Comparative genomics of cis-regulatory elements in Cytokinin Response Factors across the plant taxa

by

Rachel Victoria Powell

A dissertation submitted to the Graduate Faculty of
Auburn University
In partial fulfillment of the
Requirements for the Degree of
Master of Science

Auburn, Alabama
August 4, 2018

Keywords: Cytokinin Response Factors, MEME analysis,
cis-regulatory elements, transcription factors

Approved by

Aaron Rashotte, Chair, Associate Professor of Biological Sciences
Leslie Goertzen, Associate Professor of Biological Science
Rita Graze, Assistant Professor of Biological Sciences

Abstract

Expression patterns of genes are controlled by short regions of DNA in promoter regions known as cis-regulatory elements. How expression patterns change due to alterations in cis-regulatory elements after genes duplicate are not well studied in plants. Over 300 promoter sequences from a small, well-conserved family of plant transcription factors known as Cytokinin Response Factors (CRFs) were examined for conserved motifs throughout the evolutionarily diverged clades seen in Angiosperms. Both general CRF plant family as well as distinct clade specific motifs were found. Once identified, significantly enriched motifs were then compared to known transcription factor binding sites to elucidate potential functional roles. Additionally, presence of similar motifs shows that levels of conservation exist between different CRFs across the kingdom plantae likely occurring through processes of neo- or sub-functionalization. Furthermore, significant patterns of motif conservation are seen within and between CRF clades suggesting cis-regulatory regions have been conserved through CRF evolution.

Acknowledgements

First, I'd like to thank both Drs. Aaron Rashotte and Les Goertzen, without whom this project would not have been completed. Many thanks to Aaron for helping me explore and develop my research skills and easily adapting as my path has changed throughout the last two years. I owe my newfound love for bioinformatics to Les and our weekly meetings and credit everything I can do on the command line to him. I'd also like to thank Dr. Karthik Aghoram from Meredith College for introducing me to research and guiding me through my first independent research project. Your guidance during my undergraduate years at Meredith College are what led me to where I am today.

Many thanks to my parents for their emotional (and financial) support. Thank you for answering my phone calls and helping me destress, even when you have no understanding of what I'm actually researching. Your support and love means the world to me. To all of my friends in Raleigh, thank you tremendously for your support from afar. To all of my friends in Auburn, thank you for being my friend through the crazy, giving feedback on figures, and supporting me with Diet Coke and trivia.

To everyone who has been a part of this journey- thank you!

Table of Contents

Abstract ii

Acknowledgements iii

Introduction 1

Methods 11

Collection of sequences and analysis for presence of common motifs 11

Functional categorization of identified common motifs 11

Motif conservation alignment 12

Results 13

Introduction of experimental procedure 13

Conservation of motifs within each clade 15

Conservation of motifs shared between CRF clades 17

Distinct motif conservation within *Brassicaceae* 20

Functional roles of identified conserved CRF motifs 23

Evolutionary divergence of duplicated CRF upstream sequences 27

Discussion 29

Significant conservation of cis-regulatory motifs in Cytokinin Response Factors 29

Differing motifs within all plants compared to the *Brassicaceae* alone 30

Hormone-related motif roles 31

Citations 35

Supplemental 38

List of Tables

Table 1 6

Table 2 25

List of Figures

Figure 1 4

Figure 2 14

Figure 3 19

Figure 4 22

Figure 5 26

Figure 6 28

Introduction

With the ever-changing climate and the broad range of environmental conditions a plant may face, evolutionary adaptations have arisen as response mechanisms. As master regulators, transcription factor proteins can bind to many different gene targets allowing plants to spatiotemporally control gene expression and the ability to switch on or off individual genes as well as entire regulatory cascades. This regulatory control serves to help plants adapt to their environment (Riechmann et al., 2000). Transcription factor proteins function by binding to specific short, approximately five to twenty, base pair patterns called motifs or cis-regulatory elements in upstream, intron, or downstream regions of target genes. In order to physically bind transcription factors, a chemical interaction between the amino acid side chain of the transcription factor proteins and the base pair sequence of the DNA occurs, triggering the binding and subsequent effects of the transcription factor (Davidson and Peter, 2015). In many instances, the DNA motif is able bind more than one transcription factor, with differing responses (Berk and Schmidt, 1990).

Plant cis-regulatory elements are incredibly specific allowing for distinct and differential control of gene expression dependent upon the life stage, tissue or organ, and environmental conditions. Regulation of transcription is dependent upon both presence of transcription factors and the number, location, and specific combinations of cis-regulatory elements present in the promoter region of any specific gene (Hernandez-Garcia and Finer, 2014). Additionally, transcription factor binding is highly dependent upon cell type and combinatorial effects of any additional transcription factors, co-factors, and chromatin state (Davidson and Peter, 2015; Günesdogan and Surani, 2016).

The upstream promoter sequence of every gene contains three main regions: core promoter, proximal promoter, and distal promoter. The core promoter region is in the closest proximity to the gene transcriptional start site and is responsible for the binding of RNA polymerases. Additionally, the core promoter region contains the TATA box, which is a general cis-element binding site for a range of transcription factors and histones. Most transcription factors, however, bind within the proximal promoter region, which starts after the core promoter and is approximately two hundred to two hundred and fifty base pairs upstream of the transcriptional start site. The length of the proximal promoter differs amongst genes, but the use of serial deletions across these regions have been conducted to determine required elements to maintain normal phenotypic expression of the gene. Additional cis-regulatory elements that may influence the regulation of a gene come further upstream, even up to serval kilobases away, of the proximal promoter are located within the distal promoter region. Generally, regulation from transcription factors that bind within the distal promoter region have a lesser effect than those that bind within the proximal promoter region. The enhancers or silencers, which bolster or repress expression, respectively, within the distal promoter region bind activating or repressing transcription factors, and then the DNA strand shape loops back to the core promoter region to intervene with the core promoter region to subsequently increase or decrease expression levels (Hernandez-Garcia and Finer, 2014; Lee and Young, 2000). Transcription factors are also able to bind within intron and downstream regions in similar manners to the upstream region, but the true regulatory effects are less studied.

Since the discovery of cis-regulatory elements, there has been much debate about whether the linear order of cis-regulatory elements in the promoter plays a critical role in affecting gene function. Additionally, the location or distance of cis-regulatory elements relative to the gene

they are regulating can also be a defining characteristic in terms of level of transcript regulation. When certain binding sites are located in close proximity, transcription factors can work in conjunction with other bound transcription factors to bolster their affects (Davidson and Peter, 2015). However, there are many differing opinions on whether cis-regulatory element position and orientation is more important than the simple presence of all the necessary binding sites (Hernandez-Garcia and Finer, 2014; Ludwig et al., 2005; Yáñez-Cuna et al., 2013). Regardless, a majority (86%) of all *A. thaliana* transcription factor binding sites were found to be located from 1000 base pairs upstream of the transcriptional start site to 200 base pairs downstream from the end of the coding region (Yu et al., 2016).

As organisms evolve their genomes are not static, instead genes and even whole genomes are known to duplicate. When genes duplicate there are four different potential fates: neofunctionalization, subfunctionalization, pseudogene, and conserving the original function. After a duplication, the genes would have identical roles; with neofunctionalization, the function of the duplicate gene is altered into a novel role, by accumulating advantageous mutations. Subfunctionalization maintains a subset of the original functions. Pseudogenes lose their ability to produce a functional protein due to detrimental mutations. While the fates of duplicate genes are known, the level of conservation seen in cis-regulatory regions, versus the coding sequences, remain unclear. While it is well-known that alterations in exonic gene coding sequences can have severe effects, such as a premature stop codon resulting in only a partial protein product, alterations in promoter sequences do not have as deleterious effects and are general thought of as only potentially altering patterns of gene transcription. Mutations in promoter cis-regulatory elements can lead to different spatiotemporal expression between the duplicated genes, which is described as the first step in functional differentiation between duplicated genes (Li et al., 2005;

Yáñez-Cuna et al., 2013). As distinct mutations accumulate, noticeable phenotypic changes appear between the once identical duplicated genes leading to neo- or sub-functionalization. Such a process can allow genes to become specifically localized to function during certain life stages or plant tissues (Li et al., 2005). Therefore, alterations to cis-regulatory elements, which lead to expression divergence between duplicated genes, are considered one of the primary drivers of evolution (Ferris and Whitt, 1979; Markert, 1964; Ohno, 1970).

The duplication-divergence-complementation (DDC) model explains sub-functionalization of two duplicated genes by assuming there is complementary degradation of certain cis-regulatory elements (Figure 1). After a gene duplicates, both new gene copies start out with identical cis-regulatory elements from the ancestral gene. Over time, these copies will relax conservation of specific cis-regulatory elements and those promoter regions will become selectively neutral. However, one of the two duplicate genes will maintain the conservation of those cis-regulatory regions to ensure required functional regulation of this gene. Therefore, between the two duplicated copies, all original ancestral cis-regulatory elements are present shared between the duplicated copies. Genes with more than two duplications within a genome can therefore have

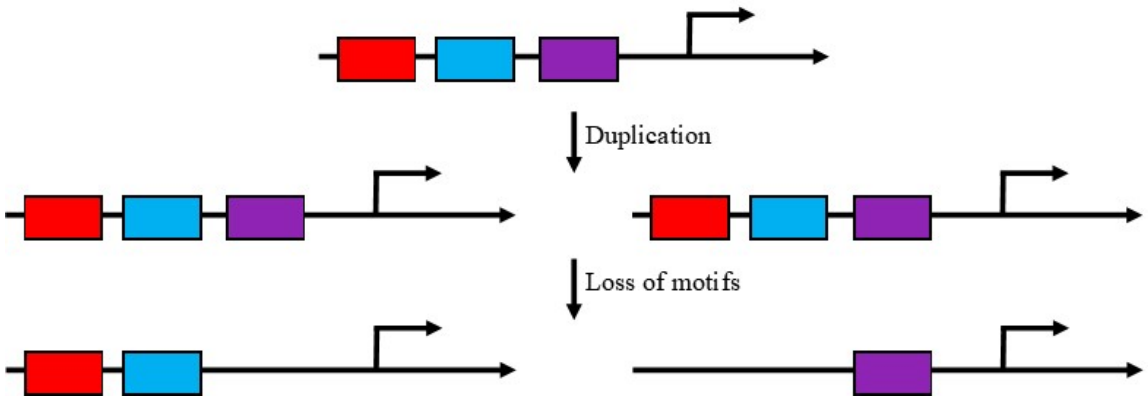


Figure 1: The DDC (divergence-deletion-complementation) model of cis-regulatory element evolution. After a gene duplicates, the two copies have complementary deletions so as to maintain all cis-regulatory elements of the ancestral gene. However, the removal of cis-regulatory elements leads to a sub-functionalization as cis-regulatory elements are critical to the spatiotemporal expression and regulation of genes.

fewer shared cis-regulatory elements due to the increased number of genes with which they have gone through complementary degradation leading to subfunctionalization (Wittkopp and Kalay, 2012). However, this is just one of the possible fates for duplicated genes, which will be examined in this thesis with a specific focus on CRFs.

Cytokinin Response Factors (CRFs) are a side branch of the cytokinin signaling pathway and a part of the AP2/ERF family of transcription factors, which are known to control a variety of developmental and environmental stress responses within all land plants (Rashotte et al., 2006; Rashotte and Goertzen, 2010). Within the AP2/ERF family, CRF1-8 were originally part of the ERF subfamily in subgroup B-5, while CRF9-12 were in subgroup B-6 (Sakuma et al., 2002). The nomenclature for ERFs was revised and CRF1-8 were placed out into Group VI, while CRF9-12 were placed as sole members within Group VI-L (Nakano et al., 2006). All genes categorized at CRFs have a conserved CRF domain at the N-terminal and the AP2/ERF domain near the middle of the gene, as well as a CRF clade specific C-terminal region (Rashotte and Goertzen, 2010; Zwack et al, 2012).

Originally, phylogenetic organization of CRF sequences from across plant taxa were divided into two distinct clades (A or B), based upon the presence (A) or absence (B) of a TEH region comprised of 13 amino acids, upstream of the canonical CRF domain (Rashotte and Goertzen, 2010). Continued phylogenetic work has indicated that the CRFs should be placed into five evolutionarily diverged groups, or clades (I, II, III, IV, V) within the Angiosperms (Supplemental Figure 2), each with unique functions, as outlined in Table 1 (Shi et al., 2012; Zwack et al., 2012). The division of CRFs into five main clades arose through gene duplications as indicated by a single CRF sequence per clade within *Amborella trichopod*, from which Clade V is placed as sister to Clades I through IV. Throughout the evolution of Angiosperms, many

duplications within individual clades resulted in one or two genes in each of Clade I, II, III, and IV per species, and Clade V having upwards of four or more CRF genes per clade (Zwack et al., 2012). The original phylogenetic work noted the duplication predates the evolutionary divergence between monocots, magnoliids, and eudicots, but after the origin of flowering land plants. The duplication events that created multiple CRF genes within each clade likely occurred prior to the diversification of Eudicots, therefore Rosids and Asterids have multiple CRF sequences per clade. In many individual species, there are additional independent CRF duplications, which is why the number of CRFs seen in each plant species varies (Rashotte and Goertzen, 2010). However, there are some CRF genes that cannot clearly be placed into any one of the five clades, due to the lack of a C-terminal protein region. One example of this is seen in *Brassicaceae*, where there are several copies of a truncated CRF gene, (labeled as "Uncladed Brassic. CRFs").

Clade	Cytokinin induced?	Regulations and Roles
I	Yes, strongly	Salt, cold, lateral root development, cytokinin (Jeon et al., 2016; Shi et al., 2014, 2012; Winter et al., 2007)
II	No	Cold, auxin, nitrogen, lateral root development (Jeon et al., 2016; Varala et al., 2018; Zwack et al., 2016)
III	Yes, strongly	Salt, oxidative stress, delayed senescence, cytokinin (Gupta and Rashotte, 2014; Rashotte et al., 2006; Zwack et al., 2013)
IV	Yes	Salt, ethylene, disease resistance, cytokinin (Gu et al., 2002; Park et al., 2001; Shi et al., 2014; Zhou et al., 1997)
V	Yes	Root and shoot growth (unpublished)
Uncladed Brassic. CRFs	No	Root development, phosphate starvation response (Ramaiah et al., 2014)

Table 1: An overview of previous research conducted on Cytokinin Response Factors, including if they are induced by cytokinin and elucidated functions for each clade. Brassicaceae lacks Clade IV sequences, but instead has a group of "uncladed" sequences, seen in their own row above.

Original research conducted on genes now classified as CRFs (then called by differing names) revealed pathogen resistance when Pti6 in tomato and Tsi1 in tobacco were overexpressed (Gu et al., 2002; Park et al., 2001; Zhou et al., 1997). Shortly after the discovery of the highly conserved CRF domain, research conducted on *A. thaliana* (At) AtCRF1- AtCRF8 showed that these are all able to form protein homo- and heterodimers with each other and also interact with Arabidopsis histidine-phosphotransferases (AHP1- AHP5), which solidified CRFs link to the cytokinin signaling pathway (Cutcliffe et al., 2011). An initial cursory cis-regulatory analysis was conducted to elucidate possible regulatory mechanisms that control CRF vascular expression. A highly conserved $(CT)^n$ motif was found multiple times within every upstream CRF sequence, which has been linked to vascular expression when found in gene promoters (Ruiz-Modrano et al., 2011; Zwack et al., 2012).

General examination of CRF expression in *A. thaliana* and *Solanum lycopersicum* (*Sl*) through Northern blot, reverse transcriptase PCR, as well as in database examinations of organ and developmental specific microarray experiments (eBAR), indicated that most CRFs are expressed in several tissues throughout the plant (Rashotte et al., 2006; Shi et al., 2012; Winter et al., 2007). Later experiments which generated promoter::GUS lines narrowed down CRF expression to primarily vascular tissue within these same tissues (Gupta and Rashotte, 2014; Shi et al., 2014; Zwack et al., 2016, 2013).

Nearly all of the direct experimental examination of CRFs has been limited to Arabidopsis (AtCRFs or often simply noted as CRFs) and tomato (SlCRFs). Results of those findings in a clade specific manner are reported here. Clade I genes showed strong expression within the vasculature of leaves, cotyledons, hypocotyls, shoot apex and roots, but there is differing expression between *At*CRF1 and *At*CRF2 within roots. In *Sl*CRF2, strong vascular expression

was present, similar to *At*CRF1 and *At*CRF2 (Zwack et al., 2012). However, *Sl*CRF2 is not limited to vascular expression and is present in leaf primordia, root tips, and flower stamen (Shi et al., 2012). Clade II *At*CRF4 was generally expressed throughout vasculature in cotyledons, hypocotyls, rosettes, and cauline leaves, inflorescence stems, sepals, petals, and primary and lateral roots. Clade I *At*CRF2 and Clade II *At*CRF4 are both expressed in the root tip.

Like Clades I and II, Clade III is also highly expressed within the vasculature tissues. However, *At*CRF5 and *At*CRF6 showed little to no expression within younger leaves but were seen as leaves began maturation starting at the distal end first and strongly in within the vasculature of fully mature leaves. Similar to Clade I, *At*CRF5 and *At*CRF6 have differing expression in the roots (Zwack et al., 2012). *Sl*CRF5 was widely present across developmental stages; expression levels were the highest in leaves of older plants and stems of younger plants and within the hypocotyl. Both *Sl*CRF5 and *At*CRF6 have increasingly stronger expression levels as plant leaves mature, but levels are subsequently reduced as leaves begin senescing (Gupta and Rashotte, 2014). Because *Arabidopsis* lacks Clade IV, tomato was used to analyze tissue expression for Clade IV. *Sl*CRF1 was expressed in the vascular tissue of roots, stems, leaves, and fruit. In leaves, unlike Clades I-III, expression was limited to the mid-vein and first- and second-order laterals; no secondary vascular tissue had strong expression of *Sl*CRF1. Additionally, no expression was found in root or shoot apices, floral tissues, and there was an overall lack of tissue specificity, unlike other CRF clades (Shi et al., 2012; Zwack et al., 2012). No GUS staining experiments have been conducted and published to date on Clade V CRF mutants. Preliminary data suggests little to no expression of CRF9 within most tissues of *A. thaliana*.

Several other noteworthy studies have been published in recent years analyzing the functional characteristics of CRFs. AHK3, a cytokinin receptor, perceives the presence of cytokinin and

subsequently phosphorylates downstream type-B ARRs. A normal feedback loop occurs between Type-A ARRs and the phosphorylated type-B ARRs. In *Arabidopsis,* CRF6, which is cytokinin induced, has been linked to a delay in senescence, due to the interaction through the cytokinin receptor Arabidopsis Histidine Kinase (AHK) 2 and 3. When CRF6 was knocked out, *A. thaliana* senesced when placed in the dark. However, when CRF6 was overexpressed, there is no upstream cis-regulatory region controlling the spatiotemporal expression of CRF6, therefore allowing for it to be consistently induced and delay senescence. Interestingly, however, the overexpression of CRF6 caused faster development compared to *crf6*, suggesting its proper expression and localization is key to normal plant development (Zwack et al., 2013).

Auxin and cytokinin work antagonistically to control plant growth and development; cytokinin influences auxin transport by changing the expression levels of auxin transporters, which subsequently alters auxin levels throughout the plant. In a study looking at the interaction between CRFs and auxin, Šimášková et al., 2015 found *At*CRF2, *At*CRF3, and *At*CRF6 transcriptionally control the expression of genes that produce PIN-FORMED auxin transporters. At the start of the signaling cascade AHK2 and AHK3 produce AHPs, which move into the nuclear membrane. Inside, AHPs cause Arabidopsis Response Regulators ARR1, ARR10, and ARR12 bind to a conserved cis-regulatory element, thus triggering the transcription of *At*CRF2. *At*CRF3 is also transcriptionally activated, but the mechanisms are not yet understood. Afterwards, *At*CRF2 and *At*CRF6 bind to a GCC box in the cis-regulatory elements of PIN1 and PIN7, which affects primary root gravitropism, lateral root initiation and formation, and lateral root primordium. Additionally, PIN transcription varies dependent upon individual CRFs, where *At*CRF3 may balance out the effects of *At*CRF2 and *At*CRF6, which is supported by previous

research on CRF interactions (Cutcliffe et al., 2011; Rashotte et al., 2006; Šimášková et al., 2015; Weirauch et al., 2014).

Mostly recently, CRF4 was found to be critical in nitrogen signaling and uptake. Within five minutes, CRF4 helps to regulate the nitrogen signaling network, which is a novel role for CRF4. Additionally, CRF4 binds to CRF2 in shoots and CRF3 and CRF4 in roots, which suggests self-regulation within Clade II CRF genes. The self-binding and binding to other CRFs indicates CRF proteins are able to bind to the promoter regions of CRF genes (Varala et al., 2018).

Although there have been several studies of CRFs in *Arabidopsis* and tomato, there has been little direct examination of CRFs outside of these two species. Therefore, it is difficult to extrapolate research currently done to plants as a whole. Research conducted in this thesis was done to broaden this scope by utilizing fifty-nine fully sequenced land plants ranging from *Embryophytes* to *Brassicaceae*, to perform a detailed analysis of cis-regulatory regions for each CRF clade. Results should provide insight into whether each clade functions in an independent manner. By using this *in silico* approach, more plant species can be examined than could otherwise be analyzed in the laboratory. Motif analyses of 1000 basepairs upstream sequences may reveal novel roles and key cis-regulatory elements for individual CRF clades and should provide help in understanding the full network of interactions of CRFs. Additionally, by using a broad sampling of plants, the evolution of cis-regulatory elements for CRFs can be elucidated.

