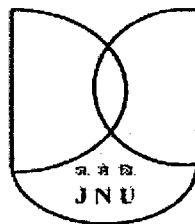


**Investigation of miRNA mediated regulation of the cell cycle
gene '*cdc2aAt*' in *Arabidopsis thaliana*
-A Computational Approach**

A thesis submitted in partial fulfillment of the requirements for
the award of the degree of

Master of Technology

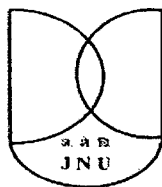


Ravi Kumar Singh
07 / 75 / MT / 10

Centre for Computational Biology and Bioinformatics
School of Information Technology
Jawaharlal Nehru University
New Delhi – 110067

May 2009

*dedicated
to
my parents*



**CENTRE FOR COMPUTATIONAL BIOLOGY AND BIOINFORMATICS
SCHOOL OF INFORMATION TECHNOLOGY
JAWAHARLAL NEHRU UNIVERSITY
NEW DELHI – 110067.**

CERTIFICATE

Certified that this M.Tech. dissertation titled “ **Investigation of miRNA mediated regulation of the cell cycle gene 'cdc2aAt' in *Arabidopsis thaliana* -A Computational Approach** “ has been carried out by **Mr. Ravi Kumar Singh** under my guidance and supervision. This work is original and has not been submitted elsewhere for award of any degree or diploma.

Prof. Indira Ghosh
Dean
School of Information
Technology,
JNU, New Delhi.

Dr. A. Krishnamachari
Supervisor
School of Information
Technology,
JNU, New Delhi.

**Investigation of miRNA mediated regulation of the cell cycle
gene ' *cdc2aAt* ' in *Arabidopsis thaliana* -
A Computational Approach**

(ABSTRACT)

Study and Prediction of microRNA function (target prediction and its validation) and its regulation has been considered a very challenging problem. There are some good computational approaches which are used for investigating such types of problems. Evidences suggests that plant cell cycle gene *cdc2aAt* is down-regulated during different stress conditions. The role of miRNAs in down-regulation of *cdc2aAt* and its possible mechanism is unknown. We investigated this aspect through *in silico* approach. Comparative computational analysis shows that miR447 family and miR161.1 are the putative candidates which targets *cdc2aAt*. Analysis of upstream sequence of these miRNAs reveals that it contain stress responsive cis-elements, such as AtMYC2, AtMYB and other members of MYB family, SED, ASR1, DRE-like promoters, GCC-box motif etc. These result further support our hypothesis that miR447 family and miR161.1 are very likely to be involved in the regulation of *cdc2aAt* during different stress conditions.

Acknowledgement

First and foremost I express my sincere and deepest gratitude to my supervisor, Dr. A. Krishnamachari, for being a constant source of inspiration, support and guidance throughout the dissertation work. I thank him for introducing me to this interesting area of research. I am grateful to him for patiently listening to my all ideas, meticulously going over cumbersome results, giving valuable direction to my work, and always being there to talk. Without his constant support and motivation this work could not have been completed successfully.

My sincere thanks to our Dean Prof. Indira Ghosh to make available whatever resources required to carry out this study.

I am grateful to my all teachers for their valuable guidance and academic helps.

My special thanks to Prof. Rahul Roy and Prof. N. Parimala for giving their precious academic suggestions.

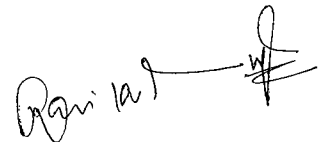
I would like to acknowledge my colleagues for their support to carry out this work.

I would also like to acknowledge Dr. Mitali Mukerji (Scientist E-II) Lab at IGIB, New Delhi for providing facility to use TRANSFAC PROFESSIONAL which help in successful completion of study.

I would also like to thank the staff of the school for their support and help.

I am also grateful to U.G.C. and D.B.T. for providing financial assistance.

I owe great gratitude for my family especially my parents who always love, motivate and support me for higher education. Love and blessing of my parents and siblings gives me immense strength to complete my project.



Ravi Kumar Singh

Table of Contents

List of Figures	
List of Tables	
1. Introduction	1
1.1 microRNAs	3
1.2 <i>cdc2aAt</i>	3
1.3 Target	4
1.4 Transcription Factors	5
1.5 Promoter	6
1.6 <i>cis</i> -elements	6
1.7 Motif	7
1.8 Consensus Sequences	7
1.9 RNA silencing	8
1.10 Objective of the study	11
2. Materials and Methods	12
2.1 Data Sources	12
2.1.1 microRNAs Data	12
2.1.2 Core Cell Cycle Gene Data	12
2.1.3 Whole Genome Sequence Data	12
2.1.4 Differential Expression Data	12
2.1.5 <i>Cis</i> -Element Data	13
2.2 Computational Tools and Methodologies	14
2.2.1 psRNATarget	14
2.2.2 MicroInspector	16
2.2.3 miRanda	17
2.2.4 RNAhybrid	18
2.2.5 Upstream Analysis	19
2.2.5.1 MatchTM Program	19

2.2.5.2 Pattern Matching Program	19
3. Results and Discussion	22
4. Conclusions	39
References	41
Appendix A	
Appendix B	
Appendix C	

List of Figures

1. miRNAs are induced and regulated by environmental biotic and abiotic stresses	2
2. Graphical representation of expression profile of <i>cdc2aAt</i> in root system during different stress conditions	4
3. Approximate base pairing pattern between binding site on the target and miRNA	5
4. Schematic representation of eukaryotic promoter	7
5. A model for miRNA biogenesis, regulation of miRNA gene regulation, and the pathway of miRNA function	10
6. Web interface of psRNA Target	15
7. A flow-chart diagram of the steps for identifying stress responsive cis-elements in the upstream region of miRNAs genes	21
8. Secondary structure of candidate miRNAs and its complementary target sequence duplex	24
9. Schematic representation of putative complementary sites of candidate miRNAs on <i>cdc2aAt</i> transcript	25
10. Schematic representation of the details of the positions of stress response cis-elements in the upstream of miRNAs genes (from Match TM result)	30
11. Schematic representation of the details of the positions of stress response cis-elements in the upstream of miRNAs genes (from Pattern matching result) ..	33
12. Expression profile of some common stress responsive TFs during different stress condition	37
13. Proposed miRNAs mediated regulatory pathways for <i>cdc2aAt</i>	38

List of Tables

1. List of putative microRNAs for cdc2aAt predicted by different predicting tools 22
2. The minimum free energy value of some selected candidate miRNAs and its target (cdc2aAt) duplex 22
3. List of possible core cell cycle genes co-suppressed by miRNAs 26
4. List of putative miRNAs for other downregulated core cell cycle genes. 27
5. Stress-response miRNAs identified by microarray analysis 28
6. List of putative stress response and some other common cis-elements in the upstream of the miRNAs genes (from MatchTM result) 29
7. List of putative stress response cis-elements in the upstream of the miRNAs genes (from Pattern Matching Algorithm) 32
8. Putative stress responsive binding elements in the upstream of miR447 family, miR161.1, miR396 family, miR168 family, miR171 family 35

Introduction

CHAPTER 1 : Introduction

During course of evolution, progressive development of diversified and complex gene regulatory systems emerges at each and every hierarchy within the cellular machinery. These regulatory systems contains different regulatory factors and its families which gives rise to different genetic circuits. A decade ago, a new remarkable unified mechanistic view of a gene regulation system at post-transcriptional level has been deciphered (Lee et al., 1993; Wightman et al., 1993). This involves a large and growing class of approximately 21-23 nucleotide long non-coding RNAs, known as microRNAs (miRNAs), a class of small silencing RNAs, along with other different factors, which play important regulatory role in both plants and animals (Chen and Rajewsky, 2007; Takuno and Innan, 2008).

Plant growth and development is affected by various types of environmental biotic and abiotic stresses such as drought, salinity, extreme temperatures, chemical toxicity, oxidative stress, pests, micro pathogens etc (Levitt J., 1980; Azaizeh & Steudle, 1991; Gueta-Dahan et al.,1997). These stresses gives rise to a series of changes in the plant body that affects its growth and productivity, resulting in an adaptive growth response. These changes are mainly consequences of alterations in cellular responses, such as production of stress responsive proteins (Adams et al., 1992; Bohnert et al., 1995; Jain & Selvaraj, 1997; Jian-Kang, 2002; Wangxia, 2003). It has been shown that environmental stresses interferes at different levels of gene expression including transcriptional and post-transcriptional level (Jain & Selvaraj, 1997; Ishitani et al., 1997; Jaglo-ottosen et al.,1998). It is known in the literature that there is activation of a large number of microRNAs and other small regulatory RNAs during environmental stresses in eukaryotes (Fig:1) (Chiou 2007; Jones-Rhoades & Bartel, 2004; Sunkar & Zhu, 2004; Zhang B. H. et. al.,2005). It is also been clearly established that non-coding RNAs (ncRNAs) such as microRNAs (miRNAs), has a major role in endogenous gene silencing (Aukerman & Sakari,

2003; Bartel,2004; Chen, 2004; Erica and O'Driscoll, 2008).

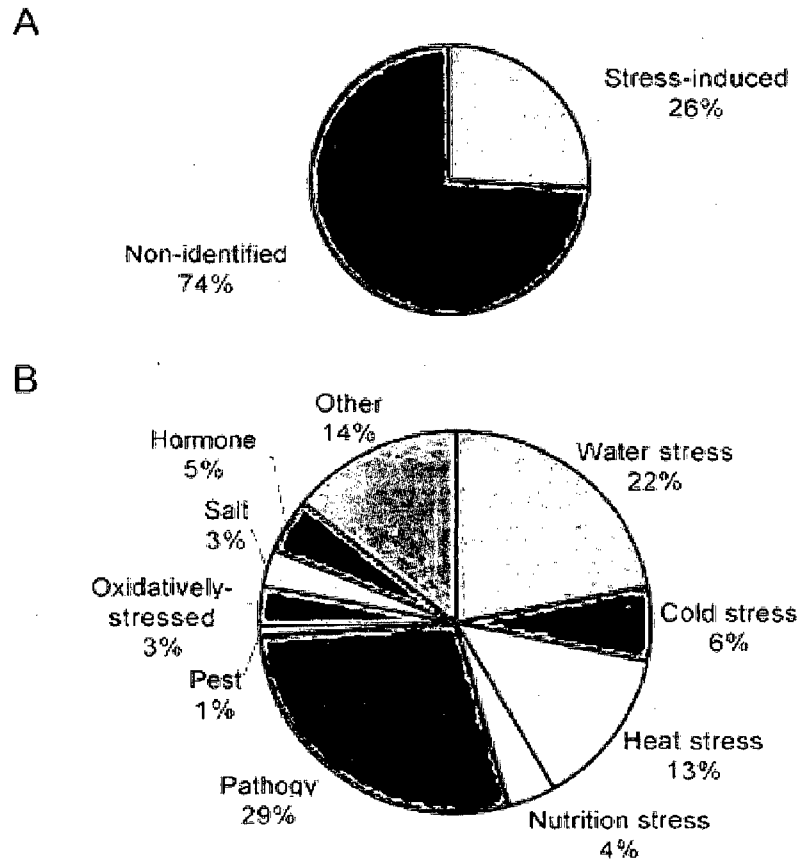


Fig 1 : miRNAs are induced and regulated by environmental biotic and abiotic stresses. (A) 26% of ESTs were obtained from stress-induced tissues. (B) ESTs containing miRNAs were obtained from tissue by various biotic and abiotic environmental stresses (Source - Zhang et al. 2005).

Study of gene regulation by miRNAs and its pathways during different stress conditions have been a challenging problem for biologists. Many attempts have been made by *in silico* approach to understand how different miRNAs assign and regulate their targets and what could be the possible pathways for that regulation (Tsang et al.,2007; Zhou et al.,2007; Joung & Zhangjun, 2008). There are several target prediction software tools which predict targets for different miRNAs on different biophysical parameters (Dai & Zhao, 2009; Rehmsmeier et al., 2004;

Rusinov et al., 2005). However, most of them are designed for animal models and there is a need to develop a methodology for the plant systems.

Background

1.1 MicroRNAs

MicroRNAs (miRNAs) are specialized class of ~22-nt long, non coding, endogenous, single stranded regulatory RNAs that regulate the expression of target gene at the post-transcriptional level in a sequence specific manner. It usually down-regulate the target gene and rarely up-regulate also (Bartel 2004; Chen 2005; Bhattacharyya et al. 2006; Vasudevan et al. 2007). It plays numerous crucial roles at each major stage of development such as proliferation , differentiation, cell death and interaction between viruses and host cell. In plants miRNAs are also involved in different physiological processes, such as auxin signaling, phase transition, flowering and leaf development (Chen 2005; Jones-Rhoades et al 2006; Mallory & Vaucheret 2006; Voinnet 2009).

microRNAs genes are typically present at intergenic areas however it can also be present within introns of known genes, particularly in animals (Baskerville S. and Bartel D.P.,2005). These intronic miRNA gene can down-regulate the host protein-coding gene by targeting the untranslated region (UTR)(Li Sung-Chou et. al., 2007). In plants, some genes may be clustered and transcribed together (Reinhart B.J. et.al ,2002 ;Jones-Rhoades M.W. et al., 2006). In an organism, it is estimated that the total number of miRNA gene may be up to ~1% of all the genes. But recent data suggests that, in human, it is around 3% of all genes(Hennessy & O'Driscoll 2008).

1.2 *cdc2aAt* (cell cycle gene)

This is a class-A type cyclin dependent kinase (CDK) gene which is constitutively expressed during the cell cycle and functionally involved at the G1-

to-S and G2-to-M phase transition. There is a transient decrease in the expression of *cdc2aAt* during salt stress condition specially in the root system (BursSENS et al. 2000). Stress-response microarray data also shows that the expression level of *cdc2aAt* decreases during different stress conditions (Fig - 2).

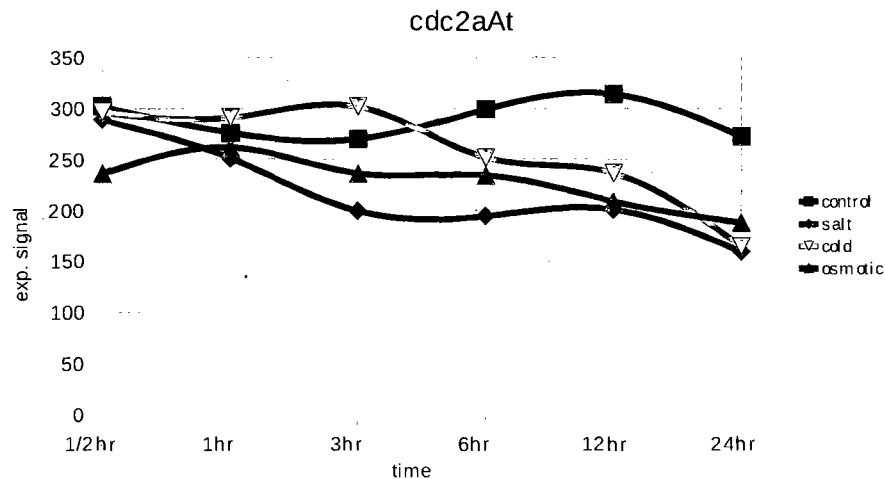


Fig 2 : Graphical representation of expression profile of *cdc2aAt* in root system during different stress conditions (Source : NASC Arrays).

1.3 Target

A target is either a gene or its transcript to which miRNAs act upon. Due to small size of miRNAs it is a challenging task to predict a target with high specificity. It has been assumed that miRNAs recognize their targets through base pairing, in general. It has been shown that there is a high complementarity between plant miRNAs and their mRNA targets than in the animal (Rhoades M. W. et. al., 2002; Bartel 2004). This aspect was exploited in computational prediction of targets in plants.

miRNAs binds to mRNAs in a sequence specific manner. Evidences suggests that in animals, target site is predominantly found in 3' UTR of protein coding genes (Rajewsky N. 2006). While in plant they are predominantly found in coding regions, some times in 3' UTR and rarely in 5' UTR (Wang X. J. et. al. 2004; Jones-Rhoades and Bartel, 2004). On the basis of pattern, the binding site can be

classified in – (1) **5' -dominant canonical**, have perfect base pairing to the 5' end seed region and extensive base pairing to the 3' end with bulge in the middle. - (2) **5'- dominant seed only**, have perfect base pairing to the seed region and limited to the 3' end. - (3) **3'- compensatory**, have extensive pairing to the 3' end and a mismatch or wooble in the seed region (Maziere and Enright, 2007; Brennecke et. al., 2005).

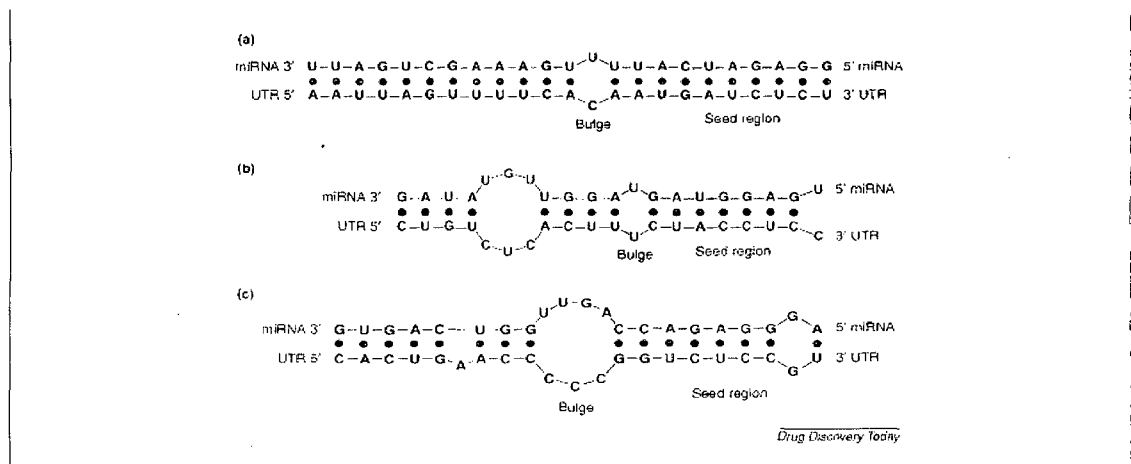


Fig 3 : Approximate base pairing pattern between binding site on the target and miRNA (a) Canonical sites (b) Dominant seed sites (c) Compensatory sites (Source - Maziere et al., 2007).

1.4 Transcription Factors

Transcription Factors (TFs) are specific proteins or protein complexes that generally bind to specific DNA sequences in upstream of genes. Primarily, TFs could be either general but diverse components of large multi-proteins pre-initiation transcription complex or are sequence specific DNA binding proteins. General TFs required for promoter recognition by RNA polymerase machines and its binding and thus helps in maintaining the basal level of rate of transcription. Sequence specific TFs generally binds to the specific sites in the promoter regions of the selective genes, called *cis*-elements, and hence causing increase or decrease in the rate of transcription from the basal level (Lewin 2004).

1.5 Promoter

Eukaryotic promoters are highly diverse in the context of structure and position. A typical promoter is a regulatory region of DNA sequence that lie upstream (5' region) of the gene. It contains specific DNA sequences for recognition and binding of TFs. A typical promoter can have three distinct parts (Fig: 2) (Lewin 2004; Weaver & Hedrick 1997).

Core Promoter : It is just immediate upstream of TSS and minimal portion of the promoter required for guiding the polymerase to properly initiate transcription. It contains TATA-box or initiator sequence through which it interacts with basal transcription factors along with polymerase to make pre-initiation complex.

Proximal Promoter : It is upstream portion of the core promoter. It contains binding sites for some general TFs which regulate the binding of core promoters to their components. It contains *cis*-elements for sequence specific DNA binding TFs which selectively activate or suppress the genes expression during diverse environmental conditions. Generally co-regulated genes have such common specific TFs binding elements and their arrangements (Lewin 2004).

Distal Promoter : It contain binding elements for that TFs which greatly affects transcription. These *cis*-element either enhance the rate of transcription, called enhancer, or reduce the rate of transcription, called silencer, through interaction with basal transcription complex. It generally located up to thousand bp upstream from the TSS or even in the non-translated region of open reading frame. They have major role in the tissue-specific expressed genes (Lewin 2004).

1.6 *cis*-elements

cis-elements are specific DNA sequences which lies within the promoter of a gene and are the preferential binding sites of TFs. By binding with TFs, these *cis*-

elements influences the rate of transcription of respective genes. *cis*-elements may be either **enhancers**, that enhance the rate of transcription or **silencers**, that reduce the rate of transcription (Lewin 2004).

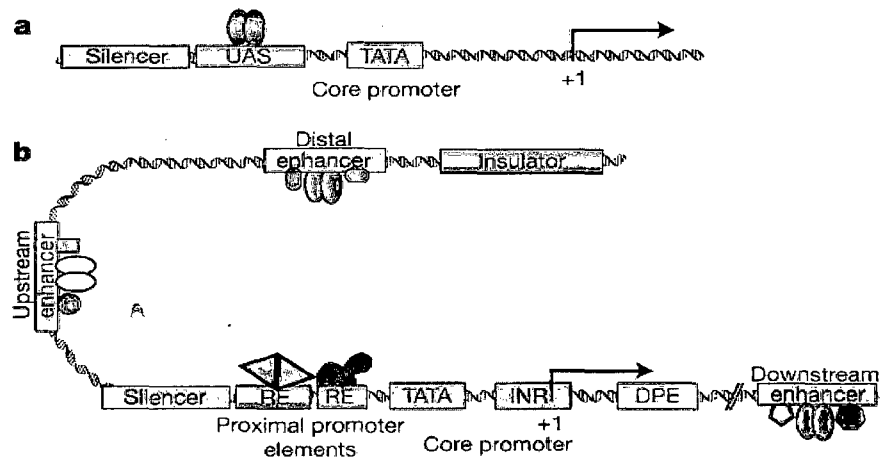


Fig:4 - Schematic representation of eukaryotic promoter (a) Simple eukaryotic transcriptional unit. (b) A complex arrangement of multiple clustered enhancer modules interspersed with silencer and insulator elements which can be located 10–50kb either upstream or downstream of a composite core promoter containing TATA box (TATA), Initiator sequences (INR), and downstream promoter elements (DPE) (Source - Levine and Tjian, 2003).

1.7 Motif

A motif is a pattern of nucleotides that has a biological significance. It is usually found in upstream of coding regions and rarely in downstream or even within untranslated parts of the transcribed gene. The pattern of nucleotides in a motif is generally widespread (Cristianini & Hahn 2007).

1.8 Consensus Sequences

A *cis*-element can be defined as a regular expression. Regular expression is a flexible representation of a type of pattern that allows for a wide range of substitution at one position, repeating of a position and other features. A regular

expression for a TFs binding element is also known as consensus sequence (Mount 2004).

1.9 RNA silencing

RNA silencing is a coordinated biochemical process of sub-cellular events which has been termed post-transcriptional gene silencing mechanism (predominantly used to describe in plant system), quelling (in fungi) and RNA interference (usually in vertebrates and invertebrates) (Susi et al. 2004).

miRNA gene transcribed in to a long sequence of more than several hundred nucleotides called primary miRNA (pri-miRNA) (Bartel D.P.,2004). The processing of miRNAs from the pri-miRNAs requires DCL1, a Dicer-Like protein (nuclear-localized RnaseIII enzyme) . Maturation of miRNAs occur in atleast two distinct steps. In 1st step pri-miRNA converts into pre-miRNA and in 2nd step pre-miRNA matures in to miRNA. During the processing from pre-miRNA to miRNA, DCL1 cuts preferentially at specific positions in the miRNA stem-loop precursor (Jones-Rhoades M. W. et. al., 2006).

The degree of complementarity between miRNA and target sequence play a major role in the function of miRNAs. The perfect sequence complementarity is observed in some plants (Rhoades et al. 2002) while the base-pairing interaction is imperfect in animals (Lee et al. 1993).

miRNAs are thought to control the gene expression either by destabilization of target mRNA, which facilitates cleavage (primarily in plant), by specific inhibition of its protein synthesis, possibly through effect on mRNA stability and compartmentalization (generally in mammals) (Hennessy and O'Driscoll 2008) or by directing the gene modifications (such as DNA methylation) that encode the targets protein. Some recent data suggest that miRNAs block translation after its initiation, perhaps either by inhibiting polypeptide elongation or even by degrading the nascent peptide (Matranga & Zamore 2007; Axtell and Bowman 2008).

Biochemical evidences suggests that target mRNA cleavage is the predominant

mechanism of miRNA-guided gene expression regulation in plants (Palatnik et al. 2003). All silencing complexes (RISCs) comprises at least three components - an Argonaute (AGO) protein, either a miRNA or siRNA and some other proteins which are complex specific (Bartel 2004). AGO contain a PAZ domain and a PIWI domain. The 3' end of miRNA is held in the groove of hydrophilic cleft of the PAZ domain and a groove in the PIWI domain hold the mRNA. Here both miRNA and mRNA aligns with each other in the fashion of Watson-Crick base pairing. The phosphodiester bond between the nucleotides of target mRNA that pair to nucleotides 10th and 11th position of the miRNA and falls near the active DDA site, seems to be cleaved by the PIWI domain of AGO1 which is guided by the miRNA (Bartel 2004; Qi Y. et al. 2005; Jones-Rhoades M. W., 2006). Exact base pairing in the center and at the 5' side of the match(i.e from 3rd-10th nucleotide) is required for the target cleavage (Mallory et al. 2004). Here, the orientation of miRNA in the RISC and the spatial dynamics of pairing may also impact on the degree/efficiency of silencing(Bartel 2004). Following mRNA cleavage, the cleavage products start to degrade from the cleaved end. In most cases, the 3'end of cleaved mRNA is degraded from 5' – to – 3' end by *Arabidopsis thaliana* EXORIBONUCLEASE4 (AtXRN4). But it is thought that other fragment of cleavage may degrade through a different pathway (Souret et al. 2004).

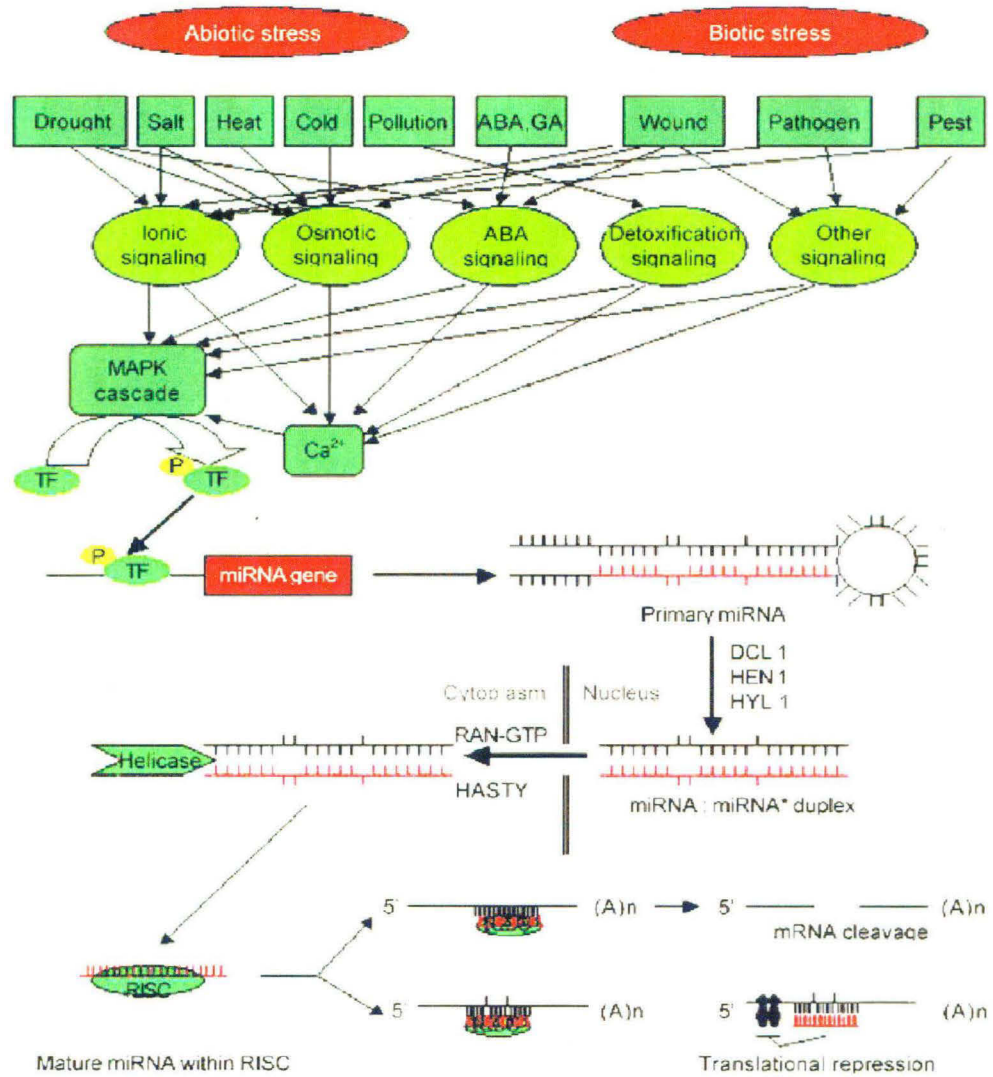


Fig 5 : A model for miRNA biogenesis, regulation of miRNA gene regulation, and the pathway of miRNA function (Source - Zhang et al. 2005).

1.10 Objective of the study

Experimental results shows that upon salt stress , *cdc2aAt*, a class-A type cyclin dependent kinase (CDK) gene, expression decreases particularly in the root vascular cylinder in *Arabidopsis* (Bursens et al. 200).

An attempt has been made to explore the possibility and feasibility of finding the role of miRNAs in the regulating the cell cycle gene *cdc2aAt* during stress conditions by *in silico* approach. Naturally, the first step is to find out the candidate miRNAs which may regulate cell cycle gene *cdc2aAt*. The second step is to do analysis on the upstream sequences of the shortlisted miRNAs genes to find out specific TFs binding sites. These analysis may provides clues about the expression of miRNAs in response to different stress conditions. Efforts have been also made to perform analysis for other down regulated core cell cycle genes.

In this dissertation, chapter 2 describes different data sources in brief and outlines different existing miRNA target prediction computational tools and programs that have been used in this study. We also describe methodologies of our study. Chapter 3 provides results and discussion. Chapter 4 gives the conclusion of our study.

Materials
and
Methods

CHAPTER 2 : Materials and Methods

2.1 Data Sources

2.1.1 MicroRNAs Data

MiRBase (Griffiths-Jones et al. 2005) (Release 12.0) (<http://microrna.sanger.ac.uk/>) is our primary source for miRNAs data. Description of about 204 mature miRNAs of *Arabidopsis thaliana* is provided and updated. The genomic coordinates of miRNAs genes were taken from TAIR8 and NCBI.

2.1.2 Core Cell Cycle Gene Data

The core cell cycle genes of *Arabidopsis thaliana* were retrieved from the “Cell cycle” section within the Gene families at TAIR (*Arabidopsis* Core Cell Cycle Gene Families). In this section, total 61 core cell cycle genes are mentioned (submitted by – Vandepoele 2002).

2.1.3 Whole Genome Sequence Data

The complete sequence of all the five chromosomes of *Arabidopsis thaliana* was taken from FTP Directory of NCBI (ftp://ftp.ncbi.nih.gov/genomes/Arabidopsis_thaliana/). The upstream sequences for analysis of regulatory cis-elements were retrieved from these sequences.

2.1.4 Differential Expression Data

Expression data for the *Arabidopsis thaliana* were obtained from The Nottingham Arabidopsis Stock Centre (NASC) microarrays database (<http://www.arabidopsis.info/>). There are about 80 different samples of expression data out of which 9 are of different abiotic stress experiments as follows : Cold, Osmotic, Salt, Drought, Genotoxic, Oxidative, UV-B, Wounding, Heat. Separate slides of expression data is available for root and shoot over a series of six different time periods (½ hr, 1hr, 3hr, 6hr, 12hr, 24hr) with two replicate for

each. Each slide has 22,814 genes.

