ACCEPTED VERSION

http://hdl.handle.net/2440/99237

Genome-Wide analysis of Alternative Splicing Events in *Hordeum vulgare:* highlighting retention of intron-based splicing and its possible function through network analysis

Bahman Panahi [a]*, Seyed Abolghasem Mohammadi[a], Reyhaneh Ebrahimi khaksefidi[b], Jalil Fallah[c], Esmaeil Ebrahimie[d,e,f,g]

[a]Department of Plant Breeding and Biotechnology, Faculty of Agriculture, University of Tabriz, Tabriz, Iran

[b]Department of Plant Breeding and Biotechnology, Faculty of Agriculture, Shahrekord University, Shahrekord, Iran

[c]Department of Biology, School of Basic Science, University of Qom, Qom, Iran

[d] School of Information Technology and Mathematical Sciences, Division of Information Technology, Engineering and the Environment, University of South Australia, Adelaide, Australia

[e] Institute of Biotechnology, Shiraz University, Shiraz, Iran

[f] Department of Genetics and Evolution, School of Biological Sciences, The University of Adelaide, Adelaide, Australia

[g] School of Biological Sciences, Faculty of Science and Engineering, Flinders University, Adelaide, Australia

* Corresponding author

Department of Plant Breeding and Biotechnology

University of Tabriz

Iran

Emails: panahibahman@ymail.com, panahibahman@tabrizu.ac.ir, Tel: +989147029760

Abstract

In this study, using homology mapping of assembled expressed sequence tags against the genomic data, we identified alternative splicing events in barley. Results demonstrated that intron retention is frequently associated with specific abiotic stresses. Network analysis resulted in discovery of some specific sub networks between miRNAs and transcription factors in genes with high number of alternative splicing, such as cross talk between *SPL2*, *SPL10* and *SPL11* regulated by *miR156* and *miR157* families. To confirm the alternative splicing events, elongation factor protein (MLOC_3412) was selected followed by experimental verification of the predicted splice variants by qRT-PCR. Our novel integrative approach opens a new avenue for functional annotation of alternative splicing through regulatory-based network discovery.

Keyword: Alternative splicing, Functional diversity, *Hordeum vulgare*, Network discovery, qRT-PCR.

Abbreviations: AS: Alternative Splicing, GO: Gene Ontology, PUT: Plant Unique Transcripts.

Introduction

The presence of introns within a protein-coding gene can generate more than one mRNA isoform through an event called as Alternative Splicing (AS). This process increases the transcriptome plasticity and proteome diversity without increasing the gene content [1]. In some cases, alternatively spliced pre-mRNAs may yield thousands of splice variants [2]. To date, four main types of AS are known: exon skipping, alternative 5' splice site, alternative 3' splice sites and intron retention [3].

In potential, large number of splice isoforms can be created by these mechanisms, both singularly or in combination [4]. Change in coding regions alter characteristics of protein structure such as binding properties, intracellular localization, stability, enzymatic and signaling activities [5]. In contrast, change in 3' or 5' UTRs may affect message stability and target location. Some splice isoforms contain a Premature Termination Codon (PTC). These ones are non-functional and often not translated, but targeted for Nonsense-Mediated Decay (NMD) which recognized mRNAs containing pre-mature termination codons for degradation. This process, termed RUST (Regulated Unproductive Splicing and Translation), act in regulation of protein expression by generating NMD-targeted isoforms [1]. The majority of plant AS events have not been functionally characterized, but evidence suggests that AS participates in vital plant functions including stress response, disease resistance, photosynthesis and flowering [1,6,7]. However, besides few observations, the knowledge on AS in plants is still limited [3]. Several reports showed the patterns of tissue-specific splicing [8]. However, the majority of these functions are still remain uncovered in *Hordeum vulgare* and verification of the AS events, may highlight some important biological phenomena involvement, specific to grass lineage.

AS has an important role in formation and regulation of protein-protein interactions (PPIs), proteins-transcription factors (TFs) and TFs-miRNA interactions [9,10] which emphasize the

3

importance of employing high-throughput approaches to identify the interaction targets of specific coding regions specified by alternative exons, and integrate this information into existing knowledge of gene and protein interaction networks [11].

Recent high-throughput based experiments on transcriptional and splicing regulation provide a pair-wise relationship between a specific regulatory factor and its targets [12,13]. However, the complex interaction between genes and environment are governing the cellular response which cannot be completely understood in the level of individual interactions. The complex interactions can be unrevealed through the intricate interplay between the different regulators and their target genes [14]. It should be noted that understanding the complex interactions between the diverse regulators in the cell is crucial for unraveling the gene regulatory network in multi cellular organisms [14]. Interestingly, integrating transcription and phosphorylation networks in different species suggested a positive correlation between the species complexity and the degree of cooperation in the network [15]. TFs and miRNAs are the major commanders of networks which have the potential role in regulation a large set of gene expression [16-18]

Genome-wide analysis using Expressed Sequence Tags (ESTs) and available genomic sequences offer new possibilities in global genome analysis and detection of AS events [16]. ESTs represent the functional transcripts expressed from active genes in the cells, tissues or conditions investigated and generally give an overview of the developmental and physiological processes that function in these conditions [19-23]. Homology based-mapping of ESTs against genomic sequences acts as a corner stone approach to discover and delineate AS variants [24]. As example, alignment of cDNAs/ESTs to their corresponding genome sequences showed that approximately 30% mRNAs of *Arabidopsis* and *Oryza* are alternatively spliced [25]. In contrast, only 6.3% of expressed genes were alternatively spliced in *Brachypodium distachyon* [26]. In *Sorghum, Arabidopsis* and *Oryza* intron

4

retention is shown to be the major AS event occurring for at least 35% of genes [18,27,28] which indicated different rates of AS events among *Brachipodium*, *Arabidopsis* and *Oryza*.

Another advance in functional genomics analysis is obtained by development of the concept Gene Ontology (GO), GO classification and GO network interaction in addition to gene network interaction [29]. This new concept can provide more comprehensive analytical approach to elucidate key genes, biological pathways and AS mechanism in barley.

As a diploid, inbreeding, temperate crop, barley (*Hordeum vulgare*) has traditionally been considered as a model for plant genetic research [30]. It has also been suggested as a model species for study of the *Triticeae* to understand the function and regulation of complex cereal crop such as wheat genomes. The genome of *H. vulgare* was recently sequenced and a total of 26,159 genes were predicted as a part of genome analysis pipeline [30]. Altogether, the availability of genomic sequence of *H. vulgare* and the availability of EST database (http://www.plantgdb.org/) [31] provide an opportunity to perform genome wide computational based analysis of AS.

In the present study, we investigated the landscape of the AS events and classified them according to function ontology. In addition, integrated network analysis of transcriptional, post transcriptional and splicing regulation was employed. Our results, for the first time, unraveled that some specific networks between TFs and miRNAs regulate splicing process in high AS genes, but these subnetworks are absent in genes with low AS events. This work reports an unprecedented evaluation of AS in *H. vulgare* as a model plant, and presents an estimate of AS extent and its complexity.

Materials and methods

Sequence data

5

To increase the accuracy of analysis and decrease the rate of false positive in identification of AS isoforms, we used tentative unique sequences (also called a Plant Unique Transcripts or PUTs) instead of individual ESTs. The PUTs sets were retrieved by assembling 531,159 ESTs and mRNA sequences (version 169a) downloaded from PlantGDB (http://www.plantgdb.org/) [31]. In brief, ESTs were analyzed for the contaminating vectors and then homo polymer tails were removed. Vmatch program [32] was used to identify contaminations and repetitive elements by comparison of the mRNA sequences to vector, bacterial and repeat databases. Cleaned EST sequences were first clustered by the PaCE program [33] and then for each clusters, clustering algorithm (CAP3) [34] is used to perform the assembly. In order to minimize such potential false negatives, the above resulted CAP3 contigs/singlets are self-clustered using the Vmatch program. These CAP3 contigs/singlets were designed as PUT. A total of 134,482 high quality tentative unique sequences were used to analyze the diversity of alternative transcripts. The genome sequence of *H. vulgare* was downloaded from Plants Ensembl (http://plants.ensembl.org/index.html).

Homology mapping and identification of alternative splicing isoforms

Homology mapping was performed by mapping the PUTs (hereafter referred as ESTs) of *H. vulgare* to the corresponding genome sequence. Identification of AS isoforms was carried out by ASFinder (http://proteomics.ysu.edu/tools/ASFinder.html/) [35]. ASFinder uses SIM4 program to map ESTs to the corresponding genome [36]. Subsequently, ASFinder identifies those ESTs that are mapped to the same genomic location but have variable exon–intron boundaries as AS isoforms. Three thresholds were used for mapping, 1) a minimum of 97% identity of aligned ESTs with the genomic sequences 2) a minimum of 80 bp of aligned length 3) 85% of EST sequence aligned to the genome. The outputs of ASFinder were

subsequently analyzed, and AS events were identified using Astalavista server (http://genome.crg.es/astalavista/) [37].

Functional impact of AS genes versus Reference Genome

GO Analysis carried out using data obtained from Ensembl through the BioMart tool (http://plants.ensembl.org/biomart). In order to perform comparison between a gene list and its genome, we used hyper-geometric distribution [38]. This comparison reveals whether a particular GO in a gene list is over-represented or under-represented. We have presented all the common GO groups between sample and its reference genome in a chart as observed protein numbers next to the expected protein number. Eventually, *K–S* test is used to compare all GO groups at once between sample and reference genome. Expected protein number of each GO group in sample i, represented by E (GOi), is mean of hyper-geometric distribution [29].

$$E(GOi) = \frac{Sample\ size}{Genome\ size} \times \text{GOi}$$

Functional impact of alternative splicing

Detection of overrepresented and underrepresented GOs provides a new avenue for unraveling of underlying differential biological pathways and regulatory networks in genes with different rate of AS. This is essential to identify important candidate genes for complex biological systems such as AS.