Methods

Collection of sequences and analysis for presence of common motifs

Cytokinin Response Factors (CRF) sequences were identified via tBLASTn on Phytozome (https://phytozome.jgi.doe.gov) (Supplemental Figure 1) using known *A. thaliana* CRF sequences (Figure 2.1, Supplemental Information 1)**.** Sequences were verified as CRFs by checking for the presence of previously identified conserved CRF domain (Rashotte and Goertzen, 2010) and AP2/ERF domain within the coding sequence(Licausi et al., 2013; Mizoi et al., 2012), after which 1000 base pairs upstream of the transcriptional start site plus any 5' untranslated regions and the entire exon sequence were collected (Fig. 2.2 and 2.3). The collected exon and upstream sequences were sorted into the five distinct groupings ("clades") by searching the exon sequence on NCBI BLAST and comparing to known CRF genes (Fig. 2.4). Varying subsets, including clade groups and plant family groups, were analyzed using MEME suite tools (Bailey et al., 2015) to identify common motifs in these sequences. For each subset of upstream sequences, motifs that appeared once in all sequences (OOPS) or in most, but not all, sequences (ZOOPS) were collected. Motifs were allowed to be from 5 to 25 base pairs in length, with an E-value less than 0.05, the default parameter for MEME (Fig. 2.5). For validation, all sequences collected were shuffled and run through MEME using identical parameters, which kept the same sequence composition as the upstream sequences analyzed in this study.

Functional categorization of identified common motifs

Collected motifs were run through TomTom (Gupta et al., 2007), a part of MEME Suite, via the JASPAR Core Plants (2018) database (Khan et al., 2018) using default parameters (Fig. 2.6). For the top three results from TomTom, Uniprot IDs were collected, with a p-value significance cut

off value of 0.01 (Fig. 2.7). These Uniprot IDs were then used to collect biological GO terms for analysis of potential functions for each motif and for analysis of each of the CRF groups. PANTHER (Mi et al., 2016) was used to analyze GO terms for statistically significant overrepresentation for each clade using default parameters (Fig. 2.8).

Motif conservation alignment

For every plant containing a specific significant motif, the 5 to 25 base pair motif plus 10 base pairs on either side were loaded into Jalview (Waterhouse et al., 2009). Sequences were then aligned using Clustal (Larkin et al., 2007), within Jalview, using default parameters. Color settings were changed to highlight based on percent identity, which illustrates level of conservation based upon sequences provided for each alignment (Fig. 2.9).

Results

Introduction of experimental procedure

To conduct the motif analysis, a novel pipeline was created (Figure 2). Fifty-nine land plant species, ranging from *Embryophytes* to *Brassicaceae*, were searched for a total of 346 CRF sequences (Fig. 2.1-2.3, Table 2). For each CRF clade, a MEME run found the top five motifs (Fig. 2.5), which were then compared to known transcription factor binding sites, using Tomtom: JASPAR (Fig. 2.6). To extrapolate functional roles, UniProt and Panther GO analysis were utilized. Separately, the sequences for each motif were aligned in Jalview to identify the level of conservation at each position (Fig. 2.9). Statistically significant identified motifs had E-values ranging from 9.1×10^{-34} to 1.6×10^{-280}, while shuffled sequences all had E-values above 8.4×10^{6}.

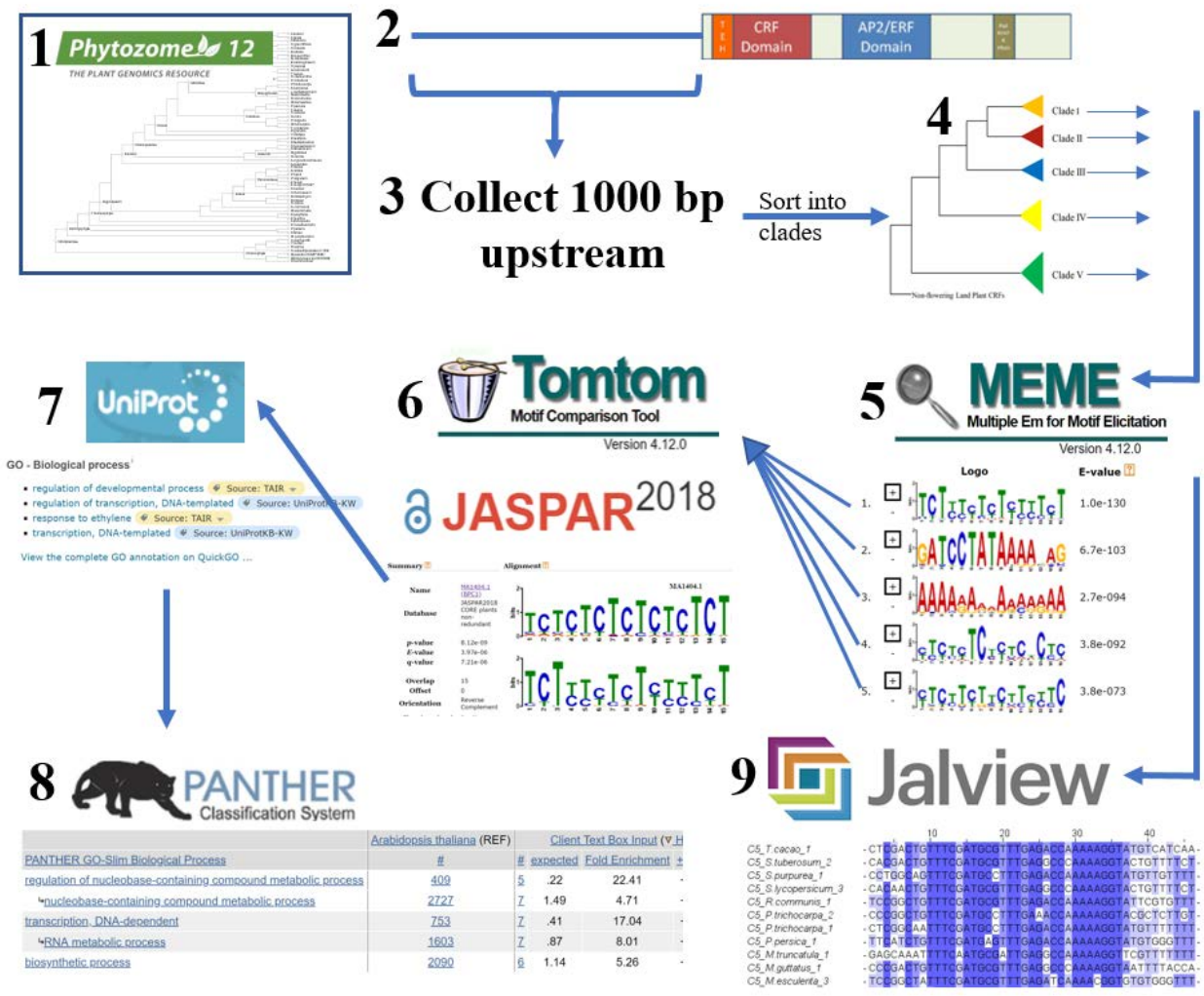


*Figure 2: A workflow overview. (1) Cytokinin Response Factor sequences were collected from Phytozome and then checked for the conserved CRF and AP2/ERF domains (2). 1000 bp upstream for each CRF sequence were collected and the amino acid sequences were compared to known A. thaliana CRF sequences (3) to place into their respective clade (4). All upstream sequences for each clade were run through MEME (5). Resulting motifs were searched through Tomtom JASPAR Core plants (6) to find matches to known motifs. For each motif, the biological GO terms were collected (7) and PANTHER GO analysis was performed (8). Additionally, an alignment for each motif was created using Jalview to show percent conserved at each position (9). *

Table 2: Plants used to isolate CRF sequences and their respective upstream sequences for this study from Phytozome.

Amaranthus hypochondriacus	Citrus sinensis	Panicum hallii
Amborella trichopod	Cucumis sativus	Panicum virgatum
Anacardium occidentale	Daucus carota	Phaseolus vulgaris
Ananas comosus	Eucalyptus grandis	Physcomitrella patens
Aquilegia coerulea	Eutrema salsugineum	Populus trichocarpa
Arabidopsis halleri	Fragaria vesca	Prunus persica
Arabidopsis lyrata	Glycine max	Ricinus communis
Arabidopsis thaliana	Gossypium raimondii	Salix purpurea
Asparagus officinalis	Hordeum vulgare	Setaria italica
Boechera stricta	Kalanchoe fedtschenkoi	Setaria viridis
Brachypodium distachyon	Kalanchoe laxiflora	Solanum lycopersicum
Brachypodium stacei	Linum usitatissimum	Solanum tuberosum
Brassica oleracea capitata	Malus domestica	Sorgum bicolor
Brassica rapa	Manihot esculenta	Sphagnum fallax
Capsella grandiflora	Marchantia polymorpha	Spirodela polyrhiza
Capsella rubella	Medicago truncatula	Theobroma cacao
Carica papaya	Mimulus guttatus	Trifolium pratense
Chenopodium quinoa	Olea europaea	Zea mays
Cicer arientinum	Oryza sativa	Zostera marina
Citrus clementina		

Conservation of motifs within each clade

Two types of MEME runs were used to elucidate motifs found in every upstream sequence in a given data set (OOPS) and motifs found within most, but not all, upstream sequences (ZOOPS). When running each CRF clades individually, the most common motif result is a $(CT)^n$ or $(GA)^n$, which is found in all clades, in every sequence, except for the uncladed *Brassicaceae* sequences. The examination of Clade I found that four of the top five motifs were either $(CT)^n$ or $(GA)^n$ for OOPS and ZOOPS MEME runs. The remaining top motif for Clade I OOPS is also returned with Clade I ZOOPS settings, with a conserved pattern of "GATCCTATAAA" and a noticeable lack of conservation flanking either side. The remaining Clade I OOPS motif has stronger

conservation across the 25 bp motif, with the main motif pattern of “TCACGTGAC”. E-value for Clade I OOPS motifs range from $2.4x10^{-107}$ to $8.1x10^{-240}$ and ZOOPS values from $2.3x10^{-81}$ to $6.9x10^{-168}$.

In the examination of Clade II OOPS MEME results, only one of the top five motifs returned was $(CT)^n$ or $(GA)^n$ with a strong E-value of $1.3x10^{-207}$, the second highest value. The top motif is seen in both Clade II OOPS and ZOOPS, with a strong nucleotide pattern of “GATCCTATAA” followed by a degradation of nucleotide conservation, with E-values of $1.1x10^{-243}$ and $2.0x10^{-215}$. The top motif for Clade II ZOOPS has conservation seen throughout the 25 bp motif, with eight nucleotides strongly conserved with no variability, “TCACGTGA”. Of the 85 Clade II sequences given, the top motif was seen in 44 of the sequences with an E-value of $1.6x10^{-280}$. Clade II ZOOPS had two other motifs with noticeable conservation, one featuring “ATGYGGCG” with significant degradation of conservation flanking either side, and the other motif having much stronger conservation and a conserved core of “CTGANTCAGCA”.

Much like CRF Clade I and II upstream sequences, Clade III OOPS and ZOOPS runs have $(CT)^n$ or $(GA)^n$ motifs in the top five motifs, ranked first and second, respectively. The top motif found in Clade III upstream CRF sequences has weak conservation followed by a relatively conserved core of “RARAWGCGGMNAGYCGYY” with a strong E-value of $1.9x10^{-107}$. A less conserved, but still nearly identical, version of the motif is seen in the OOPS run and ranked third with an E-value of $4.2x10^{-69}$, meaning that all Clade III CRF upstream sequences have this motif. The second and third ranked motifs for Clade III OOPS and ZOOPS were also identical, “TTNCTTGG” followed by several non-conserved nucleotides and “RYCAAG” at the other end. The most conserved motif seen in Clade III upstream sequences, “CNTTTTGACTCTTC”, was ranked fifth and is seen in 65% of Clade III sequences. The fifth ranked motif seen in all Clade

III sequences was a run of 16 A's with lower conservation at 3 positions and an E-value of $9.1x10^{-34}$.

Clade V motifs had the strongest E-values, ranging from $4.2x10^{-124}$ to $1.8x10^{-279}$ for ZOOPS motifs and $4.2x10^{-97}$ to $1.2x10^{-231}$ for OOPS motifs. For both Clade V ZOOPS and OOPS, the fourth ranked motif was the highly conserved $(CT)^n$ or $(GA)^n$ motif seen in all clades. Interestingly, the motifs called for both OOPS and ZOOPS was identical, seen in every Clade V upstream sequence. While other clades had $(CT)^n$ or $(GA)^n$ motifs in both OOPS and ZOOPS, they varied slightly. In the remaining four motifs, an overall pattern of regions of conservation connected by several bases with no conservation of base pairs at each site.

When looking exclusively at the upstream sequences of the *Brassicaceae*, uncladed CRFs had motifs with E-values that were overall weaker than Clades I-V, ranging from $6.5x10^{-4}$ to $2.7x10^{-23}$ for OOPS and $7.2x10^{-21}$ to $2.7x10^{-23}$ for ZOOPS. The motifs and patterns within *Brassicaceae* are rather unique compared to other clades and are discussed in more detail in subsequent sections.

Conservation of motifs shared between CRF clades

When looking across all plant taxa samples, three instances arose where identical motifs were independently identified in two different clades, indicating they have been conserved from the duplication events from which CRFs arose (Fig. 3). Two of the motifs were seen in all Clades I and II sequences. Figure 3A was the second most statistically significant motif for both Clade I and Clade II results, with an E-value of $6.4x10^{-136}$ and $1.1x10^{-243}$, respectively. Additionally, Clade II sequences have a higher level of conservation flanking either side of the motif in both Figure 3A and B. In Figure 3C, almost all sequences feature a "GC" at position 30 and 31,

except for the Clade II *Brassicaceae* sequences. *Brassicaceae* are uniquely derived and have many unique motifs compared to the MEME runs comprised of all plant taxa within an individual clade.

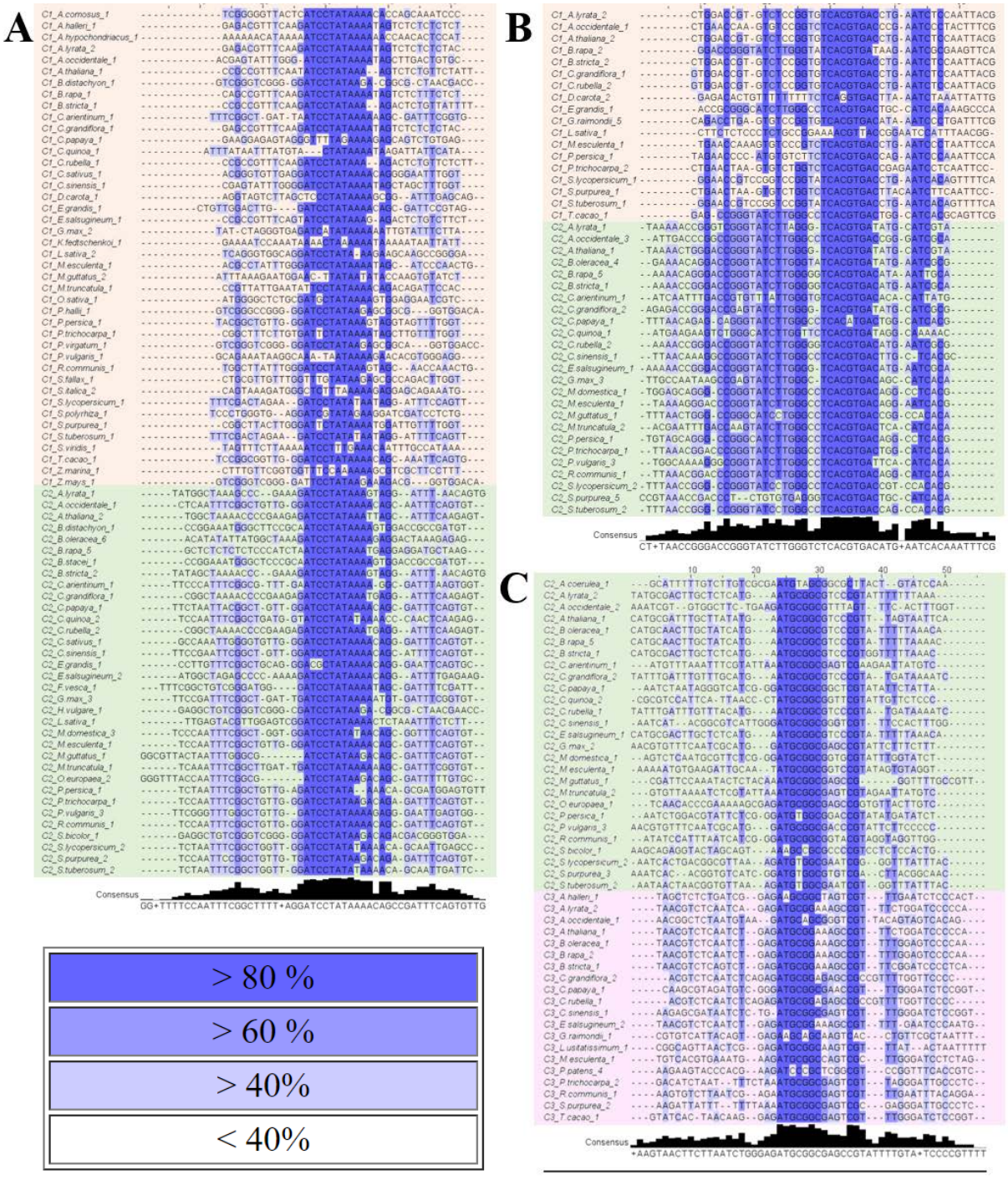
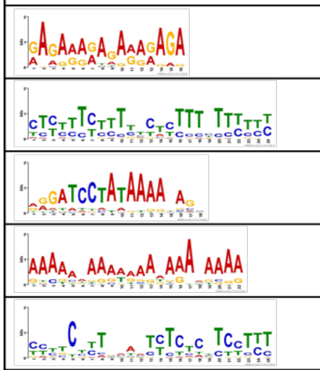
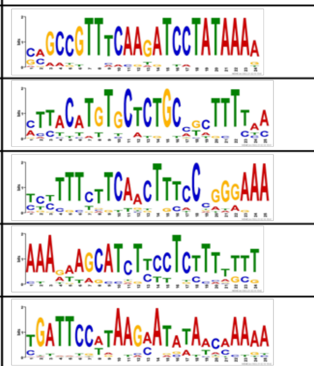
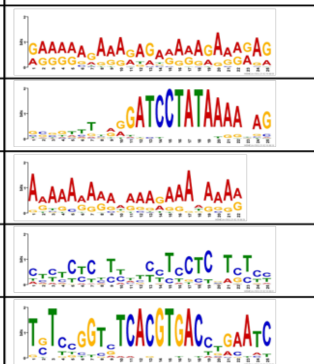
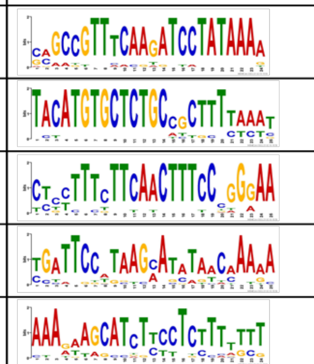
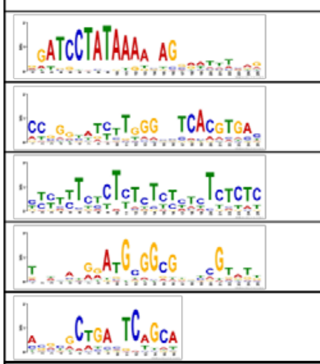
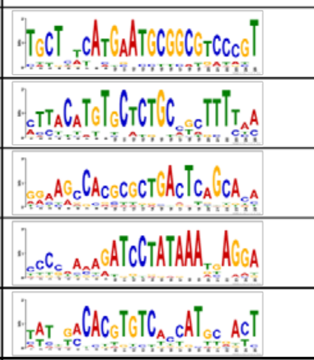
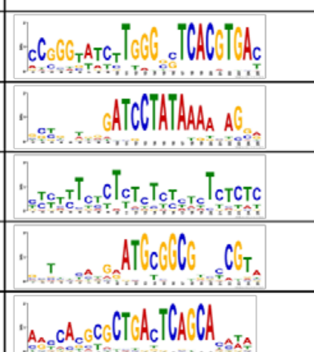
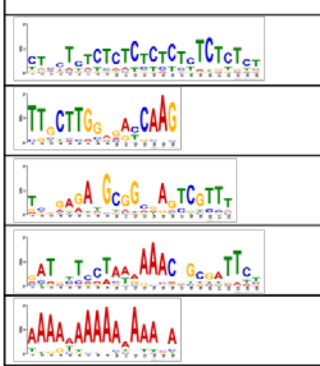
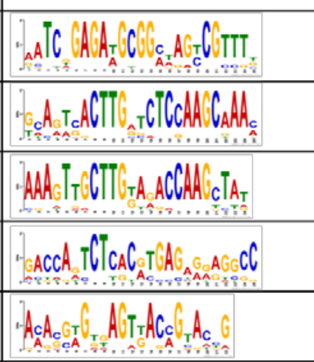
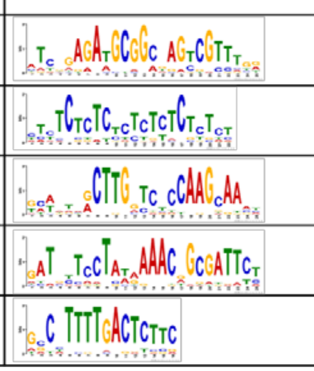
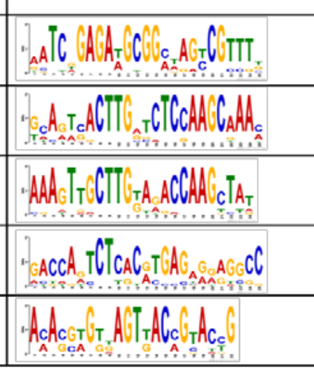


Figure 3: Alignment illustrating the percent conservation at each site. The darker the shade of blue, the higher the level of conservation at each site, with the most common color above representing more than 80% of sequences have the same nucleotide at that position. Both (A) and (B) are motifs seen in Clade I (orange) and Clade II (green) CRF sequences and are in the top five motifs returned by MEME. (C) is a motif seen in both Clade II and Clade III (pink) CRF sequences.