Stress-response miRNAs expression data was reproduced from the work of Liu H., et al. 2008 (Table- 5).

2.1.5 *Cis*-Element Data

We collect the experimentally verified TFs binding elements from **bindinglist** of **AtcisDB(<http://arabidopsis.med.ohio-state.edu/AtcisDB/bindingSiteContent.jsp>)** and **AGRIS** database. This data is slightly old hence it contain information of a very limited TFs binding elements. There is a brief description of about 100 TFs binding consensus sequences. Out of which, 38 are stress responsive elements which mediates in response against various stress conditions to respective genes either directly or indirectly (A complete list of stress-responsive binding elements is given in Appendix 1.

2.2 Computational Tools and Methodologies

We have used four different miRNAs target prediction tools for analysis of our problem. They are (1) **psRNATarget**, (2) **miRanda**, (3) **RNAhybrid**, (4) **MicroInspector**.

The available miRNAs target prediction computational tools are based on some general characteristics feature of miRNA-target pairing – (i) nearly perfect (in plants) or imperfect (in animals) base pairing between miRNA and its target mRNA sequences i.e. degree of complementary, (ii) Position of binding site on the target i.e. base pairing pattern such as, in plants, miRNA preferentially binds in coding regions of target genes, (iii) Thermodynamic stability of miRNA-mRNA duplex i.e. minimum free energy of the duplex must be minimum. However, some tools also use some specific features such as cross-species conservation of miRNA-target along lineage.

The purpose of using four different miRNAs target prediction tools was to get the most suitable potential candidate miRNAs for core cell cycle gene *cdc2aAt* for all tools and perform a comparative analysis among them. As all tools using some distinct bio-physical features of miRNA-target duplex for prediction hence we will get different result for different tools.

psRNATarget and **MicroInspector** are the web-based integrated computing system while **miRanda** is stand-alone version. **RNAhybrid** has both stand-alone and online version. While **RNAhybrid** and **MicroInspector** can be used for any organism, **psRNATarget** is mainly for plant models. **miRanda** is primarily developed for animal models.

Based on available principles of miRNA-target interactions in plant systems, we did some modifications and set the parameters for **MicroInspector**, **miRanda** and **RNAhybrid** to filter out false hits.

2.2.1 psRNATarget

This server is a developed version of the *miRU: an automated plant miRNA*

target prediction server.

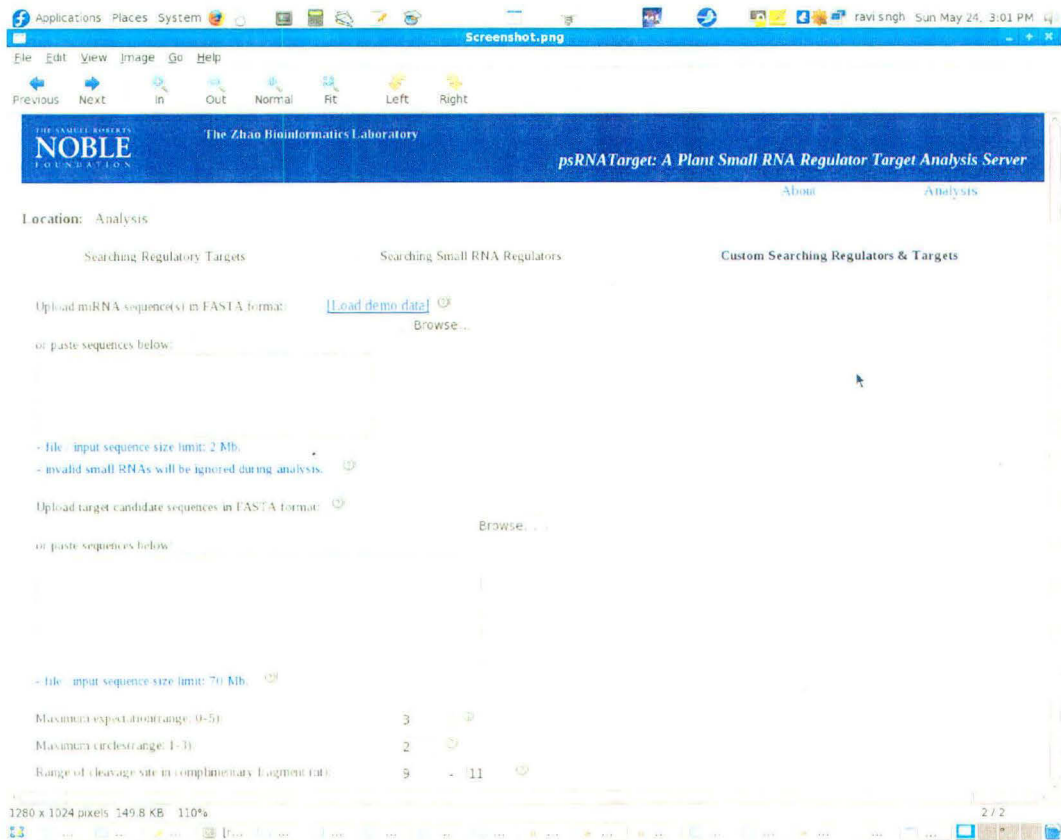


Fig 6 : Web interface of **psRNATarget** .

The **psRNATarget** accepts miRNA mature sequences and target sequence as inputs and reports all potential complementary regions between miRNA and target sequences using an improved iterative parallel Smith-Waterman algorithm and a weighted scoring schema. Each mismatch penalized according to the mismatch type and position in the miRNA.

In addition **psRNATarget** is able to search whether the user-submitted sequences are potential targets of published miRNA sequences and search user-submitted small RNA against user-submitted target candidates. The **psRNATarget** server automatically synchronizes with up-to-date miRNAs dataset of **miRBase**.

There is also up-to-date databases of Plant UniGene libraries and gene model transcript libraries.

We set the *maximum expectation* value – 5 and *maximum circles* – 3. *Maximum expectation* is the threshold of the score. The program gives output only those hits which will have score below 5. The *maximum circles* is the limit of maximum times of calling for searching the complementary site in the target sequence. Increasing *maximum circles* will increase the running time and hence give most possible complementary regions under *maximum expectation*.

2.2.2 MicroInspector

The user-submitted target sequence is scanned and analysed simultaneously and independently with two window of 6 nt of miRNA. The 1st window represents nucleotides 1-6 and 2nd represent 2-7 from the 5' of the miRNA. As window slid through the target sequence (by steps of 1 nt), program perform complementarity analysis. When the program identifies a domain having 5 Watson-Crick base pairs or 4 Watson-Crick base pair with at least one additional G:U pair, it extracts a 32-nt sequence of the target terminating at the nucleotide that is complementary to 5' nucleotide of the 1st window. Subsequently, the miRNA and target sequence duplex are subjected to a pair-wise hybridisation folding algorithm of the Vienna RNA secondary structure programming library (RNAlib) from the Vienna RNA 1.5 version package for calculation of thermodynamic properties of the duplex.

Post-filter analysis eliminates any hit characterised by

- two unpaired nucleotides on either end of miRNA,
- duplex with low folding energy due to formation of intramolecular stable structure,
- if duplex have too large interior or bulge loops or if large loops are located near the end of secondary structure.

We set the hybridization temperature threshold value (16°C) and free energy

threshold value (-30 kcal/mol) for miRNA-target duplex. The program give the result with lower energy than the threshold value.

2.2.3 miRanda

This prediction algorithm is similar to Smith-Waterman algorithm, uses a position-weighted matrix to emphasize binding of the miRNA's 5'-end segment. Scores based on the complementarity of nucleotides between miRNA and target sequence (A=U or G≡C). Scoring matrix also allows G=U 'wobble' pairs and uses affine penalties for gap-opening and gap-extension. Complementarity scores (+ve and -ve values) at the first eleven positions are multiplied by a scaling factor so as to reflect the observed 5' – 3' asymmetry. Finally, the following empirical rules are applied , from 5'end of the miRNA:

- no mismatches at positions 2 to 4.
- fewer than five mismatches between positions 3 to 12.
- at least one mismatch between position 9 and L-5 (L=total alignment length).
- fewer than two mismatches in the last five positions of the alignment.

With these parameters, algorithm optimizes the score, summed over all aligned positions and find all non-overlapping hybridization alignment in decreasing order of score down to some cutoff value (default value 80) (John et al. 2004).

The thermodynamical properties of the duplex is estimated by using folding routines from the Vienna 1.3 RNA Secondary structure programming library (RNAlib). The miRNA sequence and 3'UTR target sequence from a hybridization alignment are joined into a single sequence with an eight base sequence linker containing artificial 'X' bases that can not basepair. This strand-linker-strand configuration assumes the phase space entropy of strand-strand association. The minimum energy of this structure with the last matching base pair 'constrained' is

then calculated using RNALib.

We set the energy threshold at -20 kcal/mol and scaling parameter at 2 (scaling parameter was set at 3 when there was no significant hit on 2). Program multiply the complementarity score (+ve and -ve values), at the first eleven positions, by scaling factor, so as to more effectively detect alignments of this type. Many known examples of miRNA:Target duplexes are highly complementary in this region.

The affine penalties for gap-opening and gap-extension was (-8) and (-2) respectively. The score threshold was 95.

2.2.4 RNAhybrid

The algorithm core of this tool is an extension of the classical RNA secondary structure prediction to two sequences. Using the Dynamic Programming Technique, the program calculates the most favourable MFE between miRNA sequence and target mRNA sequence at all possible start positions. Intramolecular hybridization are not allowed. Bulge loops and internal loops are restricted to a constant maximum length in either sequence. The program predicts optimal, non-overlapping hits using user-defined parameters such as energy or *p*-value threshold. The user may fix hybridization to contain perfect helices in the 5'-part of the miRNA.

The energy threshold was set on (-20 kcal/mol). Internal loops were allowed only with a maximum of two nucleotides in each strand. Bulges with no more than one nucleotide were permitted. Helix constraint was taken from 2-to-7 and 8-to-12.

2.2.5 Upstream Analysis

To analyze the regulatory pathways for regulation of target genes by miRNAs we study the upstream sequences of miRNAs genes by using two different computational approach. First, we used the **MatchTM** program of **TRANSFAC PROFESSIONAL 12.1**. Secondly, we used a **pattern matching program** for scanning the upstream of miRNAs genes against the known TFs binding elements in plants.

2.2.5.1 MatchTM program

MatchTM uses a library of mononucleotide weight matrices from **TRANSFAC[®] Professional** to search potential transcription factor binding sites in sequences provided by the user. We use only "high quality matrices" from plant group which ensures good quality matrices. For core similarity of matrices we put threshold value of 0.950 and for entire matrix similarity keep it at 0.900. This means that the most informative positions must be equal to or more than 95% similar while the TFBS in total should not score below 90% similarity so that we get a low number of false positive matches.

However after submission of queries we get a huge list of TFS binding elements and also multiple sites for a particular elements in a particular upstream sequence. So for getting more reliable result we filter only those binding elements which has core match value equal to 1.000 and matrix match value equal to or above than 0.900.

All possible result of **MatchTM** for the candidate miRNAs of *cdc2aAt* are given in Appendix 2.

2.2.5.2 Pattern Matching Program

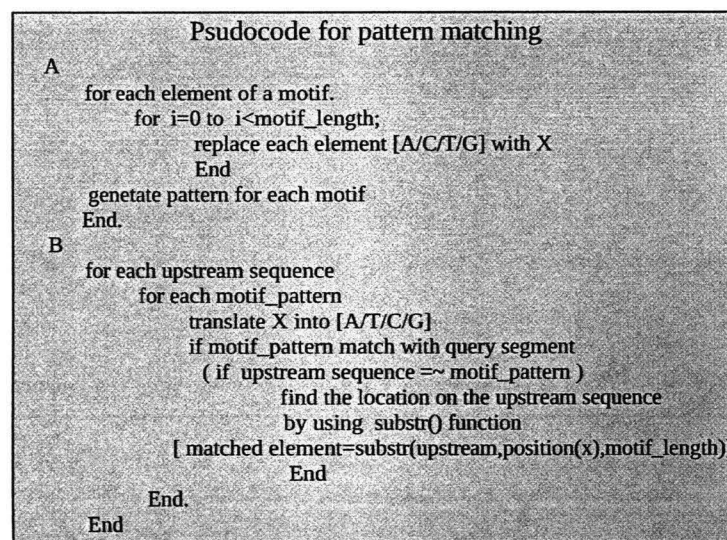
There are very limited experimental validated data about the stress responsive *cis*-elements in plants. Hence it is very difficult to predict the putative

cis-binding elements in a given DNA sequence by using statistical and / or computational methods. The well known motif finding methods such as PSSM algorithm or HMM algorithm also require a minimum sufficient amount of information regarding consensus sequences of a element to generate training set. Hence we developed a **pattern matching program** to identify putative TFs binding elements in the upstream of miRNAs genes.

To use **pattern matching program** for our analysis, we collected the experimental verified stress responsive TFs binding elements from **binding list of AtcisDB** and **AGRIS** database. Then different patterns for each binding elements is generated by mutating each site in that element. After that, 1000nt upstream of each miRNA gene was scanned against each generated pattern.

The length of string for some TFs binding elements is very short such as for ARF it is just 6. Hence it could be possible that the **pattern matching program** scan many sites in the target sequence against these motifs. Also against the long string of GT-promoter BS and MRE program could not scan any sites. All possible match site for different stress response elements in the upstream of candidate miRNAs of *cdc2aAt* is given in Appendix 3.

The complete procedure for upstream analysis is given in the Fig-7 as flow-chart diagram.



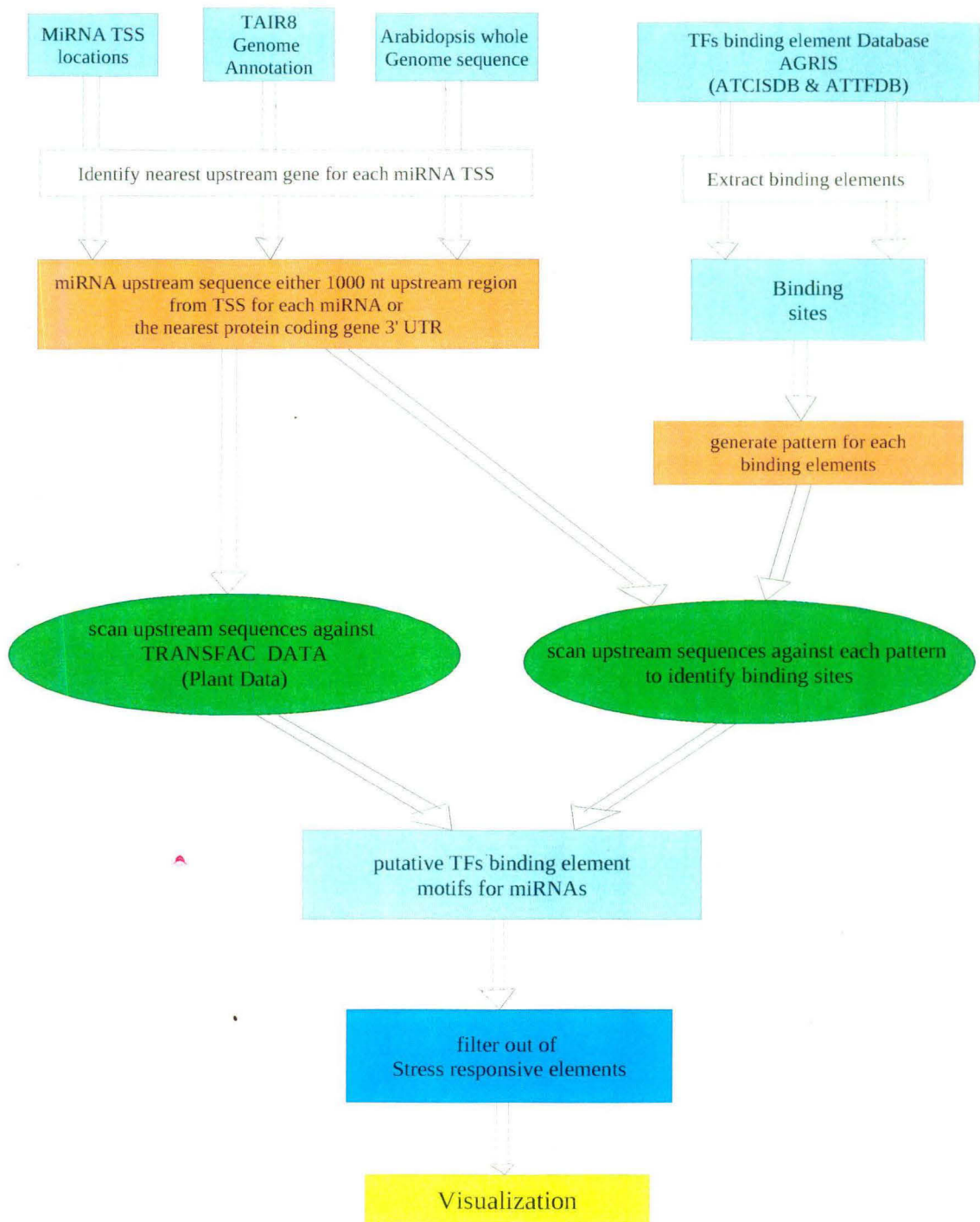


Fig 7 : A flow-chart diagram of the steps for identifying stress responsive cis-elements in the upstream region of miRNAs genes.

TH-16220

570285
S:646
In



Results
and
Discussion

CHAPTER 4 : Results and Discussion

Comparative analysis of results [Table-1] shows that miR447 family is predicted as candidate miRNAs for the cell cycle gene *cdc2aAt* by all four predicting tools. While miR161.1 is predicted by psRNATarget, RNAhybrid and MicroInspector, miR163 is predicted by psRNATarget and MicroInspector. psRNATarget and RNAhybrid both predicted miR396a & b.

psRNATarget	RNAhybrid	miRanda	MicroInspector
miR447a,b,c, miR1888, miR414, miR161.1, miR163, miR172a,b, miR852, miR396a,b, mir835-5p, miR835-3p	miR447a,b, miR396a,b, miR161.1, miR172a,b, miR1888, miR835-3p, miR395b,c,f, miR414	miR447a,b,c, miR861-5p, miR168a,b, miR833-3p, miR843, miR848, miR847, miR854a,b,c	miR447a,b,c, miR164a,b,c, miR163, miR161.1, miR834, miR395a,b,c,d,e,f, miR167d, miR851-5p, miR833-3p, miR840, miR398b,c, miR390a,b, miR402, miR854a,b,c,d

(The miRanda results shows for both scaling parameter 2 & 3)

Table 1 : List of putative microRNAs for *cdc2aAt* predicted by different predicting tools.

Based on minimum free energy value of miRNA and its target duplex [Table - 2], miR161.1 could be a potential candidate for *cdc2aAt* followed by miR447 family.

miRNAs	Minimum free energy
miR161.1	-28.4 kcal/mol
miR447a	-27.7 kcal/mol
miR447b	-27.7 kcal/mol
miR447c	-26.3 kcal/mol
miR1888	-20.2 kcal/mol
miR414	-22.3 kcal/mol
miR835-3p	-24.3 kcal/mol
miR396a	-23.6 kcal/mol
miR396b	-21.8 kcal/mol
miR172a	-22.6 kcal/mol
miR172b	-22.6 kcal/mol
miR163	-24.9 kcal/mol
miR852	-20.1 kcal/mol

Table 2 : The minimum free energy value of some selected candidate miRNAs and its target (*cdc2aAt*) duplex .

Minimum free energy is not the only criteria for selecting the candidate miRNA for a particular target. The symmetry of secondary structure of miRNAs and its complementary target sequence duplex also contribute significantly. After careful inspection of secondary structure of duplex of all the possible candidate miRNAs and its target, it may be seen that the duplex of miR161.1 is slightly more symmetrical than that of others (Fig-8). Hence this also support the inference that miR161.1 is a strong candidate for *cdc2aAt*. The consensus prediction of all the four tools points to miR447 family. Hence miR161.1 and miR447 family could be a putative candidate which contribute in the regulation of *cdc2aAt*.

In animal, miRNAs target sites are usually found in 3' UTRs whereas in plants they are predominantly binds in coding regions and rarely in UTRs (Rajewsky N. 2006; Wang X. J. et. al., 2004). Our results also support this hypothesis (Fig- 9).

It might be possible that there would be more discussion on the candidature of miR447 family and miR161.1 for *cdc2aAt*. Because all the available *in silico* approach for miRNA target prediction, including above four, based on the available information of structural and functional aspects of miRNAs and their targets. Till date, a very limited data resources and evidences are available about the structure, function and role of microRNA, so the research pertains to microRNA is currently under active area of investigation, specially in plant systems.

Some experimental evidences suggests that the efficiency of a miRNA also depends on many other biochemical and biophysical properties of RISC such as degree of site accessibility, spatial dynamics of pairing etc (Bartel 2004; Kertesz 2007). Hence to develop an efficient and high throughput method there is a need of some more information about different aspects of miRNAs.

But after analyzing the above result we may say that miR447 family and / or miR161.1 could be the putative candidates for *cdc2aAt*.

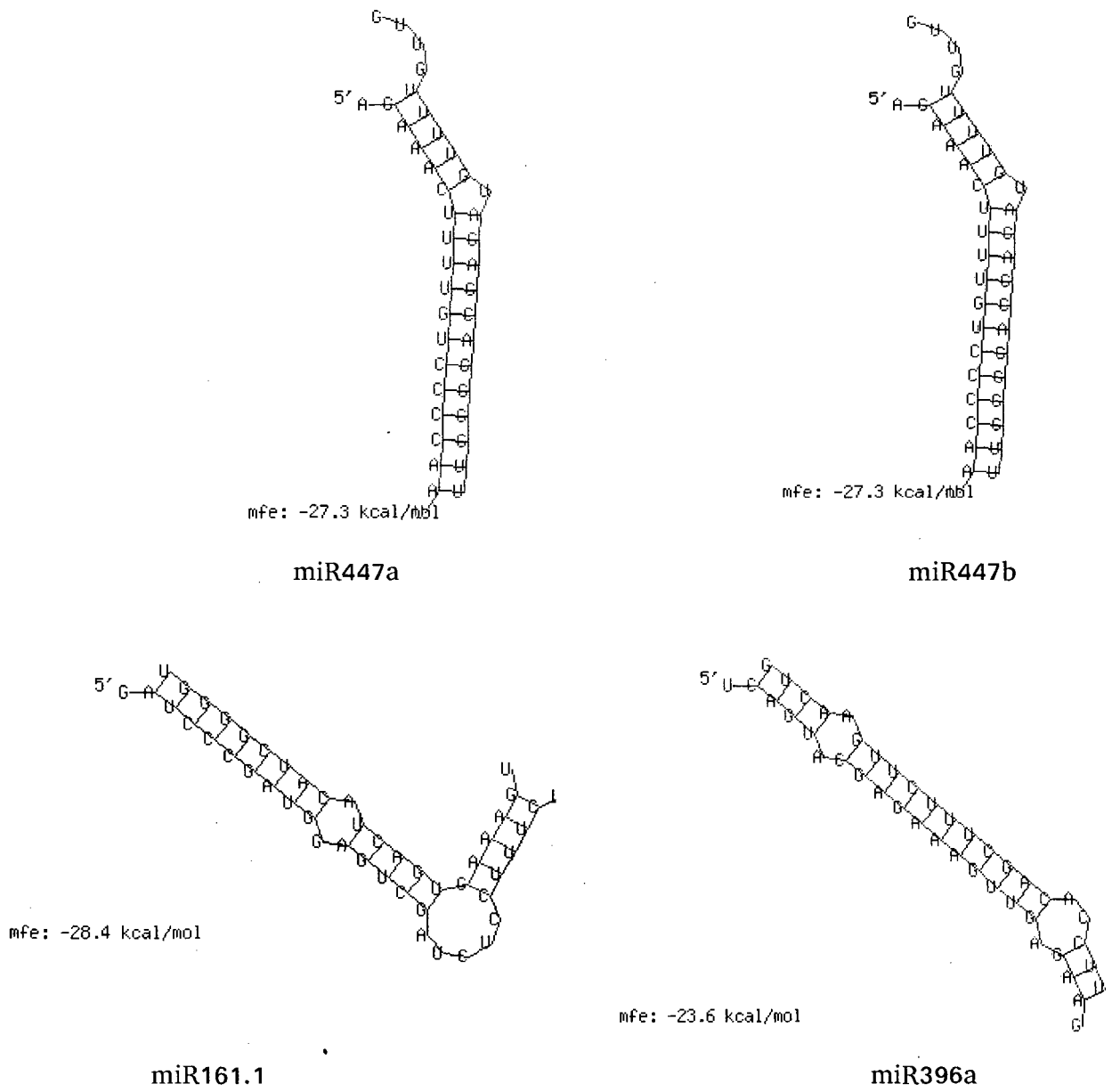


Fig 8: Secondary structure of candidate miRNAs and its complementary target sequence duplex.

Experimental results also shows that miRNAs which targets *cdc2aAt* also regulate other 6 to 7 core cell cycle genes lends support to our prediction [Table-3].

	447a	447b	447c	1888	414	835-5p	835-3p	396a	396b	172a	172b	161	163	852
AT3g48750	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
AT1g47220	✓	✓			✓									
AT1g47230			✓	✓										
AT1g47870				✓				✓	✓					
AT1g80370				✓				✓	✓	✓	✓			
AT2g17620	✓	✓						✓	✓					
AT3g01330	✓	✓	✓											
AT4g35620	✓	✓	✓			✓								
AT4g37630				✓	✓									✓
AT5g11300	✓	✓				✓								
AT5g22220			✓							✓	✓	✓		
AT5g48820	✓	✓												

Table 3 : List of possible core cell cycle genes co-suppressed by miRNAs

Signal values from the microarray data derived during salt stress conditions, reveals that 20 core cell cycle genes are down-regulated. The next logical step is to see whether these 20 genes are controlled by microRNAs or not. By using all four miRNAs target prediction computational tools with validated parameters we found that there are some putative miRNAs which target those 20 core cell cycle genes [Table-4].

If these candidate miRNAs regulate the expression of *cdc2aAt* and other core cell cycle genes during stress conditions, there must be an optimum concentration of miRNAs that could target genes more efficiently. The micro-array experiment data of microRNAs [Table-5] shows that few microRNAs are up-regulated during different abiotic stress conditions such as miR396, miR168, miR171 etc. The expression of miR396 gene increases up to three folds in different stress conditions. While expression of miR168 increases up to two folds, miR171 increases up to one and half fold(Liu et al. 2008). The work carried by Liu et al. is not exhaustive because it gives expression idea of a very few microRNAs.

Table 4 : List of putative microRNAs for other down-regulated core cell cycle genes. (Value in parentheses show the minimum free energy of miRNA-target duplex)

Core cell cycle gene (Genomic locus) (Protein name)		Candidate miRNAs (m.f.e.)	
AT1G02970	WEE1	miR823(-29.3),	miR837-5p(-24.7)
AT1G16330	CYCB3;1	miR779.1(-27.8),	miR169a(-23.5)
AT1G20930	CDKB2;2	miR843(-28.2),	miR414(-23.3)
AT1G44110	CYCA1;1	miR854(-27.4),	miR842(-26.1)
AT1G76310	CYCB2;4	miR169(-29.9),	miR869.1(-27.1)
AT1G80370	CYCA2;4	miR396a(-26.3),	miR156a(-25.8)
AT2G17620	CYCB2;1	miR834(-29.8),	miR837-5p(-29.6)
AT2G26760	CYCB1;4	miR831(-27.6),	miR838(-25.6)
AT2G27960	CKS1	miR854(-31.2),	miR405a(-28.8)
AT2G27970	CKS2	miR405(-28.9),	miR447a(-23.8)
AT3G01330	DEL3	miR162(-33.3),	miR447(-25.9)
AT3G11520	CYCB1;3	miR854(-31.8),	miR402(-27.8)
AT3G50630	KRP2	miR414(-30.2),	miR408(-24.5)
AT4G35620	CYCB2;2	miR834(-31.6),	miR156(-28.7)
AT4G37490	CYCB1;1	miR847(-32.3),	miR414(-25.6)
AT5G11300	CYCA2;2	miR869.2(-26.9),	miR857(-22.2)
AT5G27620	CYCH;1	miR854(-35.1),	miR831(-26.8)
AT5G43080	CYCA3;1	miR831(-25.0),	miR413(-24.4)
AT5G67260	CYCD3;2	miR854(-30.5),	miR414(-26.9)

Treatment	Fold change
300 mM NaCl	
miR396	2.99
miR168	1.91
miR167	1.87
miR165	1.82
miR319	1.81
miR159	1.63
miR394	1.58
miR156	1.56
miR393	1.55
miR171	1.50
miR158	1.46
miR169	1.30
200 mM mannitol	
miR396	2.59
miR168	1.91
miR167	1.84
miR408	1.73
miR171	1.55
miR157	1.42
miR393	1.36
4° C	
miR396	3.95
miR397	2.08
miR172	1.88
miR169	1.79
miR408	1.73
miR168	1.71
miR171	1.60
miR393	1.56
miR319	1.54
miR165	1.50
miR400	1.36

Table 5 : Stress-response miRNAs identified by microarray analysis
(Source : Liu H., et al. 2008)

It has been found that stress responsive up-regulated genes have stress responsive elements in their upstream (Abe et al. 1997). If these miRNAs are over-expressed during stress conditions then it could be expected that upstream of miRNAs genes would have some common stress responsive *cis*-elements.

The upstream analysis of miRNAs genes by using **MatchTM** program of **TRANSFAC PROFESSIONAL 12.1** shows that indeed there are some stress responsive elements other than some unique binding sites that might helps in up-regulation during different stress conditions [Table-6]. The major stress responsive elements are for - AtMYB, MYBAS1, GAMYB, SED, ASR1, ERF, GBP, GT-1 etc. While other common binding elements are for PBF, members of Dof, members of

Name of miRNAs	Stress responsive <i>cis</i> -elements								Non-stressed common <i>cis</i> -elements										
	AtMYB	MYBAS1	GAMYB	SED	ASR1	ERF	GBP	GT-1	PBF	CPRF	TGA1a	RITA-1	OCSBF	Knox3	BPC1	RAV1	C1	Opaque2	Dof
miR447a	√	√	√	√		√	√	√	√	√	√	√	√	√	√	√	√	√	√
miR447b			√		√		√	√	√	√	√	√	√	√	√	√	√	√	√
miR447c		√					√	√	√	√	√	√		√	√	√	√	√	√
MiR161.1		√	√		√				√						√		√		√
miR396a		√	√	√	√			√	√					√	√	√	√		√
miR396b	√	√		√			√		√	√	√	√		√		√	√	√	√
miR414		√	√	√					√					√	√	√	√		√
miR163	√	√	√	√	√			√	√	√	√	√	√	√	√	√	√	√	√
miR835-5p			√				√		√	√	√	√	√	√	√		√	√	√
miR852		√		√				√	√					√	√	√			√
miR172a		√	√						√						√	√	√		√
miR172b		√		√	√		√		√	√	√	√	√	√	√	√	√	√	√
miR395b		√		√	√		√	√	√	√	√	√	√	√	√	√	√	√	√
miR395f		√		√			√	√							√	√	√		√
miR395d		√	√	√	√		√	√	√	√	√	√			√	√	√	√	√
miR395e		√	√	√	√		√	√	√	√	√	√			√	√	√	√	√

Table 6 : List of putative stress response and some other common *cis*-elements in the upstream of the miRNAs genes (from **MatchTM** result with **matrix similarity cut-off=0.90** and **core similarity cut-off=0.95**).

