BMC
Plant Biology

RESEARCH ARTICLE

Regulatory network rewiring for secondary metabolism in *Arabidopsis thaliana* under various conditions

Qi Lv[1], Rong Cheng[1] and Tieliu Shi[1,2]*

Abstract

Background: Plant secondary metabolites are critical to various biological processes. However, the regulations of these metabolites are complex because of regulatory rewiring or crosstalk. To unveil how regulatory behaviors on secondary metabolism reshape biological processes, we constructed and analyzed a dynamic regulatory network of secondary metabolic pathways in Arabidopsis.

Results: The dynamic regulatory network was constructed through integrating co-expressed gene pairs and regulatory interactions. Regulatory interactions were either predicted by conserved transcription factor binding sites (TFBSs) or proved by experiments. We found that integrating two data (co-expression and predicted regulatory interactions) enhanced the number of highly confident regulatory interactions by over 10% compared with using single data. The dynamic changes of regulatory network systematically manifested regulatory rewiring to explain the mechanism of regulation, such as in terpenoids metabolism, the regulatory crosstalk of RAV1 (AT1G13260) and ATHB1 (AT3G01470) on HMG1 (hydroxymethylglutaryl-CoA reductase, AT1G76490); and regulation of RAV1 on epoxysqualene biosynthesis and sterol biosynthesis. Besides, we investigated regulatory rewiring with expression, network topology and upstream signaling pathways. Regulatory rewiring was revealed by the variability of genes' expression: pathway genes and transcription factors (TFs) were significantly differentially expressed under different conditions (such as terpenoids biosynthetic genes in tissue experiments and E2F/DP family members in genotype experiments). Both network topology and signaling pathways supported regulatory rewiring. For example, we discovered correlation among the numbers of pathway genes, TFs and network topology: one-gene pathways (such as δ-carotene biosynthesis) were regulated by a fewer TFs, and were not critical to metabolic network because of their low degrees in topology. Upstream signaling pathways of 50 TFs were identified to comprehend the underlying mechanism of TFs' regulatory rewiring.

Conclusion: Overall, this dynamic regulatory network largely improves the understanding of perplexed regulatory rewiring in secondary metabolism in Arabidopsis.

Keywords: Regulatory network, Rewiring, Secondary metabolism, Arabidopsis

* Correspondence: tieliushi01@gmail.com
[1]Center for Bioinformatics and Computational Biology, Shanghai Key Laboratory of Regulatory Biology, the Institute of Biomedical Sciences and School of Life Science, East China Normal University, Shanghai 200241, China
[2]Shanghai Institute of Plant Physiology and Ecology, Shanghai Institutes for Biological Sciences, Chinese Academy of Sciences, Shanghai 200031, China

Background

The researches on mechanism, function and evolution of plant secondary metabolism were traced back to about 60 years ago [1,2]. Secondary metabolic pathways lead to tens of thousands of products involved in various biological responding processes, under stimuli of specific external environmental stress elicitors as well as signal molecules of normal growth and development [3,4]. Secondary metabolisms of Arabidopsis are classified into five major groups (Additional file 1: Table S1): nitrogen-containing secondary compounds biosynthesis (NSCB), terpenoids biosynthesis (TB), sugar derivatives biosynthesis (SDB), phenylpropanoid derivatives biosynthesis (PDB) and flavonoids biosynthesis (FB) in AraCyc database [5]. Most nitrogen containing compounds, playing important roles in biological responses in plant defense and human nutrition [6,7], are regulated by MYB and bHLH members in transcription levels [8]. Sugar secondary derivatives, members of low molecular weight metabolites (mainly cyclic sugar alcohols), are associated with osmotic stress in higher plants [9]. Phenylpropanoids are constitutive compounds in certain tissues [10] or responding factors induced by stresses (such as UV, wounding, pathogen attack, low temperature and low iron level) [10-13]. And these metabolites are regulated by AtMYB21 (AT3G27810), AtMYB4 (AT4G38620), HY5 (AT5G11260), and CIP7 (AT4G27430) [8]. Flavonoids, a major metabolic branch derived from phenylalanine and malonyl coenzyme A, are regulated by MYB and bHLH family members [14,15]. Terpenoids, the largest secondary metabolic family irreplaceable in inner communication with: environment; plant growth; and development [16-18], are regulated by AP2/ERF, bHLH and MYB members [19]. The significant functions of these compounds make their regulators critical targets in genetic engineering applications for improving plant qualities, and for enzymes engineering TF is one kind of candidates [20]. However, metabolic engineering primarily concentrates on production of only one metabolite or a single metabolic gene and normally generates unexpected metabolic consequences–because metabolic pathways in plant intertwine one another to form a complex network; and perturbation of a single gene in the network usually have extensively effects on metabolic flux [21,22]. Therefore, regulatory mechanisms of biosynthetic genes are too complex to comprehensively reveal because of 'biodiversity' or 'chemodiversity', asking for system analysis rather than independent experiments.

The first sequenced flowering plant *Arabidopsis thaliana* is widely used as a model to systematically study gene function and physiology in plant science [23,24]. With high-throughput technologies such as microarray, Chip-chip etc., numerous data have been generated in this model plant, making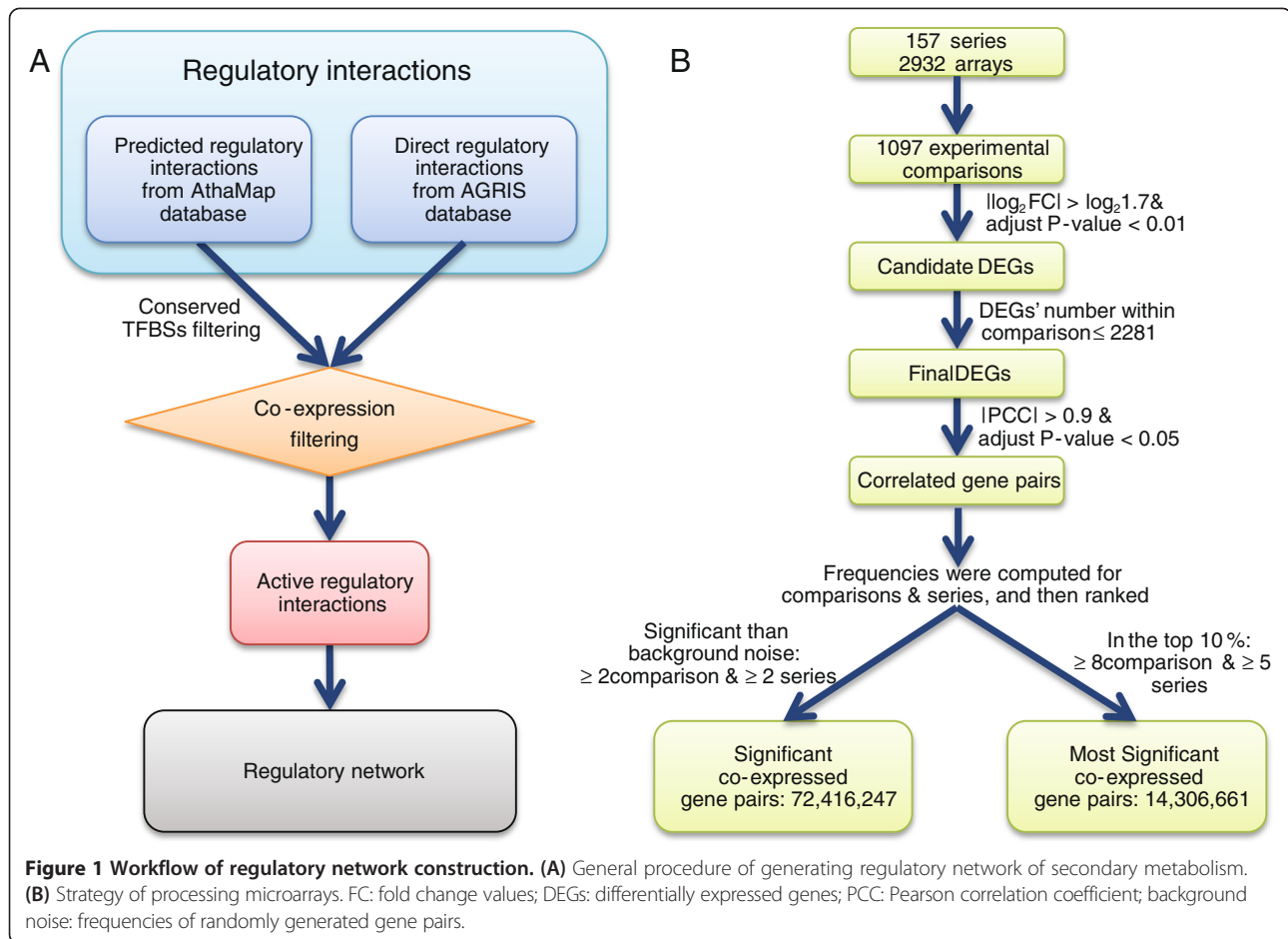 it possible to explore biological mechanisms in plant developmental and environmental responses on genomic scale. Among these technologies, gene microarray aims to investigate expression of genes on a large scale in various treatments or developmental stages [25-28]. Many approaches used microarrays in systematic analysis of regulation over whole genome. For instance, Bayesian was applied to build dynamic regulatory network over time series microarrays, presuming causal relationships between TFs and target genes [29,30]. Other studies generated co-expression data from microarrays and then utilized function specific cis-elements (obtained from multiple sequence alignments on promoter regions of co-expressed genes) to reconstruct regulatory network, assuming that co-expressed genes are co-regulated by the same TFs [31]. Also, researchers used microarrays in expression quantitative trait locus (eQTLs) analysis to identify hot spot regions where regulatory genes locate. For example, researchers built genetic regulatory network in flowering and single gene mutants in Arabidopsis [32,33] and identified effects of TFs on multiple metabolic phenotypes [34]. However, these studies focusing on regulatory network–mainly stress (drought, cold, dehydration, etc.) or development (flowering, seed maturation, etc.) specific [35-38]–are limited in: types of experiments, sizes of networks, families of TFs and numbers of target genes. Besides, most available regulatory databases only addressed on their particular regulatory information (Additional file 2: Table S2). These limitations in regulatory network analysis and database specificity make it insufficient to systemically study regulatory mechanism–neglecting dynamic changes, biological responses and regulatory rewiring or crosstalk between regulations. However, systematic researches of transcriptional regulations on metabolic pathways are still fewer than function studies (only focusing on one or several TFs) [39,40].