Distinct motif conservation within *Brassicaceae*

A majority of work conducted on CRFs has been on *A. thaliana*, creating difficulties when extrapolating results to other plants. In order to analyze the differences in cis-regulatory regions between the entire plant taxa and just *Brassicaceae*, to which *A. thaliana* is a member, *Brassicaceae* alone were analyzed using the pipeline (Fig. 2). While the conservation of cis-regulatory motifs seen within and even between clades for all land plants is significant, the conservation of cis-regulatory motifs seen exclusively in *Brassicaceae* species is even more considerable. The motifs found when evaluating based on forcing each plant species to have every motif (OOPS) typically results in having a few nucleotides of strong conservation surrounded by regions of nucleotide variation (Figure 4). The Clade I OOPS motifs for the all plant species MEME runs were primarily TCTC repeats. However, when just *Brassicaceae* were analyzed for motifs, only one of the top five motifs was a TCTC repeat, indicating more complex motif patterns are being conserved rather than this simple repeat. Evolution has allowed for the divergence and mutations of cis-regulatory regions that result in non-critical regions disintegrating into regions with no nucleotide conservation. However, when running MEME using exclusively *Brassicaceae* CRF upstream sequences, the motifs are characterized by stronger nucleotide conservation at each site with significantly less wobble or nucleotide variation seen compared to previous runs using all plants. Clades I, II, III and V CRF upstream sequences all have regions of considerable nucleotide conservation (Fig. 4). Additionally, the motifs illustrate the differences between Clade IV and the uncladed *Brassicaceae* CRFs.

Clade I OOPS	Clade I Brassicaceae OOPS	Clade I ZOOPS	Clade I Brassicaceae ZOOPS
			
Clade II OOPS	Clade II Brassicaceae OOPS	Clade II ZOOPS	Clade II Brassicaceae ZOOPS
			
Clade III OOPS	Clade III Brassicaceae OOPS	Clade III ZOOPS	Clade III Brassicaceae ZOOPS
			

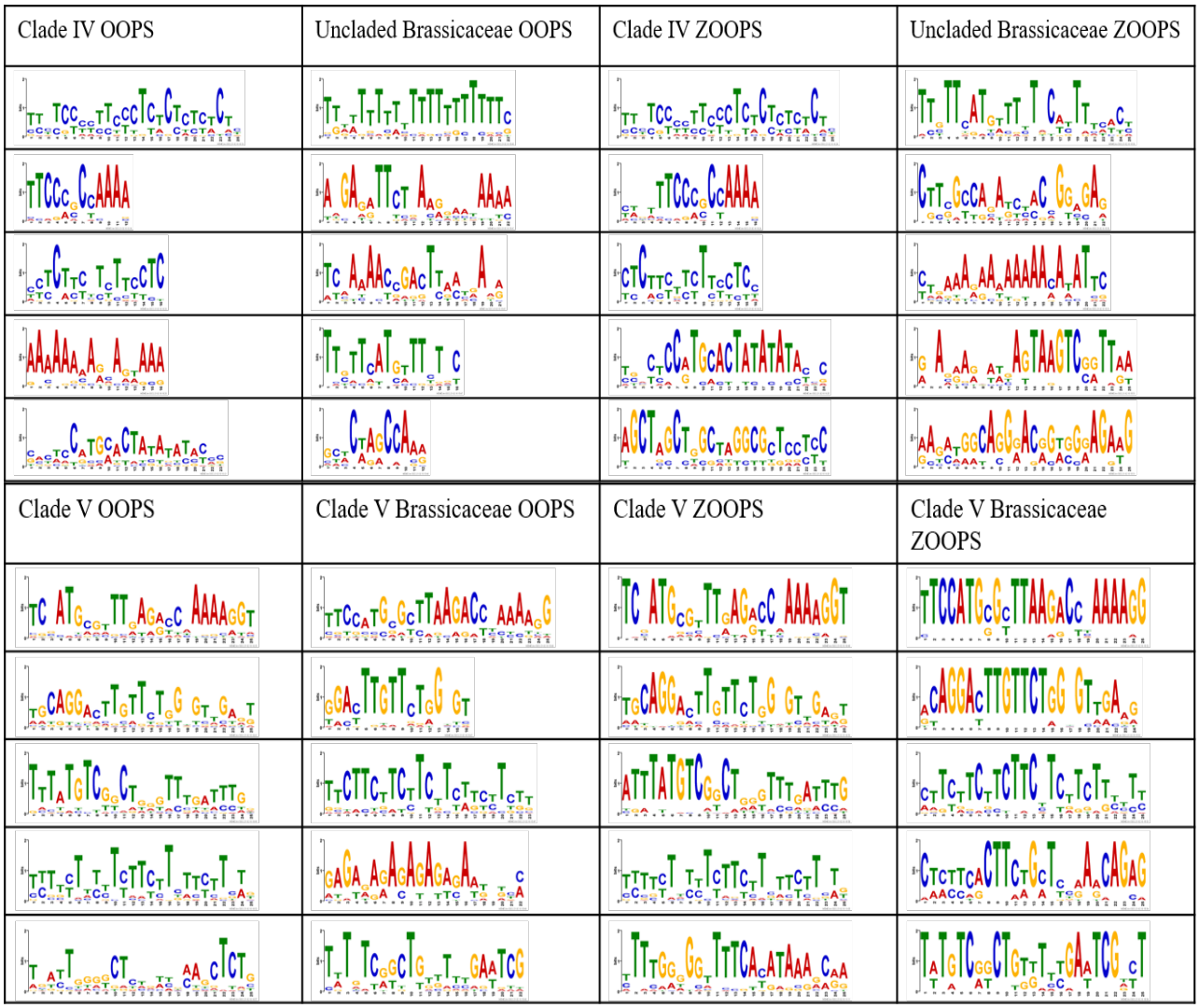


Figure 4: Comparison of motifs between all plants and only Brassicaceae, where every CRF sequence given must have every motif MEME finds (OOPS) and allowing for some sequences not to have every motif found (ZOOPS). The level of conservation varies drastically between the two different groups, as noted by the decreased frequency where a position can wobble between nucleotides at that given site.

Functional roles of identified conserved CRF motifs

The top five motifs that were identified as previously described for each individual clade were further analyzed by utilizing Tomtom, Uniprot, and Panther DB were used determine how found motifs correspond to known transcription factor binding sites and identify any potential functions as previously determined experimentally (Table 2, Supplemental Table 1). Transcription factor binding sites tended to fall into six functionally based categories: hormone, development, flowering/leaf senescence, light, transcription regulation, or stress response (Figure 5). We found the category of hormone-related function was the greatest for all motifs on average of 58% in Clades I-V (Fig. 5). All motifs examined were found as linked to the regulation of transcription, however, that is expected given they are transcription factor binding sites.

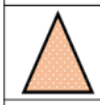
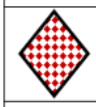
A manual in-depth analysis of Clade I motifs indicates that expression of Clade I CRF sequences should be responsive to ethylene, abscisic acid, and brassinosteroid. Additionally, motifs were found that have been connected to root and overall plant development, along with flowering time and stomatal movement. Of the fifteen motifs analyzed for Clade I, 52% related to hormones and 25% influenced or helped to regulate development (Fig. 5). Clade II shows large amounts of overlap with Clade I motifs, with three of the top five motifs featured in both clades. Clade II, similar to Clade I, has 54% of motifs relating to hormones (Fig. 5). Unique to Clades II and III are another motif that aids in the regulation of root growth as well as stomatal movement, potentially bolstering the affect Clade II sequences could have on stomatal movements and root growth. The only motif seen exclusively in Clade II aids in RNA splicing and processing, DNA repair, regulation of cell cycle and differentiation, along with the response to auxin, bacteria, fungus, and nematodes. Nearly a quarter of all motifs (23%) for Clade II aid with stress response, the highest of Clades I-V (Fig. 5). The motifs in Clade III indicated response to many of the

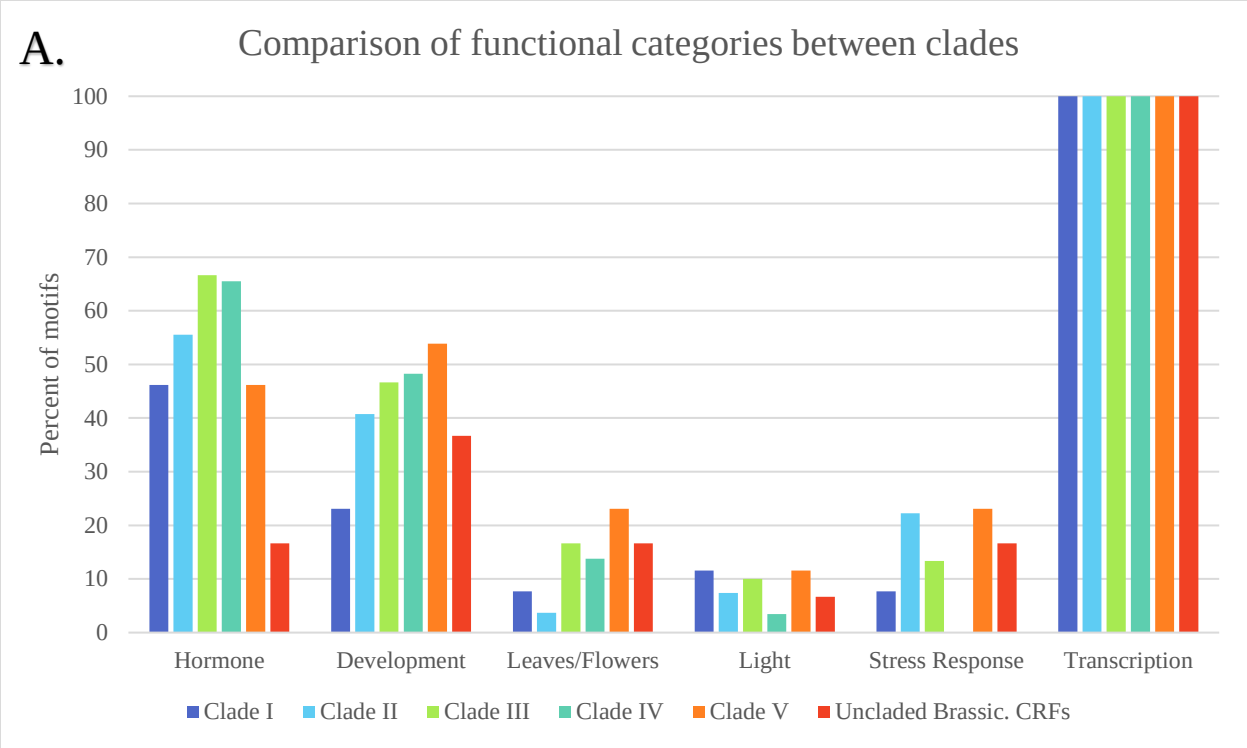
major hormones including cytokinin, ethylene, auxin, and abscisic acid. Clade III motifs have a stronger developmental influence, with 46% of motifs relating to development, as well as, 73% and 20% motifs influencing development or flowering/leaves, respectively (Fig. 5).

Clade IV motifs had the largest percentage where the only biological processes listed related strictly to transcription. As a result, no motifs were found to relate to light or stress response. Of the remaining motifs, 66% interacted with one of the major plant hormones, slightly above average, and 40% influenced development (Fig. 5). Ethylene was the regulated hormone in 75% of the hormone motifs, with abscisic acid and auxin being regulated in the remaining 25% of motifs. Unique to Clade V motifs are the regulation of gibberellic acid (GA) along with the development of many plant organs, with development making up the top percentage of categories at 66% (Fig. 5). Hormone and leaves/flowering are both seen in 42% of motifs, which is the highest percentage for leaves/flowering compared to Clades I-IV (Fig. 5). Motifs seen in Clade V lack any overlap with other clades, except for the top motif seen in all CRF sequences, which is expected given it is sister to the other CRF clades.

When all CRF upstream sequences were run together, 54% of motifs related to hormone regulation, slightly below the overall average of 58%. Development was seen in 27% of motifs, while flowering/leaf senescence, light, and stress response were all only seen in 9% of potential transcription factor roles. Matching Clades I-V, ethylene is overwhelmingly the most commonly seen hormone regulated, with cytokinin seen once, and no other plant hormones seen in resulting transcription factor binding site matches (Fig. 5).

*Table 3: A summary of functions found for each motif, represented by a symbol. Functions were broken down into "regulates", "response to", and "development of" categories. An * indicates that motif is present in both OOPS and ZOOPS MEME runs, meaning all sequences in that respective clade has that motif.*

Symbol	**Potential Functions**	Clades
	**Regulates:** development, flowering time, photoperiodism, transcription; **response to:** ethylene	ALL*
	**Regulates:** transcription, cell size, stomatal movement, senescence, pectin, root development, cell differentiation	1*, 2*
	**Regulates:** transcription, ABA and BR signaling	1, 2
	**Regulates:** root and seed growth, stomatal movement, transcription, circadian rhythm, senescence; **response to:** cytokinin, ethylene, JA, red/far red light	2, 3
	**Regulates:** transcription, RNA splicing, DNA repair, mRNA processing, cell cycle and differentiation; **response to:** auxin, bacteria, fungus, nematodes	2
	**Regulates:** transcription, heterochrony, and development; **response to:** ABA and auxin; retrograde signaling, plastid organization, cell differentiation	3
	**Regulates:** transcription, nutrient levels, cytokinin-/ABA-activated signaling; **response to:** abscisic acid, nitrogen, phosphate starvation; primary root development	3
	**Regulates:** transcription and defense response; **response to:** chitin, bacterium, salicylic acid, wounding, fungus; JA mediated signaling	3
	**Regulates:** auxin and GA synthesis, transcription, seed maturation, oilbody biogenesis; **response to:** water, cold, (red) light stimulus; seed germination, cell fate, red light; **development of:** flowers, ovules, and embryos; specification of floral organ identity and number	5
	Predicted protein	5
	**Regulates:** GA synthesis, transition from vegetative to reproductive phase; **response to:** cold and auxin	5
	**Regulates:** floral organ abscission, floral meristem identity, seed maturation, short-day photoperiodism, transcription, starch synthase SS4; **response to:** auxin stimulus, floral organ abscission, fruit abscission, fruit dehiscence, GA catabolism; **development of:** chloroplast, petals, carpels, sepals, stamen, flowers, embryos	5



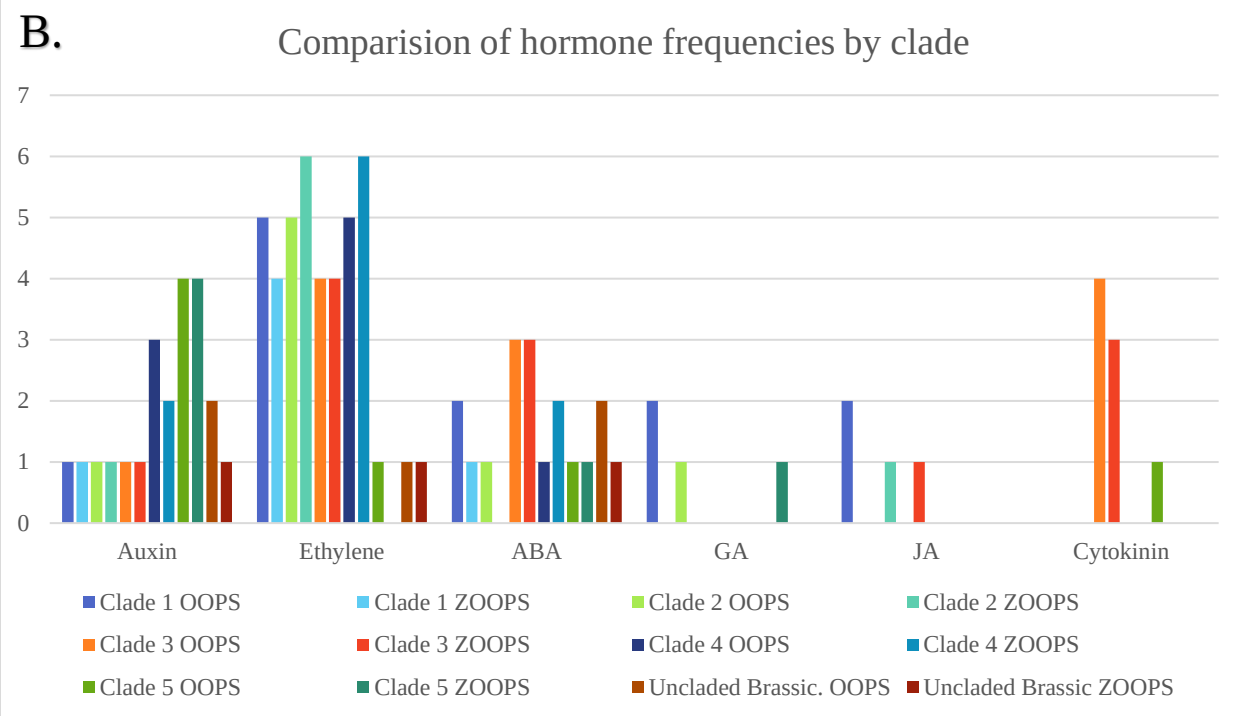


Figure 5: (A) The six main categories the roles for each ZOOPS motif matched to, based upon Tomtom, Uniprot, and Panther GO analysis, based upon frequency in the top 15 Tomtom results for each Clade sequences. Motifs could be classified into more than one category, as transcription factors often have many roles and the top three matches for each motif were examined. (B) A breakdown of which hormones are seen most commonly within each clade and overall for both OOPS and ZOOPS.

Evolutionary divergence of duplicated CRF upstream sequences

In order to examine the potential change in cis-element motif patterns, duplicated CRF genes from individual species were identified and compared. While this type of examination is regularly conducted on motifs or domains of protein sequences, it is rarely done for promoter sequences possibly because promoter sequences and intragenic are not thought to be as conserved as exon coding sequences. Despite this, our analysis of 346 CRF sequences from 59 plant species revealed an abundant conservation and divergence of cis-element motifs, which fall into four main patterns of duplication events. In two of the duplication event patterns, the upstream sequences of duplicated CRF genes have the same top five motifs present: one having motifs in the same order, while the other is in a different order (Figures 6B and 6C). In the third identified pattern, each duplicated promoter sequence has a unique motif not seen in the other copy, while the remaining motifs are identical between the copies. This pattern could indicate a sub- or neo-functionalization between CRF gene copies, as transcription factor binding sites in the upstream sequence can play large roles in gene identity. In the fourth identified duplication event pattern, the duplicated promoter sequences have serial or complementary deletions, similar to the DDC model that together make up what is likely the “original” motif pattern (Fig. 6A). From the patterns that emerged, those having identical sequence motifs (Fig. 6B) or serial deletions of motifs (Fig. 6A) are the most commonly found. In plant species that had more than the average number of CRF sequences per clade, more than one pattern was commonly identified, indicating there is not necessarily one evolutionary model that is at work in cis-regulatory regions. However, regardless of the pattern, the statistical significance of each motif and the level of pattern conservation between plant species indicates cis-regulatory motifs can and should be examined for duplication and divergence.

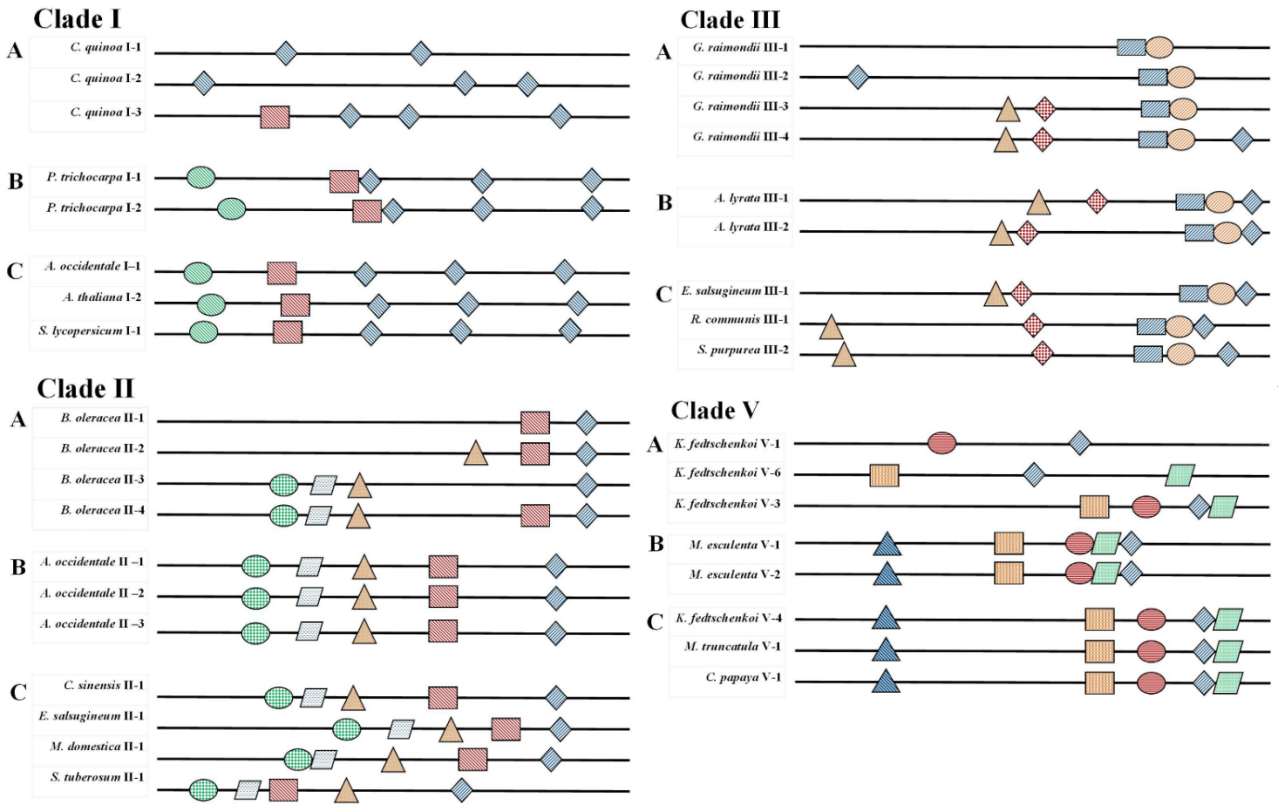


Figure 6: All (A) panels illustrate evolutionary divergence in duplicated CRF sequences. (B) sequences are duplicated CRF sequenes where both copies have retained matching motifs. (C) a collection of different plant species, all with matching copies of motifs found for their respective clade. All shapes based upon Table 3.