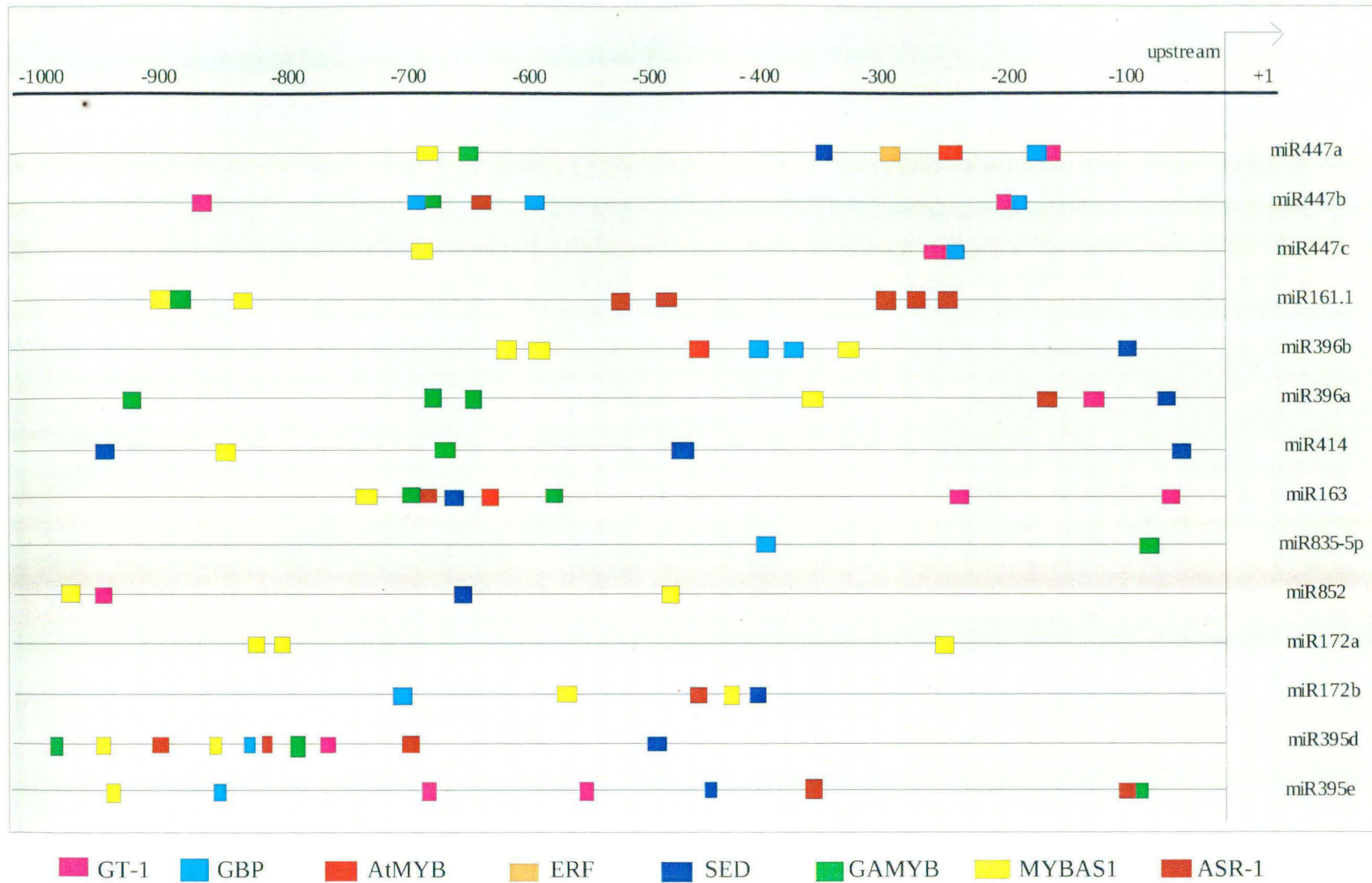


Fig 10 : Schematic representation of the details of the positions of stress response *cis*-elements in the upstream of miRNAs genes (from **MatchTM** result).

PBF, CPRF, TGA1a, RITA-1, OCSBF, Knox3, BPC1, RAV1, C1, Opaque2.

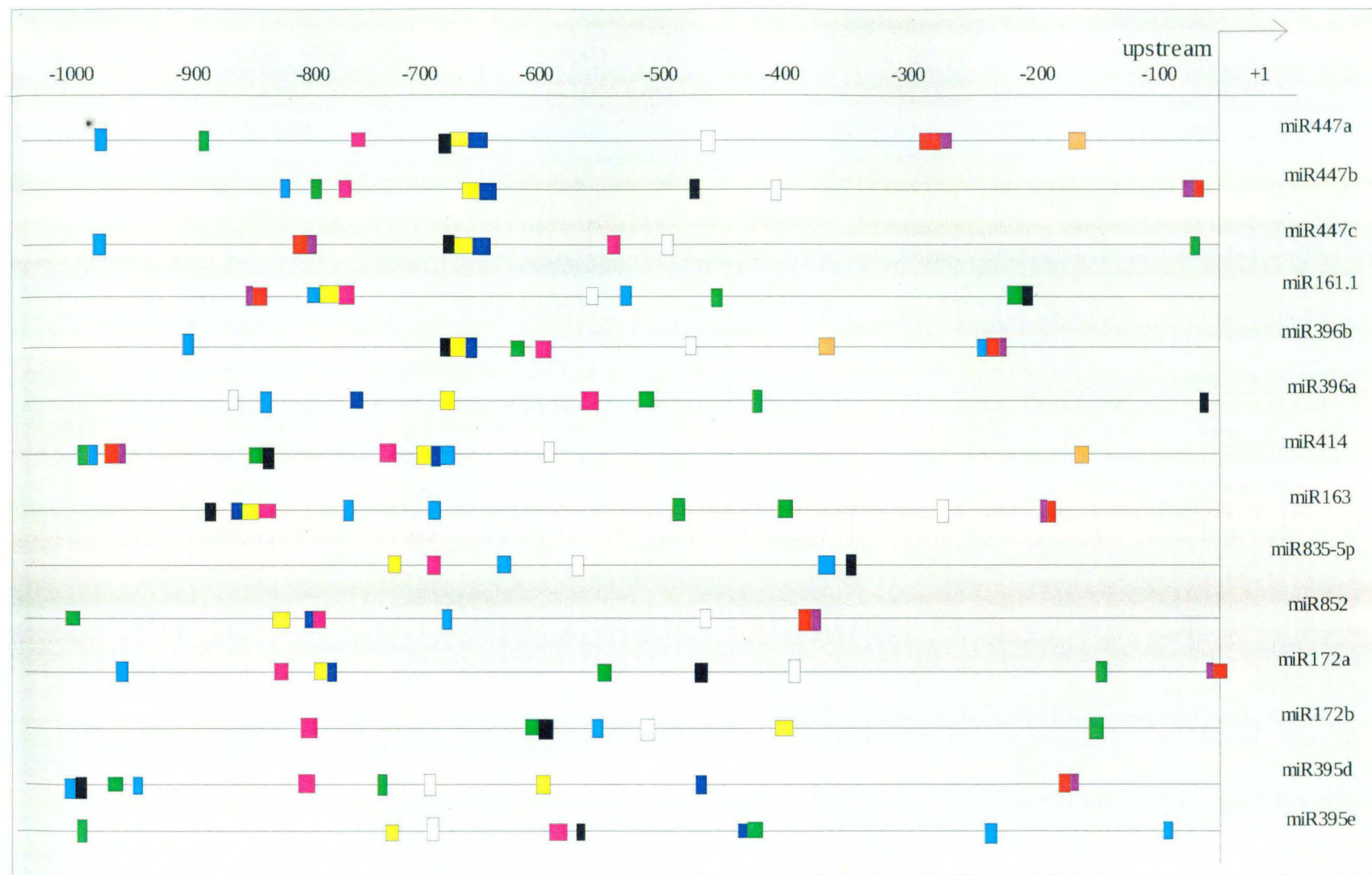
Fig-10 shows the binding pattern of different stress responsive elements in the upstream of different candidate miRNAs genes for *cdc2aAt*.

AtMYB element has role in drought- and abscisic acid- regulated gene expression (Abe et al. 1997). This element found in miR447a, miR163 and miR396b. MYB possesses an R2R3 DNA binding domain and is known to regulate the expression of salt- and dehydration- responsive genes. MYB element found in miR447a,b, miR414, miR852, miR161.1, miR163, miR172a,b, miR396a,b, miR395b,c,f,a,d and miR395e. GAMYB in the cooperation with OsDof3 influences the promoter activity of the Gibberellic acid(GA) – responsive genes (Washio K. 2003). GAMYB binding elements found in miR447a,b, miR161.1, miR396a, miR414, miR163, miR835-5p, miR172a, miR395d and miR395e. SED is a heat-stress-associated protein which is similar to dehydration-responsive family protein. SED element found in miR447a, miR414, miR852, miR163, miR172b, miR396a,b, miR395b,c,f,a,d and miR395e. ASR1(Abscisic acid stress ripening) is a water- and salt- stress regulated gene which encodes a zinc-dependent DNA binding protein. It becomes over-expressed under water- and salt- stress conditions. ASR1 element identified in miR447b, miR161.1, miR163, miR172b, miR396a, miR395b,a,d and miR395e. ERF is a ethylene response factor. This element found only in miR447a. GBP is basic penta cysteine 1. GBP element found in miR447a,b,c, miR835-5p, miR172b, miR396b , miR395b,d and miR395e. GT-1 is light response element. This motif found in miR447 family, miR396a, miR163, miR852 and miR395b,f,d and miR395e.

Analysis by **pattern matching program** shows that the stress responsive elements such as AtMYC2, AtMYB2, ABRE-like, ARF, DPBF, GCC-box, members of MYB family and SORLIP family etc. are prominently found in upstreams of candidate miRNAs genes [Table-7].

Element name	Name of microRNAs													
	miR447a	miR447b	miR447c	miR396b	miR396a	miR163	MiR161.1	miR414	miR835-5p	miR852	miR172a	miR172b	miR395d	miR395e
DREB1&2									✓				✓	
AtMYB2				✓	✓	✓	✓	✓		✓	✓	✓	✓	✓
AtMYC2	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
CBF1														
ABFs									✓					
ABRE					✓			✓						
ABRE like	✓			✓				✓						
CBF2	✓													
HSEs														
ARF	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
ACE														
DPBF	✓	✓	✓	✓	✓	✓		✓		✓	✓		✓	✓
DRE-like														
EIL														
ERE														
GCC-box	✓	✓	✓	✓		✓	✓	✓		✓	✓		✓	
GTpromoter														
LTRE				✓		✓	✓	✓	✓	✓		✓	✓	✓
MRE														
OBP1(Dof)														
ERF1	✓	✓	✓	✓		✓	✓	✓		✓	✓		✓	
GBF	✓													
MYB1					✓		✓			✓				
MYB2														
MYB3	✓		✓	✓		✓	✓	✓				✓	✓	
MYB4	✓	✓	✓	✓	✓	✓	✓	✓	✓		✓	✓	✓	✓
I Box	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
PRHA														
SORLREP1												✓		
SORLREP2							✓							✓
SORLREP3	✓		✓		✓	✓	✓	✓			✓	✓	✓	✓
SORLREP4											✓		✓	
SORLREP5														
SORLIP1	✓	✓	✓	✓	✓	✓	✓	✓	✓		✓		✓	✓
SORLIP2	✓		✓		✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
SORLIP3									✓					✓
SORLIP4		✓		✓										
SORLIP5			✓						✓	✓	✓	✓	✓	

Table 7: List of putative stress response cis-elements in the upstream of the miRNAs genes (Result from **Pattern Matching Algorithm**)



AtMYC2 AtMYB2 ABRElike ARF DPBF GCCbox LTRE ERF1 MYB lbox SORLREP3 SORLIP1

Fig 11 : Schematic representation of the details of the positions of stress response *cis*-elements in the upstream of miRNAs genes (from **Pattern Matching** result).

AtMYC2 binding sites were originally found in a drought-responsive gene, which increases responsiveness through higher sensitivity to the plant hormone ABA. AtMYC2 element found in almost all candidate miRNAs of *cdc2aAt* including miR447 family and miR161.1. ABRE like elements usually found in dehydration and low temp. responses genes. It found in miR447a, miR396b and miR414. ARF (Auxin response factors) found in upstream of almost all candidate miRNAs. DPBF-1 and 2 (Dc3 Promoter-Binding Factor-1 and 2) is the basic regions of two basic leucine zipper proteins which binds in the proximal promoter regions of ABA-induced genes (Kim SY et al., 1997). DPBF-1&2 element also found in all except miR161.1 and miR172b. GCC-box promoter motif found in dehydration and low temp. responses genes. This motif also found in most of the candidate miRNAs. MYB elements are highly ubiquitous and found in different stress responsive genes. These elements found in almost all candidate miRNAs genes. SORLREP and SORLIP are the light-response elements and also found in almost all miRNAs.

The binding patterns for these elements are shown in Fig-11.

From the results of **MatchTM** and **pattern matching program** it clear that stress responsive elements AtMYB2, AtMYC2, MYBAS1 and other members of MYB, SED, ASR-1, GBP, GT-1, ARF, DPBF etc. are the prominent stress responsive elements found in upstream of different candidate miRNAs genes [Table 6 & Table 7]. Other non-stressed TFs such as PBF, BPC1, RAV1, members of Dof, C1, E2F binding motif, CBF binding site etc. are also common between these.

Comparative analysis [Table-8] of upstream of miR447 family, miR161.1, miR396 family, miR168 family and miR171 family shows that they contain some common stress response elements such as AtMYC2, AtMYB2, MYBAS1, GAMYB, ASR1, ERFGT-1, ARF and DPBF.

miRNAs	Name of Element										
	AtMYC2	AtMYB2	MYBAS1	GAMYB	SED	ASR1	ERF	GBP	GT-1	ARF	DPBF
miR447a	✓	✓	✓	✓	✓		✓	✓	✓	✓	✓
miR447b	✓			✓		✓	✓	✓	✓	✓	✓
miR447c	✓		✓				✓	✓	✓	✓	✓
MiR161.1	✓	✓	✓	✓		✓	✓			✓	
miR396a	✓	✓	✓	✓	✓	✓			✓	✓	✓
miR396b	✓	✓	✓		✓		✓	✓		✓	✓
miR168a	✓		✓	✓		✓		✓	✓	✓	✓
miR168b		✓					✓			✓	
miR171a	✓	✓	✓		✓		✓		✓		✓
miR171b	✓	✓	✓	✓	✓		✓		✓	✓	
miR171c			✓	✓	✓	✓				✓	✓

Table 8: Putative stress-responsive binding elements in the upstream of miR447 family, miR161.1, miR396 family, miR168 family, miR171 family.

Hence from the above analysis we may say that the expression pattern of miR447 family and miR161.1 will be more or less similar to miR396 or miR168 or miR171 under different stress conditions.

Expression profile from microarray data shows that some other common TFs along with stress responsive TFs are also up-regulated during different stress conditions (Fig-12) which further support our hypothesis that miR447 family, miR161.1 genes and others are up-regulate during stress conditions.

Regulation of a gene depends upon combinatorial action of different *cis*-elements and the interaction of the factors binding to them. The above stress responsive binding elements may regulate miRNAs genes either directly or indirectly via different regulatory pathways during different stress conditions. Elements for AtMYC and AtMYB2 in the promoter of *rd22* gene directly binds with their factors, such as rd22BP1(MYC) and AtMYB2(MYB) respectively, under dehydration-stressed conditions mediated by abscisic acid(ABA) (Abe H. et al., 1997). While it has been shown that under different mineral nutrient excessive condition the level of

gibberellic acid (GAs) increases up to 27% (Battal P., 2004) which may leads to expression of GAs responsive genes regulated by GAMYB factors.

The binding patterns of different stress response TFs (Fig 10 & Fig 11) shows that AtMYC, MYBAS, GAMYB and ARF is highly ubiquitous and predominantly lie in the distal promoter regions. Pattern matching program shows AtMYC is flanked with DPBF and MYB. There is also highly repeating of AtMYC, ARF, MYB4, SORLIP2 and others motifs in the upstream of all miRNAs genes.

MatchTM result show that in the upstream of miR447 family GBP lies in the upstream of core promoter and flanked by GT-1. This specificity may give a unique regulatory pathway for miR447 family. Our proposed regulatory pathway for *cdc2aAt* has been depicted in fig 13.

Hence we may say that the different stress responsive binding elements in the upstream of miRNAs genes may have direct or indirect role in their regulation.

Our computational analysis indicates that linkage between miR447 family and the cell cycle gene *cdc2aAt*. This prediction has to be validated by experimental means.

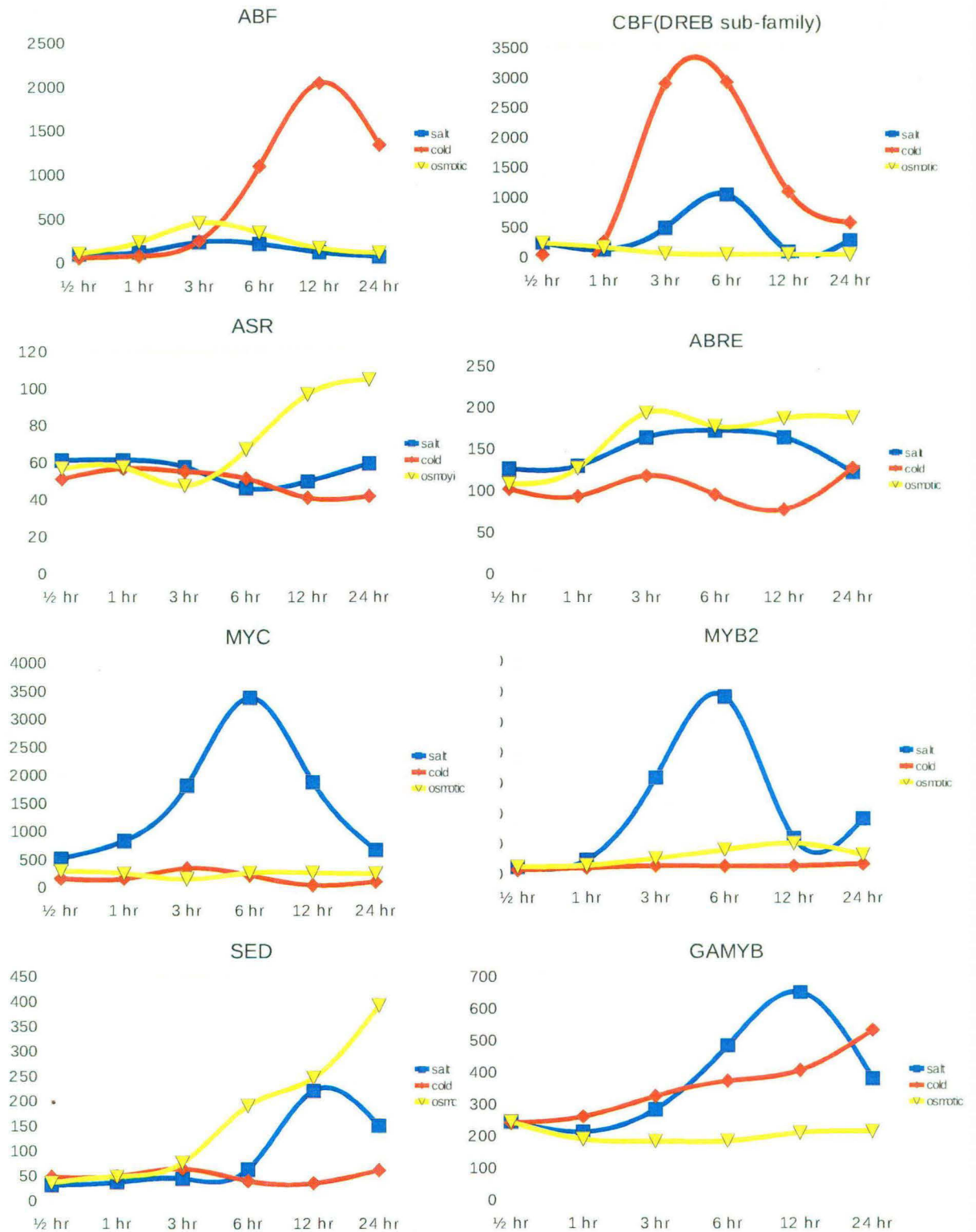


Fig 12 : Expression profile of some common stress responsive TFs during different stress condition (Source: NASC array). [Blue – salt, Red – cold , Yellow – Osmotic]

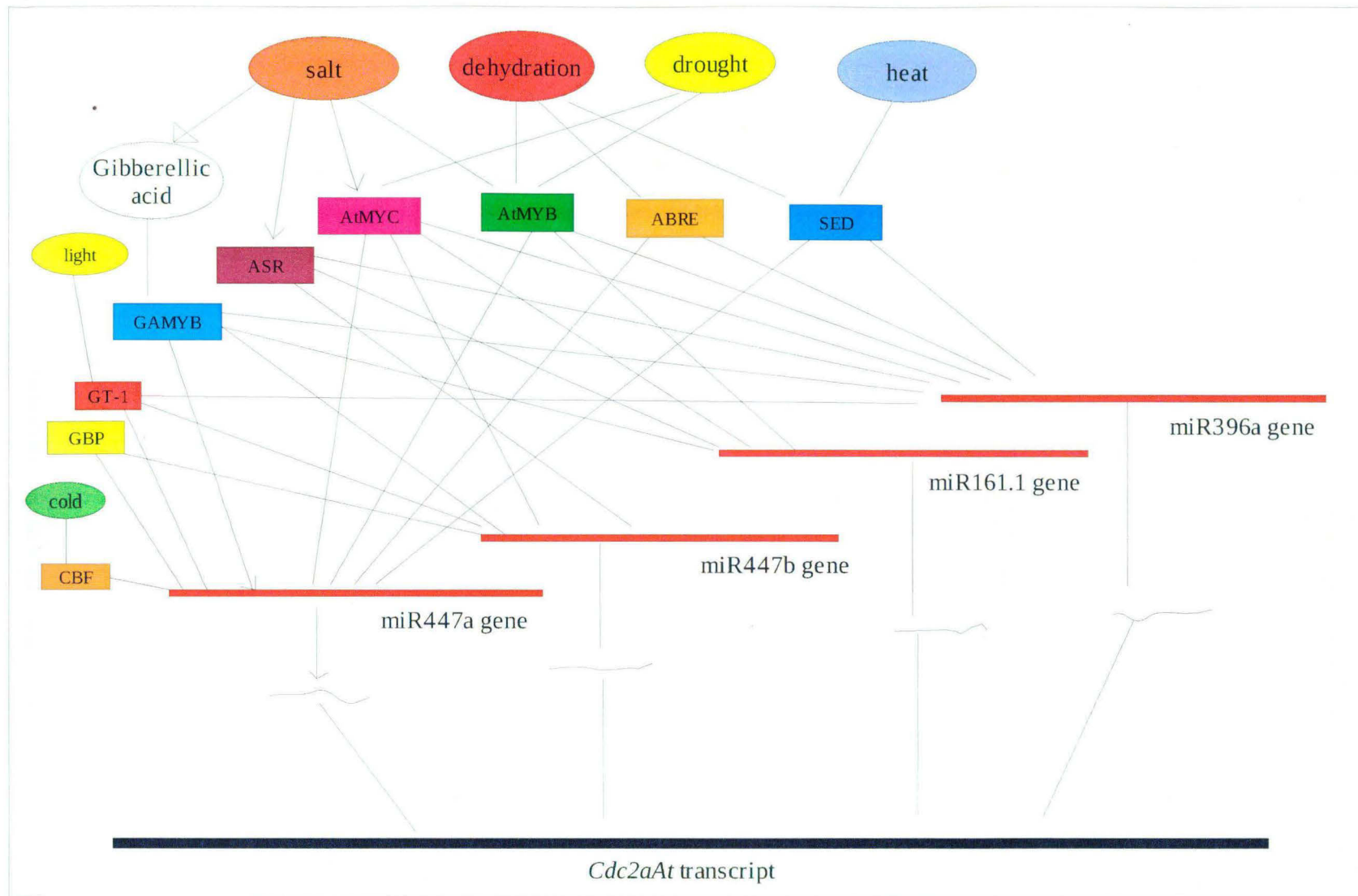


Fig 13: Proposed miRNAs mediated regulatory pathways for *cdc2aAt*.

Conclusion

CHAPTER 5 : Conclusion

It is known that cell cycle gene *cdc2aAt* is down-regulated during stress conditions. The role of miRNAs in the down-regulation of cell cycle gene *cdc2aAt* and its possible mechanism is not clear. We investigated this aspect through *in silico* approach. Comparative computational analysis shows that **miR447 family** and **miR161.1** are the putative candidates which targets *cdc2aAt*. The published data shows that the expression of a few miRNAs, such as miR396, increases up to three fold during stress conditions. Analysis of upstream sequence of these miRNAs shows that it contain stress responsive cis-elements, such as AtMYB, MYBAS1, ASR-1, SED, CBF etc., and also share some common elements feature of the miR396. Hence it could be possible that the expression profile of **miR447 family** and **miR161.1** will also show its up-regulation during different stress conditions. These analysis further support our hypothesis that **miR447 family** and **miR161.1** are very likely to be involved in the regulation of *cdc2aAt* and other stressed down-regulated core cell cycle genes.

Published expression profile data shows up-regulation of some TFs during stress conditions which may directly or indirectly regulates the expression of miRNAs. But these data are very old and hence it could not be possible to obtain information regarding some predicted stress induced TFs.

We conclude that the **miR447 family** and/or **miR161.1** are the strong potential candidates which may target cell cycle gene *cdc2aAt* and there is a need for wet lab verification to confirm our findings. Future course of investigation will be to study all cell cycle related genes in *Arabidopsis*.

The summary of our findings can be briefly stated as follows :

(a) **miRNA447 family** and **miR161.1** are the potential candidates for the plant cell cycle gene *cdc2aAt* during stress conditions.

(b) Upstream analysis shows that stress responsive TFs bind to the upstream regions of candidate miRNAs to regulate them.

(c) It is evident that each of these miRNAs is having a unique stress responsive elements and (or) may have a different regulatory pathway.

(d) Out of 20 down-regulated cell cycle genes 3 are targeted by miR447 families.

References

References

- Abe H., Yamaguchi-Shinozaki K., Urao T., Iwasaki T., Hosokawa D. and Shinozaki K. (1997).** Role of Arabidopsis MYC and MYB Homologs in Drought-and Abscisic Acid-Regulated Gene Expression. *The Plant Cell* 9(10):1859-1868.
- Adams P., Thomas J. C., Vernon D. M., Bohnert H. J. & Jensen R. G. (1992).** Distinct cellular and organismic response to salt stress. *Plant Cell Physiol.* 33:1215-1223.
- Aukerman M. J. and Sakari H. (2003).** Regulation of flowering time and floral organ identity by a MicroRNA and its APETALA2-like target genes. *Plant Cell* 15 : 2730–2741.
- Axtell M. J. and Bowman J. L. (2008).** Evolution of plant microRNAs and their targets. *Trends in Plant Science* 13: 343-349.
- Azaizeh H, Steudle E. (1991).** Effects of salinity on water transport of excised maize (*Z.mays* L.) roots. *Plant Physiology* 97:1136-1145.
- Bartel D.P. (2004).** MicroRNAs: genomics, biogenesis, mechanism, and function, *Cell* 116:281-297.
- Baskerville S. and Bartel D.P. (2005).** Microarray profiling of microRNAs reveals frequent coexpression with neighboring miRNAs and host genes. *RNA* 11: 241-47.
- Battal P. (2004).** Effects of Some Mineral Nutrients on Gibberellic Acid Levels in Maize Plants (*ZEA MAYS* L). *Economic Botany* 58(2):195-203.
- Bhattacharyya S. N., Habermacher R., Martine U., Closs E. I. & Filipowicz W. (2006).** Relief of microRNA-Mediated Translational Repression in Human Cells Subjected to Stress. *Cell* 125:1111-1124.
- Bohnert H. J., Nelson D. E. & Jensen R. G. (1995).** Adaptation to environmental stress. *Plant Cell* 7:1099-1011.
- Bray E. A. (1997).** Plant responses to water deficit. *Trends Plant Sci.* 2:48-54.
- Brennecke J., Stark A., Russell R. B., Cohen S. M. (2005).** Principles of microRNA-target recognition. *PLoS Biology* 3: e85.
- Burssens S., Himanen K., van de Cotte B., Beeckman T., Van Montagu M., Inze D. & Verbruggen N. (2000).** Expression of cell cycle regulatory genes and morphological alterations in response to salt stress in *Arabidopsis thaliana*. *Planta* 211 : 632-640.
- Chen X. (2004).** A microRNA as a translational repressor of APETALA2 in *Arabidopsis* flower development. *Science* 303:2022-25.
- Chen, X. (2005)** microRNA biogenesis and function in plants. *FEBS Lett.* 579:5923–5931.

- Chiou Tzyy-Jen** (2007). The role of miRNAs in sensing nutrient stress. *Plant, Cell and Environment* 30: 323-332.
- Cristianini N. & Hahn M. M.** (2007). *Introduction to Computational Genomics : A Case Studies Approach*. Cambridge University Press.
- Dai X. & Zhao P. X.** psRNATarget: A Plant Small RNA Regulator Target Analysis Server, manuscript under preparation.
- Erica H. and O'Driscoll L.** (2008). Molecular medicine of microRNAs: structure, function and implications for diabetes. *Expert Reviews in Molecular Medicine* 10: e24 CUP.
- Griffiths-Jones S., Grocock R. J., Dongen S. v., Bateman A. & Enright A. J.** (2005). miRBase: microRNA sequences, targets and gene nomenclature. *Nucleic Acids Research* 34:D140-D144. (<http://microrna.sanger.ac.uk/>)
- Gueta-Dahan Y., Yaniv Z., Zilinskas B. A. & Ben-Hayyim G.** (1997). Salt and oxidative stress : similar and specific responses and their relationship to salt tolerance in Citrus. *Planta* 203:460-469.
- Hennessy E. & O'Driscoll L.** (2008). Molecular medicine of microRNAs : structure, function and implications for diabetes. *Expert Reviews in Molecular Medicine* 10: e24 CUP.
- Ishitani M., Xionig L., Stevenson B. & Zhu J-K.** (1997). Genetic analysis of osmotic and cold stress signal transduction in *Arabidopsis* : interactions and convergence of abscisic acid-dependent and abscisic acid pathways. *Plant Cell* 9:1935-1949.
- Jaglo-Ottosen K. R., Gilmour S. J., Zarka D.G., Schabenberger O. & Thomashow M. F.** (1998). *Arabidopsis* CBF1 overexpression induces COR genes and enhances freezing tolerance. *Science* 280:104-106.
- Jain R. K. & Selvaraj G.** (1997). Molecular genetic improvement of salt tolerance in plants. *Biotech. Annu. Rev.* 3:245-267.
- Jian-Kang Zh** (2002). Salt and Drought Stress Signal Transduction in Plants. *Annu. Rev. Plant Biol.* 53: 247-73.
- John B., Enright A. J., Aravin A., Tuschl T., Sander C., Marks D.S.** (2004). Human MicroRNA Targets. *PLoS Biol* 2(11): e363.
- Jones-Rhoades M. W., Bartel D. P. and Bartel B.** (2006). MicroRNAs and Their Regulatory Roles in Plants. *Annu. Rev. Plant Biol.* 57: 19-53.
- Joung J.-G. & Fei Z.** (2009). Identification of microRNA regulatory modules in *Arabidopsis* via a probabilistic graphical model. *BIOINFORMATICS* 25(3):387-393.
- Kertesz M., Iovino N., Unnerstall U., Gaul U. & Segal E.** (2007). The role of site accessibility in microRNA target recognition. *Nature Genetics* 39:1278-1284.

Kim S. Y., Chung H. J. and Thomas T. L. (1997). Isolation of a novel class of bZIP transcription factors that interact with ABA-responsive and embryo-specification elements in the Dc3 promoter using a modified yeast one-hybrid system. *the plant journal* 11(6):1237-1251.

