TECHNICAL UPDATE

Visualization of molecular processes associated with seed dormancy and germination using MapMan

Ronny V.L. Joosen[1]*, Wilco Ligterink[1], Bas J. W. Dekkers[1,2] and Henk W.M. Hilhorst[1]

[1]Laboratory of Plant Physiology, Wageningen University, Droevendaalsesteeg 1, 6708 PB Wageningen, The Netherlands; [2]Department of Molecular Plant Physiology, Utrecht University, Padualaan 8, 3584 CH Utrecht, The Netherlands

Abstract

Seed dormancy and germination involve the concerted operation of molecular and biochemical programmes. It has become feasible to study these processes in great detail, using the current methods for transcriptome, proteome and metabolome analysis. Yet, the large amounts of data generated by these methods are often dazzling and demand efficient tools for data visualization. We have used the freely available PageMan/MapMan package (http://MapMan.gabipd.org) to visualize transcriptome and metabolome changes in *Arabidopsis thaliana* seeds during dormancy and germination. Using this package we developed two seed-specific MapMan pathways, which efficiently capture the most important molecular processes in seeds. The results demonstrated the usefulness of the PageMan/MapMan package for seed research.

Keywords: dormancy, MapMan, metabolome, molecular processes, seed germination, transcriptome

Introduction

The Arabidopsis community has developed comprehensive databases for gene description, annotation and expression analysis (Brazma *et al.*, 2003; Toufighi *et al.*, 2005; Zimmermann *et al.*, 2005). The available information is not limited to transcriptome but is expanded to proteome and metabolome data as well (De Vos *et al.*, 2007; Meyer *et al.*, 2007; Baerenfaller *et al.*, 2008). Arabidopsis is also an important model for seed science

and provides valuable insight into the processes underlying germination, dormancy and stress resistance (Finkelstein *et al.*, 2008; Holdsworth *et al.*, 2008a, b).

The seed represents a critical stage in the plant life cycle. After fertilization the embryo is formed, which is surrounded by the endosperm and seed coat (Liu *et al.*, 2005). Seeds acquire desiccation tolerance and dormancy during maturation to survive under harsh conditions after seed dispersal (Bewley, 1997). Germination at an appropriate time is a critical step for the initiation of the plant life cycle (Huang *et al.*, 2010; Moyers and Kane, 2010). Seeds are equipped with accurate sensors for water, light and temperature to monitor optimal seasonal timing for germination, successful seedling establishment and further plant development (Franklin and Quail, 2010). Upon imbibition, protein synthesis and DNA transcription are resumed and cell wall expansion and degradation facilitate the penetration of the radicle through the endosperm and seed coat (Nonogaki *et al.*, 2007). Finally, energy sources are remobilized to enable the fast growth of the emerging seedling (Nonogaki, 2006). The concerted operation of these molecular processes is organized by plant hormones, hormone- or photoreceptors and transcription factors (Holdsworth *et al.*, 2008a). Modern '-omics' tools can provide valuable insight into the function and regulation mechanisms of these molecular processes (Joosen *et al.*, 2009).

Many tools to analyse transcriptome, proteome or metabolome data rely on approaches to detect co-expression or co-existence. Such clustering methods and principal component analysis are efficient tools to summarize data and detect groups of genes, proteins and metabolites with similar behaviour (Rensink and Hazen, 2006). However, more insights into multiple biological processes can be captured by organizing annotations in such a way that profiling datasets are integrated with pre-existing biological knowledge

*Correspondence
Email: Ronny.Joosen@wur.nl

All microarray data were normalized using MAS 5.0 (Hennig, 2003) and raw expression values were filtered to display expression above a background value of 50 in four or more experiments. The initial screening filter yielded 11,443 seed-expressed genes (see supplementary Table S1, available online only at http://journals.cambridge.org). The PageMan tool v0.12 (http://MapMan.gabipd.org; Usadel *et al.*, 2006) was used to identify functional categories with significant enrichment or depletion of up-regulated genes. Within the PageMan package we made use of a Wilcoxon test combined with Benjamin–Hochberg filtering to calculate P values for enriched categories. The obtained P values were transformed to z-scores and plotted as a heat map (Fig. 1). Only significant functional categories are shown in the figure. Because we intended to select the categories with a general role only in dormancy and germination, we excluded mutant transcriptome datasets from the PageMan analysis.

