant Cell Physiol* 50: 1232–1248

Mitsuda N, Umemura Y, Ikeda M, Shikata M, Koyama T, Matsui K, Narumi T, Aida R, Sasaki K, Hiyama T, et al. (2008) FioreDB: a database of phenotypic information induced by the chimeric repressor silencing technology (CRES-T) in *Arabidopsis* and floricultural plants *Plant Biotechnol* 25: 37–43

Mockler TC, Michael TP, Priest HD, Shen R, Sullivan CM, Givan SA, McEntee C, Kay SA, Chory J (2007) The DIURNAL project: DIURNAL and circadian expression profiling, model-based pattern matching, and promoter analysis. *Cold Spring Harb Symp Quant Biol* 72: 353–363

Narumi T, Aida R, Koyama T, Yamaguchi H, Sasaki K, Shikata M, Nakayama M, Ohme-Takagi M, Ohtsubo N (2011) *Arabidopsis* chimeric TCP3 repressor produces novel floral traits in *Torenia fournieri* and *Chrysanthemum morifolium*. *Plant Biotechnol* 28: 131–140

Obayashi T, Kinoshita K, Nakai K, Shibaoka M, Hayashi S, Saeki M, Shibata D, Saito K, Ohta H (2007) ATTED-II: a database of co-expressed genes and *cis* elements for identifying co-regulated gene groups in *Arabidopsis*. *Nucleic Acids Res* 35: D863–869

Parkinson H, Kapushesky M, Kolesnikov N, Rustici G, Shojatalab M, Abeygunawardena N, Berube H, Dylag M, Emam I, Farne A, et al. (2009) ArrayExpress update—from an archive of functional genomics experiments to the atlas of gene expression. *Nucleic Acids Res* 37: D868–872

Paterson AH, Bowers JE, Bruggmann R, Dubchak I, Grimwood J, Gundlach H, Haberer G, Hellsten U, Mitros T, Poliakov A, et al. (2009) The *sorghum bicolor* genome and the diversification of grasses. *Nature* 457: 551–556

Rasko DA, Myers GS, Ravel J (2005) Visualization of comparative genomic analyses by BLAST score ratio. *BMC Bioinformatics* 6: 2–2

Rensing SA, Lang D, Zimmer AD, Terry A, Salamov A, Shapiro H, Nishiyama T, Perroud PF, Lindquist EA, Kamisugi Y, et al. (2008) The *Physcomitrella* genome reveals evolutionary insights into the conquest of land by plants. *Science* 319: 64–69

Riano-Pachon DM, Ruzicic S, Dreyer I, Mueller-Roeber B (2007) PlnTFDB: an integrative plant transcription factor database. *BMC Bioinformatics* 8: 42–42

Sage-Ono K, Ozeki Y, Hiyama S, Higuchi Y, Kamada H, Mitsuda N, Ohme-Takagi M, Ono M (2011) Inductive double flower in *Pharbitis nil* using class-C MADS-box transcription factor with chimeric repressor silencing technology (CRES-T). *Plant Biotechnol* 28: 153–165

Saitou N, Nei M (1987) The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol Biol Evol* 4: 406–425

Schmid M, Davison TS, Henz SR, Pape UJ, Demar M, Vingron M, Scholkopf B, Weigel D, Lohmann JU (2005) A gene expression map of *Arabidopsis thaliana* development. *Nat Genet* 37: 501–506

Schmutz J, Cannon SB, Schlueter J, Ma J, Mitros T, Nelson W, Hyten DL, Song Q, Thelen JJ, Cheng J, et al. (2010) Genome sequence of the palaeopolyploid soybean. *Nature* 463: 178–183

Shikata M, Narumi T, Yamaguchi H, Sasaki K, Aida R, Oshima Y, Takiguchi Y, Ohme-Takagi M, Mitsuda N, Ohtsubo N (2011) Efficient production of novel floral traits in torenia by collective transformation with chimeric repressors of *Arabidopsis* transcription factors. *Plant Biotechnol* 28: 189–199

Swarbreck D, Wilks C, Lamesch P, Berardini TZ, Garcia-Hernandez M, Foerster H, Li D, Meyer T, Muller R, Ploetz L, et al. (2008). The Arabidopsis information resource (TAIR): gene structure and function annotation. *Nucleic Acids Res* 36: D1009–1014

Tanaka T, Antonio BA, Kikuchi S, Matsumoto T, Nagamura Y, Numa H, Sakai H, Wu J, Itoh T, Sasaki T, et al. (2008) The rice annotation project database (RAP-DB): 2008 update. *Nucleic Acids Res* 36: D1028–1033

Tanaka Y, Yamamura T, Oshima Y, Mitsuda N, Koyama T, Ohme-Takagi M, Terakawa T (2011) Creating ruffled flower petals in *Cyclamen persicum* by expression of the chimeric cyclamen TCP repressor. *Plant Biotechnol* 28: 141–147

Tuskan GA, Difazio S, Jansson S, Bohlmann J, Grigoriev I, Hellsten U, Putnam N, Ralph S, Rombauts S, Salamov A, et al. (2006) The genome of black cottonwood, *Populus trichocarpa* (Torr. & Gray). *Science* 313: 1596–1604

Zeller G, Henz SR, Widmer CK, Sachsenberg T, Ratsch G, Weigel D, Laubinger S (2009) Stress-induced changes in the *Arabidopsis thaliana* transcriptome analyzed using whole-genome tiling arrays. *Plant J* 58: 1068–1082