INTRODUCTION

Despite the laudable aim of the Arabidopsis 2010 project we remain a long way from knowing the function of every gene of this plant, notwithstanding unprecedented research effort with recent estimates suggesting in the region of 50% of genes are functionally annotated by gene homology and between 10 and 15% have an experimentally verified biological function (Saito et al., 2008; Tohge and Fernie, 2010, Mutwil et al., 2011). The simplicity of homology searches means that at least for dicots the number of genes annotated by homology in rice and soybean or in the more recently published maize (Schnable et al., 2009), poplar (Tuskan et al., 2006), or tomato (Tomato Genome Consortium, 2012) genomes remains reasonable. However, the proportion of genes for which function has been verified experimentally is, at least in most of these species, negligible rendering predictive gene annotation and subsequent validation thereof a vital task for genomics both in model and crop species.

The development and widespread adoption of unbiased RNA sequencing (RNAseq) approaches by plant researchers (Bao et al., 2011; Matsa et al., 2011; Hamilton and Buell, 2012; Lohse et al., 2012) effectively increases the scale of this task since it circumvents the need for in depth a priori knowledge that was a pre-requisite for microarray hybridizations. Despite the fact that we have as yet not reached satisfactory levels of gene annotation several approaches – all of which are based on a common principal – have recently greatly facilitated gene annotation. This is particularly the case of pathways under strict transcriptional regulation such as cell wall associated genes and those involved in the various pathways of secondary metabolism as well as leading to the classification of process-associated gene including those linked to cold stress and jasmonate signaling, operon-like genes and seed germination (Hannah et al., 2005; McGrath et al., 2005; Tohge et al., 2005; Saito et al., 2008; Senivasasangendra et al., 2008; Mutwil et al., 2009, Obayashi et al., 2009; Usadel et al., 2009; Ogata et al., 2010; Tohge and Fernie, 2010; Basal et al., 2011; Wada et al., 2012).

Whole genome sequencing, the relative ease of transcript profiling by the use of microarrays and latterly RNA sequencing approaches have facilitated the capture of vast amounts of transcript data. However, despite the enormous progress made in gene annotation a substantial proportion of genes remain to be annotated at the functional level. Considerable progress has, however, been made by searching for transcriptional coordination between genes of known function and non-annotated genes on the premise that such co-expressed genes tend to be functionally related. Here we review progress made following this approach as well as its expansion to include phenotypic information from other levels of cellular organization such as proteomic and metabolomic data as well as physiological and developmental phenotypes.

Keywords: gene annotation, network analysis, gene expression, plant metabolism, correlation analysis

PREDICTION OF THE FUNCTION OF ARABIDOPSIS GENES

In spite of the clear advantage of biological co-expression network approaches based on gene expression, protein interaction, and genetic interactions for microorganisms such as yeast (see an example, Zhang et al., 2005), co-expression network approaches in plant research have largely been developed solely on the basis of microarray data. This has revealed clear correlations between genes in multiple biosynthetic pathways (Tohge et al., 2007; Morshed et al., 2011; Mutwil et al., 2011). In addition, Arabidopsis thaliana is currently the most useful model plant for integrative analysis due to the availability of several resources such as knockout mutants, cDNA library, tag counts of ESTs, microarray, data and metabolite profiling data. Furthermore, several co-expression gene network analyses and integrative analysis with metabolite profiles have been used to understand the transcriptional correlation networks and discover novel gene functions in this species (Noji et al., 2006; Saito et al., 2008; Ma et al., 2009; Tohge and Fernie, 2010; Mutwil et al., 2011). For this purpose, several web-based
co-expression applications, for example ATTED-II (Obayashi et al., 2009, 2011), AraNet (Hwang et al., 2011), Expression Angler of the Bio-Array Resource (BAR; Tooufighi et al., 2005), Cress-Express (Srinivasasainagendra et al., 2008), CBSD (Steinhauer et al., 2004), KappaViewer (Sakurai et al., 2011), GeneCAT (Murvil et al., 2008), GeneVestigator (Zimmermann et al., 2004), OryzaExpress (Hamada et al., 2011), and VirtualPlant (Katari et al., 2010) have been developed (Table 1).

