

Cite this: DOI: 10.1039/c3mb70072k

## Comparative analyses of stress-responsive genes in *Arabidopsis thaliana*: insight from genomic data mining, functional enrichment, pathway analysis and phenomics†

Mahantesha Naika,<sup>a,b</sup> Khader Shameer<sup>†,a</sup> and Ramanathan Sowdhamini<sup>\*a</sup>

Biotic and abiotic stresses adversely affect agriculture by reducing crop growth and productivity worldwide. To investigate the abiotic stress-responsive genes in *Arabidopsis thaliana*, we compiled a dataset of stress signals and differentially upregulated genes ( $\geq 2.5$  fold change) from Stress-responsive transcription Factors DataBase (STIFDB) with additional set of stress signals and genes curated from PubMed and Gene Expression Omnibus. A dataset of 3091 genes differentially upregulated due to 14 different stress signals (abscisic acid, aluminum, cold, cold-drought-salt, dehydration, drought, heat, iron, light, NaCl, osmotic stress, oxidative stress, UV-B and wounding) were curated and used for the analysis. Details about stress-responsive enriched genes and their association with stress signals can be obtained from STIFDB2 database <http://caps.ncbs.res.in/stifdb2>. The gene–stress-signal data were analyzed using an enrichment-based meta-analysis framework consisting of two different ontologies (Gene Ontology and Plant Ontology), biological pathway and functional domain annotations. We found several shared and distinct biological processes, cellular components and molecular functions associated with stress-responsive genes. Pathway analysis revealed that stress-responsive genes perturbed the pathways under the “Metabolic pathways” category. We also found several shared and stress-signal specific protein domains, suggesting functional mechanisms regulating stress-response. Phenomic characteristics of abiotic stress-responsive genes were ascertained for several stresses and found to be shared by multiple stresses in both anatomy and temporal categories of Plant Ontology. We found several constitutive stress-responsive genes that are differentially upregulated due to perturbation of different stress signals, for example a gene (AT1G68440) involved in phenylpropanoid metabolism and polyamine catabolism as responsive to seven different stress signals. We also performed structure–function prediction of five genes associated responsive to multiple abiotic stress signals. We envisage that results from our analysis that provide insight into functional repertoire, metabolic pathways and phenomic characteristics common and specifically associated with stress signals would help to understand abiotic stress regulome in *Arabidopsis thaliana* and may also help to develop an improved plant variety using molecular breeding and genetic engineering techniques that are rapidly stress-responsive and tolerant.

Received 20th February 2013,  
Accepted 10th April 2013

DOI: 10.1039/c3mb70072k

[www.rsc.org/molecularbiosystems](http://www.rsc.org/molecularbiosystems)

## Introduction

Plants are sessile organisms, and constantly exposed to a wide range of abiotic stresses and have to adapt to variable

environmental stimuli. Attaining a balanced state between normal and extreme environmental signals and biological responses is crucial for optimal growth of a plant.<sup>1</sup> Crop growth and development are constrained by various biotic and abiotic stresses which result in reduced crop productivity and losses worldwide. For example, more than 10% of arable land is affected by drought and salinity, which results in more than 50% decline in the average yield worldwide.<sup>2</sup> Stress response and stress tolerance or susceptibility towards abiotic stresses are often mediated by complex biological processes. Stress may occur singularly or in combination and induce cellular damage at multiple stages of plant development and growth and may

<sup>a</sup> National Centre for Biological Sciences (TIFR), GKVK Campus, Bangalore, 560065, India. E-mail: mini@ncbs.res.in

<sup>b</sup> Department of Plant Biotechnology, University of Agricultural Sciences, GKVK Campus, Bangalore, 560065, India

† Electronic supplementary information (ESI) available. See DOI: 10.1039/c3mb70072k

‡ Present address: Division of Biomedical Statistics and Informatics, Division of Cardiovascular Diseases, Mayo Clinic, Rochester, MN 55905, USA.

induce phenotypic lethality.<sup>3</sup> However, plants have developed mechanisms to perceive the subtle changes of growth conditions, and activate distinct signal transduction cascades, which in turn activate stress-responsive genes and ultimately lead to transcriptional reprogramming of metabolism, physiology, morphology and survival.<sup>4</sup> Several aspects of stress response are known, but connections between the core group of genes, functional processes, pathways, protein domains or the phenotype shared between multiple stress signals remain elusive. We used an integrated analytical pipeline that combined data curation,<sup>5</sup> genomic data mining,<sup>6</sup> Gene Ontology (GO) term enrichment analysis<sup>7,8</sup> and pathway analysis<sup>9</sup> coupled with phenomic enrichment analysis using Plant Ontology (PO)<sup>10–16</sup> to understand various aspects of abiotic stress-responsive events in *Arabidopsis thaliana*.

Abiotic stress refers to a collective term to define a class of a large number of stress signals that could induce several molecular and phenotypic changes in the plants. Abiotic factors that could induce stress in the plant phenotype include abscisic acid (ABA), temperature, UV-radiation, humidity, light intensity, heat, dehydration, level of minerals like aluminum, CO<sub>2</sub> level, drought, salt, osmotic stress, oxidative stress etc.<sup>17</sup> Abiotic stresses affect various aspects of plant growth and perturb the growth potential.<sup>18</sup> Abiotic stresses were shown to impact transcriptional reprogramming that leads to expression of stress-responsive genes, influence multiple signalling pathways<sup>4,19–28</sup> and may also affect photosynthesis.<sup>29</sup> Various transcription factors were also shown to influence stress response in plants.<sup>30–32</sup> Previous studies have shown that microRNA (miRNA) and small interfering RNA (siRNA) in plants may also influence stress response and stress tolerance in plants including *A. thaliana*.<sup>33,34</sup> DNA methylation, an epigenetic mechanism, also plays a crucial role in stress responses.<sup>35–37</sup> Various signaling pathways were also attributed as mediators of stress response in *A. thaliana*.<sup>38–42</sup> Abiotic stresses have shown to affect multiple stages of plant growth or morphology. Abiotic stresses perturb plant secondary metabolite production, affect seed germination,<sup>34</sup> leaf senescence,<sup>38</sup> induce leaf rolling,<sup>43</sup> limit crop productivity, delay developmental and growth rate of plants. Abiotic stress response in plant induces production of reactive oxygen species (ROS),<sup>44</sup> transcriptional regulation, epigenetic modification,<sup>21</sup> activation of signal transduction pathways in endoplasmic reticulum,<sup>26</sup> calcium signaling,<sup>45</sup> metabolic perturbations<sup>40</sup> and regulation of plant hormones including auxin.<sup>42,46</sup> Different stresses have shown to adapt different biological directions to respond to stress. The type of stress and the plant growth stage at which the plant was exposed to a particular stress primarily drove stress response mechanisms. In *Arabidopsis thaliana* mitochondrial electron transport,<sup>47</sup> signaling events,<sup>19,26,38,48</sup> molecular cross-talks,<sup>41</sup> regulatory events,<sup>49</sup> miRNAs and siRNAs<sup>34</sup> co-ordinate the stress-responsive pathways. Early responses have also shown to help the plant to adapt to extreme environmental conditions<sup>27,50</sup> and gain stress tolerance. At the signal transduction level apart from drought, cold, salinity and heavy metals etc., one of the transduction pathways that plant encounters is by ABA signalling. ABA and salt stresses induce ATCBG expression, and specific knock-out mutants are highly

sensitive to ABA and salt treatments. Earlier findings suggest that this protein was an ABA and a salt stress-related signal transducer.<sup>51</sup> In *Oryza sativa*, salt stress was responsible for upregulation of the OsMCSU gene in root tissues and could be mediated by both ABA-dependent and ABA-independent signaling pathways.<sup>52</sup> Similarly in *Arabidopsis thaliana*, ALDH7B4, ATTSPo, ANAC019, AZF2 and ATADH1 were induced by salt stress and abscisic acid and identified that the expansion domain is highly related to both abiotic stress and abscisic acid.<sup>53</sup> In general, there is a direct relationship between various abiotic stresses like cold, drought, salinity, ABA and heavy metals etc. and stress response in plants. It will be valuable to identify additional genes and understand common and unique functional mechanisms associated with genes responsive to various abiotic stresses for crop improvement.

Abiotic stress responses were attributed to complex molecular events mediated by multiple genes and pathways.<sup>54</sup> Abiotic stress response events in plants were investigated using a variety of experimental methodologies<sup>1,27,43,55–58</sup> using genetic screening,<sup>58–60</sup> large-scale and targeted genomic studies,<sup>18,23,50,57,60–63</sup> proteomic studies,<sup>37,42,57,64–66</sup> systems<sup>26,50,54,57,63,67–69</sup> and computational<sup>42,49,50,53,55,58,67</sup> approaches. Previous studies suggest that abiotic stress-responsive pathways are highly specialized for individual stress signals. It still remains unanswered whether the stress-responsive genes could participate in similar biological processes, molecular function or colocalize in the same cellular compartment, involve in same functional pathways, encode common functional domains and finally induce similar phenotypes. Therefore, there is a need to understand the complex relationship between stress-signals, stress-responsive genes, stress tolerance mechanisms and the molecular and phenotypic characteristics. Understanding common and stress-specific genes, functional repertoire, pathways and phenotypes associated with various stress signals would help to design plants that could tolerate adverse environmental conditions and may ultimately help to reduce crop productivity losses. Application of bioinformatics approaches has been recommended as a complimentary approach to experimental studies.<sup>70</sup> We performed large-scale analysis of a dataset consisting of 14 stress signals and 3091 genes to understand functional and phenomic roles of genes responding to different types of stresses.

## Materials and methods

We attempted to address the important question of characterizing functional modules, biological pathways, plant phenotypes associated with various stress signals by characterizing functionally cohesive and distinct genes differentially upregulated due to various abiotic stress signals by performing functional enrichment analysis using annotations from multiple open access bioinformatics resources (DAVID 6.7,<sup>71</sup> Ontologizer 2.0<sup>72</sup>) and databases (PubMed, Gene Expression Omnibus (GEO),<sup>73–75</sup> Gene Ontology (GO),<sup>7</sup> Kyoto Encyclopedia of Genes and Genomes (KEGG),<sup>76</sup> Pfam<sup>77</sup> and Plant Ontology (PO)).<sup>10–14</sup> The analytical approach used in this study was designed as four distinct modules: (i) data curation, (ii) enrichment analysis (iii) structure-function prediction and (iv) biological inference and data visualization.

## Data curation strategy

An initial list of 14 different abiotic stress signals was created. These stress signals were used as queries in GEO to obtain corresponding gene expression studies in *Arabidopsis thaliana*. Gene lists for each stress signals were retrieved from respective publications *via* the PubMed link in the GEO page for a gene expression study; for gene expression studies without publication the gene list was obtained from the GEO files directly. Genes with a fold-change at the expression level  $\geq 2.5$  were retained in the stress-signal-gene lists. We have used the gene lists provided by the authors and did not assign any explicit weights to genes as expression analysis often have a high-degree of variability and assigning weights to genes identified by multiple groups may induce additional bias to the study. After data curation a single, non-redundant gene-stress signal file is created and used for the analysis (see ESI†).

## Biological enrichment analyses

Enrichment analysis is performed using DAVID 6.7 (GO enrichment analysis, pathway enrichment analysis, functional domain enrichment analysis) and Ontologizer 2.0 (phenotype enrichment analysis). We generated 14 different non-redundant gene-stress signal files using The Arabidopsis Information Resource (TAIR)<sup>78</sup> identifiers of the individual gene; these individual lists were used for stress signal specific enrichment analysis in DAVID 6.7 and Ontologizer 2.0. In this way we were able to ascribe significantly enriched GO terms, pathways, functional domains and plant phenotypes to various stress signals. To perform enrichment analysis using DAVID requires an input “gene list” and a “background”. As DAVID do not accept TAIR identifiers in default settings, for the GO term, the KEGG pathway and Pfam enrichment analysis we first converted TAIR identifiers to UniProt identifiers<sup>79</sup> using BioMart<sup>80</sup> and used the UniProt identifier list as the input “gene list”; “background” was defined as genes in *Arabidopsis thaliana* genome. We have used default setting in DAVID 6.7 for analysis. GO term enrichment using biological process, cellular compartment and molecular function terms in GO was performed using GO terms integrated in DAVID 6.7 (“GO FAT”). Pathway enrichment analysis using KEGG annotations and protein domain enrichment analyses were performed using annotation data integrated in DAVID 6.7. To deal with false positives associated with analysis using large number of genes, we filtered the output from both DAVID 6.7 on corrected *P*-values. For DAVID 6.7 results we filtered the initial output files using the Bonferroni correction method ( $P < 0.05$ ).

## Phenomic enrichment analysis using Plant Ontology

Phenotype enrichment analysis using PO terms were performed using Ontologizer 2.0. While GO annotation extensively capture generic molecular information pertaining to genes and gene products, nevertheless it does not capture organism specific features like plant phenotypes or disease phenotypes. To better illustrate plant-specific properties in an ontology framework, PO project is launched as a complementary resource to catalog phenotypic information on plants. PO is a semantic framework

that capture phenotype data sets from biological and genomic experiments. Initially it was designed as two categories as “plant anatomical entity” and “plant structure development stage”. These categories were also referred simply as “structure” and “temporal” or “growth” by TAIR curators. As per current release: PO recommends usage of a single ontology file for ontology based studies, but at the time of this writing the .obo and association files were provided by TAIR in two separate files. We have performed PO term enrichment analysis using two separate categories using Ontologizer 2.0 in order to obtain a better understanding of shared and distinct phenotypes associated with stress-responsive genes. 3 major files were required: an open bio-ontology format file that describes the ontology, an association file that provides annotation for each gene using the ontology in .obo file, a study set file identifiers of genes to test for enrichment and a population file with identifiers to define the background for statistical estimation. PO.obo file was downloaded from OBO foundry,<sup>81</sup> two associations were obtained from TAIR. Each gene list associated with the stress signal was used as the input and background is defined using identifiers derived from the genes in the association file obtained from TAIR. A total of 21 601 genes and 1166 terms (325 unique terms) were available for the anatomy category. A total of 18 416 genes and 282 terms (71 unique terms) were available for temporal.obo. A total of 28 enrichment tests were performed using two categories of PO and 14 stress signal specific gene lists. “Term-For-Term” (default) were used to perform the calculations, multiple-testing correction was set to “Bonferroni” and 1000 resampling steps were performed for multiple-testing correction. Results from Ontologizer 2.0 were filtered after multiple testing correction ( $P < 0.05$ ). A flow-chart of the analytical approach employed in this study is provided in Fig. 1(a).