To achieve this goal, we performed GO analysis for AS genes. Analysis of gene sets and significance estimation is carried out using g: Profiler (http://biit.cs.ut.ee/gprofiler/index.cgi) [39]. This web server uses hyper geometric distribution for significance estimation.

Network analysis of alternative spliced genes: based on High and Low AS genes sets

7

To understand the dynamics of complex process such as AS in barley, information about functions of proteins and underlying pathways are needed. This information is scattered throughout the scientific publications. some software systems, such as Pathway Studio, can bring the relevant information together and organize them for further analysis [40].

A range of interaction sets, derived by text mining from literature, are deposited in the Pathway Studio database such as regulation, expression, promoter binding, molecular transport, protein modification, binding, molecular synthesis, chemical reaction, direct regulation, miRNA effect, protein complex, and small molecule function. Furthermore, we updated the network database Pathway Studio 10 using Medscan language programming by literature mining. Based on AS genes as input, different possible pathways to which these genes may contribute were called from the updated database of the Pathway Studio package. The gene set enrichment concept was used to distinguish and select important pathways. Gene set enrichment measures the enrichment of different networks/pathways by an imported list of genes/miRNAs and highlights the statistically enriched pathways at p≤0.05, based on various statistical tests and particularly Fisher's exact test. Gene set enrichment approach, implemented in Pathway Studio 10 software [41] was used to construct different networks which are statistically significant. Gene set enrichment measures the enrichment of different networks/pathways by the imported list of genes and highlights the statistically enriched sub networks at p=0.05[40]. To construct the interaction network with two sets of genes, high and low AS genes, RESNET Plant database of Pathway Studio software v.10 (Elsevier) was used. This database includes new aliases for genes in the model plant and the other plants including tobacco, tomato and barley[40,41]. This software assembles data through MedScan (text mining tool) and process assumed information by natural language processing [40]. The language expounds them to logical concepts and extracts functional relationships miRNAs, TFs, proteins and cellular processes [41]. RESNET database were updated using MedScan,

especially for literatures on AS genes. To predict interaction networks, the software makes different groups of genes and finds a relation between a protein and its group using algorithms such as Fisher's Exact Test [40]. The software is equipped with several layout algorithms for drawing links and visualization of the network [40]. At first, we calculated median-test for transcript numbers of all of the genes by Minitab 16. Median value (which was 2) used as a threshold for division of AS genes in two sets: genes with transcript number >2 as a high AS genes and genes with 2 transcript number as a low AS genes.

Three networks were constructed based on AS isoform number; network for high AS genes sets, network for low AS genes sets and network for interaction of high and low genes sets. To provide comprehensive information on different interactions which may happen in constructed networks some special sub networks of miRNA-TF, TF-TF, miRNA-TF-Gene were subtracted in two genes sets networks. Union selection, physical and direct interaction algorithms were used to make statistical network based on AS genes.

To evaluate and compare of obtained results of pathway studio software with other tools and algorithms, interaction of nodes (proteins) in both gene sets were reanalyzed by STRING version 10 (http://string-db.org/). This online server provides critical assessment of protein-protein interaction based on prediction and clustering algorithms, including K-means and MCL [42]. Moreover, AtmiRNET (http://AtmiRNET.itps.ncku.edu.tw/) was used to reanalyze interaction between miRNAs and TFs [43].

Prediction of miRNA for AS genes

miRNAs bind to their targets with perfect or near-perfect complementary [17]. We used this characteristic to predict potential miRNA for identified AS targets [44]. miRNAs were predicted using psRNA target online sever (http://plantgrn.noble.org/psRNATarget/?function=2) with the default parameters.

Plant Material and Salt stress treatment

9

The *H. vulgare* plants were sown in sand and watered daily with a half-strength Hoagland nutrient solution under normal conditions (16 h of light, 25 °C; 8 h of dark, 15 °C). Five weeks old seedlings were transferred to hydroponic chambers and treated in half-strength Hoagland nutrient solutions with 300 mM NaCl. Shoot and root samples were collected after 6, 12, 24, 48 and 72 after treatments and briefly rinsed in distilled water, blotted dry, immediately frozen in liquid nitrogen before stored at -80 °C.

Semi quantitative Reverse Transcription PCR (qRT-PCR)

Total RNA was extracted from root and shoot tissues using the Fast PureTM RNA Kit (Takara, Japan; Cat#9190). RNA was treated with RNase-free Dnase I (Fermentas) to remove DNA contamination. The quantity and quality of purified RNA was evaluated using a nano-drop spectrophotometer and agarose gel electrophoresis. Less than 350 ng of total RNA samples of treated and controlled shoot and root tissues were used for RT-PCR reaction. cDNA synthesis was carried out using RNA PCR Kit (AMV, Takara, Japan). The oligo dT-adaptor primer was used for cDNA synthesis at 42 °C, for 60 min.

Two primers were used for amplification of V2: F, 5-GUAACUGGAGCGAAAAUUGGGAACC-3 and R, 3-AAUCAGGAUUGCUUUGUUUUAUCUU-5

And for amplification of V 1 the following pair primers utilized:

F, 5-GUAACUGGAGCGAAAAUUGGGAACC-3

R, 3-AUUUAUGCA CCGGGCUUUCCAGGGU-5

The primers sequence of 18 s rRNA as a housekeeping gene was:

F, 5 - ATGATAACTCGACGGATCGC-3

R, 3-CTTGGATGTGGTAGCCGTTT-5

The PCR condition was: 4 min at 94 °C (one cycle), 40 sec at 94 °C, 30 sec at 54 °C, 30 sec at 72 °C (37 cycles), and 10 min at 72 °C (one cycle). The PCR product was purified followed by sequencing. 18srRNA gene used as a house-keeping gene for internal control and measures the expression level of the target gene (EF, MLOC_3412) under treated and untreated conditions.

Result and Discussion

Detection of alternative splicing events

Result showed that 51% of barley genes, serine/threonine-protein kinase schematically presented as an example of these genes undergone AS events (Figure 1). On the basis of the putative mapping, in the AS landscape 4% Exons skipping, 7% alternative donor sites, 17% alternative acceptor sites, 54% introns retention and 18% others (complex events) were observed (Figure 2). The observed high abundance of intron retention is consistent with previous reports in organism such as *Medicago truncatula* (39%), *Populus trichocarpa* (34%), *Sorghum bicolor* (41%), *Arabidopsis thaliana* (56%), *Oryza sativa* (54%), *Chlamydomonas reinhardtii* (50%), *Brachypodium distachyon* (55%) [28,45-47].

Among the different classes of alternative splicing events in barley, intron retention was the most prevalent which is associated with specific abiotic stresses [1]. Prevalence of intron retention type in *H. vulgare* supports the intron definition model, which introns are identified by the splicing machinery splicesomes. It has been stated that plant introns retention is not only derived just as a cause of incomplete splicing, but also they are maintained as potential cytoplasmic translatable transcripts [48]. As most of the intron retention events in *Arabidopsis* happened in PTC+ isoforms, it possibly means that intron retention, as well as the other nonsense-generating splicing events play a significant role in RUST-regulated gene

11

expression. The coupling of AS with NMD appears to be widespread among eukaryotes [49] that their coupling was showed in *AFC2*, encoding a highly conserved *LAMMER kinase* which phosphorylates splicing factors, establishing a complex loop in AS regulation ([50])

Recently, evidence showed that AS play a possible role in regulating the level of functional transcripts RUST in addition to generating proteome diversity [51].

Deep transcriptome sequencing showed that PTC1 is prevalent in barley [30]. PTC1 and the NMD pathway which leads to rapid degradation of PTC1 transcripts have a central role in expression regulation of numerous genes [52]. It has been estimated that up to 72% of AS events in human spliceosomal factors introduce in frame PTCs [53], while more than one-third of PTC+ splicing isoforms could be targeted for NMD [54].

Despite the low sequence homology in the retained intron, an example of such an event is conserved IntronR4 in the *CCA1/LHY* mRNAs of *Arabidopsis*, *Populus*, *Oryza* and *Brachypodium* [47,55,56]. This intron retention event resulted in accumulation of high levels of PTC+ transcripts [1]. Recent study in barley was shown that about 10% of PTC+ transcripts caused by intron retention [30]. Interestingly, most part of intron retention transcripts of *Arabidopsis* do not leave the nucleus and so avoid NMD [57,58].

Features of exons and introns

Following the mapping, we calculated the lengths of internal exons and introns. The sizes of exons were from 10 to 3241 bp. Nevertheless, the majority of exons had a size below 500 bp (90%) and 60% lied between 10 and 200 bp (Figure 3 A). The exon length is in line with the previous reports in *Arabidopsis* and *Sorghum* [18,28]. The shorter exons have high potential for AS events by exon shuffling events coupled with the exon duplication and may promote the genome and transcriptomic complexity [59]. However, in compared with exon size, the distributions of intron ranged from 13 bp to 1900 bp. Eighty four percentages of introns had a

12

size range of 100 to 500 bp (Figure 3 B). The average intron size in *Hordeum* was close to *Sorghum* [18], whereas longer than the average intron size of *Arabidopsis* [28]. These data also verified the hypothesis that AS events in barley is line with the intron definition model. Intron and exon lengths reflect the constraints imposed by splicing recognition and frequency of AS [18]. Introns flanking alternatively spliced exons are long, whereas constitutively spliced exons are flanked by short introns [60]. Also, enlarged exons lead to exon skipping, but if the flanking introns are short, the enlarged exon is included [61].

GOs of AS genes versus Reference Genome

Comparing gene lists with their expected genome-wide protein distribution can give insight into potential biological significance [29]. Figure 4 (A and B) shows an example of this capability. Here, under ''Molecular Function'' catalytic activity and binding molecular functions are significantly higher than of its expected genome distribution in the AS affected genes. It may be of functional importance in revealing the genes in various biological pathways.