One of the best examples of co-expression analysis in Arabidopsis is cellulose synthase (CESA) genes in the secondary cell wall metabolism (Brown et al., 2005; Persson et al., 2005b), and the primary wall hemicellulose xyloglucan (Cocorun et al., 2007). These studies used the three major secondary wall CESA genes as the baits to construct networks and find novel functional genes displaying similar expression patterns. The CESA gene network has been in several publications (Persson et al., 2005a; Murvil et al., 2008, 2009, 2011; Ruprecht et al., 2011). A second successful example of the co-expression approach is that of plant secondary metabolism, since this is of the directly regulated at the transcriptional level by a range of different transcription factors including the MYB transcription factors. Since a framework of flavonoid co-expression network was constructed for identify the flavonol-3′-O-methyltransferase (AtOMT1; Tohge et al., 2007), such co-expression network approaches have been expanded to find other flavonoid biosynthetic genes such as flavonol-7-O-cinnamoyltransferase (AtFCY6009) and flavonol-3-O-arabinoxygenase transferase (AtFLO2004; Yonkura-Sakakibara et al., 2007, 2008; Tohge and Fernie, 2010). In addition, this approach was also utilized in the identification of glucosinolate MYB regulators (AtMYB29 and AtMYB29; Hirai et al., 2007), monolignol transporter (AtABC29) involved in lignin biosynthesis (Alejandro et al., 2012) and novel signaling related candidate genes and transporters following the exposure of Arabidopsis to UV-B (Tohge et al., 2011a).

In addition to its utility in understanding the regulation of individual metabolic pathways or even metabolic networks co-expression analysis has also been applied at a much broader level to look at tissue-specific transcriptional networks (Song et al., 2010) and at diverse biological processes including seed germination and dark-induced senescence (Araújo et al., 2011), however these remain to be functionally verified. The combined findings were thus taken to suggest that seed dormancy is an adaptive trait that arose evolutionarily late and evolved by coopting existing biosynthetic pathways regulating cellular phase transitions and abiotic stress response genes. During dark-induced senescence there is a dramatic switch from respiration of sugars to respiration of protein which is underpinned by dramatic transcriptional reprogramming of metabolism (Araújo et al., 2010; Araújo et al., 2011), including the degradation of heme and branched chain amino acids by as yet undefined pathways. In this case the co-expression response was able to provide a high number of candidate genes involved in this process (Araújo et al., 2011), however these remain to be functionally verified.

**PREDICTION OF THE FUNCTION OF CROP GENES: WITHIN SPECIES COMPARISONS**

Although considerably fewer microarray experiments have been reported for crop species, with the possible exceptions of rice, several examples exist of the power of the approach in standalone network analyses for rice and tomato (Ficklin et al., 2010; Ozaki et al., 2010; Rohmann et al., 2011; Sakurai et al., 2011; Fukushima et al., 2012). We will here shortly review these studies and highlight the important knowledge inference for studies in tomato and the grasses. In tomato the most comprehensive

---

**Table 1 | Co-expression databases presented in this article.**

| Species        | Database URL                                                      |
|----------------|------------------------------------------------------------------|
| Arabidopsis thaliana | ATTED-II http://atted.jp/                                           |
| Arabidopsis thaliana | AraNet http://www.functionalnet.org/aranaet/                        |
| Arabidopsis thaliana | BAR http://bar.utoronto.ca/welcome.htm                             |
| Arabidopsis thaliana | COP http://webs2.kazusa.or.jp/kagiana/cop                          |
| Arabidopsis thaliana | CBSD http://cbdb.mpimp-golm.mpg.de/cbldb/abc/ash.html            |
| Arabidopsis thaliana | CressExpress http://www.cressexpress.org/index                     |
| Arabidopsis thaliana | GeneVestigator http://genevestigator.com/angelator.py             |
| Arabidopsis thaliana | GeneCAT http://genecat.mpg.de/cbg-bio/Angelator.py                |
| Arabidopsis thaliana | Seaweed http://froot.cs.nott.ac.uk/arabidopsis/                     |
| Oryza sativa     | OryzaExpress http://bioinf.mind.meiji.ac.jp/plice_network_public/script/index.html |
| Oryza sativa     | ATTED-II http://atted.jp/                                           |
| Populus trichocarpa | COP http://webs2.kazusa.or.jp/kagiana/cop                          |