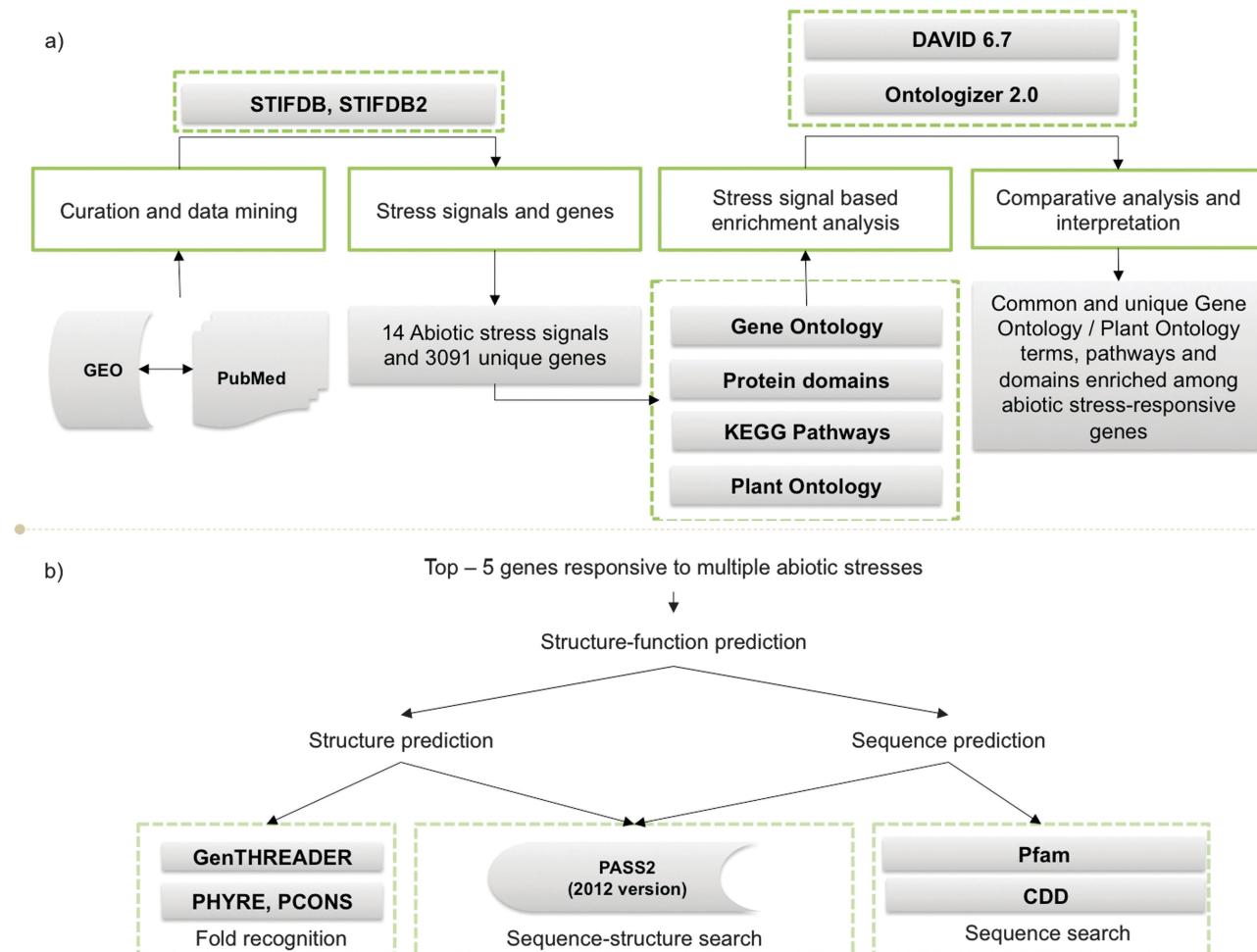
## Structure-function prediction of constitutive genes responsive to multiple abiotic stress signals

A bioinformatics pipeline was implemented using multiple tools (Fig. 1(b)) for structure–function prediction. We integrated tools from protein superfamily database PASS2.<sup>82</sup> Sensitive sequence searches were performed against the latest version of PASS2 database (version 4, release 2012) using PSI-BLAST<sup>83</sup> and HMMER<sup>84</sup> to identify structural homologues, followed by fold-recognition using GenTHREADER,<sup>85</sup> PHYRE 2<sup>86,87</sup> and PCONS.<sup>88</sup> We also consulted Pfam database<sup>77</sup> and Conserved Domain Database<sup>89</sup> for annotating the sequence of top genes for the presence of functional domains.

We have also consulted STRING – a database and derived interactome of experimentally characterized and predicted protein–protein interaction data of the top constitutive gene. The first level of network was analyzed using enrichment routine available in STRING database to recognise collective functional characteristics of networks associated with the top constitutive gene.

## Comparative analysis, visualization and biological interpretation

Relationship between genes associated with different stress-signals and annotation terms were derived after multiple testing correction.



**Fig. 1** Flow-chart of analytical pipelines used for the (a) analysis methodology used for data curation of abiotic stress-responsive genes using STIFDB2 and stress signal based enrichment analyses and (b) structure–function prediction of five genes responsive to multiple abiotic stress-signals.

For functional and pathway enrichment results from DAVID 6.7 and phenomic enrichment results were also filtered using the Bonferroni correction ( $P < 0.05$ ). Information on shared and specific genes, GO terms, biological pathways protein domains and PO terms were identified using custom Perl scripts. Shared and specific genes, protein domains, GO terms and PO terms were visualized using Circos,<sup>90</sup> REVIGO<sup>91</sup> and Gephi (URL: <https://gephi.org/>).

## Results

### Data curation

The data used in this analysis were obtained from Stress-responsive Transcription Factor DataBase version 2 (STIFDB2).<sup>92</sup> STIFDB2 was developed using a bioinformatics pipeline that combined biocuration,<sup>5</sup> genomic data mining<sup>93</sup> and stress-responsive transcription factor binding site prediction using the STIF algorithm.<sup>31</sup> Following the extensive literature survey, we created a list of 14 different abiotic stress signals as follows: abscisic acid, aluminum, cold, cold-drought-salt, dehydration, drought, heat, iron, light, NaCl, osmotic stress, oxidative stress,

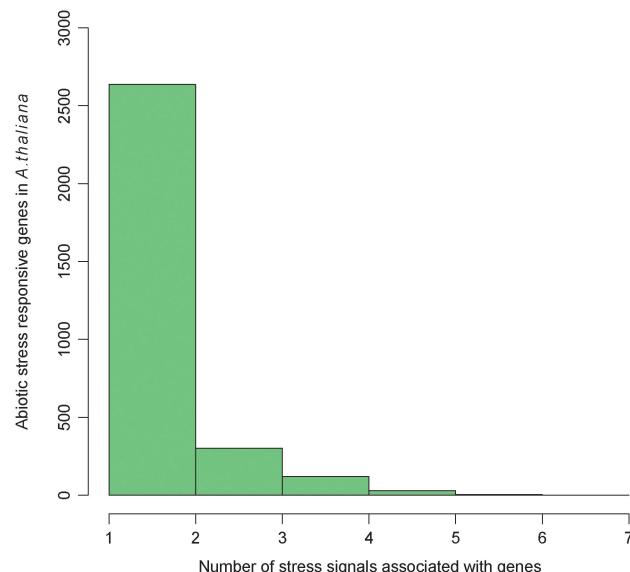
UV-B and wounding. These stress signals were separately used as search terms to query GEO and retrieve 17 different gene expression datasets. Differentially upregulated genes ( $\geq 2.5$  fold) were retrieved from GEO files or respective publications for 4948 genes and 14 different stress signal specific gene lists. The curated list has a total of 3091 unique genes. The list of 14 stress signals and GEO dataset identifiers used to curate stress-based gene lists is provided (see Table 1).

### Genes associated with multiple abiotic stress signals

To understand common functional or regulatory players underlying stress-response, we investigated whether any genes were consistently overexpressed between different stresses. A histogram of genes and stress signal distribution is provided (Fig. 2). We found that 60.8% of curated genes is associated with a single stress response. This is consistent with previous findings that a majority of stress-responsive mechanisms were driven by unique molecular mechanisms: 24.3%, 9.7%, 6.4%, 0.9% genes were responsive to two, three, four and five different abiotic stress signals. We observed that only a small percentage of common genes were responsive to different types of abiotic stresses.

**Table 1** Abiotic stress signals and corresponding GEO accession identifiers and PubMed identifiers used for curation of stress based gene lists

Abiotic stress-signal	PubMed identifier	GEO accession identifier	Genes
Abscisic acid (ABA)	21426425	GSE23301, GSE9646	700
Aluminum	—	GSE7334	25
Cold	17376166, 18625610, 21227933	GSE5621, GSE9646	1078
Cold-drought-salt (CDS)	17376166, 18625610, 21227933	GSE5621, GSE9646	41
Dehydration (DEHY)	—	GSE28493	93
Drought (DROU)	20553421, 20807999, 17376166, 18625610, 21227933	GSE5624, GSE19700 GSE24177, GSE9646	815
Heat	20139171	GSE19603	64
High light (LIGHT)	21531897	GSE22671	147
Iron deficiency (IRON)	20675571	GSE21625	674
NaCl	21227933, 18625610 21821598, 17376166	GSE5623, GSE16765 GSE9646	1040
OSMOTIC-STRESS (O-S)	19906889	GSE16474	59
Oxidative stress (OX-S)	11553744	GSE39570	40
UV-B	17376166	GSE5626	122
Wounding (WOUND)	21241326	GSE26374	50

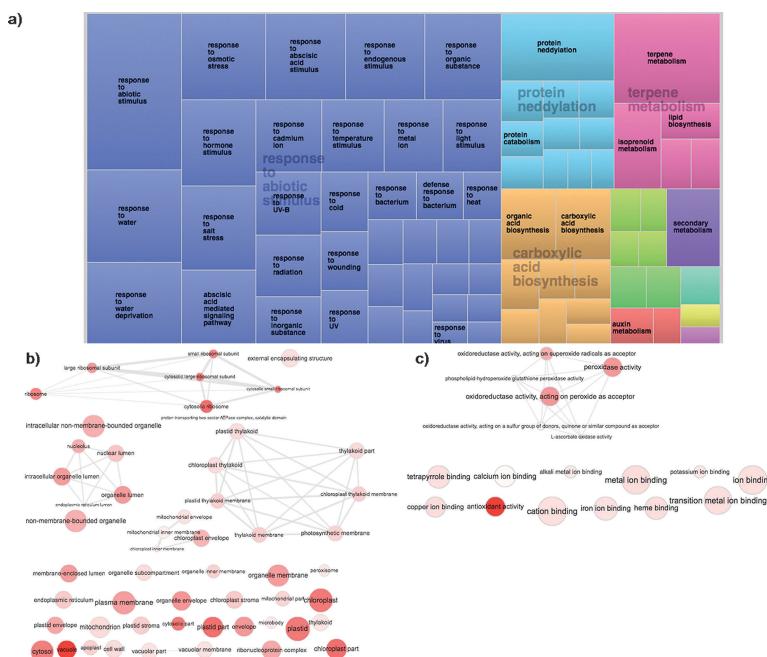
**Fig. 2** Distribution of the abiotic stress signal and stress-responsive genes.

We found a poorly characterized, hypothetical protein coding gene (AT1G68440) is upregulated as responsive to seven different stresses. Four genes (AT1G59590, AT5G58770, AT1G49450 and AT3G53110) were responsive to a set of six different abiotic stress signals surveyed in this study. AT1G68440 was annotated to be expressed during multiple growth and developmental stages as follows: 4 anthesis stage, 4 leaf senescence stage, C globular stage, D bilateral stage, E expanded cotyledon stage, F mature embryo stage, LP.02 two leaves visible stage, LP.04 four leaves visible stage, LP.06 six leaves visible stage, LP.08 eight leaves visible stage, LP.10 ten leaves visible stage, LP.12 twelve leaves visible stage and petal differentiation and expansion stage. Gene was also expressed in various parts of plant structure: carpel, caulin leaf, collective leaf structure, cotyledon, flower, hypocotyl, inflorescence meristem, leaf apex, leaf lamina base, pedicel, petal, petiole, plant embryo, plant ovule, pollen, root, seed, sepal, shoot apex,

shoot system, stamen, stem and vascular leaf.<sup>94</sup> Recent computational analyses on *Arabidopsis thaliana* genes suggest that this gene is also associated with various biological processes like the phenylpropanoid metabolic process, the polyamine catabolic process and the cellular modified amino acid biosynthetic process.<sup>95</sup>

### GO term enrichment analysis

Different gene lists have varying degrees of GO term associations. For example, see Fig. 3 for a reduced visualization of 3 largest lists of terms identified from the analysis. Significantly enriched biological process terms were identified for genes associated with the following 10 different stress signals: ABA (Table S1, ESI†), COLD (Table S2, ESI†), COLD-DROUGHT-SALT (Table S3, ESI†), DEHYDRATION (Table S4, ESI†), DROUGHT (Table S5, ESI†), IRON (Table S6, ESI†), LIGHT (Table S7, ESI†), NaCl (Table S8, ESI†), OXIDATIVE-STRESS (Table S9, ESI†), UV-B (Table S10, ESI†). Significantly enriched cellular compartment terms were identified for genes associated with following 9 different stress signals: ABA (Table S11, ESI†), COLD (Table S12, ESI†), COLD-DROUGHT-SALT (Table S13, ESI†), DEHYDRATION (Table S14, ESI†), DROUGHT (Table S15, ESI†), IRON (Table S16, ESI†), LIGHT (Table S17, ESI†), NaCl (Table S18, ESI†) and OXIDATIVE-STRESS (Table S19, ESI†). Significantly enriched molecular function terms were identified for genes associated with following 7 different stress signals: ABA (Table S20, ESI†), COLD (Table S21, ESI†), COLD-DROUGHT-SALT (Table S22, ESI†), DROUGHT (Table S23, ESI†), LIGHT (Table S24, ESI†), NaCl (Table S25, ESI†) and OXIDATIVE-STRESS (Table S26, ESI†). A total of 476 terms with 219 unique terms under the biological process category were significant. A total of 259 terms with 87 unique terms under the cellular compartment category were significantly enriched. A total of 56 terms with 48 unique terms under the molecular function category were enriched in the curated gene lists. Shared and specific biological process terms associated with 10 different stress signals (Fig. 4), cellular compartment terms associated with nine different stress signals (Fig. 5) and molecular



**Fig. 3** Reduction and visualization of highly populated list of terms derived from GO term enrichments. (a) ABA: 112 “biological process” terms visualized as a tree map; (b) LIGHT: 56 cellular component terms and (c) OXIDATIVE-STRESS: 23 molecular function terms visualized as graph where nodes are GO terms and edges are relationships between terms derived from GO diacyclic graph and visualized using Cytoscape. Size of squares or circles indicates inclusion of similar terms using semantic similarity algorithm (Resnik, normalized method).

functions terms associated with seven different stress signals were visualized using Gephi (Fig. 6).

