HRGRN: A Graph Search-Empowered Integrative Database of Arabidopsis Signaling Transduction, Metabolism and Gene Regulation Networks

Xinbin Dai1, Jun Li1,2, Tingsong Liu3 and Patrick Xuechun Zhao1,*

1Plant Biology Division, The Samuel Roberts Noble Foundation, 2510 Sam Noble Parkway, Ardmore, OK 73401, USA
2Present address: Department of Genomics Medicine, University of Texas MD Anderson Cancer Center, Houston, TX 77030, USA.
*Corresponding author: Email: pzhao@noble.org; Fax, +1-580-224-4743.
(Received September 1, 2015; Accepted December 7, 2015)

The biological networks controlling plant signal transduction, metabolism and gene regulation are composed of not only tens of thousands of genes, compounds, proteins and RNAs but also the complicated interactions and coordination among them. These networks play critical roles in many fundamental mechanisms, such as plant growth, development and environmental response. Although much is known about these complex interactions, the knowledge and data are currently scattered throughout the published literature, publicly available high-throughput data sets and third-party databases. Many ‘unknown’ yet important interactions among genes need to be mined and established through extensive computational analysis. However, exploring these complex biological interactions at the network level from existing heterogeneous resources remains challenging and time-consuming for biologists. Here, we introduce HRGRN, a graph search-empowered integrative database of Arabidopsis signal transduction, metabolism and gene regulatory networks. HRGRN utilizes Neo4j, which is a highly scalable graph database management system, to host large-scale biological interactions among genes, proteins, compounds and small RNAs that were either validated experimentally or predicted computationally. The associated biological pathway information was also specially marked for the interactions that are involved in the pathway to facilitate the investigation of cross-talk between pathways. Furthermore, HRGRN integrates a series of graph path search algorithms to discover novel relationships among genes, compounds, RNAs and even pathways from heterogeneous biological interaction data that could be missed by traditional SQL database search methods. Users can also build subnetworks based on known interactions. The outcomes are visualized with rich text, figures and interactive network graphs on web pages. The HRGRN database is freely available at http://plantgrn.noble.org/hrgrn/.

Keywords: Biological network • Gene regulatory network • Graph database • Graph search • Metabolic pathway • Plant signal transduction.

Abbreviations: CPI, compound–protein interaction; CR, chromatin regulator; EC, Enzyme Commission; KEGG, Kyoto Encyclopedia of Genes and Genomes; miRNA, microRNA; PPI, protein—protein interaction; SQL, structured query language; sRNA, small non-coding RNA; TAIR10, The Arabidopsis Information Resource Release 10 (TAIR10); TF, transcription factor; TR, transcriptional regulator.

Introduction

The biological networks controlling plant signal transduction, metabolism and gene regulation play critical roles in many plant fundamental mechanisms, such as growth, development and environmental response. In these networks, tens of thousands of genes, proteins, compounds and RNAs engage in numerous complex interactions. These interactions provide essential information for modeling biological networks, and can be classified into multiple categories based on biological function. Several examples include protein–protein interactions (PPIs), compound–protein interactions (CPIs), the regulatory effect of transcription factors (TFs) or small non-coding RNAs (sRNAs) towards downstream target genes, the catalytic effect of an enzyme on its substrate/product compound in a chemical reaction and the relationship between a transporter protein and its substrate. To fully understand the function of genes at the network level, scientists need to consider comprehensively these heterogeneous interactions in the vicinity of the genes.

However, information about these interactions is scattered among the published literature, high-throughput data sets and third-party databases. For example, the Arabidopsis thaliana Protein Interaction Network (AtPIN) (http://atpin.bioinfojug.net/) houses information about Arabidopsis PPIs that are either experimentally validated or computationally predicted (Brandao et al. 2009). The Kyoto Encyclopedia of Genes and Genomes (KEGG) (http://www.genome.jp/kegg/) compiles enzyme-catalytic chemical reactions in terms of metabolic pathways, which links a substrate/product compound with an enzyme (Kanehisa et al. 2014). Computation-based prediction technologies also provide valuable clues for understanding these interactions. For instance, the Transporter Classification Database (TCDB) (http://www.tcdb.org/) (Saier et al. 2014) hosts the standard sequence for each transporter family, thus enabling researchers to connect transporters with...
Gene function at the level of pathway or network.

on biological functions will be a valuable resource for studying integrating these heterogeneous biological interactions based on biological functions will be a valuable resource for studying gene function at the level of pathway or network.