---

"fpls-03-00248" — 2012/11/6 — 22:03 — page 2 — #2
study was that performed by Fukushima et al. (2012) who con-
structed co-response networks from 327 tomato Affymetrix arrays.
Although this dataset was substantially smaller than that regularly
used for Arabidopsis a number of important conclusion could
be drawn including biologically relevant co-expression networks
including DNA endoreduplication, response to cold, jasmonate-
associated metabolic processes, and the ubiquitous photosynthetic
gene cluster. The study also revealed that duplicated genes often
displayed differential co-expression when tissue-type was studied
a fact highlighted by genes of lycopene and flavonoid biosyn-
thesis (Fukushima et al., 2012). In two more targeted analyses
co-expression analysis was also linked to metabolite levels in
tomato fruit (Rohrmann et al., 2011; Lee et al., 2012), however,
we will return to these studies later when discussing layering in
other phenotypes to aid annotation strategies.

In addition to these recent studies in tomato there have also
been studies in barley (Hordeum vulgare; Faccoli et al., 2005;
Mochida et al., 2011; Tohge et al., 2011b), wheat (Manickavelu
et al., 2012), rice (Fukushima et al., 2009; Lee et al., 2009; Ficklin
et al., 2010; Cháls et al., 2011; Hamada et al., 2011), maize (Fick-
lin and Feltrin, 2011), poplar (Populus spp.; Ogata et al., 2010),
and tobacco (Nicotiana tabacum; Edwards et al., 2010). Studies in rice
revealed that gene co-expression analysis facilitated elucidation
of gene function. With the study of Ficklin et al. (2011) returning
45 co-expressed gene modules and 76 cofunctional gene clusters
some of which were enriched for previously characterized mutant
phenotypes thus providing strong hints toward molecular func-
tions of unknown genes within the clusters with similar outcomes
being achieved for the other species mentioned above.

PREDICTION OF THE FUNCTION OF CROP GENES: BETWEEN
SPECIES COMPARISONS

Whilst the above described studies show that there is consider-
able benefit from co-expression analysis in species such as tomato
for which genome scale microarray platforms do not yet exist
another approach that has been demonstrated to be highly pow-
erful is combining comparisons of gene cluster networks and
sequence homology as a method of assigning gene function and
was recently published under the acronym PlaNet (Mutwil et al.,
2011). PlaNet builds on the concept first published in 2008 by the
same group which already described the search for barley gene
orthologs of annotated Arabidopsis genes (Mutwil et al., 2008).
PlaNet extended this to include the crop species barley, medicago,
poplar, rice, soybean, and wheat, and used a comparative network
algorithm to estimate similarities between network structures.
The algorithm was exemplified using the canonical the photosystem I
reaction center (PSA-D) family gene-related networks as well as
those related to chalcone synthase suggesting that the rapid trans-
fer of knowledge between species will be possible. That this is so
was recently also demonstrated by the same group in a study of
secondary wall cellulose biosynthesis (Ruprecht et al., 2011). In
this study, the authors compared co-expressed gene vicinity net-
works of primary and secondary wall CESAs in all species housed
in PlaNet to identify those genes consistently co-regulated with
cellulose biosynthesis. In addition to the expected polysaccharide
acting enzymes, they also found many gene families associated
with cytoskeleton, signaling, transcriptional regulation, oxidation,
and protein degradation. Based on these analyses, they selected
and biochemically analyzed T-DNA insertion lines correspond-
ing to approximately 20 genes from gene families that re-occur in
the co-expressed gene vicinity networks of secondary wall CESAs
across the seven species. One of the mutants, corresponding to a
pinosylvin reductase gene, was subsequently characterized as dis-
playing disturbed xylem morphology and containing lower levels
of lignin than the wild-type.