Biological process (BP) GO terms are analogous to genes because they have regulatory relationships with each other that can be used to construct a directed acyclic network [38]. For this, classifying a large number of AS affected genes in a small number of BP GO classes can help to the decrease the complexity and unraveling of common and specific pathways for two sets of genes that impacted with different rate of AS.

To achieve this goal, analysis of GO based on BP were done for AS genes (Figure 4 B). As shown, metabolic and cellular processes were higher than its expected genome distribution.

Comparative molecular GO analysis for high and low AS genes revealed that they have similar pattern in the distribution of different GO based on hyper-geometric test. Only

13

differences were observed in the molecular transducer activity and structural molecular activity.

Functional Ontology of AS genes

Functional ontology analysis of AS genes in molecular function category showed that catalytic activity and binding proteins are the most prevalent ontology. About 45% of predicted proteins coded by the alternatively spliced genes had catalytic activity (Figure 5). About 24% of the proteins had binding activity including nucleotide, protein and RNA binding (Figure 5). AS genes grouped in binding activity are of particular functional importance in splicing biological pathways. It has been suggested that the binding of the proteins to *cis*-regulatory sequences in exons and introns and associated splicing regulators may regulate the loading of the splicing machinery to splice site [3]. Another example is transcription factors which are known to affect the expression of genes. Interestingly, several transcription factors such as *ANAC*, *MYB* and *WRKY* families themselves are subjected to splicing regulation. The exact mechanism by which alternative isoforms of TFs are regulated and how they affect downstream targets remains unknown [62]. Computational analysis has identified several exonic and intronic splicing enhancer sequences in *Arabidopsis*. It is suggested that these sequences together with *SR* proteins likely play important roles in regulating alternative splicing in response to various stresses [63]. Recently, it has been reported that a splice variant of *Arabidopsis Indeterminate Domain14* transcription factor regulates the function of the *IDD14α* in starch metabolism by acting as an inhibitor forming hetero dimers, which explains the role of the AS in regulation of the transcription factor activity [64]. The occurrence of AS in transcription factors intimates to another level of relationship between transcription and splicing [63]. Our studies are also consistent with the

14

proposal that various stages of genetic information transmission display extensive coupling at various levels.

GO analysis for biological process in AS genes showed that most of the alternative spliced genes are involved in metabolic process (Figure S1 A). Result of surveying of different levels of GOs in biological process showed that some classes of GOs, under biological process, are specific for high and low AS genes. For example nitrogen compound in metabolic process were found only in high AS sets while macromolecule metabolic process class were specific to low AS genes (Figure S1 B, C).

Networks revealed a cross talk between AS and miRNA

Two separate networks for high and low AS genes were built. Results showed that the most of AS affected genes haven't any connection with together in the networks. Closer inspection revealed that AS acts as a regulatory link and possibly regulates the interactions of miRNAs and environmental stress.

Interaction network of splicing machinery genes in high and low AS

Figure 6, combined network of high and low AS, highlighted the possible following interactions: 1) *U1-70K* as a major splicing element 2) interaction between U1-70K and SC-35 (ATSC-35 in the network) 3) cross-talk between miRNA biogenesis and AS 4) response to stress by *SR* family 5) regulation of SR1 by high and low AS.

U1-70K is one of the *U1 snRNP-specific* proteins implicated in regulating basic splicing and AS of pre-mRNAs [65]. Network analysis revealed that *U1-70K* is a hub protein in network of AS genes which interact with splicing factors of *SR* family proteins including *SR45*, *SR33*, *SRZP21*, *SRZ22*, *SR1* and *SC-35* in high splicing gene set of barley (Figure 6). Results

15

showed that *SR45*, *SR33*, *SR1*, *SRZ21* connects with *AFC2*, a kinase with more than two isoforms in barley. It was also previously reported in *Prunus scoparia* [66].

Golovkin and Reddy (2003) reported that *SC-35* interact directly with *U1-70K* and affects its splice site selection characteristics [67]. The presence of *SC-35* in connection with *U1-70K* and other splicing factors such as *RSZ33* and *SIP1A* in network of two gene sets (high and low splicing) indicates that *SC-35* may be an important splicing factors in the barley splicing networks. Also we showed that this protein connected directly with RNA polymerase II which regulated by *DCL1* (Figure 6). Splicing and transcription are usually linked [68] and miRNAs are post-transcriptional regulators of regulatory networks [69]. *DCL1* (*Dicer Like proteins*) which is a RNA helicase involved in miRNAs processing and biogenesis [70] and categorized in high AS gene set. Expression of *DCL1* prohibited by *miR162A*, *miR162B*, *miR168A*, *miR172A* and activated by *miR414* (Figure 6). In addition, some of genes involve in the miRNA biogenesis such as *SE* and *HEN1* were affected by splicing. *DCL1* and *HEN1* belong to the high AS gene in the barley whereas *SE* was a common in two gene sets. As a result, it probably is a linkage between miRNA processing and splicing.

Common and specific miRNAs and TFs in high and low AS genes

Network analysis exhibited that some of miRNAs and TFs were common in the two sets of genes (Table S1) for example *miR156A*, *miR156B*, *miR399A*, *miR408*, *miR414*, *miR778A*, *miR823A*, *miR827A*, *miR828A*, *miR858A*, *miR156*, *miR157* and some TFs as a *ABF1, ABF2, ABF4, ABI4, AG* and *AGF1* that highlighted its vital regulatory roles in AS process.

Interestingly, as shown in Table S1, some regulatory molecules were specific for high and low AS gene sets. For example *miR171C* and *miR169J* were specific for low and high AS genes respectively.

16

For more dissection of interaction between TFs and miRNAs that regulate AS events, we draw sub-networks between miRNAs and some TFs that specific for high AS gene set.

Sub-network of *NF-YA10* and *miR169* as a clue for crosstalk of stress and high AS events

All of *NF-YA* family members were belonged to high AS genes. They delayed flowering and reduced growth of plants for the save energy in abiotic stress and strongly can be induced by various stress condition [71]. *NF-YA2* and *NF-YA10* are present in the high AS genes and they produce more than two alternatively spliced variants regulated by *miR169* family members, in post-transcriptional manner. Interestingly all of *miR169* members targeted a high AS genes except *miR169A* which were common between networks of low and high AS gene set.

MiR169 family, except of *miR169A*, regulates *NF-YA10* (Figure7). Decreased levels of *miR169* and over expressed levels of *NF-YA2* and *NF-YA10* enhanced tolerance to various abiotic stresses and make balance in sucrose/starch and cell elongation [71]. Consequently adaptive response to adverse environmental conditions can be achieved. Altogether, it seems that *miR169* family members play important roles in AS events and adaptive process by regulation of *NF-Y10*.

Sub-network of *REV* and *miR166* as a part of auxin signaling system

GO analysis revealed that AS is in metabolic process and response to stress. As auxin signaling is one of the mechanisms of response to stress [72], sub-network was constructed to show specific sub-network in auxin signaling found solely in high AS genes. *REV* is a *homeobox-leucine* zipper protein that exhibits a gene structure with more than two alternative transcriptional variants in transcripts. This TF plays an important role in laying down leaf polarity, root and vasculature [73]. *REV* is a part of a complex regulatory system involving

17

auxin, transcription factors, miRNAs and carrier complex and is the key regulatory gene in plant morphogenesis [74]. Transcriptional variant of *REV* was regulated by *miR165* and *miR166* families. It is indicated that *REV* may be regulated in the transcriptional and post-transcriptional levels and may be play a role in linking between post-transcriptional gene silencing and production of proteins (Figure 8). *REV* increased auxin polar flowering vascular bundles [75] and it showed negative correlation with *miR165/166* [76]. This may suggest alternation expression of *REV* by *miR165/166* during stress caused induce levels of auxin in the vascular and show tolerance to stress.

SPLs as important TFs in AS system

SPL2, 6, 10, 11, *MBA10.13* (*squamosa promoter-binding-like protein 13*) and *MFB16.6* (*squamosa promoter-binding-like protein 16*) are specific for high AS gene sets which regulated by *miR156* and *miR157* family members [77]. These TFs located in the nucleus and can bind to DNA with their zinc ion binding motifs to regulate the gene expression (Figure 9). *SPL2 and MFB16.6* are involved in vegetative and reproduction phases and regulate the floral transitions [78]. *miR172B* regulated by *SPL10* [79] also, *MiR172B* was indirectly regulated by *miR156/157* family members (Figure 9). *MiR172B* regulated transcription factors (*SNZ, AP2, SMZ, K21H1.22, RAP2.7* and *TOE2*) that are involved in ethylene mediated signaling pathway and organism development [77] which they are common in high and low AS. *SPL2* and *SPL10* regulated by *DCL1* [80] so these TFs may be involve in the gene silencing. *SPL2* suppressed *YUC2* and *YUC6*; these genes involved in the auxin biosynthesis process [81] and generate more than two alternatively spliced variant. Hence, *SPLs* may regulate auxin and ethylene signaling and gene silencing in vegetative and reproductive phase in the barley.

18

Also analysis of interaction between AS genes with *K*-means and MCL algorithms in STRING server, revealed significant score (confidence score > 0.9) for each sub networks (Table S2) and results were consistent with Pathway Studio outputs (Figure S2 A, B, C, D) (Figure S3 A, B) (Table S3)

Experimental validation of alternative splicing events

To validate the accuracy of bioinformatics analysis, elongation factor family protein (*MLOC_3412*) with identified alternative splicing events were selected. The decisive factor in the choice of a sample gene was the potential modification of protein domains or protein sequence motifs by alternative splicing [82]. As a consequence of this approach, elongation factor family protein (*MLOC_3412*) selected. For amplification of the variant1 of *MLOC_3412*, one primer resides on exon 13, which is lacking in the mRNA of the variant2; the second primer was on exon 9 which was common in both splicing variants (Figure 10). For verification of the second splice variant, one primer only binds to the exon-exon junction of exons 11 and 12. Amplification of splicing variants of this gene was shown in Figure 10 A. The tissue-specific pattern was shown for both splice variants (Figure 10 B and C). As a result, the expression levels of both variants in the shoot tissue were higher than the root tissue during the times of stress (Figure 10 C). However, the relative expression of variant 2 was higher than variant 1 which revealed is; the variant2 is relatively major transcript in response to salt stress in the *H. vulgare.*

The findings of this study are mostly obtained from bioinformatics and computational biology analysis which suggest new possibilities and approaches in the study of alternative splicing. However, more experimental support and evidence are clearly needed to be undertaken in future studies.