Discussion

Cis-regulatory motifs are evolutionarily conserved in Cytokinin Response Factors

Prior to this study, limited research had been conducted analyzing the evolutionary conservation of cis-regulatory motifs within a highly duplicated gene family. While done extensively on protein sequences, the notion that cis-regulatory regions are not highly conserved combined with the lack of fully sequenced genomes seems to have limited researchers interest from conducting such efforts. However, this study proves motif analyses on cis-regulatory elements should be conducted as they provide valuable information about the regulation of genes. In addition to the importance of the motifs found as part of this research pipeline to the investigation of CRF gene function (Fig. 2), the level of conservation identified in this study alongside of duplication and divergence of cis-regulatory motifs indicates the need for additional broad research focused on cis-regulatory regions of duplicated genes of many different gene families.

The duplications of CRF genes has spanned across evolutionary time originating with an ancestral CRF found in species after the emergence of plants on land and becoming widespread within Angiosperms. This expansion has also occurred through individual gene duplications within a single plant species, which lend cis-regulatory regions to high levels of mutation and divergence, since intragenic regions are not as highly conserved as genetic coding sequences. The simple principle of having at least five evolutionary preserved motifs of up to twenty-five base pairs shows a significant level of conservation of the upstream region similar to that found in the coding region of CRFs.

Even more potentially significant are the motifs which have been conserved between Clades I and II and Clades II and III (Fig. 3), as these duplications occurred in the evolutionary timeframe ranging from the origin of Angiosperms to present day and illustrate the level of conservation

seen within the cis-regulatory motifs found in this study. By analyzing these differing motifs present within Clade I and II and Clade II and III, the divergence of cis-regulatory elements between the clades occurring after their duplication can be elucidated. This pipeline created for this study of CRFs could easily be adapted and applied to the upstream and downstream cis-regulatory regions of other groups of conserved duplicated genes to analyze how the cis-regulatory motifs have been conserved or diverged across an evolutionary landscape.

Differing motifs within all plants compared to the *Brassicaceae* alone

While the motif conservation of cis-regulatory elements seen across all land plants is significant considering evolutionary forces, an examination specifically focused on the *Brassicaceae* revealed an even stronger level of conservation present in the cis-regulatory elements found in this group. Due to the high level of conservation seen in the *Brassicaceae*, a dichotomy arises as this same level of conservation is not present when comparing the Brassica-specific elements to those found in other plant families. As such, this creates issues when trying to extrapolate research results from well-studied members of the *Brassicaceae*, specifically *A. thaliana*, to other plants in other families. When trying to determine the functional role of a specific unstudied CRF gene within any plant, the upstream motifs are a critical component to consider, as transcription factors influence the spatiotemporal expression and regulation of said gene. If the upstream regions, where transcription factors bind, are drastically different between plant species, genetic differences should be expected.

Interesting differences were found when examining functional roles attributed to specific CRF clades from *Brassicaceae* alone sequences run through the pipeline versus all plant species create (Fig. 5 and Supplemental Figure 3). Specific differences could be seen within the breakdown of a hormone functional attributed role. When attributing a functional role for the plant hormone

ethylene from all plant sequences (Fig. 5), this was found most commonly in Clades II and IV sequences but not seen in Clade V sequences. However, a similar examination analyzing sequences from the *Brassicaceae* identified ethylene as most common attribute in Clade V sequences. Likewise, for the plant hormone ABA, Clade II sequences were found to not have any motifs related to ABA, but in *Brassicaceae* Clade II had the highest number of motifs relating to ABA. These results do not indicate that there are no motifs for these hormones to be found within the examined 1000 base pairs upstream region, just that the motifs are not within the five highest ranking motifs decided by MEME. While such differences may be expected due to evolutionary selection forces unique to the *Brassicaceae* among plants, this difference suggests the need for researchers to branch out of experimental studies focused on the model system, *A. thaliana*, and expand into less traditional plants in a broader range of different plant families.

Hormone-related motif roles

Previous research has shown CRF Clades I, III, IV, and V genes are inducible by cytokinin and the MEME motifs support this result (Gupta and Rashotte, 2014; Rashotte et al., 2006; Shi et al., 2012; Zwack et al., 2013). Within the top five motif results, both Clade III and Clade V sequences have motifs that indicate cytokinin-related transcription factors can bind. While the top five motifs for Clade I sequences did not have a motif relating to cytokinin, within the top ten motifs one was very similar to the canonical GCC box, falling into the AP2/ERF transcription factor family and is known to aid with the response to cytokinin (Supplemental Figure 4) (Fujimoto et al., 2000; Rashotte et al., 2006; Weirauch et al., 2014). Not every motif found by MEME is necessarily a transcription factor binding site and motifs not within the top five

MEME results are still worth looking at for matching motifs to previously conducted experimental results.

The remaining five plant hormones, auxin, ethylene, abscisic acid (ABA), gibberellic acid (GA), and jasmonic acid (JA), support the intricate and highly complicated interaction between plant hormones. CRFs are a side branch of the cytokinin signaling pathway and it is expected they would share an equally complex interaction with the other plant hormones. Of the major plant hormones, auxin is the only plant hormone seen in every MEME run, OOPS and ZOOPS, in every clade and the uncladed *Brassicaceae* CRF sequences. Auxin and cytokinin are antitheses, working in conjunction to regulate many plant processes. While no research has been conducted to understand how auxin influences CRFs, an abundance of research has been conducted linking auxin and cytokinin and the interaction with CRFs and known auxin transporters (Šimášková et al., 2015).

Ethylene and ABA were both seen in all MEME runs, except one, Clade V ZOOPS and Clade II ZOOPS respectively. Recall, however, that ZOOPS allows for not all upstream sequences to have each motif, while OOPS requires all upstream sequences to have every motif. Therefore, Clade V ZOOPS and Clade II ZOOPS MEME runs not having ethylene and ABA-related motifs, respectively, is not significant as Clade V OOPS and Clade II OOPS both had these motifs. To test the true interaction between the major plant hormones and CRFs, experiments utilizing overexpressing and knock out plants and treating them with doses of each hormone will help to understand if there is any interaction between transcription factors with known roles and regulations with each hormone and CRFs.

Development-related motif roles

Development was the second most frequent category for motif roles to be classified in, which is one of the main functions of AP2/ERF transcription factors (Licausi et al., 2013). Several papers published recently have analyzed the phenotypic effects of both overexpressing and knock out lines of CRFs. *At*CRF1, *At*CRF2, *At*CRF3, *At*CRF5, and *At*CRF6 have all been linked to help regulate primary and lateral root and shoot growth, embryo development, leaf senescence, hypocotyl elongation, and rosette size (Jeon et al., 2016; Raines et al., 2016; Šimášková et al., 2015; Zwack et al., 2013), indicating that the proper spatiotemporal expression of CRF genes are critical for normal plant development. The transcription factors that bind to the upstream regions of the CRF sequences analyzed in this study play a key role in maintaining spatiotemporal expression of each CRF, and therefore, helping to regulate normal development.

Conclusions and future directions

Through the creation of this novel pipeline, promoter regions can be analyzed for cis-regulatory elements to indicate potential functions of a gene and reveal conservation or divergence of cis-regulatory elements. Over 300 promoter regions for CRF genes were collected and analyzed with MEME Suite to find numerous conserved cis-regulatory elements. The motifs for each individual CRF clade can be examined for patterns of conservation, sub-, or neo-functionalization. Many of the previously conducted experiments on CRFs support the motifs elucidates though this novel pipeline, bolstering the significance. The conservation of motifs within each CRF clade, even when species span across Angiosperms, is considerable, especially given its long been thought cis-regulatory regions are not conserved like coding regions. However, further confirmation is still necessary.

Validation of what transcription factors are binding to each motif should still be conducted using a combination of both serial deletion of promoter regions and bioinformatic techniques to capture transcription factors binding to cis-regulatory regions. By using a variety of Angiosperms, the evolutionary changes of cis-regulatory regions can continue to be understood and later applied to other plant gene families. Additionally, the novel pipeline can be applied to the downstream regions relative to CRF gene coding regions. All motifs, upstream and downstream, can then be directly examined for potentially novel functionally predicted roles for CRFs. Outside of the CRF transcription factor family, this innovative pipeline can be applied to upstream, downstream, or intron regions of gene for families within plants or families outside of plants.

Citations

Bailey, T.L., Johnson, J., Grant, C.E., Noble, W.S., 2015. The MEME Suite. Nucleic Acids Res. 43, W39–W49. https://doi.org/10.1093/nar/gkv416

Cutcliffe, J.W., Hellmann, E., Heyl, A., Rashotte, A.M., 2011. CRFs form protein–protein interactions with each other and with members of the cytokinin signalling pathway in Arabidopsis via the CRF domain. J. Exp. Bot. 62, 4995–5002. https://doi.org/10.1093/jxb/err199

Davidson, E.H., Peter, I.S., 2015. Chapter 1 - The Genome in Development, in: Genomic Control Process. Academic Press, Oxford, pp. 1–40. https://doi.org/10.1016/B978-0-12-404729-7.00001-0

Ferris, S.D., Whitt, G.S., 1979. Evolution of the differential regulation of duplicate genes after polyploidization. J. Mol. Evol. 12, 267–317. https://doi.org/10.1007/BF01732026

Fujimoto, S.Y., Ohta, M., Usui, A., Shinshi, H., Ohme-Takagi, M., 2000. Arabidopsis Ethylene-Responsive Element Binding Factors Act as Transcriptional Activators or Repressors of GCC Box–Mediated Gene Expression. Plant Cell 12, 393–404. https://doi.org/10.1105/tpc.12.3.393

Gu, Y.-Q., Wildermuth, M.C., Chakravarthy, S., Loh, Y.-T., Yang, C., He, X., Han, Y., Martin, G.B., 2002. Tomato Transcription Factors Pti4, Pti5, and Pti6 Activate Defense Responses When Expressed in Arabidopsis. Plant Cell 14, 817–831. https://doi.org/10.1105/tpc.000794

Günesdogan, U., Surani, M.A., 2016. Chapter Twenty-Eight - Developmental Competence for Primordial Germ Cell Fate, in: Wassarman, P.M. (Ed.), Current Topics in Developmental Biology, Essays on Developmental Biology, Part B. Academic Press, pp. 471–496. https://doi.org/10.1016/bs.ctdb.2015.11.007

Gupta, S., Rashotte, A.M., 2014. Expression patterns and regulation of SlCRF3 and SlCRF5 in response to cytokinin and abiotic stresses in tomato (Solanum lycopersicum). J. Plant Physiol. 171, 349–358. https://doi.org/10.1016/j.jplph.2013.09.003

Gupta, S., Stamatoyannopoulos, J.A., Bailey, T.L., Noble, W.S., 2007. Quantifying similarity between motifs. Genome Biol. 8, R24. https://doi.org/10.1186/gb-2007-8-2-r24

Hernandez-Garcia, C.M., Finer, J.J., 2014. Identification and validation of promoters and cis-acting regulatory elements. Plant Sci. 217–218, 109–119. https://doi.org/10.1016/j.plantsci.2013.12.007

Jeon, J., Cho, C., Lee, M.R., Binh, N.V., Kim, J., 2016. CYTOKININ RESPONSE FACTOR2 (CRF2) and CRF3 Regulate Lateral Root Development and Response to Cold Stress in Arabidopsis. Plant Cell Online tpc.00909.2015. https://doi.org/10.1105/tpc.15.00909

Khan, A., Fornes, O., Stigliani, A., Gheorghe, M., Castro-Mondragon, J.A., van der Lee, R., Bessy, A., Chèneby, J., Kulkarni, S.R., Tan, G., Baranasic, D., Arenillas, D.J., Sandelin, A., Vandepoele, K., Lenhard, B., Ballester, B., Wasserman, W.W., Parcy, F., Mathelier, A., 2018. JASPAR 2018: update of the open-access database of transcription factor binding profiles and its web framework. Nucleic Acids Res. 46, D260–D266. https://doi.org/10.1093/nar/gkx1126

Larkin, M.A., Blackshields, G., Brown, N.P., Chenna, R., McGettigan, P.A., McWilliam, H., Valentin, F., Wallace, I.M., Wilm, A., Lopez, R., Thompson, J.D., Gibson, T.J., Higgins, D.G., 2007. Clustal W and Clustal X version 2.0. Bioinformatics 23, 2947–2948. https://doi.org/10.1093/bioinformatics/btm404

Lee, T.I., Young, and R.A., 2000. Transcription of Eukaryotic Protein-Coding Genes. Annu. Rev. Genet. 34, 77–137. https://doi.org/10.1146/annurev.genet.34.1.77

Li, W.-H., Yang, J., Gu, X., 2005. Expression divergence between duplicate genes. Trends Genet. 21, 602–607. https://doi.org/10.1016/j.tig.2005.08.006

Licausi, F., Ohme-Takagi, M., Perata, P., 2013. APETALA2/Ethylene Responsive Factor (AP2/ERF) transcription factors: mediators of stress responses and developmental programs. New Phytol. 199, 639–649. https://doi.org/10.1111/nph.12291

Ludwig, M.Z., Palsson, A., Alekseeva, E., Bergman, C.M., Nathan, J., Kreitman, M., 2005. Functional Evolution of a cis-Regulatory Module. PLOS Biol. 3, e93. https://doi.org/10.1371/journal.pbio.0030093

Markert, C.L., 1964. Cellular differentiation—an expression of differential gene function. Congential Malform. 163–174.
Mi, H., Poudel, S., Muruganujan, A., Casagrande, J.T., Thomas, P.D., 2016. PANTHER version 10: expanded protein families and functions, and analysis tools. Nucleic Acids Res. 44, D336–D342. https://doi.org/10.1093/nar/gkv1194
Mizoi, J., Shinozaki, K., Yamaguchi-Shinozaki, K., 2012. AP2/ERF family transcription factors in plant abiotic stress responses. Biochim. Biophys. Acta BBA - Gene Regul. Mech., Plant gene regulation in response to abiotic stress 1819, 86–96. https://doi.org/10.1016/j.bbagrm.2011.08.004
Nakano, T., Suzuki, K., Fujimura, T., Shinshi, H., 2006. Genome-Wide Analysis of the ERF Gene Family in Arabidopsis and Rice. Plant Physiol. 140, 411–432. https://doi.org/10.1104/pp.105.073783
Ohno, S., 1970. Evolution by Gene Duplication. Springer-Verlag.
Park, J.M., Park, C.-J., Lee, S.-B., Ham, B.-K., Shin, R., Paek, K.-H., 2001. Overexpression of the Tobacco Tsi1 Gene Encoding an EREBP/AP2–Type Transcription Factor Enhances Resistance against Pathogen Attack and Osmotic Stress in Tobacco. Plant Cell 13, 1035–1046. https://doi.org/10.1105/tpc.13.5.1035
Raines, T., Shanks, C., Cheng, C.-Y., McPherson, D., Argueso, C.T., Kim, H.J., Franco-Zorrilla, J.M., López-Vidriero, I., Solano, R., Vaňková, R., Schaller, G.E., Kieber, J.J., 2016. The cytokinin response factors modulate root and shoot growth and promote leaf senescence in Arabidopsis. Plant J. 85, 134–147. https://doi.org/10.1111/tpj.13097
Ramaiah, M., Jain, A., Raghothama, K.G., 2014. Ethylene Response Factor070 regulates root development and phosphate starvation-mediated responses. Plant Physiol. 164, 1484–1498. https://doi.org/10.1104/pp.113.231183
Rashotte, A.M., Goertzen, L.R., 2010. The CRF domain defines Cytokinin Response Factor proteins in plants. BMC Plant Biol. 10, 74. https://doi.org/10.1186/1471-2229-10-74
Rashotte, A.M., Mason, M.G., Hutchison, C.E., Ferreira, F.J., Schaller, G.E., Kieber, J.J., 2006. A subset of Arabidopsis AP2 transcription factors mediates cytokinin responses in concert with a two-component pathway. Proc. Natl. Acad. Sci. 103, 11081–11085. https://doi.org/10.1073/pnas.0602038103
Riechmann, J.L., Heard, J., Martin, G., Reuber, L., Jiang, C.-Z., Keddie, J., Adam, L., Pineda, O., Ratcliffe, O.J., Samaha, R.R., Creelman, R., Pilgrim, M., Broun, P., Zhang, J.Z., Ghandehari, D., Sherman, B.K., Yu, G.-L., 2000. Arabidopsis Transcription Factors: Genome-Wide Comparative Analysis Among Eukaryotes. Science 290, 2105–2110. https://doi.org/10.1126/science.290.5499.2105
Ruiz-Modrano, R., Xoconostle-Cázares, B., Ham, B.-K., Li, G., Lucas, W.J., 2011. Vascular expression in Arabidopsis is predicted by the frequency of CT/GA-rich repeats in gene promoters. Plant J. 67, 130–144. https://doi.org/10.1111/j.1365-313X.2011.04581.x
Sakuma, Y., Liu, Q., Dubouzet, J.G., Abe, H., Shinozaki, K., Yamaguchi-Shinozaki, K., 2002. DNA-Binding Specificity of the ERF/AP2 Domain of Arabidopsis DREBs, Transcription Factors Involved in Dehydration- and Cold-Inducible Gene Expression. Biochem. Biophys. Res. Commun. 290, 998–1009. https://doi.org/10.1006/bbrc.2001.6299
Shi, X., Gupta, S., Rashotte, A.M., 2014. Characterization of two tomato AP2/ERF genes, <Emphasis Type="Italic">SlCRF1</Emphasis> and <Emphasis Type="Italic">SlCRF2</Emphasis> in hormone and stress responses. Plant Cell Rep. 33, 35–45. https://doi.org/10.1007/s00299-013-1510-6
Shi, X., Gupta, S., Rashotte, A.M., 2012. Solanum lycopersicum cytokinin response factor (SlCRF) genes: characterization of CRF domain-containing ERF genes in tomato. J. Exp. Bot. 63, 973–982. https://doi.org/10.1093/jxb/err325
Šimášková, M., O'Brien, J.A., Khan, M., Van Noorden, G., Ötvös, K., Vieten, A., De Clercq, I., Van Haperen, J.M.A., Cuesta, C., Hoyerová, K., Vanneste, S., Marhavý, P., Wabnik, K., Van Breusegem, F., Nowack, M., Murphy, A., Friml, J., Weijers, D., Beeckman, T., Benková, E.,

2015. Cytokinin response factors regulate *PIN-FORMED* auxin transporters. Nat. Commun. 6, 8717. https://doi.org/10.1038/ncomms9717

Varala, K., Marshall-Colón, A., Cirrone, J., Brooks, M.D., Pasquino, A.V., Léran, S., Mittal, S., Rock, T.M., Edwards, M.B., Kim, G.J., Ruffel, S., McCombie, W.R., Shasha, D., Coruzzi, G.M., 2018. Temporal transcriptional logic of dynamic regulatory networks underlying nitrogen signaling and use in plants. Proc. Natl. Acad. Sci. 201721487. https://doi.org/10.1073/pnas.1721487115

Waterhouse, A.M., Procter, J.B., Martin, D.M.A., Clamp, M., Barton, G.J., 2009. Jalview Version 2—a multiple sequence alignment editor and analysis workbench. Bioinformatics 25, 1189–1191. https://doi.org/10.1093/bioinformatics/btp033

Weirauch, M.T., Yang, A., Albu, M., Cote, A.G., Montenegro-Montero, A., Drewe, P., Najafabadi, H.S., Lambert, S.A., Mann, I., Cook, K., Zheng, H., Goity, A., van Bakel, H., Lozano, J.-C., Galli, M., Lewsey, M.G., Huang, E., Mukherjee, T., Chen, X., Reece-Hoyes, J.S., Govindarajan, S., Shaulsky, G., Walhout, A.J.M., Bouget, F.-Y., Ratsch, G., Larrondo, L.F., Ecker, J.R., Hughes, T.R., 2014. Determination and Inference of Eukaryotic Transcription Factor Sequence Specificity. Cell 158, 1431–1443. https://doi.org/10.1016/j.cell.2014.08.009

Winter, D., Vinegar, B., Nahal, H., Ammar, R., Wilson, G.V., Provart, N.J., 2007. An "Electronic Fluorescent Pictograph" Browser for Exploring and Analyzing Large-Scale Biological Data Sets. PLOS ONE 2, e718. https://doi.org/10.1371/journal.pone.0000718

Wittkopp, P.J., Kalay, G., 2012. Cis-regulatory elements: molecular mechanisms and evolutionary processes underlying divergence. Nat. Rev. Genet. 13, 59–69. https://doi.org/10.1038/nrg3095

Yáñez-Cuna, J.O., Kvon, E.Z., Stark, A., 2013. Deciphering the transcriptional cis-regulatory code. Trends Genet. 29, 11–22. https://doi.org/10.1016/j.tig.2012.09.007

Yu, C.-P., Lin, J.-J., Li, W.-H., 2016. Positional distribution of transcription factor binding sites in *Arabidopsis thaliana*. Sci. Rep. 6, 25164. https://doi.org/10.1038/srep25164