Therefore, we developed a method to construct a dynamic regulatory network significant in biological function by integrating regulatory interactions, large-scale microarray data and evolutionary conservation of TFBSs (Figure 1A). This dynamic network is efficient in systematically exploring regulatory rewiring (or crosstalk) on pathways to explain the mechanism of regulation. We investigated the regulatory rewiring with expression, network topology and upstream signaling pathways, which largely improves the understanding of perplexed regulatory rewiring mechanism in secondary metabolism.

Results

Dynamic regulatory network reconstruction with co-expression data and regulatory interactions

We reconstructed the regulatory network of secondary metabolism in *Arabidopsis thaliana* through combining co-expressed gene pairs with regulatory interactions

Figure 1 Workflow of regulatory network construction. (A) General procedure of generating regulatory network of secondary metabolism. **(B)** Strategy of processing microarrays. FC: fold change values; DEGs: differentially expressed genes; PCC: Pearson correlation coefficient; background noise: frequencies of randomly generated gene pairs.

(Figure 1A): either 422,967 predicted regulatory interactions from AthaMap [41], which were then filtered by conserved transcription factor binding sites (TFBSs); or 10,653 directly experiment-proved ones from AGRIS [42,43].

We first used conserved TFBSs (see "Methods") to increase the confidence of regulatory interactions predicted in AthaMap from the perspective of evolution. As expected, poplar (the closest species to Arabidopsis among four used organisms in the evolutionary tree) had more conserved TFBSs. In contrast, we did not find any conserved TFBSs among TB orthologous genes in chlamydomonas (the farthest species to Arabidopsis among four used organisms in the evolutionary tree)–possibly owing to large evolutionary distance between them. This verifies the rationale of our results in conserved TFBSs computing.

Next, 72,416,247 significantly co-expressed gene pairs and 14,306,661 most significantly co-expressed gene pairs were obtained from microarray analysis (Figure 1B). Based on these co-expressed gene pairs, we identified a substantial amount of active regulatory interactions to construct regulatory network. 28% of regulatory interactions from AthaMap were maintained after being filtered with

significantly co-expressed gene pairs. At the same time, among the regulatory interactions from AGRIS database, about 39% were significantly co-expressed and 6% were most significantly co-expressed–consistent with the fundamental assumption of regulatory interaction prediction: expression patterns of TFs and their target genes were similar.

To validate our filtering strategies of AthaMap data, we compared the proportion of direct regulatory interactions (AGRIS) in the predicted ones (AthaMap) across different data filtering strategies (Figure 2A) with five TFs: FUS3 (AT3G26790), AtLEC2 (AT1G28300), AG (AT4G18960), AGL15 (AT5G13790) and HY5 (AT5G11260) in both AGRIS and AthaMap. After being filtered by only TFBSs alignments or significant co-expressed gene pairs, the percentages of direct regulatory interactions were 5.46% and 4.26% respectively. When being filtered by both TFBSs alignments and significantly co-expressed gene pairs, the fraction of direct regulatory interactions increased to 12.10%. After adding evolutionary conservation filtering, this percentage reached 14.53%. Therefore, our filtering methods were efficient in predicting highly confident regulatory interactions.
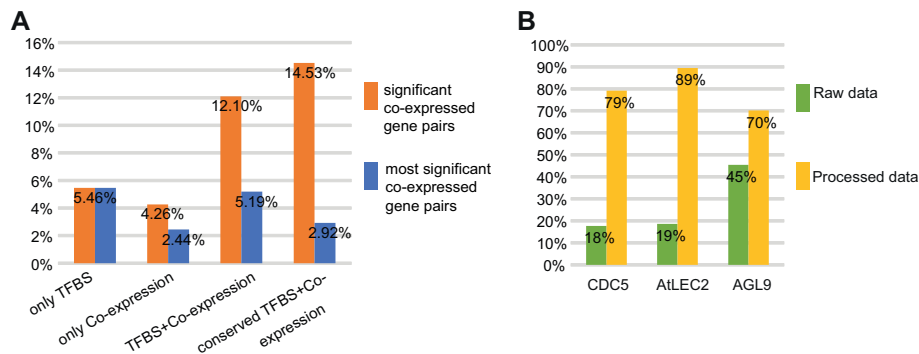
Figure 2 Validation of filtering method in identifying high confident regulatory interactions. (A) The percentages of experiment-confirmed regulatory interactions in regulatory interactions predicted by different data. To validate the filtering strategies of AthaMap data, the proportion of direct regulatory interactions (AGRIS) in the predicted ones (AthaMap) across different data filtering strategies were computed with the regulatory interactions of five TFs: FUS3 (AT3G26790), AtLEC2 (AT1G28300), AG (AT4G18960), AGL15 (AT5G13790) and HY5 (AT5G11260) present in both AGRIS and AthaMap. **(B)** Literature evidences about the function of TFs' target pathways that were consistent with phenotypes of TFs' mutant. The function of TFs' target pathways were mined in literatures. The phenotypes of TFs' mutant were obtained from AtPID (Arabidopsis thaliana Protein Interactome Database). The percentages of target pathways whose functions were consistent to TFs' mutant phenotypes were computed.