Lee R. C., Feinbaum R. L. & Ambros V. (1993). The *C. elegans* heterochronic gene *lin-4* encodes small RNAs with antisense complementarity to *lin-14*. *Cell* 75:843-854.

Levitt J. (1980). *Responses of plant to Environmental Stress Chilling, Freezing and High Temperature Stresses*. New York: Academic. 2nd ed.

Levine M. & Tjian R. (2003). Transcription regulation and animal diversity. *Nature* 424:147-151.

Lewin B. (2004). *GENESVIII*. Pearson Prentice Hall .

Li Sung-Chou, Tang P., Lin Wen-Chang (2007). Intronic MicroRNA : Discovery and Biological Implications. *DNA and Cell Biology* 26(4) : 195-207.

Liu H.-H., Tian X., Li Y.-J., Wu C.-A. & Zheng C.-C. (2008). Microarray-based analysis of stress-regulated microRNAs in *Arabidopsis thaliana*. *RNA* 14:836-843.

Mallory A. C., Reinhart B. J., Jones-Rhoades M. W., Tang G., Zamore P. D., Barton M. K. and Bartel D. P. (2004). MicroRNA control of PHABULOSA in leaf development : importance of pairing to the microRNA 5' region. *EMBO J.* 23 : 3356-64.

Mallory A. C. and Vaucheret H. (2006). Functions of microRNAs and related small RNAs in plants. *Nature Genetics* Vol. 38:S31-S36.

Matranga C. and Zamore P. D. (2007). Small Silencing RNAs. *Current Biology* Vol.17:R789-R793.

Matthew W. Jones-Rhoades and David P. Bartel (2004). Computational Identification of Plant MicroRNAs and Their Targets, Including a Stress-Induced miRNA. *Molecular Cell* 14: 787-799.

Maziere P. and Enright A. J. (2007). Prediction of microRNA targets. *Drug Discovery Today* 12 : 452-458

Mount D. (2004). *Bioinformatics: Sequence and Genome Analysis*. 2nd ed. CSHL Press.

Palatnik J. F., Allen E., Wu X., Schommer C., Schwab R., et al. (2003). Control of leaf morphogenesis by microRNAs. *Nature* 425 : 257-63.

Qi Y., Denli A. M. & Hannon G. J. (2005). Biochemical specialization within *Arabidopsis* RNA silencing pathways. *Molecular Cell* 19 : 421-28.

Rajewsky N. (2006). MicroRNA target predictions in animals. *Nature Genetics* 38 : S8-13.

Rehmsmeier M., Steffen P., Hochsmann M. & Giegerich R. (2004). Fast and effective prediction of microRNA/target duplexes. *RNA* 10:1507-1517.

- Reinhart B.J., Weinstein E.G., Rhoades M.W., Bartel B. and Bartel D.P. (2002).** MicroRNAs in Plants. *Genes Dev.* 16: 1616-26.
- Rhoades M. W., Reinhart B. J., Lim L. P., Burge C. B., Bartel B. and Bartel D. P. (2002).** Prediction of Plant MicroRNA Targets. *Cell* 110 : 513-520.
- Rusinov V., Baev V., Minkov I. N. & Tabler M. (2005).** MicroInspector: a web tool for detection of miRNA binding sites in an RNA sequence. *Nucleic Acid Res.* 33:696-700.
- Souret F. F., Kastenmayer J. P., Green P.J. (2004).** AtXRN4 degrades mRNA in *Arabidopsis* and its substrates include selected miRNA targets. *Molecular Cell* 15 : 173-83.
- Sunkar R. and Zhu J.-K. (2004)** Novel and Stress-Regulated MicroRNAs and Other Small RNAs from *Arabidopsis*. *The Plant Cell* 16: 2001-2019.
- Susi P., Hohkuri M., Wahlroos T. and Kilby N. J. (2004).** Characteristics of RNA silencing in plants: similarities and differences across kingdoms. *Plant Molecular Biology* 54: 157-174.
- Tsang J. , Zhu J. & Oudenaarden A.v (2007).** MicroRNA-Mediated Feedback and Feedforward Loops Are Recurrent Network Motifs in Mammals. *Molecular Cell* 26:753–767.
- Vandepoele K, Raes J, De Veylder L, Rouze P, Rombauts S, Inze D. (2002)** Genome-wide analysis of core cell cycle genes in *Arabidopsis*. *Plant Cell* 4: 903-916.
(<http://www.arabidopsis.org/browse/genefamily/index.jsp>)
- Vasudevan, S., Tong Y. & Steitz J. A. (2007).** Switching from repression to activation: microRNAs can up-regulate translation. *Science* 318:1931–1934.
- Voinnet O. (2009).** Origin, Biogenesis, and Activity of Plant MicroRNAs. *Cell* 136:669-687.
- Wang X. J., Reyes J. L., Chua N. H., Gassterland T. (2004).** Prediction and identification of *Arabidopsis thaliana* microRNAs and their mRNA targets. *Genome Biology* 5 : R65.
- Wangxia W., Basia V. & Arie A. (2003).** Plant responses to drought, salinity and extreme temperatures: towards genetic engineering for stress tolerance. *Planta* 218: 1-14.
- Washio K. (2003).** Functional Dissections between GAMYB and Dof Transcription Factors Suggest a Role for Protein-Protein Associations in the Gibberellin-Mediated Expression of the *RAmy1A* Gene in the Rice Aleurone. *Plant Physiol.* 133(2):850-863.
- Weaver R. F. & Hedrick. P. W. (1997).** *Genetics* . Wm. C. Brown Publishers.
- Wightman B., Ha I. & Ruvkun G. (1993).** Post-transcriptional regulation of the heterochronic gene *lin-14* by *lin-4* mediates temporal pattern formation in *C. elegans*. *Cell* 75:855–862.
- Zhang B.H., Pan X.P., Wang Q.L., Cobb G. P. and Anderson T. A. (2005).**

Identification and characterization of new plant microRNAs using EST analysis. *Cell Research* 15:336-360.

Zhou Y., Ferguson J., Chang J. T. & Kluger Y. (2007). Inter- and intra-combinatorial regulation by transcription factors and microRNAs. *BMC Genomics* 8:396-406.

WEB SOURCES

Differential Expression Data :

<http://affymetrix.arabidopsis.info/narrays/experimentbrowse.pl>

Cis-Element Data :

<http://arabidopsis.med.ohio-state.edu/AtcisDB/bindingSiteContent.jsp>

<http://arabidopsis.med.ohio-state.edu/>

miRNAs Data :

<http://microrna.sanger.ac.uk/>

Whole Genome sequence Data :

ftp://ftp.ncbi.nih.gov/genomes/Arabidopsis_thaliana/

Core Cell Cycle Gene Data :

<http://www.arabidopsis.org/browse/genefamily/cellcycle.jsphttp>

psRNATarget :

<http://bioinfo3.noble.org/miRU2/index.php?function=function3>

MicroInspector :

<http://mirna.imbb.forth.gr/microinspector/index.html>

Appendix

Appendix A

**Table : List of stress responsive TFs binding *cis*-elements.
(source: AGRIS and AtcisDB)**

Name	Consensus motif	Title
AtMYB2 BS in RD22 (MYB)	CTAACCA	Drought-responsive element, mediated through higher sensitivity to the plant hormone ABA.
AtMYC2 BS in RD22 (BHLH)	CACATG	Drought-responsive element, mediated through higher sensitivity to the plant hormone ABA.
DREB1&2 BS in rd29a (AP2-EREBP)	TACCGACAT	Improving plant drought, salt, and freezing tolerance by gene transfer of a single stress-inducible transcription factor
ABFs binding site motif	CACGTGGC	A plant leucine zipper protein that recognizes an abscisic acid response element
ABRE binding site motif	(C/T)ACGTGGC	ABFs, a family of ABA-responsive element binding factors
ABRE like binding site motif	(C/G/T)ACGTG(G/T) (A/C)	Molecular responses to dehydration and low temperature
CBF2 binding site motif (AP2-EREBP)	CCACGTGG	The cis-regulatory element involved in ABA and water-stress responses
HSEs binding site motif	AGAANNTTCT	Heat stress response element
ARF binding site motif	TGTCTC	Binding element of auxin response factor
ACE promoter motif	GACACGTAGA	UV/blue light-response elements
DPBF1&2 binding site motif	ACACNNG	ABA-responsive and embryo-specification elements in the Dc3 promoter
DRE-like promoter	(A/G/T) (A/G)CCGACN(A/T)	Response to Environmental Stresses
EIL binding site in ERF	TTCAAGGGGGCAT GTATCTTGAA	ETHYLENE-INSENSITIVE3 and ETHYLENE-RESPONSE element
ERE promoter motif	TAAGAGCCGCC	ethylene-responsive region in the promoter of a tobacco class I chitinase gene
GCC-box promoter motif	GCCGCC	Molecular responses to dehydration and low temperature
GT-promoter motif	TGTGTGGTTAATAT G	light-responsive elements upstream of the rbcS-3A gene.
LTRE promoter motif	ACCGACA	low-temperature-induced response element

MRE motif in CHS (MYB)	TCTAACCTACCA	IdentificationUV/blue light-response elements in the Arabidopsis thaliana chalcone synthase promoter
OBP-1-4-5 BS in GST6 (C2C2-Dof)	TACACTTTTGG	promoter of a H2O2-inducible, Arabidopsis glutathione S-transferase gene contains closely linked OBF- and OBP1-binding RT
ERF1 BS in AtCH1-B	GCCGCC	ETHYLENE-INSENSITIVE3 and ETHYLENE-RESPONSE element
GBF1/2/3 BS in ADH1	CCACGTGG	a maize G-box binding factor that is induced by hypoxia
MYB1 binding site motif (MYB)	(A/C)TCC(A/T)ACC	G-box-binding factor
MYB2 binding site motif (MYB)	TAAct(G/C)GTT	MYB transcription factors
MYB3 binding site motif (MYB)	TAActAAC	response to environmental stresses.
MYB4 binding site motif (MYB)	A(A/C)C(A/T)A(A/C)C	response to environmental stresses
I box promoter motif	GATAAG	An evolutionarily conserved protein binding sequence upstream of a plant light-regulated gene
PRHA BS in PAL1	TAATTGACTCAATT A	auxin-induced expression element of the Arabidopsis prha homeobox gene
SORLREP1	TT(A/T)TACTAGT	key promoter motifs involved in the network of light-regulated gene expression
SORLREP2	ATAAAACGT	key promoter motifs involved in the network of light-regulated gene expression
SORLREP3	TGTATATAT	key promoter motifs involved in the network of light-regulated gene expression
SORLREP4	CTCCTAATT	key promoter motifs involved in the network of light-regulated gene expression
SORLREP5	TTGCATGACT	key promoter motifs involved in the network of light-regulated gene expression
SORLIP1	AGCCAC	key promoter motifs involved in the network of light-regulated gene expression
SORLIP2	GGGCC	key promoter motifs involved in the network of light-regulated gene expression
SORLIP3	CTCAAGTGA	key promoter motifs involved in the network of light-regulated gene expression
SORLIP4	GTATGATGG	key promoter motifs involved in the network of light-regulated gene expression
SORLIP5	GAGTGAG	key promoter motifs involved in the network of light-regulated gene expression

Appendix B

Result of MatchTM program

Scanning	sequence	ID:	icrona_miR447a			
matrix identifier	position	(strand)	core match	matrix match	sequence	factor name
P\$PBF_01	103	(+)	1	0.99	gaIAAAAGgga	PBF
P\$DOF_Q2	103	(+)	1	0.94	gaIAAAAGgga	Dof
P\$PBF_Q2	106	(+)	1	1	aAAAGG	PBF
P\$GT1_Q6	160	(+)	1	0.91	GTGAACA	GT-1
P\$GBP_Q6	165	(+)	1	0.91	cacCACGTgag	GBP
P\$CPRF3_Q2	166	(+)	1	0.92	acCACGTiga	CPRF-3
P\$CPRF2_Q2	166	(+)	1	0.94	acCACGTiga	CPRF-2
P\$O2_Q2	166	(+)	1	0.97	acCACGTiga	Opaque-2
P\$TGA1A_Q2	166	(+)	1	0.96	acCACGTiga	TGA1a
P\$CPRF3_Q1	166	(+)	1	0.93	acCACGTiga	CPRF-3
P\$CPRF2_Q1	166	(+)	1	0.94	acCACGTiga	CPRF-2
P\$RITA1_Q1	167	(+)	1	0.97	cCACGTI	RITA-1
P\$OCSBF1_Q1	168	(+)	1	1	CACGT	OCSBF-1
P\$KNOX3_Q1	206	(+)	1	0.96	attaTGACAaaa	Knox3
P\$ATMYB77_Q1	234	(+)	1	0.9	aatgggCGGTTcg	AtMYB-77
P\$ERF2_Q1	283	(+)	1	0.93	tgCCGCC	ERF2
P\$SED_Q2	333	(+)	1	0.99	ctcICCTTTI	SED
P\$BPC1_Q2	436	(+)	1	0.99	AGAAAt	BPC1
P\$PBF_Q1	446	(+)	1	0.97	gcgAAAAGgga	PBF
P\$DOF_Q2	446	(+)	1	0.95	gcgAAAAGgga	Dof
P\$PBF_Q2	449	(+)	1	1	aAAAGG	PBF
P\$BPC1_Q2	459	(+)	1	1	AGAAAa	BPC1
P\$GAMYB_Q2	629	(+)	1	0.91	ttggIACAACaca	GAMYB
P\$RAV1_Q1	632	(+)	1	0.95	gtaCAACacata	RAV1
P\$C1_Q2	662	(+)	1	0.93	aaAACTAacat	C1
P\$MYBAS1_Q1	664	(+)	1	0.98	aaCTAACat	MYBAS1
P\$BPC1_Q2	684	(+)	1	0.99	AGAAAc	BPC1
P\$RAV1_Q2	853	(+)	1	0.99	ttcACCTGccat	RAV1

Scanning	sequence	ID:	icrona_miR447b			
matrix identifier	position	(strand)	core match	matrix match	sequence	factor name
P\$GBP_Q6	183	(+)	1	0.91	tcaCACGTgtgt	GBP
P\$HBP1A_Q2	184	(+)	1	0.91	cacACGTGtg	HBP-1a
P\$CPRF_Q2	184	(+)	1	0.92	caCACGTgtg	CPRF-1
P\$TAF1_Q2	184	(+)	1	0.95	cacACGTGtg	TAF-1
P\$CPRF3_Q2	184	(+)	1	0.98	caCACGTgtg	CPRF-3
P\$CPRF2_Q2	184	(+)	1	0.99	caCACGTgtg	CPRF-2
P\$O2_Q2	184	(+)	1	0.95	caCACGTgtg	Opaque-2
P\$TGA1A_Q2	184	(+)	1	0.99	caCACGTgtg	TGA1a
P\$EMBP1_Q2	184	(+)	1	0.92	cacACGTGtg	EmBP-1b
P\$CPRF1_Q1	184	(+)	1	0.94	caCACGTgtg	CPRF-1
P\$TAF1_Q1	184	(+)	1	0.97	cacACGTGtg	TAF-1
P\$CPRF3_Q1	184	(+)	1	0.98	caCACGTgtg	CPRF-3
P\$CPRF2_Q1	184	(+)	1	0.99	caCACGTgtg	CPRF-2
P\$RITA1_Q1	185	(+)	1	0.98	aCACGTg	RITA-1
P\$bHLH66_Q1	185	(+)	1	0.92	aCACGTgt	bHLH66
P\$OCSBF1_Q1	186	(+)	1	1	CACGT	OCSBF-1
P\$GT1_Q6	193	(+)	1	0.91	GTGAAGa	GT-1
P\$ALFIN1_Q2	194	(+)	1	0.96	tgaagaGTGGGgagt	Alfin-1
P\$PBF_Q1	236	(+)	1	0.96	cggAAAAGtcg	PBF
P\$DOF_Q2	236	(+)	1	0.94	cggAAAAGtcg	Dof
P\$DOF1_Q1	253	(+)	1	0.98	ctTAAAGaat	Dof1
P\$PEND_Q1	376	(+)	1	0.96	tAAGAAAt	PEND
P\$PBF_Q1	393	(+)	1	0.99	cagAAAAGaga	PBF
P\$DOF_Q2	393	(+)	1	0.95	cagAAAAGaga	Dof
P\$BPC1_Q2	394	(+)	1	1	AGAAAa	BPC1
P\$BPC1_Q2	401	(+)	1	1	AGAAAa	BPC1
P\$C1_Q2	466	(+)	1	0.92	gaAACTAaatt	C1
P\$BPC1_Q2	492	(+)	1	0.99	AGAAAt	BPC1
P\$GBP_Q6	581	(+)	1	0.91	agaCACGTgtt	GBP
P\$EMBP1_Q2	582	(+)	1	0.96	gaCACGTgtt	EmBP-1
P\$HBP1A_Q2	582	(+)	1	0.92	gacACGTGtt	HBP-1a

Sheet1

P\$CPRF_Q2	582	(+)	1	0.97	gaCACGTgtt	CPRF-1
P\$TAF1_Q2	582	(+)	1	0.97	gacACGTGtt	TAF-1
P\$CPRF3_Q2	582	(+)	1	0.98	gaCACGTgtt	CPRF-3
P\$CPRF2_Q2	582	(+)	1	0.99	gaCACGTgtt	CPRF-2
P\$O2_Q2	582	(+)	1	0.95	gaCACGTgtt	Opaque-2
P\$TGA1A_Q2	582	(+)	1	0.98	gaCACGTgtt	TGA1a
P\$EMBP1_Q2	582	(+)	1	0.93	gacACGTGtt	EmBP-1b
P\$CPRF1_Q1	582	(+)	1	0.97	gaCACGTgtt	CPRF-1
P\$TAF1_Q1	582	(+)	1	0.98	gacACGTGtt	TAF-1
P\$CPRF3_Q1	582	(+)	1	0.99	gaCACGTgtt	CPRF-3
P\$CPRF2_Q1	582	(+)	1	0.99	gaCACGTgtt	CPRF-2
P\$TGA1B_Q1	582	(+)	1	0.95	gaCACGTgtt	TGA1b
P\$RITA1_Q1	583	(+)	1	0.98	aCACGTg	RITA-1
P\$bHLH66_Q1	583	(+)	1	0.92	aCACGTgt	bHLH66
P\$OCSBF1_Q1	584	(+)	1	1	CACGT	OCSBF-1
P\$ASR1_Q1	621	(+)	1	1	ACCCA	ASR-1
P\$C1_Q2	648	(+)	1	0.92	aaAACTAtact	C1
P\$KNOX3_Q1	661	(+)	1	0.96	tgatTGACAata	Knox3
P\$GAMYB_Q2	668	(+)	1	0.92	caataACAACata	GAMYB
P\$RAV1_Q1	671	(+)	1	0.96	taaCAACAtatc	RAV1
P\$GBP_Q6	679	(+)	1	0.92	tatCACGTtcaa	GBP
P\$CPRF3_Q2	680	(+)	1	0.91	atCACGTtca	CPRF-3
P\$CPRF2_Q2	680	(+)	1	0.93	atCACGTtca	CPRF-2
P\$O2_Q2	680	(+)	1	0.93	atCACGTtca	Opaque-2
P\$TGA1A_Q2	680	(+)	1	0.96	atCACGTtca	TGA1a
P\$CPRF3_Q1	680	(+)	1	0.92	atCACGTtca	CPRF-3
P\$CPRF2_Q1	680	(+)	1	0.93	atCACGTtca	CPRF-2
P\$RITA1_Q1	681	(+)	1	0.97	ICACGTt	RITA-1
P\$OCSBF1_Q1	682	(+)	1	1	CACGT	OCSBF-1
P\$BPC1_Q2	690	(+)	1	0.99	AGAAAc	BPC1
P\$TEIL_Q1	786	(+)	1	0.92	ATGTAtat	TEIL
P\$GT1_Q6	852	(+)	1	0.91	GTAAACA	GT-1
P\$RAV1_Q2	878	(+)	1	0.99	ttaCCTGccat	RAV1

Scanning	sequence	ID:	icrorna_miR447c			
matrix identifier	position	(strand)	core match	matrix match	sequence	factor name
P\$BZR1_Q1	38	(+)	1	0.9	CGTGct	BZR1
P\$TAF1_Q2	106	(+)	1	0.92	cttACGTGat	TAF-1
P\$TAF1_Q1	106	(+)	1	0.94	cttACGTGat	TAF-1
P\$ATHB5_Q1	136	(+)	1	0.92	ctaTTATTg	ATHB-5
P\$GBP_Q6	230	(+)	1	0.91	tcaCACGTgtgt	GBP
P\$HBP1A_Q2	231	(+)	1	0.91	cacACGTgtg	HBP-1a
P\$CPRF_Q2	231	(+)	1	0.92	caCACGTgtg	CPRF-1
P\$TAF1_Q2	231	(+)	1	0.95	cacACGTgtg	TAF-1
P\$CPRF3_Q2	231	(+)	1	0.98	caCACGTgtg	CPRF-3
P\$CPRF2_Q2	231	(+)	1	0.99	caCACGTgtg	CPRF-2
P\$O2_Q2	231	(+)	1	0.95	caCACGTgtg	Opaque-2
P\$TGA1A_Q2	231	(+)	1	0.99	caCACGTgtg	TGA1a
P\$EMBP1_Q2	231	(+)	1	0.92	cacACGTgtg	EmBP-1b
P\$CPRF1_Q1	231	(+)	1	0.94	caCACGTgtg	CPRF-1
P\$TAF1_Q1	231	(+)	1	0.97	cacACGTgtg	TAF-1
P\$CPRF3_Q1	231	(+)	1	0.98	caCACGTgtg	CPRF-3
P\$CPRF2_Q1	231	(+)	1	0.99	caCACGTgtg	CPRF-2
P\$RITA1_Q1	232	(+)	1	0.98	aCACGTg	RITA-1
P\$bHLH66_Q1	232	(+)	1	0.92	aCACGTgt	bHLH66
P\$OCSBF1_Q1	233	(+)	1	1	CACGT	OCSBF-1
P\$GT1_Q6	240	(+)	1	0.91	GTAAAGA	GT-1
P\$ALFIN1_Q2	241	(+)	1	0.96	tgaagaGTGGGgagt	Allin1
P\$PBF_Q1	283	(+)	1	0.96	cggAAAAGtcg	PBF
P\$DOF_Q2	283	(+)	1	0.94	cggAAAAGtcg	Dof
P\$DOF1_Q1	300	(+)	1	0.98	cttTAAAGaga	Dof1
P\$DOF1_Q1	346	(+)	1	0.98	cacTAAAGtct	Dof1
P\$RAV1_Q1	443	(+)	1	0.99	tagCAACAaaaa	RAV1
P\$PBF_Q1	453	(+)	1	0.98	aaaAAAAGaaa	PBF
P\$DOF_Q2	453	(+)	1	1	aaaAAAAGaaa	Dof
P\$BPC1_Q2	459	(+)	1	0.99	AGAAAt	BPC1
P\$PBF_Q1	471	(+)	1	0.96	gaaAAAAGatg	PBF
P\$DOF_Q2	471	(+)	1	0.98	gaaAAAAGatg	Dof
P\$PBF_Q1	482	(+)	1	0.97	tagAAAAGtct	PBF
P\$DOF_Q2	482	(+)	1	0.94	tagAAAAGtct	Dof
P\$BPC1_Q2	483	(+)	1	1	AGAAAa	BPC1
P\$TEIL_Q1	538	(+)	1	0.93	ATGTAtgt	TEIL
P\$C1_Q2	667	(+)	1	0.93	aaAACTAacat	C1
P\$MYBAS1_Q1	669	(+)	1	0.98	aaCTAACat	MYBAS1

Sheet1

P\$BPC1_Q2	688	(+)	1	0.99	AGAAAc	BPC1
P\$SBF1_01	782	(+)	1	0.92	ctatATTTAAaaag	SBF-1
P\$DOF2_01	788	(+)	1	0.99	ttaaAAAGCtg	Dof2
P\$DOF3_01	788	(+)	1	0.98	ttaaAAAGCtg	Dof3
P\$PBF_01	788	(+)	1	0.97	ttaaAAAGCtg	PBF
P\$DOF_Q2	788	(+)	1	0.97	ttaaAAAGCtg	Dof
P\$SPF1_Q2	813	(+)	1	0.92	gcATAGTata	SPF1
P\$P_01	836	(+)	1	1	acCTACCag	P
P\$AGP1_01	943	(+)	1	0.94	agAGATCigt	AGP1
P\$ARR10_01	945	(+)	1	0.98	AGATCtg	ARR10

matrix identifier	position	(strand)	core match	matrix match	sequence	factor name
P\$SED_Q2	46	(+)	1	0.97	ttctCCTTTa	SED
P\$C1_Q2	65	(+)	1	0.95	aaAACTActag	C1
P\$SBF1_01	103	(+)	1	0.92	ttatTTAAAtta	SBF-1
P\$BPC1_Q2	191	(+)	1	1	AGAAa	BPC1
P\$C1_Q2	201	(+)	1	0.95	caAACTAatta	C1
P\$PBF_01	208	(+)	1	0.97	attAAAAGtat	PBF
P\$DOF_Q2	208	(+)	1	0.95	attAAAAGtat	Dof
P\$ATHB1_01	242	(+)	1	0.94	acataATTATccg	Athb-1
P\$ATHB5_01	245	(+)	1	0.92	taaTTATTc	ATHB-5
P\$PBF_01	284	(+)	1	0.97	tttAAAAGaag	PBF
P\$DOF_Q2	284	(+)	1	0.95	tttAAAAGaag	Dof
P\$PBF_01	293	(+)	1	0.96	agaAAAAGtta	PBF
P\$BPC1_Q2	293	(+)	1	1	AGAAa	BPC1
P\$DOF_Q2	293	(+)	1	0.98	agaAAAAGtta	Dof
P\$ABZ1_01	312	(+)	1	0.95	aagttACGTGgtcc	ABZ1
P\$TAF1_Q2	314	(+)	1	0.94	gttACGTGgt	TAF-1
P\$EMBP1_Q2	314	(+)	1	0.93	gttACGTGgt	EmBP-1b
P\$TAF1_01	314	(+)	1	0.96	gttACGTGgt	TAF-1
P\$DOF1_01	361	(+)	1	0.98	tttTAAAGctg	Dof1
P\$DOF2_01	361	(+)	1	1	tttTAAAGctg	Dof2
P\$DOF3_01	361	(+)	1	0.97	tttTAAAGctg	Dof3
P\$KNOX3_01	369	(+)	1	0.96	ctgtTGACAaaa	Knox3
P\$PBF_01	374	(+)	1	0.97	gacAAAAGaaa	PBF
P\$DOF_Q2	374	(+)	1	0.94	gacAAAAGaaa	Dof
P\$BPC1_Q2	380	(+)	1	1	AGAAa	BPC1
P\$PBF_01	381	(+)	1	0.97	gaaAAAAGaaa	PBF
P\$DOF_Q2	381	(+)	1	1	gaaAAAAGaaa	Dof
P\$BPC1_Q2	387	(+)	1	1	AGAAa	BPC1
P\$PBF_Q2	400	(+)	1	0.96	cAAAGG	PBF
P\$C1_Q2	422	(+)	1	0.94	tAACTAgata	C1
P\$SED_Q2	460	(+)	1	0.92	cgaaCCTTTt	SED
P\$C1_Q2	634	(+)	1	0.96	caAACTActta	C1
P\$KNOX3_01	653	(+)	1	0.98	ctgtTGACAact	Knox3
P\$GAMYB_Q2	654	(+)	1	0.95	ttgtGACAACTga	GAMYB
P\$KNOX3_01	660	(+)	1	0.98	caactGACAagt	Knox3
P\$SBF1_01	717	(+)	1	0.92	lcatctTTAAAtatt	SBF-1
P\$RAV1_01	740	(+)	1	0.95	actCAACAgatc	RAV1
P\$AGP1_01	745	(+)	1	0.95	acAGATCitt	AGP1
P\$ARR10_01	747	(+)	1	0.96	AGATCtt	ARR10
P\$MYBAS1_01	832	(+)	1	0.95	gaCTAACaaaa	MYBAS1
P\$BPC1_Q2	862	(+)	1	1	AGAAa	BPC1
P\$DOF1_01	879	(+)	1	0.97	ggcTAAAGitt	Dof1
P\$SED_Q2	932	(+)	1	0.94	acaaCCTTTg	SED
P\$KNOX3_01	936	(+)	1	0.97	ccitTGACAatca	Knox3

matrix identifier	position	(strand)	core match	matrix match	sequence	factor name
P\$BZR1_01	14	(+)	1	0.9	CGTGCt	BZR1
P\$BPC1_Q2	24	(+)	1	0.99	AGAAAt	BPC1
P\$PBF_01	34	(+)	1	0.97	aaaAAAAGgta	PBF
P\$DOF_Q2	34	(+)	1	0.98	aaaAAAAGgta	Dof
P\$PBF_Q2	37	(+)	1	1	aAAAGG	PBF
P\$GAMYB_Q2	64	(+)	1	0.91	gacgaACAACtct	GAMYB
P\$BPC1_Q2	120	(+)	1	0.99	AGAAAt	BPC1
P\$BPC1_Q2	143	(+)	1	1	AGAAa	BPC1
P\$DOF2_01	188	(+)	1	0.98	attcAAAGCta	Dof2
P\$DOF3_01	188	(+)	1	0.98	attcAAAGCta	Dof3
P\$P_01	194	(+)	1	0.92	agCTACCTa	P