The very same seven species used for the PlaNet study were
used in an independent study to generate a pipeline within the
BAR software suite (Toufighi et al., 2005) to rank ortholog predic-
tions based on sequence and expression profile similarity with the
best fitting on this criteria being defined as the expressolog (Patel
et al., 2012). Interestingly, global analyses revealed that orthologs
with the highest sequence similarity do not necessarily exhibit
the highest expression pattern similarity. Moreover, other putative
orthologs show highly distinct expression patterns suggesting they
may need re-annotating or at best to be given a more specific anno-
tation. A similar comprehensive comparison between maize and
rice was additionally recently carried out using the IsoRank tool
(Ficklin and Feltrin, 2011). It thus appears likely that both these
tools as well as PlaNet will likely greatly aid translational efforts
to translate the huge knowledge we have gained from Arabidopsis
studies into crop species.

LAYERING IN OTHER PHENOTYPES TO AID ANNOTATION
STRATEGIES

The above examples have by and large only relied on data
from transcript profiling and have neither harnessed informa-
tion derived from other molecular approaches, such as proteomics
and metabolomics, nor indeed of end-phenotypes such as total
yield and harvest indexes. Several recent studies have however
incorporated such data collected in order to complement tran-
scriptomic efforts of gene functional annotation (Hirai et al., 2007;
Horan et al., 2008; Yonekura-Sakakibara et al., 2008; Subiza et al.,
2009; Allen et al., 2010; Tohge and Fernie, 2010; Araujo et al.,
2011; Rohrmann et al., 2011; Tohge et al., 2011a,b). Returning
to the tomato examples mentioned above, in order to exploit the
impact of tomato genetic diversity on carotenoids, Lee et al.
(2012) used Solanum pennellii introgression lines as a source of
defined natural variation and as a resource for the identifica-
tion of candidate regulatory genes. For this purpose ripe fruits
were analyzed for numerous fruit metabolites and transcriptome
profiles generated using a 12,000 unigene oligoarray. Correlation
analysis between carotenoid content and gene expression profiles
revealed 953 carotenoid-correlated genes. A subnetwork analysis
of carotenoid-correlated transcription narrowed this down to 38
candidates. One of which, Solanum lycopersicum ethylene response
factor 6 (SERRP6), was subsequently functionally characterized
revealing that it indeed influences carotenoid biosynthesis and
additional ripening phenotypes. In a similar approach Rohrmann
et al. (2011) developed a quantitative real-time PCR platform
allowing accurate quantification of the expression level of approx-
imately 1000 tomato transcription factors. In addition to utilizing
this novel approach, they performed cDNA microarray analysis
and metabolite profiling of primary and secondary metabolites
using gas chromatography–mass spectrometry (GC–MS) and
liquid chromatography–mass spectrometry (LC–MS), respectively. Applying these platforms to pericarp material harvested throughout fruit development and studying both wild-type Solanum lycopersicum cv. Ailsa Craig and the fpl1 (high pigment) mutant which is functionally deficient in the tomato homolog of the negative regulator of the light signal transduction gene UV-DAMAGED DNA BINDING PROTEIN 1 (ZDBU) from Arabidopsis. They chose this particular mutant since it had previously been shown to harbor dramatic alterations in the content of several important fruit metabolites but relatively little impact on other ripening phenotypes. The combined dataset was extensively mined for co-responsive metabolites and transcription factors, and, where possible, the respective transcriptional expression network underlying this control. Two further studies in tomato merit discussion here. Moutet et al. (2009), used a combination of metabolite profiling and transcript profiling to identify candidate for the key factor of fruit composition and development. More recently Osoiri et al. (2011), used a combination of transcriptionomics, proteomics, and metabolomics alongside network computation to assess ripening across a range of classical ripening mutants and recently extended this analysis to compare ripening in tomato with that in pepper (Osoiri et al., 2012).