To verify that our processed regulatory interactions are more confident than raw regulatory interactions, we first compared the numbers of target pathways in these two datasets. We observed that the numbers of target pathways in our result were smaller than that in both raw dataset and random dataset (Additional file 3: Figure S1A). And the raw dataset was not significantly different from random dataset compared with our processed data (Additional file 3: Figure S1B). Furthermore, to validate the reliability of our proposed method in eliminating low confident data in raw dataset, we mined function of pathways (containing predicted target genes) from literatures and checked their consistency with TFs' mutant-phenotypes in ATPID database [44]. Here we only chose TFs with simple mutant-phenotypes (Additional file 4: Table S3) to make the result more precise, ignoring complex mutant-phenotypes associated with multiple functions. The percentages of literature evidences in our processed data were higher than that in the raw dataset (Figure 2B): for each TF, functions of more than half target pathways in processed regulatory interactions were correlated with mutant-phenotype. In our results, CDC5 (AT1G09770, cell division cycle 5, a MYB family member) regulated 24 pathways, and 19 of them were correlated with embryo development in literatures–in accordance with embryonic defect, the phenotype of CDC5 mutant in ATPID. AGL9 (AT1G24260, agamous-like 9, a member of SEP3 family), whose mutant-phenotype was about flowers in ATPID, regulated 57 pathways in our result, and 40 of target pathways were associated with flowers in literatures. AtLEC2 mutant affected normal embryonic and cotyledonal development in ATPID, 17 of the 19 pathways which were predicted to be regulated by AtLEC2 were related to this TF's mutant-phenotype. These results indicate that our method is efficient

in identifying highly confident regulatory interactions from raw dataset.

Regulatory rewiring under diverse conditions in TB

Based on the dynamic regulatory network, we analyzed regulatory rewiring to demonstrate the dynamic changes of regulation. In pathway level, according to the regulation of TFs on target genes, we classified regulations into three types: positive, negative or both positive and negative. Positive or negative regulation of a TF on a pathway signifies that the regulations on different pathway genes are constant and don't change with experiments. Both positive and negative regulation, considered as inconstant and reciprocal regulation on different target enzymatic genes, maintains the balance of metabolic flux within pathways. For example, abscisic acid glucose ester biosynthesis contains only one reaction, and the reaction is catalyzed by abscisic acid glucosyltransferase that is encoded by more than 20 genes. We found both positive and negative regulations of RAV1 (AT1G13260, an AP2/B3 domain TF) on this pathway, which could keep abscisic acid glucosyltransferase steady.

Since TB is critical to plant, we took two examples in TB sub-network to illustrate regulatory rewiring under diverse experimental conditions. Generally, the occurrence of rewiring is caused by either regulatory interactions between TFs or regulatory alterations under different conditions.

One example is the rewiring between the regulatory crosstalk of RAV1 and ATHB1 (AT3G01470, a HD-ZIP family member) on HMG1 (AT1G76490, a hydroxy methylglutaryl-CoA reductase). RAV1 positively regulated HMG1 gene in independence (Additional file 3: Figure S2A) under many conditions: the grown stage of leaves; tocopherols mutant VTE1 (vitamin E deficient 1) gene; tocopherols mutant VTE2 (vitamin E deficient 2)

gene; leaves responses to *Phytophthora infestans* and COP9 (constitutive photomorphogenic 9) signalosome mutants grown in dark, etc. However, the positive regulation role of RAV1 on HMG1 gene altered when ATHB1 promoted the expression of HMG1 gene (Additional file 3: Figure S2B) under a few conditions, like ABA1 (zeaxanthin epoxidase) gene mutant and hypoxia stress. We found that the distance between two TFs' binding sites of HMG1 gene was within 200 bps, suggesting that the binding of ATHB1 on HMG1 gene affect normal binding of RAV1, and thereby change RAV1's usual regulatory function. It implies that the interactions between these two TFs result in the alteration of RAV1 regulation on HMG1 gene.
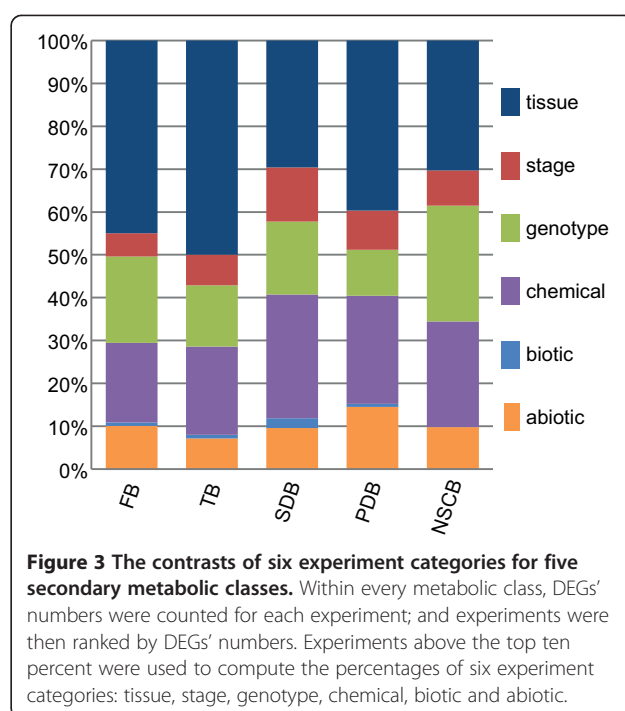
The example of epoxysqualene biosynthesis also demonstrates the rewiring of regulatory crosstalk (Additional file 3: Figure S3). Epoxysqualene biosynthesis pathway is the upstream pathway, leading to sterol biosynthesis (a major class of triterpenoids) and other triterpenoids biosynthesis. The regulations of RAV1 on epoxysqualene biosynthesis and sterol biosynthesis pathways were the same in multiple experiments: Phytophthora infestans plants, water treatment on leaves for 24 hours, $MgCl_2$ treatment on leaves for 12 hours, etc. But RAV1 regulated the two pathways differently under a few conditions (such as seedling and CSN3 gene mutant in dark): negatively regulated epoxysqualene biosynthesis but positively regulated sterol biosynthesis, because of the activation of some downstream biosynthetic genes regulated by RAV1 in sterol biosynthesis pathway. The change of RAV1's regulation would affect normal metabolic distribution of sterol-related and the other triterpenoids, indicating the importance of regulatory rewiring in controlling metabolic flux.

Variability of genes' expressions revealing regulatory rewiring

Significant variability in expressions of TF-encoding and pathway genes could further reveal regulatory rewiring by providing gene activities and functions specific to experimental conditions. We investigated the changes of gene activities and functions through differentially expressed genes (DEGs) (see "Methods").

Biosynthetic genes of FB, PDB and TB were significantly differentially expressed in tissue experiments (Figure 3), suggesting their dramatic changes in biological development of specific tissues. On the contrary, genes in NSCB were significantly differentially expressed in genotype experiments, indicating notable activities and biological function of nitrogen-containing compounds in plants of different genotypes.

The expressions of TF-encoding genes were also different under different experimental conditions. On one hand, genes–primary in WRKY(Zn), NAC, AP2/EREBP

Figure 3 The contrasts of six experiment categories for five secondary metabolic classes. Within every metabolic class, DEGs' numbers were counted for each experiment; and experiments were then ranked by DEGs' numbers. Experiments above the top ten percent were used to compute the percentages of six experiment categories: tissue, stage, genotype, chemical, biotic and abiotic.

and MYB TF families–were widely differentially expressed (Additional file 3: Figure S4), indicating their global roles in regulations of downstream TFs or target enzymatic genes. In WRKY(Zn) family, WRKY18 (AT4G31800) and WRKY40 (AT1G80840) (which were induced by pathogen) [45], were greatly differentially expressed in genotype and chemical experiments; whereas WRKY6 (AT1G62300, associated with leaf senescence and defense) [46] was significantly differentially expressed in tissue and chemical treatments. Compared with WRKY(Zn) family, stress induced NAC family members (ANAC072, AT4G27410; ANAC019, AT1G52890; ANAC055, AT3G15500) [47] were differentially expressed in grown stages. On the other hand, TFs only differentially expressed in a few experiments included E2F/DP family (regulating core cell cycle) [48], C2H2(Zn) family (controlling flowering, germination and root development) [49-51] and ABI3/VP1 family (governing seed maturation) [52]. Those TFs were possibly more specific to particular conditions. For instance, a member of E2F/DP family (E2Ff, AT3G01330) was more specific to genotype experiments; whereas a member of C2H2(Zn) family (ID1, AT1G25250) was more typical to tissues.