Sheet1

P\$DOF1_01	289	(+)	1	0.98	taTAAAGatg	Dof1
P\$C1_Q2	305	(+)	1	0.94	taAACTAtaat	C1
P\$ABZ1_01	387	(+)	1	0.9	ttctCACGTGact	ABZ1
P\$GBP_Q6	388	(+)	1	0.95	tcCACGTgact	GBP
P\$HBP1A_Q2	389	(+)	1	0.96	ctCACGTGac	HBP-1a
P\$TAF1_Q2	389	(+)	1	0.97	ctCACGTGac	TAF-1
P\$CPRF3_Q2	389	(+)	1	0.99	ctCACGTgac	CPRF-3
P\$CPRF2_Q2	389	(+)	1	1	ctCACGTgac	CPRF-2
P\$O2_Q2	389	(+)	1	0.97	ctCACGTgac	Opaque-2
P\$TGA1A_Q2	389	(+)	1	1	ctCACGTgac	TGA1a
P\$CG1_Q6	389	(+)	1	0.92	ctCACGTgact	CG1
P\$CPRF1_01	389	(+)	1	0.91	ctCACGTgac	CPRF-1
P\$TAF1_01	389	(+)	1	0.98	ctCACGTGac	TAF-1
P\$CPRF3_01	389	(+)	1	1	ctCACGTgac	CPRF-3
P\$CPRF2_01	389	(+)	1	1	ctCACGTgac	CPRF-2
P\$TGA1B_01	389	(+)	1	0.94	ctCACGTgac	TGA1b
P\$RITA1_01	390	(+)	1	1	TCACGTg	RITA-1
P\$BHLH66_01	390	(+)	1	0.92	TCACGTga	bHLH66
P\$OCSBF1_01	391	(+)	1	1	CACGT	OCSBF-1
P\$AGP1_01	557	(+)	1	0.95	ttAGATCtgc	AGP1
P\$ARR10_01	559	(+)	1	0.98	AGATCtg	ARR10
P\$SBF1_01	731	(+)	1	0.92	atttgaTTAAaatt	SBF-1
P\$PBF_01	846	(+)	1	0.96	tggAAAAGaac	PBF
P\$DOF_Q2	846	(+)	1	0.95	tggAAAAGaac	Dof
P\$KNOX3_01	864	(+)	1	0.96	gcaaTGACAaac	Knox3
P\$PBF_01	873	(+)	1	0.96	aacAAAAGatt	PBF
P\$DOF_Q2	873	(+)	1	0.93	aacAAAAGatt	Dof
P\$PBF_Q2	916	(+)	1	0.96	cAAAGG	PBF
P\$O2_Q3	926	(+)	1	0.93	GATGAgatgt	Opaque-2

matrix identifier	position	(strand)	core match	matrix match	sequence	factor name
P\$KNOX3_01	47	(+)	1	0.97	tgaaTGACAtct	Knox3
P\$KNOX3_01	237	(+)	1	0.98	tttgTGACAatt	Knox3
P\$BPC1_Q2	306	(+)	1	0.99	AGAAAc	BPC1
P\$MYBAS1_01	470	(+)	1	0.96	tgCCAACTtca	MYBAS1
P\$BPC1_Q2	528	(+)	1	1	AGAAa	BPC1
P\$PBF_01	529	(+)	1	0.95	gaaAAAAGggt	PBF
P\$DOF_Q2	529	(+)	1	0.98	gaaAAAAGggt	Dof
P\$PBF_Q2	532	(+)	1	1	aAAAGG	PBF
P\$SED_Q2	623	(+)	1	1	gtttCCTTTt	SED
P\$PEND_01	683	(+)	1	0.9	gAAGAAgt	PEND
P\$PBF_Q2	703	(+)	1	0.96	cAAAGG	PBF
P\$KNOX3_01	831	(+)	1	0.97	ggtttGACAaatt	Knox3
P\$GT1_Q6	925	(+)	1	0.97	GTGAaaa	GT-1
P\$SPF1_Q2	941	(+)	1	0.95	aaATAGTatc	SPF1
P\$RAV1_01	948	(+)	1	0.93	atcCAACtctg	RAV1
P\$MYBAS1_01	948	(+)	1	0.98	atCCAACatct	MYBAS1

matrix identifier	position	(strand)	core match	matrix match	sequence	factor name
P\$PBF_01	8	(+)	1	0.99	cagAAAAGgga	PBF
P\$DOF_Q2	8	(+)	1	0.95	cagAAAAGgga	Dof
P\$BPC1_Q2	9	(+)	1	1	AGAAa	BPC1
P\$PBF_Q2	11	(+)	1	1	aAAAGG	PBF
P\$BPC1_Q2	43	(+)	1	1	AGAAa	BPC1
P\$DOF2_01	49	(+)	1	0.99	caagAAAGCta	Dof2
P\$DOF3_01	49	(+)	1	0.99	caagAAAGCta	Dof3
P\$BPC1_Q2	51	(+)	1	1	AGAAAg	BPC1
P\$DOF1_01	55	(+)	1	0.99	agcTAAAGaga	Dof1
P\$BPC1_Q2	114	(+)	1	0.99	AGAAAt	BPC1
P\$C1_Q2	214	(+)	1	0.95	aaAACTAtctc	C1
P\$ASR1_01	242	(+)	1	1	ACCCA	ASR-1
P\$ASR1_01	263	(+)	1	1	ACCCA	ASR-1
P\$PBF_01	274	(+)	1	0.97	catAAAAGtac	PBF
P\$DOF_Q2	274	(+)	1	0.95	catAAAAGtac	Dof
P\$ASR1_01	283	(+)	1	1	ACCCA	ASR-1
P\$TEIL_01	437	(+)	1	0.92	ATGTAtat	TEIL
P\$PBF_01	458	(+)	1	0.97	ataAAAAGtac	PBF
P\$DOF_Q2	458	(+)	1	1	ataAAAAGtac	Dof
P\$ASR1_01	467	(+)	1	1	ACCCA	ASR-1

Sheet1

P\$ASR1_01	505	(+)	1	1	ACCCA	ASR-1
P\$PEND_01	628	(+)	1	0.96	tAAGAAit	PEND
P\$MYBAS1_01	819	(+)	1	0.95	acCTAACtccg	MYBAS1
P\$GAMYB_01	868	(+)	1	0.92	CAACCttc	GAMYB
P\$MYBAS1_01	886	(+)	1	0.98	ctCTAACtcat	MYBAS1

Scanning	sequence	ID:	nicroma_miR163				
matrix identifier	position	(strand)	core match	matrix match	sequence	factor name	
P\$DOF2_01	34	(+)	1	0.98	tcccAAAGCcg	Dof2	
P\$DOF3_01	34	(+)	1	0.98	tcccAAAGCcg	Dof3	
P\$DOF1_01	48	(+)	1	0.99	attTAAAGtga	Dof1	
P\$GT1_Q6	55	(+)	1	0.97	GTGAAaa	GT-1	
P\$PBF_01	56	(+)	1	0.96	tgaAAAAGaag	PBF	
P\$DOF_Q2	56	(+)	1	0.99	tgaAAAAGaag	Dof	
P\$O2_Q2	76	(+)	1	0.93	atggTGACGTgaa	Opaque-2	
P\$ABZ1_01	77	(+)	1	0.93	tgggACGTGaaac	ABZ1	
P\$HBP1A_Q2	79	(+)	1	0.91	gtgACGTGaa	HBP-1a	
P\$TAF1_Q2	79	(+)	1	0.94	gtgACGTGaa	TAF-1	
P\$TAF1_01	79	(+)	1	0.96	gtgACGTGaa	TAF-1	
P\$TGA1B_01	79	(+)	1	0.94	gtGACGTgaa	TGA1b	
P\$PBF_01	172	(+)	1	0.96	gatAAAAGact	PBF	
P\$DOF_Q2	172	(+)	1	0.93	gatAAAAGact	Dof	
P\$DOF1_01	218	(+)	1	0.97	lggTAAAGacg	Dof1	
P\$GT1_Q6	220	(+)	1	0.91	GTAAGaa	GT-1	
P\$PBF_01	258	(+)	1	0.98	cagAAAAGigt	PBF	
P\$DOF_Q2	258	(+)	1	0.94	cagAAAAGigt	Dof	
P\$BPC1_Q2	259	(+)	1	1	AGAAAa	BPC1	
P\$KNOX3_01	314	(+)	1	0.97	gcttTGACAAtt	Knox3	
P\$KNOX3_01	411	(+)	1	0.96	aaaaTGACAata	Knox3	
P\$DOF2_01	455	(+)	1	0.99	tttaAAAGCct	Dof2	
P\$DOF3_01	455	(+)	1	0.98	tttaAAAGCct	Dof3	
P\$PBF_01	455	(+)	1	0.97	tttAAAAGcct	PBF	
P\$DOF_Q2	455	(+)	1	0.93	tttAAAAGcct	Dof	
P\$SPF1_Q2	511	(+)	1	0.96	taATAGTaga	SPF1	
P\$GAMYB_Q2	566	(+)	1	0.95	tataaACAACaaa	GAMYB	
P\$RAV1_01	569	(+)	1	0.95	aaaCAACAaata	RAV1	
P\$ATMYB77_01	606	(+)	1	0.97	tgtagaCAGTTag	AtMYB-77	
P\$AGP1_01	626	(+)	1	0.9	caAGATCaaa	AGP1	
P\$PBF_01	636	(+)	1	0.97	taaAAAAGacc	PBF	
P\$DOF_Q2	636	(+)	1	0.98	taaAAAAGacc	Dof	
P\$SED_Q2	641	(+)	1	0.93	aagaCCTTTt	SED	
P\$ASR1_01	665	(+)	1	1	ACCCA	ASR-1	
P\$GAMYB_Q2	667	(+)	1	0.93	ccacgACAACgac	GAMYB	
P\$MYBAS1_01	715	(+)	1	0.99	acCCAACgacc	MYBAS1	
P\$ASR1_01	715	(+)	1	1	ACCCA	ASR-1	
P\$BPC1_Q2	770	(+)	1	1	AGAAAa	BPC1	
P\$HBP1B_Q6	787	(+)	1	0.92	ttcgtacaCGTCAat	HBP-1b	
P\$CPRF3_Q2	791	(+)	1	0.96	taCACGTcat	CPRF-3	
P\$CPRF2_Q2	791	(+)	1	0.96	taCACGTcat	CPRF-2	
P\$O2_Q2	791	(+)	1	0.96	taCACGTcat	Opaque-2	
P\$TGA1B_Q2	791	(+)	1	0.93	taCACGTcat	TGA1b	
P\$TGA1A_Q2	791	(+)	1	0.98	taCACGTcat	TGA1a	
P\$HBP1A_Q6_01	791	(+)	1	0.91	tacaCGTCAat	HBP-1a	
P\$CPRF3_01	791	(+)	1	0.98	taCACGTcat	CPRF-3	
P\$CPRF2_01	791	(+)	1	0.97	taCACGTcat	CPRF-2	
P\$RITA1_01	792	(+)	1	0.96	aCACGTc	RITA-1	
P\$OCSBF1_01	793	(+)	1	1	CACGT	OCSBF-1	
P\$BPC1_Q2	844	(+)	1	0.99	AGAAAa	BPC1	
P\$C1_Q2	845	(+)	1	0.92	gaAACTAtttt	C1	
P\$C1_Q2	854	(+)	1	0.96	ttAACTAaccac	C1	
P\$MYBAS1_01	856	(+)	1	0.94	aaCTAACacgg	MYBAS1	
P\$PBF_01	923	(+)	1	0.97	atcAAAAGtct	PBF	
P\$DOF_Q2	923	(+)	1	0.92	atcAAAAGtct	Dof	
P\$PBF_Q2	943	(+)	1	0.96	cAAAGG	PBF	

Scanning	sequence	ID:	icroma_miR172a				
matrix identifier	position	(strand)	core match	matrix match	sequence	factor name	
P\$RAV1_01	23	(+)	1	0.92	tgtCAACAatca	RAV1	
P\$RAV1_01	53	(+)	1	0.93	aaaCAACAatca	RAV1	
P\$MYBAS1_01	223	(+)	1	0.95	tgCTAACacaa	MYBAS1	
P\$RAV1_01	228	(+)	1	0.92	acaCAACAacat	RAV1	

Sheet1

P\$GAMYB_Q2	228	(+)	1	0.92	acacaCAACaItt	GAMYB
P\$RAV1_01	231	(+)	1	0.93	caaCAACAIttc	RAV1
P\$BPC1_Q2	251	(+)	1	1	AGAAAa	BPC1
P\$DOF1_01	267	(+)	1	0.97	tttTAAAGtga	Dof1
P\$ATHB5_01	323	(+)	1	0.93	ccaTTATTg	ATHB-5
P\$PBF_01	381	(+)	1	0.97	cagAAAAGtac	PBF
P\$DOF_Q2	381	(+)	1	0.96	cagAAAAGtac	Dof
P\$BPC1_Q2	382	(+)	1	1	AGAAAa	BPC1
P\$C1_Q2	451	(+)	1	0.95	tcAACTaaacg	C1
P\$ATHB1_01	466	(+)	1	0.95	actaaATTATtact	Athb-1
P\$ATHB5_01	469	(+)	1	0.91	aaaTTATTa	ATHB-5
P\$ATHB1_01	555	(+)	1	0.93	actaaATTATttc	Athb-1
P\$ATHB5_01	558	(+)	1	0.92	aaaTTATTi	ATHB-5
P\$DOF1_01	576	(+)	1	0.97	tgtTAAAGgct	Dof1
P\$PBF_Q2	579	(+)	1	0.99	tAAAGG	PBF
P\$DOF1_01	584	(+)	1	0.98	gctTAAAGcct	Dof1
P\$DOF2_01	584	(+)	1	0.99	gctTAAAGCtt	Dof2
P\$DOF3_01	584	(+)	1	0.99	gctTAAAGCtt	Dof3
P\$LIM1_01	596	(+)	1	0.93	CCACCAaatgtt	LIM1
P\$AGP1_01	685	(+)	1	0.93	atAGATCcaa	AGP1
P\$ARR10_01	687	(+)	1	0.93	AGATCca	ARR10
P\$PEND_01	710	(+)	1	0.92	tAAGAAGA	PEND
P\$PEND_01	713	(+)	1	0.9	gAAGAAGt	PEND
P\$MYBAS1_01	782	(+)	1	0.96	tcCTAACTcca	MYBAS1
P\$MYBAS1_01	797	(+)	1	0.98	tcCCAACtctc	MYBAS1
P\$RAV1_01	885	(+)	1	0.95	atCAACAtatg	RAV1
P\$TEIL_01	894	(+)	1	0.92	ATGTAtat	TEIL
P\$RAV1_01	915	(+)	1	0.95	aatCAACAtata	RAV1

Scanning	sequence	ID:	icroma_miR172b			
matrix identifier	position	(strand)	core match	matrix match	sequence	factor name
P\$RAV1_01	3	(+)	1	0.95	catCAACAaaag	RAV1
P\$PBF_01	7	(+)	1	0.98	aacAAAAGaaa	PBF
P\$DOF_Q2	7	(+)	1	0.94	aacAAAAGaaa	Dof
P\$BPC1_Q2	13	(+)	1	0.99	AGAAAt	BPC1
P\$SHAHB4_01	53	(+)	1	0.94	tAATGAtat	HAHB-4
P\$DOF2_01	93	(+)	1	0.99	tataAAAGCaa	Dof2
P\$DOF3_01	93	(+)	1	0.99	tataAAAGCaa	Dof3
P\$PBF_01	93	(+)	1	0.98	tataAAAGcaa	PBF
P\$DOF_Q2	93	(+)	1	0.95	tataAAAGcaa	Dof
P\$RAV1_01	135	(+)	1	0.91	ggiCAACAagtt	RAV1
P\$AGP1_01	261	(+)	1	0.95	tcAGATCtita	AGP1
P\$ARR10_01	263	(+)	1	0.96	AGATCtt	ARR10
P\$SED_Q2	393	(+)	1	0.96	actcCCTTTc	SED
P\$MYBAS1_01	402	(+)	1	0.98	ctCTAACTagt	MYBAS1
P\$C1_Q2	404	(+)	1	0.95	ctAACTAgct	C1
P\$ASR1_01	422	(+)	1	1	ACCCA	ASR-1
P\$SBF1_01	459	(+)	1	0.91	atattTTAAAatt	SBF-1
P\$PBF_01	490	(+)	1	0.96	tgaAAAAGtta	PBF
P\$DOF_Q2	490	(+)	1	0.98	tgaAAAAGtta	Dof
P\$AGP1_01	536	(+)	1	0.91	acAGATCgaa	AGP1
P\$C1_Q2	549	(+)	1	0.97	ttAACTAaccg	C1
P\$MYBAS1_01	551	(+)	1	0.99	aaCTAACcoggt	MYBAS1
P\$C1_Q2	574	(+)	1	0.94	tcAACTAttat	C1
P\$O2_Q2	662	(+)	1	0.94	cttatGACGTaaa	Opaque-2
P\$PBF_01	669	(+)	1	0.95	cgtIAAAAGgac	PBF
P\$DOF_Q2	669	(+)	1	0.95	cgtIAAAAGgac	Dof
P\$PBF_Q2	672	(+)	1	1	aAAAGG	PBF
P\$KNOX3_01	687	(+)	1	0.97	aagtTGACAcgt	Knox3
P\$HBP1B_Q6	688	(+)	1	0.94	agttgacaCGTCag	HBP-1b
P\$GBP_Q6	691	(+)	1	0.92	tgaCACGTcagc	GBP
P\$EMBP1_Q2	692	(+)	1	0.93	gaCACGTcag	EmBP-1
P\$CPRF3_Q2	692	(+)	1	0.97	gaCACGTcag	CPRF-3
P\$CPRF2_Q2	692	(+)	1	0.97	gaCACGTcag	CPRF-2
P\$O2_Q2	692	(+)	1	0.95	gaCACGTcag	Opaque-2
P\$TGA1B_Q2	692	(+)	1	0.97	gacACGTcag	TGA1b
P\$TGA1A_Q2	692	(+)	1	0.99	gaCACGTcag	TGA1a
P\$HBPA1_Q6_01	692	(+)	1	0.98	gacaCGTCag	HBP-1a
P\$CPRF3_01	692	(+)	1	0.98	gaCACGTcag	CPRF-3
P\$CPRF2_01	692	(+)	1	0.97	gaCACGTcag	CPRF-2
P\$TGA1B_01	692	(+)	1	0.94	gaCACGTcag	TGA1b
P\$RITA1_01	693	(+)	1	0.96	aCACGTc	RITA-1
P\$OCSBF1_01	694	(+)	1	1	CAGT	OCSBF-1
P\$BPC1_Q2	815	(+)	1	1	AGAAAa	BPC1

Sheet1

P\$AGP1_01	822	(+)	1	0.91	agAGATCgaa	AGP1
P\$PBF_01	827	(+)	1	0.95	tcgAAAAGatt	PBF
P\$DOF_Q2	827	(+)	1	0.94	tcgAAAAGatt	Dof
P\$AGP1_01	836	(+)	1	0.99	ttAGATCtat	AGP1
P\$ARR10_01	838	(+)	1	0.91	AGATCta	ARR10
P\$BPC1_Q2	857	(+)	1	0.99	AGAAAt	BPC1
P\$BZR1_01	934	(+)	1	1	CGTGCg	BZR1
P\$DOF2_01	972	(+)	1	0.98	gaagAAAAGCtc	Dof2
P\$DOF3_01	972	(+)	1	0.99	gaagAAAAGCtc	Dof3
P\$BPC1_Q2	974	(+)	1	1	AGAAAg	BPC1

Scanning	sequence	ID:	icrorna_miR396a			sequence	factor name
matrix identifier	position	(strand)	core match	matrix match			
P\$PBF_01	18	(+)	1	0.98	taaAAAAGaaa	PBF	
P\$DOF_Q2	18	(+)	1	0.99	taaAAAAGaaa	Dof	
P\$BPC1_Q2	24	(+)	1	1	AGAAa	BPC1	
P\$C1_Q2	31	(+)	1	0.92	aaAACTAaata	C1	
P\$DOF1_01	41	(+)	1	0.98	agaTAAAGtta	Dof1	
P\$SED_Q2	48	(+)	1	1	gttaCCTTTt	SED	
P\$RAV1_02	92	(+)	1	0.93	attACCTGatac	RAV1	
P\$GT1_Q6	115	(+)	1	1	GTGAAta	GT-1	
P\$DOF1_01	117	(+)	1	0.98	gaaTAAAGtga	Dof1	
P\$PBF_01	135	(+)	1	0.96	accAAAAGaag	PBF	
P\$DOF_Q2	135	(+)	1	0.94	accAAAAGaag	Dof	
P\$ASR1_01	158	(+)	1	1	ACCCA	ASR-1	
P\$SPF1_Q2	337	(+)	1	0.93	aaATAGTacc	SPF1	
P\$MYBAS1_01	343	(+)	1	0.98	taCCAACttc	MYBAS1	
P\$BPC1_Q2	474	(+)	1	1	AGAAa	BPC1	
P\$RAV1_01	592	(+)	1	0.95	catCAACAtagt	RAV1	
P\$SPF1_Q2	597	(+)	1	0.99	acATAGTatt	SPF1	
P\$SPF1_Q2	609	(+)	1	0.94	ttATAGTatt	SPF1	
P\$KNOX3_01	622	(+)	1	0.97	tataTGACAAcg	Knox3	
P\$GAMYB_Q2	623	(+)	1	0.91	atatgACAACgat	GAMYB	
P\$KNOX3_01	657	(+)	1	0.97	cataTGACAact	Knox3	
P\$GAMYB_Q2	658	(+)	1	0.9	atatgACAACtat	GAMYB	
P\$C1_Q2	663	(+)	1	0.93	acAACTAtagt	C1	
P\$BPC1_Q2	758	(+)	1	0.99	AGAAAt	BPC1	
P\$BPC1_Q2	804	(+)	1	1	AGAAa	BPC1	
P\$DOF2_01	809	(+)	1	0.99	aaaaAAAGCca	Dof2	
P\$DOF3_01	809	(+)	1	0.99	aaaaAAAGCca	Dof3	
P\$PBF_01	809	(+)	1	0.98	aaaAAAAGcca	PBF	
P\$DOF_Q2	809	(+)	1	0.98	aaaAAAAGcca	Dof	
P\$PBF_01	842	(+)	1	0.95	lgtIAAAAgaat	PBF	
P\$DOF_Q2	842	(+)	1	0.94	lgtIAAAAgaat	Dof	
P\$GAMYB_Q2	905	(+)	1	0.91	ttctACAACttc	GAMYB	

Scanning	sequence	ID:	icrorna_miR396b			sequence	factor name
matrix identifier	position	(strand)	core match	matrix match			
P\$PBF_01	40	(+)	1	0.97	catAAAAGata	PBF	
P\$DOF_Q2	40	(+)	1	0.93	catAAAAGata	Dof	
P\$SBF1_01	48	(+)	1	0.96	atattgTTAAaatg	SBF-1	
P\$BPC1_Q2	78	(+)	1	1	AGAAa	BPC1	
P\$SED_Q2	90	(+)	1	0.97	ttacCCTTTa	SED	
P\$PBF_01	117	(+)	1	0.97	ctcAAAAGata	PBF	
P\$DOF_Q2	117	(+)	1	0.92	ctcAAAAGata	Dof	
P\$TGA1B_Q2	274	(+)	1	0.95	catACGTcAc	TGA1b	
P\$HBP1A_Q6_01	274	(+)	1	0.92	cataCGTcAc	HBP-1a	
P\$DOF1_01	303	(+)	1	0.97	gcaTAAAGcct	Dof1	
P\$DOF2_01	303	(+)	1	0.99	gcatAAAGCct	Dof2	
P\$DOF3_01	303	(+)	1	0.99	gcatAAAGCct	Dof3	
P\$MYBAS1_01	310	(+)	1	0.95	gcCTAACc:aaa	MYBAS1	
P\$BPC1_Q2	323	(+)	1	0.99	AGAAc	BPC1	
P\$C1_Q2	324	(+)	1	0.91	gaAACTAgict	C1	
P\$PIF3_01	349	(+)	1	0.91	tattcaCACGTgca:aaa	PIF3	
P\$GBP_Q6	352	(+)	1	0.94	tcaCACGTgcaa	GBP	
P\$HBP1A_Q2	353	(+)	1	0.9	cacACGTgca	HBP-1a	
P\$CPRF_Q2	353	(+)	1	0.91	caCACGTgca	CPRF-1	
P\$TAF1_Q2	353	(+)	1	0.94	cacACGTgca	TAF-1	
P\$CPRF3_Q2	353	(+)	1	0.97	caCACGTgca	CPRF-3	

Sheet1

P\$CPRF2_Q2	353	(+)	1	0.98	caCACGTgca	CPRF-2
P\$O2_Q2	353	(+)	1	0.94	caCACGTgca	Opaque-2
P\$TGA1A_Q2	353	(+)	1	0.98	caCACGTgca	TGA1a
P\$EMBP1_Q2	353	(+)	1	0.94	cacACGTGca	EmBP-1b
P\$CPRF1_Q1	353	(+)	1	0.94	caCACGTgca	CPRF-1
P\$TAF1_Q1	353	(+)	1	0.97	cacACGTGca	TAF-1
P\$CPRF3_Q1	353	(+)	1	0.97	caCACGTgca	CPRF-3
P\$CPRF2_Q1	353	(+)	1	0.98	caCACGTgca	CPRF-2
P\$RITA1_Q1	354	(+)	1	0.98	aCACGTg	RITA-1
P\$BHLH66_Q1	354	(+)	1	0.96	aCACGTgc	bHLH66
P\$OCSBF1_Q1	355	(+)	1	1	CACGT	OCSBF-1
P\$BZR1_Q1	357	(+)	1	0.92	CGTGCa	BZR1
P\$ABZ1_Q1	389	(+)	1	0.9	ctcacACGTgtggc	ABZ1
P\$GBP_Q6	390	(+)	1	0.92	tcaCACGTgtgg	GBP
P\$HBP1A_Q2	391	(+)	1	0.91	cacACGTgtg	HBP-1a
P\$CPRF_Q2	391	(+)	1	0.92	caCACGTgtg	CPRF-1
P\$TAF1_Q2	391	(+)	1	0.95	cacACGTgtg	TAF-1
P\$CPRF3_Q2	391	(+)	1	0.98	caCACGTgtg	CPRF-3
P\$CPRF2_Q2	391	(+)	1	0.99	caCACGTgtg	CPRF-2
P\$O2_Q2	391	(+)	1	0.95	caCACGTgtg	Opaque-2
P\$TGA1A_Q2	391	(+)	1	0.99	caCACGTgtg	TGA1a
P\$EMBP1_Q2	391	(+)	1	0.92	cacACGTgtg	EmBP-1b
P\$CPRF1_Q1	391	(+)	1	0.94	caCACGTgtg	CPRF-1
P\$TAF1_Q1	391	(+)	1	0.97	cacACGTgtg	TAF-1
P\$CPRF3_Q1	391	(+)	1	0.98	caCACGTgtg	CPRF-3
P\$CPRF2_Q1	391	(+)	1	0.99	caCACGTgtg	CPRF-2
P\$TRAB1_Q2	392	(+)	1	0.98	acACGTgtggc	TRAB1
P\$RITA1_Q1	392	(+)	1	0.98	aCACGTg	RITA-1
P\$BHLH66_Q1	392	(+)	1	0.92	aCACGTgt	bHLH66
P\$OCSBF1_Q1	393	(+)	1	1	CACGT	OCSBF-1
P\$ATMYB15_Q1	477	(+)	1	0.91	glacGGTAGgttg	AtMYB-15
P\$MYBAS1_Q1	580	(+)	1	0.98	gtCTAACttt	MYBAS1
P\$MYBAS1_Q1	601	(+)	1	0.98	atCTAAAtca	MYBAS1
P\$C1_Q2	603	(+)	1	0.98	ctAACTAtcta	C1
P\$SBF1_Q1	713	(+)	1	0.91	ttacaTTAAAaaa	SBF-1
P\$SPF1_Q2	734	(+)	1	0.94	gcATAGTatt	SPF1
P\$PBF_Q1	779	(+)	1	0.97	acaAAAAGaaa	PBF
P\$DOF_Q2	779	(+)	1	1	acaAAAAGaaa	Dof
P\$BPC1_Q2	785	(+)	1	1	AGAAAa	BPC1
P\$DOF1_Q1	792	(+)	1	0.97	gcaTAAAGaaa	Dof1
P\$BPC1_Q2	798	(+)	1	0.99	AGAAAc	BPC1
P\$C1_Q2	924	(+)	1	0.94	tcaACTAaat	C1
P\$RAV1_Q1	948	(+)	1	0.92	tcCACAAattc	RAV1

Scanning	sequence	ID:	microrna_miR395b			
matrix identifier	position	(strand)	core match	matrix match	sequence	factor name
P\$ARR10_Q1	11	(+)	1	0.91	AGATCat	ARR10
P\$DOF2_Q1	30	(+)	1	0.98	aaggAAAGCtt	Dof2
P\$DOF3_Q1	30	(+)	1	0.98	aaggAAAGCtt	Dof3
P\$DOF1_Q1	63	(+)	1	0.97	ggTTAAAGgaa	Dof1
P\$PBF_Q2	66	(+)	1	0.99	IAAAGG	PBF
P\$PBF_Q1	77	(+)	1	0.96	gtaAAAAGaat	PBF
P\$GT1_Q6	77	(+)	1	0.97	GTAAAAA	GT-1
P\$DOF_Q2	77	(+)	1	0.99	gtaAAAAGaat	Dof
P\$PBF_Q1	98	(+)	1	0.98	aagAAAAGicc	PBF
P\$DOF_Q2	98	(+)	1	0.95	aagAAAAGicc	Dof
P\$BPC1_Q2	99	(+)	1	1	AGAAAa	BPC1
P\$BPC1_Q2	128	(+)	1	1	AGAAAg	BPC1
P\$PBF_Q2	129	(+)	1	0.97	gAAAGG	PBF
P\$GBP_Q6	140	(+)	1	0.91	aaaCACGTtagag	GBP
P\$CPRF3_Q2	141	(+)	1	0.92	aaCACGTaga	CPRF-3
P\$CPRF2_Q2	141	(+)	1	0.95	aaCACGTaga	CPRF-2
P\$O2_Q2	141	(+)	1	0.96	aaCACGTaga	Opaque-2
P\$TGA1A_Q2	141	(+)	1	0.96	aaCACGTaga	TGA1a
P\$CPRF3_Q1	141	(+)	1	0.93	aaCACGTaga	CPRF-3
P\$CPRF2_Q1	141	(+)	1	0.95	aaCACGTaga	CPRF-2
P\$RITA1_Q1	142	(+)	1	0.97	aCACGTa	RITA-1
P\$OCSBF1_Q1	143	(+)	1	1	CACGT	OCSBF-1
P\$ASR1_Q1	182	(+)	1	1	ACCCA	ASR-1
P\$DOF2_Q1	262	(+)	1	0.99	gagaAAAGCtt	Dof2
P\$DOF3_Q1	262	(+)	1	0.99	gagaAAAGCtt	Dof3
P\$PBF_Q1	262	(+)	1	0.96	gagAAAAGcct	PBF
P\$DOF_Q2	262	(+)	1	0.94	gagAAAAGcct	Dof

Sheet1

P\$BPC1_Q2	263	(+)	1	1	AGAAAa	BPC1
P\$KNOX3_01	311	(+)	1	0.97	iatTTGACAtt	Knox3
P\$DOF1_01	334	(+)	1	0.98	gatTAAAGaaa	Dof1
P\$BPC1_Q2	340	(+)	1	0.99	AGAAAt	BPC1
P\$C1_Q2	357	(+)	1	0.95	aaAACTAgctg	C1
P\$MYBAS1_01	397	(+)	1	0.95	taCCAAClaga	MYBAS1
P\$C1_Q2	399	(+)	1	0.95	ccAACTAgata	C1
P\$PBF_01	415	(+)	1	0.95	lggAAAAAGgtg	PBF
P\$DOF_Q2	415	(+)	1	0.94	tggAAAAAGgtg	Dof
P\$PBF_Q2	418	(+)	1	1	aAAAGG	PBF
P\$DOF1_01	447	(+)	1	0.97	tgtTAAAGtcg	Dof1
P\$SED_Q2	479	(+)	1	0.97	cttaCCTTTg	SED
P\$C1_Q2	606	(+)	1	0.97	ttaACTAcctt	C1
P\$P_01	608	(+)	1	0.92	aaCTACCTt	P
P\$SED_Q2	609	(+)	1	0.94	actaCCTTTg	SED
P\$PBF_01	698	(+)	1	0.98	taaAAAAAGagt	PBF
P\$DOF_Q2	698	(+)	1	0.98	taaAAAAAGagt	Dof
P\$DOF1_01	711	(+)	1	0.98	atgTAAAGaat	Dof1
P\$GT1_Q6	713	(+)	1	0.91	GTAAAg	GT-1
P\$RAV1_01	805	(+)	1	0.94	gccCAACAttc	RAV1
P\$MYBAS1_01	805	(+)	1	0.98	gcCCAACAtt	MYBAS1
P\$CPRF3_Q2	841	(+)	1	0.9	caCACGTtct	CPRF-3
P\$CPRF2_Q2	841	(+)	1	0.93	caCACGTtct	CPRF-2
P\$O2_Q2	841	(+)	1	0.92	caCACGTtct	Opaque-2
P\$TGA1A_Q2	841	(+)	1	0.97	caCACGTtct	TGA1a
P\$CPRF3_01	841	(+)	1	0.91	caCACGTtct	CPRF-3
P\$CPRF2_01	841	(+)	1	0.93	caCACGTtct	CPRF-2
P\$RITA1_01	842	(+)	1	0.95	aCACGTt	RITA-1
P\$OCSBF1_01	843	(+)	1	1	CACGT	OCSBF-1
P\$PCF5_01	879	(+)	1	0.92	ttGGTCCcat	PCF5
P\$PBF_01	904	(+)	1	0.96	tgaAAAAAGgaa	PBF
P\$DOF_Q2	904	(+)	1	0.99	tgaAAAAAGgaa	Dof
P\$PBF_Q2	907	(+)	1	1	aAAAGG	PBF