Staying with the integration of transcriptomic, proteomic, and metabolomic data we recently combined data from all three platforms to infer function within the tonoplast proteome (Tohge et al., 2011b). In order to do so we performed metabolic profiling of both primary and secondary metabolites in highly purified vacuoles of barley or the protoplast preparations from which they were isolated. This gave us quantitative data on 59 primary metabolites for which we knew the exact chemical formulae. This data was then compared to the 88 tonoplast proteins reported for barley (Endler et al., 2008) and evaluating there co-expression using PlaNet. This strategy allowed us to putatively assign transport function for phenylpropanoids, flavonoids, storage proteins, and mucin-like acid, as well as a potential transport system for phytoxidiphosphates.

Proteomic data are also an important component of the interaction networks that form part of the CORNET tools (De Bodt et al., 2010, 2012) which combine co-expression analysis with protein–protein interaction searches. The latter is similar to other tools compared to the 88 tonoplast proteins reported for barley (Endler et al., 2008) and evaluating there co-expression using PlaNet. This strategy allowed us to putatively assign transport function for phenylpropanoids, flavonoids, storage proteins, and mucin-like acid, as well as a potential transport system for phytoxidiphosphates.

Proteomic data are also an important component of the interaction networks that form part of the CORNET tools (De Bodt et al., 2010, 2012) which combine co-expression analysis with protein–protein interaction searches. The latter is similar to other tools compared to the 88 tonoplast proteins reported for barley (Endler et al., 2008) and evaluating there co-expression using PlaNet. This strategy allowed us to putatively assign transport function for phenylpropanoids, flavonoids, storage proteins, and mucin-like acid, as well as a potential transport system for phytoxidiphosphates.

**REFERENCES**

Alessandro, S., Lee, Y., Tohge, T., Stand, D., Osoiri, S., Park, J., et al. (2012). ArabiC28 is a monolignol transporter involved in lignin biosynthesis. *Cell Rep.* 10, 1207–1212.

Allen, E., Moing, A., Ebbels, T. M., Marchant, R., Jamieson, A. D., Rank, D., et al. (2010). Correlation network analysis reveals a sequential retroconversion of metabolic and transcriptional states during germination and gene-metabolite relationships in developing seedlings of Arabidopsis. *BMC Syst. Biol.* 4, 62. doi:10.1186/1752-0509-4-62

Araujo, W. L., Ishizaki, K., Nansen, N., Larson, T. R., Tohge, T., Krahnert, I., et al. (2010). Identification of the 3-hydroxyacetonephosphate isomerase and 1,2-Cyclohexane dicarboxylic acid dehydrogenase as alternative electron donors linking isocitrate to reactive oxygen species. *Biochem. Biophys. Res. Commun.* 389, 1285–1290.

Araya, W. L., Ishizaki, K., Nansen, N., Larson, T. R., Tohge, T., Krahnert, I., et al. (2011). Identification of a range of catalytic mutants provides evidence that phytanoyl-CoA dehydrogenase A does not act as a substrate for the electron-transfer flavoprotein/electron-transfer flavoprotein ubiquinone oxidoreductase complex in Arabidopsis during dark-induced senescence. *Plant Physiol.* 157, 55–69.

Bao, E., Jiang, T., Kaloshian, I., and Goodacre, R. (2012). Identification of novel genes in Arabidopsis involved in secondary cell wall formation using expression profiling and reverse genetics. *Plant Cell Rep.* 31, 2201–2213.

Bass, G. W., Lan, H., Glärn, E., Gibbs, D. J., Gerrens, T., Krzanov- gos, N., et al. (2011). Genome-wide network model capturing seed germination reveals coordinated regulation of plant cellular phase transitions. *Proc. Natl. Acad. Sci. USA.* 108, 9779–9784.