### Protein domain enrichment analysis

Conserved protein domains encoded in genes responsive to various abiotic stress signals were derived from protein domain enrichment analyses using Pfam annotations. A total of 37 unique protein domains were associated with the following 9 different stress signals: ABA (Table S27, ESI†), COLD (Table S28, ESI†), COLD–DROUGHT–SALT (Table S29, ESI†), DEHYDRATION (Table S30, ESI†), DROUGHT (Table S31, ESI†), LIGHT (Table S32, ESI†), NaCl (Table S33, ESI†), OXIDATIVE-STRESS (Table S34, ESI†) and WOUNDING (Table S35, ESI†). A circular representation of 37 protein domains and stress signals were created using Circos<sup>90</sup> to visualize shared and stress specific protein domains (Fig. 7).

### Biological pathway enrichment analysis

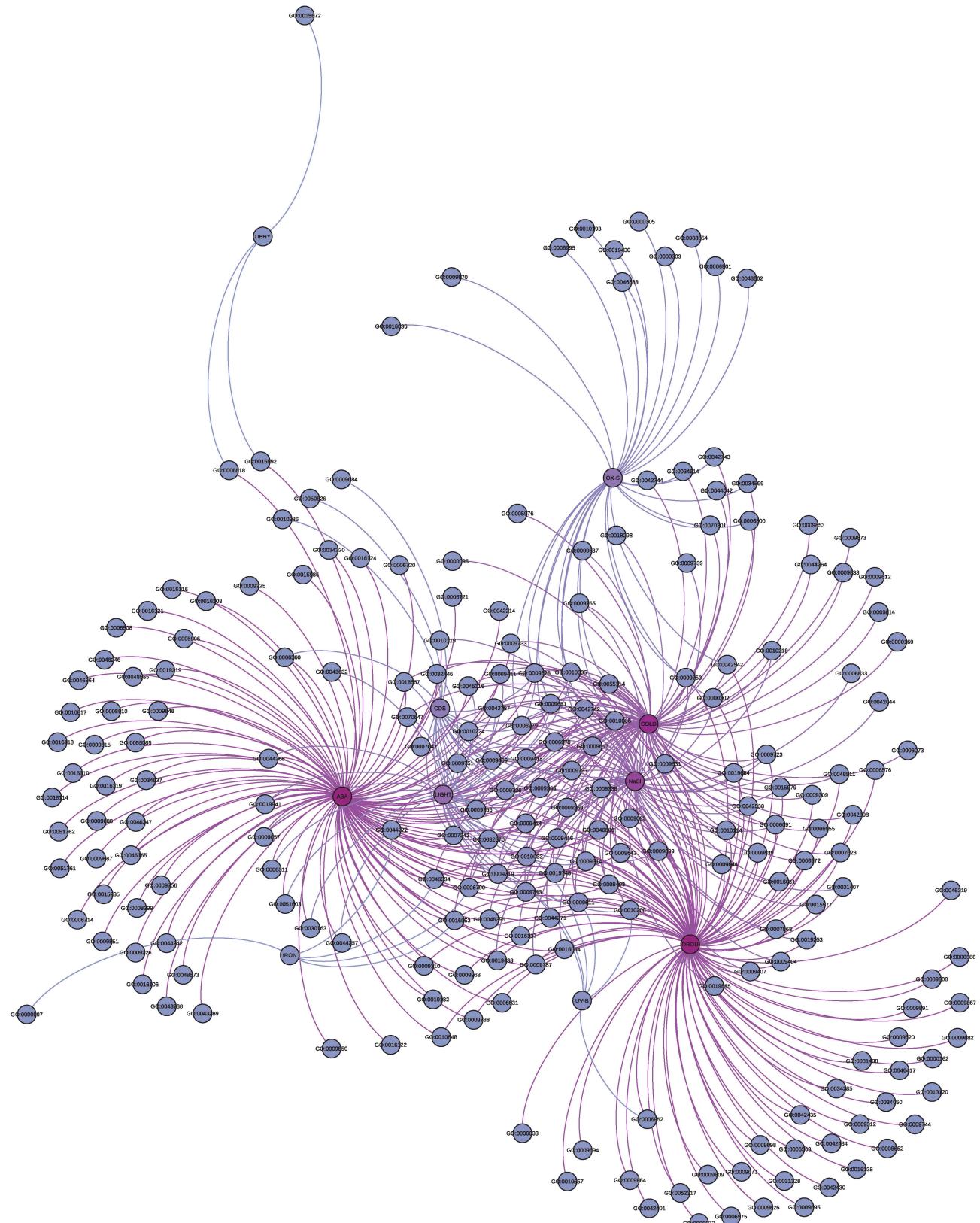
Pathway enrichment analysis revealed association of nine different KEGG pathways and genes responsive to following six stress signals: COLD–DROUGHT–SALT, DROUGHT, NaCl, OXIDATIVE-STRESS and UV-B. To understand the class of pathways, we mapped the KEGG pathway identifiers to BRITE hierarchy and noted that eight out of nine pathways were associated with metabolism. Pathways that mediate metabolism of lipids (fatty acid metabolism pathway, arachidonic acid metabolism pathway and alpha-linolenic acid metabolism pathway), energy (pathways of photosynthesis, photosynthesis – antenna proteins and carbon fixation in photosynthetic organisms), carbohydrates (ascorbate and aldarate metabolism)

and other amino acids (glutathione metabolism pathway) were found to be significant. Statistically significant KEGG pathways after multiple testing correction and associated stress signals are summarized in Table 2.

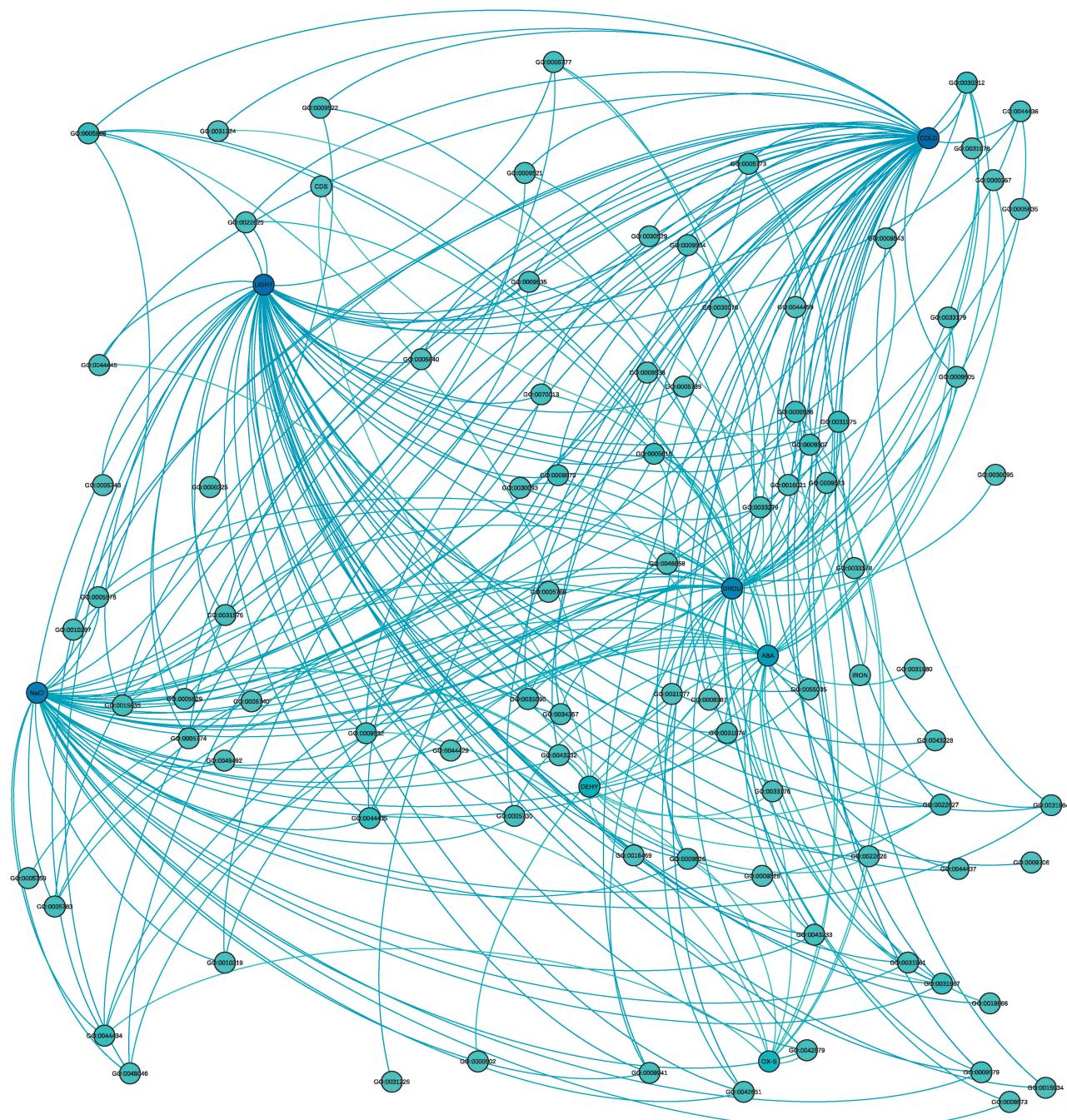
### Phenomic enrichment analysis using plant anatomy and temporal annotations

Phenotype enrichment analyses revealed significantly enriched anatomy and temporal annotations associated with genes responsive to various abiotic stresses. The anatomy subset of PO was significantly enriched among genes responsive to the following 10 stress signals: ABA (Table S36, ESI†), COLD (Table S37, ESI†), COLD–DROUGHT–SALT (Table S38, ESI†), DEHYDRATION (Table S39, ESI†), DROUGHT (Table S40, ESI†), IRON (Table S41, ESI†), LIGHT (Table S42, ESI†), NaCl (Table S43, ESI†), OXIDATIVE-STRESS (Table S44, ESI†) and UV-B (Table S45, ESI†). The temporal subset of PO was significantly enriched among genes responsive to following 8 stress signals: ABA (Table S46, ESI†), COLD (Table S47, ESI†), DEHYDRATION (Table S48, ESI†), DROUGHT (Table S49, ESI†), IRON (Table S50, ESI†), LIGHT (Table S51, ESI†), NaCl (Table S52, ESI†) and OXIDATIVE-STRESS (Table S53, ESI†). PO anatomy (Fig. 8) and temporal (Fig. 9) terms enriched among gene lists shared by different stress signals are visualized using Gephi.

The multi-step enrichment analysis revealed both shared and distinct functional roles, conserved domains, biological pathways and phenotypic characteristics mediated stress-responsive genes. A table summarizing total number of enriched GO terms, KEGG pathways, Pfam domains and PO terms under anatomy and temporal category is provided (Table 3). Individual annotations



**Fig. 4** Modular architecture of shared and specific GO terms (biological process category) associated with 10 different stress signals (ABA, COLD, CDS, DEHYD, DROUGHT, IRON, LIGHT, NaCl, OX-S, UV-B). Edges (represented as curved lines) indicate relationship of a biological process term (only identifier is shown;  $P < 0.05$ ) derived from genes associated with a stress-signal. Nodes are colored by degrees of shared biological process terms (nodes with lower degree will have lighter color compared to nodes with higher degree.)



**Fig. 5** Modular architecture of shared and specific GO terms (cellular compartment category) and abiotic stress signals. Edges (represented as curved lines) indicate relationship of a cellular compartment terms (only identifier is shown;  $P < 0.05$ ) derived from genes associated with a stress-signal. Nodes are colored by degrees of shared cellular compartment terms (nodes with lower degree will have lighter color compared to nodes with higher degree).

pertaining to genes associated with stress signals are provided in ESI† (see Tables S1–S53).

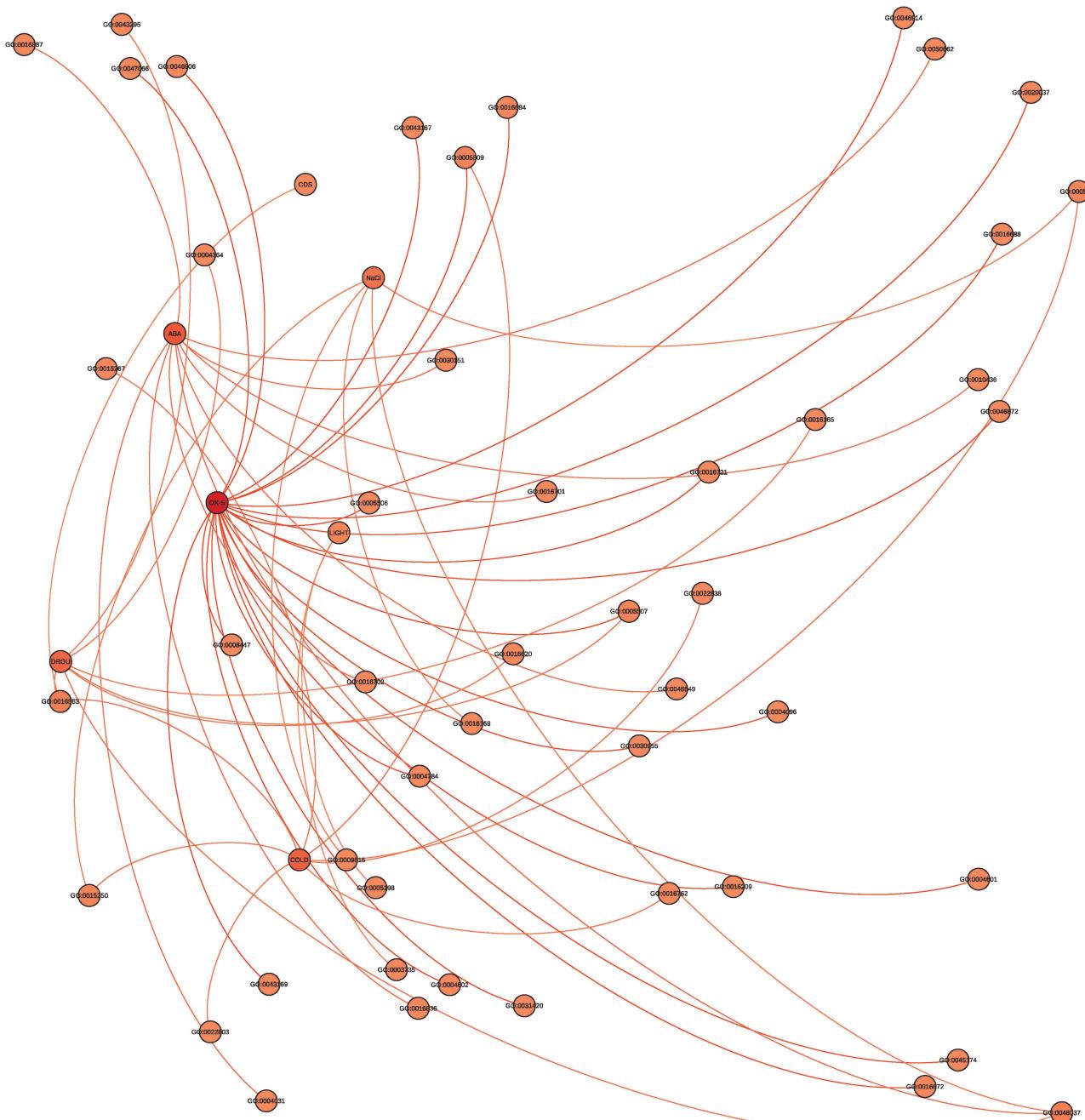
#### Structure-function annotation of genes responsive to multiple stress signals

Results from the structure–function prediction pipeline used to annotate the structure–function properties of the five genes are summarized in Table S54 (ESI†). Downstream biochemical and

experimental studies could help to identify the specific role of genes associated with multiple stresses could help to understand abiotic stress response in *Arabidopsis thaliana*.