To create a detailed view of the enriched functional categories that we identified with PageMan, we made two custom pathway images (using CorelDRAW graphics suite X4, www.corel.com) which can be used in the MapMan tool v3.5.0 (http://MapMan. gabipd.org). First, we created an 'Arabidopsis seed – Molecular Networks' diagram including all enriched functional categories (see Figs 2–4). The diagram of hormonal regulation was adopted from Finkelstein *et al.* (2002) and was simplified to depict hormone signalling. More detailed information about hormone signalling can be found in Kucera *et al.* (2005) and Holdsworth *et al.* (2008a). Two functional categories describing genes that were linked to dormancy or germination were added to the mapping file (data derived from Taggit ontology; Carrera *et al.*, 2007). Second, we created an 'Arabidopsis seed – Cell wall Networks' diagram that allows a focused view of cell wall changes (synthesis, modification, degradation and proteins). For this second diagram some subdivisions were made within the original 'Cell wall' bin. The 'Cell wall' bin 10.5.1, 'Cell wall Proteins AGP (arabinogalactan proteins)' was further divided to 'AGPs', 'FLA (fasciclin-like arabinogalactan proteins)' and 'AGP Other' and 'Cell wall' bin 10.7, 'Cell wall Modification' was subdivided to 'Expansin A', 'Expansin B' and 'Xyloglucan' (see Fig. 5). All these files are freely available at http://mapman.gabipd. org/web/guest/mapmanstore. Both transcript and metabolite levels can be visualized with this user-friendly package. All individual genes within a functional category are represented as a square box and their expression levels are shown in a colour (blue–red) scale. Metabolites are represented as coloured circles (Fig. 2B). Users can load raw expression levels as well as expression ratios. We calculated expression ratios by dividing Log2 expression values and subtracting −1 for scaling around 0 ('log2−1').

A two-tailed paired t-test was used to calculate P values for all expression ratios (see supplementary Table S2, available online only at http://journals.cambridge.org). AGI codes (The Arabidopsis Genome Initiative, 2000) or metabolite names are used to match the data with a mapping file that contains the functional categorization of genes and metabolites.

Here we show the power of efficient data visualization of changes in transcriptome and/or metabolome using four examples: (1) dry seeds versus imbibed seeds resulting in germination (dry versus 24-h-imbibed); (2) dormant imbibed versus non-dormant germinating seeds (PD24h versus LIG); (3) 24-h-imbibed, stored Ler versus 24-h-imbibed, stored *cts-1* seeds; and (4) embryo versus endosperm tissue.

Results and discussion

To examine the efficiency of MapMan data visualization, expression ratios were calculated for dry versus 24-h-imbibed seeds (Fig. 2A). In the diagram global transcriptome changes are obvious at a first glance. For example, strong up-regulation of genes related to amino acid biosynthesis ('Amino acid'), energy metabolism ('Energy') and cell wall modification ('Cell wall') in 24-h-imbibed seeds were visualized (red squares). In contrast, transcripts related to late embryogenesis abundant (LEA) and seed storage proteins ('Seed storage proteins') rapidly decline (blue squares). Also a decline in stress-related transcripts was observed. Most likely these transcripts accumulated at the end of seed maturation and dehydration and were rapidly lost upon imbibition. In this figure we combined both transcript and metabolite level ratios for dry versus germinating seeds, which allowed us to analyse changes at the metabolite levels in relation to transcriptional changes. For example, several enzymes in the tricarboxylic acid (TCA) cycle in the 'Energy' category were up-regulated in 24-h-imbibed seeds (Fig. 2A), which is consistent with higher levels of TCA intermediates, such as citrate, iso-citrate, 2-oxoglutarate and malate, known to occur in imbibed Arabidopsis seeds (Fait *et al.*, 2006). In Fig. 2B, we depict an example; the concomitant accumulation of malate and a transcript (*FUM2*) encoding a fumarase that catalyses the conversion of fumarate to malate. This particular example should be interpreted with some caution since transcript levels of Arabidopsis Columbia (Col) seeds were compared with metabolite levels of stratified Arabidopsis Wassilewskija (Ws) seeds in this case. However, this type of analysis opened the possibility of combining transcript and metabolite data using MapMan.