Scanning	sequence	ID:	microna_miR395c			
matrix identifier	position	(strand)	core match	matrix match	sequence	factor name
P\$SBF1_01	101	(+)	1	0.91	tgatatTAAAAaat	SBF-1
P\$DOF1_01	142	(+)	1	0.97	gcaTAAAGtat	Dof1
P\$GT1_Q6	172	(+)	1	0.97	GTAAAAA	GT-1
P\$C1_Q2	275	(+)	1	0.93	taAACTAaatt	C1
P\$SBF1_01	297	(+)	1	0.91	tttacaTTAATatc	SBF-1
P\$SBF1_01	333	(+)	1	0.92	ttattaTTAATlat	SBF-1
P\$BPC1_Q2	354	(+)	1	1	AGAAAa	BPC1
P\$SBF1_01	359	(+)	1	0.92	attaagTTAATAaa	SBF-1
P\$GT1_Q6	412	(+)	1	1	GTAAATA	GT-1
P\$PBF_01	508	(+)	1	0.98	gctAAAAAGaga	PBF
P\$DOF_Q2	508	(+)	1	0.94	gctAAAAAGaga	Dof
P\$MYBAS1_01	541	(+)	1	0.95	tcCCAAClata	MYBAS1
P\$C1_Q2	543	(+)	1	0.98	ccAACTAtatg	C1
P\$SED_Q2	555	(+)	1	0.96	tggCCTTTa	SED
P\$GT1_Q6	624	(+)	1	0.91	GTAAACA	GT-1
P\$PBF_01	686	(+)	1	0.97	gagAAAAAGaca	PBF
P\$DOF_Q2	686	(+)	1	0.94	gagAAAAAGaca	Dof
P\$BPC1_Q2	687	(+)	1	1	AGAAAa	BPC1
P\$PCF5_01	874	(+)	1	0.92	ttGGTCCcat	PCF5
P\$PBF_01	899	(+)	1	0.96	tgaAAAAAGaag	PBF
P\$DOF_Q2	899	(+)	1	0.99	tgaAAAAAGaag	Dof
P\$KNOX3_01	971	(+)	1	0.97	atcaTGACAgag	Knox3
P\$BPC1_Q2	991	(+)	1	1	AGAAAa	BPC1

Scanning	sequence	ID:	microna_miR395f			
matrix identifier	position	(strand)	core match	matrix match	sequence	factor name
P\$BPC1_Q2	5	(+)	1	1	AGAAAa	BPC1
P\$RAV1_01	16	(+)	1	0.93	gacCAACAcgac	RAV1
P\$MYBAS1_01	16	(+)	1	0.95	gaCCAACacga	MYBAS1
P\$PBF_01	110	(+)	1	0.95	tgaAAAAAGaat	PBF
P\$DOF_Q2	110	(+)	1	0.99	tgaAAAAAGaat	Dof
P\$TAF1_Q2	166	(+)	1	0.93	tagACGTGaa	TAF-1
P\$EMBP1_02	166	(+)	1	0.93	tagACGTGaa	EmBP-1b

Sheet1

P\$TAF1_01	166	(+)	1	0.95	tagACGTGaa	TAF-1
P\$PBF_01	215	(+)	1	0.99	aagAAAAGtgc	PBF
P\$DOF_Q2	215	(+)	1	0.95	aagAAAAGtgc	Dof
P\$BPC1_Q2	216	(+)	1	1	AGAAaA	BPC1
P\$SED_Q2	235	(+)	1	0.97	tttCCTTTc	SED
P\$BPC1_Q2	261	(+)	1	1	AGAAAg	BPC1
P\$SED_Q2	365	(+)	1	0.95	tttCCTTTg	SED
P\$PBF_01	444	(+)	1	0.97	agaAAAAGaca	PBF
P\$BPC1_Q2	444	(+)	1	1	AGAAaA	BPC1
P\$DOF_Q2	444	(+)	1	0.98	agaAAAAGaca	Dof
P\$GT1_Q6	487	(+)	1	0.97	GTAAaA	GT-1
P\$MYBAS1_01	509	(+)	1	0.95	tcCTAACataa	MYBAS1
P\$C1_Q2	527	(+)	1	0.95	taAACTAattg	C1
P\$BPC1_Q2	551	(+)	1	1	AGAAaA	BPC1
P\$DOF1_01	586	(+)	1	0.98	cgITAAAGgcc	Dof1
P\$PBF_Q2	589	(+)	1	0.99	tAAAGG	PBF
P\$PBF_01	697	(+)	1	0.98	gaaAAAAGagt	PBF
P\$DOF_Q2	697	(+)	1	0.98	gaaAAAAGagt	Dof
P\$AGP1_01	712	(+)	1	0.95	taAGATCtgc	AGP1
P\$ARR10_01	714	(+)	1	0.98	AGATCtg	ARR10
P\$AGP1_01	728	(+)	1	0.94	aaAGATCttg	AGP1
P\$ARR10_01	730	(+)	1	0.96	AGATCtt	ARR10
P\$GT1_Q6	737	(+)	1	0.91	GTGAAGa	GT-1
P\$E2F_Q2	827	(+)	1	0.97	tarTTCCGca	E2F
P\$PCF5_01	870	(+)	1	0.92	atGGTCcctc	PCF5
P\$SED_Q2	875	(+)	1	0.95	ccctCCTTTa	SED
P\$PBF_01	972	(+)	1	0.98	tgcAAAAGaga	PBF
P\$DOF_Q2	972	(+)	1	0.92	tgcAAAAGaga	Dof
P\$DOF1_01	989	(+)	1	0.98	atgTAAAGtca	Dof1

Scanning	sequence	ID:	icroma_miR395a			sequence	factor name
matrix identifier	position	(strand)	core match	matrix match			
P\$BPC1_Q2	93	(+)	1	1	AGAAaA	BPC1	
P\$MYBAS1_01	142	(+)	1	0.95	tcCCAACgtaa	MYBAS1	
P\$AG_Q2	153	(+)	1	0.9	TTTCGgattaaagaaa	AG	
P\$DOF1_01	158	(+)	1	0.98	galTAAAGaaa	Dof1	
P\$BPC1_Q2	164	(+)	1	1	AGAAaA	BPC1	
P\$MYBPH3_Q2	201	(+)	1	0.92	caaaggTAGTTaa	MYB.Ph3	
P\$PBF_Q2	201	(+)	1	0.96	cAAAGG	PBF	
P\$DOF1_01	284	(+)	1	0.98	gacTAAAGaaa	Dof1	
P\$BPC1_Q2	290	(+)	1	1	AGAAAg	BPC1	
P\$PBF_Q2	331	(+)	1	0.96	cAAAGG	PBF	
P\$SED_Q2	392	(+)	1	0.98	ttaCCTTTt	SED	
P\$RAV1_01	534	(+)	1	0.91	tatCAACAtcca	RAV1	
P\$SBF1_01	572	(+)	1	0.91	taaaaaTTAAaAaaa	SBF-1	
P\$PBF_01	587	(+)	1	0.96	attAAAAGatt	PBF	
P\$DOF_Q2	587	(+)	1	0.93	attAAAAGatt	Dof	
P\$KNOX3_Q1	593	(+)	1	0.96	agatTGACAagc	Knox3	
P\$BPC1_Q2	619	(+)	1	0.99	AGAAAI	BPC1	
P\$DOF2_01	638	(+)	1	0.99	actaAAAGCat	Dof2	
P\$DOF3_01	638	(+)	1	0.99	actaAAAGCat	Dof3	
P\$PBF_01	638	(+)	1	0.96	actAAAAGcat	PBF	
P\$DOF_Q2	638	(+)	1	0.95	actAAAAGcat	Dof	
P\$ABZ1_01	667	(+)	1	0.91	tcttACGTGtitt	ABZ1	
P\$TAF1_Q2	669	(+)	1	0.93	tctACGTGtt	TAF-1	
P\$TAF1_01	669	(+)	1	0.95	tctACGTGtt	TAF-1	
P\$SED_Q2	681	(+)	1	0.96	gcatCCTTTc	SED	
P\$DOF2_01	697	(+)	1	0.99	gtgaAAAGCcg	Dof2	
P\$DOF3_01	697	(+)	1	0.99	gtgaAAAGCcg	Dof3	
P\$PBF_01	697	(+)	1	0.97	gtgAAAAGccg	PBF	
P\$GT1_Q6	697	(+)	1	0.97	GTGAaAa	GT-1	
P\$DOF_Q2	697	(+)	1	0.94	gtgAAAAGccg	Dof	
P\$RAV1_01	719	(+)	1	0.93	cttCAACAttcg	RAV1	
P\$SED_Q2	744	(+)	1	0.98	cttCCTTTa	SED	
P\$ASR1_01	754	(+)	1	1	ACCCA	ASR-1	
P\$DOF2_01	774	(+)	1	0.99	tcaaAAAGCtt	Dof2	
P\$DOF3_01	774	(+)	1	0.98	tcaaAAAGCtt	Dof3	
P\$PBF_01	774	(+)	1	0.95	tcaAAAAGcct	PBF	
P\$DOF_Q2	774	(+)	1	0.97	tcaAAAAGcct	Dof	
P\$SED_Q2	854	(+)	1	0.97	ttatCCTTTa	SED	
P\$C1_Q2	869	(+)	1	0.94	ttaACTAtagt	C1	
P\$DOF1_01	884	(+)	1	0.99	ataTAAAGaaa	Dof1	
P\$BPC1_Q2	890	(+)	1	0.99	AGAAAc	BPC1	

Sheet1

P\$C1_Q2	896	(+)	1	0.93	gtAACTActtt	C1
P\$SPF1_Q2	959	(+)	1	0.94	tcATAGTaca	SPF1
P\$BPC1_Q2	980	(+)	1	0.99	AGAAAc	BPC1

Scanning	sequence	ID:	icromna_miR395d			
matrix identifier	position	(strand)	core match	matrix match	sequence	factor name
P\$SBF1_01	18	(+)	1	0.93	gaatgtTTAAAaat	SBF-1
P\$DOF1_01	42	(+)	1	0.97	tgaTAAAGaag	Dof1
P\$PBF_01	77	(+)	1	0.95	cgtAAAAGgta	PBF
P\$DOF_Q2	77	(+)	1	0.93	cgtAAAAGgta	Dof
P\$PBF_Q2	80	(+)	1	1	aAAAGG	PBF
P\$PBF_Q2	111	(+)	1	0.96	cAAAGG	PBF
P\$TEIL_01	146	(+)	1	0.93	ATGTAtgt	TEIL
P\$LIM1_01	161	(+)	1	0.93	CCACCatatata	LIM1
P\$RAV1_02	275	(+)	1	0.92	tgaACCTGgtcg	RAV1
P\$PBF_01	301	(+)	1	0.98	attAAAAGaca	PBF
P\$DOF_Q2	301	(+)	1	0.94	attAAAAGaca	Dof
P\$C1_Q2	397	(+)	1	0.93	aaAACTAcatt	C1
P\$SED_Q2	476	(+)	1	0.98	ftaaCCTTTt	SED
P\$SBF1_01	482	(+)	1	0.94	ttttgtTTAAAaaa	SBF-1
P\$CDC5_01	530	(+)	1	0.97	ggcTCAGCgat	CDC5
P\$RAV1_02	582	(+)	1	0.91	tcgACCTGatga	RAV1
P\$C1_Q2	592	(+)	1	0.93	gaAACTAtata	C1
P\$C1_Q2	618	(+)	1	0.94	acAACTAatgg	C1
P\$SPF1_Q2	646	(+)	1	0.91	taATAGTiga	SPF1
P\$ASR1_01	686	(+)	1	1	ACCCA	ASR-1
P\$AGP1_01	690	(+)	1	0.94	agAGATCttg	AGP1
P\$ARR10_01	692	(+)	1	0.96	AGATCtt	ARR10
P\$TEIL_01	705	(+)	1	0.95	ATGTAcgt	TEIL
P\$GT1_Q6	742	(+)	1	0.91	GTGAAGa	GT-1
P\$PBF_01	755	(+)	1	0.96	aggAAAAGaag	PBF
P\$DOF_Q2	755	(+)	1	0.96	aggAAAAGaag	Dof
P\$GAMYB_Q2	777	(+)	1	0.91	ttgtcACAACtct	GAMYB
P\$ASR1_01	793	(+)	1	1	ACCCA	ASR-1
P\$GBP_Q6	799	(+)	1	0.9	gttCACGTcca	GBP
P\$CPRF3_Q2	800	(+)	1	0.96	ttCACGTcca	CPRF-3
P\$CPRF2_Q2	800	(+)	1	0.95	ttCACGTcca	CPRF-2
P\$O2_Q2	800	(+)	1	0.95	ttCACGTcca	Opaque-2
P\$TGA1A_Q2	800	(+)	1	0.97	ttCACGTcca	TGA1a
P\$CPRF3_01	800	(+)	1	0.97	ttCACGTcca	CPRF-3
P\$CPRF2_01	800	(+)	1	0.96	ttCACGTcca	CPRF-2
P\$RITA1_01	801	(+)	1	0.98	tCACGTc	RITA-1
P\$OCSBF1_01	802	(+)	1	1	CACGT	OCSBF-1
P\$MYBAS1_01	831	(+)	1	0.98	caCTAACgtac	MYBAS1
P\$ASR1_01	879	(+)	1	1	ACCCA	ASR-1
P\$AGP1_01	905	(+)	1	0.99	atAGATCtaa	AGP1
P\$ARR10_01	907	(+)	1	0.91	AGATCtaa	ARR10
P\$AGP1_01	912	(+)	1	0.94	taAGATCtta	AGP1
P\$ARR10_01	914	(+)	1	0.96	AGATCtt	ARR10
P\$MYBAS1_01	932	(+)	1	0.95	ttCTAACcata	MYBAS1
P\$BPC1_Q2	945	(+)	1	0.99	AGAAAt	BPC1
P\$C1_Q2	967	(+)	1	0.94	gtAACTAgcaa	C1
P\$RAV1_01	972	(+)	1	0.95	tagCAACAacac	RAV1
P\$GAMYB_Q2	972	(+)	1	0.98	tagcaCAACaca	GAMYB
P\$RAV1_01	975	(+)	1	0.96	caaCAACAcaat	RAV1

Scanning	sequence	ID:	icromna_miR395e			
matrix identifier	position	(strand)	core match	matrix match	sequence	factor name
P\$GAMYB_Q2	71	(+)	1	0.96	tagcgACAACcca	GAMYB
P\$ASR1_01	79	(+)	1	1	ACCCA	ASR-1
P\$DOF1_01	127	(+)	1	0.97	tttTAAAGgag	Dof1
P\$PBF_Q2	130	(+)	1	0.99	tAAAGG	PBF
P\$SBF1_01	231	(+)	1	0.93	taatatTTAATatt	SBF-1
P\$AGP1_01	277	(+)	1	0.94	caAGATCttt	AGP1
P\$ARR10_01	279	(+)	1	0.96	AGATCtt	ARR10
P\$AGP1_01	293	(+)	1	0.94	gcAGATCtta	AGP1
P\$ARR10_01	295	(+)	1	0.96	AGATCtt	ARR10
P\$BPC1_Q2	335	(+)	1	0.99	AGAAAc	BPC1
P\$ASR1_01	339	(+)	1	1	ACCCA	ASR-1
P\$DOF1_01	347	(+)	1	0.97	taaTAAAGtat	Dof1

Sheet1

P\$PBF_01	355	(+)	1	0.97	tatAAAAGtat	PBF
P\$DOF_Q2	355	(+)	1	0.95	tatAAAAGtat	Dof
P\$C1_Q2	386	(+)	1	0.93	aaAACTAaacg	C1
P\$BPC1_Q2	398	(+)	1	1	AGAAAa	BPC1
P\$SED_Q2	416	(+)	1	0.98	atggCCTTTa	SED
P\$BPC1_Q2	512	(+)	1	0.99	AGAAAc	BPC1
P\$GT1_Q6	532	(+)	1	0.97	GTGAAaa	GT-1
P\$C1_Q2	550	(+)	1	0.92	aaAACTAaatt	C1
P\$PBF_01	620	(+)	1	0.97	cacAAAAGaag	PBF
P\$DOF_Q2	620	(+)	1	0.94	cacAAAAGaag	Dof
P\$PBF_Q2	640	(+)	1	0.96	cAAAGG	PBF
P\$GT1_Q6	664	(+)	1	0.97	GTAAAAA	GT-1
P\$DOF2_01	665	(+)	1	0.99	taaaAAAGCaa	Dof2
P\$DOF3_01	665	(+)	1	0.99	taaaAAAGCaa	Dof3
P\$PBF_01	665	(+)	1	0.98	taaaaaAGcaa	PBF
P\$DOF_Q2	665	(+)	1	0.99	taaaaaAGcaa	Dof
P\$TEIL_01	700	(+)	1	0.92	ATGTAat	TEIL
P\$BPC1_Q2	740	(+)	1	0.99	AGAAAc	BPC1
P\$TGA1B_Q2	741	(+)	1	0.91	gaaACGTct	TGA1b
P\$DOF2_01	760	(+)	1	0.99	gcaaaaAGCtg	Dof2
P\$DOF3_01	760	(+)	1	0.99	gcaaaaAGCtg	Dof3
P\$PBF_01	760	(+)	1	0.96	gcaaaaAGCtg	PBF
P\$DOF_Q2	760	(+)	1	0.98	gcaaaaAGCtg	Dof
P\$PBF_Q2	770	(+)	1	0.97	gAAAGG	PBF
P\$PBF_01	774	(+)	1	0.95	ggaAAAAGaat	PBF
P\$DOF_Q2	774	(+)	1	0.99	ggaAAAAGaat	Dof
P\$SBF1_01	792	(+)	1	0.92	cttttTTAAaaaa	SBF-1
P\$DOF1_01	818	(+)	1	1	accTAAAGccc	Dof1
P\$DOF2_01	818	(+)	1	1	acctAAAGCcc	Dof2
P\$DOF3_01	818	(+)	1	0.99	acctAAAGCcc	Dof3
P\$RAV1_02	829	(+)	1	0.93	ttaACCTGaatt	RAV1
P\$GBP_Q6	838	(+)	1	0.92	attCACGTctaa	GBP
P\$CPRF3_Q2	839	(+)	1	0.96	ttCACGTcta	CPRF-3
P\$CPRF2_Q2	839	(+)	1	0.96	ttCACGTcta	CPRF-2
P\$O2_Q2	839	(+)	1	0.96	ttCACGTcta	Opaque-2
P\$TGA1B_Q2	839	(+)	1	0.9	ttCACGTcta	TGA1b
P\$TGA1A_Q2	839	(+)	1	0.98	ttCACGTcta	TGA1a
P\$CPRF3_01	839	(+)	1	0.98	ttCACGTcta	CPRF-3
P\$CPRF2_01	839	(+)	1	0.97	ttCACGTcta	CPRF-2
P\$RITA1_01	840	(+)	1	0.98	ICACGTc	RITA-1
P\$OCSBF1_01	841	(+)	1	1	CACGT	OCSBF-1
P\$MYBAS1_01	907	(+)	1	0.99	ccCTAACactc	MYBAS1
P\$TEIL_01	938	(+)	1	0.98	ATGTAact	TEIL

Appendix C

Result of Pattern Matching program

upstream sequence: [miR447a]

```
AACATTCITTCACATTGTCACATTAACAAACAAAAACAAAAAACAATTCITTCACATTGCAGCAAAATGTGGGATAGTTA
GATAAAGGGCGTCGCTGAGATAAAGGGATTACTAAAACTGGGCCCATTTGTGTTAAACITTTATAGGCCCAATTTGTGAACACC
ACGTTGAGITCGACTTTTTTAGTAGGCAGTTGTCCAAATATGACAAAAAATAAAATAAAATAAAATGGGCGGTTCCGGCTGAG
ACAGCTAATCGGGTATTGCGGCCGTGTGATGCCGCCGTACAATTATATCCATGGTCATGGCTCATCATTAGTCGCACTGCTCTCC
TTTTCTCAAAGTTTAAATTCGACATTTGGTAAAATGATGAAACCTCGATGGAACCTGCTCTCTTTATGGAATCACGGAATGGACAA
ATAATCAAAATCAGAAATCGAAGCGAAAAGGGAGGAGAAAAACGCAGATTGGAGGATTGGGGACAGATTAGATACTGTTGAA
TGCATCACTCTAATGCTATCAGCCTATTAATAGCGTCCTATATTTTCGAAGACTTTTAATGTTTAGGGTTATGGATTTTCGAGCG
AAGCATGGAGAGATGTTGAATTGGATACTATAGGATTTGGTACAACACATACATATGTTCTGCTTCTGCAAAAATAACATATCAA
GTTTCAGAGAAAACAGTAAGTCGTTGAATATTTTATATCCATTCACAGCTTCTCTTTTGGATCATGTCCTGTTTGTCTTGACCA
TCTTCTTCTTAAAGAGGATGGACAATATAAAAACTGGAGCCTTCTTTTCTATGAAATGCTTATCATCGCGGAGTTGATCTGTT
AATTACCTGCCAATGGATGCTTTTTTTATATATACTTCACTGTTCAATTCAGATGCTTTAGAAGGTTTCGGGAGTAGCTAGAGA
ATCTGGTATCTTCAGTCTTCAATTCAGTACTTGGTATCAGTTCGTCATTGTATATCAACA
```

```
==== motif: [19] -> CA[ATCG]ATG === AtMYC2
==== pattern found: CATATG (position = 643)
==== pattern found: CAGATG (position = 901)
==== motif: [20] -> CAC[ATCG]TG
==== pattern found: CACCTG (position = 854)
==== motif: [22] -> CACAT[ATCG]
==== pattern found: CACATT (position = 9)
==== pattern found: CACATT (position = 17)
==== pattern found: CACATT (position = 56)
==== pattern found: CACATA (position = 637)
```

```
==== motif: [52] -> CACGT[ATCG]GA === ABRE-like
==== pattern found: CACGTTGA (position = 167)
```

```
==== motif: [61] -> CCACGT[ATCG]G === CBF2
==== pattern found: CCACGTTG (position = 166)
```

```
==== motif: [74] -> T[ATCG]TCTC === ARF
==== pattern found: TTTCTC (position = 339)
==== motif: [75] -> TG[ATCG]CTC
==== pattern found: TGGCTC (position = 310)
==== motif: [77] -> TGTC[ATCG]C
==== pattern found: TGTCAC (position = 14)
==== motif: [78] -> TGTCT[ATCG]
==== pattern found: TGTCTT (position = 743)
```

```
==== motif: [92] -> ACA[ATCG]ATG === DPBF
==== pattern found: ACATATG (position = 642)
==== motif: [95] -> ACACAT[ATCG]
==== pattern found: ACACATA (position = 636)
```

```
==== motif: [139] -> [ATCG]CCGCC === GCC-box
==== pattern found: GCCGCC (position = 283)
==== motif: [140] -> G[ATCG]CGCC
==== pattern found: GCCGCC (position = 283)
==== motif: [141] -> GC[ATCG]GCC
==== pattern found: GCCGCC (position = 270)
==== pattern found: GCCGCC (position = 283)
==== motif: [142] -> GCC[ATCG]CC
==== pattern found: GCCGCC (position = 283)
==== motif: [143] -> GCCG[ATCG]C
==== pattern found: GCCGCC (position = 283)
==== motif: [144] -> GCCGC[ATCG]
==== pattern found: GCCGCC (position = 283)
```


=====

upstream sequence: [miR447b]

TTGAGCTTTGTTTTCTGATTTTCATTTGTATAACCGCAGCCTCCTTAGTCCTTACCCCTTATGTGATCCCGCAGGTTACGGGCTT
CTTATTGTTGGAGTTTTGACTAGTTCATATCTTGATTTTCTATTGCTGTTTGTAAATGAAGATGATGATGAAAAAATTACACCAC
CGGCTCATCTCACACGTGTGTAAGAGTGGGGAGTCCATGGGATCTTCAAGCGGGGTCCAAGCGGAAAAAGTCGAAAGTGCTTTA
AAGAATTGCATCGTTGAAATTCCTTGTCTTTATTTGATCACTA TAGTTTTGGCGTTTACTTTCTGCTCGTAATAAACCGCGGCTA
GTGTTAAGGTTTTTTTTTCGTTGGTTGTAATGTAAGAATTTTACTGCAGCAGAAAAAGAGAAAAACGTTTTAACAGTCAAAATGTTAG
CTAATTTGAATGTTTTGAAAGAATTTTTTAAAAATGAAACTAAATTAGAAGAGAGTTGGAAAGAAATCCCAAAAATCCCAAACT
TAGCTTTGGGTACAAAGTCTGGGGTTTTAAATGCTACAAGCCTATTATAACGACCTCTTTTTTCGAAGACACGCTGTTTAGAGTATG
GAGAGATGTTTCTTTAGTACACCCATATACATATGTTGTTCTGCAAACTATACTCTTGATTGACAATAACAACATATCACGT
TCAAGAAACCAGCAAGTCGTTGGATATTTATATCCATTGAATGCTTTCTTCTTTTGTATCATGCTCTGTTTCTTTTACCATTCCCTTC
TTACTTAGAGGATGTATATAGAAGCTGGAGCCTTCTTTGTTTCTATAAAATGCTGCTTATCATCGTGGAGTTGGTTCTGTAAACATT
TGAAAATTCGAACAGTTTACCTGCCATTGGATGCTTTGTTTCAATTCAGGTGCGTTAGAAGGTTTGCAGAGTAGCTAGAGAA
TCTCGTATCTTACITTTCTGCTACTTGGTATCAGCITTCGTACTTTATATCAACA

===== motif: [19] -> CA[ATCG]ATG === AtMYC2
===== pattern found: CAAATG (position = 419)
===== pattern found: CATATG (position = 629)
===== motif: [20] -> CAC[ATCG]TG
===== pattern found: CACGTG (position = 185)
===== pattern found: CACGTG (position = 583)
===== pattern found: CACCTG (position = 879)
===== motif: [21] -> CACA[ATCG]G
===== pattern found: CACACG (position = 183)

===== motif: [76] -> TGT[ATCG]TC === ARF
===== pattern found: TGTTTC (position = 606)
===== pattern found: TGTTTC (position = 638)
===== pattern found: TGTTTC (position = 753)
===== pattern found: TGTTTC (position = 811)
===== pattern found: TGTTTC (position = 898)
===== motif: [78] -> TGTCT[ATCG]
===== pattern found: TGTCTT (position = 283)
===== pattern found: TGTCTT (position = 748)

===== motif: [92] -> ACA[ATCG]ATG === DPBF
===== pattern found: ACATATG (position = 628)
===== motif: [93] -> ACAC[ATCG]TG
===== pattern found: ACACGTG (position = 184)
===== pattern found: ACACGTG (position = 582)

===== motif: [141] -> GC[ATCG]GCC === GCC-box
===== pattern found: GCAGCC (position = 36)
===== motif: [142] -> GCC[ATCG]CC
===== pattern found: GCCTCC (position = 39)

===== motif: [192] -> GC[ATCG]GCC === ERF1
===== pattern found: GCAGCC (position = 36)
===== motif: [193] -> GCC[ATCG]CC
===== pattern found: GCCTCC (position = 39)

===== motif: [235] -> AACTAA[ATCG] === MYB4
===== pattern found: AACTAAA (position = 467)

===== motif: [237] -> G[ATCG]TAAG === Ibox
===== pattern found: GTTAAG (position = 346)
===== motif: [238] -> GA[ATCG]AAG
===== pattern found: GAAAAG (position = 237)
===== pattern found: GAAAAG (position = 394)

===== motif: [284] -> TGTATATA[ATCG] === SORLREP3
===== pattern found: TGTATATAG (position = 786)

===== motif: [306] -> AG[ATCG]CAC === SORLIP1

=====
===== pattern found: AGACAC (position = 580)
===== motif: [308] -> AGCC[ATCG]C
===== pattern found: AGCCTC (position = 38)

=====
===== motif: [311] -> G[ATCG]GCC === SORLIP2
===== pattern found: GAGCC (position = 801)
===== motif: [312] -> GG[ATCG]CC
===== pattern found: GGTCC (position = 227)
===== motif: [313] -> GGG[ATCG]C
===== pattern found: GGGTC (position = 226)
===== motif: [314] -> GGGC[ATCG]
===== pattern found: GGGCT (position = 81)

=====
===== motif: [338] -> GAGTG[ATCG]G === SORLIP5
===== pattern found: GAGTGGG (position = 197)

=====

upstream sequence: [miR447c]

TCTTTTATGGCGTTCTGGCTACAATACAGCATAACATCGTGCTTTCACGCATCATAAGCCTTGTATTTATTATCTTATTTATTATGA
CCGGAACCCCTCCTTACCCCTTACGTGATCCCGCAGGTTTACGGCCTTCCTATTATTGGAGTTTTGACTAGTTCATATCTTGATTTTC
TATFGCTGTTTGTAAATGAAGATGATGATGAAAAACATTACACCACCGGCTCATCTCACACGTGTGGAAGAGTGGGGAGTCCAT
GGGATCTTCAAGCGGGTCCAAGCGGAAAAGTCGAAAGTGCTTTAAAGAGATTGCCATCGTTTCAAATTAGCTTTTCCTTATTTGGG
TCACTAAAGTCTTGGCATTTACTTTCTGCTTGTAAATAAACCGCGGCTACTGTAAATGTTTTTTCTTGGTGTGTTGTACAATTGTA
TAGTAATGTAATAGCAACAAAAAAGAAATCGAAGCGAAAAAGATGTAGAAAAGTCTGGATTTTAAATTGCTACTAGTCT
CGACCTCTATATTCGAGGACTTATGTATGTTTAGAGTAATGGAATTTTTGAGTGAAGCAGGAGAGATGTTGAAAATTGGATACT
ATATGGTGTATTAGACTACGCAGGAGCTGATTTGGTACACACATACATATGTTCTGCTTTTCAAATAACATATCAAGTTC
AAGAAACCAGTAAGTCGTTTATCTATTCATGCTTTCTTTTGAATCATGCCTTGTGTTTGTGTTTGAACATATCTTCTACTTAG
TGGATGGACTATATTTAAAAGCTGCAGCCCTCTTTTAGCATAGTATAACCGCTGATGTACACCTACCAGCTTGATAACTCTTTT
TCGTGGTTTCTGTTACTCGTTTCTGTTTGTACAGATACITCTTGTTCAAATTTTCAGATGCTTTAGAAGGTTTTCCGGAGTGGCTAGAG
ATCTGTTATCTGTATGAACAGCTACTTGGTATCAGCTTCGTCATTTTATCAACA

=====
===== motif: [19] -> CA[ATCG]ATG === AtMYC2
===== pattern found: CATATG (position = 647)
===== pattern found: CAGATG (position = 912)
===== motif: [20] -> CAC[ATCG]TG
===== pattern found: CACGTG (position = 232)
===== motif: [21] -> CACA[ATCG]G
===== pattern found: CACACG (position = 230)
===== motif: [22] -> CACAT[ATCG]
===== pattern found: CACATA (position = 641)

=====
===== motif: [73] -> [ATCG]GTCTC === ARF
===== pattern found: AGTCTC (position = 510)

=====
===== motif: [92] -> ACA[ATCG]ATG === DPBF
===== pattern found: ACATATG (position = 646)
===== motif: [93] -> ACAC[ATCG]TG
===== pattern found: ACACGTG (position = 231)
===== motif: [95] -> ACACAT[ATCG]
===== pattern found: ACACATA (position = 640)

=====
===== motif: [141] -> GC[ATCG]GCC === GCC-box
===== pattern found: GCAGCC (position = 797)

=====
===== motif: [192] -> GC[ATCG]GCC === ERF1
===== pattern found: GCAGCC (position = 797)

=====
===== motif: [221] -> [ATCG]AACTAAC === MYB3
===== pattern found: AACTAAC (position = 667)