To explore the relationship between TFs and pathway genes, we clustered them by expression profiles under different experiment categories and took TB as an example (Additional file 3: Figure S5). Firstly, we could not distinguish TFs and biosynthetic genes by two separated clusters. Secondly, one cluster was a union of similarly expressed genes, and contained both TFs and enzymatic genes, suggesting potential regulations between them in

a cluster. Thirdly, clusters were different under five experiment categories, indicating variability of potential regulations depending on experimental conditions. This variability reflects biodiversity to a certain extent, which is in accordance with regulatory rewiring–emphasizing the reasons of biological complexity.
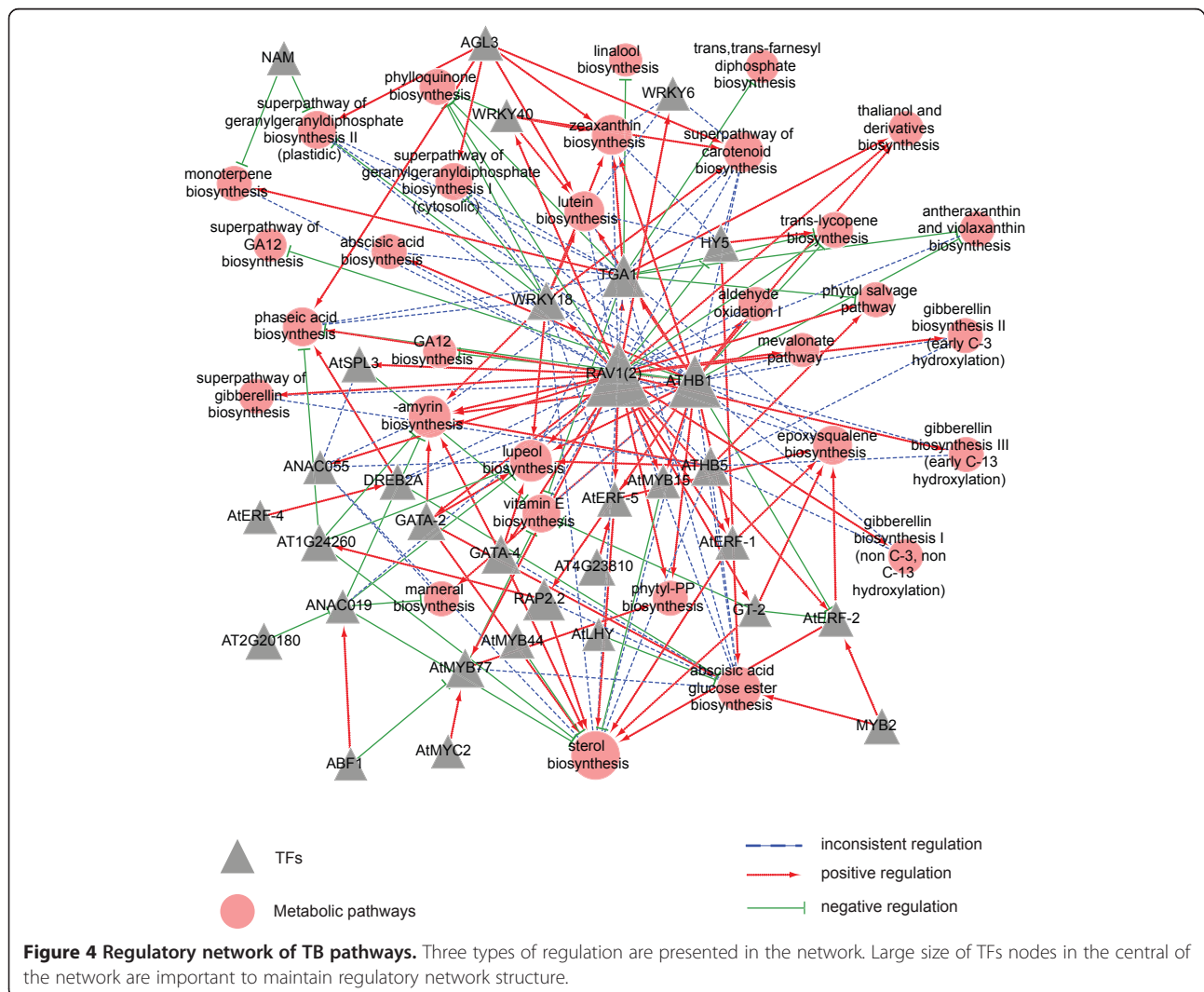
Explanation of regulatory rewiring by network topology

Network topological properties could efficiently explain regulatory rewiring based on network structure (Figure 4, Additional file 3: Figure S6-S9). We compared the general network properties of five secondary metabolic classes (Table 1). The numbers of TFs, genes and regulatory interactions in TB were the highest, indicating that the regulation of TB was the most complicated. In contrast, these topology properties in SDB were the lowest, revealing that the regulation of SDB was the simplest. The number of positive regulations was larger than that of negative regulations, and the number of inconsistent regulations was the least, which was discovered in network motifs as well. In addition, the network motifs (the basic component of the network) with high frequencies contained more positive regulations than negative regulations and inconsistent regulations, showing that most regulations were invariant (Additional file 3: Figure S10).

Since high degree nodes are important in maintaining network structure whereas closeness gives a rough indication of how well a node connects the network, we used degree and closeness to explore the functions of pathways and TFs. We found that the nodes with high degree were also high in closeness (Additional file 3: Figure S11A). Interestingly, almost all nodes with both high degrees and closeness were TFs, while pathways always had low degree and closeness, indicating that TFs play a key role in maintaining the structure of network.

It is intriguing which factor affects the numbers of TFs in network. We discovered that the number of TFs was

Figure 4 Regulatory network of TB pathways. Three types of regulation are presented in the network. Large size of TFs nodes in the central of the network are important to maintain regulatory network structure.

Table 1 Network properties of five secondary metabolic regulatory networks

	No. of TFs	No. of genes	Average degree	No. of negative regulation	No. of positive regulation	No. of inconsistent regulation
FB	25	57	4.15	47	101	22
NSCB	14	25	4.05	14	42	23
PDB	22	38	4.40	22	85	25
SDB	7	4	3.45	7	9	3
TB	32	72	4.12	58	118	38

FB, flavonoids biosynthesis; NSCB, nitrogen-containing secondary compounds biosynthesis; PDB, phenylpropanoid derivatives biosynthesis; SDB, sugar derivatives biosynthesis; TB, terpenoids biosynthesis; TFs, transcription factor; Average degree, the mean value of nodes' degree to present network density.

correlated with the number of genes in this pathway (Additional file 3: Figure S11B): the fewer genes in a pathway, the fewer TFs regulating this pathway, suggesting simpler regulation of this pathway. Besides, similar distributions of TFs' numbers (Additional file 3: Figure S12) and pathway genes' numbers (Additional file 3: Figure S13) in terpenoids metabolism also demonstrate the relevance; and the two numbers were all correlated with topological property. For example, one gene pathways (such as β-caryophtllene biosynthesis, δ-carotene biosynthesis and arabidiol biosynthesis) were regulated by a small number of TFs, indicating simple regulations; and these pathways were not critical to metabolic network, because of their low degrees in topology (shown as small nodes with light color in Additional file 3: Figure S12-13). However, pathways with more genes (like genranylgeranyl diphosphate, nonaprenyl diphosphate biosynthesis and epoxysqualene biosynthesis) were regulated by more than 20 TFs and were hub pathways in terpeniod metabolism (shown as large nodes with deep color in Additional file 3: Figure S12-13). The regulations of these pathways were complicated so that perplexing regulatory rewiring often occurred and led metabolic fluxes flowing into disparate downstream pathways. In other words, regulatory rewiring happened with changing conditions, and in turn affected metabolic flux within inner pathway or between different downstream pathways. These examples demonstrate that both simple and complex regulations can adapt to function of metabolic pathways, either specific or extensive.