Arabidopsis seeds show certain levels of primary dormancy immediately after seed harvest. In our
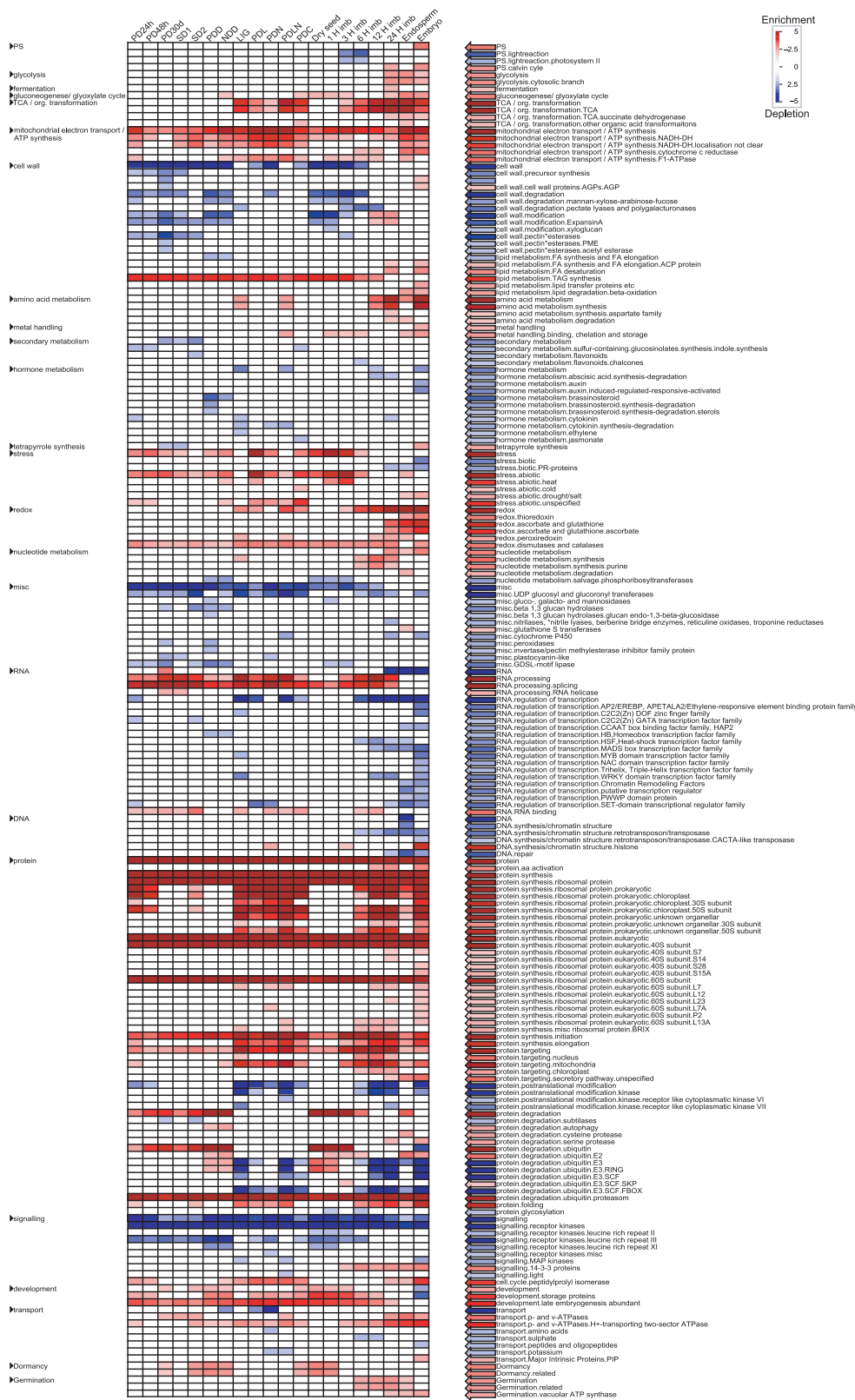
Figure 1. PageMan display of coordinated changes of gene function categories during Arabidopsis seed dormancy and germination. The Affymetrix AtH1 normalized gene expression data were subjected to an analysis to identify overrepresented functional categories using PageMan. The red colour indicates significant enrichment of up-regulated genes, blue indicates significant depletion of up-regulated genes. Only significant gene function categories are shown. See Table 1 for a detailed description of sample abbreviations.
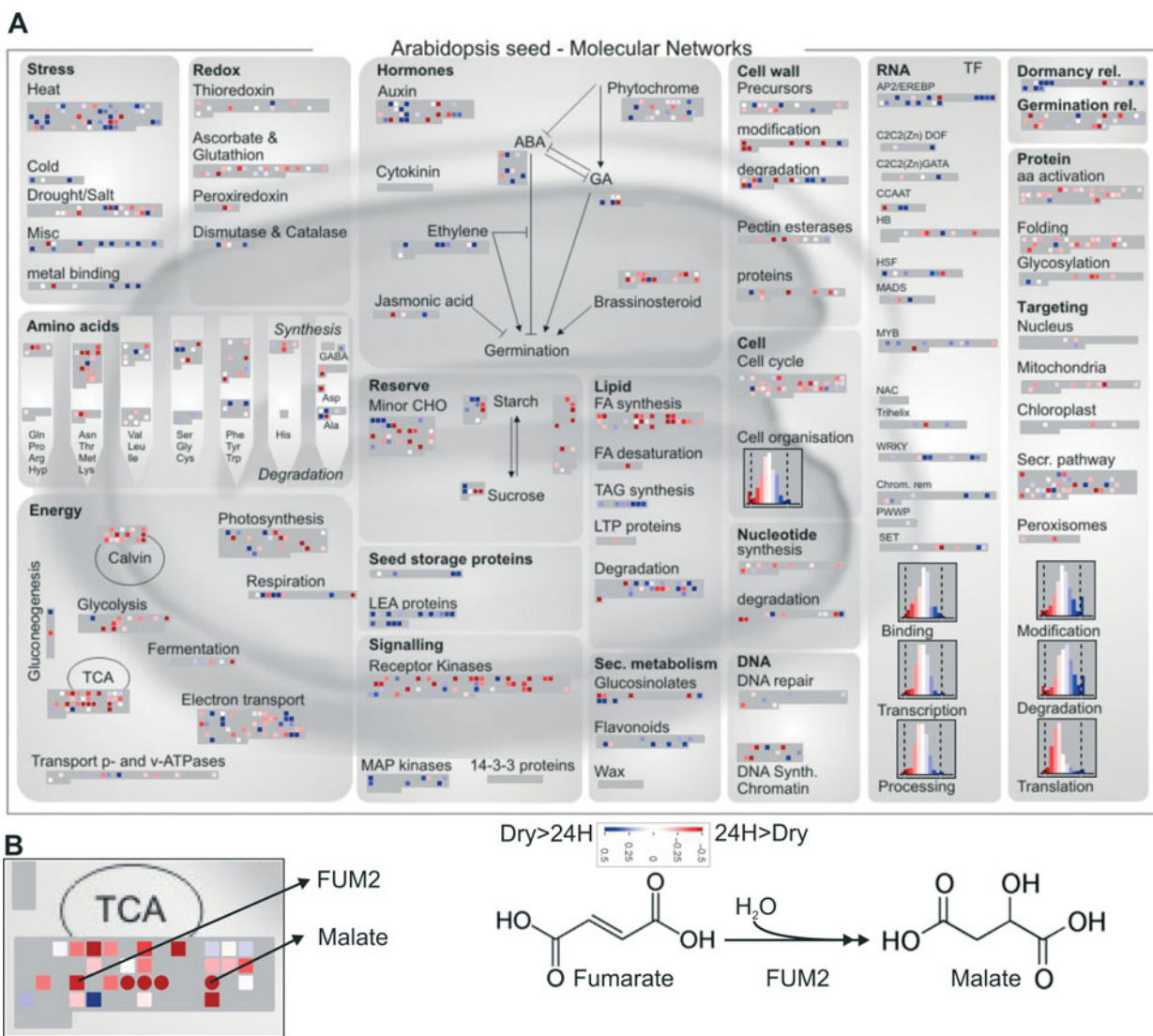
Figure 2. (A) MapMan Arabidopsis seed – Molecular Networks map. The Molecular Networks map shows differences in both transcript (coloured squares) and metabolite (coloured circles) levels. Squares and circles can be clicked to retrieve gene or metabolite data. (B) Example of a close-up view of one category. The TCA cycle in the 'Energy' category is shown to depict the accumulation of malate (red circle) and the up-regulation of a gene (*FUM2*) encoding a fumarase, which catalyses one of the steps of the TCA cycle and converts fumarate to malate, in 24-h-imbibed seeds. Log2−1 ratios are used to express relative levels of transcripts and metabolites in dry versus 24-h-imbibed seeds using a colour scale. Red, higher levels in 24-h-imbibed seeds; blue, higher levels in dry seeds. Only ratios with a *P* value < 0.05 are presented.