#### Network-based enrichment analysis of AT1G68440

To understand the functional role of AT1G68440, we retrieved known and predicted interacting partners from STRING (version 9.05). The confidence score was set to 0.150 to gather



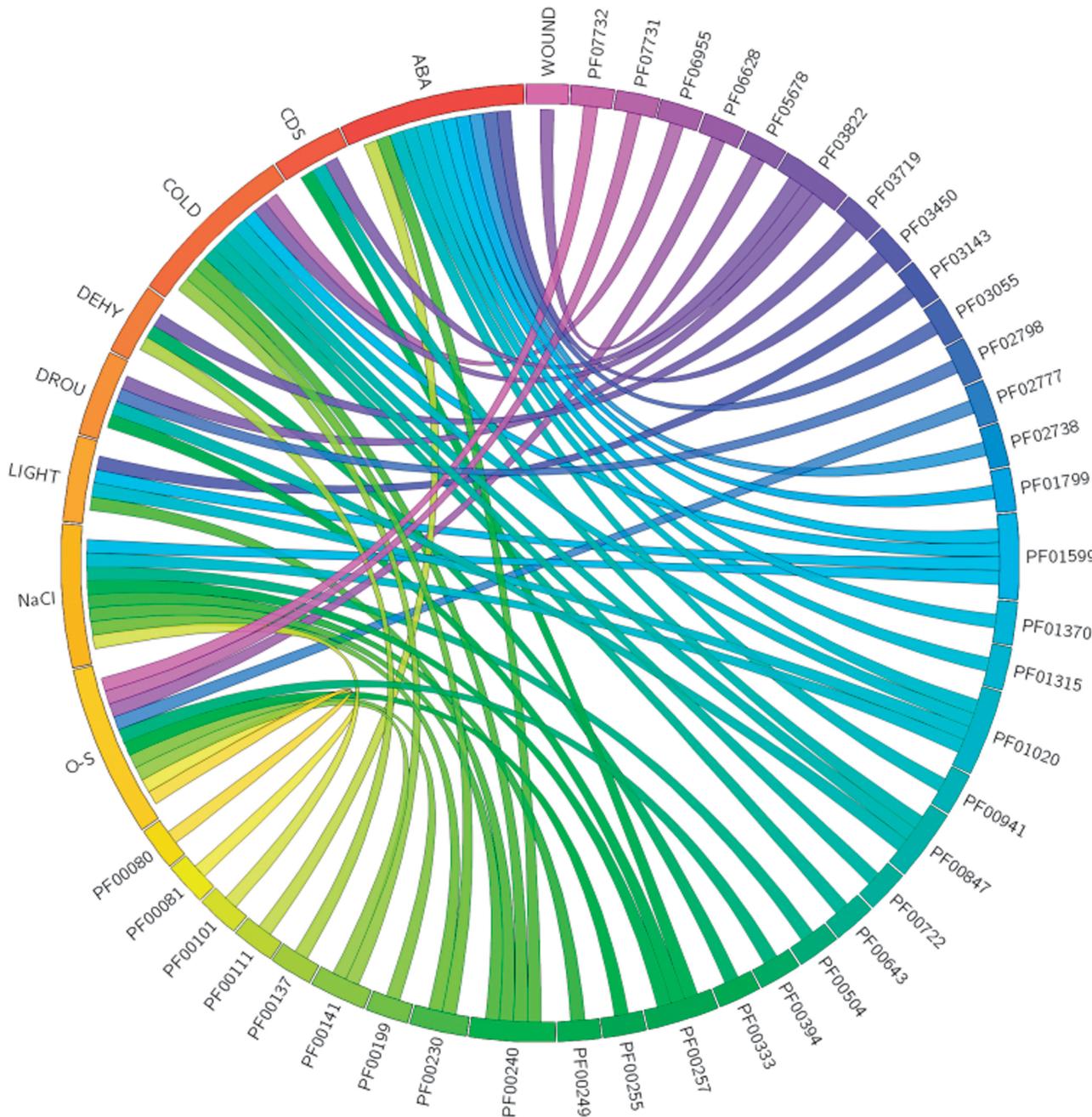
**Fig. 6** Modular architecture of shared and specific GO terms (molecular function category) and abiotic stress signals. Edges (represented as curved lines) indicate relationship of a molecular function term (only identifier is shown;  $P < 0.05$ ) derived from genes associated with a stress-signal. Nodes are colored by degrees of shared molecular function terms (nodes with lower degree will have lighter color compared to nodes with higher degree).

a wide array of known or predicted interacting partners and not more than 50 interactors were retrieved. Using STRING enrichment routine, we tested whether the network is enriched with genes involved in abiotic or biotic related ontology terms. We found that three GO terms related to stress response, tolerance or adaptation were statistically significant (Bonferroni corrected;  $P < 0.05$ ). Three biological process terms: response to stress (GO:0006950,  $P = 9.17 \times 10^{-4}$ ); response to chemical stimulus (GO:0042221,  $P = 1.93 \times 10^{-2}$ ) and response to stimulus

(GO:0050896,  $P = 2.38 \times 10^{-2}$ ) were associated with 21 (Fig. 11), 19 and 25 genes in the interactome. These results indicate that AT1G68440 could play a vital role as a regulator of several genes involved in stress response in *Arabidopsis thaliana*.

#### Stress signals, stress-responsive genes and biological roles

A brief summary of different stress signals, overexpressed stress-responsive genes and corresponding enrichment analyses results are mentioned here.



**Fig. 7** Shared and distinct relationship between 37 protein domains and eight different abiotic stress signals (abscisic acid (ABA), COLD–DROUGHT–SALT (CDS), COLD, dehydration (DEHY), drought (DROU), light, NaCl and oxidative-stress) visualized using Circos. Legend: PF00080: Sod\_Cu, PF00081: Sod\_Fe\_N, PF00101: RuBisCO\_small, PF00111: Fer2, PF00137: ATP-synt\_C, PF00141: peroxidase, PF00199: catalase, PF00230: MIP, PF00240: ubiquitin, PF00249: Myb\_DNA-binding, PF00255: GSHPx, PF00257: dehydrin, PF00333: Ribosomal\_S5, PF00394: Cu-oxidase, PF00504: Chlороa\_b-bind, PF00643: zf-B\_box, PF00722: Glyco\_hydro\_16, PF00847: AP2, PF00941: FAD\_binding\_5, PF01020: Ribosomal\_L40e, PF01315: Ald\_Xan\_dh\_C, PF01370: epimerase, PF01599: Ribosomal\_S27, PF01799: Fer2\_2, PF02738: Ald\_Xan\_dh\_C2, PF02777: Sod\_Fe\_C, PF02798: GST\_N, PF03055: RPE65, PF03143: GTP\_EFTU\_D3, PF03450: CO\_deh\_flav\_C, PF03719: Ribosomal\_S5\_C, PF03822: NAF, PF05678: VQ, PF06628: Catalase-rel, PF06955: XET\_C, PF07731: Cu-oxidase\_2, PF07732: Cu-oxidase\_3.

**Abscisic acid.** 700 genes were identified to be associated with ABA. GO Term enrichment analysis suggested 112 terms associated with the biological process category, 28 terms were associated with the cellular compartment category and 10 terms were enriched under molecular functions. 11 protein domains were associated with upregulated genes respond to ABA. The genes also influence the plant phenotype in multiple levels of

plant structure developmental stage and plant anatomy. No KEGG pathways were associated with genes perturbed due to ABA.

**Aluminum, heat and osmotic stress.** Lists of 25, 64 and 59 genes were identified from the curation and literature mining for aluminum, heat and osmotic stresses, respectively. Functional enrichment, pathway analysis or PO analysis did not

**Table 2** KEGG pathways associated abiotic stress signals

KEGG ID: pathway name	Class/KEGG BRITE hierarchy	P-value <sup>a</sup>				
		CDS	DROUGHT	NaCl	OXIDATIVE	UV-B
ath00053: Ascorbate and aldarate metabolism	Metabolism; carbohydrate metabolism	—	—	—	$4.8 \times 10^{-14}$	—
ath00071: Fatty acid metabolism	Metabolism; lipid metabolism	—	$2.3 \times 10^{-2}$	—	—	—
ath00195: Photosynthesis	Metabolism; energy metabolism	—	—	$4.4 \times 10^{-2}$	—	—
ath00196: Photosynthesis – antenna proteins	Metabolism; energy metabolism	—	—	$2.6 \times 10^{-5}$	—	—
ath00480: Glutathione metabolism	Metabolism; metabolism of other amino acids	—	—	—	$1.2 \times 10^{-18}$	—
ath00590: Arachidonic acid metabolism	Metabolism; lipid metabolism	—	—	—	$5.4 \times 10^{-11}$	—
ath00592: Alpha-linolenic acid metabolism	Metabolism; lipid metabolism	—	—	—	—	$4.6 \times 10^{-2}$
ath00710: Carbon fixation in photosynthetic organisms	Metabolism; energy metabolism	—	$7.7 \times 10^{-4}$	—	—	—
ath04650: Natural killer cell mediated cytotoxicity	Organismal systems; immune system	$2.2 \times 10^{-2}$	—	—	—	—

Pathways mediated by abiotic stress-responsive genes. Pathways with a *p*-value  $\leq 0.05$  after multiple testing correction using the Bonferroni method are considered as significant. DS: COLD–DROUGHT–SALT stress; —: pathway not associated with genes upregulated as responsive to the abiotic stress signal or no significant enrichment for a pathway. <sup>a</sup> Bonferroni adjusted *P*-values.

reveal any terms with significant *p*-values with these gene lists (after the Bonferroni correction). This could be due to low number of genes identified as responsive to these signals or due to unavailability of annotations for the genes.

**Cold.** Largest number of stress-responsive genes was identified from data curation and literature survey as responsive to cold in our analysis. GO Term enrichment analysis suggested 86 terms associated with the biological process category and 8 terms associated with molecular functions were significant. No pathway or protein domains were associated with the upregulated genes respond to cold. The genes were associated with 70 anatomy and 13 temporal PO terms.

**Cold–drought–salt.** This stress signal is a combination of three different abiotic signals and 41 genes were found to be responsive with  $\geq 2.5$  fold. Lower number of genes extracted from this study indicates that combination of stress-signals may activate different sets of stress-responsive genes compared to single stress-signals. Genes associated with this combination signals had significantly enriched GO terms under all three categories. A KEGG pathway (natural killer cell mediated cytotoxicity pathway, KEGG ID: ath04650) and 3 Pfam domains (dehydrin,<sup>96</sup> Apetala 2 (AP2)<sup>97</sup> and NAF<sup>98</sup> domains) were also enriched among the genes. Only the PO anatomy subset had significant association, no temporal terms were significant.

**Dehydration.** A total of 93 dehydration responsive genes were identified from integrated curation and literature search. Significantly enriched annotations were obtained from all analyses except molecular function and KEGG pathways.

**Drought, NaCl and oxidative-stress.** Drought (815 genes), NaCl (1040 genes) and oxidative stress were the three different abiotic stresses for which significant annotations were obtained from all analyses. This could also be attributed to the large number of genes associated with these abiotic stress signals and the quality of annotations available for cold, NaCl and oxidative-stress-responsive genes.

**Iron.** Curation using iron retrieved 147 genes as stress response. Significant enrichments were obtained from all

annotations except molecular functions, KEGG pathways and Pfam domains.

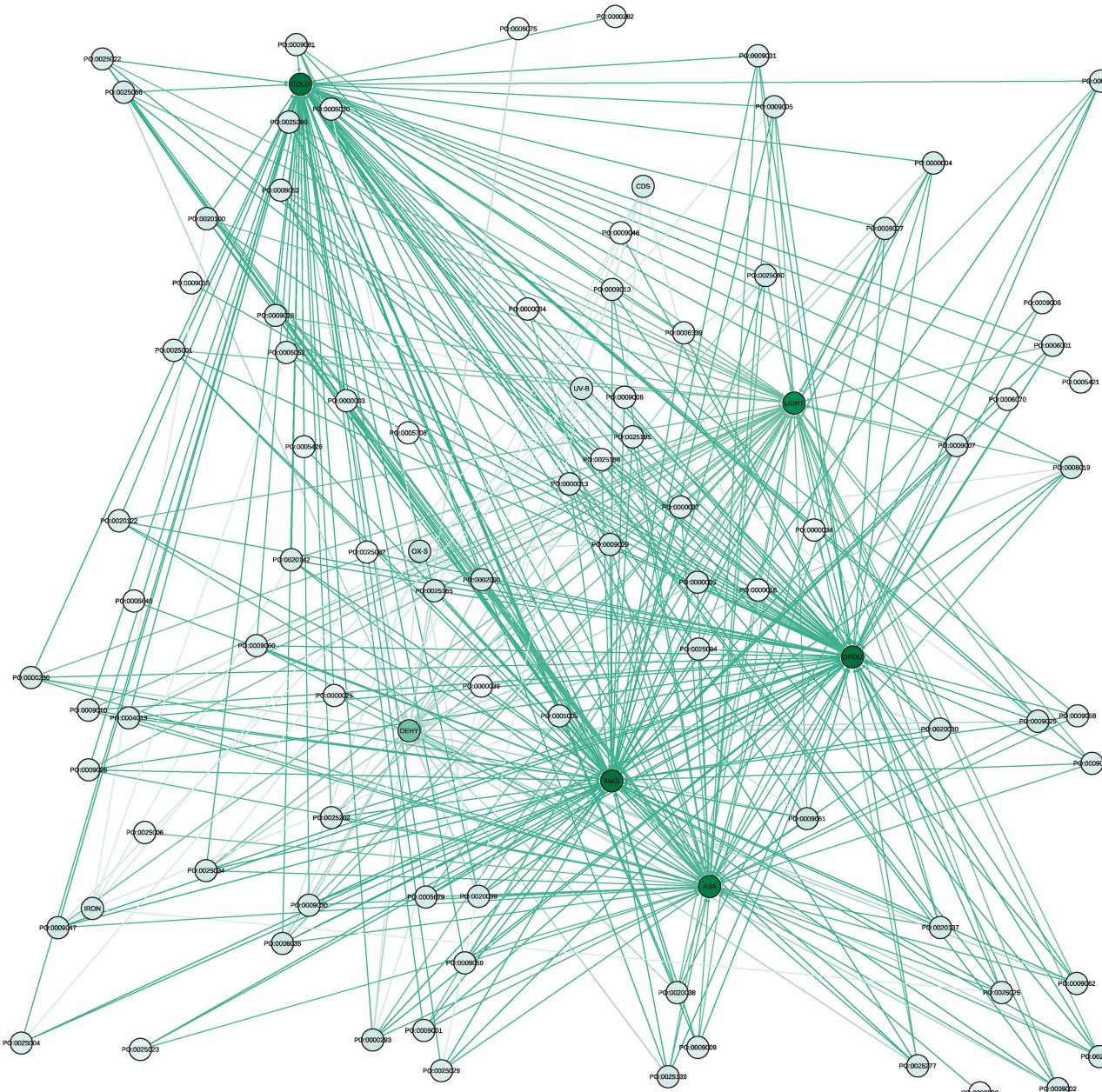
**Light.** 674 genes were identified as responsive to light in the curation step. All categories except KEGG pathway annotations had significant associations with the gene lists.