In conclusion, widespread occurrence of AS and the range of functional gene groups supports an essential role for AS in development, physiological, metabolism, and response to environmental conditions on barley. Comparative analysis between networks of two gene sets showed that some common TFs, miRNAs, kinases have a controlling role in the networks of two genes sets. These common regulators are vital in controlling of AS systems and AS related process in barley. Interestingly, results indicated that most of the high AS genes controlled by low splicing gene such as *DCL1, LEC2, CDF1* (Figure 6, 7, 8, 9). For the first time, this study introduces the Network-based approach in study of AS. This study provides a platform for future investigates on analysis of AS in different developmental stages.

References

[1]   Filichkin, S.A., Priest, H.D., Givan, S.A., Shen, R., Bryant, D.W., Fox, S.E., Wong, W.-K. and Mockler, T.C. (2010). Genome-wide mapping of alternative splicing in *Arabidopsis thaliana*. Genome Research 20, 45-58.

[2]   Chen, W.-H., Lv, G., Lv, C., Zeng, C. and Hu, S. (2007). Systematic analysis of alternative first exons in plant genomes. BMC Plant Biology 7, 55.

[3]   Reddy, A.S.N. (2007). Alternative Splicing of Pre-Messenger RNAs in Plants in the Genomic Era. Annual Review of Plant Biology 58, 267-294.

[4]   Black, D.L. (2003). Mechanisms of alternative pre-messenger RNA splicing. Annual Review of Biochemistry 72, 291-336.

[5]   Stamm, S., Ben-Ari, S., Rafalska, I., Tang, Y., Zhang, Z., Toiber, D., Thanaraj, T. and Soreq, H. (2005). Function of alternative splicing. Gene 344, 1-20.

[6]   Quesada, V., Macknight, R., Dean, C. and Simpson, G.G. (2003). Autoregulation of FCA pre-mRNA processing controls *Arabidopsis* flowering time. EMBO J 22, 3142-3152.

[7]   Ferrier-Cana, E., Macadre, C., Sevignac, M., David, P., Langin, T. and Geffroy, V. (2005). Distinct post-transcriptional modifications result into seven alternative transcripts of the CC–NBS–LRR gene JA1tr of *Phaseolus vulgaris*. Theoretical and Applied Genetics 110, 895-905.

[8]   Kriechbaumer, V., Wang, P., Hawes, C. and Abell, B.M. (2012). Alternative splicing of the auxin biosynthesis gene YUCCA4 determines its subcellular compartmentation. The Plant Journal 70, 292-302.

[9]   Kopelman, N.M., Lancet, D. and Yanai, I. (2005). Alternative splicing and gene duplication are inversely correlated evolutionary mechanisms. Nature Genetics 37

[10]   Su, Z., Wang, J., Yu, J., Huang, X. and Gu, X. (2006). Evolution of alternative splicing after gene duplication. Genome Research 16, 182-189.

[11]   Blencowe, B.J. (2006). Alternative splicing: new insights from global analyses. Cell 126, 37-47.

[12]   Valouev, A., Johnson, D.S., Sundquist, A., Medina, C., Anton, E., Batzoglou, S., Myers, R.M. and Sidow, A. (2008). Genome-wide analysis of transcription factor binding sites based on ChIP-Seq data. Nature Methods 5, 829-834.

[13]   Raj, B., O'Hanlon, D., Vessey, John P., Pan, Q., Ray, D., Buckley, Noel J., Miller, Freda D. and Blencowe, Benjamin J. (2011). Cross-Regulation between an Alternative Splicing Activator and a Transcription Repressor Controls Neurogenesis. Molecular cell 43, 843-850.

[14]   Kosti, I., Radivojac, P. and Mandel-Gutfreund, Y. (2012). An Integrated Regulatory Network Reveals Pervasive Cross-Regulation among Transcription and Splicing Factors. PLoS Computational Biology 8, e1002603.

[15]   Bhardwaj, N., Carson, M.B., Abyzov, A., Yan, K.-K., Lu, H. and Gerstein, M.B. (2010). Analysis of Combinatorial Regulation: Scaling of Partnerships between Regulators with the Number of Governed Targets. PLoS Comput Biol 6, e1000755.

[16]   Shalgi, R., Lieber, D., Oren, M. and Pilpel, Y. (2007). Global and local architecture of the mammalian microRNA–transcription factor regulatory network. PLoS Computational Biology 3, e131.

[17]   Panahi, B., MOHAMMADI, S.A. and EBRAHIMIE, E. (2013). Identification of miRNAs and their potential targets in halophyte plant *Thellungiella halophila*. Biotechnologia 94, 285-290.

[18]   Panahi, B., Abbaszadeh, B., Taghizadeghan, M. and Ebrahimie, E. (2014). Genome-wide survey of Alternative Splicing in *Sorghum bicolor*. Physiology and Molecular Biology of Plants

[19]   Bakhtiarizadeh, M.R., Ebrahimi, M. and Ebrahimie, E. (2011). Discovery of EST-SSRs in lung cancer: Tagged ESTs with SSRs lead to differential amino acid and protein expression patterns in cancerous tissues. PloS One 6, e27118.

[20]   Bakhtiarizadeh, M.R., Arefnejad, B., Ebrahimie, E. and Ebrahimi, M. (2012). Application of functional genomic information to develop efficient EST-SSRs for the chicken (*Gallus gallus*). Genetics and Molecular Research 11, 1558-1574.

[21]   Bakhtiarizadeh, M.R., Moradi-Shahrbabak, M. and Ebrahimie, E. (2013). Underlying functional genomics of fat deposition in adipose tissue. Gene 521, 122-128.

[22]   Shamloo-Dashtpagerdi, R., Razi, H., Lindlöf, A., Niazi, A., Dadkhodaie, A. and Ebrahimie, E. (2013). Comparative analysis of expressed sequence tags (ESTs) from *Triticum monococcum* shoot apical meristem at vegetative and reproductive stages. Genes & Genomics 35, 365-375.

[23]   Bakhtiarizadeh, M.R., Moradi-Shahrbabak, M. and Ebrahimie, E. (2014). Transcriptional regulatory network analysis of the over-expressed genes in adipose tissue. Genes & Genomics 36, 105-117.

[24]   Gupta, S., Zink, D., Korn, B., Vingron, M. and Haas, S.A. (2004). Genome wide identification and classification of alternative splicing based on EST data. Bioinformatics 20, 2579-2585.

[25]   Campbell, M., Haas, B., Hamilton, J., Mount, S. and Buell, C.R. (2006). Comprehensive analysis of alternative splicing in rice and comparative analyses with *Arabidopsis*. BMC Genomics 7, 327.

[26]   Walters, B., Lum, G., Sablok, G. and Min, X.J. (2013). Genome-wide landscape of alternative splicing events in *Brachypodium distachyon*. DNA Research 20, 163-171.

[27]   Xiao, Y.-L. et al. (2005). Analysis of the cDNAs of hypothetical genes on *Arabidopsis* chromosome 2 reveals numerous transcript variants. Plant Physiology 139, 1323-1337.

[28]   Wang, B.-B. and Brendel, V. (2006). Genomewide comparative analysis of alternative splicing in plants. Proceedings of the National Academy of Sciences 103, 7175-7180.

[29]   Fruzangohar, M., Ebrahimie, E., Ogunniyi, A.D., Mahdi, L.K., Paton, J.C. and Adelson, D.L. (2013). Comparative GO: A web application for comparative gene ontology and gene ontology-based gene selection in bacteria. PloS One 8, e58759.

[30]   Consortium, I.B.G.S. (2012). A physical, genetic and functional sequence assembly of the barley genome. Nature 491, 711-716.

[31]   Duvick, J., Fu, A., Muppirala, U., Sabharwal, M., Wilkerson, M.D., Lawrence, C.J., Lushbough, C. and Brendel, V. (2008). PlantGDB: a resource for comparative plant genomics. Nucleic Acids Research 36, D959-D965.

[32]   Kurtz, S., Choudhuri, J.V., Ohlebusch, E., Schleiermacher, C., Stoye, J. and Giegerich, R. (2001). REPuter: the manifold applications of repeat analysis on a genomic scale. Nucleic Acids Research 29, 4633-4642.

[33]   Kalyanaraman, A., Aluru, S., Kothari, S. and Brendel, V. (2003). Efficient clustering of large EST data sets on parallel computers. Nucleic Acids Research 31, 2963-2974.

[34]   Huang, X. and Madan, A. (1999). CAP3: A DNA sequence assembly program. Genome Research 9, 868-877.

[35]   Min, X.J. (2013). ASFinder: a tool for genome–wide identification of alternatively splicing transcripts from EST–derived sequences. International Journal of Bioinformatics Research and Applications 9, 221-226.

[36]   Florea, L., Hartzell, G., Zhang, Z., Rubin, G.M. and Miller, W. (1998). A Computer Program for Aligning a cDNA Sequence with a Genomic DNA Sequence. Genome Research 8, 967-974.

[37]   Foissac, S. and Sammeth, M. (2007). ASTALAVISTA: dynamic and flexible analysis of alternative splicing events in custom gene datasets. Nucleic Acids Research 35, W297-W299.