Graph theory has been applied in the analysis of biological networks (Ren and Liu 2013). Graph theory is the study of graphs, which are mathematical structures used to model pair-wise relationships between objects in computer science and mathematics. A ‘graph’ in this context is made up of ‘nodes’ and ‘edges’ (or ‘relationships’) that connect them (Cover and Thomas 1991). Compared with other network analysis algorithms, such as Bayesian integration (Lee et al. 2010), graph theory is capable of modeling complicated biological networks. With this method, biological entities such as genes, proteins, small compounds and RNAs are represented as nodes, and the interactions among nodes (termed edges or relationships) denote biological relationships among the biological entities. Traversal algorithms of graphical models have been used to mine valuable relationships across networks (Stavrakas et al. 2015) that might be omitted by traditional relational database search methods. However, a graph-based database for analyzing the biological networks that control signal transduction, metabolism and gene regulation is still lacking for the end-users, i.e. biologists.

We developed a graph-based database named HRGRN to search and discover the relationships among genes, proteins, compounds and small RNAs in plant signal transduction, metabolism and gene regulatory networks. The database not only includes genes, proteins, small RNAs and compounds as nodes, but, more importantly, defines comprehensive types of edges to model the interactions between nodes. These include the interactions between proteins; compounds and proteins; TFs and their downstream target genes; small RNAs and their target genes; kinases and downstream genes; transporters and substrates; as well as between substrate/product compounds and enzymes. In addition, the genes with similar expression patterns (also called co-expressed genes) are connected by edges, which will provide in-depth insight into gene–gene relationships. These pre-defined edges were deduced from various data sources, such as PPIs from the third party databases based on two-hybrid experiments and homolog analysis, co-expressed gene pairs from computational analysis of transcriptome data and sRNA–target gene pairs from both experimental validation and computational analysis. Using graph traversal algorithms, the built-in analysis tools can facilitate users to discover novel interactions between genes and/or pathways and build subnetworks for specified nodes based on known interactions. HRGRN offers rich graphics and user-friendly web interfaces that allow users to search, analyze and visualize their results.

Design and Implementation

In our HRGRN database, graphical models were used to reflect the complex biological networks that control plant signal transduction, metabolism and gene expression. We defined the entities within a network as nodes, including coding and non-coding genes (producing proteins and regulatory small RNAs, respectively), substrates of transporter, catalytic products of enzymes (i.e. compounds) and gene family/protein complexes. We defined eight types of edges that link the above nodes based on their biological function (described in the next section). Most edges possessed a property indicating whether the relationship was validated or predicted. In addition, we highlighted the relationships in hormone signaling and gene regulatory networks for pathway cross-talk analysis.

Compilation of nodes

The HRGRN stores approximately 50,000 nodes, including the genes from The Arabidopsis Information Resource Release 10 (TAIR10) (https://www.arabidopsis.org/), compounds from the KEGG database, miRNAs from mirBase (http://www.mirbase.org/) (Kozomara and Griffiths-Jones 2014) and gene families/protein complexes curated from the literature. Within them, regulatory elements such as chromatin regulators (CRs), TFs, transcriptional regulators (TRs), miRNAs, trans-acting small interfering RNAs (ta-siRNAs) and kinases are highlighted due to their important regulatory functions within the networks. The TF/CR genes were annotated using our online plant TF and CR categorization and analysis tool, PlantTFcat (http://plantgrn.noble.org/PlantTFcat/) (Dai and Zhao 2011), and other important elements are highlighted referring to TAIR10 official annotation. In addition, we marked transporters because of the important roles they play in shuttling molecules across membranes.