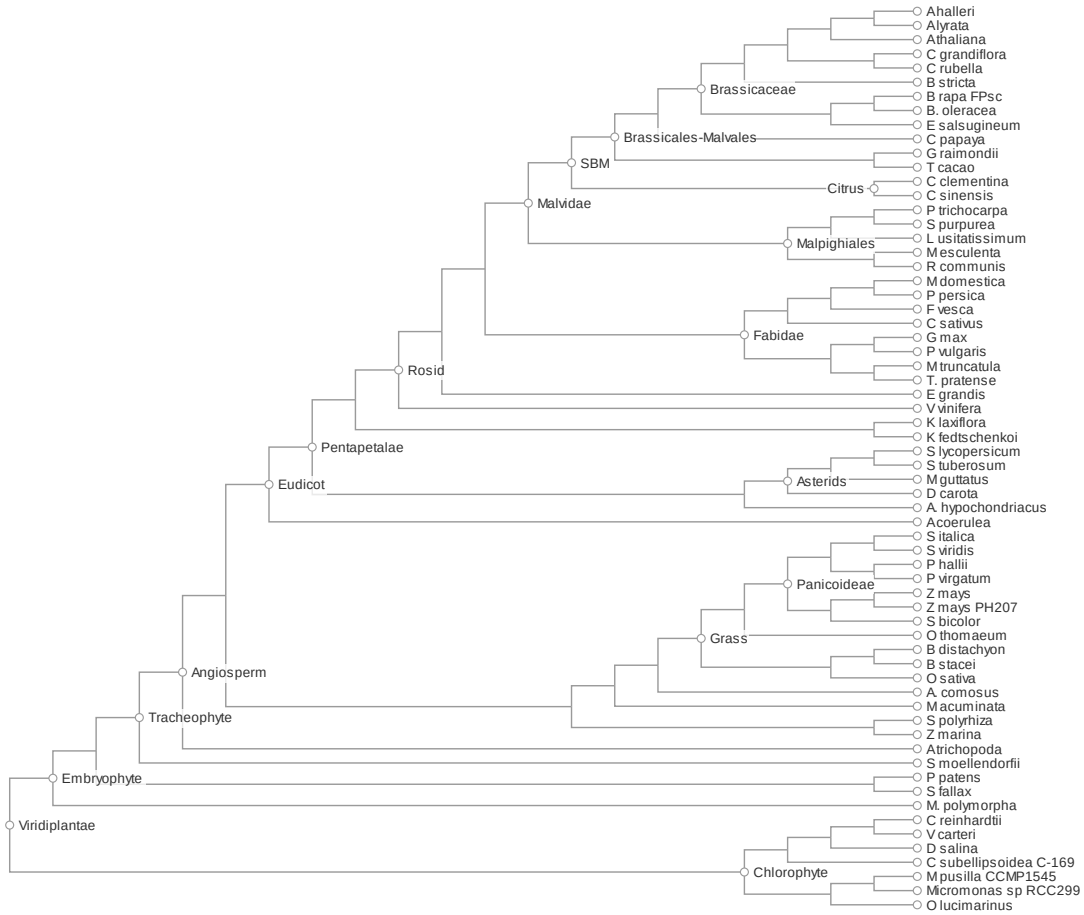
Zhou, J., Tang, X., Martin, G.B., 1997. The Pto kinase conferring resistance to tomato bacterial speck disease interacts with proteins that bind a cis-element of pathogenesis-related genes. EMBO J. 16, 3207–3218. https://doi.org/10.1093/emboj/16.11.3207

Zwack, P.J., Compton, M.A., Adams, C.I., Rashotte, A.M., 2016. Cytokinin response factor 4 (CRF4) is induced by cold and involved in freezing tolerance. Plant Cell Rep. 35, 573–584. https://doi.org/10.1007/s00299-015-1904-8

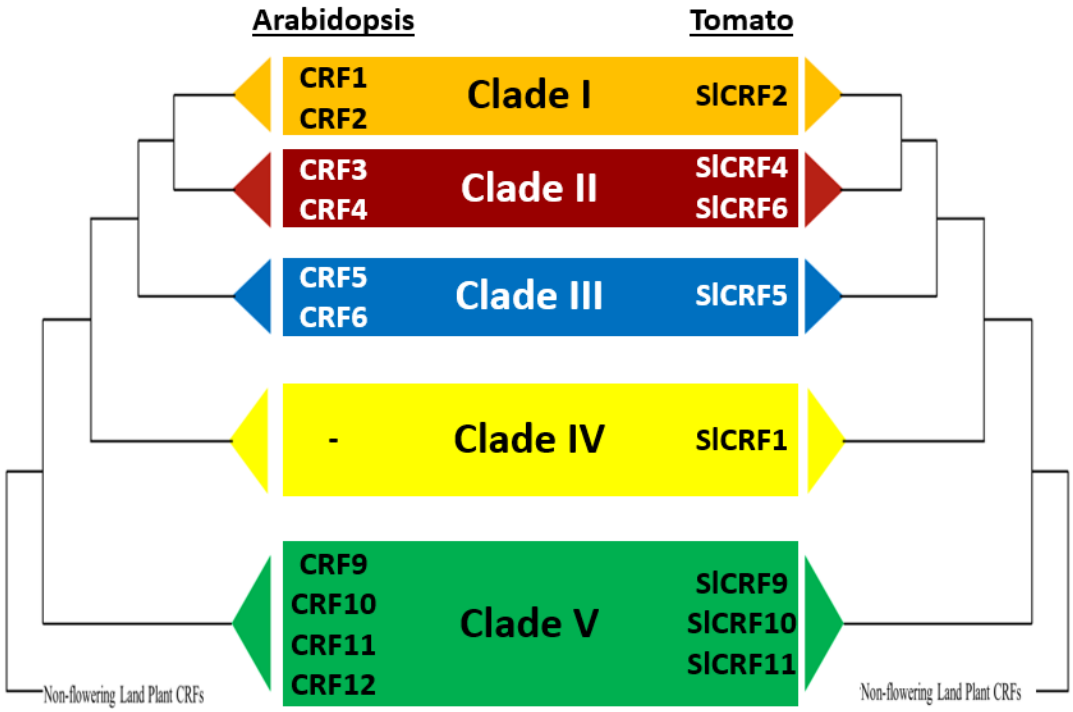
Zwack, P.J., Robinson, B.R., Risley, M.G., Rashotte, A.M., 2013. Cytokinin Response Factor 6 Negatively Regulates Leaf Senescence and is Induced in Response to Cytokinin and Numerous Abiotic Stresses. Plant Cell Physiol. 54, 971–981. https://doi.org/10.1093/pcp/pct049

Zwack, P.J., Shi, X., Robinson, B.R., Gupta, S., Compton, M.A., Gerken, D.M., Goertzen, L.R., Rashotte, A.M., 2012. Vascular Expression and C-Terminal Sequence Divergence of Cytokinin Response Factors in Flowering Plants. Plant Cell Physiol. 53, 1683–1695. https://doi.org/10.1093/pcp/pcs110

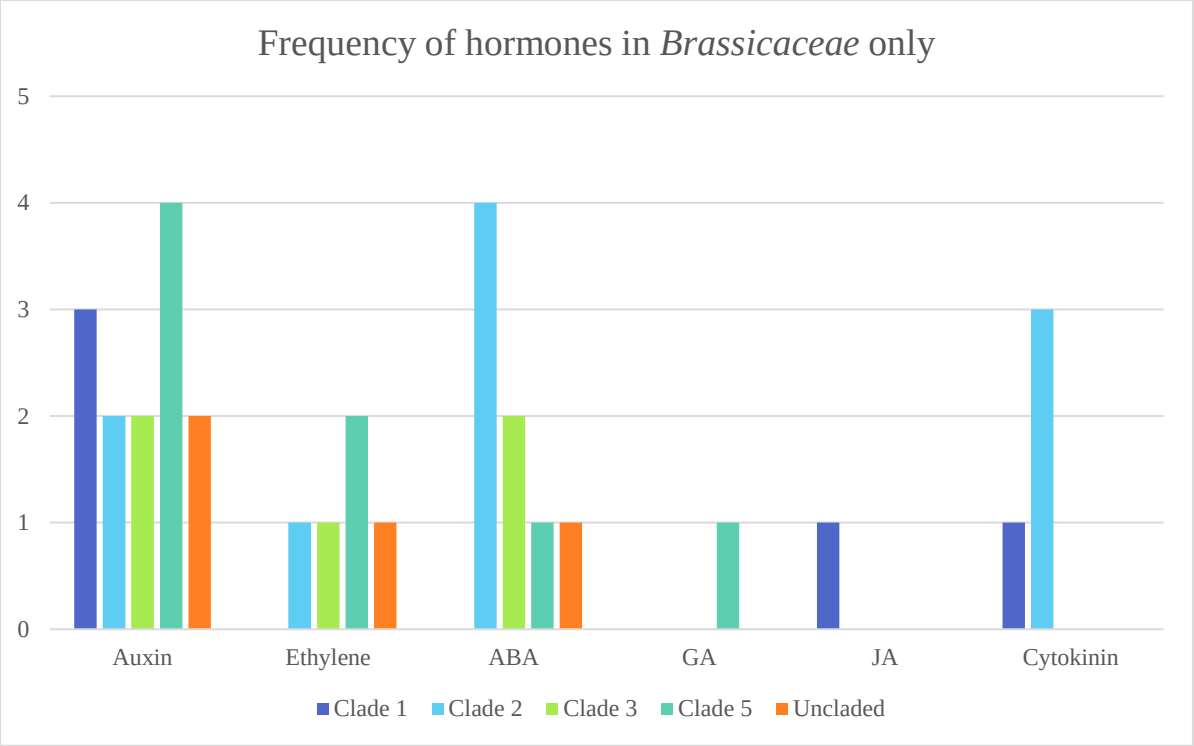
Supplemental



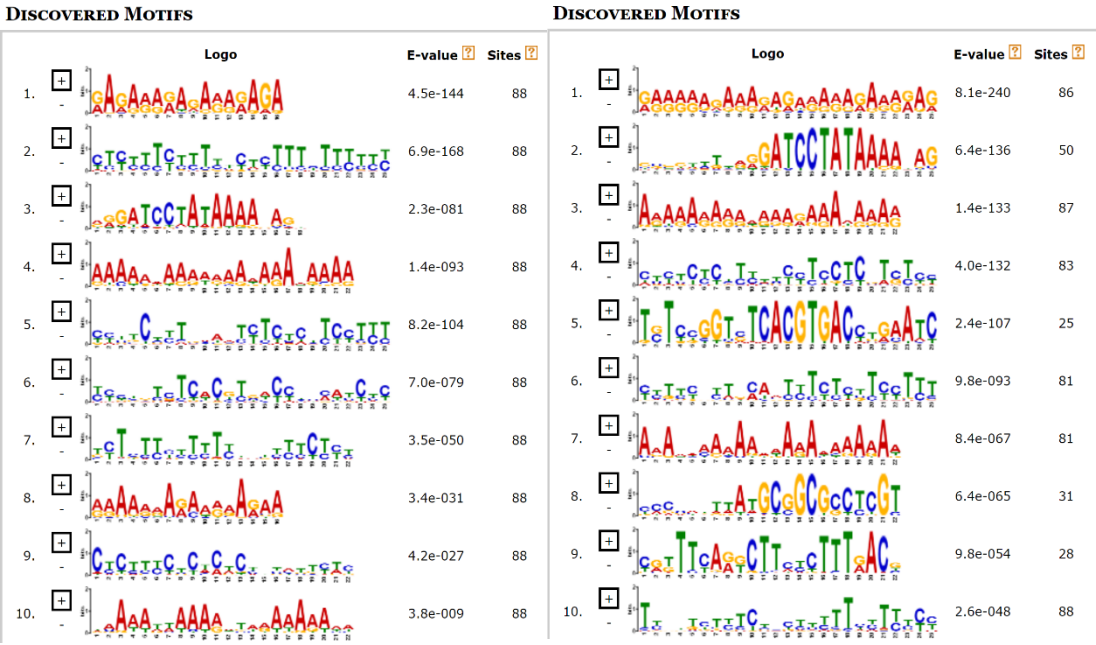
Supplemental Figure 1: A phylogeny of plant genomes available from Phytozome



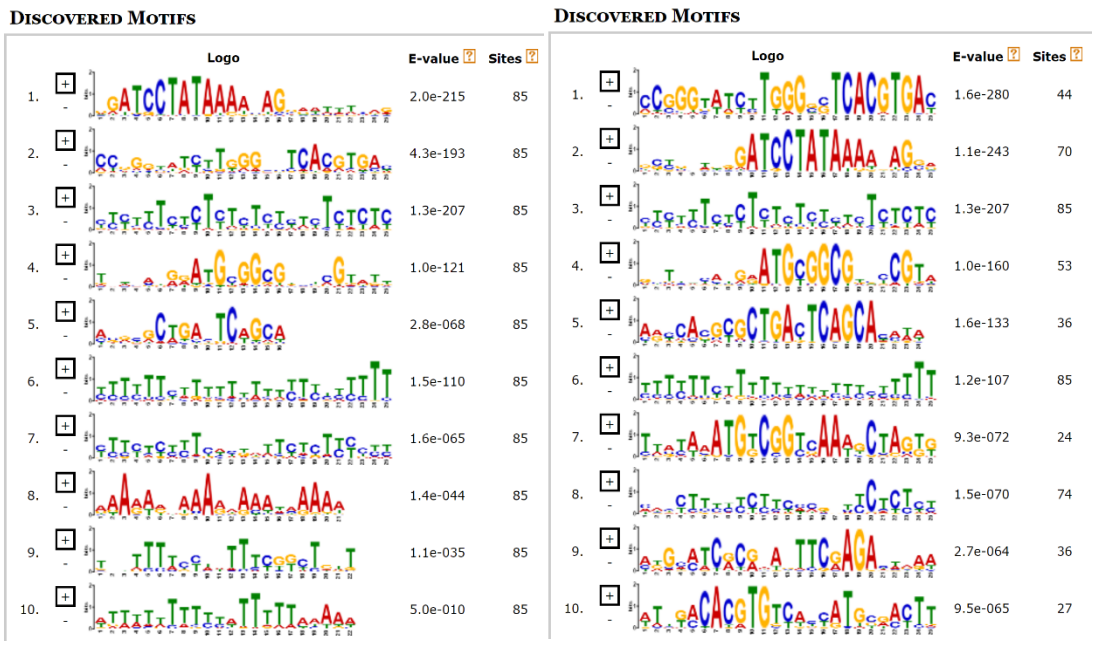
Supplemental Figure 2: Arabidopsis and tomato CRFs by clade



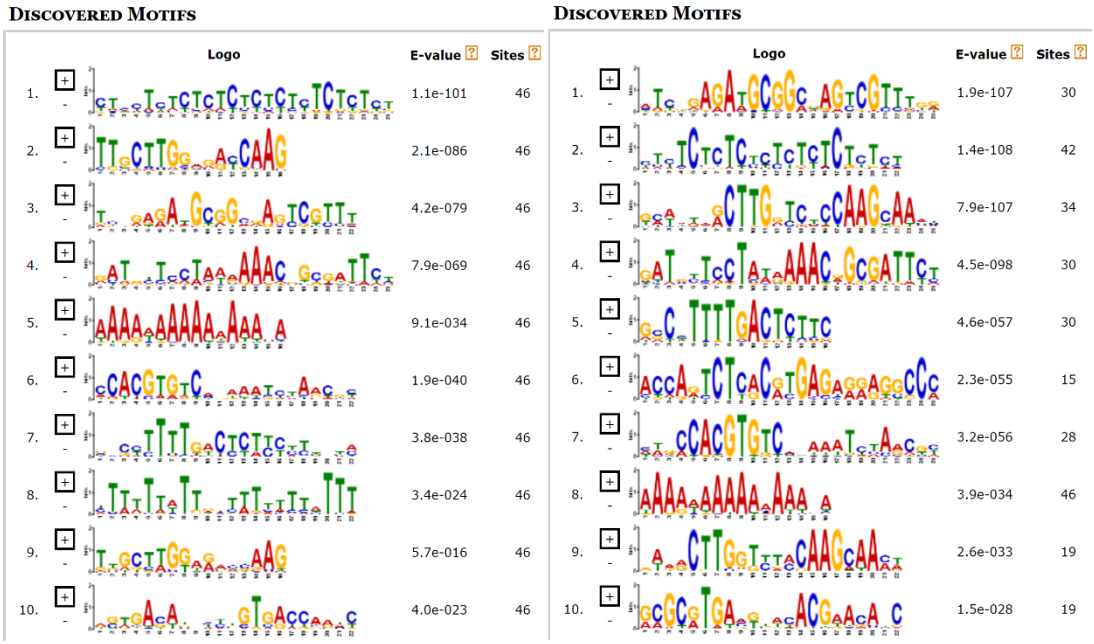
Supplemental Figure 3: A breakdown of hormone frequencies seen in Tomtom results based upon MEME runs with only Brassicaceae.



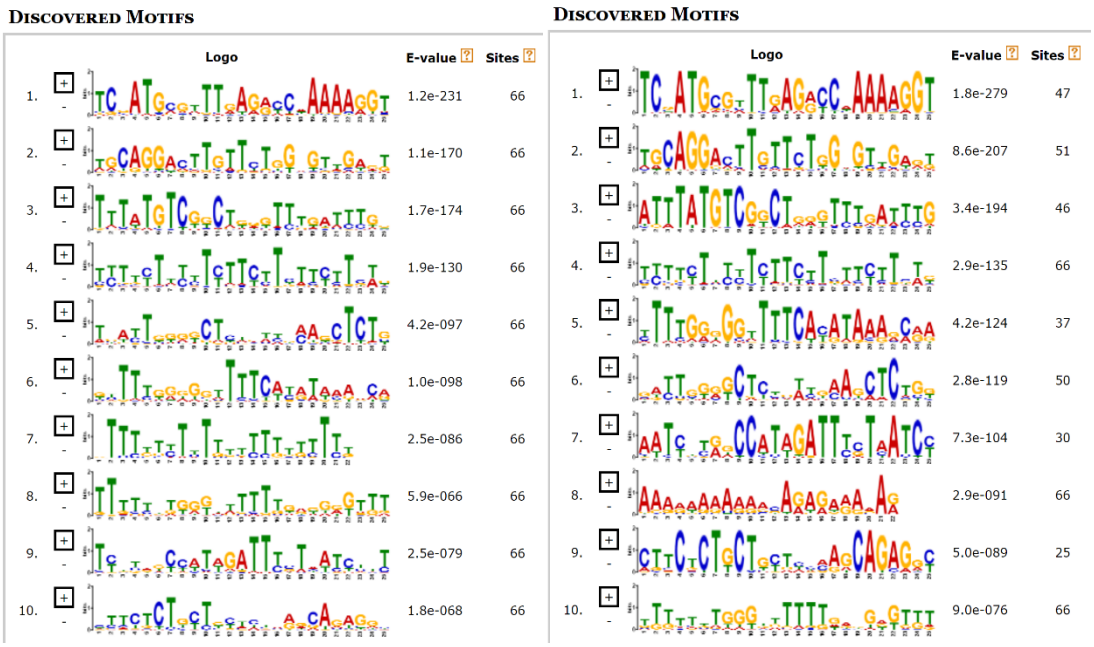
Supplemental Figure 4: Clade I top 10 motifs



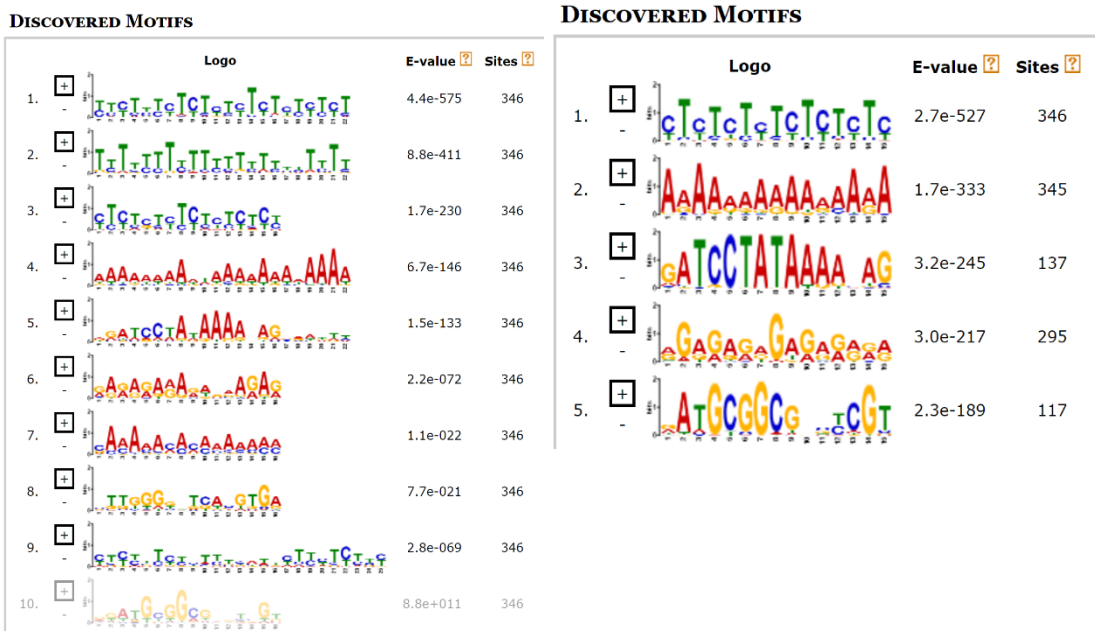
Supplemental Figure 5: Clade II top 10 motifs