[38]   Fruzangohar, M., Ebrahimie, E. and Adelson, D.L. (2014). Application of Global Transcriptome Data in Gene Ontology Classification and Construction of a Gene Ontology Interaction Network. bioRxiv

[39]   Reimand, J., Arak, T. and Vilo, J. (2011). g: Profiler—a web server for functional interpretation of gene lists (2011 update). Nucleic Acids Research 39, W307-W315.

[40]   Hosseinpour, B., HajiHoseini, V., Kashfi, R., Ebrahimie, E. and Hemmatzadeh, F. (2012). Protein Interaction Network of *Arabidopsis thaliana* Female Gametophyte Development Identifies Novel Proteins and Relations. PloS One 7, e49931.

[41]   Nikitin, A., Egorov, S., Daraselia, N. and Mazo, I. (2003). Pathway studio—the analysis and navigation of molecular networks. Bioinformatics 19, 2155-2157.

[42]   Szklarczyk, D. et al. (2014). STRING v10: protein–protein interaction networks, integrated over the tree of life. Nucleic Acids Research, gku1003.

[43]   Chien, C.-H., Chiang-Hsieh, Y.-F., Chen, Y.-A., Chow, C.-N., Wu, N.-Y., Hou, P.-F. and Chang, W.-C. (2015). AtmiRNET: a web-based resource for reconstructing regulatory networks of *Arabidopsis* microRNAs. Database 2015, bav042.

[44] Jones-Rhoades, M.W. and Bartel, D.P. (2004). Computational Identification of Plant MicroRNAs and Their Targets, Including a Stress-Induced miRNA. Molecular Cell 14, 787-799.

[45] Baek, J.-M., Han, P., Iandolino, A. and Cook, D. (2008). Characterization and comparison of intron structure and alternative splicing between *Medicago truncatula*, *Populus trichocarpa*, *Arabidopsis* and rice. Plant Molecular Biology 67, 499-510.

[46] Labadorf, A., Link, A., Rogers, M., Thomas, J., Reddy, A. and Ben-Hur, A. (2010). Genome-wide analysis of alternative splicing in *Chlamydomonas reinhardtii*. BMC Genomics 11, 114.

[47] Sablok, G., Gupta, P.K., Baek, J.-M., Vazquez, F. and Min, X. (2011). Genome-wide survey of alternative splicing in the grass *Brachypodium distachyon*: a emerging model biosystem for plant functional genomics. Biotechnology Letters 33, 629-636.

[48] Ner-Gaon, H., Leviatan, N., Rubin, E. and Fluhr, R. (2007). Comparative Cross-Species Alternative Splicing in Plants. Plant Physiology 144, 1632-1641.

[49] Mühlemann, O., Eberle, A.B., Stalder, L. and Zamudio Orozco, R. (2008). Recognition and elimination of nonsense mRNA. Biochimica et Biophysica Acta (BBA) - Gene Regulatory Mechanisms 1779, 538-549.

[50] Marquez, Y., Brown, J.W., Simpson, C., Barta, A. and Kalyna, M. (2012). Transcriptome survey reveals increased complexity of the alternative splicing landscape in *Arabidopsis*. Genome Research 22, 1184-1195.

[51] Lareau, L.F., Inada, M., Green, R.E., Wengrod, J.C. and Brenner, S.E. (2007). Unproductive splicing of SR genes associated with highly conserved and ultraconserved DNA elements. Nature 446, 926-929.

[52] Morello, L. and Breviario, D. (2008). Plant spliceosomal introns: not only cut and paste. Current Genomics 9, 227-238.

[53] Saltzman, A.L., Kim, Y.K., Pan, Q., Fagnani, M.M., Maquat, L.E. and Blencowe, B.J. (2008). Regulation of Multiple Core Spliceosomal Proteins by Alternative Splicing-Coupled Nonsense-Mediated mRNA Decay. Molecular and Cellular Biology 28, 4320-4330.

[54] Lewis, B.P., Green, R.E. and Brenner, S.E. (2003). Evidence for the widespread coupling of alternative splicing and nonsense-mediated mRNA decay in humans. Proceedings of the National Academy of Sciences 100, 189-192.

[55] Lynch, M. and Conery, J.S. (2000). The evolutionary fate and consequences of duplicate genes. Science 290, 1151-1155.

[56] Tuskan, G.A. et al. (2006). The genome of black cottonwood, *Populus trichocarpa* (Torr. & Gray). Science 313, 1596-1604.

[57] Kalyna, M. et al. (2012). Alternative splicing and nonsense-mediated decay modulate expression of important regulatory genes in Arabidopsis. Nucleic Acids Research 40, 2454-2469.

[58] Göhring, J., Jacak, J. and Barta, A. (2014). Imaging of Endogenous Messenger RNA Splice Variants in Living Cells Reveals Nuclear Retention of Transcripts Inaccessible to Nonsense-Mediated Decay in *Arabidopsis*. The Plant Cell Online 26, 754-764.

[59] Palusa, S.G. and Reddy, A.S. (2010). Extensive coupling of alternative splicing of pre-mRNAs of serine/arginine (SR) genes with nonsense-mediated decay. New Phytologist 185, 83-89.

[60] Fox-Walsh, K.L., Dou, Y., Lam, B.J., Hung, S.-p., Baldi, P.F. and Hertel, K.J. (2005). The architecture of pre-mRNAs affects mechanisms of splice-site pairing. Proceedings of the National Academy of Sciences of the United States of America 102, 16176-16181.

[61] Sterner, D.A., Carlo, T. and Berget, S.M. (1996). Architectural limits on split genes. Proceedings of the National Academy of Sciences 93, 15081-15085.

[62] Reddy, A.S., Marquez, Y., Kalyna, M. and Barta, A. (2013). Complexity of the alternative splicing landscape in plants. The Plant Cell Online 25, 3657-3683.

[63] Pertea, M., Mount, S.M. and Salzberg, S.L. (2007). A computational survey of candidate exonic splicing enhancer motifs in the model plant *Arabidopsis thaliana*. BMC Bioinformatics 8, 159.

[64] Seo, P.J., Kim, M.J., Ryu, J.-Y., Jeong, E.-Y. and Park, C.-M. (2011). Two splice variants of the IDD14 transcription factor competitively form nonfunctional heterodimers which may regulate starch metabolism. Nature Communications 2, 303.

[65] Palusa, S.G., Ali, G.S. and Reddy, A.S. (2007). Alternative splicing of pre-mRNAs of *Arabidopsis* serine/arginine-rich proteins: regulation by hormones and stresses. The Plant Journal 49, 1091-1107.

[66] Alimohammadi, A., Shiran, B., Martinez-Gomez, P. and Ebrahimie, E. (2013). Identification of water-deficit resistance genes in wild almond *Prunus scoparia* using cDNA-AFLP. Scientia Horticulturae 159, 19-28.

[67] Golovkin, M. and Reddy, A.S.N. (2003). Expression of U1 Small Nuclear Ribonucleoprotein 70K Antisense Transcript Using APETALA3 Promoter Suppresses the Development of Sepals and Petals. Plant Physiology 132, 1884-1891.

[68] Ali, G.S., Golovkin, M. and Reddy, A.S.N. (2003). Nuclear localization and in vivo dynamics of a plant-specific serine/arginine-rich protein. The Plant Journal 36, 883-893.

[69] Jin, D., Wang, Y., Zhao, Y. and Chen, M. (2013). MicroRNAs and Their Cross-Talks in Plant Development. Journal of Genetics and Genomics 40, 161-170.

[70] Kurihara, Y., Takashi, Y. and Watanabe, Y. (2006). The interaction between DCL1 and HYL1 is important for efficient and precise processing of pri-miRNA in plant microRNA biogenesis. RNA 12, 206-212.

[71] Leyva-González, M.A., Ibarra-Laclette, E., Cruz-Ramírez, A. and Herrera-Estrella, L. (2012). Functional and Transcriptome Analysis Reveals an Acclimatization Strategy for Abiotic Stress Tolerance Mediated by *Arabidopsis* NF-YA Family Members. PloS One 7, e48138.

[72] Sunkar, R., Li, Y.-F. and Jagadeeswaran, G. (2012). Functions of microRNAs in plant stress responses. Trends in Plant Science 17, 196-203.

[73] Kumari, G., Kusumanjali, K., Srivastava, P. and Das, S. (2013). Isolation and expression analysis of miR165a and REVOLUTA from *Brassica* species. Acta Physiologiae Plantarum 35, 399-410.

[74] Porth, I. et al. (2013). Extensive functional pleiotropy of REVOLUTA substantiated through forward genetics. Plant Physiology 164, 548-54.

[75] Ratcliffe, O.J., Riechmann, J.L. and Zhang, J.Z. (2000). INTERFASCICULAR FIBERLESS1 Is the Same Gene as REVOLUTA. The Plant Cell Online 12, 315-317.

[76] Zhong, R. and Ye, Z.-H. (2007). Regulation of HD-ZIP III genes by microRNA 165. Plant Signaling & Behavior 2, 351-353.

[77] Gustafson, A.M., Allen, E., Givan, S., Smith, D., Carrington, J.C. and Kasschau, K.D. (2005). ASRP: the *Arabidopsis* small RNA project database. Nucleic Acids Research 33, D637-D640.

[78] Shikata, M., Koyama, T., Mitsuda, N. and Ohme-Takagi, M. (2009). *Arabidopsis* SBP-box genes SPL10, SPL11 and SPL2 control morphological change in association with shoot maturation in the reproductive phase. Plant and Cell Physiology 50, 2133-2145.

[79] Wu, G., Park, M.Y., Conway, S.R., Wang, J.-W., Weigel, D. and Poethig, R.S. (2009). The Sequential Action of miR156 and miR172 Regulates Developmental Timing in *Arabidopsis*. Cell 138, 750-759.

[80] Ronemus, M., Vaughn, M.W. and Martienssen, R.A. (2006). MicroRNA-targeted and small interfering RNA–mediated mRNA degradation is regulated by Argonaute, Dicer, and RNA-dependent RNA polymerase in *Arabidopsis*. The Plant Cell Online 18, 1559-1574.