**UV-B.** 122 genes were identified as responsive to UV-B from literature survey and curation step. These genes had significantly enriched annotations under the biological process subset of GO and the anatomy subset of PO.

**Wounding.** With a list of 50 genes responsive to wounding, a single Pfam domain annotation (VQ motif) was found significant after correction using the Bonferroni method. VQ motifs are abundant in plant proteins and mostly composed of low complexity regions and thus may involve in a position dependent role including protein–protein interactions.<sup>99,100</sup>

#### Constitutive stress responsive genes in *Arabidopsis thaliana*

We found a subset of 1284 upregulated genes to be responsive to more than one stress signal (see ESI† S2). We hypothesise that genes that are responsive to multiple stress signals can be considered as “constitutive stress-responsive genes” and may play a significant role in stress response and tolerance. Understanding the genetic, functional and biochemical roles of these genes could help to unravel the stress response code in plants. Among the large list of genes, we focused on top five genes. Among the top five genes, AT1G68440 was differentially upregulated as responsive to seven different stress signals like ABA, drought, heat, light, NaCl, osmotic stress and wounding. This gene was recently reported to be involved in key metabolic processes like the cellular modified amino acid biosynthetic process (GO:0042398), phenylpropanoid metabolic process (GO:0009698) and polyamine catabolic process (GO:0006598).<sup>95</sup> Both phenylpropanoids<sup>101–104</sup> and polyamines<sup>105–110</sup> play key roles in stress response and tolerance in plants, including *Arabidopsis thaliana*. Hitherto, no studies have reported the constitutive stress-responsive role of AT1G68440 as an important constitutive gene responsive to diverse types of stresses. These biological terms have common parent biological process



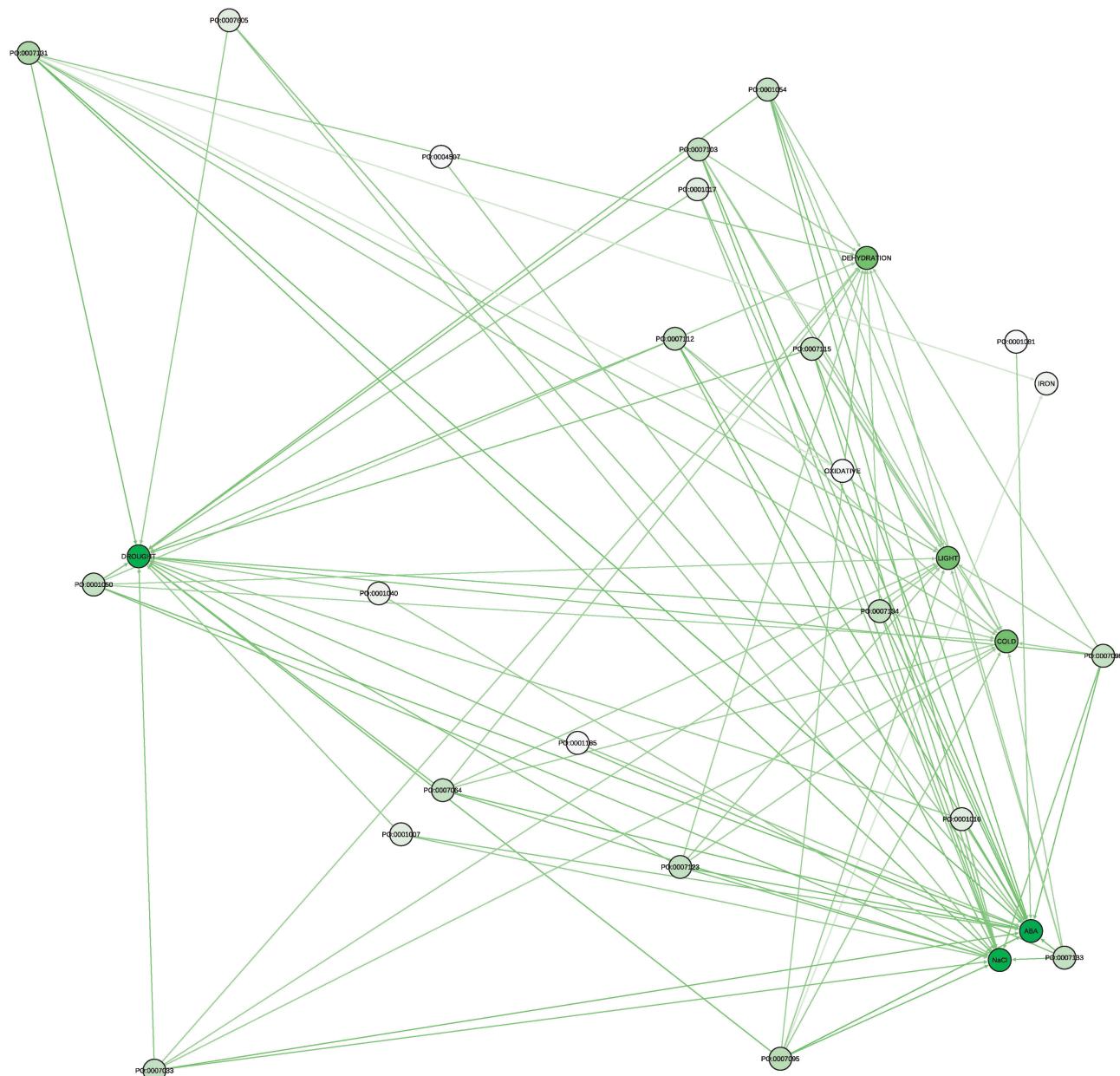
**Fig. 8** Relationship between PO terms (anatomy) and abiotic stress signals derived from PO term enrichment analysis using abiotic stress-responsive genes. Edges indicate relationship of a PO anatomy term (only identifier is shown;  $P < 0.05$ ) derived from genes associated with a stress-signal. Nodes are colored by degrees of shared PO anatomy terms (nodes with lower degree will have lighter color compared to nodes with higher degree).

term “cellular metabolic process” and semantic similarity of the terms in biological process subset of GO is computed using G-SESAME<sup>111</sup> (Fig. 10) to understand the semantic variance of GO terms. The three terms were highly co-occurring based on non-IEA annotations compiled in Quick-GO database.<sup>112</sup> This indicates the functional concordance of AT1G68440 in playing a key role in plant stress response. While structure-function analysis of constitutive stress-responsive genes did not reveal conclusive results, we performed network-based enrichment analysis on AT1G68440 and it revealed that this gene interacts

directly or indirectly with different genes involved in stress response in *Arabidopsis thaliana*.

#### Concordance of data mining results with existing findings

Our comparative meta-analysis using 14 stress signals and 3091 differentially upregulated genes is a primary attempt to characterize stress-responsive functions, pathways, domain and phenomic features. While we focused on major findings from the study as the important aspect of constitutive stress-responsive genes in *Arabidopsis thaliana*, our results are also



**Fig. 9** Relationship between PO terms (temporal) and abiotic stress signals derived from PO term enrichment analysis using abiotic stress-responsive genes. Edges indicate relationship of a PO temporal terms (only the identifier is shown;  $P < 0.05$ ) derived from genes associated with a stress-signal. Nodes are colored by degrees of shared PO temporal terms (nodes with lower degree will have lighter color compared to nodes with higher degree).

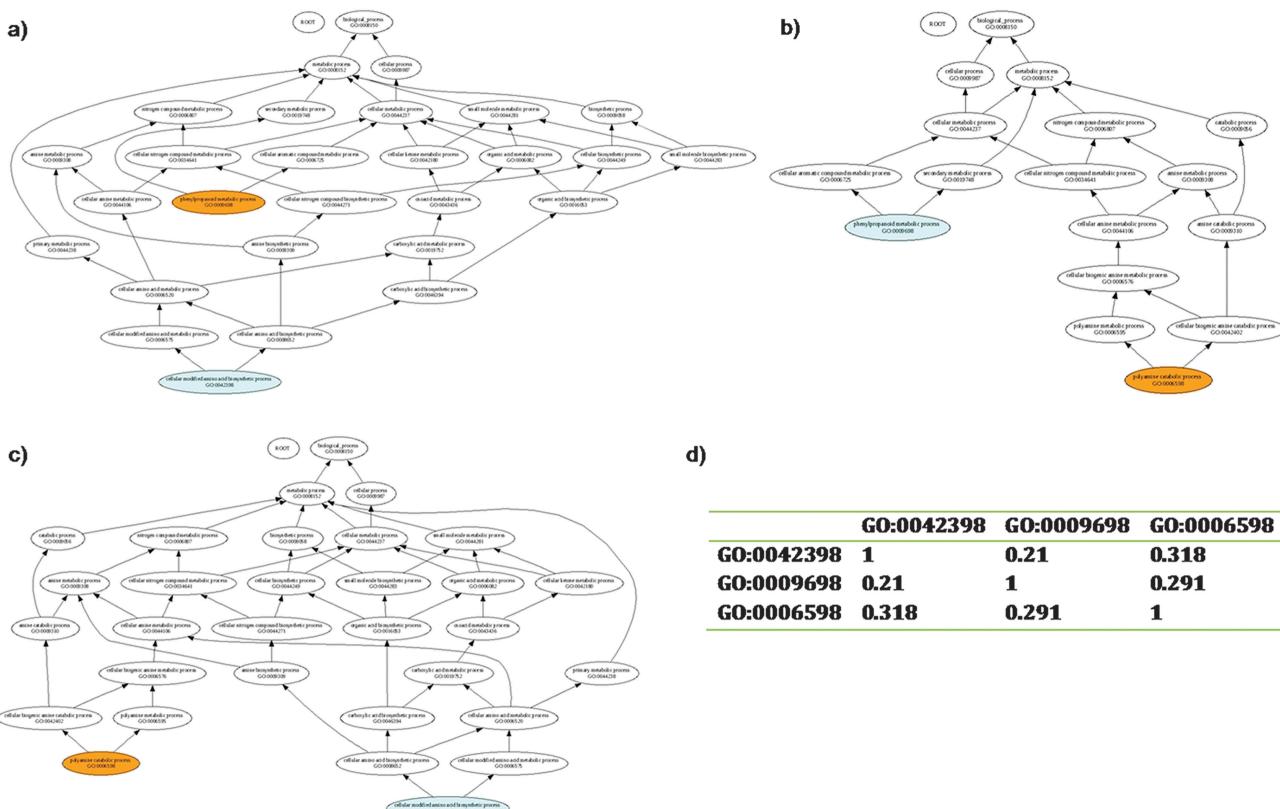
concordant with previous reports on abiotic and biotic stress responsive studies. For example, stress-activated mitogen activated protein kinases (also known as ATMPKs) are considered as key regulators of abiotic stress response in *Arabidopsis thaliana*.<sup>113,114</sup> To illustrate concordance of our finding with respect to the existing literature, we compiled MAPKs curated from the literature. We found eight ATMPKs or related gene family members (MAP3KA, MPK7, MAPKKK14, MEKK1, MEK1, MAPKKK16 and MKK2) to be responsive to various stress signals. Five of these (MPK7, MAPKKK14, MEKK1, MEK1 and MEK1) were associated with cold stress and two genes (MEKK1 and MAPKKK16) were associated with salt-induced stress.

GO term enrichment is a routine analytical approach used to understand functional cohesiveness of differentially expressed genes.<sup>115,116</sup> To the best of our knowledge, no previous reports have provided an extensive functional, pathway or phenomic terms associated with 14 different stress signals. The roles of metabolomic pathways, in response to multiple stress signals, have been experimentally validated earlier.<sup>117</sup> However, the role of specific pathways as responsive to different signals was not extensively reported. The role of members of the dehydrin gene family in abiotic stress response and tolerance was experimentally verified in previous studies.<sup>96,118</sup> Large-scale study on specific associations between protein domains and various

**Table 3** Abiotic stress signals, responsive genes and significant annotations identified from enrichment analysis pipeline

Stress signals	Genes	Gene Ontology terms				KEGG pathways	Pfam domains	Plant Ontology (PO) terms	
		BP	CC	MF				Anatomy	Temporal
ABA	700	112	28	10	—	—	11	65	20
Aluminum	25	—	—	—	—	—	—	—	—
Cold	1078	86	—	8	—	—	10	70	13
CDS <sup>a</sup>	41	22	2	1	1	—	3	7	—
Dehydration	93	3	14	—	—	—	3	30	13
Drought	815	110	43	7	2	—	4	71	18
Heat	64	—	—	—	—	—	—	—	—
Iron	147	5	2	—	—	—	—	7	2
Light	674	37	56	2	—	—	1	57	13
NaCl	1040	66	46	5	2	—	8	72	18
Osmotic-stress	59	—	—	—	—	—	—	—	—
Oxidative-stress	40	29	10	23	3	—	10	12	1
UV-B	122	6	—	—	1	—	—	9	—
Wounding	50	—	—	—	—	—	1	—	—

<sup>a</sup> CDS: Cold-drought-salt; —: no significant association or  $P > 0.1$ .