Brown, D. M., Ziel, L. A. H., Ellis, J., Goodacre, R., and Turner, S. R. (2005). Identification of novel genes in Arabidopsis involved in secondary cell wall formation using expression profiling and reverse genetics. *Plant Cell Rep.* 24, 1021–1029.

Chan, W. S., Goodacre, R., and Turner, S. R. (2003). Identification of novel genes in Arabidopsis involved in secondary cell wall formation using expression profiling and reverse genetics. *Plant Cell Rep.* 24, 1021–1029.

CONCLUSION

It is hopefully clearly apparent from this mini-review that co-expression analyses are a very powerful tool in gene annotation not only in model systems such as Arabidopsis and rice but also in less well characterized plant species. To date, it has found great utility in improving our understanding of pathways which are known to be regulated at the transcriptional level such as cell wall biosynthesis and various pathways of secondary metabolism, however, recent examples also demonstrate its utility in elucidating novel players in various developmental processes. The guilt-by-association response is clearly powerful even in stand-alone single species approaches. However, the increasing availability of data from multiple species and at multiple different levels of the cellular hierarchy will likely facilitate the adoption of integrative genomics approaches by many more laboratories in the near future. Even some 6–9 years ago the power of combining transcript and metabolite profiling for (candidate) gene discovery was demonstrated for non-sequenced species (Urbanzczik-Wodniak et al., 2003; Rische et al., 2006). Recent developments in RNA sequencing (Schneeberger and Weigel, 2011), will likely render this considerably easier in the near future.

ACKNOWLEDGMENTS

Research activity of Takayuki Tohge is supported by the Alexander von Humboldt Foundation. Funding from the Max-Planck Society (to Takayuki Tohge and Alison R. Fernie) is gratefully acknowledged.
Choe, J., Lalonde, S., Obroljak, P., Nesteriuc, V.-A., Park, S., Vilar-Airola, M., et al. (2012). Uncovering Arabidopsis membrane protein interactions enriched in transporters using state-dependent co-expression gene annotation. Front Plant Sci. 249–49.

Chin, D.-K., Davidson, R. M., and Feltus, F. A. (2011). Characterizing gene co-expression modules in Oryza sativa based on a graph-clustering approach. Plant Biotechnol. 28, 485–493.

Fukushima, A., Kanayama, S., and Aris, M. (2009). Characterizing gene co-expression modules in Oryza sativa based on a graph-clustering approach. Curr. Genet. 44, 197–202.

Fukushima, A., Nakashima, T., Harakata, M., Hikosaka, S., Saito, K., Goris, E., et al. (2012). Exploring tomato gene functions based on co-expression modules using graph clustering and differential coexpression analysis. Plant Physiol. 158, 1497–1502.

Hatakeyama, K., Hongo, K., Swiss, S., Shimizu, A., Nagamura, T., Abe, R., et al. (2013). (E)-Corioliopine is an integrated database of gene expression networks and omics annotations in rice. Plant Cell Physiol. 52, 226–239.

Hammond, J. P., and Bass, R. C. (2012). Advances in plant genome sequencing. Plant J. 70, 177–190.

Eshuis, K. A., Hug, A. G., and Hän, D. K. (2003). A global survey of gene regulation during cold acclimation in Arabidopsis thaliana. Planta 216, 179–190. doi: 10.1007/s00425-002-0820-6.

Hira, M., Yagura, M., Sueda, Y., Yoda, T., Obaya, T., Suzuki, A., et al. (2007). Omics-based identification of Arabidopsis Mth transcription factors regulating aliphatic glucosinolate biosynthesis. Proc. Natl. Acad. Sci. USA. 104, 6474–6483.

Horner, K., Jong, C., Bailey-Serres, J., Mitller, R., Shilton, C., Hart, J. P., et al. (2008). Ameliorating genes of known and unknown function by large-scale coexpression analysis. Plant Physiol. 147, 46–57.

Huang, S., Rob, J. V., Marinescu, E. M., and Lee, I. (2011). Systematic prediction of gene function in Arabidopsis thaliana using a probabilistic functional gene network. PLoS One 6, 14924–14939.