[81] Li, L.C. et al. (2008). SPOROCYTELESS modulates YUCCA expression to regulate the development of lateral organs in *Arabidopsis*. New Phytologist 179, 751-764.

[82] Kianianmomeni, A., Ong, C.S., Rätsch, G. and Hallmann, A. (2014). Genome-wide analysis of alternative splicing in Volvox carteri. BMC Genomics 15, 1117.

Figures

Figure 1: Transcript variants of Serine/threonine-protein kinase in *Hordeum vulgare*. Alternative splicing (AS) in this gene produces 13 transcripts per gene. Intron retention occurred in all of the variants; also exon skiping resulted in a variant in Serine/threonine-protein kinase for example in MLOC_53969.8.

Figure 2: Distribution of different type of Alternative splicing (AS) **events in *H. vulgare*.** Intron R: Intron retention, AltD:Alternative donor splice site, AltA: Alternative acceptor splice site, ExonS: Exon skipping. Others: Mutually exclusive exons, Alternative promoters, Alternative poly (A) and other trans forms categorized as a Others or complex events. Intron retention is the prevalent type of AS in *H. vulgare.* In contrast, Exon skiping is a rarest form.

Figure 3: Distribution of internal exon and intron size in Alternative splicing (AS) events in *H. vulgare*. A) The x-axis indicates the size of internal exons and the y-axis indicates the frequency of internal exons. B) The x-axis indicates the size of introns and the y-axis indicates the frequency of internal exons. Bin sizes are right inclusive (e.g. bin 100 comprises sequences of lengths 1–100 bp). More than 50 % of exons sizes are below 200 bp (A). The average intron size is 433 bp(B).

Figure 4. Gene Onyology (GO) distribution of Alternative spliced (AS) genes (blue bars) vs. Reference genome (green bars). The Y-axis is the percentage of genes mapped by the term, and represents the abundance of the GO term. The X-axis is the definition of GO terms. A) Molecular function for AS genes. B) Biological process for AS genes. 'Catalytic activity and binding' in Molecular function and 'Metabolic process' in Biological Process of AS genes is higher than its expected number of whole Genome in both sets of genes. Catalytic activities including of transferase activity, hydrolase activity, oxidoreductase activity, RNA splicing factor activity, lyse activity, ligase activity, isomerase activity genes. 'Binding' also included binding to DNA, RNA (interestingly consist a proteins that including in mRNA processing).

Figure 5. Gene Onyology (GO) distribution of Alternative spliced (AS) genes in Molecular function. The Y-axis is the percentage of genes mapped by the term, and

26

represents the abundance of the GO term. The X-axis is the definition of GO terms. Catalytic activity and binding are prevalence group in molecular function categories.

Figure 6. Subnetwork of cross-talk between splicing machinery genes derived from high Alternative spliced (AS) and low AS sets in *Hordeum vulgare*. Network includes interaction splicing machinery, miRNAs, TFs, stress, cell process, complex and functional class**.** Network constructed using Pathway Studio version 10 (ELSEVIER) based on high and low AS genes. To identify splicing machinery genes of high and low AS genes, genes that found specifically in the high AS, low AS and common in the two gene sets network are highlighted in yellow, blue and pink. Yellow boxes under the network are Gene Ontology (GO) which reveled AS involved in many cell processes. U1-70 k is a hub protein in this network that is shown by red flash. ⊕Represents positive-regulated and⊢represents negative-regulated

Figure 7. Interaction between miR169 family and *NF-YA10*. Network includes interaction of miRNAs and transcription factors (TFs). *NF-YA10* is in the high **Alternative spliced (**AS) genes and regulated by miR169 family members (except miR169a). High AS, low AS and common in the two genes sets in network highlighted as yellow, blue and pink.⊢Represent as a negative-regulated

Figure 8: Interaction of *REV* and miR165, miR166 and its impact on various developmental processes. Network contains miRNAs, transcription factors (TFs), cell process, complex and small molecule. Network constructed using Pathway Studio version 10 (ELSEVIER) based on high and low AS genes. *REV* is highlighted in yellow and TFs belonging to low AS gene are blue highlighted. Yellow boxes under the network are Gene Ontology (GO) which reveled involvement of *REV* in many cell processes. ⊕ represents positive-regulated and⊢ represents negative-regulated.

Figure 9: Cross-talk between *SPLs* and miR156/miR157 family in high splicing gene set. Network includes interaction between miRNAs, transcription factors (TFs) and proteins. Network is constructed using Pathway Studio version 10 (ELSEVIER) based on SPLs that are high Alternative spliced (AS) genes. Yellow, blue and pink highlights are genes that are in high, low and common in two set genes, respectively. ⊕ represents positive-regulated and ⊢ represents negative-regulated.

Figure 10: Alternative Splicing (A) and expression pattern (B and C) of Elongation factor (EE) MLOC_3412 in *H. vulgare*. EF has two splicing variant which variant 2 is a major transcript in response to salt stress in barley
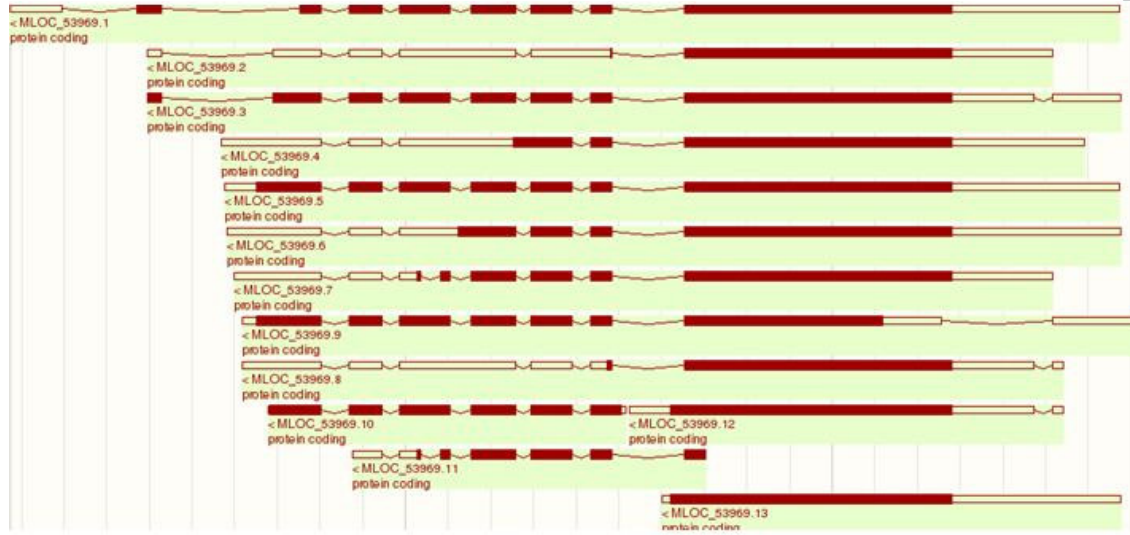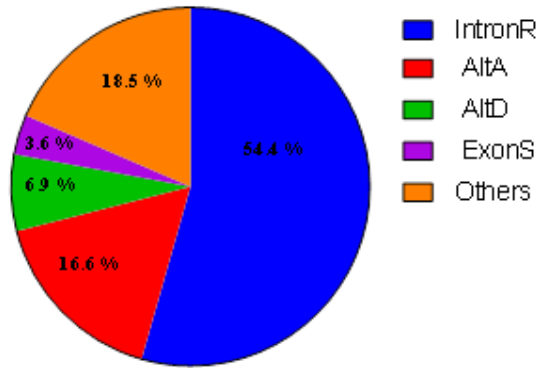
Fig. 1