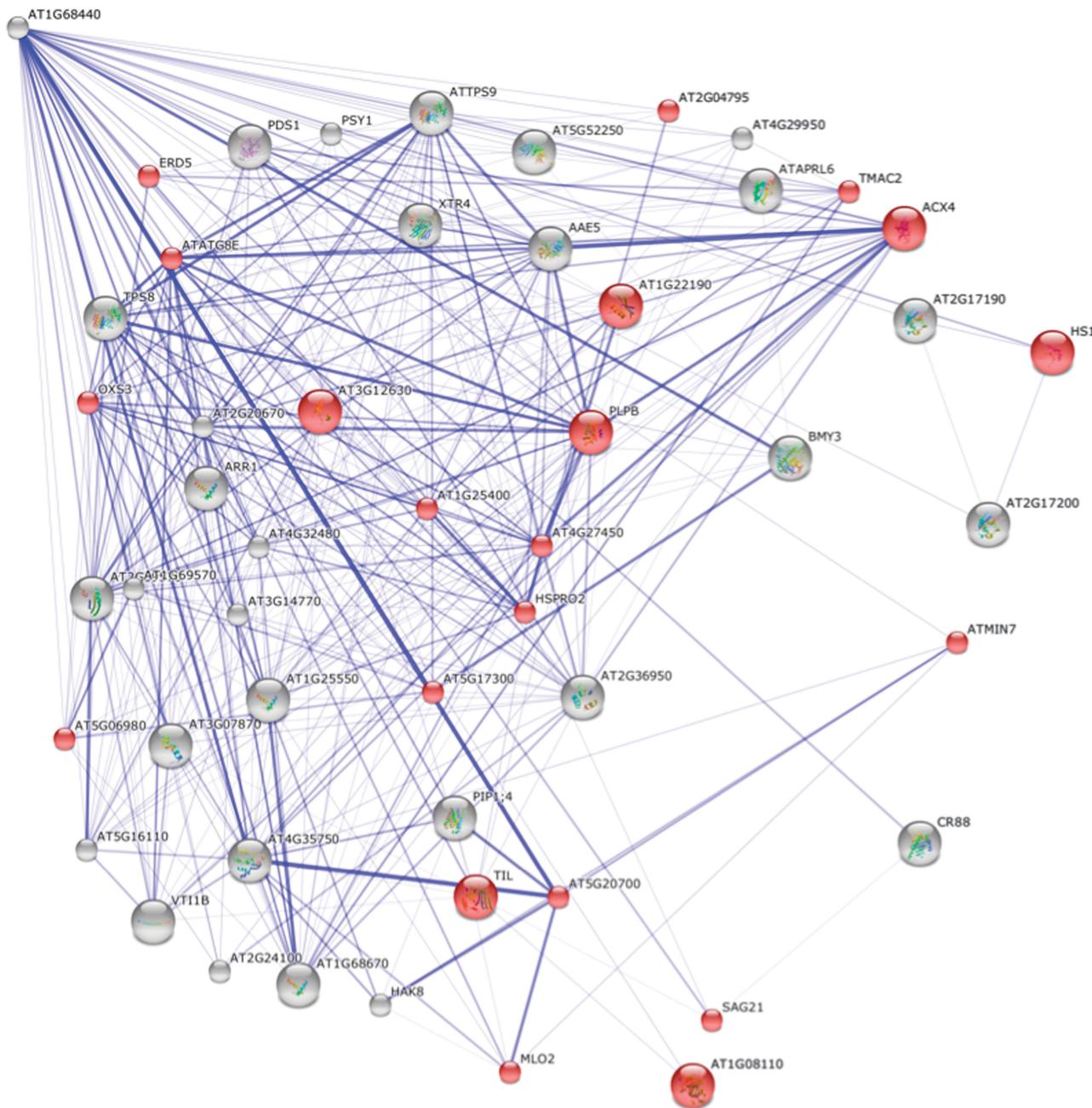
**Fig. 10** Semantic similarity between GO terms (cellular modified amino acid biosynthetic process (GO:0042398), phenylpropanoid metabolic process (GO:0009698) and polyamine catabolic process (GO:0006598)) associated with AT1G68440. Semantic distances were computed using G-SESAME between (a) GO:0042398 and GO:0006598; (b) GO:0006598 and GO:0009698; (c) GO:0009698 and GO:0042398; (d) matrix of semantic distance between different GO terms.

stress signals were not reported before. From a phenomic perspective, several studies have experimentally characterized stress-related phenotypes in plant model systems including *Arabidopsis thaliana*, but phenomic characteristics corresponding to stress responsive genes using Plant Ontology and enrichment analysis have not been reported before. In a similar way, our results obtain high concordance with the existing

literature reports on stress response, tolerance and adaptation in *Arabidopsis thaliana* and also provide several novel insights.

## Discussion

Abiotic stress responses are attributed to various genes and pathways and previous studies suggest that abiotic



**Fig. 11** Visualization of first-degree interactome of AT1G68440 retrieved from STRING (v 9.05). Red nodes are genes annotated with the GO term “response to stress (GO:0006950,  $P = 9.17 \times 10^{-4}$ )” in the first-degree interactome of AT1G68440, remaining nodes are colored in grey. Nodes with a structural logo indicate that protein structural data are available for the node. Edge thickness indicates the STRING score (thinner edges have low score and thicker edges have high score to associate interaction between two connecting nodes).

stress-responsive pathways are highly specialized for individual stress signals. In this work we attempted to characterize the functional basis of various stress responses by performing functional enrichment analysis using annotations from GO, KEGG, Pfam and PO. In the post-genome era, multiple high-throughput techniques has been used to study abiotic stress, stress response and stress tolerance in plants. To the best of our knowledge, our study is the first to combine 14 different signals and used a collective approach to identify functional repertoire,

biological pathways and phenotypic aspects of abiotic stress response in plants. Our multi-step enrichment analysis helped to learn different aspects of abiotic stress-responsive genes in *Arabidopsis thaliana*. GO terms helped to understand the basic set of biological processes, cellular component and molecular function mediated by genes. Protein domain enrichment helped to find evolutionarily conserved protein domains associated with the stress-responsive genes. Pathway enrichment analysis helped to find the pathways associated with genes

perturbed by different stress signals. Enrichment analysis using PO terms enabled the association of genes with plant structures, growth and developmental stages and helped us to understand the phenotypic role of genes response to abiotic stress signals.

Enrichment analysis provides a rapid framework for inferring the biological significance of the gene list from experimental and computational studies. Nevertheless, such high-throughput analysis approaches could be impacted due to false positive associations.<sup>116</sup> We also noted that multiple testing correction steps used to control false positives in our study reduced terms in different genomic and phenomic enrichment analyses we performed in the study (Fig. 12). To deal with the false-positive stress-responsive gene list and GO terms, pathways, protein domains or PO terms, we have reported the results after multiple-testing correction.

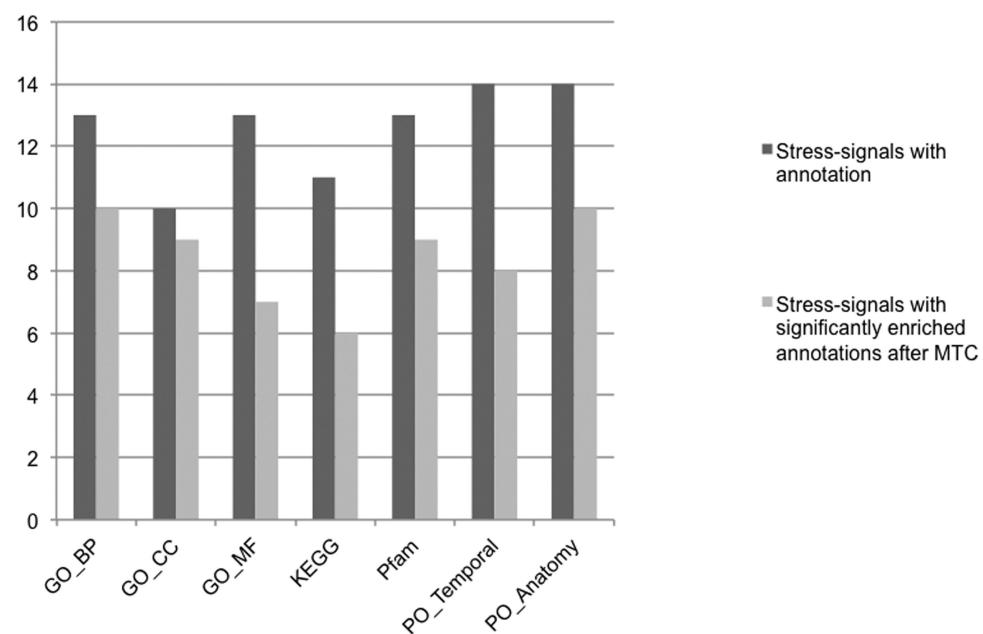
GO terms shared by genes responding to multiple stress signals under the biological process category were all related to response processes. GO biological process terms provided similar and distinct functional aspects of stress-responsive genes. Terms like “response to organic substance”, “response to endogenous stimulus”, “response to salt stress”, “response to abiotic stimulus”, “response to oxidative stress”, “response to osmotic stress”, “defense response to bacterium”, “response to metal ions”, “response to inorganic substance” and “response to carbohydrate stimulus” were significantly enriched in genes associated with multiple stresses. Cellular colocalization patterns of stress-responsive genes were obtained from GO cellular component terms. Cellular compartment terms like “envelope”, “organelle envelope”, “plastid part”, “chloroplast part”, “external encapsulating structure”, “plasma membrane”, “cell wall”, “chloroplast stroma” and “plastid” were terms shared by multiple stresses. Genes implicated in

molecular functions including “transcription activator activity”, “cofactor binding”, “copper ion binding”, “calcium ion binding” “water channel activity” and “water transport activity” were associated with more than one stress signals.

We identified nine different molecular pathways enriched among the gene lists. Irrespective of pathway enrichment analysis using large genes only a small of significant associations were observed. KEGG based pathway analysis retrieved pathways for 11 different signals, but after the Bonferroni correction only six stress signals retained significant pathway annotations ( $n = 9$ ). Out of the 9 pathways; all pathways except one was classified to be part of “Metabolism” under the KEGG BRITE pathway classification system. This implicates that diverse metabolic pathways are activated as responsive to stresses.

Genes that encode protein domains, like NAF domains (domain encoded in plant-specific subgroup of serine-threonine protein kinases),<sup>98</sup> ribosomal protein S27a domain,<sup>119</sup> ribosomal L40e domain, AP2 domain (a member of a large family of transcription factors and involved in the ABC model of flower development<sup>97</sup>), dehydrin domains (associated with protection of membranes from damage due to ABA, salt, low temperatures and drought stresses<sup>96,120</sup>), ubiquitin domain, major intrinsic protein domains (found in plant tonoplast intrinsic proteins), peroxidase domain (found in heme-containing enzymes that use hydrogen peroxide as the electron acceptor to catalyse a number of oxidative reactions), were associated with genes responsive to more than one stress signals.

Phenotype enrichment analysis using PO anatomy terms enabled the characterization of the plant structure regions where abiotic stress-responsive genes expressed. Significantly enriched PO anatomy terms like “androecium”, “stamen”, “shoot epidermal cell”, “shoot epidermis”, “stomatal complex”,



**Fig. 12** Total number of annotations retrieved from initial enrichment analysis and annotations retained after multiple testing correction.

"guard cell", "cotyledon", "stem", "leaf" etc. were associated with more than two different stress signals. Growth and developmental stages at which stress-responsive genes expressed were identified using PO temporal terms associated with the gene lists. PO temporal terms like "seedling development stage", "LP.08 eight leaves visible", "A vegetative growth", "leaf production", "LP.06 six leaves visible", "LP.04 four leaves visible", "1 main shoot growth", "LP.10 ten leaves visible" and "LP.02 two leaves visible" were enriched among gene lists for two or more stress signals. The shared phenotypic trend observed among genes responsive to multiple stress signals suggests that irrespective of diversity in stress-responsive pathways, target phenotypes including plant structure or growth and development stages could be common.

Structure-function annotation approach provided several clues to understand functional mechanisms mediated by constitutive genes responsive to multiple stresses. Expanding similar computational approaches to the full list of multi-stress-responsive genes could help to understand specific and collective roles of these genes in stress response and tolerance mechanisms.

We found a large collection of constitutive gene responsive to stress signals, the top gene AT1G68440 involved in important metabolic processes like phenylpropanoid metabolic process (GO:0009698) and polyamine catabolic process (GO:0006598) was found to be responsive to perturbations from seven different stress signals. Functional studies targeting the genes responsive to multiple stress signals could help to delineate components shared and specifically involved in stress pathways and regulatory mechanisms.

In summary, we performed an extensive annotation driven enrichment analysis to characterize shared and distinct functional repertoire, pathways and phenotypes enriched in genes respond to 14 different abiotic stresses. We found various annotation terms that are shared and specific to multiple to abiotic stress signals. We also found a hypothetical gene to be differentially upregulated as responsive to multiple stresses. Full list of genes responsive to various signals, GO biological process, cellular compartment, molecular function, PO anatomy and PO temporal identifiers and corresponding stress signals are provided in ESI,<sup>†</sup> S2.

We have considered a stringent fold-change threshold of 2.5 in order to capture differentially upregulated genes responding to different stress-signals. Stress responsive genes expressed below 2.5 fold were excluded in this analysis. Similar to other high-throughput data analyses studies, this stringent expression criteria is a limiting factor of our study. Also, we used Bonferroni as a method for multiple testing corrections. The Bonferroni method is considered as one of the stringent statistical measures that could have led to low number of enriched terms retained after multiple testing corrections. Also the top five constitutive stress responsive genes we identified from our study would need additional experimental validation to understand its role in stress response, adaptation and tolerance.

## Conclusion

Worldwide, the plant productivity is severely affected due to various abiotic stresses and the demand for food is

exponentially growing due to rapid population growth. According to Food and Agriculture Organization of the United Nations (FAO), the World would need >70% food by the year 2050. Developing stress-tolerant varieties of plants to maintain the supply and demand of food crops is crucial for food security. Since plants are sessile, they must adapt to adverse environmental conditions and adapt to abiotic stress via multiple stress-responsive mechanisms to gain optimal growth and yield. Identifying the core set of genes and functional roles and pathways associated with abiotic stresses would help to identify the molecular mechanisms behind abiotic stress, stress response and stress tolerance. Current challenge in dealing with abiotic stress response is mainly in delineating stress-specific genes and pathways. Effective understanding of such key pathways and molecular connections would help to develop plants with traits that confer tolerance to abiotic stresses.

We performed an extensive functional genomics meta-analysis of differentially upregulated genes due to abiotic stresses. We used functional enrichment analyses using GO terms, pathway annotations and found several shared and specific functional cues pertaining to abiotic stress response. Further analysis was performed using Plant Ontology annotation to understand plant-specific phenomic features associated with genes differentially upregulated due to the individual signal. The identification of functional repertoire, biological pathways and plant-specific phenotypes of genes perturbed due to the abiotic stress signal would help to understand stress-tolerance and design crop plants that could adapt to stress conditions. Experimental validation of constitutive genes responding to multiple stress signals, including AT1G68440 and other constitutive stress responsive genes, could provide additional insight into the stress response code and regulatory pathways in *Arabidopsis thaliana*.