Compilation of edges

Edges, representing biological relationships between nodes, were compiled into the database using the various data sources below.

TFs and target genes. The data set of TF–target genes was downloaded from the TRANSFAC database (http://www.biobase-international.com/gene-regulation) (Fogel et al. 2005), which houses around 400 validated TF and target gene pairs.

miRNAs and target genes. The database hosts 2,300 miRNA-related edges (i.e. miRNA–target gene pairs) as predicted by the plant small RNA target analysis server, psRNATarget (http://plantgrn.noble.org/psRNATarget/) (Dai and Zhao 2011) analysis server with an expected value of < 3 and UPE value < 25.0. Approximately 300 of their edges have been validated by the literature.
Protein–protein interactions (PPIs). Approximately 27,000 experimentally validated and 90,000 predicted PPIs were collected from the Biogrid (Chatr-aryamonti et al. 2015) database and the AtPin (Brandao et al. 2009) database, respectively.

Compound–protein interactions (CPIs). A typical CPI in HRGRN consists of the interaction between a signal compound and its receptor in a signal transduction pathway. We curated approximately 20 CPIs from the literature.

Molecular transport. We classified transporter genes into families using the comparative and integrative transporter annotation system analysis pipeline, Ciport (http://bioinfo3.noble.org/ciport/) (Li et al. 2008) and the TCDB database (Saier et al. 2014). Subsequently, substrates (e.g. compounds or ions) were linked to their corresponding transporter-coding genes according to the description of its family.

Chemical reactions. Enzyme-catalytic chemical reactions were gathered from the KEGG database (Kanehisa et al. 2014). We downloaded the reference sequences for each Enzyme Commission (EC) number and performed a similarity search against these sequences using the BLASTP program (Altschul et al. 1990) with an e-value against these sequences using the BLASTP program (Altschul et al. 1990) with an e-value $< 1 \times 10^{-6}$ for all Arabidopsis TAIR10 gene models. The best BLASTP hits were used to annotate enzyme-coding genes with EC numbers. This information was used to connect substrates/products with corresponding genes, according to the descriptions of the EC terms.

Protein modifications. Protein phosphorylation is catalyzed by kinases, which regulate the functions of downstream target genes. We collected approximately 3,000 kinase–target gene pairs from the Arabidopsis protein phosphorylation site database, PhosPhAt (http://phoshat.mpimp-golm.mpg.de) (Zulawski et al. 2013).

Pairing genes with a similar expression pattern. Transcriptome data can be used to associate co-expressed genes through expression pattern analysis. We manually curated and selected a total of 29 microarray experiments, including approximately 300 hybrid assays from the ArrayExpress database (Parkinson et al. 2007). All experiments were performed using the Affymetrix 22k chip and were designed to include control and hormone-treated samples from a variety of tissues and organs. We normalized the raw data using the ’affy’ module of the R (https://www.r-project.org/) software package.

We further analyzed the normalized expression data using our high-performance gene association network reconstruction and analysis server, GPLEXUS (http://plantgrn.noble.org/GPLEXUS/) (Li et al. 2014). The GPLEXUS calculates a mutual information value (also called a gene association value) based on the Spearman correlation coefficient between genes over multiple samples (see equation 1 in Li et al. 2014). The value can be used to estimate the similarity of gene expression pattern across the microarray experiments. The higher the value, the more similar the expression pattern between the pair of genes. We set the default cut-off threshold to 0.5 in order to keep those edges that represented gene pairs with similar expression patterns at reasonable confidence. A similar expression pattern may imply a potential relationship between genes. We populated 35,877 such edges (among 12,173 genes) into the HRGRN database.