Fig. 2
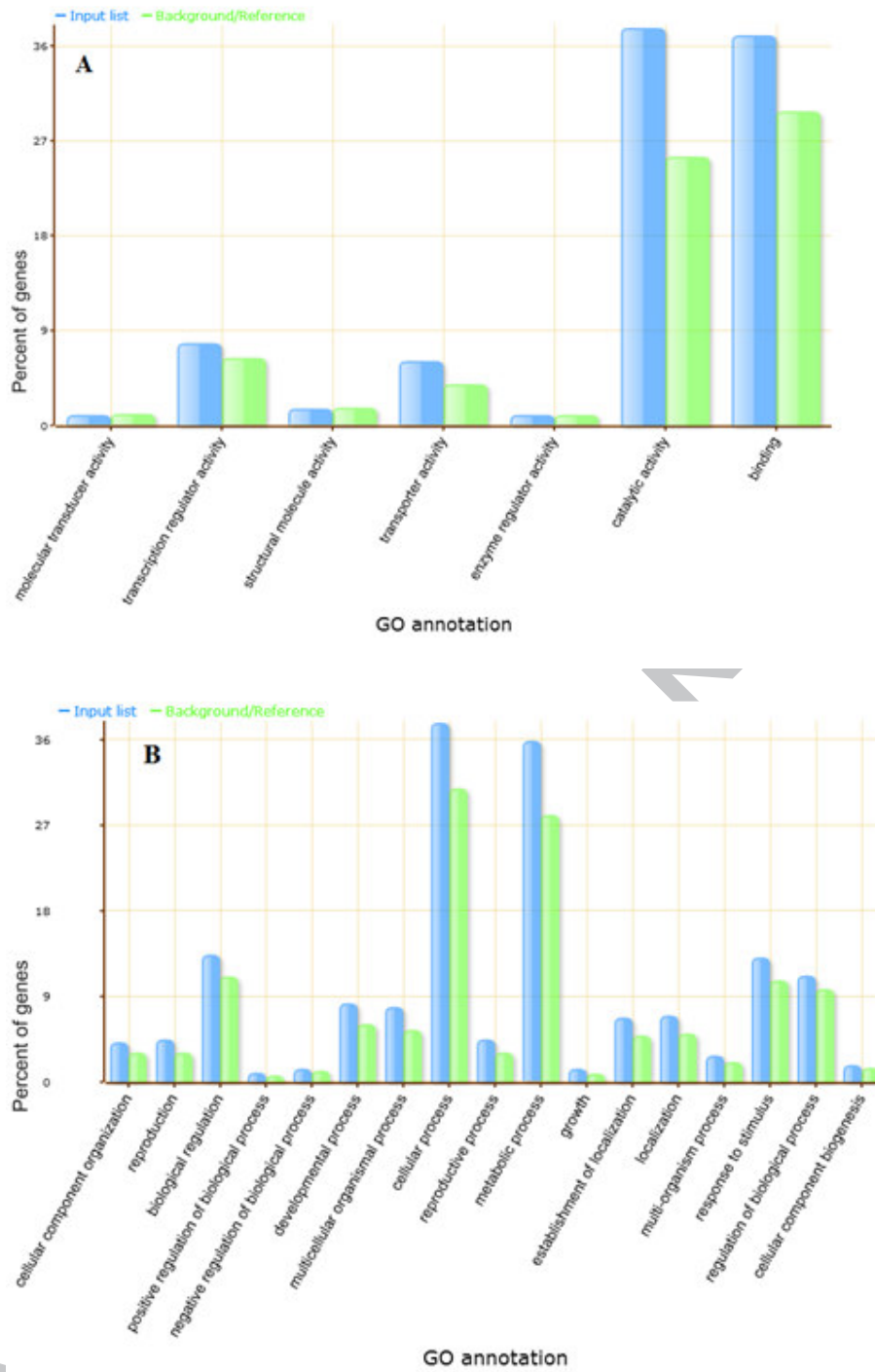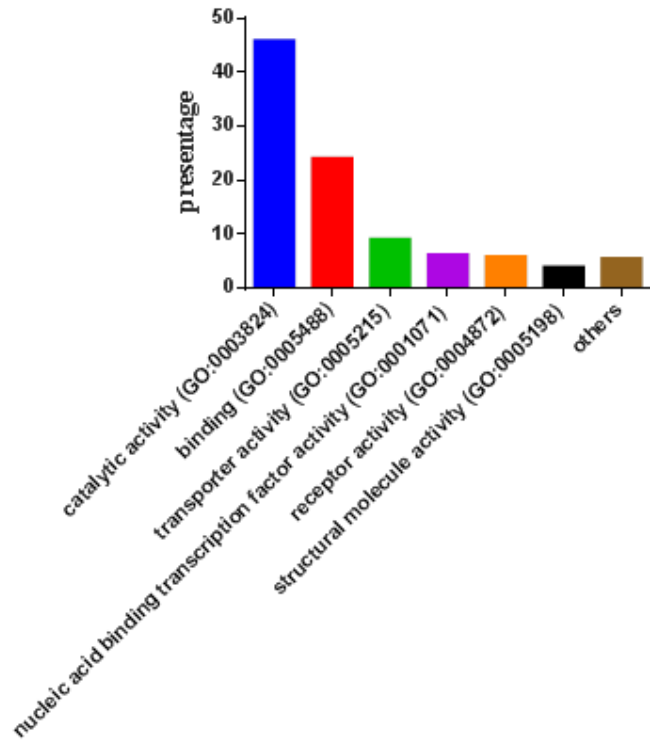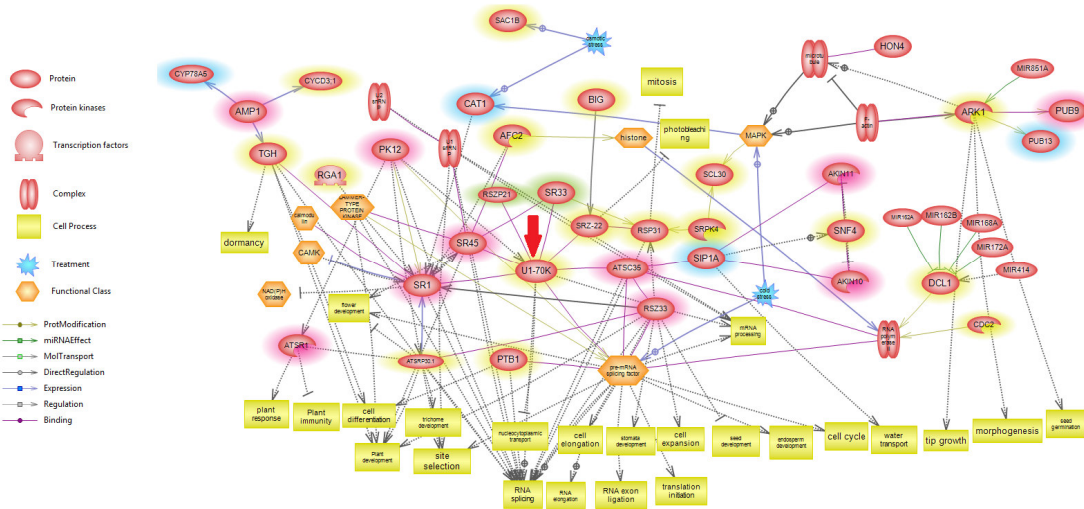
Fig. 3