Contribution of TFs' upstream signaling pathways to regulatory rewiring

Since signaling pathways regulate the activities of TFs, they contribute to TF's regulatory rewiring. To define upstream signaling influences on TFs, we computed significances of expression correlations between plant signaling pathway genes and TFs (see "Methods", Additional file 3: Figure S14). Plant signaling molecules are mainly metabolites, such as jasmonate (fatty acid derivatives biosynthesis), ethylene (methionine biosynthesis), brassinosteroid (terpenoids biosynthesis) and cytokinin

(terpenes biosynthesis). We found that TFs involved in certain signaling pathways were truly significantly co-expressed with related signaling pathway genes. For example, AtMYC2 (AT1G32640, a MYC-related transcriptional activator), in the downstream of jasmonate signaling pathway, was significantly correlated with genes in this pathway; ARR1 (AT3G16857, response regulator 1) and ARR2 (AT4G16110, response regulator 1), activated by cytokinin indirectly, were significantly co-expressed with cytokinin signaling pathway genes. These results indicate the efficiency of predicting TFs' upstream signaling pathways. Totally we found that 50 TFs were significantly correlated with 3 signaling pathways (Additional file 5: Table S4). Among the three signaling pathways, jasmonate and cytokinin signaling pathways were correlated with 45 TFs and 31 TFs respectively while ethylene signaling pathway was only correlated with 12 TFs: suggesting global regulatory function of jasmonate and cytokinin compared to ethylene. In addition, 11 TFs were correlated with three signaling pathways, and 16 TFs were associated with both jasmonate and cytokinin signaling pathways—indicating complicated regulations of signaling pathways on these TFs. Besides, the rest 23 TFs were correlated with only one signaling pathway, implying specific regulations of signaling pathways on these TFs. For example, RAV1 was significantly correlated with cytokinin signaling pathway, demonstrating potential regulation of cytokinin on RAV1. Here, the identification and summary of potential signaling pathways for TFs could largely improve the understanding of TFs' regulatory rewiring.
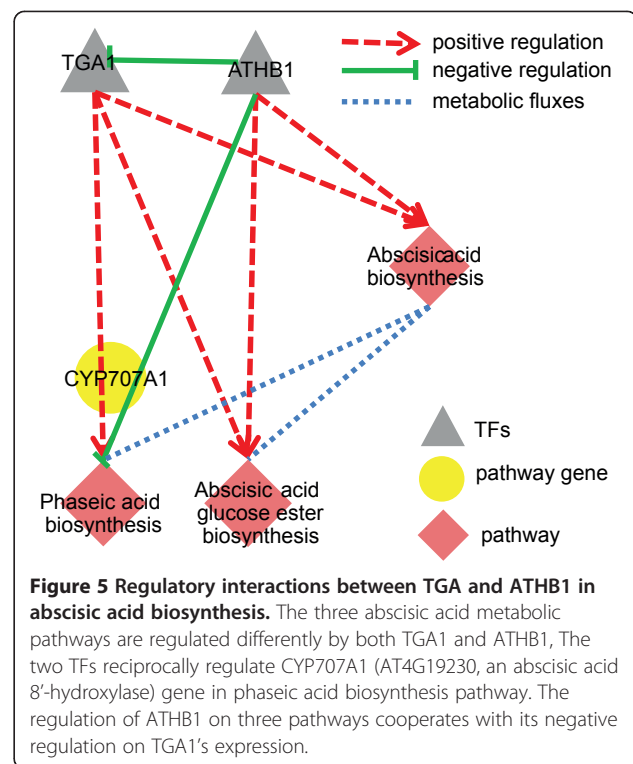
Discussion

Here we presented a method to construct dynamic regulatory network of secondary metabolic pathways. Based on the dynamic regulatory network, we systematically explored complicated regulatory rewiring or crosstalk occurring under distinct experimental conditions, and investigated the relationships between regulatory rewiring and expression, network topology and signaling pathway to unveil the complex regulatory mechanism. The major assumption of our method is that active regulatory interactions are co-expressed which was also

applied in previous studies [53-55]. The activation and inhibition effects can be distinguished by correlation coefficient of TFs and their target genes. Our method of integrating multiple data is efficient in identifying high confident regulatory interactions. As described in result, regulatory interactions determined by three types of data were more reliable than those predicted by single or two type(s) of data. However, most significant co-expression relations are not efficient in prediction because strict criteria of co-expression would filter out gene pairs only co-expressed in a few experiments (this is also why we manipulated arrays within experimental comparisons but not whole arrays).

Based on regulatory alterations at pathway level (rather than genes), we mined regulatory rewiring to comprehensively understand the regulation mechanism of biological metabolic fluxes. For instance, co-regulation of two pathways may attribute to multi-functions of pathway genes shared by these two pathways. Besides, some TFs were considered as dominant regulators because of their frequently changed activities: such as RAV1, which widely regulates growth and developmental genes in Arabidopsis [56]. Furthermore, our result explained the mechanism of TFs regulation on metabolic pathways. For example, flavonoid biosynthesis is influenced by AtLEC2, HY5 and AGL15 [57-59]. Based on our result we discovered these TFs' potential target genes, encoding flavanone 3β-hydroxylase, acetyl-CoA synthetase, 4-coumarate-CoA ligase and naringenin chalcone synthase. And the target genes encoding flavanone 3β-hydroxylase and 4-coumarate-CoA ligase could contribute to regulatory rewiring of HY5, while genes encoding 4-coumarate-CoA ligase and acetyl-CoA synthetase might be the reason of AGL15's regulatory rewiring.

Moreover, our work benefits plant metabolic engineering. A persuasive example is the regulatory crosstalk in abscisic acid metabolism (Figure 5). Abscisic acid biosynthesis is followed by two downstream pathways, abscisic acid glucose ester biosynthesis and phaseic acid biosynthesis. The three pathways are regulated differently by both TGA1(AT5G65210, a bZIP family member) and ATHB1: TGA1 positively regulates three pathways, whereas ATHB1 negatively regulates phaseic acid biosynthesis and positively regulates the other pathways. ATHB1 and TGA1 reciprocally regulate CYP707A1 (AT4G19230, an abscisic acid 8'-hydroxylase) gene in phaseic acid biosynthesis pathway; and the distance between two TFs' binding sites on CYP707A1 promoter is within 200 bps–indicating spatial physical effects of the two TFs on their normal binding processes [55,60]. Besides, regulation of ATHB1 on three pathways cooperates with its negative regulation on TGA1's expression, suggesting ATHB1 should be a key factor in abscisic acid metabolic regulation. Practically, we could overexpress ATHB1 to

Figure 5 Regulatory interactions between TGA and ATHB1 in abscisic acid biosynthesis. The three abscisic acid metabolic pathways are regulated differently by both TGA1 and ATHB1, The two TFs reciprocally regulate CYP707A1 (AT4G19230, an abscisic acid 8'-hydroxylase) gene in phaseic acid biosynthesis pathway. The regulation of ATHB1 on three pathways cooperates with its negative regulation on TGA1's expression.

increase the yield of abscisic acid glucose ester (playing a potential physiological role under water stress) [61] and inhibit phaseic acid metabolic branch at the same time. In conclusion, this example of pathway crosstalk provides a reference to biologists on how to control metabolic products to improve desired plant traits in metabolic engineering.