Supplemental Figure 6: Clade III top 10 motifs



Supplemental Figure 7: Clade V top 10 motifs



Supplemental Figure 8: All CRF upstream sequences motifs

Supplemental Information 1: Arabidopsis Cytokinin Response Factors

>C1_AT_AT4G11140_AA // AtCRF1

METEKKVSLPRILRISVTDPYATDSSSDEEEEVDFDALSTKRRRVKKYVKEVVLDSVVSD
KEKPMKKKRKKRVVTVPVVVTTATRKFRGVRQRPWGKWAAEIRDPSRRVRVWLGTF
DTAEEAAIVYDNAAIQLRGPNAELNFPPPPVTENVEEASTEVKGVSDFIIGGGECLRSPVS
VLESPFSGESTAVKEEFVGVSTAEIVVKKEPSFNGSDFSAPLFSDDDVFGFSTSMSESFGG
DLFGDNLFADMSFGSGFGFGSGSGFSSWHVEDHFQDIGDLFGSDPVLTV*

>C1_AT_AT4G23750_AA // AtCRF2

MEAEKKMVLPRIKFTEHKTNTTTIVSELTNTHQTRILRISVTDPDATDSSSDDEEEEHQRF
VSKRRRVKKFVNEVYLDSGAVVTGSCGQMESKKRQKRAVKSESTVSPVVSATTTTTGE
KKFRGVRQRPWGKWAAEIRDPLKRVRLWLGTYNTAEEAAMVYDNAAIQLRGPDALTN
FSVTPTTATEKKAPPPSPVKKKKKKNNKSKKSVTASSSISRSSSNDCLCSPVSVLRSPFAV
DEFSGISSSPVAAVVVKEEPSMTTVSETFSDFSAPLFSDDDVFDFRSSVVPDYLGGDLFGE
DLFTADMCTDMNFGFDFGSGLSSWHMEDHFQDIGDLFGSDPLLAV*

>C2_AT_AT4G27950_AA // AtCRF3

MMMDEFMDLRPVKYTEHKTVIRKYTKKSSMERKTSVRDSARLVRVSMTDRDATDSSS
DEEEFLFPRRRVKRLINEIRVEPSSSSTGDVSASPTKDRKRINVDSTVQKPSVSGQNQKKY
RGVRQRPWGKWAAEIRDPEQRRRIWLGTFATAEEAAIVYDNAAIKLRGPDALTNFTVQ
PEPEPVQEQEQEPESNMSVSISESMDDSQHLSSPTSVLNYQTYVSEEPIDSLIKPVKQEFLE
PEQEPISWHLGEGNTNTNDDSFPLDITFLDNYFNESLPDISIFDQPMSPIQPTENDFFNDLM
LFDSNAEEYYSSEIKEIGSSFNDLDDSLISDLLLV*

>C2_AT_AT5G53290_AA // AtCRF4

MDEYIDFRPLKYTEHKTSMTKYTKKSSEKLSGGKSLKKVSICYTDPDATDSSSDEDEEDF
LFPRRRVKRFVNEITVEPSCNNVVTGVSMKDRKRLSSSSDETQSPASSRQRPNNKVSVSG
QIKKFRGVRQRPWGKWAAEIRDPEQRRRIWLGTFETAEEAAVVYDNAAIRLRGPDALT
NFSIPPQEEEEEEEPEPVIEEKPVIMTTPTPTTSSSESTEEDLQHLSSPTSVLNHRSEEIQQV
QQPFKSAKPEPGVSNAPWWHTGFNTGLGESDDSFPLDTPFLDNYFNESPPEMSIFDQPM
DQIFCENDDIFNDMLFLGGETMNIEDELTSSSIKDMGSTFSDFDDSLISDLLVA*

>C3_AT_AT2G46310_AA // AtCRF5

MKSRVRKSKYTVHRKITSTPFDGFPKIVKIIVTDPCATDSSSDEENDNKSVAPRVKRYVD
EIRFCDEDDEPKPARKAKKKSPAAAAENGGDLVKSVVKYRGVRQRPWGKFAAEIRDPS
SRTRLWLGTFATAEEAAIGYDRAAIRIKGHNAQTNFLTPPPSPTTEVLPETPVIDLETVSG
CDSARESQISLCSPTSVLRFSHNDETEYRTEPTEEQNPFFLPDLFRSGDYFWDSEITPDPLF
LDEFHQSLLPNINNNNTVCDKDTNLSDSFPLGVIGDFSSWDVDEFFQDHLLDK*

>C3_AT_AT3G61630_AA // AtCRF6

MERRTRRVKFTENRTVTNVAATPSNGSPRLVRITVTDPFATDSSSDDDDNNNVTVVPRV
KRYVKEIRFCQGESSSSTAARKGKHKEEESVVVEDDVSTSVKPKKYRGVRQRPWGKFA
AEIRDPSSRTRIWLGTFVTAEEAAIAYDRAAIHLKGPKALTNFLTPPTPTPVIDLQTVSAC
DYGRDSRQSLHSPTSVLRFNVNEETEHEIEAIELSPERKSTVIKEEEEESSAGLVFPDPYLLP

DLSLAGECFWDTEIAPDLLFLDEETKIQSTLLPNTEVSKQGENETEDFEFGLIDDFESSPW
DVDHFFDHHHHSFD*

>C78_AT_AT1G22985_AA // AtCRF7

MKRIVRISFTDMEATDSSSSEDESPPSSRRRGKKLVKEIVIDHSDPPEVGKTRFKIRIPASL
LAARNTTANKKKFRGVRQRPWGKWAAEIRCGRVKGRPERIWLGTFETAEEAALAYDN
AAIQLIGPDAPTNFGRPDVDSAVVKKQDSDASGGASEEVV*

>C78_AT_AT1G71130_AA // AtCRF8

MKRIIRISFTDAEATDSSSDEDTEERGGASQTRRRGKRLVKEIVIDPSDSADKLDVCKTRF
KIRIPAEFLKTAKTEKKYRGVRQRPWGKWVAEIRCGRGACKGRRDRLWLGTFNTAEEA
ALAYDNASIKLIGPHAPTNFGLPAENQEDKTVIGASEVARGA*

>C5_AT_AT1G49120_AA // AtCRF9

MISFREENIDLNLIKTISVICNDPDATDSSSDDESISGNNPRRQIKPKPPKRYVSKICVPTLI
KRYENVSNSTGNKAAGNRKTSSGFKGVRRRPWGKFAAEIRNPFEKKRKWLGTFPTEEE
AAEAYQKSKREFDERLGLVKQEKDLVDLTKPCGVRKPEEKEVTEKSNCKKVNKRIVTD
QKPFGCGYNADHEEEGVISKMLEDPLMTSSIADIFGDSAVEANDIWVDYNSVEFISIVDD
FKFDFVENDRVGKEKTFGFKIGDHTKVNQHAKIVSTNGDLFVDDLLDFDPLIDDFKLED
FPMDDLGLLGDPEDDDFSWFNGTTDWIDKFL*

>C5_AT_AT1G68550_AA // AtCRF10

MVAIRKEQSLSGVSSEIKKRAKRNTLSSLPQETQPLRKVRIIVNDPYATDDSSSDEEELKV
PKPRKMKRIVREINFPSMEVSEQPSESSSQDSTKTDGKIAVSASPAVPRKKPVGVRQRKW
GKWAAEIRDPIKKTRTWLGTFDTLEEAAKAYDAKKLEFDAIVAGNVSTTKRDVSSSETS
QCSRSSPVVPVEQDDTSASALTCVNNPDDVSTVAPTAPTPNVPAGGNKETLFDFDFTNL
QIPDFGFLAEEQQDLDFDCFLADDQFDDFGLLDDIQGFEDNGPSALPDFDFADVEDLQL
ADSSFGFLDQLAPINISCPLKSFAAS*

>C5_AT_At3g25890_AA // AtCRF11

MAERKKRSSIQTNKPNKKPMKKKPFQLNHLPGLSEDLKTMRKLRFVVNDPYATDYSSS
EEEERSQRRKRYVCEIDLPFAQAATQAESESSYCQESNNNGVSKTKISACSKKVLRSKAS
PVVGRSSTTVSKPVGVRQRKWGKWAAEIRHPITKVRTWLGTYETLEQAADAYATKKLE
FDALAAATSAASSVLSNESGSMISASGSSIDLDKKLVDSTLDQQAGESKKASFDFDFADL
QIPEMGCFIDDSFIPNACELDFLLTEENNNQMLDDYCGIDDLDIIGLECDGPSELPDYDFS
DVEIDLGLIGTTIDKYAFVDHIATTTPTPLNIACP*

>C5_AT_At1g25470_AA // AtCRF12

MKSFVKPERDSLLRTVRIVFTDPDATDDSSSSSDEWLPKPRKVKRFVHEITFLPQVSESSQ
DRSNAVKTPRRKSTRQFKYPVGVRPRPSGKFAAEILNPFTKTKKWLGTYETPAEAEKAY
VDKKVEYDALASSGSAVSSSVVTVTSQCLRSPTSASVSCVSADDLSKEKTSLNKDVAAS
GDSTTKEVFTTFDFSDVKIPDLRFLAAEEDSMVSNANGAELDFDCFLTDSNILLDDYSLL
ENDINFSRFENSLPSELPDCDFTEMEFQLDDFKFAYTDHLTTPPLGLV*

Supplemental Table 1: Full UniProt results

Clade	Run Type	Rank	E-value	Sites	UniProt ID	Gene Name	P-value	GO-terms from Uniprot
1	O	1	4.5E-144	88	Q9SKD0	BPC1	3.50E-09	Regulation of developmental process, regulation of transcription, response to ethylene, transcription
1	O	1	4.5E-144	88	Q8L999	BPC6	5.14E-09	Regulation of transcription, response to ethylene, transcription
1	O	1	4.5E-144	88	F4IUI3	BPC5	2.61E-08	Regulation of transcription, response to ethylene, transcription
1	O	2	6.9E-168	88	Q9FGD6	DOF5.8	5.85E-10	Regulation of transcription, transcription
1	O	2	6.9E-168	88	Q9SEZ3	CDF5	8.04E-10	Flower development, negative regulation of long- and short-day photoperiodism, reulgation of transcription, transcription
1	O	2	6.9E-168	88	Q9M2U1	DOF3.6	4.27E-09	Regulation of transcription, transcription
1	O	3	2.3E-81	88	Q9ZPX0	GATA20	2.99E-03	Cell differentiation, transcription, transcription regulation
1	O	3	2.3E-81	88	C0SVG5	RVE5	1.36E-02	Response to abscisic acid, response to auxin, response to cadmium ion, response to ethylene, response to gibberllin, response to jasmonic acid, response to salicylic acid, response to salt stress, transcription
1	O	3	2.3E-81	88	Q8H0W3	RVE6	1.53E-02	Regulation of circadian rhythm, response to abscisic acid, response to cadmium ion, response to ethylene, response to gibberllin, response to jasmonic acid, response to salicylic acid, response to salt stress, transcription
1	O	4	1.4E-93	88	Q9FGD6	DOF5.8	3.55E-11	Regulation of transcription, transcription
1	O	4	1.4E-93	88	Q9M2U1	DOF3.6	1.89E-10	Regulation of transcription, transcription
1	O	4	1.4E-93	88	Q9LZ56	DOF5.1	2.37E-10	Regulation of transcription, transcription
1	O	5	8.2E-104	88	P68350	DOF1.5	7.87E-05	Regulation of transcription, seed coat development, transcription
1	O	5	8.2E-104	88	XXX	XXX	1.56E-04	XXX
1	O	5	8.2E-104	88	O82155	DOF1.7	1.66E-04	Regulation of transcription, response to chitin, transcription
1	Z	1	8.1E-240	86	Q9SKD0	BPC1	8.24E-11	Regulation of developmental process, regulation of transcription, response to ethylene, transcription
1	Z	1	8.1E-240	86	F4IUI3	BPC5	2.90E-10	Regulation of transcription, response to ethylene, transcription
1	Z	1	8.1E-240	86	Q8L999	BPC6	1.32E-09	Regulation of transcription, response to ethylene, transcription
1	Z	2	6.4E-136	50	Q9ZPX0	GATA20	1.47E-03	Cell differentiation, transcription, transcription regulation
1	Z	2	6.4E-136	50	XXX	XXX	XXX	XXX
1	Z	2	6.4E-136	50	XXX	XXX	XXX	XXX
1	Z	3	1.4E-133	87	Q9M2U1	DOF3.6	2.82E-11	Regulation of transcription, transcription
1	Z	3	1.4E-133	87	Q9LZ56	DOF5.1	2.04E-10	Regulation of transcription, transcription
1	Z	3	1.4E-133	87	Q9SEZ3	CDF5	2.14E-10	Flower development, negative regulation of long- and short-day photoperiodism, reulgation of transcription, transcription
1	Z	4	4E-132	83	Q39088	DOF3.4	6.37E-05	Cell wall modification, positive regulation of cell cycle, positive regulation of transcription, response to auxin, response to salicylic acid, transcription
1	Z	4	4E-132	83	Q8L999	BPC6	7.72E-05	Regulation of transcription, response to ethylene, transcription
1	Z	4	4E-132	83	Q9FGD6	DOF5.8	1.09E-04	Regulation of transcription, transcription

	Z	5	2.4E-107	25	A9RZ73	xxx	9.69E-06	predicted protein
1	Z	5	2.4E-107	25	Q9CAA4	BIM2	5.79E-05	Positive brassinosteroid signaling protein, regulation of transcription, transcription
1	Z	5	2.4E-107	25	Q9M7Q5	ABF1	2.06E-04	Abscisic acid-activated signaling pathway, transcription

2	O	1	2E-215	85	Q9ZPX0	GATA20	2.20E-03	Cell differentiation, transcription, transcription regulation
2	O	1	2E-215	85	Q6R053	MYB56	6.15E-03	Brassinosteroid mediated signaling pathway, cell differentiation, cellular response to brassinosteroid stimulus, endothelial cell proliferation, integument development, negative regulation of cell division, negative resulation of long-day photoperiodism, positive regulation of transcription, quiescent center organization, regulation of seed growth, regulation of transcription, regulation of transcription from RNA polymperase II promoter, transcription
2	O	1	2E-215	85	Q9LSL6	DOF5.7	6.78E-03	Guard cell differentiation, positive regulation of transcription, regulation of cell wall pectin metabolic process, regulation of transcription, stomatal movement
2	O	2	4.3E-193	85	A9RZ73	PHYPADRAFT_72483	9.24E-06	Predicted protein
2	O	2	4.3E-193	85	Q9CAA4	BIM2	5.75E-05	Transcription, transcription regulation
2	O	2	4.3E-193	85	Q9LT89	TCP19	8.85E-05	Regulation of defense response, regulation of transcription, transcription
2	O	3	1.3E-207	85	Q9SKD0	BPC1	5.52E-16	Regulation of developmental process, regulation of transcription, response to ethylene, transcription
2	O	3	1.3E-207	85	F4JUI3	BPC5	5.55E-15	Regulation of transcription, response to ethylene, transcription
2	O	3	1.3E-207	85	Q8L999	BPC6	1.11E-13	Regulation of transcription, response to ethylene, transcription
2	O	4	1E-121	85	P93007	ERF112	4.48E-03	Ethylene-activated signaling pathway, transcription
2	O	4	1E-121	85	O80339	ERF3	4.52E-03	Defense response, ethylene-activated signaling pathway, negative regulation of ethylene-activated signaling pathway, transcription
2	O	4	1E-121	85	O24646	HY5	4.86E-03	Abscisic acid-activated signaling pathway, gibberellic acid mediated signaling pathway, positive gravitropism, positive regulation of anthyocyanin metabolic process, positive regulation of circadian rhythm, red/far-red light phototransduction, red/far-red light signaling pathway, regulation of absciscic acid-activated signaling pathway, regulation of photomorphogeneiss, response to abscisic acid, response to far red light, response to karrikin, response to red light, response to UV-B, transcription
2	O	5	2.8E-68	85	P92948	CDC5	6.53E-04	Cell cycle, cell differentiation, defense response signaling pathway, defense response to bacterium, defense response to fungus, DNA repair, innate immune response, mRNA processing, regulation of transcription, regulation of transcription from RNA polymerase II promoter, RNA splicing, transcription
2	O	5	2.8E-68	85	O22900	WRKY23	1.60E-03	Response to auxin, response to nematode, transcription
2	O	5	2.8E-68	85	Q9SKD9	WRKY46	3.27E-03	Lateral root development, regulation of transcription, response to chitin, transcription
2	Z	1	1.6E-280	44	A9RZ73	PHYPADRAFT_72483	4.70E-06	Predicted protein

2	Z	1	1.6E-280	44	Q9CAA4	BIM2	3.77E-05	Transcription, transcription regulation
2	Z	1	1.6E-280	44	Q9LEZ3	BIM1	6.81E-05	Regulation of transcription, transcription
2	Z	2	1.1E-243	70	Q9ZPX0	GATA20	1.70E-03	Cell differentiation, transcription, transcription regulation
2	Z	2	1.1E-243	70	Q64647	TCP9	7.16E-03	Negative regulation of leaf senescence, regulation of cell size, regulation of transcription, root development, transcription
2	Z	2	1.1E-243	70	Q9LSL6	DOF5.7	9.78E-03	Guard cell differentiation, positive regulation of transcription, regulation of cell wall pectin metabolic process, regulation of transcription, stomatal movement
2	Z	3	1.3E-207	85	Q9SKD0	BPC1	5.52E-16	Regulation of developmental process, regulation of transcription, response to ethylene, transcription
2	Z	3	1.3E-207	85	F4JUI3	BPC5	5.55E-15	Regulation of transcription, response to ethylene, transcription
2	Z	3	1.3E-207	85	Q8L999	BPC6	1.11E-13	Regulation of transcription, response to ethylene, transcription
2	Z	4	1E-160	53	Q38Q40	ERF112	3.53E-03	Ethylene-activated signaling pathway, transcription
2	Z	4	1E-160	53	Q9LND1	ERF094	4.77E-03	Ethylene-activated signaling pathway, jasmonic acid and ethylene-dependent systemic resistance, response to ethylene, response to jasmonic acid, transcription
2	Z	4	1E-160	53	P93007	ERF112	7.09E-03	Ethylene-activated signaling pathway, transcription
2	Z	5	1.6E-133	36	P92948	CDC5	6.82E-04	Cell cycle, cell differentiation, defense response signaling pathway, defense response to bacterium, defense response to fungus, DNA repair, innate immune response, mRNA processing, regulation of transcription, regulation of transcription from RNA polymerase II promoter, RNA splicing, transcription
2	Z	5	1.6E-133	36	A9TN45	PHYPADRAFT_48267	1.48E-03	Predicted protein
2	Z	5	1.6E-133	36	Q22900	WRKY23	1.48E-03	Response to auxin, response to nematode, transcription

3	O	1	1.1E-101	46	Q8L999	BPC6	1.91E-15	Regulation of transcription, response to ethylene, transcription
3	O	1	1.1E-101	46	Q9SKD0	BPC1	5.62E-13	Regulation of developmental process, regulation of transcription, response to ethylene, transcription
3	O	1	1.1E-101	46	F4JUI3	BPC5	1.70E-11	Regulation of transcription, response to ethylene, transcription
3	O	2	2.1E-86	46	Q01593	ABI3	1.97E-03	Abscisic acid-activated signaling pathway, embryo development ending in seed dormancy, mitochondira-nucleus signaling pathway, plastid organization, positive regulation of transcription, response to abscisic acid, response to auxin, transcription
3	O	2	2.1E-86	46	D7L3Y2	ARALYDRAFT_897773	3.77E-03	Predicted protein; cell differentiation, cotyledon morphogenesis, positive regulation of development, heterochronic, regulation of transcription, transcription
3	O	2	2.1E-86	46	Q9LEZ9	TCP17	5.91E-03	Cell differentiation, leaf morphogenesis, positive regulation of development, heterochronic, regulation of transcription, transcription
3	O	3	4.2E-79	46	Q9ZWJ9	ARR2	3.69E-03	Cellular response to cytokinin stimulus, cytokinin-activated signaling pathway, ethylene-activated signaling pathway, leaf senescence, regulation of root meristem growth, regulation of seed growth, regulation of stomatal movement, response to cytokinin, response to etylene, root development, transcription

3	O	3	4.2E-79	46	Q9SWG3	FAR1	1.17E-02	Far-red light signaling pathway, positive regulation of circadian rhythm, positive regulation of transcription, red/far-red phototransduction, red/far-red light signaling pathway, response to red/far red light
3	O	3	4.2E-79	46	Q8L9Y3	ARR14	1.29E-02	Regulation of transcription, transcription
3	O	4	7.9E-69	46	Q9FX67	HRS1	1.32E-03	Cellular response to nitrogen compound, cellular response to phosphate starvation, negative regulation of abscisic acid-activated signaling pathway, primary root development, regulation of response to nutrient levels, regulation of transcription, transcription
3	O	4	7.9E-69	46	Q8L9Y3	ARR14	7.03E-03	Cytokinin-activated signaling pathway, transcription
3	O	4	7.9E-69	46	Q8VZS3	HHO2	8.63E-03	Phosphate ion homeostasis, regulation of transcription, response to abscisic acid, transcription
3	O	5	9.1E-34	46	Q9M2U1	OBF3/DOF3.6	4.65E-09	Regulation of transcription, transcription
3	O	5	9.1E-34	46	Q9SEZ3	CDF5	8.03E-09	Flower development, negative regulation of long-/short-day photoperiodism, regulation of transcription, transcription
3	O	5	9.1E-34	46	Q9LZ56	DOF5.1	8.85E-08	Regulation of transcription, transcription
3	Z	1	1.9E-107	30	Q8L9Y3	ARR14	1.05E-02	Regulation of transcription, transcription
3	Z	1	1.9E-107	30	Q9ZWJ9	ARR2	1.29E-02	Cellular response to cytokinin stimulus, cytokinin-activated signaling pathway, ethylene-activated signaling pathway, leaf senescence, regulation of root meristem growth, regulation of seed growth, regulation of stomatal movement, response to cytokinin, response to etylene, root development, transcription
3	Z	1	1.9E-107	30	Q9SWG3	FAR1	1.37E-02	Far-red light signaling pathway, positive regulation of circadian rhythm, positive regulation of transcription, red/far-red phototransduction, red/far-red light signaling pathway, response to red/far red light
3	Z	2	1.4E-108	42	Q9SKD0	BPC1	4.35E-17	Regulation of developmental process, regulation of transcription, response to ethylene, transcription
3	Z	2	1.4E-108	42	F4JUI3	BPC5	1.21E-15	Regulation of transcription, response to ethylene, transcription
3	Z	2	1.4E-108	42	Q8L999	BPC6	6.94E-15	Regulation of transcription, response to ethylene, transcription
3	Z	3	7.9E-107	34	Q01593	ABI3	4.32E-03	Abscisic acid-activated signaling pathway, embryo development ending in seed dormancy, mitochondira-nucleus signaling pathway, plastid organization, positive regulation of transcription, response to abscisic acid, response to auxin, transcription
3	Z	3	7.9E-107	34	D7L3Y2	ARALYDRAFT_897773	6.97E-03	Predicted protein; cell differentiation, cotyledon morphogenesis, positive regulation of development, heterochronic, regulation of transcription, transcription
3	Z	3	7.9E-107	34	Q9LEZ9	TCP17	8.37E-03	Cell differentiation, leaf morphogenesis, positive regulation of development, heterochronic, regulation of transcription, transcription
3	Z	4	4.5E-98	30	Q9FX67	HRS1	1.82E-03	Cellular response to nitrogen compound, cellular response to phosphate starvation, negative regulation of abscisic acid-activated signaling pathway, primary root development, regulation of response to nutrient levels, regulation of transcription, transcription
3	Z	4	4.5E-98	30	Q8VZS3	HHO2	5.72E-03	Phosphate ion homeostasis, regulation of transcription, response to abscisic acid, transcription
3	Z	4	4.5E-98	30	Q8L9Y3	ARR14	9.18E-03	Regulation of transcription, transcription