Fig. 4
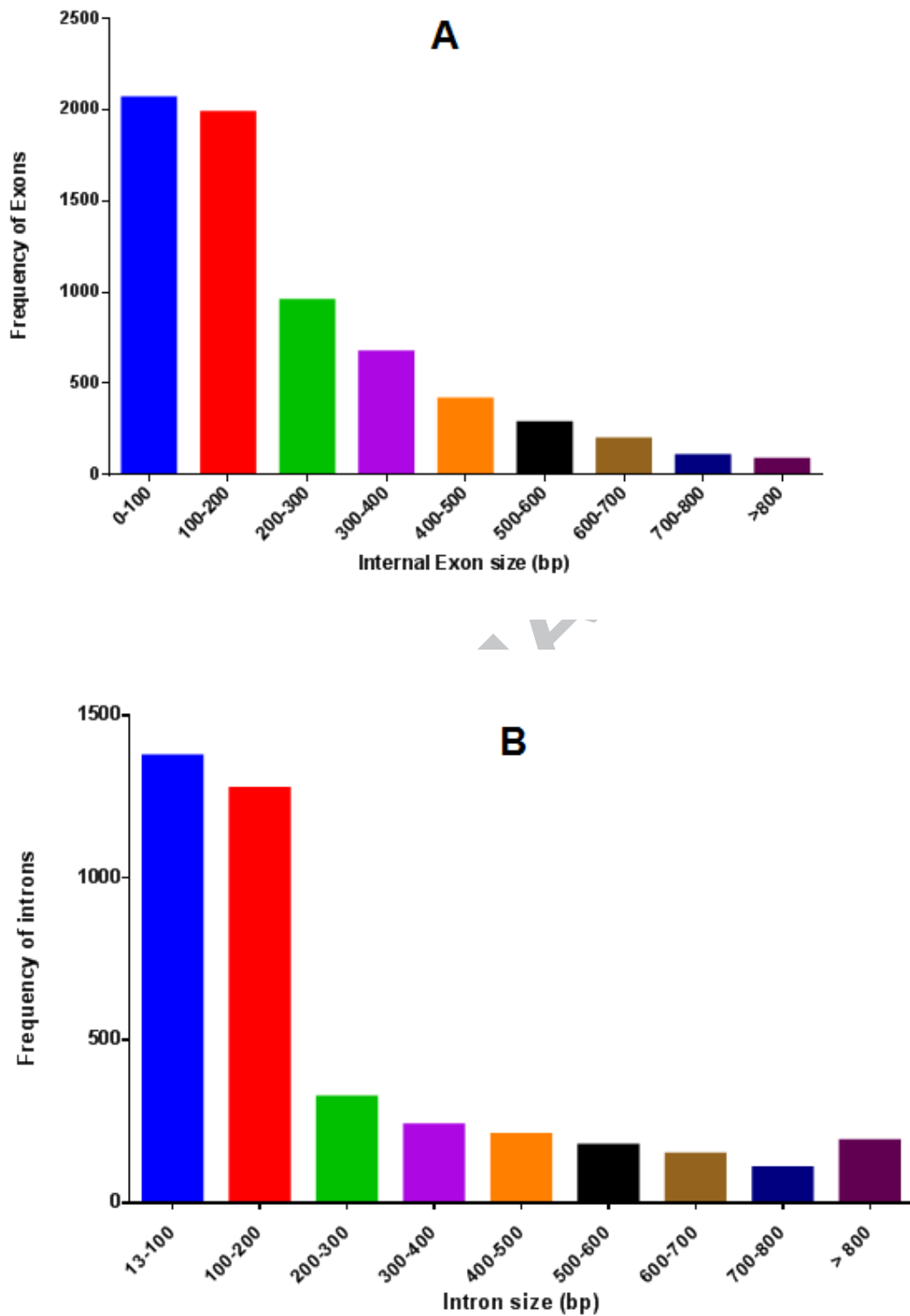
Fig. 5

Fig. 6
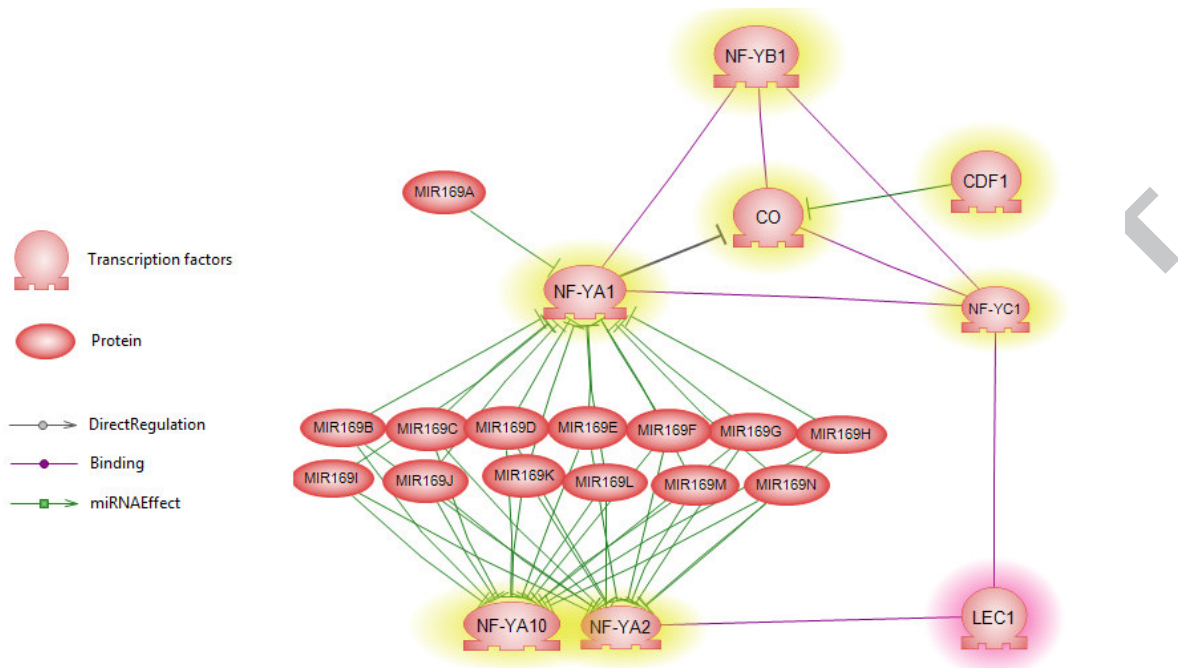
Fig. 7

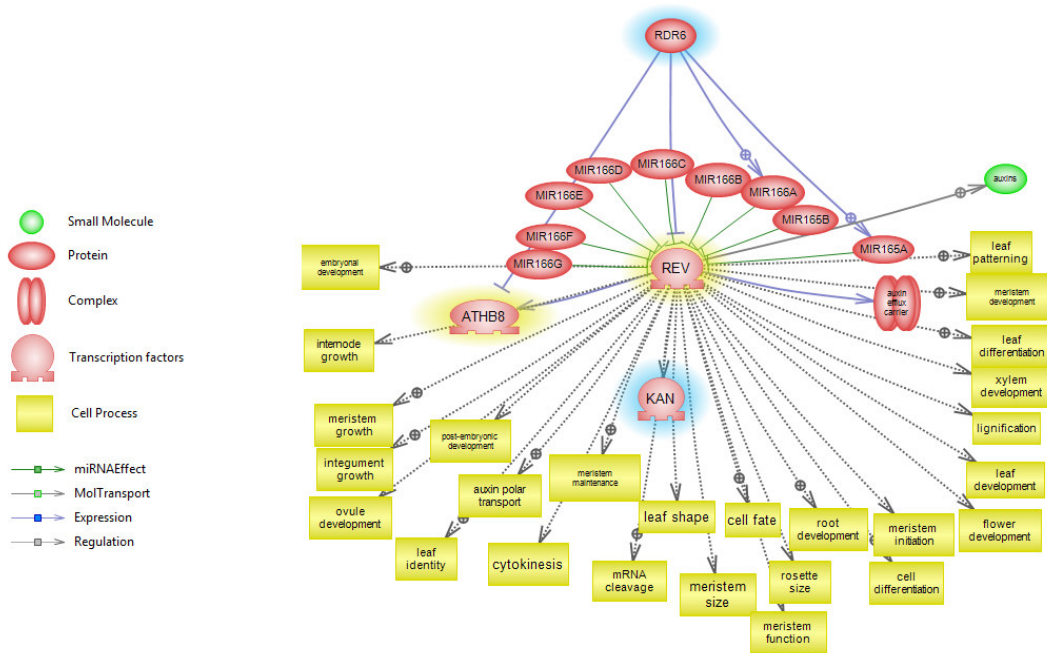Fig. 8
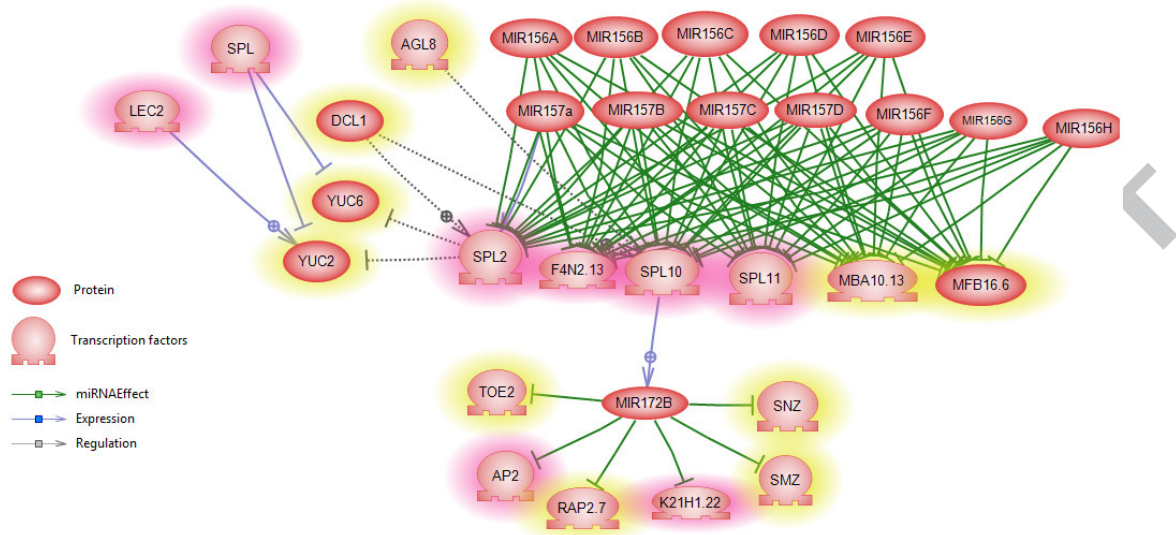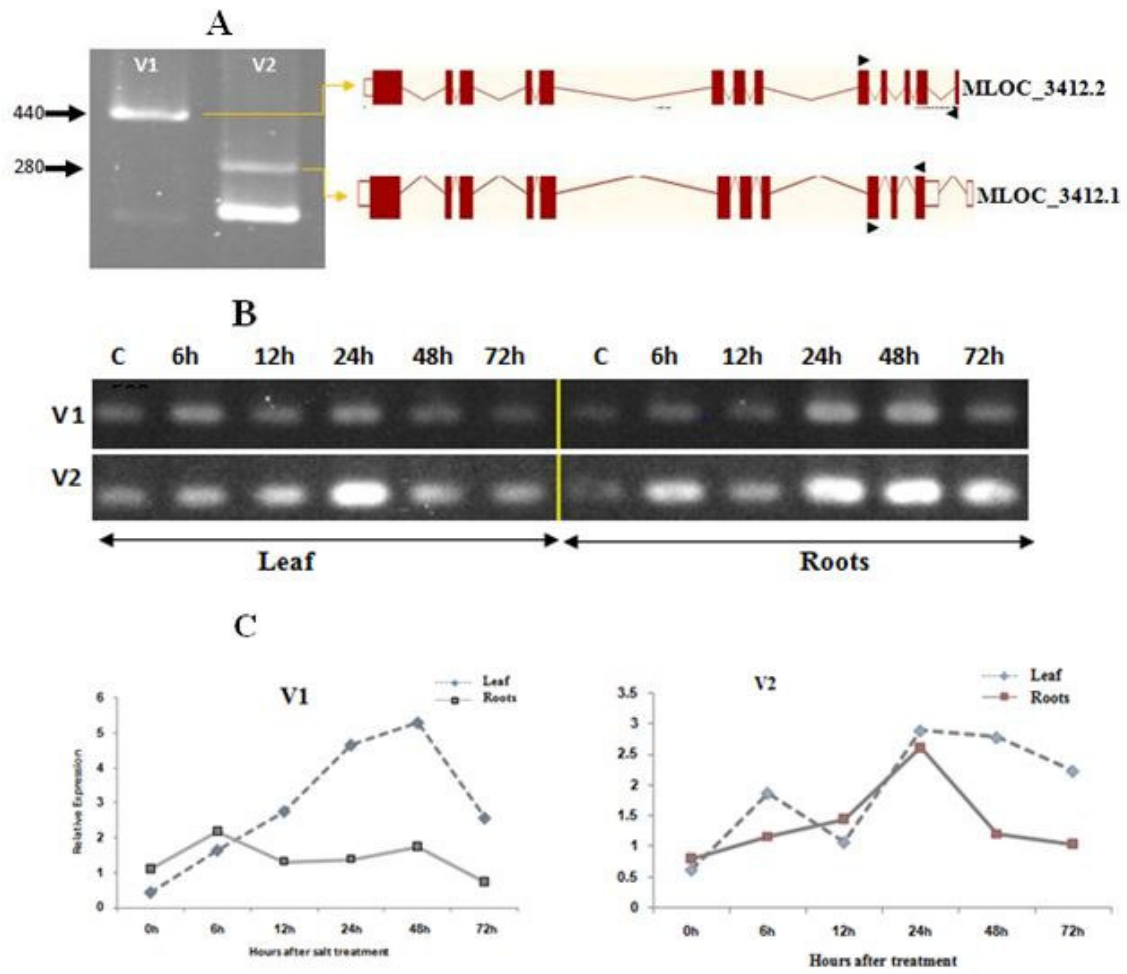
Fig. 9

Fig. 10

>Results demonstrated that intron retention is frequently associated with specific abiotic stresses

>Network analysis resulted in discovery of some specific sub networks between miRNAs and transcription factors in genes with high number of alternative splicing,

>Interaction between high and low alternative splicing genes highlighted the role of U1-70K as a hub (central) protein.