## References

- 1 N. J. Atkinson and P. E. Urwin, The interaction of plant biotic and abiotic stresses: from genes to the field, *J. Exp. Bot.*, 2012, **63**(10), 3523–3543.
- 2 E. A. Bray, J. Bailey-Serres and E. Weretilnyk, *Responses to abiotic stresses*, in *Biochemistry and Molecular Biology of Plants*, 2000, pp. 1158–1203.
- 3 V. Chinnusamy, K. Schumaker and J. K. Zhu, Molecular genetic perspectives on cross-talk and specificity in abiotic stress signalling in plants, *J. Exp. Bot.*, 2004, **55**(395), 225–236.
- 4 L. Xiong and J. K. Zhu, Abiotic stress signal transduction in plants: molecular and genetic perspectives, *Physiol. Plant.*, 2001, **112**(2), 152–166.
- 5 D. Howe, M. Costanzo, P. Fey, T. Gojobori, L. Hannick, W. Hide, D. P. Hill, R. Kania, M. Schaeffer and S. St Pierre, et al., Big data: the future of biocuration, *Nature*, 2008, **455**(7209), 47–50.
- 6 X. M. Fernandez-Suarez and E. Birney, Advanced genomic data mining, *PLoS Comput. Biol.*, 2008, **4**(9), e1000121.
- 7 The Gene Ontology: enhancements for 2011. *Nucleic Acids Res.*, 2012, **40** (Database issue), D559–D564.

- 8 W. Huang da, B. T. Sherman and R. A. Lempicki, Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists, *Nucleic Acids Res.*, 2009, **37**(1), 1–13.
- 9 T. Kelder, B. R. Conklin, C. T. Evelo and A. R. Pico, Finding the right questions: exploratory pathway analysis to enhance biological discovery in large datasets, *PLoS Biol.*, 2010, **8**(8), 1–5.
- 10 P. Jaiswal, S. Avraham, K. Ilic, E. A. Kellogg, S. McCouch, A. Pujar, L. Reiser, S. Y. Rhee, M. M. Sachs and M. Schaeffer, *et al.*, Plant Ontology (PO): a Controlled Vocabulary of Plant Structures and Growth Stages, *Comp. Funct. Genomics*, 2005, **6**(7–8), 388–397.
- 11 K. Ilic, E. A. Kellogg, P. Jaiswal, F. Zapata, P. F. Stevens, L. P. Vincent, S. Avraham, L. Reiser, A. Pujar and M. M. Sachs, *et al.*, The plant structure ontology, a unified vocabulary of anatomy and morphology of a flowering plant, *Plant Physiol.*, 2007, **143**(2), 587–599.
- 12 Q. He, A. Berg, Y. Li, C. E. Vallejos and R. Wu, Mapping genes for plant structure, development and evolution: functional mapping meets ontology, *Trends Genet.*, 2010, **26**(1), 39–46.
- 13 S. Avraham, C. W. Tung, K. Ilic, P. Jaiswal, E. A. Kellogg, S. McCouch, A. Pujar, L. Reiser, S. Y. Rhee and M. M. Sachs, *et al.*, The Plant Ontology Database: a community resource for plant structure and developmental stages controlled vocabulary and annotations, *Nucleic Acids Res.*, 2008, **36**(Database issue), D449–D454.
- 14 The Plant Ontology Consortium and plant ontologies, *Comp. Funct. Genomics*, 2002, **3**(2), 137–142.
- 15 R. T. Furbank and M. Tester, Phenomics – technologies to relieve the phenotyping bottleneck, *Trends Plant Sci.*, 2011, **16**(12), 635–644.
- 16 S. J. Roy, E. J. Tucker and M. Tester, Genetic analysis of abiotic stress tolerance in crops, *Curr. Opin. Plant Biol.*, 2011, **14**(3), 232–239.
- 17 A. Ramakrishna and G. A. Ravishankar, Influence of abiotic stress signals on secondary metabolites in plants, *Plant Signaling Behav.*, 2011, **6**(11), 1720–1731.
- 18 N. Sreenivasulu, S. K. Sopory and P. B. Kavi Kishor, Deciphering the regulatory mechanisms of abiotic stress tolerance in plants by genomic approaches, *Gene*, 2007, **388**(1–2), 1–13.
- 19 J. W. Walley and K. Dehesh, Molecular mechanisms regulating rapid stress signaling networks in *Arabidopsis*, *J. Integr. Plant Biol.*, 2010, **52**(4), 354–359.
- 20 W. R. Swindell, The association among gene expression responses to nine abiotic stress treatments in *Arabidopsis thaliana*, *Genetics*, 2006, **174**(4), 1811–1824.
- 21 A. P. Santos, T. Serra, D. D. Figueiredo, P. Barros, T. Lourenco, S. Chander, M. M. Oliveira and N. J. Saibo, Transcription regulation of abiotic stress responses in rice: a combined action of transcription factors and epigenetic mechanisms, *Omics*, 2011, **15**(12), 839–857.
- 22 K. Nakashima, Y. Ito and K. Yamaguchi-Shinozaki, Transcriptional regulatory networks in response to abiotic stresses in *Arabidopsis* and grasses, *Plant Physiol.*, 2009, **149**(1), 88–95.
- 23 A. Matsui, J. Ishida, T. Morosawa, Y. Mochizuki, E. Kaminuma, T. A. Endo, M. Okamoto, E. Nambara, M. Nakajima and M. Kawashima, *et al.* *Arabidopsis* transcriptome analysis under drought, cold, high-salinity and ABA treatment conditions using a tiling array, *Plant Cell Physiol.*, 2008, **49**(8), 1135–1149.
- 24 S. Ma and H. J. Bohnert, Integration of *Arabidopsis thaliana* stress-related transcript profiles, promoter structures, and cell-specific expression, *Genome Biol.*, 2007, **8**(4), R49.
- 25 T. A. Long, H. Tsukagoshi, W. Busch, B. Lahner, D. E. Salt and P. N. Benfey, The bHLH transcription factor POPEYE regulates response to iron deficiency in *Arabidopsis* roots, *Plant Cell*, 2010, **22**(7), 2219–2236.
- 26 J. X. Liu, R. Srivastava, P. Che and S. H. Howell, Salt stress responses in *Arabidopsis* utilize a signal transduction pathway related to endoplasmic reticulum stress signaling, *Plant J.*, 2007, **51**(5), 897–909.
- 27 S. Gonzalez-Perez, J. Gutierrez, F. Garcia-Garcia, D. Osuna, J. Dopazo, O. Lorenzo, J. L. Revuelta and J. B. Arellano, Early transcriptional defense responses in *Arabidopsis* cell suspension culture under high-light conditions, *Plant Physiol.*, 2011, **156**(3), 1439–1456.
- 28 C. Zou, K. Sun, J. D. Mackaluso, A. E. Seddon, R. Jin, M. F. Thomashow and S. H. Shiu, *Cis*-regulatory code of stress-responsive transcription in *Arabidopsis thaliana*, *Proc. Natl. Acad. Sci. U. S. A.*, 2011, **108**(36), 14992–14997.
- 29 M. M. Chaves, J. Flexas and C. Pinheiro, Photosynthesis under drought and salt stress: regulation mechanisms from whole plant to cell, *Ann. Bot.*, 2009, **103**(4), 551–560.
- 30 K. Singh, R. C. Foley and L. Onate-Sanchez, Transcription factors in plant defense and stress responses, *Curr. Opin. Plant Biol.*, 2002, **5**(5), 430–436.
- 31 A. S. Sundar, S. M. Varghese, K. Shameer, N. Karaba, M. Udayakumar and R. Sowdhamini, STIF: identification of stress-upregulated transcription factor binding sites in *Arabidopsis thaliana*, *Bioinformation*, 2008, **2**(10), 431–437.
- 32 K. Shameer, S. Ambika, S. M. Varghese, N. Karaba, M. Udayakumar and R. Sowdhamini, STIFDB-Arabidopsis Stress Responsive Transcription Factor DataBase, *Int. J. Plant Genomics*, 2009, **2009**, 583429.
- 33 H. H. Liu, X. Tian, Y. J. Li, C. A. Wu and C. C. Zheng, Microarray-based analysis of stress-regulated microRNAs in *Arabidopsis thaliana*, *RNA*, 2008, **14**(5), 836–843.
- 34 B. Khraiwesh, J. K. Zhu and J. Zhu, Role of miRNAs and siRNAs in biotic and abiotic stress responses of plants, *Biochim. Biophys. Acta*, 2012, **1819**(2), 137–148.
- 35 X. Zhang, Dynamic differential methylation facilitates pathogen stress response in *Arabidopsis*, *Proc. Natl. Acad. Sci. U. S. A.*, 2012, **109**(32), 12842–12843.
- 36 V. Chinnusamy, Z. Gong and J. K. Zhu, Abscisic acid-mediated epigenetic processes in plant development and stress responses, *J. Integr. Plant Biol.*, 2008, **50**(10), 1187–1195.
- 37 M. Bohmer and J. I. Schroeder, Quantitative transcriptomic analysis of abscisic acid-induced and reactive oxygen

- species-dependent expression changes and proteomic profiling in *Arabidopsis* suspension cells, *Plant J.*, 2011, **67**(1), 105–118.
- 38 A. Wingler and T. Roitsch, Metabolic regulation of leaf senescence: interactions of sugar signalling with biotic and abiotic stress responses, *Plant Biol.*, 2008, **10**(suppl. 1), 50–62.
- 39 S. Neill, R. Barros, J. Bright, R. Desikan, J. Hancock, J. Harrison, P. Morris, D. Ribeiro and I. Wilson, Nitric oxide, stomatal closure, and abiotic stress, *J. Exp. Bot.*, 2008, **59**(2), 165–176.
- 40 J. Krasensky and C. Jonak, Drought, salt, and temperature stress-induced metabolic rearrangements and regulatory networks, *J. Exp. Bot.*, 2012, **63**(4), 1593–1608.
- 41 V. Chinnusamy, K. Schumaker and J. K. Zhu, Molecular genetic perspectives on cross-talk and specificity in abiotic stress signalling in plants, *J. Exp. Bot.*, 2004, **55**(395), 225–236.
- 42 A. Daszkowska-Golec, *Arabidopsis* seed germination under abiotic stress as a concert of action of phytohormones, *Omics*, 2011, **15**(11), 763–774.
- 43 A. Kadioglu, R. Terzi, N. Saruhan and A. Saglam, Current advances in the investigation of leaf rolling caused by biotic and abiotic stress factors, *Plant Sci.*, 2012, **182**, 42–48.
- 44 J. Dechorganat, O. Patrit, A. Krapp, M. Fagard and F. Daniel-Vedele, Characterization of the Nrt2.6 Gene in *Arabidopsis thaliana*: A Link with Plant Response to Biotic and Abiotic Stress, *PLoS One*, 2012, **7**(8), e42491.
- 45 S. Mahajan and N. Tuteja, Cold, salinity and drought stresses: an overview, *Arch. Biochem. Biophys.*, 2005, **444**(2), 139–158.
- 46 K. Shibasaki, M. Uemura, S. Tsurumi and A. Rahman, Auxin response in *Arabidopsis* under cold stress: underlying molecular mechanisms, *Plant Cell*, 2009, **21**(12), 3823–3838.
- 47 L. Zsigmond, G. Rigo, A. Szarka, G. Szekely, K. Otvos, Z. Darula, K. F. Medzihradszky, C. Konec, Z. Konec and L. Szabadoss, *Arabidopsis* PPR40 connects abiotic stress responses to mitochondrial electron transport, *Plant Physiol.*, 2008, **146**(4), 1721–1737.
- 48 L. Xiong, K. S. Schumaker and J. K. Zhu, Cell signaling during cold, drought, and salt stress, *Plant Cell*, 2002, **14**(Suppl.), S165–S183.
- 49 H. Less, R. Angelovici, V. Tzin and G. Galili, Coordinated gene networks regulating *Arabidopsis* plant metabolism in response to various stresses and nutritional cues, *Plant Cell*, 2011, **23**(4), 1264–1271.
- 50 A. Harb, A. Krishnan, M. M. Ambavaram and A. Pereira, Molecular and physiological analysis of drought stress in *Arabidopsis* reveals early responses leading to acclimation in plant growth, *Plant Physiol.*, 2010, **154**(3), 1254–1271.
- 51 K. Jayasekaran, K. N. Kim, M. Vivekanandan, J. S. Shin and S. H. Ok, Novel calcium-binding GTPase (AtCBG) involved in ABA-mediated salt stress signaling in *Arabidopsis*, *Plant Cell Rep.*, 2006, **25**(11), 1255–1262.
- 52 P. M. Huang, C. J. Chen and S. J. Wang, Tissue-specific regulation of rice molybdenum cofactor sulfurase gene in response to salt stress and ABA, *Acta Physiol. Plant*, 2009, **31**(3), 545–551.
- 53 G. Yu, W. Zang, X. Yang, L. Wang, Z. Tang and P. Luo, Meta-analysis of *Arabidopsis thaliana* under abscisic acid and salt stress, *J. Med. Plants Res.*, 2011, **5**(24), 5889–5893.
- 54 G. R. Cramer, K. Urano, S. Delrot, M. Pezzotti and K. Shinozaki, Effects of abiotic stress on plants: a systems biology perspective, *BMC Plant Biol.*, 2011, **11**, 163.
- 55 J. Kilian, D. Whitehead, J. Horak, D. Wanke, S. Weinl, O. Batistic, C. D'Angelo, E. Bornberg-Bauer, J. Kudla and K. Harter, The AtGenExpress global stress expression data set: protocols, evaluation and model data analysis of UV-B light, drought and cold stress responses, *Plant J.*, 2007, **50**(2), 347–363.
- 56 R. Desikan, S. AH-Mackerness, J. T. Hancock and S. J. Neill, Regulation of the *Arabidopsis* transcriptome by oxidative stress, *Plant Physiol.*, 2001, **127**(1), 159–172.
- 57 M. Debnath, M. Pandey and P. S. Bisen, An omics approach to understand the plant abiotic stress, *Omics*, 2011, **15**(11), 739–762.
- 58 A. Harb and A. Pereira, Screening *Arabidopsis* genotypes for drought stress resistance, *Methods Mol. Biol.*, 2011, **678**, 191–198.
- 59 N. Galpaz and M. Reymond, Natural variation in *Arabidopsis thaliana* revealed a genetic network controlling germination under salt stress, *PLoS One*, 2010, **5**(12), e15198.
- 60 D. Fleury, S. Jefferies, H. Kuchel and P. Langridge, Genetic and genomic tools to improve drought tolerance in wheat, *J. Exp. Bot.*, 2010, **61**(12), 3211–3222.
- 61 M. Seki, M. Satou, T. Sakurai, K. Akiyama, K. Iida, J. Ishida, M. Nakajima, A. Enju, M. Narusaka and M. Fujita, et al., RIKEN *Arabidopsis* full-length (RAFL) cDNA and its applications for expression profiling under abiotic stress conditions, *J. Exp. Bot.*, 2004, **55**(395), 213–223.
- 62 M. Seki, M. Narusaka, H. Abe, M. Kasuga, K. Yamaguchi-Shinozaki, P. Carninci, Y. Hayashizaki and K. Shinozaki, Monitoring the expression pattern of 1300 *Arabidopsis* genes under drought and cold stresses by using a full-length cDNA microarray, *Plant Cell*, 2001, **13**(1), 61–72.
- 63 T. Hirayama and K. Shinozaki, Research on plant abiotic stress responses in the post-genome era: past, present and future, *Plant J.*, 2010, **61**(6), 1041–1052.
- 64 M. I. Qureshi, S. Qadir and L. Zolla, Proteomics-based dissection of stress-responsive pathways in plants, *J. Plant Physiol.*, 2007, **164**(10), 1239–1260.
- 65 K. Kosova, P. Vitamvas, I. T. Prasil and J. Renaut, Plant proteome changes under abiotic stress – contribution of proteomics studies to understanding plant stress response, *J. Proteomics*, 2011, **74**(8), 1301–1322.
- 66 S. Amme, A. Matros, B. Schlesier and H. P. Mock, Proteome analysis of cold stress response in *Arabidopsis thaliana* using DIGE-technology, *J. Exp. Bot.*, 2006, **57**(7), 1537–1546.