Website implementation

The HRGRN database was developed with a series of open-source software, including the Resin Java web server (http://caucho.com/), Neo4j database management system (http://neo4j.com/), Cytoscape.js (http://js.cytoscape.org/) and Groovy language interpreter (http://www.groovy-lang.org/). In the back end, nodes and edges were organized by Neo4j, which is a highly scalable native graph database management system developed in Java. The open-source graph database management system, Neo4j, was designed specifically to host graphical data represented by nodes and edges. Furthermore, Neo4j natively supports graph traversal algorithms for various graph search applications, such as finding the shortest path between nodes; therefore, the path searches were significantly accelerated by 3–4 orders of magnitude when compared with graph searches implemented using traditional relational database management systems because the latter need to join to the whole graph path table itself multiple times and thus exponentially increase the demand of both computing time and memory. We implemented several graph path search algorithms that allow users to customize search preferences with considerations for specific aims. We developed the front-end web interfaces using the Groovy language, Cytoscape.js libraries and state-of-the-art hypertext markup language version 5 (HTML5) technologies, which display the graph search results as rich text and interactive network graphs.

Interface and Function

Visualization of nodes and edges

HRGRN executes node, path and biological pathway searches, and subnetwork reconstructions over nodes. The results are visualized on web pages with specific schemas. In the network graph views, the node shape distinguishes important nodes according to their types, such as TFs/TRs/CRs, transporters, kinases and small RNAs. Users can also specify a background color to highlight one or more specific nodes. Meanwhile, the types of edges (i.e. the biological types of relationships) are distinguished by color, and the edge direction (i.e. positive or negative interactions) is represented by the shape of the arrow (e.g. regular arrow or T-shaped arrow). Finally, the evidence of the edge (i.e. validated or predicted interaction) is denoted by the style of line (solid or dashed line) (see the legend of Fig. 1).

Auto-completion

We implemented an auto-completion function to assist users in locating their node of interest from tens of thousands of hosted nodes. Once users input keywords for nodes (e.g. ARF), the web page will list all candidate nodes on-site for
further selection. This function is available on all related web pages.

Search potential paths between nodes

In HRGRN, a node may represent a gene, compound and gene family/protein complex. Users may search possible paths between nodes under the ‘Search Relationship’ menu. Here, one path is composed of a series of direct or indirect relationships between two nodes. Users may specify two nodes with the help of auto-completion, the maximum steps allowed between nodes, edge type (biological relationship) and evidence (validated/predicted) as search preferences (Fig. 2). After submission, the HRGRN database will return an interactive graph displaying all possible paths between the two nodes (Fig. 1). We also extended the path search algorithms to find potential relationships (cross-talk) between biological pathways (currently only for hormone pathways; see other analysis pages under the ‘Search Relationship’ menu).

Building a node-centralized subnetwork

HRGRN can display a node-centralized hierarchical subnetwork. Users need to select a node, such as a gene, small RNA or compound, under the ‘Search Nodes’ menu. Next, after clicking the detail link, HRGRN will generate an interactive subnetwork graph for the specified node in the ‘Neighborhood Graph View tag’ (Fig. 3). In the graph view, the specified node is placed in the center of the subnetwork with the directly related (i.e. one step away) and indirectly related (i.e. two steps away) nodes circling the center node in hierarchical fashion. Clicking any node within the interactive graph, followed by the detailed link in the ‘clicked node’ panel on the right, can also generate the node-centralized subnetwork for that node.

Building a subnetwork for nodes of interest

Under the ‘Build Subnetwork’ menu, users can construct subnetworks based on submitted nodes by utilizing built-in graph search algorithms (Fig. 4).

Customize and download networks

Users may customize the network graph by dragging and right-clicking the nodes and edges. Nodes can be highlighted with the color shown on the ‘Graph Customization’ panel at the right side of the graph. The customized figure can be saved using the ‘Export’ button in the ‘Graph Customization’ panel.