We also notice that some TFs are not included in our result. This limitation attributes to restricted data sources and incomplete knowledge of regulation mechanism. Firstly, regulatory interactions collected from two databases are incomplete. For example, both AthaMap and AGRIS databses didn't collect some well-known TFs (such as MYB28, MYB29, MYB34, MYB90, MYB12, MYB11, MYB4, MYB58, MYB63, TTG1, MYB75, etc.) and complete regulatory interactions (such as TT8 and TTG2, only have one target gene respectively in AGRIS database). Besides, the number of available microarrays is limited and insufficient to cover various experiments where TFs function–so that some TFs were filtered in microarrays analysis. Secondly, even if data collected in the databases were complete, not all functional TFs can meet our basic assumption and show significance in co-expression, because present knowledge of regulation is unable to definitely identify how well TFs' expressions reflect their function switches. Therefore, the incompleteness of both data-collection and regulation mechanisms impacts the results, which is a common issue in systematic analysis.

Conclusion

This systematic network-based bioinformatics approach largely improves the understanding of perplexed regulatory rewiring mechanism in secondary metabolism and provides useful references for biological experiments, especially metabolic engineering. The approach of reconstructing regulatory network and analyzing regulatory rewiring can be applied to comprehend the whole metabolism in Arabidopsis.

Methods

Data preparation

157 Affymetrix Arabidopsis ATH1 Genome Array platform (GPL198) series of microarrays with complete annotation and more than two duplications (Additional file 6: Table S5) were downloaded from Gene Expression Omnibus (GEO) [62]. Based on annotations, the experimental conditions could be clustered into six categories: biotic stresses, abiotic stresses, genotype experimental comparisons, chemical treatments, tissue experimental comparisons and grown stages. 1,097 experimental comparisons between two treatments within series (batches of experiments) were made for further analysis (Additional file 7: Table S6). Metabolic genes, enzymes and pathways were collected from AraCyc, a biochemical pathway database for Arabidopsis [5] (Additional file 8: Table S7). TFBSs of TFs and 422,967 regulatory interactions predicted by TFBSs alignments were collected from AthaMap [41,63]. 10,653 direct regulatory interactions between TFs and target genes confirmed by experiments from Arabidopsis gene regulatory information server (AGRIS), were used as both positive data and supplement of predicted regulatory interactions. 5 signaling pathways involving 93 genes in Arabidopsis were collected from Database of Cell Signaling (http://stke.sciencemag.org/cgi/collection/pw_plants) (Additional file 9: Table S8).

Microarray data processing

Raw microarrays were preprocessed with RMA function in Bioconductor. Then Limma package in Bioconductor was applied to compute fold change values and P-values of all genes in each experimental comparison. We ranked genes in a descendent order by their absolute fold change values and then selected the top ten percent genes (absolute values of fold change larger than 1.7) with adjusted P-value less than 0.01 as candidate differentially expressed genes (DEGs). Raw P-values were adjusted by Benjamini & Hochberg method in p.adjust function. We kept the number of final DEGs no more than 2,281 (ten percent of all genes designed on GPL198 platform) to make DEGs more meaningful in both technological and biological sense, in consideration of microarray assumption that only a small number of genes are differentially expressed under different conditions.

Computation of significantly and most significantly co-expressed gene pairs

For each experimental comparison, we calculated Pearson Correlation Coefficient (PCC) for each differentially expressed gene pairs using cor function in R. Then we tested the correlation coefficient by cor.test function in R and maintained gene pairs with absolute value of correlation coefficient bigger than 0.9 and adjusted P-value less than 0.05 as correlated gene pairs. Raw P-values were adjusted by Benjamini & Hochberg method in p.adjust function. Considering that the correlated gene pairs appearing only in one experimental comparison or one series might occur by chance, we further measured the statistical significance of correlated gene pairs as the following procedure similar with previous method [64]. We randomly generated gene pairs, keeping the same degree distribution and number of correlated gene pairs within one experimental comparison or series; and then calculated average frequencies of randomly generated gene pairs for experimental comparisons and series respectively. For each experimental comparison or series, DEGs were designated with non-duplicate genes randomly selected; and randomly co-expressed genes pairs were generated by replacing DEGs of original co-expressed gene pairs by the DEGs' designated random genes. The frequencies of these random gene pairs were then counted, with mean value being defined as one random frequency. After repeating this procedure for 100 times, the mean values of random frequencies in experimental comparisons (1.38) and series (1.63) were obtained respectively, with standard deviation less than 0.01. Therefore, correlated gene pairs present in 2 or more experimental comparisons and series were defined as significantly co-expressed gene pairs. Most significantly co-expressed gene pairs were defined as the top ten percent of all correlated gene pairs, the frequencies of which were sorted in a descendent order.

Regulatory network reconstruction and analysis

Co-expressed gene pairs obtained above were used to filter regulatory interaction pairs from both AthaMap and AGRIS, resulting in two sets of co-expressed regulatory interaction pairs. The set of co-expressed regulatory interaction pairs from AthaMap was further filtered by conserved TFBSs. The conserved TFBSs were those detected by sequence alignments in the upstream 3000 bps of transcription starting sites of four sequenced species (Populus trichocarpa, Sorghum bicolor, Brachypodium distachyon and Chlamydomonas reinhardtii) whose genome data were available from Phytozome database (http://www.phytozome.net/). If TFBSs of TFs with

co-expressed regulatory interactions were conserved in these four organisms, related co-expressed regulatory interaction pairs were maintained. As the procedure described in Figure 1, we obtained complete regulatory interactions used for building the regulatory network of secondary metabolism in Arabidopsis. The final network of metabolic pathways was constructed by mapping enzymatic genes to pathways (Additional file 10: Table S9 and Additional file 11: Table S10). Then the topology properties of this network were computed by functions in igraph package in R, and network motif analysis was carried out by FANMOD. Frequent regulatory patterns were defined as the regulatory interaction pairs significantly simultaneously occurred.

The significance of co-expression relationship between signaling pathways and TFs were tested by fisher.test function in R. Finally, the signaling pathways with adjusted P-value less than 0.05 were regarded as significantly correlated with TFs. Raw P-values were adjusted by Benjamini & Hochberg method in p.adjust function.

Additional files

Additional file 1: Table S1. Classification of secondary metabolism.

Additional file 2: Table S2. List of Arabidopsis regulatory databases.

Additional file 3 Figure S1. Significant performance of processed regulatory interactions. **Figure S2.** The regulatory influence of RAV1 on HMG1. **Figure S3.** Regulatory rewirings of RAV1 on epoxysqualene biosynthesis and sterol biosynthesis. **Figure S4.** The contrasts of six experiment categories for differentially expressed TFs. **Figure S5.** The expression profiles of TF-encoding genes and TB genes under five categories of experimental conditions. **Figure S6.** Regulatory network of SDB metabolic pathways. **Figure S7.** Regulatory network of PDB metabolic pathways. **Figure S8.** Regulatory network of FB metabolic pathways. **Figure S9.** Regulatory network of NSCB metabolic pathways. **Figure S10.** Motifs with highest frequencies in secondary metabolic regulatory network. **Figure S11.** Relationship of TF-encoding genes and pathways in the network. **Figure S12.** Distribution of TFs' numbers in TB metabolic pathway network. **Figure S13.** Distribution of pathway genes' numbers in TB metabolic pathway network. **Figure S14.** Significance of co-expression relationship between signal pathways and TFs.

Additional file 4: Table S3. List of predicted regulatory interactions, TF mutant phenotypes in ATPID and pathways' function mined from literatures.

Additional file 5: Table S4. Significantly correlated TFs and signaling pathways.

Additional file 6: Table S5. Information of applied microarray data.

Additional file 7: Table S6. Information of experimental comparisons in microarray analysis.

Additional file 8: Table S7. Information of AraCyc pathways.

Additional file 9: Table S8. 5 signaling pathways of Arabidopsis in Database of Cell Signaling.

Additional file 10: Table S9. Regulatory network of secondary metabolic pathways (filtered by significantly co-expressed gene pairs).

Additional file 11: Table S10. Regulatory network of secondary metabolic pathways (filtered by most significantly co-expressed gene pairs).

Abbreviations

DEGs: Differentially expressed genes; FB: Flavonoids biosynthesis; NSCB: Nitrogen-containing secondary compounds biosynthesis; PDB: Phenylpropanoid derivatives biosynthesis; SDB: Sugar derivatives biosynthesis; TB: Terpenoids biosynthesis; TFs: Transcription factors; TFBSs: Transcription factor binding sites.