second example, we visualized changes in molecular processes that are affected by dormancy (Fig. 3). Therefore, we plotted the expression ratios for primary dormant seeds that were imbibed for 24 h (PD24h) and seeds that were afterripened for 120 d and imbibed for 24 h in the dark with a 4 h pulse of red light (LIG). The PD24h seeds will not complete germination, in contrast to the LIG-treated seeds which do complete germination. When dormant and non-dormant seeds were compared, obvious transcriptional differences were observed in the gene clusters 'Cell wall', 'Stress', 'Secondary metabolism' and 'Hormones'. Surprisingly, relatively small changes were observed

in Taggit gene clusters 'Dormancy-related' and 'Germination-related'.

In Figs 2 and 3, the same sets of data were used for the initial selection of gene function categories in PageMan. Because this could potentially lead to a self-fulfilling, re-detection of differentially regulated genes, we also analysed a transcriptome dataset which was not used for our pathway selection. We compared transcript profiles of 24-h-imbibed, stored wild-type Landsberg erecta (Ler) seeds (Ler AR) and 24-h-imbibed, stored *comatose* (*cts*)-1 mutant seeds (*cts-1* S) using our MapMan diagram (Fig. 4). As expected, *CTS-1* levels were strongly reduced in the mutant.

Figure 3. MapMan Seed – Molecular Networks map showing differences in transcript (coloured squares) levels. Log2−1 ratios are used to express relative levels of transcript in the primary dormant seeds (PD24h) versus afterripened seeds that were imbibed for 24 h (LIG). Red, higher levels in LIG seeds; blue, higher levels in PD24h. Only ratios with a P value < 0.05 are presented.

Consistent with the results described by Carrera *et al.* (2007), effects on the production of anthocyanin pigment 2 protein (PAP2 = *MYB90*), GA-responsive GAST1 protein homologues (*GASA-1*, *GASA-4*) and the flavonoid pathway were clearly visible (Fig. 4, 'RNA' and 'Hormones'). Theodoulou *et al.* (2005) described the jasmonic acid (JA)-deficient phenotype of the *cts-1* mutant. Our results suggest that up-regulation of a seed-specific JA biosynthesis gene (putative 12-oxophytodienoic acid reductase, OPR) could be part of this mechanism (Fig. 4, 'Hormones'). The up-regulation of several photosynthesis pathway genes in the 'Energy' category in the mutant is noteworthy and might be an intriguing starting point for new research.