Case Studies

Path discovery between genes/pathways

HRGRN is able effectively to identify potential paths between genes which might be missed by structured query language (SQL) query in traditional relational databases. Taking IAA28 (AT5G25890) and ATCUL3 (AT1G26830) as examples, no direct relationship between them was found based on the published information. However, after we input both gene names under the ‘Search Relationship/Node to Node’ page with the assistance of auto-completion (Fig. 2), the generated diagram indicates that ARF12 (AT1G34310) can link both genes because it is a validated target of IAA28 (Li et al. 2011) with an expression pattern similar to ATCUL3 (Fig. 1). Such integrative analysis of both experimentally validated and computationally validated targets can provide deep insight into the interactions within biological pathways.
predicted data for heterogeneous biological interactions is a unique feature of the HRGRN database. The result provides a clue to build hypotheses that associate two different hormone pathways since IAA28 and ATCUL3 belong to the Auxin pathway and the Ethylene pathway, respectively. The method can be extended to uncover potential cross-talk between biological pathways by iteratively searching all possible relationships between genes in different pathways (see other items in the ‘Search Relationship’ menu).

Subnetwork analysis among nodes
Based on what is already known about interactions, users can use the HRGRN to construct a gene interaction subnetwork for a group of specified genes. In this case study, we used WRKY33 and the F box genes AFB1, AFB2, AFB3 and TIR1 as examples. Under the ‘Build Sub-network’ menu, users can submit these nodes, change the maximum search steps from default 3 to 2 and leave other options as their defaults to generate the subnetwork (Fig. 4), which displays the comprehensive relationships among the five genes with different colors, line shapes and arrow types to represent biological type, evidence and direction of the interactions, respectively. The WRKY33 TF regulates many pathways, including those involved in defense, salicylic acid (SA) signaling and ethylene (ET)–jasmonate (JA)-mediated cross-communication, and camalexin biosynthesis (Petersen et al. 2008, Jiang and Deyholos 2009, Birkenbihl et al. 2012). The F box proteins function as auxin receptors to regulate Auxin/IAA degradation and collectively mediate auxin responses throughout plant development (Dharmasiri et al. 2005). Based on pre-populated background knowledge in HRGRN, the database introduced additional genes, such as miR393 and other coding genes, into the subnetwork to link the five specified genes.

Discussion
As discussed above, the interactions constituting biological networks are heterogeneous since they represent various biological processes and functions, and the data are scattered across a variety of repositories, including published literature, databases and individual high-throughput data sets. Some data require further analysis using computation-based prediction algorithms. Using traditional SQL databases to integrate and analyze such complicated data sets is challenging. As we demonstrated in the case study section, HRGRN is able to discover new relationships between genes, which might be missed by traditional relational database approaches, by integrating these complicated data. Such a capability is one of the major advantages of the graph search-empowered database.

As an integrative biological interaction database, HRGRN still has room for improvement. The plant signaling transduction, metabolism and gene regulation networks are composed of a series of pathways dedicated to specific mechanisms. Within organisms, these pathways also interact with each other through cross-talk mechanisms. From the perspective of a graph model, these pathways are composed of edges and nodes. Thus, the biological pathway information of nodes and edges will assist researchers in discovering new genes and new interactions that are involved in specific pathways, and even performing cross-talk analysis between certain pathways. Referring to published literature and public databases, such as KEGG and the Arabidopsis Hormone Database 2.0 (http://ahd.cbi.pku.edu.cn/) (Jiang et al. 2011), we have marked the edges which are involved in hormone signal transduction and the gene regulatory network. The improvement provides valuable information for discovering novel cross-talk between hormone pathways. We plan to add more pathway information for nodes and edges with an emphasis on development and...
environmental response pathways. For example, the data set from ATTED-II (Obayashi et al. 2014, Aoki et al. 2015) will be included in HRGRN to provide condition-specific co-expression and derived gene–gene associations or pathways.