Kawamura, S., Nosokawa, S., Asai, F. F., Ners, D., Kolter, J., Thompson, J. L., et al. (2010). VirtualPlant—a software platform to support systems biology research. Plant Physiol. 152, 508–515.

Liu, J., Jin, J.-G., and McQueen, R., Chang, M.-Y., and Xu, Z., Tian, J., et al. (2010). Combined transcription, protein diversity and metabolite profiling in tomato fruit reveals that the ethylene response factor (ERF) plays an important role in ripening and sugar accumulation. Plant J. 70, 191–204.

Lee, T. H., Kim, T. Y., Pham, T. T. T., Song, S. I., Kim, J.-K., Kang, K. Y., et al. (2009). Bioarraynet: a database for correlating gene expression from transcriptome profiling, and its application to the analysis of co-expressed genes in rice. Plant Physiol. 151, 15–33.

Lee, M., Joung, J.-G., McQuinn, R., Chung, M.-Y., and Xu, Z., Tian, J., et al. (2010). Combined transcription, protein diversity and metabolite profiling in tomato fruit reveals that the ethylene response factor (ERF) plays an important role in ripening and sugar accumulation. Plant J. 70, 191–204.

Lee, T. H., Kim, T. Y., Pham, T. T. T., Song, S. I., Kim, J.-K., Kang, K. Y., et al. (2009). Bioarraynet: a database for correlating gene expression from transcriptome profiling, and its application to the analysis of co-expressed genes in rice. Plant Physiol. 151, 15–33.

Lohse, M., Bolzer, A., Nagel, A., Forni, A., Linn, J., Stitt, M., et al. (2012). ReBioNet: a user-friendly, integrated software solution for RNA-seq based transcriptomics. Nucl. Acid Res. 40, W262–W267.

Manickavelu, A., Kawaura, K., Ohku, K., Shi, L., Tofukura, Y., Saitou, N., et al. (2012). Comprehensive functional analysis of expressed sequence tags in common wheat (Triticum aestivum). DNA Res. 19, 169–175.

Miao, L., Van Hauter, J. J., Dax, S., and Dickerson, J. A. (2009). Arabidopsis gene co-expression network and its functional modules. BMC Bioinformatics 10, 346. doi: 10.1186/1471-2105-10-346.

Miras, A. J., Yota, T. H., Bada, G. J., Zhang, Y., Chatterjee, T., Teha, T., et al. (2011). Tissue- and cell-type specific transcriptional profiling of expanding tomato fruit provides insights into metabolic and regulatory specialization and tissue formation. Plant Cell 23, 3993–4010.

McGrath, K. C., Donahue, R., Mann, J. M., Schreck, P. M., Edgert, C. J., Mclaran, D. J., et al. (2005). Repressor-and activator-type ethylene response factors functioning in cotyledon outgrowth and disease resistance identified via a genome-wide screen of Arabidopsis transcription factor gene expression. Plant Physiol. 139, 494–509.

Matsuda, H., Uehara-Yamaguchi, Y., Yoshikawa, T., Kikuchi, S., and Shinji, C. I., Maclean, D. J., et al. (2005). RiceArrayNet: a functional genomics database for rice. Plant Physiol. 139, 213–219.

Matsuda, H., Uehara-Yamaguchi, Y., Yoshikawa, T., Kikuchi, S., and Shinji, C. I., Maclean, D. J., et al. (2005). RiceArrayNet: a functional genomics database for rice. Plant Physiol. 139, 213–219.

Matias, A. J., Yota, T. H., Bada, G. J., Zhang, Y., Chatterjee, T., Teha, T., et al. (2011). Tissue- and cell-type specific transcriptional profiling of expanding tomato fruit provides insights into metabolic and regulatory specialization and tissue formation. Plant Cell 23, 3993–4010.

Miras, A. J., Yota, T. H., Bada, G. J., Zhang, Y., Chatterjee, T., Teha, T., et al. (2011). Tissue- and cell-type specific transcriptional profiling of expanding tomato fruit provides insights into metabolic and regulatory specialization and tissue formation. Plant Cell 23, 3993–4010.