- 67 K. Urano, Y. Kurihara, M. Seki and K. Shinozaki, 'Omics' analyses of regulatory networks in plant abiotic stress responses, *Curr. Opin. Plant Biol.*, 2010, **13**(2), 132–138.
- 68 M. A. Moreno-Risueno, W. Busch and P. N. Benfey, Omics meet networks – using systems approaches to infer regulatory networks in plants, *Curr. Opin. Plant Biol.*, 2010, **13**(2), 126–131.
- 69 T. A. Long, S. M. Brady and P. N. Benfey, Systems approaches to identifying gene regulatory networks in plants, *Annu. Rev. Cell Dev. Biol.*, 2008, **24**, 81–103.
- 70 S. Y. Rhee, J. Dickerson and D. Xu, Bioinformatics and its applications in plant biology, *Annu. Rev. Plant Biol.*, 2006, **57**, 335–360.
- 71 W. Huang da, B. T. Sherman and R. A. Lempicki, Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources, *Nat. Protocols*, 2009, **4**(1), 44–57.
- 72 S. Bauer, S. Grossmann, M. Vingron and P. N. Robinson, Ontologizer 2.0 – a multifunctional tool for GO term enrichment analysis and data exploration, *Bioinformatics*, 2008, **24**(14), 1650–1651.
- 73 R. Edgar, M. Domrachev and A. E. Lash, Gene Expression Omnibus: NCBI gene expression and hybridization array data repository, *Nucleic Acids Res.*, 2002, **30**(1), 207–210.
- 74 T. Barrett and R. Edgar, Gene expression omnibus: microarray data storage, submission, retrieval, and analysis, *Methods Enzymol.*, 2006, **411**, 352–369.
- 75 T. Barrett, D. B. Troup, S. E. Wilhite, P. Ledoux, C. Evangelista, I. F. Kim, M. Tomashevsky, K. A. Marshall, K. H. Phillippy and P. M. Sherman, *et al.*, NCBI GEO: archive for functional genomics data sets – 10 years on, *Nucleic Acids Res.*, 2011, **39**(Database issue), D1005–D1010.
- 76 M. Kanehisa, S. Goto, Y. Sato, M. Furumichi and M. Tanabe, KEGG for integration and interpretation of large-scale molecular data sets, *Nucleic Acids Res.*, 2012, **40**(Database issue), D109–D114.
- 77 M. Punta, P. C. Coggill, R. Y. Eberhardt, J. Mistry, J. Tate, C. Boursnell, N. Pang, K. Forslund, G. Ceric and J. Clements, *et al.*, The Pfam protein families database, *Nucleic Acids Res.*, 2012, **40**(Database issue), D290–D301.
- 78 P. Lamesch, T. Z. Berardini, D. Li, D. Swarbreck, C. Wilks, R. Sasidharan, R. Muller, K. Dreher, D. L. Alexander and M. Garcia-Hernandez, *et al.*, The Arabidopsis Information Resource (TAIR): improved gene annotation and new tools, *Nucleic Acids Res.*, 2012, **40**(Database issue), D1202–D1210.
- 79 Reorganizing the protein space at the Universal Protein Resource (UniProt). *Nucleic Acids Res.*, 2012, **40**(Database issue), D71–D75.
- 80 S. Haider, B. Ballester, D. Smedley, J. Zhang, P. Rice and A. Kasprzyk, BioMart Central Portal – unified access to biological data, *Nucleic Acids Res.*, 2009, **37**(Web server issue), W23–W27.
- 81 B. Smith, M. Ashburner, C. Rosse, J. Bard, W. Bug, W. Ceusters, L. J. Goldberg, K. Eilbeck, A. Ireland and C. J. Mungall, *et al.*, The OBO Foundry: coordinated evolution of ontologies to support biomedical data integration, *Nat. Biotechnol.*, 2007, **25**(11), 1251–1255.
- 82 A. Gandhimathi, A. G. Nair and R. Sowdhamini, PASS2 version 4: an update to the database of structure-based sequence alignments of structural domain superfamilies, *Nucleic Acids Res.*, 2012, **40**(Database issue), D531–D534.
- 83 S. F. Altschul, T. L. Madden, A. A. Schaffer, J. Zhang, Z. Zhang, W. Miller and D. J. Lipman, Gapped BLAST and PSI-BLAST: a new generation of protein database search programs, *Nucleic Acids Res.*, 1997, **25**(17), 3389–3402.
- 84 S. R. Eddy, Profile hidden Markov models, *Bioinformatics*, 1998, **14**(9), 755–763.
- 85 L. J. McGuffin and D. T. Jones, Improvement of the GenTHREADER method for genomic fold recognition, *Bioinformatics*, 2003, **19**(7), 874–881.
- 86 R. M. Bennett-Lovsey, A. D. Herbert, M. J. Sternberg and L. A. Kelley, Exploring the extremes of sequence/structure space with ensemble fold recognition in the program Phyre, *Proteins*, 2008, **70**(3), 611–625.
- 87 L. A. Kelley and M. J. Sternberg, Protein structure prediction on the Web: a case study using the Phyre server, *Nat. Protocols*, 2009, **4**(3), 363–371.
- 88 P. Larsson, M. J. Skwark, B. Wallner and A. Elofsson, Improved predictions by Pcons.net using multiple templates, *Bioinformatics*, 2011, **27**(3), 426–427.
- 89 A. Marchler-Bauer, C. Zheng, F. Chitsaz, M. K. Derbyshire, L. Y. Geer, R. C. Geer, N. R. Gonzales, M. Gwadz, D. I. Hurwitz and C. J. Lanczycki, *et al.*, CDD: conserved domains and protein three-dimensional structure, *Nucleic Acids Res.*, 2013, **41**(Database issue), D348–D352.
- 90 M. Krzywinski, J. Schein, I. Birol, J. Connors, R. Gascoyne, D. Horsman, S. J. Jones and M. A. Marra, Circos: an information aesthetic for comparative genomics, *Genome Res.*, 2009, **19**(9), 1639–1645.
- 91 F. Supek, M. Bosnjak, N. Skunca and T. Smuc, REVIGO summarizes and visualizes long lists of gene ontology terms, *PLoS One*, 2011, **6**(7), e21800.
- 92 M. Naika, K. Shameer, O. K. Mathew, R. Gowda and R. Sowdhamini, STIFDB2: an updated version of plant Stress-responsive Transcription Factor DataBase with additional stress signals, stress-responsive transcription factor binding sites and stress-responsive genes in *Arabidopsis* and rice, *Plant Cell Physiol.*, 2013, **54**(2), e8.
- 93 A. Faro, D. Giordano and C. Spampinato, Combining literature text mining with microarray data: advances for system biology modeling, *Briefings Bioinf.*, 2012, **13**(1), 61–82.
- 94 M. Schmid, T. S. Davison, S. R. Henz, U. J. Pape, M. Demar, M. Vingron, B. Scholkopf, D. Weigel and J. U. Lohmann, A gene expression map of *Arabidopsis thaliana* development, *Nat. Genet.*, 2005, **37**(5), 501–506.
- 95 K. S. Heyndrickx and K. Vandepoele, Systematic identification of functional plant modules through the integration of complementary data sources, *Plant Physiol.*, 2012, **159**(3), 884–901.
- 96 T. Puhakainen, M. W. Hess, P. Makela, J. Svensson, P. Heino and E. T. Palva, Overexpression of multiple

- dehydrin genes enhances tolerance to freezing stress in *Arabidopsis*, *Plant Mol. Biol.*, 2004, **54**(5), 743–753.
- 97 J. L. Riechmann and E. M. Meyerowitz, The AP2/EREBP family of plant transcription factors, *Biol. Chem.*, 1998, **379**(6), 633–646.
- 98 V. Albrecht, O. Ritz, S. Linder, K. Harter and J. Kudla, The NAF domain defines a novel protein–protein interaction module conserved in  $\text{Ca}^{2+}$ -regulated kinases, *EMBO J.*, 2001, **20**(5), 1051–1063.
- 99 A. Coletta, J. W. Pinney, D. Y. Solis, J. Marsh, S. R. Pettifer and T. K. Attwood, Low-complexity regions within protein sequences have position-dependent roles, *BMC Syst. Biol.*, 2010, **4**, 43.
- 100 Y. Cheng, Y. Zhou, Y. Yang, Y. J. Chi, J. Zhou, J. Y. Chen, F. Wang, B. Fan, K. Shi and Y. H. Zhou, *et al.*, Structural and functional analysis of VQ motif-containing proteins in *Arabidopsis* as interacting proteins of WRKY transcription factors, *Plant Physiol.*, 2012, **159**(2), 810–825.
- 101 C. M. Fraser and C. Chapple, The phenylpropanoid pathway in *Arabidopsis*, *Arabidopsis Book*, 2011, **9**, e0152.
- 102 T. Vogt, Phenylpropanoid biosynthesis, *Mol. Plant*, 2010, **3**(1), 2–20.
- 103 R. A. Dixon and N. L. Paiva, Stress-Induced Phenylpropanoid Metabolism, *Plant Cell*, 1995, **7**(7), 1085–1097.
- 104 K. Hahlbrock and D. Scheel, Physiology and Molecular Biology of Phenylpropanoid Metabolism, *Annu. Rev. Plant Physiol. Plant Mol. Biol.*, 1989, **40**(1), 347–369.
- 105 M. D. Groppa and M. P. Benavides, Polyamines and abiotic stress: recent advances, *Amino Acids*, 2008, **34**(1), 35–45.
- 106 T. Kusano, T. Berberich, C. Tateda and Y. Takahashi, Polyamines: essential factors for growth and survival, *Planta*, 2008, **228**(3), 367–381.
- 107 R. Alcazar, T. Altabella, F. Marco, C. Bortolotti, M. Reymond, C. Koncz, P. Carrasco and A. F. Tiburcio, Polyamines: molecules with regulatory functions in plant abiotic stress tolerance, *Planta*, 2010, **231**(6), 1237–1249.
- 108 S. S. Gill and N. Tuteja, Polyamines and abiotic stress tolerance in plants, *Plant Signaling Behav.*, 2010, **5**(1), 26–33.
- 109 S. S. Hussain, M. Ali, M. Ahmad and K. H. Siddique, Polyamines: natural and engineered abiotic and biotic stress tolerance in plants, *Biotechnol. Adv.*, 2011, **29**(3), 300–311.
- 110 F. Marco, R. Alcazar, A. F. Tiburcio and P. Carrasco, Interactions between polyamines and abiotic stress pathway responses unraveled by transcriptome analysis of polyamine overproducers, *Omics*, 2011, **15**(11), 775–781.
- 111 Z. Du, L. Li, C. F. Chen, P. S. Yu and J. Z. Wang, G-SESAME: web tools for GO-term-based gene similarity analysis and knowledge discovery, *Nucleic Acids Res.*, 2009, **37**(Web Server issue), W345–349.
- 112 E. C. Dimmer, R. P. Huntley, Y. Alam-Faruque, T. Sawford, C. O'Donovan, M. J. Martin, B. Bely, P. Browne, W. Mun Chan and R. Eberhardt, *et al.*, The UniProt-GO Annotation database in 2011, *Nucleic Acids Res.*, 2012, **40**(Database issue), D565–570.
- 113 T. Mizoguchi, N. Hayashida, K. Yamaguchi-Shinozaki, H. Kamada and K. Shinozaki, ATMPKs: a gene family of plant MAP kinases in *Arabidopsis thaliana*, *FEBS Lett.*, 1993, **336**(3), 440–444.
- 114 K. Ichimura, T. Mizoguchi, R. Yoshida, T. Yuasa and K. Shinozaki, Various abiotic stresses rapidly activate *Arabidopsis* MAP kinases ATMPK4 and ATMPK6, *Plant J.*, 2000, **24**(5), 655–665.
- 115 B. H. van den Berg, C. Thanthiriwatte, P. Manda and S. M. Bridges, Comparing gene annotation enrichment tools for functional modeling of agricultural microarray data, *BMC Bioinf.*, 2009, **10**(suppl. 11), S9.
- 116 A. Subramanian, P. Tamayo, V. K. Mootha, S. Mukherjee, B. L. Ebert, M. A. Gillette, A. Paulovich, S. L. Pomeroy, T. R. Golub and E. S. Lander, *et al.*, Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles, *Proc. Natl. Acad. Sci. U. S. A.*, 2005, **102**(43), 15545–15550.
- 117 L. Rizhsky, H. Liang, J. Shuman, V. Shulaev, S. Davletova and R. Mittler, When defense pathways collide. The response of *Arabidopsis* to a combination of drought and heat stress, *Plant Physiol.*, 2004, **134**(4), 1683–1696.
- 118 M. Nylander, J. Svensson, E. T. Palva and B. V. Welin, Stress-induced accumulation and tissue-specific localization of dehydrins in *Arabidopsis thaliana*, *Plant Mol. Biol.*, 2001, **45**(3), 263–279.
- 119 K. L. Redman and M. Rechsteiner, Identification of the long ubiquitin extension as ribosomal protein S27a, *Nature*, 1989, **338**(6214), 438–440.
- 120 T. J. Close, A. A. Kortt and P. M. Chandler, A cDNA-based comparison of dehydration-induced proteins (dehydrins) in barley and corn, *Plant Mol. Biol.*, 1989, **13**(1), 95–108.