In addition, since the experimental identification of TFs and target genes is inefficient and time-consuming, HRGRN only hosts approximately 400 validated TF–target interactions, which is much less than the actual interactions observed in the studied organism. Thus, we plan to supplement predicted TF–target interaction data in the future. The association analysis between identified TF-binding motifs and potential target gene promoter regions may predict the interactions between TFs and their target genes (Dai et al. 2007). We will evaluate the performance of a top-down Gaussian graphic models (GGM)
algorithm, because it can reportedly find potential TF target genes from transcriptome data (Wei et al. 2013). We also plan to analyze comprehensively short small RNA regulatory cascade pathways in Arabidopsis utilizing our psRNAMiner (Dai and Zhao 2008) and psRNATarget servers (Dai and Zhao 2011). To provide services to wide plant research communities, we are currently working with the Araport (https://www.araport.org/) (Hanlon et al. 2015) development team to integrate our graph search functions using Remote Procedure Calls (RPC) protocols, which will enable potential users of Araport to search novel interactions among genes and pathways using our HRGRN database. We believe that these improvements will make HRGRN an essential resource for plant science researchers.

Funding

This work was supported by the National Science Foundation [grants DBI: 0960897 and DBI: 1458597 to P.X.Z.]; the Samuel Roberts Noble Foundation.

Acknowledgments

The authors would like to thank Ms. Zhaohong Zhuang for her assistance with collection, compilation and curation of the hormone-related Arabidopsis microarray data for development of the HRGRN database.

Disclosures

The authors have no conflicts of interest to declare.