Nishida, K., Kasahara, K., and Inze, D. (2010). CORNET: a user-friendly tool for data mining and integration. Plant Cell Physiol. 51, 1187–1197.

Obata, S., Vahani, A., Parsa, S., Vilar-Airola, M., et al. (2010). Systems biology of tomato fruit development: combined transcript, protein, and metabolic analysis of tomato transcrip- tion factor (NtERF026) overexpression plants reveals novel regulatory interactions. Plant Physiol. 157, 405–423.

Ohtsuka, S., Naka, H., Kusabira, K., and Kusabira, K. (2011). ATTED-II update: condition-specific gene coexpression to extend coexpression analyses and applications to a broad range of flowering plants. Plant Cell Physiol. 52, 213–229.

Oztar, Y., Suzuki, H., Sakurai, N., and Shibata, D. (2010). Col: a database for characterizing co-expressed gene modules with biological information in plants. Bioinformatics 26, 1267–1268.

Ozono, S., Alba, R., Dassanico, C. M., Lopez-Casado, L., Lohse, M., Zaner, M. L., et al. (2011). System biology of tomato fruit development: combined transcript, protein, and metabolic analysis of tomato transcrip- tion factor (NtERF026) overexpression plants reveals novel regulatory interactions. Plant Physiol. 157, 405–423.

Ooi, S., Alba, R., Dassanico, C. M., Lopez-Casado, L., Lohse, M., Zaner, M. L., et al. (2011). System biology of tomato fruit development: combined transcript, protein, and metabolic analysis of tomato transcrip- tion factor (NtERF026) overexpression plants reveals novel regulatory interactions. Plant Physiol. 157, 405–423.

Ooi, S., Alba, R., Dassanico, C. M., Lopez-Casado, L., Lohse, M., Zaner, M. L., et al. (2011). System biology of tomato fruit development: combined transcript, protein, and metabolic analysis of tomato transcrip- tion factor (NtERF026) overexpression plants reveals novel regulatory interactions. Plant Physiol. 157, 405–423.

Ooi, S., Alba, R., and Dassanico, C. M., Lopez-Casado, L., Lohse, M., Zaner, M. L., et al. (2011). System biology of tomato fruit development: combined transcript, protein, and metabolic analysis of tomato transcrip-
Networks for gene annotation

Patel, R. V., Nakah, H. K., Brotz, R., and Proctor, N. J. (2012). BAR expression identification: expression profile similarity ranking of homologous genes in plant species. Plant J. 71, 1038–1050.

Persson, S., Wu, H., Mihle, J., Page, G., and Somerville, C. (2006a). Large-scale coexpression analysis reveals novel genes involved in cellulose biosynthesis. Plant Cell (Rockville) 2005, 176.

Persson, S., Wu, H. B., Mihle, J., Page, G. P., and Somerville, C. (2005b). Identification of genes required for cellulose synthesis by regression analysis of public microarray data sets. Proc. Natl. Acad. Sci. U.S.A. 102, 8633–8638.

Rischler, H., Oresic, M., Sappanen-Rohrmann, J., Tohge, T., Alba, R., Saito, K., Hirai, M. Y., and Yonekura-Sakakibara, C., Mutwil, M., Saxe, Persson, S., Wei, H., Milne, J., Page, G., Patel, R. V., Nahal, H. K., Breit, R., and in Plant Systems Biology, a specialty of Frontiers in Plant Science. 2011; published online: 08 October 2012; published: 17 August 2012; accepted: 25 July 2012; paper pending acceptance: 20 October 2012; pending publication: 18 November 2012.

Copyright © 2012 Tohge and Fernie. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in other forums, provided that the original authors and source are credited and any copyright notices are included and subject to any copyright notices concerning any third-party graphics etc.

Several gene co-expression and metabolomics -- "fpls-03-00248" -- 2012/11/6 -- 22:03 -- page 6 -- #6