3	Z	5	4.6E-57	30	Q9SAH7	WRKY40	3.42E-05	Defense response to bacterium, defense response to fungus, negative regulation of transcription, regulation of defense response, response to chitin, reponse to molecule of bacterial origin, response to salicylic acid, response to wounding, transcription
3	Z	5	4.6E-57	30	Q9LX86	F12M12_40	6.17E-04	Regulation of transcription, transcription
3	Z	5	4.6E-57	30	Q8VWQ5	WRKY50	7.17E-04	Defense response to fungus, jasmonic acid mediated signaling pathway, transcription

4	O	1	1.2E-35	20	Q9SKD0	BPC1	2.57E-12	Regulation of developmental process, regulation of transcription, response to ethylene, transcription
4	O	1	1.2E-35	20	F4JUI3	BPC5	1.88E-10	Regulation of transcription, response to ethylene, transcription
4	O	1	1.2E-35	20	Q8L999	BPC6	5.55E-10	Regulation of transcription, response to ethylene, transcription
4	O	2	2.8E-21	20	Q9FNY0	E2FA	6.34E-04	Cell cycle, positive regulation of meitoic cell cycle, positive regulation of transcription, transcription
4	O	2	2.8E-21	20	Q9SW63	RAP2-10	1.54E-03	Ethylene-activated signaling pathway, transcription
4	O	2	2.8E-21	20	Q9FGV1	ARF8	9.17E-03	Auxin-activated signaling pathway, flower development, response to auxin, transcription
4	O	3	5.2E-10	20	Q9SKD0	BPC1	1.69E-04	Regulation of developmental process, regulation of transcription, response to ethylene, transcription
4	O	3	5.2E-10	20	Q8L999	BPC6	2.22E-04	Regulation of transcription, response to ethylene, transcription
4	O	3	5.2E-10	20	XXX	RAMOSA1	2.22E-04	
4	O	4	0.00023	20	Q9M2U1	OBP3/DOF3.6	5.83E-08	Regulation of transcription, transcription
4	O	4	0.00023	20	Q9FGD6	DOF5.8	8.51E-08	Regulation of transcription, transcription
4	O	4	0.00023	20	Q9SEZ3	CDF5	3.14E-07	Flower development, negative regulation of long-/short-day photoperiodism, regulation of transcription, transcription
4	O	5	0.0015	20	Q1PFR7	LEC2	1.55E-03	Embryo development ending in seed dormancy, positive regulation of auxin biosynthesis process, positive regulation of transcription, seed maturation, seed oilbody biogenesis, somatic embryogenesis, transcription
4	O	5	0.0015	20	Q01593	ABI3	1.75E-03	Abscisic acid-activated signaling pathway, embryo development ending in seed dormancy, mitochondira-nucleus signaling pathway, plastid organization, positive regulation of transcription, response to abscisic acid, response to auxin, transcription
4	O	5	0.0015	20	Q8LFV3	CDF3	7.04E-03	Flower development, regulation of transcription, transcription
4	Z	1	1.2E-35	20	Q9SKD0	BPC1	2.57E-12	Regulation of developmental process, regulation of transcription, response to ethylene, transcription
4	Z	1	1.2E-35	20	F4JUI3	BPC5	1.88E-10	Regulation of transcription, response to ethylene, transcription
4	Z	1	1.2E-35	20	Q8L999	BPC6	5.55E-10	Regulation of transcription, response to ethylene, transcription
4	Z	2	4.9E-23	19	Q9FNY0	E2FA	8.09E-04	Cell cycle, positive regulation of meitoic cell cycle, positive regulation of transcription, transcription
4	Z	2	4.9E-23	19	Q9SW63	RAP210	9.62E-04	Ethylene-activated signaling pathway, transcription
4	Z	2	4.9E-23	19	Q9FJ93	DREB1D	4.41E-03	Abscisic acid-activated signaling pathway, glucosinolate metabolic process, transcription
4	Z	3	1.1E-11	20	Q8L999	BPC6	7.81E-05	Regulation of transcription, response to ethylene, transcription

4	Z	3	1.1E-11	20	Q9SKD0	BPC1	9.77E-05	Regulation of developmental process, regulation of transcription, response to ethylene, transcription
4	Z	3	1.1E-11	20	Q9LX86	F12M12_40	1.37E-04	Regulation of transcription, response to chitin
4	Z	4	1.5E-12	10	Q01593	ABI3	2.39E-03	Abscisic acid-activated signaling pathway, embryo development ending in seed dormancy, mitochondira-nucleus signaling pathway, plastid organization, positive regulation of transcription, response to abscisic acid, response to auxin, transcription
4	Z	4	1.5E-12	10	Q1PFR7	LEC2	3.26E-03	Embryo development ending in seed dormancy, positive regulation of auxin biosynthesis process, positive regulation of transcription, seed maturation, seed oilbody biogenesis, somatic embryogenesis, transcription
4	Z	4	1.5E-12	10	Q8LFV3	CDF3	4.93E-03	Flower development, regulation of transcription, transcription
4	Z	5	4.5E-11	7	Q9M2Y9	RAX3	9.69E-03	Cell differentiation, multicellular organism development, regulation of transcription from RNA polymerase II promoter, response to salicylic acid, transcription
4	Z	5	4.5E-11	7	Q8VWG0	bZIP52	1.87E-02	Regulation of transcription
4	Z	5	4.5E-11	7	XXX	XXX	XXX	

5	O	1	1.2E-231	66	Q1PFR7	LEC2	1.98E-03	Embryo development ending in seed dormancy, positive regulation of auxin biosynthesis process, positive regulation of transcription, seed maturation, seed oilbody biogenesis, somatic embryogenesis, transcription
5	O	1	1.2E-231	66	Q9ZPY0	DOF2.5	1.07E-02	Cellular response to red light, cellular response to water stimulus, positive regulation of gibberellin biosynthetic process, positive regulation of seed germination, red light signaling pathway, response to cold, response to light stimulus, seed germination, transcription
5	O	1	1.2E-231	66	Q80928	DOF2.4	1.36E-02	Regulation of trascription, transcription
5	O	2	1.1E-170	66	Q5DW98	PEND	5.40E-03	Predicted protein
5	O	2	1.1E-170	66	Q9ZUL3	IDD5	1.35E-02	Positive regulation of starch synthase SS4, development of chloroplast and startch granules, regulation of transcription, transcription
5	O	2	1.1E-170	66	Q9FM03	DOF5.6	1.49E-02	Phloem or xylem histogenesis, positive regulation of transcription, procambium histogenesis, regulation of transcription, transcription
5	O	3	1.7E-174	66	Q9SGJ6	DREB1E	2.95E-05	Regulation of gibberellin biosynthesis process, regulation of timing of transcition from vegetative to reproductive phase, transcription
5	O	3	1.7E-174	66	Q9SYS6	DRE1C	9.59E-05	Cold acclimation, response to cold, transcription
5	O	3	1.7E-174	66	Q9FGV1	ARF8	1.77E-04	Auxin-activated signaling pathway, flower development, response to auxin, transcription
5	O	4	1.9E-130	66	O82155	DOF1.7	2.09E-09	Regulation of transcription, response to chitin, transcription
5	O	4	1.9E-130	66	Q9FGD6	DOF5.8	5.03E-09	Regulation of transcription, transcription
5	O	4	1.9E-130	66	Q39088	DOF3.4	6.98E-09	Cell wall modification, positive regulation of cell cycle, positive regulation of transcription, response to auxin, response to salicylic acid, transcription

5	O	5	4.2E-97	66	O49397	ARR10	8.49E-03	Cellular response to cytokinin stimulus, cytokinin-activated signaling pathway, maintenance of shoot apical meristem identity, primary root development, regulation of anthocyanin metabolic process, regulation of chlorophyll biosynthetic process, regulation of cytokinin-activated signaling pathway, regulation of root meristem growth, regulation of seed growth, response to cytokinin, response to water deprevivation, root development, shoot system development, transcription
5	O	5	4.2E-97	66	Q9FJW5	TRB2	1.24E-02	Nucleosome assembly, response to abscisic acid, response to auxin, cadmium ion, response to ethylene, response to gibberellin, response to jasmonic acid, response to salicylic acid, response to salt stress, telomeric loop formation, transcription
5	O	5	4.2E-97	66	Q6DBP8	GATA11	1.89E-02	Cell differentiation, regulation of transcription, transcription
5	Z	1	1.8E-279	47	Q1PFR7	LEC2	4.67E-03	Embryo development ending in seed dormancy, positive regulation of auxin biosynthesis process, positive regulation of transcription, seed maturation, seed oilbody biogenesis, somatic embryogenesis, transcription
5	Z	1	1.8E-279	47	Q9ZPY0	DOF2.5	1.33E-02	Cellular response to red light, cellular response to water stimulus, positive regulation of gibberellin biosynthetic process, positive regulation of seed germination, red light signaling pathway, response to cold, response to light stimulus, seed germination, transcription
5	Z	1	1.8E-279	47	O22456	Sep3	1.36E-02	Cell fate specification, flower development, plant ovule development, positive regulation of transcription from RNA polymerase II promoter, specification of floral organ identity, specification of floral organ number, transcription
5	Z	2	8.6E-207	51	Q5DW98	PEND	5.42E-03	Predicted protein
5	Z	2	8.6E-207	51	XXX	XXX	XXX	
5	Z	2	8.6E-207	51	XXX	XXX	XXX	
5	Z	3	3.4E-194	46	Q9SGJ6	DREB1E	7.57E-06	Regulation of gibberellin biosynthesis process, regulation of timing of transcition from vegetative to reproductive phase, transcription
5	Z	3	3.4E-194	46	Q9SYS6	DRE1C	6.87E-05	Cold acclimation, response to cold, transcription
5	Z	3	3.4E-194	46	Q9FGV1	ARF8	1.87E-04	Auxin-activated signaling pathway, flower development, response to auxin, transcription
5	Z	4	2.9E-135	66	O82155	DOF1.7	2.80E-09	Regulation of transcription, response to chitin, transcription
5	Z	4	2.9E-135	66	Q9FGD6	DOF5.8	3.83E-09	Regulation of transcription, transcription
5	Z	4	2.9E-135	66	Q39088	DOF3.4	1.50E-08	Cell wall modification, positive regulation of cell cycle, positive regulation of transcription, response to auxin, response to salicylic acid, transcription
5	Z	5	4.2E-124	37	P29383	AGL3	5.35E-03	Carpel development, cell differentiation, maintenance of floral meristem identity, petal development, positive regulation of transcription from RNA polymerase II promoter, sepal development, stamen development, transcription
5	Z	5	4.2E-124	37	Q38847	AGL15	1.20E-02	Cellular response to auxin stimulus, embryo development ending in seed dormancy, floral organ abscission, fruit abscission, fruit dehiscence, gibberellin catabolic process, negative regulation of floral organ abscission, negative regulation of flower development, negative regulation of gene expression, negative regulation of seed maturation, negative regulation of short-day photoperiodism, negative/positive regulation of transcription, positive regulation of transcription from RNA polymerase II promoter, somatic embryogenesis, transcription

5	Z	5	4.2E-124	37	Q9ZUL3	IDD5	1.88E-02	Positive regulation of starch synthase SS4, development of chloroplast and startch granules, regulation of transcription, transcription

78	O	1	2.7E-23	20	Q9M2U1	OBP3/DOF 3.6	1.60E-10	Regulation of transcription, transcription
78	O	1	2.7E-23	20	Q9FGD6	DOF5.8	7.64E-09	Regulation of transcription, transcription
78	O	1	2.7E-23	20	Q9SEZ3	CDF5	1.40E-08	Flower development, negative regulation of long-/short-day photoperiodism, regulation of transcription, transcription
78	O	2	1.4E-17	20	F4JRB0	HHO5	3.78E-04	Floral organ formation, negative regulation of gene expression, regulation of transcription, specification of plant organ identity, transcription
78	O	2	1.4E-17	20	F4JRB0	HHO5	1.59E-03	Floral organ formation, negative regulation of gene expression, regulation of transcription, specification of plant organ identity, transcription
78	O	2	1.4E-17	20	Q9FGD6	DOF5.8	3.50E-03	Regulation of transcription, transcription
78	O	3	2.5E-20	20	Q9FFH3	NUC	2.91E-04	Asymmetic cell division, leaf senescence, flowering photoperiodism, positive regulation of transcription, regulation of timining of transcition from vegetative to reproductive phase, transcription
78	O	3	2.5E-20	20	Q9ZWA6	MGP	3.26E-04	Asymmetic cell division, regulation of transcription, transcription
78	O	3	2.5E-20	20	O80917	DREB2E	4.40E-04	Abscisic acid-activated signaling pathway, cellular response to heat, positive regulation of transcrition, response to salt stress, response to water deprivation, transcription
78	O	4	0.00012	20	Q9FGD6	DOF5.8	2.74E-04	Regulation of transcription, transcription
78	O	4	0.00012	20	Q39088	DOF3.4	3.52E-04	Cell wall modification, positive regulation of cell cycle, positive regulation of transcription, response to auxin, response to salicyilc acid, transcription
78	O	4	0.00012	20	Q9LZ56	DOF5.1	5.13E-04	Regulation of transcription, transcription
78	O	5	0.00065	20	Q9SEZ3	CDF5	6.49E-04	Flower development, negative regulation of long-/short-day photoperiodism, regulation of transcription, transcription
78	O	5	0.00065	20	Q9FKA0	NAC92	7.65E-04	Lateral root development, leaf senescence, positive regulation of age-related resistance, positive regulation of DNA binding transcription factor activity, positive regulation of leaf senescence, positive regulation of programmed cell death, regulation of gene expression, regulation of seed germination, response to abscisic acid, response to auxin, response to etylene, response to hydrogen peroxide, response to oxidative stress, response to salt, response to salt stress, stress-induced premature senescence, transcription
78	O	5	0.00065	20	Q9S7L2	MYB98	1.56E-03	Embryo sac development, pollen tube guidance, regulation of embryo sac central cell differentiation, regulation of synergid differentiation, regulation of transcription, regulation of transcription from RNA polymerase II promoter, transcription
78	Z	1	2.7E-23	20	Q9M2U1	OBP3/DOF 3.6	1.60E-10	Regulation of transcription, transcription
78	Z	1	2.7E-23	20	Q9LZ56	DOF5.1	7.64E-09	Regulation of transcription, transcription
78	Z	1	2.7E-23	20	Q9SEZ3	CDF5	1.40E-08	Flower development, negative regulation of long-/short-day photoperiodism, regulation of transcription, transcription

78	Z	2	1.9E-22	12	Q38Q39	ERF027	1.20E-03	Ethylene-activated signaling pathway, glucosinolate metabolic process, transcription
78	Z	2	1.9E-22	12	Q38914	ANT	2.35E-03	Cell differentiation, flower development, glucosinolate metabolic process, maintenance of shoot apical meristem identity, regulation of cell proliferation, transcription
78	Z	2	1.9E-22	12	Q8VZP4	GATA10	2.37E-03	Cell differentiation, regulation of transcription, transcription
78	Z	3	8.2E-20	15	F4JRB0	HHO5	5.56E-04	Floral organ formation, negative regulation of gene expression, regulation of transcription, specification of plant organ identity, transcription
78	Z	3	8.2E-20	15	Q9FGD6	DOF5.8	8.57E-04	Regulation of transcription, transcription
78	Z	3	8.2E-20	15	F4JRB0	HHO5	1.56E-03	Floral organ formation, negative regulation of gene expression, regulation of transcription, specification of plant organ identity, transcription
78	Z	4	1.7E-20	19	Q9ZWA6	MGP	2.35E-04	Asymmetic cell division, leaf senescence, flowering photoperiodism, positive regulation of transcription, regulation of timining of transcition from vegetative to reproductive phase, transcription
78	Z	4	1.7E-20	19	Q9FFH3	NUC	2.64E-04	Asymmetic cell division, regulation of transcription, transcription
78	Z	4	1.7E-20	19	Q80917	DREB19	2.81E-04	Abscisic acid-activated signaling pathway, cellular response to heat, positive regulation of transcrition, response to salt stress, response to water deprivation, transcription
78	Z	5	7.2E-21	12	Q39088	DOF2.2	1.96E-05	Regulation of transcription, transcription
78	Z	5	7.2E-21	12	Q9ZV33	DOF3.4	1.21E-04	Cell wall modification, positive regulation of cell cycle, positive regulation of transcription, response to auxin, response to salicyilc acid, transcription
78	Z	5	7.2E-21	12	Q9FGD6	DOF5.8	1.24E-04	Regulation of transcription, transcription

ALL	O	1	4.4e-575	346	Q9SKD0	BPC1	6.92E-13	Regulation of developmental process, regulation of transcription, response to ethylene, transcription
ALL	O	1	4.4e-575	346	F4JUI3	BPC5	3.31E-11	Regulation of transcription, response to ethylene, transcription
ALL	O	1	4.4e-575	346	Q8L999	BPC6	5.00E-11	Regulation of transcription, response to ethylene, transcription
ALL	O	2	8.8e-411	346	Q9M2U1	DOF3.6	5.05E-11	Regulation of transcription, transcription
ALL	O	2	8.8e-411	346	Q9SEZ3	CDF5	1.02E-10	Flower development, negative regulation of long-/short-day photoperiodism, regulation of transcription, transcription
ALL	O	2	8.8e-411	346	Q9LZ56	DOF5.1	1.17E-10	Regulation of transcription, transcription
ALL	O	3	1.70E-230	346	Q9SKD0	BPC1	3.87E-11	Regulation of developmental process, regulation of transcription, response to ethylene, transcription
ALL	O	3	1.70E-230	346	XXX	RAMOSA 1	8.64E-11	Predicted protein
ALL	O	3	1.70E-230	346	Q8L999	BPC6	1.01E-10	Regulation of transcription, response to ethylene, transcription
ALL	O	4	6.70E-146	346	Q9FGD6	DOF5.8	2.37E-11	Regulation of transcription, transcription
ALL	O	4	6.70E-146	346	Q9M2U1	DOF3.6	4.53E-10	Regulation of transcription, transcription
ALL	O	4	6.70E-146	346	Q9LZ56	DOF5.1	1.73E-09	Regulation of transcription, transcription
ALL	O	5	1.50E-133	346	Q9LSL6	DOF5.7	4.12E-03	Guard cell differentiation, positive regulation of transcription, regulation of cell wall pectin metabolic process, regulation of transcription, stomatal movement

ALL	O	5	1.50E-133	346	Q6R053	MYB56	4.21E-03	Brassinosteroid mediated signaling pathway, cell differentiation, cellular response to brassinosteroid stimulus, endothelial cell proliferation, integument development, negative regulation of cell division, negative resulation of long-day photoperiodism, positive regulation of transcription, quiescent center organization, regulation of seed growth, regulation of transcription, regulation of transcription from RNA polymperase II promoter, transcription
ALL	O	5	1.50E-133	346	Q9ZPX0	GATA20	8.07E-03	Cell differentiation, regulation of transcription
ALL	Z	1	2.7e-527	346	Q9SKD0	BPC1	4.07E-11	Regulation of developmental process, regulation of transcription, response to ethylene, transcription
ALL	Z	1	2.7e-527	346	Q8L999	BPC6	6.69E-11	Regulation of transcription, response to ethylene, transcription
ALL	Z	1	2.7e-527	346	xxx	RAMOSA1	1.34E-10	
ALL	Z	2	1.7e-333	345	Q9M2U1	DOF3.6	4.25E-10	Regulation of transcription, transcription
ALL	Z	2	1.7e-334	345	Q9SEZ3	CDF5	2.76E-09	Flower development, negative regulation of long-/short-day photoperiodism, regulation of transcription, transcription
ALL	Z	2	1.7e-335	345	Q9FGD6	DOF5.8	5.73E-09	Regulation of transcription, transcription
ALL	Z	3	3.20E-245	137	Q9ZPX0	GATA20	1.32E-03	Cell differentiation, regulation of transcription
ALL	Z	3	3.2E-245	137				
ALL	Z	3	3.2E-245	137				
ALL	Z	4	3E-217	295		RAMOSA1	9.5E-10	
ALL	Z	4	3.00E-217	295	Q9SKD0	BPC1	2.14E-09	Regulation of developmental process, regulation of transcription, response to ethylene, transcription
ALL	Z	4	3.00E-217	295	Q8L999	BPC6	6.50E-09	Regulation of transcription, response to ethylene, transcription
ALL	Z	5	2.30E-189	117	Q9SLH1	MYB81	6.63E-03	Cell differentiation, regulation of transcription from RNA polymerase II promoter
ALL	Z	5	2.30E-189	117	P42736	RAP2-3	6.77E-03	Cell death, ethylene-activated signaling pathway, heat acclimation, positive regulation of transcription, response to cytokinin, response to ethylene, response to jasmonic acid, response to other organism, transcription
ALL	Z	5	2.30E-189	117	P93007	ERF112	8.23E-03	Ethylene-activated signaling pathway, transcription