=====
===== motif: [229] -> [ATCG]ACTAAC === MYB4
===== pattern found: AACTAAC (position = 668)
===== motif: [230] -> A[ATCG]CTAAC
===== pattern found: AACTAAC (position = 668)
===== motif: [231] -> AA[ATCG]TAAC

=====
pattern found: AACTAAC (position = 668)
=====
motif: [232] -> AAC[ATCG]AAC
=====
pattern found: AACTAAC (position = 668)
=====
motif: [233] -> AACT[ATCG]AC
=====
pattern found: AACTAAC (position = 668)
=====
motif: [234] -> AACTA[ATCG]C
=====
pattern found: AACTAAC (position = 668)
=====
motif: [235] -> AACTAA[ATCG]
=====
pattern found: AACTAAC (position = 668)

=====
motif: [236] -> [ATCG]ATAAG === Ibox
=====
pattern found: CATAAG (position = 52)
=====
motif: [238] -> GA[ATCG]AAG
=====
pattern found: GAAAAG (position = 284)
=====
pattern found: GAAAAG (position = 483)
=====
motif: [241] -> GATAA[ATCG]
=====
pattern found: GATAAC (position = 847)

=====
motif: [281] -> TGTAT[ATCG]TAT === SORLREP3
=====
pattern found: TGTATTTAT (position = 61)

=====
motif: [307] -> AGC[ATCG]AC === SORLIP1
=====
pattern found: AGCAAC (position = 443)
=====
pattern found: AGCTAC (position = 965)

=====
motif: [310] -> [ATCG]GGCC === SORLIP2
=====
pattern found: CGGCC (position = 127)
=====
motif: [312] -> GG[ATCG]CC
=====
pattern found: GGTCC (position = 274)
=====
motif: [313] -> GGG[ATCG]C
=====
pattern found: GGGTC (position = 273)
=====
pattern found: GGGTC (position = 341)

=====
motif: [338] -> GAGTG[ATCG]G === SORLIP5
=====
pattern found: GAGTGGG (position = 244)
=====
motif: [339] -> GAGTGA[ATCG]
=====
pattern found: GAGTGAA (position = 564)

=====
upstream sequence: [miR396b]
CGAGTTAGTTGGATCGTTTTTTGTTTGTATATGTCATGTCATAAAAGATATTGTTAAAAATGGTGGAGATTTTCTTTTAGAAAAAAT
GTATTACCCCTTTATAGAGATATCAATTAAGCCTCAAAGATAACGAAGAGTTACTTGCCTAAGTATTGGTCAATATTTGATAAGGA
TTGAGTTACAATTTCTCCTTCTTAATTTCCGATACACCGTTTTTACACGAGATAGAGCCACCGGCACATCACCGATTCAAACTTC
ACTACTCTAGCGACACATACGTCACTGAACAAACATCAAAACAATGCATAAAGCCTAACCAAATGAGAACTAGTCTTATATAC
AAGTCATTATTCACACGTGCAAAACATTAAGTTCTTCAAGCTTCTTCTCTCACACGTGTGGCTTCTCTGCTCTATTGTTTCTTGTAG
ACACATTCAGTGGCTTATCAATTAATGACCGGATAAATTTTTGGGTTAGGTACGGTAGGTTGGATTACTTGAGGCCATATATATGA
CCACTATACGGTGAGCGTCTCCATATCGAGTAATGAGGTAACACATCTAAGCTAATTATCAAGTTGTCTAACTTTTGTAAAGATT
GTATCTAACTATCTAAGCACCATAAGTAATTCACATATAAATTACACATATAAAGACCAAATTTAAATGATGAGTTAGCCCGAA
CGTATACTTTTACACTAGCATAAGAGTGGTTACATTAATAAAAAAATGGGCATAGTATTTGGACTTGATGGAACCTATTATATA
AAACTGTATAACAAAAAGAAAACGCATAAAGAAAACAAAGTCTACTTTTATCTTTTCTCAGTCCCACGGCTACTAAGCAGGTTTAC
ATATCTATCACTCACTTTGTATTATATAAATAGAGTGGTCTCTTGTAGTTTCTTCTATATGCTAGTTGCTCAACTAATATCTCTC
TCTCTATCTCTCAACAATTTCTTAAAGTTTAATTAGTTTTTCAGAAAGAAGGAGAAGATGAAG

=====
motif: [10] -> [ATCG]TAACCA === AtMYB2
=====
pattern found: CTAACCA (position = 311)
=====
motif: [11] -> C[ATCG]AACCA
=====
pattern found: CTAACCA (position = 311)
=====
motif: [12] -> CT[ATCG]AACCA
=====
pattern found: CTAACCA (position = 311)
=====
motif: [13] -> CTA[ATCG]CCA
=====
pattern found: CTAACCA (position = 311)
=====
motif: [14] -> CTAA[ATCG]CA
=====
pattern found: CTAACCA (position = 311)
=====
pattern found: CTAAGCA (position = 610)
=====
pattern found: CTAAGCA (position = 838)
=====
motif: [15] -> CTAAC[ATCG]A
=====
pattern found: CTAACCA (position = 311)
=====
pattern found: CTAACCTA (position = 602)

```

===== motif: [16] -> CTAACC[ATCG]
===== pattern found: CTAACCA (position = 311)
===== motif: [17] -> [ATCG]ACATG
===== motif: [18] -> C[ATCG]CATG

===== motif: [19] -> CA[ATCG]ATG === AtMYC2
===== pattern found: CAAATG (position = 316)
===== motif: [20] -> CAC[ATCG]TG
===== pattern found: CACGTG (position = 354)
===== pattern found: CACGTG (position = 392)
===== motif: [21] -> CACA[ATCG]G
===== pattern found: CACACG (position = 352)
===== pattern found: CACACG (position = 390)
===== motif: [22] -> CACAT[ATCG]
===== pattern found: CACATC (position = 234)
===== pattern found: CACATA (position = 271)
===== pattern found: CACATT (position = 429)
===== pattern found: CACATC (position = 555)
===== pattern found: CACATA (position = 630)
===== pattern found: CACATA (position = 643)
===== pattern found: CACATA (position = 849)

===== motif: [53] -> CACGTG[ATCG]A === ABRE-like
===== pattern found: CACGTGCA (position = 354)

===== motif: [74] -> T[ATCG]TCTC === ARF
===== pattern found: TCTCTC (position = 385)
===== pattern found: TATCTC (position = 931)
===== pattern found: TCTCTC (position = 933)
===== pattern found: TCTCTC (position = 935)
===== pattern found: TCTCTC (position = 937)
===== pattern found: TATCTC (position = 943)
===== pattern found: TCTCTC (position = 945)
===== motif: [76] -> TGT[ATCG]TC
===== pattern found: TGTTC (position = 416)
===== pattern found: TGTATC (position = 597)
===== motif: [78] -> TGTCT[ATCG]
===== pattern found: TGTCTA (position = 578)
===== motif: [93] -> ACAC[ATCG]TG
===== pattern found: ACACGTG (position = 353)
===== pattern found: ACACGTG (position = 391)
===== motif: [95] -> ACACAT[ATCG]
===== pattern found: ACACATA (position = 270)
===== pattern found: ACACATT (position = 428)
===== pattern found: ACACATC (position = 554)
===== pattern found: ACACATA (position = 642)

===== motif: [142] -> GCC[ATCG]CC === GCC-box
===== pattern found: GCCACC (position = 226)

===== motif: [161] -> A[ATCG]CGACA === LTRE
===== pattern found: AGCGACA (position = 266)
===== motif: [164] -> ACCG[ATCG]CA
===== pattern found: ACCGGCA (position = 229)

===== motif: [193] -> GCC[ATCG]CC === ERF1
===== pattern found: GCCACC (position = 226)

===== motif: [223] -> TA[ATCG]CTAAC === MYB3
===== pattern found: TATCTAAC (position = 599)
===== motif: [227] -> TAACTA[ATCG]C
===== pattern found: TAACTATC (position = 603)

===== motif: [230] -> A[ATCG]CTAAC === MYB4
===== pattern found: ATCTAAC (position = 600)
===== motif: [232] -> AAC[ATCG]AAC
===== pattern found: AACAAAC (position = 285)

```


==== motif: [14] -> CTAA[ATCG]CA === AtMYB2
==== pattern found: CTAAACA (position = 509)

==== motif: [17] -> [ATCG]ACATG === AtMYC2
==== pattern found: TACATG (position = 381)
==== pattern found: CACATG (position = 742)
==== pattern found: TACATG (position = 989)
==== motif: [18] -> C[ATCG]CATG
==== pattern found: CACATG (position = 742)
==== motif: [19] -> CA[ATCG]ATG
==== pattern found: CATATG (position = 656)
==== pattern found: CACATG (position = 742)
==== pattern found: CAAATG (position = 748)
==== motif: [20] -> CAC[ATCG]TG
==== pattern found: CACTTG (position = 109)
==== pattern found: CACATG (position = 742)
==== motif: [21] -> CACA[ATCG]G
==== pattern found: CACATG (position = 742)
==== motif: [22] -> CACAT[ATCG]
==== pattern found: CACATG (position = 742)

==== motif: [42] -> TAC[ATCG]TGGC === ABRE
==== pattern found: TACATGGC (position = 381)

==== motif: [74] -> T[ATCG]TCTC === ARF
==== pattern found: TTTCTC (position = 438)
==== pattern found: TTTCTC (position = 927)
==== motif: [76] -> TGT[ATCG]TC
==== pattern found: TGTTC (position = 274)
==== motif: [77] -> TGT[ATCG]C
==== pattern found: TGTCC (position = 868)
==== motif: [78] -> TGTCT[ATCG]
==== pattern found: TGTCTA (position = 530)

==== motif: [89] -> [ATCG]CACATG === DBPF
==== pattern found: ACACATG (position = 741)
==== motif: [90] -> A[ATCG]ACATG
==== pattern found: ATACATG (position = 380)
==== pattern found: ACACATG (position = 741)
==== motif: [91] -> AC[ATCG]CATG
==== pattern found: ACACATG (position = 741)
==== motif: [92] -> ACA[ATCG]ATG
==== pattern found: ACACATG (position = 741)
==== motif: [93] -> ACAC[ATCG]TG
==== pattern found: ACACATG (position = 741)
==== motif: [94] -> ACACA[ATCG]G
==== pattern found: ACACATG (position = 741)
==== motif: [95] -> ACACAT[ATCG]
==== pattern found: ACACATG (position = 741)

==== motif: [206] -> AT[ATCG]CAACC === MYB1
==== pattern found: ATACAACC (position = 695)

==== motif: [235] -> AACTAA[ATCG] === MYB4
==== pattern found: AACTAAA (position = 32)

==== motif: [239] -> GAT[ATCG]AG === lbox
==== pattern found: GATTAG (position = 832)
==== motif: [241] -> GATAA[ATCG]
==== pattern found: GATAAA (position = 41)

==== motif: [276] -> [ATCG]GTATATAT === SORLREP3
==== pattern found: AGTATATAT (position = 398)
==== motif: [278] -> TG[ATCG]ATATAT
==== pattern found: TGAATATAT (position = 449)

=====
motif: [309] -> AGCCA[ATCG] === SORLIP1
=====
pattern found: AGCCAA (position = 814)

=====
motif: [310] -> [ATCG]GGCC === SORLIP2
=====
pattern found: TGGCC (position = 385)

=====
upstream sequence: [miR163]
TTAATAATCCTGTTATCGCAATGACCATGGGATTCCTCAAAGCCGTCCATTTAAAGTGAAAAAGAAGACAATGATGATGGTGACGGT
GAAACAAAGTGTGGACATATCCACATAAAATTGGAAAGTTAATGGATTTCGTGTTTCATTCTAAGTTTATGTTTCGATTCTTATTA
GATAAAAGACTTTTTCTGCCGCATTTATATTTCTGTGATGGTGTGGTAAAGACGGTGGAGCAGCAGATGCTGAAGATAACGT
TCAGAAAAGTGTGTTACTTATGCATATATTGCTATTTCTTTCTTTAAACGTTGGGGCTTTGACATTTTCTGAAGGTATTTTAATT
AGTTTAAATAATTGAAGATTAGTTTAGAGCTTATCTAGGGTTTTGTGACTTAGCTCACCATTTTCATAAAATGACAATATGCATCT
AATAATTTGTATCGAAATAACATCATTTAAAAGCCTGTATATTTTTATATATTGAATATGATGTA TAATTAATGCATAAATAATA
GTAGAACCCTCTTTTATTATTTATACTTATACTTGTATCATATACTTTACATAAATAAAACAACAAATAGGTAATCAATTTTCTCGTGT
GTGGTGTAGACAGTTAGGATTTAACAAGATCAAATAAAAAAGACCTTTTCAAATCAAGCCGAGACCCACGACAACGACACACTA
CCCCAATAATGTTTACACAATCATAAAATACCAACGACCGGCCAATGCGTATCCACTAGTGAATTGATACTTTTAAAGGTTAAGA
GAAAATGAGGTTTATTTTCGTACACGTCATTTGGTGTACTGCTCGACCACATTCACATGTTTTCTGAGGTCGAGAAAACATTTTA
ACTAACACGGCACITAAAATCAACTGCAAGATTTTTTGAATGGAAGACTTATTAGTTATTACCAAATCAAAGTCTTCTGATCA
TCAAAGGAAAATTAGTATAAATAAGCATAGAGCGTCCATGGATTATCACAGTTCTCAT

=====
motif: [12] -> CT[ATCG]ACCA === AtMYB2
=====
pattern found: CTCACCA (position = 397)

=====
motif: [17] -> [ATCG]ACATG === AtMYC2
=====
pattern found: CACATG (position = 824)
=====
motif: [18] -> C[ATCG]CATG
=====
pattern found: CACATG (position = 824)
=====
motif: [19] -> CA[ATCG]ATG
=====
pattern found: CAGATG (position = 237)
=====
pattern found: CACATG (position = 824)
=====
motif: [20] -> CAC[ATCG]TG
=====
pattern found: CACATG (position = 824)
=====
motif: [21] -> CACA[ATCG]G
=====
pattern found: CACATG (position = 824)
=====
motif: [22] -> CACAT[ATCG]
=====
pattern found: CACATA (position = 106)
=====
pattern found: CACATT (position = 818)
=====
pattern found: CACATG (position = 824)

=====
motif: [73] -> [ATCG]GTCTC === ARF
=====
pattern found: TGTCTC (position = 809)
=====
motif: [74] -> T[ATCG]TCTC
=====
pattern found: TGTCTC (position = 809)
=====
motif: [75] -> TG[ATCG]CTC
=====
pattern found: TGTCTC (position = 809)
=====
motif: [76] -> TGT[ATCG]TC
=====
pattern found: TGTTC (position = 136)
=====
pattern found: TGTTC (position = 154)
=====
pattern found: TGTATC (position = 436)
=====
pattern found: TGTCTC (position = 809)
=====
motif: [77] -> TGTCT[ATCG]C
=====
pattern found: TGTCTC (position = 809)
=====
motif: [78] -> TGTCT[ATCG]
=====
pattern found: TGTCTA (position = 285)
=====
pattern found: TGTCTC (position = 809)

=====
motif: [89] -> [ATCG]CACATG === DPBF
=====
pattern found: TCACATG (position = 823)

=====
motif: [143] -> GCCG[ATCG]C === GCCbox
=====
pattern found: GCCGTC (position = 40)
=====
motif: [144] -> GCCGC[ATCG]
=====
pattern found: GCCGCA (position = 189)

=====
motif: [161] -> A[ATCG]CGACA === LTRF
=====
pattern found: AACGACA (position = 673)

=====
motif: [194] -> GCCG[ATCG]C === ERF1
=====
pattern found: GCCGTC (position = 40)
=====
motif: [195] -> GCCGC[ATCG]
=====
pattern found: GCCGCA (position = 189)

=====
motif: [221] -> [ATCG]AACTAAC === MYB3
=====
pattern found: TAACTAAC (position = 854)
=====
motif: [222] -> T[ATCG]ACTAAC
=====
pattern found: TAACTAAC (position = 854)
=====
motif: [223] -> TA[ATCG]CTAAC
=====
pattern found: TAACTAAC (position = 854)
=====
motif: [224] -> TAA[ATCG]TAAC
=====
pattern found: TAACTAAC (position = 854)
=====
motif: [225] -> TAAC[ATCG]AAC
=====
pattern found: TAACTAAC (position = 854)
=====
motif: [226] -> TAACT[ATCG]AC
=====
pattern found: TAACTAAC (position = 854)
=====
motif: [227] -> TAACTA[ATCG]C
=====
pattern found: TAACTAAC (position = 854)
=====
motif: [228] -> TAACTAA[ATCG]
=====
pattern found: TAACTAAC (position = 854)

=====
motif: [229] -> [ATCG]ACTAAC === MYB4
=====
pattern found: AACTAAC (position = 855)
=====
motif: [230] -> A[ATCG]CTAAC
=====
pattern found: AACTAAC (position = 855)
=====
motif: [231] -> AA[ATCG]TAAC
=====
pattern found: AAATAAC (position = 443)
=====
pattern found: AACTAAC (position = 855)
=====
motif: [232] -> AAC[ATCG]AAC
=====
pattern found: AACTAAC (position = 855)
=====
motif: [233] -> AACT[ATCG]AC
=====
pattern found: AACTAAC (position = 855)
=====
motif: [234] -> AACTA[ATCG]C
=====
pattern found: AACTAAC (position = 855)
=====
motif: [235] -> AACTAA[ATCG]
=====
pattern found: AACTAAC (position = 855)

=====
motif: [236] -> [ATCG]ATAAG === Ibox
=====
pattern found: AATAAG (position = 960)
=====
motif: [237] -> G[ATCG]TAAG
=====
pattern found: GTTAAG (position = 763)
=====
motif: [238] -> GA[ATCG]AAG
=====
pattern found: GAAAAG (position = 259)
=====
motif: [239] -> GAT[ATCG]AG
=====
pattern found: GATTAG (position = 360)
=====
motif: [241] -> GATAA[ATCG]
=====
pattern found: GATAAA (position = 171)
=====
pattern found: GATAAC (position = 248)

=====
motif: [277] -> T[ATCG]TATATAT === SORLREP3
=====
pattern found: TTATATAT (position = 473)
=====
motif: [278] -> TG[ATCG]ATATAT
=====
pattern found: TGCATATAT (position = 276)

=====
motif: [305] -> A[ATCG]CCAC === SORLIP1
=====
pattern found: ATCCAC (position = 103)
=====
pattern found: ACCCAC (position = 664)
=====
pattern found: ATCCAC (position = 736)

=====
motif: [310] -> [ATCG]GGCC === SORLIP2
=====
pattern found: CGGCC (position = 724)
=====
motif: [313] -> GGG[ATCG]C
=====
pattern found: GGGGC (position = 310)
=====
motif: [314] -> GGGC[ATCG]

===== pattern found: GGGCT (position = 311)

=====

upstream sequence: [miR161.1]

```
TAGGGAACAGAAAAGGGAAGATGGCTAGACATCATCAGACCTAGAAAACAAGAAAGCTAAAGAGAATTACTAGATTTTGGCAT
GTGGGIAITCGGGTCGGGTTTTTCGGATTAGAAAATTTGGACCTGTATGGTTATATTAACAATTTGGGTCGGGTTGGGTTGGG
TTTACTTTGGGTCCGGTAAATTTGGIATTAACCTATAAAATAACCAAACTATCTCGGGTTGAGTACTACTACCCAACTTTGTC
CAAATATACCCAAACATACATAAAAGTACCCATAAAACCCGAAAAATCCACAATTATAACAAGTTTTTTTTCTTAAGACAATAAT
ATAAAACAAGAAGTAAACCTTGTTTGAACTTGTTCAGAACTCAAATCATTAACAAGTCACAGATTCAAGTTATTAGTGCATA
TATGTCCATATGTATATATATGTTCCGGTAATAAAAAGTACCCACGAGTATCAGATTCCGGTCCGGTAATACCCGATACCCAC
AGTTATTGAAAACAACCCGATCGGATAATTAGCCGGAACCGAAATAAACGGATTCAAGTATTTGGGTCCGAGTCCGGTTC
GGTTCCTCGAGTTCGGGTAATATGCCCCAGGCCTAAGAAATACATCGGGATTCTCACACTCGTATGCTTACGAGGACGAGCCTT
GTTGTAGTTGCAACTTTTTCCGTTTACTTGGTAGTACACCAATTAATTAGCTTCGTGTTATAAAATTGAAACTTTATTTGGGCCG
ATGTATCCACATCTTGGCCATTACTGAATTAGGGTTTCTTTGTTTTCTTCGACCTAACTCCGCCGTTTTCTCTTCTCTTCTTCT
TCATAATTATATAACAACCTTCTCTACTTATCTCTAACTCATCTTCTCTATGAAAATTCCATTGTTCTGCGCAAGCTTTGATC
AGTACTTCTCTTTGCTTGATCTCGGTTTTTGACCAGTTATTGGCTCGATCAATGCA
```

===== motif: [10] -> [ATCG]TAACCA === AtMYB2
===== pattern found: ATAACCA (position = 207)

===== motif: [18] -> C[ATCG]CATG === AtMYC2
===== pattern found: CGCATG (position = 763)
===== motif: [19] -> CA[ATCG]ATG
===== pattern found: CATATG (position = 433)
===== motif: [22] -> CACAT[ATCG]
===== pattern found: CACATC (position = 773)

===== motif: [74] -> T[ATCG]TCTC === ARF
===== pattern found: TATCTC (position = 218)
===== pattern found: TTCTC (position = 834)
===== pattern found: TTTCTC (position = 840)
===== pattern found: TCTCTC (position = 873)
===== pattern found: TATCTC (position = 882)
===== motif: [76] -> TGT[ATCG]TC
===== pattern found: TGTATC (position = 767)
===== pattern found: TGTTC (position = 920)

===== motif: [139] -> [ATCG]CCGCC === GCCbox
===== pattern found: TCCGCC (position = 825)
===== motif: [144] -> GCCGC[ATCG]
===== pattern found: GCCGCA (position = 761)
===== pattern found: GCCGCT (position = 828)

===== motif: [163] -> ACC[ATCG]ACA === LTRE
===== pattern found: ACCACA (position = 504)
===== motif: [165] -> ACCGA[ATCG]A
===== pattern found: ACCGAAA (position = 550)

===== motif: [190] -> [ATCG]CCGCC === ERF1
===== pattern found: TCCGCC (position = 825)
===== motif: [195] -> GCCGC[ATCG]
===== pattern found: GCCGCA (position = 761)
===== pattern found: GCCGCT (position = 828)

===== motif: [206] -> AT[ATCG]CAACC === MYB1
===== pattern found: ATACAACC (position = 864)

===== motif: [224] -> TAA[ATCG]TAAC === MYB3
===== pattern found: TAAATAAC (position = 204)

===== motif: [230] -> A[ATCG]CTAAC === MYB4
===== pattern found: ACCTAAC (position = 818)
===== motif: [231] -> AA[ATCG]TAAC
===== pattern found: AAATAAC (position = 205)

=====
motif: [234] -> AACTA[ATCG]C
=====
pattern found: AACTATC (position = 215)

=====
motif: [238] -> GA[ATCG]AAG == lbox
=====
pattern found: GAAAAG (position = 9)
=====
motif: [239] -> GAT[ATCG]AG
=====
pattern found: GATCAG (position = 938)
=====
motif: [241] -> GATAA[ATCG]
=====
pattern found: GATAAT (position = 536)

=====
motif: [273] -> ATAAAA[ATCG]GT == SORLREP2
=====
pattern found: ATAAAAAGT (position = 457)
=====
motif: [275] -> ATAAAACG[ATCG]
=====
pattern found: ATAAAACGG (position = 556)

=====
motif: [276] -> {ATCG}GTATATAT == SORLREP3
=====
pattern found: TGTATATAT (position = 437)
=====
motif: [277] -> T[ATCG]TATATAT
=====
pattern found: TGTATATAT (position = 437)
=====
pattern found: TATATATAT (position = 439)
=====
motif: [278] -> TG[ATCG]ATATAT
=====
pattern found: TGCATATAT (position = 419)
=====
pattern found: TGTATATAT (position = 437)
=====
motif: [279] -> TGT[ATCG]TATAT
=====
pattern found: TGTATATAT (position = 437)
=====
motif: [280] -> TGT[ATCG]ATAT
=====
pattern found: TGTATATAT (position = 437)
=====
motif: [281] -> TGTAT[ATCG]TAT
=====
pattern found: TGTATATAT (position = 437)
=====
motif: [282] -> TGTATA[ATCG]AT
=====
pattern found: TGTATATAT (position = 437)
=====
motif: [283] -> TGTATAT[ATCG]T
=====
pattern found: TGTATATAT (position = 437)
=====
motif: [284] -> TGTATATA[ATCG]
=====
pattern found: TGTATATAT (position = 437)

=====
motif: [305] -> A[ATCG]CCAC == SORLIP1
=====
pattern found: ATCCAC (position = 300)
=====
pattern found: ACCCAC (position = 466)
=====
pattern found: ACCCAC (position = 504)
=====
pattern found: ATCCAC (position = 770)
=====
motif: [306] -> AG[ATCG]CAC
=====
pattern found: AGTCAC (position = 397)
=====
motif: [307] -> AGC[ATCG]AC
=====
motif: [308] -> AGCC[ATCG]C
=====
motif: [309] -> AGCCA[ATCG]
=====
motif: [310] -> [ATCG]GGCC
=====
pattern found: AGGCC (position = 622)
=====
pattern found: GGGCC (position = 759)
=====
pattern found: TGGCC (position = 780)
=====
motif: [311] -> G[ATCG]GCC
=====
pattern found: GAGCC (position = 672)
=====
pattern found: GGGCC (position = 759)
=====
motif: [312] -> GG[ATCG]CC
=====
pattern found: GGACC (position = 122)
=====
pattern found: GGTCC (position = 178)
=====
pattern found: GGGCC (position = 759)
=====
pattern found: GGCCC (position = 781)
=====
motif: [313] -> GGG[ATCG]C
=====
pattern found: GGGTC (position = 93)
=====
pattern found: GGGTC (position = 150)
=====
pattern found: GGGTC (position = 177)
=====
pattern found: GGGTC (position = 578)
=====
pattern found: GGGCC (position = 759)
=====
motif: [314] -> GGG[ATCG]
=====
pattern found: GGGCC (position = 759)

=====
upstream sequence: [miR414]


```

===== motif: [75] -> TG[ATCG]CTC
===== pattern found: TGACTC (position = 394)
===== pattern found: TGTCTC (position = 664)
===== motif: [76] -> TG[ATCG]TTC
===== pattern found: TGTCTC (position = 664)
===== motif: [77] -> TGTC[ATCG]C
===== pattern found: TGTCTC (position = 664)
===== motif: [78] -> TGTCT[ATCG]
===== pattern found: TGTCTT (position = 655)
===== pattern found: TGTCTC (position = 664)

===== motif: [160] -> [ATCG]CCGACA === LTRE
===== pattern found: ACCGACA (position = 602)
===== motif: [161] -> A[ATCG]CGACA
===== pattern found: ACCGACA (position = 602)
===== motif: [162] -> AC[ATCG]GACA
===== pattern found: ACCGACA (position = 602)
===== motif: [163] -> ACC[ATCG]ACA
===== pattern found: ACCGACA (position = 602)
===== motif: [164] -> ACCG[ATCG]CA
===== pattern found: ACCGACA (position = 602)
===== motif: [165] -> ACCGA[ATCG]A
===== pattern found: ACCGACA (position = 602)
===== motif: [166] -> ACCGAC[ATCG]
===== pattern found: ACCGACA (position = 602)

===== motif: [231] -> AA[ATCG]TAAC === MYB4
===== pattern found: AATTAAC (position = 323)

===== motif: [238] -> GA[ATCG]AAG === Ibox
===== pattern found: GACAAG (position = 605)
===== pattern found: GAAAAG (position = 847)
===== motif: [239] -> GAT[ATCG]AG
===== pattern found: GATTAG (position = 891)
===== pattern found: GATGAG (position = 925)
===== motif: [240] -> GATA[ATCG]G
===== pattern found: GATATG (position = 937)
===== motif: [241] -> GATAA[ATCG]
===== pattern found: GATAAA (position = 539)
===== pattern found: GATAAA (position = 776)

===== motif: [305] -> A[ATCG]CCAC === SORLIP1
===== pattern found: AACCAC (position = 172)
===== pattern found: AACCAC (position = 233)
===== pattern found: AACCAC (position = 352)
===== motif: [307] -> AGC[ATCG]AC
===== pattern found: AGCTAC (position = 193)
===== motif: [308] -> AGCC[ATCG]C
===== motif: [309] -> AGCCA[ATCG]
===== motif: [310] -> [ATCG]GGCC
===== pattern found: TGGCC (position = 508)
===== pattern found: GGGCC (position = 945)
===== motif: [311] -> G[ATCG]GCC
===== pattern found: GGGCC (position = 945)
===== motif: [312] -> GG[ATCG]CC
===== pattern found: GGACC (position = 103)
===== pattern found: GGACC (position = 266)
===== pattern found: GGGCC (position = 945)
===== motif: [313] -> GGG[ATCG]C
===== pattern found: GGGAC (position = 61)
===== pattern found: GGGCC (position = 945)
===== motif: [314] -> GGGC[ATCG]
===== pattern found: GGGCT (position = 251)
===== pattern found: GGGCT (position = 581)
===== pattern found: GGGCC (position = 945)

===== motif: [319] -> CTC[ATCG]GTGA === SORLIP3
===== pattern found: CTCACGTGA (position = 388)

```

=====
motif: [337] -> GAGT[ATCG]AG === SORLIP5
=====
pattern found: GAGTCAG (position = 439)

=====
upstream sequence: [miR852]
GAGTAGTAGGCGGAATAGTCGAAAAAATGAATTCCTTTAGAGTAATTTGAATGACATCTGAATCAGACCACTGCTTTTAGGTTTAT
TGGTAATGACGAGGGAGATATGTTTCTAAGGATAGCTATAGACGATCAATTTTLAGGTGTCTAGTAGTATACATTGTTCTTTG
TTAGCTTTGGTTCTTTTGGGCACGATGTGATGCGTTGTTGCATTTACCCCTTTGTGTGGCATTTTTGTGACATTTTACTTGAAC
TGTGTTCTTGCTATAATGAAATGGGAATTCCTAAATACTAGATGATAGAACATTTGATCAAATTTGTTGTAGCCTGATAGGT
CGTATTGTTTATCAGACAGGTCCTGTTATTTATCGCAGCCGGACTGTGTTTLAGGATGCAACTCCAGATAATTGTTAGGGATTCTT
CTGGGATGATGAGTTTCGACCTAAATGGGATTTTCATGCTTGCCAACCTCAAACGTTGGAGGAGGACACTCAGACAGGAACAAT
GATTGTTTCAGTGGAGAAAAAGGTTTCTGGACTTTTGATACTTTGGTCACTTAAATAAAACAATCCCACATTAATAAACTGTTGTCA
TTACTAATTTCTTTTGGCTTGGCAGTTTCCTTTTCTGTAGTGACCGAGAGTATATCATTGGCCGGAGAATATGGGAGTCTGGGAAG
AAGTATTATTGTGTCACAAAGGTAATGTCATGCTACAAGAAGTTGTTTGGCTTCCACAGAAACAAGTGTGTAGCATTTCCACATGT
CTGAGAAATCTAAATGTTTACGCCAGAAATCTGTAGGGGACAAAACCAAAGTTAGGTTTGAACAATTCTGAATGCTGGAGA
TAATTCATGGGACCAATTGCTCTTATGTTTCTAGTTCTATTAAGTTTCATTGCATCACTTATCAGTGAAAAAATTTGATCAAAT
AGTATCCAACATCTGTTTATCTAAACATCTATTAAGTTTATGCTGAAACAAA

=====
motif: [14] -> CTAA[ATCG]CA === AtMYB2
=====
pattern found: CTAAACA (position = 965)