Authors' contributions

QL collected all the datasets, analyzed raw data, wrote the draft and participated in study design. RC prepared figures and tables, participated in data collection and analysis, and helped to revise draft. TS designed the study and helped to draft and finalized the manuscript. All authors read and approved the final manuscript.

References

1. Bourgaud F, Gravot A, Milesi S, Gontier E: **Production of plant secondary metabolites: a historical perspective.** *Plant Sci* 2001, **161**(5):839–851.
2. Hartmann T: **From waste products to ecochemicals: fifty years research of plant secondary metabolism.** *Phytochemistry* 2007, **68**(22-24):2831–2846.
3. Zhao J, Davis LC, Verpoorte R: **Elicitor signal transduction leading to production of plant secondary metabolites.** *Biotechnol Adv* 2005, **23**(4):283–333.
4. Broun P: **Transcriptional control of flavonoid biosynthesis: a complex network of conserved regulators involved in multiple aspects of differentiation in Arabidopsis.** *Curr Opin Plant Biol* 2005, **8**(3):272–279.
5. Mueller LA, Zhang P, Rhee SY: **AraCyc: a biochemical pathway database for Arabidopsis.** *Plant Physiol* 2003, **132**(2):453–460.
6. Sonderby IE, Geu-Flores F, Halkier BA: **Biosynthesis of glucosinolates-gene discovery and beyond.** *Trends Plant Sci* 2010, **15**(5):283–290.
7. Wittstock U, Halkier BA: **Glucosinolate research in the Arabidopsis era.** *Trends Plant Sci* 2002, **7**(6):263–270.
8. Davies KM, Schwinn KE: **Transcriptional regulation of secondary metabolism.** *Funct Plant Biol* 2003, **30**(9):913–925.
9. Vernon DM, Tarczynski MC, Jensen RG, Bohnert HJ: **Cyclitol production in transgenic tobacco.** *Plant J* 1993, **4**(1):199–205.
10. Dixon RA, Paiva NL: **Stress-induced phenylpropanoid metabolism.** *Plant Cell* 1995, **7**(7):1085–1097.
11. Dixon RA, Achnine L, Kota P, Liu CJ, Reddy MS, Wang L: **The phenylpropanoid pathway and plant defence-a genomics perspective.** *Mol Plant Pathol* 2002, **3**(5):371–390.
12. Vogt T: **Phenylpropanoid biosynthesis.** *Mol Plant* 2010, **3**(1):2–20.
13. Noel JP, Austin MB, Bomati EK: **Structure-function relationships in plant phenylpropanoid biosynthesis.** *Curr Opin Plant Biol* 2005, **8**(3):249–253.
14. Kubasek WL, Shirley BW, McKillop A, Goodman HM, Briggs W, Ausubel FM: **Regulation of flavonoid biosynthetic genes in germinating Arabidopsis seedlings.** *Plant Cell* 1992, **4**(10):1229–1236.
15. Koes R, Verweij W, Quattrocchio F: **Flavonoids: a colorful model for the regulation and evolution of biochemical pathways.** *Trends Plant Sci* 2005, **10**(5):236–242.
16. Bohlmann J, Gershenzon J: **Old substrates for new enzymes of terpenoid biosynthesis.** *Proc Natl Acad Sci U S A* 2009, **106**(26):10402–10403.
17. Gershenzon J, Dudareva N: **The function of terpene natural products in the natural world.** *Nat Chem Biol* 2007, **3**(7):408–414.
18. Aharoni A, Giri AP, Deuerlein S, Griepink F, De Kogel WJ, Verstappen FW, Verhoeven HA, Jongsma MA, Schwab W, Bouwmeester HJ: **Terpenoid metabolism in wild-type and transgenic Arabidopsis plants.** *Plant Cell* 2003, **15**(12):2866–2884.

19. Yang CQ, Fang X, Wu XM, Mao YB, Wang LJ, Chen XY: **Transcriptional regulation of plant secondary metabolism.** *J Integr Plant Biol* 2012, **54**(10):703–712.

20. Nascimento NC, Fett-Neto AG: **Plant secondary metabolism and challenges in modifying its operation: an overview.** *Methods Mol Biol* 2010, **643**:1–13.

21. Nielsen J, Olsson L: **An expanded role for microbial physiology in metabolic engineering and functional genomics: moving towards systems biology.** *FEMS Yeast Res* 2002, **2**(2):175–181.

22. Sweetlove LJ, Last RL, Fernie AR: **Predictive metabolic engineering: a goal for systems biology.** *Plant Physiol* 2003, **132**(2):420–425.

23. Meinke DW, Cherry JM, Dean C, Rounsley SD, Koornneef M: **Arabidopsis thaliana: a model plant for genome analysis.** *Science* 1998, **282**(5389):662. 679-682.

24. The Arabidopsis Genome Initiative: **Analysis of the genome sequence of the flowering plant Arabidopsis thaliana.** *Nature* 2000, **408**(6814):796–815.

25. Braxton S, Bedilion T: **The integration of microarray information in the drug development process.** *Curr Opin Biotechnol* 1998, **9**(6):643–649.

26. Baldwin D, Crane V, Rice D: **A comparison of gel-based, nylon filter and microarray techniques to detect differential RNA expression in plants.** *Curr Opin Plant Biol* 1999, **2**(2):96–103.

27. Bednar M: **DNA microarray technology and application.** *Med Sci Monit* 2000, **6**(4):796–800.

28. Harkin DP: **Uncovering functionally relevant signaling pathways using microarray-based expression profiling.** *Oncologist* 2000, **5**(6):501–507.

29. Yan W, Zhu H, Yang Y, Chen J, Zhang Y, Shen B: **Effects of time point measurement on the reconstruction of gene regulatory networks.** *Molecules* 2010, **15**(8):5354–5368.

30. Zou M, Conzen SD: **A new dynamic Bayesian network (DBN) approach for identifying gene regulatory networks from time course microarray data.** *Bioinformatics* 2005, **21**(1):71–79.

31. Vandepoele K, Quimbaya M, Casneuf T, De Veylder L, Van de Peer Y: **Unraveling transcriptional control in Arabidopsis using cis-regulatory elements and coexpression networks.** *Plant Physiol* 2009, **150**(2):535–546.

32. Terpstra IR, Snoek LB, Keurentjes JJ, Peeters AJ, van den Ackerveken G: **Regulatory network identification by genetical genomics: signaling downstream of the Arabidopsis receptor-like kinase ERECTA.** *Plant Physiol* 2010, **154**(3):1067–1078.

33. Keurentjes JJ, Fu J, Terpstra IR, Garcia JM, van den Ackerveken G, Snoek LB, Peeters AJ, Vreugdenhil D, Koornneef M, Jansen RC: **Regulatory network construction in Arabidopsis by using genome-wide gene expression quantitative trait loci.** *Proc Natl Acad Sci U S A* 2007, **104**(5):1708–1713.

34. Small KS, Hedman AK, Grundberg E, Nica AC, Thorleifsson G, Kong A, Thorsteindottir U, Shin SY, Richards HB, Consortium G, Investigators M, Consortium D, Soranzo N, Ahmadi KR, Lindgren CM, Stefansson K, Dermitzakis ET, Deloukas P, Spector TD, McCarthy MI, Mu TC: **Identification of an imprinted master trans regulator at the KLF14 locus related to multiple metabolic phenotypes.** *Nat Genet* 2011, **43**(6):561–564.

35. Shinozaki K, Yamaguchi-Shinozaki K, Seki M: **Regulatory network of gene expression in the drought and cold stress responses.** *Curr Opin Plant Biol* 2003, **6**(5):410–417.

36. Nakashima K, Ito Y, Yamaguchi-Shinozaki K: **Transcriptional regulatory networks in response to abiotic stresses in Arabidopsis and grasses.** *Plant Physiol* 2009, **149**(1):88–95.