References

Altschul, S.F., Gish, W., Miller, W., Myers, E.W. and Lipman, D.J. (1990) Basic local alignment search tool. J. Mol. Biol. 215: 403–410.
Aoki, Y., Okamura, Y., Tadaka, S., Kinoshita, K. and Obayashi, T. (2015) ATTED-II in 2016: a plant coexpression database towards lineage-specific coexpression. Plant Cell Physiol. 57: XXX–XXX.
Barrett, T., Wilhite, S.E., Ledoux, P., Evangelista, C., Kim, I.F., Tomashovsky, M., et al. (2013) NCBI GEO: archive for functional genomics data sets—update. Nucleic Acids Res. 41: D991–D995.
Birkenbihl, R.P., Diezel, C. and Somssich, I.E. (2012) Arabidopsis WRKY33 is a key transcriptional regulator of hormonal and metabolic responses toward Botrytis cinerea infection. Plant Physiol. 159: 266–285.
Brandao, M., Dantas, L. and Silva-Filho, M. (2009) AtPIN: Arabidopsis thaliana protein interaction network. BMC Bioinformatics 10: 454.
Chatr-aryamontri, A., Breitkreutz, B.-J., Oughtred, R., Boucher, L., Heinicke, S., Chen, D., et al. (2015) The BioGRID interaction database: 2015 update. Nucleic Acids Res. 43: D470–D478.
Cover, T.M. and Thomas, J.A. (1991) Elements of Information Theory. Wiley, New York.
Dai, X., He, J. and Zhao, X. (2007) A new systematic computational approach to predicting target genes of transcription factors. Nucleic Acids Res. 35: 4433–4440.
Dai, X. and Zhao, P.X. (2008) psRNAMiner: a plant short small RNA regulatory cascade analysis server. Nucleic Acids Res. 36: W114–W118.
Dai, X. and Zhao, P.X. (2011) psRNATarget: a plant small RNA target analysis server. Nucleic Acids Res. 39: W155–W159.
Dharmasiri, N., Dharmasiri, S., WeiJers, D., Lechner, E., Yamada, M., Hobbie, L., et al. (2005) Plant development is regulated by a family of auxin receptor F box proteins. Dev. Cell 9: 109–119.
Fogel, G.B., Weekes, D.G., Varga, G., Dow, E.R., Craven, A.M., Harlow, H.B., et al. (2005) A statistical analysis of the TRANSCR database. Biosystems 81: 137–154.
Hanlon, M.R., Vaughn, M., Mock, S., Dooley, R., Moreira, W., Stubbs, J., et al. (2015) Araport: an application platform for data discovery. Concurrenc Computat. Pract. Exper. 27: 4412–4422.
Jiang, Y. and Deyholos, M.K. (2009) Functional characterization of Arabidopsis NaCl-inducible WRKY25 and WRKY33 transcription factors in abiotic stresses. Plant Mol. Biol. 69: 91–105.
Jiang, Z., Liu, X., Peng, Z., Wan, Y., Ji, Y., He, W., et al. (2011) AHD2.0: an update version of Arabidopsis Hormone Database for plant systematic studies. Nucleic Acids Res. 39: D1123–D1129.
Kanehisa, M., Goto, S., Sato, Y., Kawashima, M., Furumichi, M. and Tanabe, M. (2014) Data, information, knowledge and principle: back to metabolism in KEGG. Nucleic Acids Res. 42: D199–D205.
Kolesnikov, N., Hastings, E., Keays, M., Melinchuk, O., Tang, Y.A., Williams, E., et al. (2015) ArrayExpress update—simplifying data submissions. Nucleic Acids Res. 43: D1113–D1116.
Kozomara, A. and Griffiths-Jones, S. (2014) miRBase: annotating high confidence microRNAs using deep sequencing data. Nucleic Acids Res. 42: D68–D73.
Lee, I., Ambaru, B., Thakkar, P., Marcotte, E.M. and Rhee, S.Y. (2010) Rational association of genes with traits using a genome-scale gene network for Arabidopsis thaliana. Nat. Biotechnol. 28: 149–156.
Li, H., Benedito, V.A., Udvardi, M.K. and Zhao, P.X. (2009) TransportTP: a two-phase classification approach for membrane transporter prediction and characterization. BMC Bioinformatics 10: 418.
Li, H., Dai, X. and Zhao, X. (2008) A nearest neighbor approach for automated transporter prediction and categorization from protein sequences. Bioinformatics 24: 1129–1136.
Li, J.-F., Bush, J., Xiong, Y., Li, L. and McCormack, M. (2011) Large-scale protein–protein interaction analysis in Arabidopsis mesophyll proplasts by split firefly luciferase complementation. PLoS One 6: e27364.
Li, J., Wei, H., Liu, T. and Zhao, P.X. (2014) CPLEXUS: enabling genome-scale gene association network reconstruction and analysis for very large-scale expression data. Nucleic Acids Res. 42: e32.
Mishra, N.K., Chang, J. and Zhao, P.X. (2014) Prediction of membrane transport proteins and their substrate specificities using primary sequence information. PLoS One 9: e100278.
Obayashi, T., Okamura, Y., Ito, S., Tadaka, S., Aoki, Y., Shirotta, M., et al. (2014) ATTED-II in 2014: evaluation of gene coexpression in agriculturally important plants. Plant Cell Physiol. 55: 66.
Parkinson, H., Kapushesky, M., Shojaatlab, M., Abeygunawardena, N., Coulson, R., Farne, A., et al. (2007) ArrayExpress—a public database of microarray experiments and gene expression profiles. Nucleic Acids Res. 35: D767–D770.
Petersen, K., Fii, B.K., Mundy, J. and Petersen, M. (2008) Downstream targets of WRKY33. Plant Signal. Behav. 3: 1033–1034.
Ren, G. and Liu, Z. (2013) NetCAD: a network analysis tool for coronary artery disease-associated PPI network. Bioinformatics 29: 279–280.
Saier, M.H., Reddy, V.S., Tamang, D.G. and Västermark, Å. (2014) The transporter classification database. Nucleic Acids Res. 42: D251–D258.
Stavrakas, V., Melas, I.N., Sakellaropoulos, T. and Alexopoulos, L.G. (2015) Stomatopedia: a comprehensive database of plant receptor F box proteins. Dev. Cell 29: 109–119.
Wei, H., Yordanov, Y., Kumaria, S., Georgieva, T. and Busov, V. (2013) Genetic networks involved in poplar root response to low nitrogen. Plant Signal. Behav. 8: e27211.
Zulawski, M., Braginets, R. and Schulze, W.Z. (2013) PhosPhAt goes kinases—searchable protein kinase target information in the plant phosphorylation site database PhosPhAt. Nucleic Acids Res. 41: D1176–D1184.