=====
motif: [17] -> [ATCG]ACATG === AtMYC2
=====
pattern found: CACATG (position = 765)
=====
motif: [18] -> C[ATCG]CATG
=====
pattern found: CACATG (position = 765)
=====
motif: [19] -> CA[ATCG]ATG
=====
pattern found: CACATG (position = 765)
=====
motif: [20] -> CAC[ATCG]TG
=====
pattern found: CACATG (position = 765)
=====
motif: [21] -> CACA[ATCG]G
=====
pattern found: CACATG (position = 765)
=====
motif: [22] -> CACAT[ATCG]
=====
pattern found: CACATT (position = 580)
=====
pattern found: CACATG (position = 765)

=====
motif: [73] -> [ATCG]GTCTC === ARF
=====
pattern found: TGTCTC (position = 873)
=====
motif: [74] -> T[ATCG]TCTC
=====
pattern found: TGTCTC (position = 873)
=====
motif: [75] -> TG[ATCG]CTC
=====
pattern found: TGTCTC (position = 873)
=====
motif: [76] -> TGT[ATCG]TC
=====
pattern found: TGTTC (position = 105)
=====
pattern found: TGTTC (position = 161)
=====
pattern found: TGTGTC (position = 695)
=====
pattern found: TGTCTC (position = 873)
=====
motif: [77] -> TGT[ATCG]C
=====
pattern found: TGTAC (position = 697)
=====
pattern found: TGTCTC (position = 873)
=====
motif: [78] -> TGTCT[ATCG]
=====
pattern found: TGTCTA (position = 142)
=====
pattern found: TGTCTG (position = 769)
=====
pattern found: TGTCTG (position = 847)
=====
pattern found: TGTCTC (position = 873)

=====
motif: [89] -> [ATCG]CACATG === DPBF
=====
pattern found: CCACATG (position = 764)

=====
motif: [141] -> GC[ATCG]GCC === GCCbox
=====
pattern found: GCAGCC (position = 378)

=====
motif: [165] -> ACCGA[ATCG]JA === LTRE
=====
pattern found: ACCGAGA (position = 642)

=====
motif: [192] -> GC[ATCG]GCC === ERF1
=====
pattern found: GCAGCC (position = 378)


```

===== pattern found: TCTCTC (position = 980)
===== pattern found: TCTCTC (position = 982)
===== pattern found: TCTCTC (position = 984)
===== pattern found: TCTCTC (position = 986)
===== pattern found: TCTCTC (position = 988)
===== pattern found: TCTCTC (position = 990)
===== pattern found: TCTCTC (position = 992)
===== pattern found: TCTCTC (position = 994)
===== motif: [76] -> TGT[ATCG]TC
===== pattern found: TGTGTC (position = 20)
===== pattern found: TGT TTC (position = 353)
===== motif: [77] -> TGT[ATCG]C
===== pattern found: TGTCCC (position = 794)

===== motif: [89] -> [ATCG]CACATG === DPBF
===== pattern found: ACACATG (position = 722)
===== motif: [90] -> A[ATCG]ACATG
===== pattern found: ACACATG (position = 722)
===== motif: [91] -> AC[ATCG]CATG
===== pattern found: ACACATG (position = 722)
===== motif: [92] -> ACA[ATCG]ATG
===== pattern found: ACACATG (position = 722)
===== pattern found: ACATATG (position = 889)
===== motif: [93] -> ACAC[ATCG]TG
===== pattern found: ACACATG (position = 722)
===== motif: [94] -> ACACA[ATCG]G
===== pattern found: ACACATG (position = 722)
===== motif: [95] -> ACACAT[ATCG]
===== pattern found: ACACATG (position = 722)
===== pattern found: ACACATC (position = 973)

===== motif: [143] -> GCCG[ATCG]C === GCCbox
===== pattern found: GCCGTC (position = 1)

===== motif: [194] -> GCCG[ATCG]C === ERF1
===== pattern found: GCCGTC (position = 1)

===== motif: [235] -> AACTAA[ATCG] === MYB4
===== pattern found: AACTAAA (position = 452)

===== motif: [236] -> [ATCG]ATAAG === lbox
===== pattern found: TATAAG (position = 674)
===== pattern found: AATAAG (position = 707)
===== motif: [238] -> GA[ATCG]AAG
===== pattern found: GAAAAG (position = 382)
===== motif: [241] -> GATAA[ATCG]
===== pattern found: GATAAT (position = 64)

===== motif: [276] -> [ATCG]GTATATAT === SORLREP3
===== pattern found: TGTATATAT (position = 40)
===== pattern found: TGTATATAT (position = 894)
===== motif: [277] -> T[ATCG]TATATAT
===== pattern found: TGTATATAT (position = 40)
===== pattern found: TCTATATAT (position = 78)
===== pattern found: TATATATAT (position = 111)
===== pattern found: TATATATAT (position = 113)
===== pattern found: TATATATAT (position = 115)
===== pattern found: TATATATAT (position = 117)
===== pattern found: TATATATAT (position = 119)
===== pattern found: TATATATAT (position = 131)
===== pattern found: TATATATAT (position = 133)
===== pattern found: TATATATAT (position = 135)
===== pattern found: TATATATAT (position = 137)
===== pattern found: TATATATAT (position = 139)
===== pattern found: TATATATAT (position = 141)
===== pattern found: TATATATAT (position = 143)
===== pattern found: TATATATAT (position = 145)
===== pattern found: TATATATAT (position = 147)
===== pattern found: TATATATAT (position = 149)

```


==== motif: [20] -> CAC[ATCG]TG === AtMYC2
==== pattern found: CACTTG (position = 378)

==== motif: [73] -> [ATCG]GTCTC === ARF
==== pattern found: TGCTCT (position = 344)
==== motif: [74] -> T[ATCG]TCTC
==== pattern found: TGCTCT (position = 344)
==== pattern found: TTTCTC (position = 398)
==== pattern found: TATCTC (position = 444)
==== pattern found: TTTCTC (position = 720)
==== pattern found: TCTCTC (position = 722)
==== pattern found: TTTCTC (position = 765)
==== pattern found: TTTCTC (position = 786)
==== motif: [75] -> TG[ATCG]CTC
==== pattern found: TGCTCT (position = 344)
==== motif: [76] -> TGT[ATCG]TTC
==== pattern found: TGCTCT (position = 344)
==== motif: [77] -> TGTCT[ATCG]C
==== pattern found: TGCTCT (position = 344)
==== motif: [78] -> TGTCT[ATCG]
==== pattern found: TGTCTA (position = 115)
==== pattern found: TGCTCT (position = 344)

==== motif: [161] -> A[ATCG]CGACA === LTRE
==== pattern found: ATCGACA (position = 531)

==== motif: [221] -> [ATCG]AACTAAC === MYB3
==== pattern found: TAACTAAC (position = 549)
==== motif: [222] -> T[ATCG]AACTAAC
==== pattern found: TAACTAAC (position = 549)
==== motif: [223] -> TA[ATCG]CTAAC
==== pattern found: TAACTAAC (position = 549)
==== motif: [224] -> TAA[ATCG]TAAC
==== pattern found: TAACTAAC (position = 549)
==== motif: [225] -> TAAC[ATCG]AAC
==== pattern found: TAACTAAC (position = 549)
==== motif: [226] -> TAACT[ATCG]AC
==== pattern found: TAACTAAC (position = 549)
==== motif: [227] -> TAACTA[ATCG]C
==== pattern found: TAACTAAC (position = 549)
==== motif: [228] -> TAACTAA[ATCG]
==== pattern found: TAACTAAC (position = 549)

==== motif: [229] -> [ATCG]AACTAAC === MYB4
==== pattern found: AACTAAC (position = 550)
==== motif: [230] -> A[ATCG]CTAAC
==== pattern found: AACTAAC (position = 550)
==== motif: [231] -> AA[ATCG]TAAC
==== pattern found: AATTAAC (position = 546)
==== pattern found: AACTAAC (position = 550)
==== motif: [232] -> AAC[ATCG]AAC
==== pattern found: AACTAAC (position = 550)
==== motif: [233] -> AACT[ATCG]AC
==== pattern found: AACTAAC (position = 550)
==== motif: [234] -> AACTA[ATCG]C
==== pattern found: AACTAAC (position = 550)
==== motif: [235] -> AACTAA[ATCG]
==== pattern found: AACTAAC (position = 550)

==== motif: [236] -> [ATCG]ATAAG === lbox
==== pattern found: TATAAG (position = 683)
==== motif: [237] -> G[ATCG]TAAG
==== motif: [238] -> GA[ATCG]AAG
==== pattern found: GAAAAG (position = 828)
==== motif: [239] -> GAT[ATCG]AG
==== pattern found: GATTAG (position = 503)
==== pattern found: GATTAG (position = 833)
==== motif: [240] -> GATA[ATCG]G
==== pattern found: GATATG (position = 485)

==== motif: [263] -> TTATAC[ATCG]AGT === SORLREP1
==== pattern found: TTATACAAGT (position = 945)

==== motif: [277] -> T[ATCG]TATATAT === SORLREP3
==== pattern found: TCTATATAT (position = 117)
==== pattern found: TATATATAT (position = 119)
==== motif: [279] -> TGT[ATCG]TATAT
==== pattern found: TGTCTATAT (position = 115)
==== motif: [283] -> TGTATAT[ATCG]T
==== pattern found: TGTATATGT (position = 151)

==== motif: [310] -> [ATCG]GGCC === SORLIP2
==== pattern found: TGGCC (position = 42)
==== pattern found: GGTC (position = 314)
==== pattern found: GGACC (position = 675)

==== motif: [334] -> G[ATCG]GTGAG === SORLIP5
==== pattern found: GTGTGAG (position = 517)

====

upstream sequence: [miR395b]

CGAGACAGAGAGATCATAAGATGATGACCAAGGAAAGCTTTTGAATATAGACGCAATTTGGGTTAAAGGAAAGAGTAAAA
GAATACGAATGTTGAAGAAAAGTCCATACGGCTTTTCACATGATAGAAAGGATGCAAAAACCGTAGAGATTCTGTTTTATAATTT
ATATGCTTTTGTACCCATAACGGTTATTTCTCACTACATTCCTGCTTG'CAATCTTTTAATATTTTTTAATTTTTTAATGTTATGA
CTATGGAGAAAAGCTTATGGATGTTGATATTTTTTCATTACCAAAAATCCATTTTATTTGACATTTTGTCCCGCACTGATTAAGA
AATGATGAAAAACGAAAACCTAGCTGCCTACATATGTTGGCGCATCGTATGATATGTACCAACTAGATATGTCATGGAAAAGGTGA
AAGTCCCCTCAAATAAATGATGTTAAAGTCGATCAGCTGAAATAAATTAATGCTTACCTTTGTATGGATCGTTTATGAAAAATAT
GAAGTAGTACCACITTCITTAGTCTATACAATATTTCTACCATATTTTCCAAAAGTAAAATCTAATAATATATACTCCACATTTCT
ATATTGTTAACTACCTTTGTATGGTAAGATTAATGCTGATCATATTTTTTCTTTAATCGGAAAATACGTTGGGAATATGAT
GAGTGGCAGCAATTAATAAAGAGTCCATGTAAGAATTTTCTGCTGTTTCAAATATTTCAAGCAAATTAATCGAAGGAATATA
GAATGAATATTTGGTCAAATAATATCTAAGGACAAGGCCCAACATTTGACTCGATTCATCTTTCAAATTTCCACACGTTCTTTAG
TTTTATGATTCGAATATGATTATTTGGTCCCATTATGAAAATTTCTCATTGAAAAAGGAAGAAGTGGAGACCCCTTAGCACTGATA
TATAAATAGGTATGAAGTTATTGTTTATATCAGGACAGAGCTACAAGAAGAGAGTCAA

==== motif: [17] -> [ATCG]ACATG === AtMYC2
==== pattern found: CACATG (position = 119)
==== motif: [18] -> C[ATCG]CATG
==== pattern found: CACATG (position = 119)
==== motif: [19] -> CA[ATCG]ATG
==== pattern found: CACATG (position = 119)
==== pattern found: CATATG (position = 371)
==== motif: [20] -> CAC[ATCG]TG
==== pattern found: CACATG (position = 119)
==== motif: [21] -> CACA[ATCG]G
==== pattern found: CACATG (position = 119)
==== pattern found: CACACG (position = 840)
==== motif: [22] -> CACAT[ATCG]
==== pattern found: CACATG (position = 119)
==== pattern found: CACATT (position = 590)

==== motif: [41] -> TA[ATCG]GTGGC === ABRE
==== pattern found: TATGTGGC (position = 373)

==== motif: [52] -> CACGT[ATCG]GA === ABRE-like
==== pattern found: CACGTAGA (position = 142)

==== motif: [64] -> A[ATCG]AATATTCT === HSEs
==== pattern found: ACAATATTCT (position = 538)
==== motif: [68] -> AGAAT[ATCG]TTCT
==== pattern found: AGAATTTTCT (position = 716)

==== motif: [74] -> T[ATCG]TCTC === ARF
==== pattern found: TTCTC (position = 195)
==== motif: [77] -> TGT[ATCG]C

```

===== pattern found: TGTCCC (position = 322)

===== motif: [79] -> [ATCG]ACACGTAGA === ACE
===== pattern found: AACACGTAGA (position = 140)

===== motif: [89] -> [ATCG]CACATG === DPBF
===== pattern found: TCACATG (position = 118)
===== motif: [92] -> ACA[ATCG]ATG
===== pattern found: ACATATG (position = 370)

===== motif: [141] -> GC[ATCG]GCC === GCCbox
===== pattern found: GCTGCC (position = 363)

===== motif: [192] -> GC[ATCG]GCC === ERF1
===== pattern found: GCTGCC (position = 363)

===== motif: [227] -> TAACTA[ATCG]C === MYB3
===== pattern found: TAACTACC (position = 606)

===== motif: [234] -> AACTA[ATCG]C === MYB4
===== pattern found: AACTAGC (position = 358)
===== pattern found: AACTACC (position = 607)
===== motif: [235] -> AACTAA[ATCG]

===== motif: [236] -> [ATCG]ATAAG === lbox
===== pattern found: CATAAG (position = 14)
===== motif: [237] -> G[ATCG]TAAG
===== pattern found: GGTAAG (position = 621)
===== motif: [238] -> GA[ATCG]AAG
===== pattern found: GAAAAG (position = 99)
===== pattern found: GAAAAG (position = 263)
===== pattern found: GAAAAG (position = 416)
===== pattern found: GACAAG (position = 798)
===== motif: [239] -> GAT[ATCG]AG
===== pattern found: GATCAG (position = 456)
===== pattern found: GATGAG (position = 681)
===== motif: [240] -> GATA[ATCG]G
===== pattern found: GATATG (position = 390)

===== motif: [279] -> TGT[ATCG]TATAT === SORLREP3
===== pattern found: TGTTTATAT (position = 963)

===== motif: [307] -> AGC[ATCG]AC === SORLIP1
===== pattern found: AGCTAC (position = 980)

===== motif: [310] -> [ATCG]GGCC === SORLIP2
===== pattern found: AGGCC (position = 802)
===== motif: [312] -> GG[ATCG]CC
===== pattern found: GGCCC (position = 803)
===== pattern found: GGTCC (position = 880)

```

```

=====
upstream sequence: [miR395f]
CTCTAGAAAACATCCGACCAACACGACGATTGTTTGAATATATAGATATTACATTATCAATATATTGTAGATACATTTTGGCGAA
TCGGTTTATAAGAGTGTAGGGTTTGAAAAAGAAATGGAGGGACTGGCGTTTATATATACAATGGAAAGCATTTGGAACITAGAC
GTGAATTCAGGTTAAGGGCTTTAGGTGAGGGAATCAATTTTTTAAGAAAAGTGCATAAAATCTTTTTCCTTTTCAGCTTTTTCGAA
GATAAGAAAAGACGTTTCTATAAACGCATAGATGCCAATGAATGTACACTCATATACATGTAAATGAACATATGATTTTGTGTTG
CTTTTTTACATGCATATCATTTTTTTTTCCTTTGTTTTTTTGTCTTTTGTGTATACACTACTGCATGCTACTTCAAATTTATGATGT
ATTATTTATAGACAGAAAAGACAATTTAGTTTTTCGGGATTATATTTTTCCACAAAAGTAAAAATATGTTTTCTATTTTCCTAACA
TAACCTTACGTAAACTAAITGATAAGACGGACAGAGAAAACCTTGTGAAAATGCAGATATATGTAAAGCGTTAAAGCCATGT
TTAGGGAAAATTTTCTCGTTTGTACTATAGTGAGATTTTTAAGATACTTTTATACITTTATTATTTGGGTTTCTAAATCAAT
TTTACATCGGAAAAAAGAGTAGAATAAGATCTGCGGGACGAAAGATCTGTGAAAGAAATGTGCTTATGTAGCACACTGATGGGAA
TATTAATATTTATCTTGGCTCCTTATTACAAAATTCGACTTGGTTCATCTCATATTTCCCGCATGTTTTGCCATTCAATGAAGA

```


=====
motif: [310] -> [ATCG]GGCC === SORLIP2
=====
pattern found: AGGCC (position = 591)
=====
motif: [312] -> GG[ATCG]CC
=====
pattern found: GGTC (position = 871)
=====
motif: [313] -> GGG[ATCG]C
=====
pattern found: GGGAC (position = 123)
=====
pattern found: GGGAC (position = 721)
=====
motif: [314] -> GGGC[ATCG]
=====
pattern found: GGGCT (position = 185)

=====
motif: [333] -> [ATCG]AGTGAG === SORLIP5
=====
pattern found: TAGTGAG (position = 632)
=====
motif: [337] -> GAGT[ATCG]AG
=====
pattern found: GAGTAA (position = 952)

=====
upstream sequence: [miR395d]
ACTTGA AAAATATATTTTCGAATGTTTAAAAATAAGTTCTGAATGATAAAGAAGTTAAGGAATCCCACCTTTTTTAATGCGTAAAAGG
TATCATATGTTGATATACTGAAGAACAAAGGGACATGTGACTCTTCCTTACTTATATGATGTCCAAATAGCCACCATATAT
ATTTTGAAGCITTTTCATTCTCTCTTTTCGTGGAAATTTAAGACAAGGGCCATCAAATTTTGATTCTTATCAAAAACGTCTGCACAT
TTTATATTACCGAAGATGAACCTGGTCGATGTTTTGGACCTAATTAAGACATTGAAATTAAGAATAACTTGTGCCATGTGTTA
AACAAAGTTGCTCGATATATTCTATGTTTCAATTAGCACACACTAGGGATTGGAAAAAACTCATTTTTAAAAATGGAGGCATTACCCC
TAATTACAAAAAACAACACAAGTTATAAACTCGGCATTAGTAATTTAACCTTTTGTTTAAAAACCGTTTGAGTGAGTTAAA
CAAGTTGTGGACTCGAGGCTCAGCGATAGAATGTGGATTACTACAAGAATATCTACTTGATGTGGTTTCGACCTGATGAACTA
TATATTAGCACTAATAATCAACAATAATGGTATGTATGCCACTCATTTAATAGTTGACTGATAATATAAACCAATCTAGATTCTGA
CCACCCAGAGATCTTGATAGATGTACGATGTTTTGTAGAAATCAAATTTTCGTCTCGTGAAGAGACAGAAGGAAAAAGAGTAC
TAATTCTCTTGTCACAACCTCTTAAACCCAAGTTCACGTCCAGGTTCAAACACTCTCCCTCACACTAACGTACATGAGATCTTTA
GTTCTCGCTATAAATTAAGATGTACCCATCTATCTTCTCAAACAAAAATAGATCTAAGATCTTATATACCTACATTTCAACCA
TACAAGAAATATTTATGTTTCATGGTGGTAACTAGCAACAACACAATAATAATTCGACTAG

=====
motif: [5] -> TACC[ATCG]ACAT === DREB1&2
=====
pattern found: TACCTACAT (position = 923)

=====
motif: [10] -> [ATCG]TAACCA === AtMYB2
=====
pattern found: CTAACCA (position = 933)
=====
motif: [11] -> C[ATCG]AACCA
=====
pattern found: CTAACCA (position = 933)
=====
motif: [12] -> CT[ATCG]AACCA
=====
pattern found: CTAACCA (position = 933)
=====
motif: [13] -> CTA[ATCG]AACCA
=====
pattern found: CTAACCA (position = 933)
=====
motif: [14] -> CTAA[ATCG]AACCA
=====
pattern found: CTAACCA (position = 933)
=====
motif: [15] -> CTAA[ATCG]AACCA
=====
pattern found: CTAACCA (position = 933)
=====
motif: [16] -> CTAACC[ATCG]AACCA
=====
pattern found: CTAACCA (position = 933)

=====
motif: [17] -> [ATCG]ACATG === AtMYC2
=====
pattern found: GACATG (position = 116)
=====
pattern found: TACATG (position = 838)
=====
motif: [19] -> CA[ATCG]ATG
=====
pattern found: CATATG (position = 88)
=====
motif: [21] -> CACA[ATCG]G
=====
pattern found: CACAAG (position = 446)
=====
pattern found: CACAAG (position = 555)
=====
motif: [22] -> CACAT[ATCG]
=====
pattern found: CACATT (position = 253)

=====
motif: [73] -> [ATCG]GTCTC === ARF
=====
pattern found: TGTCTC (position = 735)
=====
motif: [74] -> T[ATCG]TCTC
=====
pattern found: TGTCTC (position = 735)
=====
motif: [75] -> TG[ATCG]CTC
=====
pattern found: TGAATC (position = 122)
=====
pattern found: TGTCTC (position = 735)

```

===== motif: [76] -> TGT[ATCG]TC
===== pattern found: TGT TTC (position = 366)
===== pattern found: TGT CTC (position = 735)
===== motif: [77] -> TGTC[ATCG]C
===== pattern found: TGTC TC (position = 735)
===== pattern found: TGTC AC (position = 777)
===== pattern found: TGTC AC (position = 874)
===== motif: [78] -> TGTC[ATCG]
===== pattern found: TGTC TC (position = 735)

===== motif: [94] -> ACACA[ATCG]G === DPBF
===== pattern found: ACACAAG (position = 445)

===== motif: [142] -> GCC[ATCG]CC === GCCbox
===== pattern found: GCCACC (position = 159)

===== motif: [163] -> ACC[ATCG]ACA === LTRE
===== pattern found: ACCTACA (position = 924)

===== motif: [193] -> GCC[ATCG]CC === ERF1
===== pattern found: GCCACC (position = 159)

===== motif: [227] -> TAACTA[ATCG]C === MYB3
===== pattern found: TAACTAGC (position = 967)

===== motif: [229] -> [ATCG]ACTAAC === MYB4
===== pattern found: CACTAAC (position = 830)
===== motif: [234] -> AACTA[ATCG]C
===== pattern found: AACTAGC (position = 968)
===== motif: [235] -> AACTAA[ATCG]
===== pattern found: AACTAAT (position = 619)

===== motif: [236] -> [ATCG]ATAAG === lbox
===== pattern found: AATAAG (position = 28)
===== motif: [237] -> G[ATCG]TAAG
===== pattern found: GTTAAG (position = 51)
===== motif: [238] -> GA[ATCG]AAG
===== pattern found: GACAAG (position = 213)
===== pattern found: GAAAAG (position = 756)
===== motif: [241] -> GATAA[ATCG]
===== pattern found: GATAAA (position = 42)
===== pattern found: GATAAT (position = 657)

===== motif: [282] -> TGTATA[ATCG]AT === SORLREP3
===== pattern found: TGTATAGAT (position = 697)

===== motif: [286] -> C[ATCG]CCTAATT ==== SORLREP4
===== pattern found: CCCCTAATT (position = 424)

===== motif: [304] -> [ATCG]GCCAC === SORLIP1
===== pattern found: AGCCAC (position = 158)
===== pattern found: TGCCAC (position = 634)
===== motif: [305] -> A[ATCG]CCAC
===== pattern found: AGCCAC (position = 158)
===== motif: [306] -> AG[ATCG]CAC
===== pattern found: AGCCAC (position = 158)
===== motif: [307] -> AGC[ATCG]AC
===== pattern found: AGCCAC (position = 158)
===== pattern found: AGCAAC (position = 972)
===== motif: [308] -> AGCC[ATCG]C
===== pattern found: AGCCAC (position = 158)
===== motif: [309] -> AGCCA[ATCG]
===== pattern found: AGCCAC (position = 158)

```

=====
motif: [310] -> [ATCG]GGCC ==== SORLIP2
pattern found: GGGCC (position = 218)
motif: [311] -> G[ATCG]GCC
pattern found: GGGCC (position = 218)
pattern found: GTGCC (position = 330)
motif: [312] -> GG[ATCG]CC
pattern found: GGGCC (position = 218)
pattern found: GGACC (position = 293)
motif: [313] -> GGG[ATCG]C
pattern found: GGGAC (position = 114)
pattern found: GGGCC (position = 218)
motif: [314] -> GGGC[ATCG]
pattern found: GGGCC (position = 218)

=====
motif: [333] -> [ATCG]AGTGAG ==== SORLIP5
pattern found: GAGTGAG (position = 501)
motif: [334] -> G[ATCG]GTGAG
pattern found: GAGTGAG (position = 501)
motif: [335] -> GA[ATCG]TGAG
pattern found: GAGTGAG (position = 501)
motif: [336] -> GAG[ATCG]GAG
pattern found: GAGTGAG (position = 501)
motif: [337] -> GAGT[ATCG]AG
pattern found: GAGTGAG (position = 501)
motif: [338] -> GAGTG[ATCG]G
pattern found: GAGTGAG (position = 501)
motif: [339] -> GAGTGA[ATCG]
pattern found: GAGTGAG (position = 501)

=====
upstream sequence: [miR395e]
CTCAAGGGGACATTTGACTTTACATGTTATGTCTCTTTTGCAAATGCTTCTCTTACTCAACTGACCTTAGCGACAACCCAAAC
TGCATATTTATAGTCTATAGATACTTGCCAATGGAATGTTTTAAAGGAGGGACCATAGATTCTATATTCCTTCATTGAATGGCAA
AACATGCGGGAAATATGAGGATGAACCAAGTCGGAATTTGGTAATAAGGAGCCAAGATAATATTTAATATTTCCCATCAGTGTG
CTACATAAGCACATTTCTCACAAGATCTTTCGTCGCCGAGATCTTAITCTACTCTTTTTCCGATGTAAAATTGATTTAGAAACCC
AATATAATAAAGTATAAAGTATCTTAAAAATCTCACTATAGTAAACTAAACGAAGAAAATTTCCCTAAACATGGCCTTTAAC
GCTTACATATACTGCAATTTTCAAACAAGGTTTTCTCTGTCGGCTTATCAATTAGTTTACGTAAGGTTATGTTAGGATGAAATAG
AAACATATTTTTACTTTGGTGA AAAATATAATCGCGAAAAC TAAATGTCTTTTTCTGTCTATAAAATAATACATCATAAATTTGAA
GTAGCATGCAGTAGTGTATACACAAAAGAAGCAAAAAAACAAGGAAAAAAAATGATATGCATGTAAAAAAGCAAAAAACAA
AATCATAATGTTTCATTTACATGTATATGAGTGTACATTTCATTGGCATCTATGCGTTTTATAGAAAACGCTTTTCTTATCTTGCAAAAAG
CTGAAAGGAAAAAGAATTTATGCACITTTCTT'AAAAAAATGATTCCTCAACCTAAAGCCCTTAACCTGAATTCAGTCTAAGTT
CACAATGCTTCCAATTGTATATATAAACGCCAGTCCCTCCAATTCTTTTTCAAACCCTAACACTCTTATAAACCGATTGCGCAAAA
TGTATCTACAATATATGATAATGTAATATCTATATATTCAAACAATCGTCGTGTTGGTCCG

=====
motif: [14] -> CTAA[ATCG]CA ==== AtMYB2
pattern found: CTAAACA (position = 409)

=====
motif: [17] -> [ATCG]ACATG ==== AtMYC2
pattern found: TACATG (position = 20)
pattern found: AACATG (position = 172)
pattern found: AACATG (position = 412)
pattern found: TACATG (position = 696)
motif: [19] -> CA[ATCG]ATG
pattern found: CAAATG (position = 40)
pattern found: CATATG (position = 684)
motif: [21] -> CACA[ATCG]G
pattern found: CACAAG (position = 274)
motif: [22] -> CACAT[ATCG]
pattern found: CACATT (position = 265)

=====
motif: [73] -> [ATCG]GTCTC === ARF
pattern found: TGTCTC (position = 29)
motif: [74] -> T[ATCG]TCTC
pattern found: TGTCTC (position = 29)
pattern found: TTTCTC (position = 458)
motif: [75] -> TG[ATCG]CTC
pattern found: TGTCTC (position = 29)
motif: [76] -> TGT[ATCG]TC
pattern found: TGTCTC (position = 29)
pattern found: TGTATC (position = 938)

```

===== motif: [77] -> TGTC[ATCG]C
===== pattern found: TGCTCT (position = 29)
===== motif: [78] -> TGTCT[ATCG]
===== pattern found: TGCTCT (position = 29)
===== pattern found: TGTCTT (position = 559)
===== pattern found: TGTCTA (position = 569)

===== motif: [90] -> A[ATCG]ACATG === DPBF
===== pattern found: AAACATG (position = 171)
===== pattern found: AAACATG (position = 411)

===== motif: [161] -> A[ATCG]CGACA === LTRE
===== pattern found: AGCGACA (position = 71)

===== motif: [233] -> AACT[ATCG]AC ==== MYB4
===== pattern found: AACTGAC (position = 61)
===== motif: [235] -> AACTAA[ATCG]
===== pattern found: AACTAAA (position = 387)
===== pattern found: AACTAAA (position = 551)

===== motif: [236] -> [ATCG]ATAAG ==== lbox
===== pattern found: AATAAG (position = 215)
===== pattern found: CATAAG (position = 259)
===== motif: [240] -> GATA[ATCG]G
===== pattern found: GATATG (position = 654)
===== motif: [241] -> GATAA[ATCG]
===== pattern found: GATAAT (position = 228)
===== pattern found: GATAAT (position = 955)

===== motif: [275] -> ATAAAACG[ATCG] ==== SORLREP2
===== pattern found: ATAAAACGC (position = 874)

===== motif: [276] -> [ATCG]GTATATAT ==== SORLREP3
===== pattern found: TGTATATAT (position = 867)
===== motif: [277] -> T[ATCG]TATATAT
===== pattern found: TGTATATAT (position = 867)
===== pattern found: TCTATATAT (position = 967)
===== motif: [278] -> TG[ATCG]ATATAT
===== pattern found: TGTATATAT (position = 867)
===== motif: [279] -> TGT[ATCG]TATAT
===== pattern found: TGTATATAT (position = 867)
===== motif: [280] -> TGTA[ATCG]ATAT
===== pattern found: TGTATATAT (position = 867)
===== motif: [281] -> TGTAT[ATCG]TAT
===== pattern found: TGTATATAT (position = 867)
===== motif: [282] -> TGTATA[ATCG]AT
===== pattern found: TGTATATAT (position = 867)
===== motif: [283] -> TGTATAT[ATCG]T
===== pattern found: TGTATATAT (position = 867)
===== motif: [284] -> TGTATATA[ATCG]
===== pattern found: TGTATATAT (position = 867)

===== motif: [307] -> AGC[ATCG]AC === SORLIP1
===== pattern found: AGCGAC (position = 71)
===== motif: [309] -> AGCCA[ATCG]
===== pattern found: AGCCAA (position = 222)

===== motif: [310] -> [ATCG]GGCC ==== SORLIP2
===== pattern found: TGGCC (position = 416)
===== motif: [311] -> G[ATCG]GCC
===== pattern found: GAGCC (position = 221)
===== motif: [312] -> GG[ATCG]CC
===== pattern found: GGACC (position = 137)
===== motif: [313] -> GGG[ATCG]C
===== pattern found: GGGAC (position = 6)
===== pattern found: GGGAC (position = 136)

```