37. Espinosa-Soto C, Padilla-Longoria P, Alvarez-Buylla ER: **A gene regulatory network model for cell-fate determination during Arabidopsis thaliana flower development that is robust and recovers experimental gene expression profiles.** *Plant Cell* 2004, **16**(11):2923–2939.

38. To A, Valon C, Savino G, Guilleminot J, Devic M, Giraudat J, Parcy F: **A network of local and redundant gene regulation governs Arabidopsis seed maturation.** *Plant Cell* 2006, **18**(7):1642–1651.

39. Zhang X, Jiang T, Yu Y, Wu Z, Jiang S, Lu K, Feng X, Liang S, Lu Y, Wang X, Zhang D: **Arabidopsis co-chaperonin CPN20 antagonizes Mg-chelatase H subunit to derepress ABA-responsive WRKY40 transcription repressor.** *Sci China Life Sci* 2014, **57**(1):11–21.

40. Luan W, Shen A, Jin Z, Song S, Li Z, Sha A: **Knockdown of OsHox33, a member of the class III homeodomain-leucine zipper gene family, accelerates leaf senescence in rice.** *Sci China Life Sci* 2013, **56**(12):1113–1123.

41. Bulow L, Engelmann S, Schindler M, Hehl R: **AthaMap, integrating transcriptional and post-transcriptional data.** *Nucleic Acids Res* 2009, **37**(Database issue):D983–D986.

42. Palaniswamy SK, James S, Sun H, Lamb RS, Davuluri RV, Grotewold E: **AGRIS and AtRegNet. a platform to link cis-regulatory elements and transcription factors into regulatory networks.** *Plant Physiol* 2006, **140**(3):818–829.

43. Yilmaz A, Mejia-Guerra MK, Kurz K, Liang X, Welch L, Grotewold E: **AGRIS: the Arabidopsis Gene Regulatory Information Server, an update.** *Nucleic Acids Res* 2011, **39**(Database issue):D1118–D1122.

44. Li P, Zang W, Li Y, Xu F, Wang J, Shi T: **AtPID: the overall hierarchical functional protein interaction network interface and analytic platform for Arabidopsis.** *Nucleic Acids Res* 2011, **39**(Database issue):D1130–D1133.

45. Xu X, Chen C, Fan B, Chen Z: **Physical and functional interactions between pathogen-induced Arabidopsis WRKY18, WRKY40, and WRKY60 transcription factors.** *Plant Cell* 2006, **18**(5):1310–1326.

46. Robatzek S, Somssich IE: **A new member of the Arabidopsis WRKY transcription factor family, AtWRKY6, is associated with both senescence- and defence-related processes.** *Plant J* 2001, **28**(2):123–133.

47. Tran LS, Nakashima K, Sakuma Y, Simpson SD, Fujita Y, Maruyama K, Fujita M, Seki M, Shinozaki K, Yamaguchi-Shinozaki K: **Isolation and functional analysis of Arabidopsis stress-inducible NAC transcription factors that bind to a drought-responsive cis-element in the early responsive to dehydration stress 1 promoter.** *Plant Cell* 2004, **16**(9):2481–2498.

48. Vandepoele K, Raes J, De Veylder L, Rouze P, Rombauts S, Inze D: **Genome-wide analysis of core cell cycle genes in Arabidopsis.** *Plant Cell* 2002, **14**(4):903–916.

49. Devaiah BN, Nagarajan VK, Raghothama KG: **Phosphate homeostasis and root development in Arabidopsis are synchronized by the zinc finger transcription factor ZAT6.** *Plant Physiol* 2007, **145**(1):147–159.

50. Kim S, Choi K, Park C, Hwang HJ, Lee I: **SUPPRESSOR OF FRIGIDA4, encoding a C2H2-Type zinc finger protein, represses flowering by transcriptional activation of Arabidopsis FLOWERING LOCUS C.** *Plant Cell* 2006, **18**(11):2985–2998.

51. Feurtado JA, Huang D, Wicki-Stordeur L, Hemstock LE, Potentier MS, Tsang EW, Cutler AJ: **The Arabidopsis C2H2 zinc finger INDETERMINATE DOMAIN1/ENHYDROUS promotes the transition to germination by regulating light and hormonal signaling during seed maturation.** *Plant Cell* 2011, **23**(5):1772–1794.

52. Luerssen H, Kirik V, Herrmann P, Misera S: **FUSCA3 encodes a protein with a conserved VP1/AB13-like B3 domain which is of functional importance for the regulation of seed maturation in Arabidopsis thaliana.** *Plant J* 1998, **15**(6):755–764.

53. Annicotte JS, Fayard E, Swift GH, Selander L, Edlund H, Tanaka T, Kodama T, Schoonjans K, Auwerx J: **Pancreatic-duodenal homeobox 1 regulates expression of liver receptor homolog 1 during pancreas development.** *Mol Cell Biol* 2003, **23**(19):6713–6724.

54. Gu P, Goodwin B, Chung AC, Xu X, Wheeler DA, Price RR, Galardi C, Peng L, Latour AM, Koller BH, Gossen J, Kliewer SA, Cooney AJ: **Orphan nuclear receptor LRH-1 is required to maintain Oct4 expression at the epiblast stage of embryonic development.** *Mol Cell Biol* 2005, **25**(9):3492–3505.

55. Rausa FM, Tan Y, Costa RH: **Association between hepatocyte nuclear factor 6 (HNF-6) and FoxA2 DNA binding domains stimulates FoxA2 transcriptional activity but inhibits HNF-6 DNA binding.** *Mol Cell Biol* 2003, **23**(2):437–449.

56. Hu YX, Wang YH, Liu XF, Li JY: **Arabidopsis RAV1 is down-regulated by brassinosteroid and may act as a negative regulator during plant development.** *Cell Res* 2004, **14**(1):8–15.

57. Guo F, Liu C, Xia H, Bi Y, Zhao C, Zhao S, Hou L, Li F, Wang X: **Induced expression of AtLEC1 and AtLEC2 differentially promotes somatic embryogenesis in transgenic tobacco plants.** *PLoS One* 2013, **8**(8):e71714.

58. Sibout R, Sukumar P, Hettiarachchi C, Holm M, Muday GK, Hardtke CS: **Opposite root growth phenotypes of hy5 versus hy5 hyh mutants correlate with increased constitutive auxin signaling.** *PLoS Genet* 2006, **2**(11):e202.

59. Zheng Q, Perry SE: **Alterations in the Transcriptome of Soybean in Response to Enhanced Somatic Embryogenesis Promoted by Orthologs of AGAMOUS-Like15 and AGAMOUS-Like18.** *Plant Physiol* 2014, **164**(3):1365–1377.

60. Costamagna E, Garcia B, Santisteban P: **The functional interaction between the paired domain transcription factor Pax8 and Smad3 is involved in transforming growth factor-beta repression of the sodium/iodide symporter gene.** *J Biol Chem* 2004, **279**(5):3439–3446.

61. Neill SJ, Horgan R, Heald JK: **Determination of the levels of abscisic acid-glucose ester in plants.** *Planta* 1983, **157**(4):371–375.

62. Barrett T, Troup DB, Wilhite SE, Ledoux P, Rudnev D, Evangelista C, Kim IF, Soboleva A, Tomashevsky M, Marshall KA, Phillippy KH, Sherman PM, Muertter RN, Edgar R: **NCBI GEO: archive for high-throughput functional genomic data.** *Nucleic Acids Res* 2009, **37**(Database issue):D885–D890.
63. Steffens NO, Galuschka C, Schindler M, Bulow L, Hehl R: **AthaMap: an online resource for in silico transcription factor binding sites in the Arabidopsis thaliana genome.** *Nucleic Acids Res* 2004, **32**(Database issue): D368–D372.
64. Lee HK, Hsu AK, Sajdak J, Qin J, Pavlidis P: **Coexpression analysis of human genes across many microarray data sets.** *Genome Res* 2004, **14**(6):1085–1094.