1-B	O	1	1.6e-073	15	Q9FGT7	ARR18	4.26E-03	Cytokinin signaling pathway, Transcription, Transcription regulation, Two-component regulatory system
1-B	O	1	1.6e-073	15	Q9FPE8	HHO3	1.70E-02	Transcription, Transcription regulation; Probable transcription factor involved in phosphate signaling in roots.
1-B	O	1	1.6e-073	15	Q9ZPX0	GATA20	4.56E-03	Transcription, Transcription regulation; Transcriptional regulator that specifically binds 5'-GATA-3' or 5'-GAT-3' motifs within gene promoters
1-B	O	2	3.5e-050	15	O49687	MYC4	7.32E-03	Anthocyanin-containing compound biosynthetic process, jasmonic acid gene regulation, defense response, positive regulation of transcription, regulation of transcription, transcription
1-B	O	2	3.5e-050	15	Q9LSL6	DOF5.7	5.02E-03	Guard cell differentiation, positive regulation of transcription, regulation of cell wall pectin metabolic process, regulation of transcription, stomatal movement

1-B	O	2	3.5e-050	15	Q9SEZ3	CDF5	7.14E-03	Flower development, negative regulation of long-/short-day photoperiodism, regulation of transcription, transcription
1-B	O	3	6.6e-048	15	Q39088	DOF3.4	1.52E-03	Cell wall modification, positive regulation of cell cycle, positive regulation of transcription, response to auxin, response to salicyilc acid, transcription
1-B	O	3	6.6e-048	15	Q9LZ56	DOF5.1	3.45E-03	Regulation of transcription, transcription
1-B	O	3	6.6e-048	15	Q9ZV33	DOF2.2	9.24E-04	Regulation of transcription, transcription
1-B	O	4	2.4e-037	15	C0SVS4	PHL11	2.92E-04	Transcription, Transcription regulation
1-B	O	4	2.4e-037	15	F4IPE3	SGR5	4.78E-04	Circumnution, detection of gravity, floral organ morphogenesis, gravitropism, leaf morphogenesis, positive regulation of auxin biosynthetic process, regulation of auxin polar transport, regulation of startch metabolic process, regulation of transcription
1-B	O	4	2.4e-037	15	Q9LSL6	DOF5.7	5.02E-03	Guard cell differentiation, positive regulation of transcription, regulation of cell wall pectin metabolic process, regulation of transcription, stomatal movement
1-B	O	5	5.0e-032	15	F4IPE3	SGR5	4.78E-04	Circumnution, detection of gravity, floral organ morphogenesis, gravitropism, leaf morphogenesis, positive regulation of auxin biosynthetic process, regulation of auxin polar transport, regulation of startch metabolic process, regulation of transcription
1-B	O	5	5.0e-032	15	F4JN35	NTL9	1.12E-03	Cellular response to osmotic stress, negative regulation of transcription, positive regulation of defense responser to bacterium, regulation of defense response, regulation of transcription, transcription
1-B	O	5	5.0e-032	15	Q8GYC1	IDD4	2.47E-03	Assymetic cell division, root development, transcription
2-B	O	1	2.3e-153	24	Q01593	ABI3	6.16E-03	Abscisic acid-activated signaling pathway, embryo development, mitochondira-nucleus signaling pathway, plastid organization, postive regulation of transcription, response to abscisic acid, response to auxin, transcription
2-B	O	1	2.3e-153	24	Q38Q40	ERF122	8.78E-03	Ethylene signaling pathway, Transcription, Transcription regulation
2-B	O	1	2.3e-153	24	Q8L7W9	abi4	1.26E-02	Abscisic acid-activated signaling pathway, seed development, transcription
2-B	O	2	2.7e-130	24	Q8L9Y3	ARR14	6.35E-03	Cytokinin-activated signaling pathway, transcription, two-component regulatory system, activate some type-A response regulators in response to cytokinins
2-B	O	2	2.7e-130	24	Q9FXD6	ARR11	2.91E-03	Cytokinin-activated signaling pathway, regulation of root meristem growth, response to cytokinin, transcription, two-component regulatory system
2-B	O	2	2.7e-130	24	Q9ZWJ9	ARR2	1.31E-02	Cellular response to cytokinin stimulus, cytokinin-activated signaling pathway, ethylene-activated signaling pathway, leaf senescence, regulation of root meristem, regulation of seed growth, regulation of stomatal movement, response to cytokinin, response to ethylene, root development, transcription
2-B	O	3	5.2e-103	24	A9TN45	Predicted protein	9.55E-04	
2-B	O	3	5.2e-103	24	P92948	CDC5	5.01E-04	Cell cycle, cell differentiation, defense response signaling pathway, defense response to bacterium, defense response to fungus, DNA repair, innate immune response, mRNA processing, regulation of transcription, regulation of transcription from RNA polymerase II promoter, RNA splicing, transcription
2-B	O	3	5.2e-103	24	Q9LTC7	BHLH34	5.37E-04	Regulation of transcription, transcription
2-B	O	4	3.0e-088	24	F4JRB0	HHO5	6.46E-03	Floral organ formation, negative regulation of gene expression, regulation of transcription, specification of plant organ identity, transcription

2-B	O	4	3.0e-088	24	Q9LVR0	ATHB-53	3.57E-03	Regulation of auxin/cytokinin signaling during root development, response to auxin, root development, transcription
2-B	O	4	3.0e-088	24	Q9ZPX0	GATA20	2.84E-03	Transcription, Transcription regulation; Transcriptional regulator that specifically binds 5'-GATA-3' or 5'-GAT-3' motifs within gene promoters
2-B	O	5	1.1e-074	24	Q75HX9	Predicted protein	1.67E-05	
2-B	O	5	1.1e-074	24	Q9M7Q3	ABF3	1.67E-05	Abscisic-acid activated signaling pathway, response to abscisic acid, response to salt stress, response to water deprivation, transcription
2-B	O	5	1.1e-074	24	Q9SJN0	ABI5	1.54E-05	Abscisic-acid activated signaling pathway, negative regulation of seed germination, positive regulation of transcription, response to abscisic acid, response to chitin, response to gibberellin, response to salt stress, response to water deprevation, seed development, seed germination, sugar mediated signaling pathway, transcirption
3-B	O	1	2.30E-86	16	Q02994	myb.Ph3	3.80E-03	DNA binding
3-B	O	1	2.30E-86	16	Q8VZS3	HHO2	4.83E-03	DNA binding, DNA binding transcription factor activity, phosphate ion homeostasis, regulation of transcription, DNA-templated, response to abscisic acid, transcription, DNA-templated
3-B	O	1	2.30E-86	16	Q9FX84	HHO6	8.38E-03	DNA binding, DNA binding transcription factor activity, regulation of transcription, DNA-templated, transcription, DNA-templated
3-B	O	2	6.70E-86	16	Q6NKN9	BHLH74	4.11E-03	DNA binding, DNA binding transcription factor activity, protein dimerization activity, regulation of growth, regulation of transcription, DNA-templated, response to blue light, transcription, DNA-templated
3-B	O	2	6.70E-86	16	Q01593	ABI3	6.02E-03	DNA binding, DNA binding transcription factor activity, transcription factor activity, RNA polymerase II transcription factor binding, abscisic acid-activated signalling pathway, embryo development ending in seed dormancy, mitochondria-nucleus signalling pathway, plastid organization, positive regulation of transcription, DNA-templated, response to abscisic acid, response to auxin, transcription, DNA-templated
3-B	O	2	6.70E-86	16	Q5YGP8	PLT1	1.52E-02	DNA binding, DNA binding transcription factor activity, auxin-activated signaling pathway, ethylene-activated signaling pathway, pattern specification process, root development, root meristem growth, stem cell population maintenance, telomere maintenance, transcription, DNA-templated
3-B	O	3	4.30E-79	16	Q49550	DOF4.5	5.85E-03	DNA binding, DNA binding transcription factor activity, metal ion binding, regulation of transcription, DNA-templated, transcription, DNA-templated
3-B	O	3	4.30E-79	16	Q9SQQ6	NAC046	7.18E-03	DNA binding, DNA binding transcription factor activity, positive regulation of chlorophyll catabolic process, positive regulation of leaf senescence, transcription, DNA-templated
3-B	O	3	4.30E-79	16	Q9LS50	NAC058	1.18E-02	DNA binding, DNA binding transcription factor activity, transcription, DNA-templated
3-B	O	4	2.00E-72	16	A9RZ73	PHYPADRAFT_72483	2.07E-03	protein dimerization activity
3-B	O	4	2.00E-72	16	Q9CAA4	BIM2	3.31E-03	DNA binding transcription factor activity, protein dimerization activity, sequence-specific DNA binding, transcription, DNA-templated
3-B	O	4	2.00E-72	16	Q49404	BEH3	3.91E-03	DNA binding, DNA binding transcription factor activity, brassinosteroid mediated signaling pathway, regulation of transcription, DNA-templated, transcription, DNA-templated
3-B	O	5	2.20E-48	16	I3NN78	ZF2	2.74E-04	nucleic acid binding

3-B	O	5	2.20E-48	16	B9GPL8	POPTR_0002s00440g	1.52E-03	DNA binding, metal ion binding
3-B	O	5	2.20E-48	16	Q9SMX9	SPL1	1.92E-03	DNA binding, DNA binding transcription factor activity, metal ion binding, regulation of transcription, DNA-templated, transcription, DNA-templated
5-B	O	1	4.10E-57	20	Q1PFR7	LEC2	2.58E-04	DNA binding, DNA binding transcription factor activity, embryo development ending in seed dormancy, positive regulation of auxin biosynthetic process, positive regulation of transcription, DNA-templated, seed maturation, seed oilbody biogenesis, somatic embryogenesis, transcription, DNA-templated
5-B	O	1	4.10E-57	20	Q9FY74	CAMTA1	1.23E-03	calmodulin binding, sequence-specific DNA binding, transcriptional activatory activity, RNA polymerase II proximal promoter sequence-specific DNA binding, positive regulation of transcription, DNA-templated, positive regulation of transcription from RNA polymerase II promoter, response to auxin, response to freezing, response to water deprivation
5-B	O	1	4.10E-57	20	Q9LIE5	FHY3	1.41E-03	DNA binding transcription factor activity, zinc ion binding, circadian rhythm, far-red light signaling pathway, positive regulation of circadian rhythm, positive regulation of transcription, DNA-templated, red, far-red light phototransduction, red or far-red light signaling pathway, response to far red light, transcription, DNA-templated
5-B	O	2	2.20E-32	20	Q5DW98	PEND	7.67E-03	
5-B	O	2	2.20E-32	20	Q9FM03	DOF5.6	1.57E-02	DNA binding, DNA binding transcription factor activity, metal ion binding, phloem or xylem histogenesis, positive regulation of transcription, DNA-templated, procambium histogenesis, regulation of transcription, DNA-templated, transcription, DNA-templated
5-B	O	2	2.20E-32	20	Q84TE9	DOF5.3	1.67E-02	DNA binding transcription factor activity, metal ion binding, transcription regulatory region DNA binding, regulation of transcription, DNA-templated, root development, transcription, DNA-templated
5-B	O	3	3.90E-36	20	O82155	DOF1.7	1.64E-06	DNA binding transcription factor activity, metal ion binding, transcription regulatory region DNA binding, regulation of transcription, DNA-templated, response to chitin, transcription, DNA-templated
5-B	O	3	3.90E-36	20	Q9SEZ3	CDF5	3.75E-06	DNA binding, DNA binding transcription factor activity, metal ion binding, flower development, negative regulation of long-day photoperiodism, flowering, negative regulation of short-day photoperiodism, flowering, regulation of transcription, DNA-templated, transcription, DNA-templated
5-B	O	3	3.90E-36	20	Q9FGD6	DOF5.8	3.85E-06	DNA binding transcription factor activity, metal ion binding, transcription regulatory region sequence-specific DNA binding, regulation of transcription, DNA-templated, transcription, DNA-templated
5-B	O	4	1.50E-38	20			9.12E-10	
5-B	O	4	1.50E-38	20	Q9SKD0	BPC1	1.40E-07	DNA binding, sequence-specific DNA binding, regulation of developmental process, regulation of transcription, DNA-templated, response to ethylene, transcription, DNA-templated
5-B	O	4	1.50E-38	20	Q8L999	BPC6	1.78E-07	protein homodimerization activity, sequence-specific DNA binding, regulation of transcription, DNA-templated, response to ethylene, transcription, DNA-templated
5-B	O	5	2.50E-37	20	Q9FGV1	ARF8	1.58E-04	DNA binding transcription factor activity, sequence-specific DNA binding, auxin-activated signaling pathway, flower development, response to auxin, transcription, DNA-templated
5-B	O	5	2.50E-37	20	Q9SGJ6	DREB1E	4.27E-04	DNA binding, DNA binding transcription factor activity, regulation of gibberellin biosynthetic process, regulation of timing of transition from vegetative to reproductive phase, transcription, DNA-templated

5-B	O	5	2.50E-37	20	Q94JM3	ARF2	5.19E-04	DNA binding transcription factor activity, sequence-specific DNA binding, auxin-activated signaling pathway, floral organ abscission, fruit dehiscence, leaf senescence, negative regulation of cell proliferation, negative regulation of transcription, DNA-templated, plant ovule development, positive regulation of flower development, positive regulation of potassium ion import, response to abscisic acid, transcription, DNA-templated
78-B	O	1	3.40E-21	19	Q9M2U1	DOF3.6	1.22E-10	DNA binding, metal ion binding, regulation of transcription, DNA-templated, transcription, DNA-templated
78-B	O	1	3.40E-21	19	Q9LZ56	DOF5.1	5.57E-09	DNA binding, DNA binding transcription factor activity, metal ion binding, regulation of transcription, DNA-templated, transcription, DNA-templated
78-B	O	1	3.40E-21	19	Q9SEZ3	CDF5	1.22E-08	DNA binding, DNA binding transcription factor activity, metal ion binding, flower development, negative regulation of long-day photoperiodism, flowering, negative regulation of short-day photoperiodism, flowering, regulation of transcription, DNA-template
78-B	O	2	2.90E-18	19	F4JRB0	HHO5	3.93E-04	DNA binding transcription factor activity, transcription regulatory region DNA binding, floral organ formation, negative regulation of gene expression, regulation of transcription, DNA-templated, specification of plant organ identity, transcription, DNA-template
78-B	O	2	2.90E-18	19	F4JRB0	HHO6	1.85E-03	DNA binding transcription factor activity, transcription regulatory region DNA binding, floral organ formation, negative regulation of gene expression, regulation of transcription, DNA-templated, specification of plant organ identity, transcription, DNA-template
78-B	O	2	2.90E-18	19	Q9FGD6	DOF5.8	2.67E-03	DNA binding transcription factor activity, metal ion binding, transcription regulatory region sequence-specific DNA binding, regulation of transcription, DNA-templated, transcription, DNA-templated
78-B	O	3	3.60E-21	19	O80917	DREB2E	2.31E-04	DNA binding transcription factor activity, sequence-specific DNA binding, transcription regulatory region DNA binding, abscisic acid-activated signaling pathway, cellular response to heat, positive regulation of transcription, DNA-templated, response to salt stress, response to water deprivation, transcription, DNA-templated
78-B	O	3	3.60E-21	19	Q9ZWA6	MGP	2.86E-04	DNA binding transcription factor activity, metal ion binding, transcription regulatory region DNA binding, asymmetric cell division, regulation of transcription, DNA-templated, transcription, DNA-templated
78-B	O	3	3.60E-21	19	Q9FFH3	NUC	2.86E-04	DNA binding, DNA binding transcription factor activity, metal ion binding, asymmetric cell division, leaf senescence, photoperiodism, flowering, positive regulation of transcription, DNA-templated, regulation of timing of transision from vegetative to reproductive phase, regulation of transcription, DNA-templated, transcription, DNA-templated
78-B	O	4	1.60E-07	19	Q39088	DOF3.4	1.17E-03	DNA binding, DNA binding transcription factor activity, metal ion binding, transcription regulatory region DNA binding, cell wall modification, positive regulation of cell cycle, positive regulation of transcription, DNA-templated, response to auxin, response to salicylic acid, transcription, DNA-templated
78-B	O	4	1.60E-07	19	Q9FGD6	DOF5.8	3.77E-03	DNA binding transcription factor activity, metal ion binding, transcription regulatory region sequence-specific DNA binding, regulation of transcription, DNA-templated, transcription, DNA-templated
78-B	O	4	1.60E-07	19	P68350	DOF1.5	7.71E-03	DNA binding transcription factor activity, metal ion binding, transcription regulatory region DNA binding, regulation of transcription, DNA-templated, seed coat development, transcription, DNA-templated

78-B	O	5	1.10E-03	19	Q9FKA0	NAC92	4.61E-03	DNA binding transcription factor activity, protein homodimerization activity, sequence-specific DNA binding, lateral root development, leaf senescence, positive regulation of age-related resistance, positive regulation of DNA binding transcription factor activity, positive regulation of leaf senescence, positive regulation of programmed cell death, regulation of gene expression, regulation of seed germination, response to abscisic acid, response to auxin, response to ethylene, response to hydrogen peroxide, response to oxidative stress, response to salt, response to salt stress, stress-induced premature senescence, transcription, DNA-templated
78-B	O	5	1.10E-03	19	Q9SAH7	WRKY40	1.10E-02	DNA binding transcription factor activity, sequence-specific DNA binding, transcription regulatory region DNA binding, defense response to bacterium, defence response to fungus, negative regulation of transcription, DNA-templated, regulation of defense response, response to chitin, response to molecule of bacterial origin, response to salicylic acid, response to wounding, transcription, DNA-templated
78-B	O	5	1.10E-03	19	Q9SJ09	WRKY59	1.21E-02	DNA binding transcription factor activity, sequence-specific DNA binding, transcription, DNA-templated
ALL-B	O	1	3.30E-190	95	Q9SKD0	BPC1	1.27E-13	DNA binding, sequence-specific DNA binding, regulation of developmental process, regulation of transcription, DNA-templated, response to ethylene, transcription, DNA-templated
ALL-B	O	1	3.30E-190	95	F4JUI3	BPC5	2.00E-12	sequence-specific DNA binding, regulation of transcription, DNA-templated, response to ethylene, transcription, DNA-templated
ALL-B	O	1	3.30E-190	95	Q8L999	BPC6	3.25E-12	protein homodimerization activity, sequence-specific DNA binding, regulation of transcription, DNA-templated, response to ethylene, transcription, DNA-templated
ALL-B	O	2	1.90E-130	95	Q9FGD6	DOF5.8	3.58E-11	DNA binding transcription factor activity, metal ion binding, transcription regulatory region sequence-specific DNA binding, regulation of transcription, DNA-templated, transcription, DNA-templated
ALL-B	O	2	1.90E-130	95	Q9LZ56	DOF5.1	3.80E-11	DNA binding, DNA binding transcription factor activity, metal ion binding, regulation of transcription, DNA-templated, transcription, DNA-templated
ALL-B	O	2	1.90E-130	95	Q9M2U1	DOF3.6	5.53E-11	DNA binding, metal ion binding, regulation of transcription, DNA-templated, transcription, DNA-templated
ALL-B	O	3	5.60E-101	95				
ALL-B	O	3	5.60E-101	95				
ALL-B	O	3	5.60E-101	95				
ALL-B	O	4	3.50E-95	95	Q9C8P8	BHLH80	1.36E-03	core promoter sequence-specific DNA binding, DNA binding transcription factor activity, protein dimerization activity, transcriptional activator activity, RNA polymerase II transcription regulatory region sequence-specific DNA binding, cuticle development, regulation of transcription, DNA-templated, transcription from RNA polymerase II promoter
ALL-B	O	4	3.50E-95	95	Q94JL3	BHLH112	1.66E-03	DNA binding transcription factor activity, protein dimerization activity, RNA polymerase II proximal promoter sequence-specific DNA binding, sequence-specific DNA binding, transcriptional activator activity, RNA polymerase II transcription regulatory region sequence-specific DNA binding, cellular response to abscisic acid stimulus, cellular response to salt stress, cellular response to water deprivation, positive regulation of transcription, DNA-templated, regulation of proline metabolic process, regulation of reactive oxygen species metabolic process, transcription from RNA polymerase II promoter

ALL-B	O	4	3.50E-95	95	Q66GR3	BHLH130	2.87E-03	core promoter sequence-specific DNA binding, DNA binding transcription factor activity, protein dimerization activity, transcriptional activator activity, RNA polymerase II transcription regulatory region sequence-specific DNA binding, photoperiodism, flowering, regulation of transcription, DNA-templated, transcription from RNA polymerase II promoter
ALL-B	O	5	3.90E-70	95	Q9M2U1	DOF3.6	6.12E-10	DNA binding, metal ion binding, regulation of transcription, DNA-templated, transcription, DNA-templated
ALL-B	O	5	3.90E-70	95	Q9SEZ3	CDF5	1.49E-09	DNA binding, DNA binding transcription factor activity, metal ion binding, flower development, negative regulation of long-day photoperiodism, flowering, negative regulation of short-day photoperiodism, flowering, regulation of transcription, DNA-templated, transcription, DNA-templated
ALL-B	O	5	3.90E-70	95	Q9FGD6	DOF5.8	1.92E-07	DNA binding transcription factor activity, metal ion binding, transcription regulatory region sequence-specific DNA binding, regulation of transcription, DNA-templated, transcription, DNA-templated