Several studies describe the important role of the endosperm layer in the regulation of seed germination (e.g. Müller *et al.*, 2006; Penfield *et al.*, 2006). We calculated ratios of gene expression levels between endosperm and embryo to visualize affected genes and molecular processes (see supplementary Fig. F1,

available online only at http://journals.cambridge.org). As expected, genes related to photosynthesis and DNA synthesis are mainly expressed in the embryo. Since cell wall changes are crucial for the completion of germination, we created a diagram specifically to map cell wall changes and used it to highlight differences in cell wall modification between embryo and endosperm (Fig. 5). Arabinogalactan proteins (AGPs) were more linked to embryo cell walls. AGPs are a diverse class of cell wall proteins which are implicated in growth, development and plant–microbe interactions (Seifert and Roberts, 2007), and therefore it is not surprising to find these transcripts more abundantly in embryo tissues.

It is a major challenge to interpret the overwhelming amount of information currently available for genes, proteins and metabolites and understand their function in various biological processes. Combining the information about gene expression levels with known biological function of genes or gene classification can be very helpful in creating
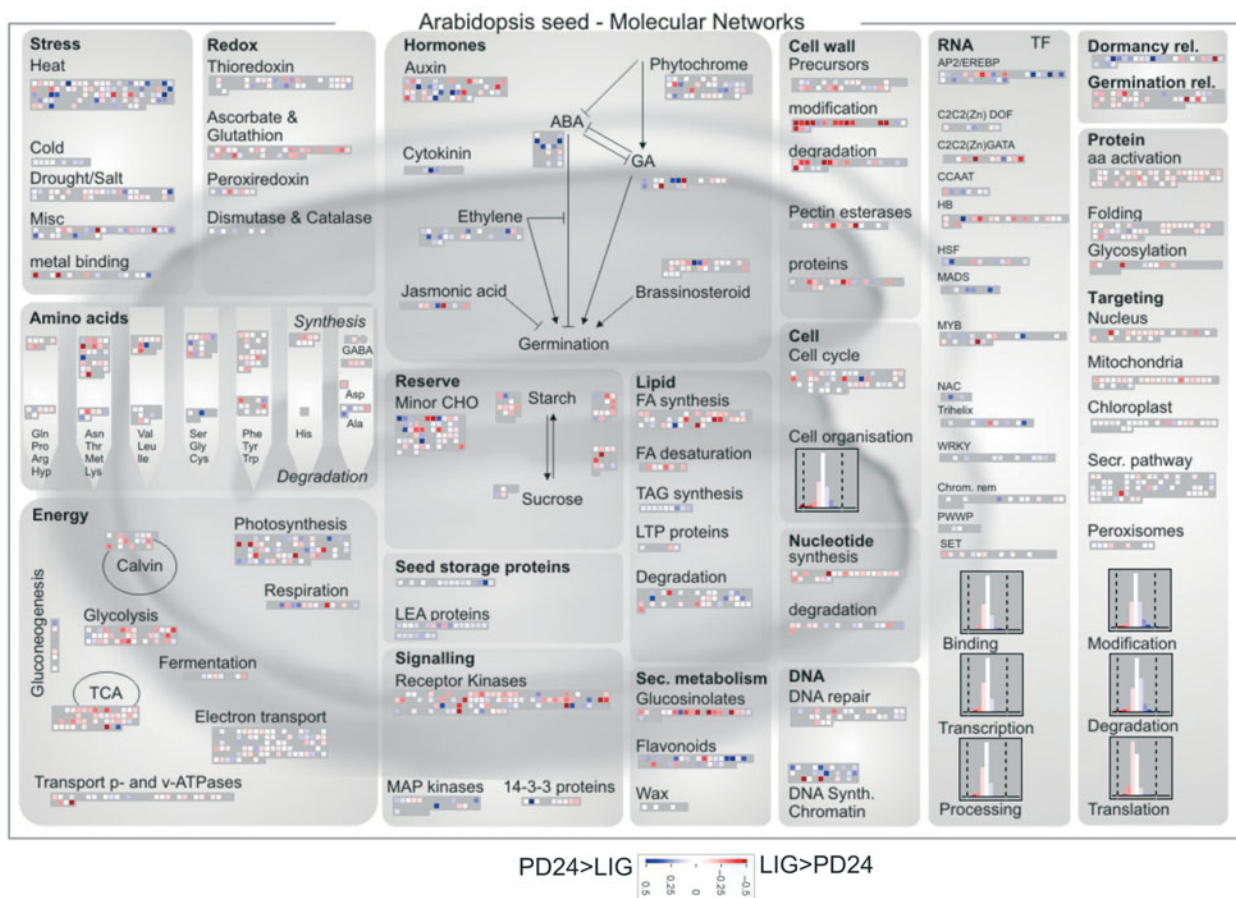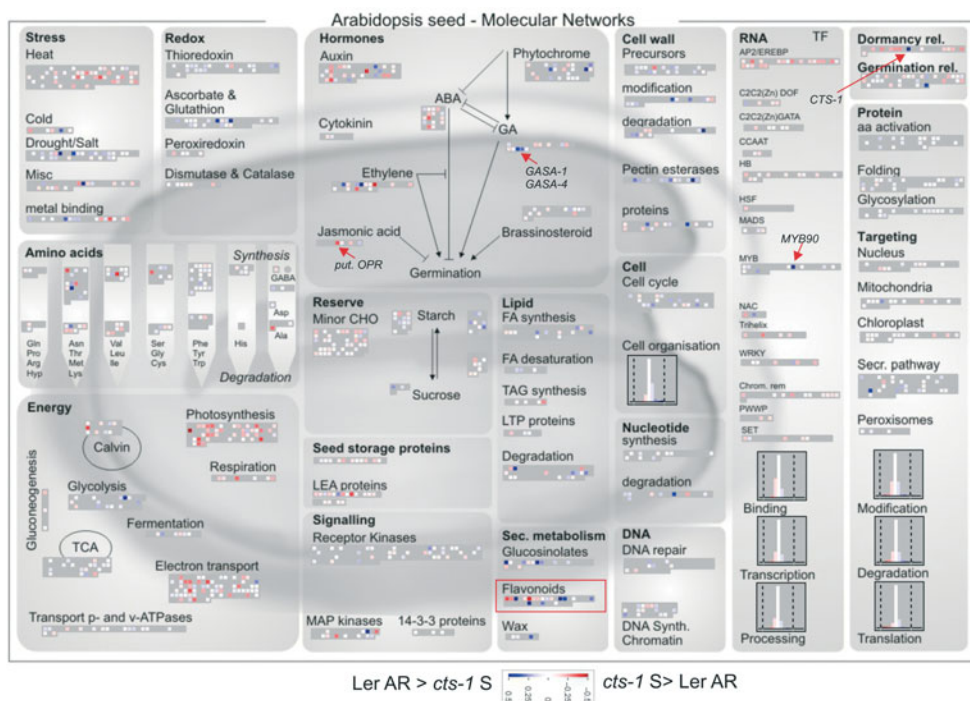
Figure 4. MapMan Seed – Molecular Networks map showing differences in transcript (coloured squares) levels. Log2–1 ratios are used to express relative transcript levels in 4-h-imbibed, stored wild-type Landsberg erecta (Ler) seeds versus 24-h-imbibed stored *cts-1* seeds. Red, higher levels in *cts-1* seeds; blue, higher levels in Ler seeds. Only ratios with a *P* value < 0.05 are presented.

a categorization or applying priority within the data. The freeware tool MapMan has proved to be an easy-to-use and helpful tool to visualize multilevel data (transcriptomics, metabolomics and proteomics).

The data used in this study (summarized in Table 1) and the way of data visualization (Fig. 1) that we examined turned out to be very informative as it clearly summarizes the molecular processes that are affected. For example, the importance of triacylglycerol (TAG) and protein synthesis during germination can readily be inferred from the diagram. This is in agreement with the role of TAG metabolism in germination control and seedling establishment (Penfield *et al.*, 2007) and the essential role described for translation in the completion of germination (Rajjou *et al.*, 2004).

When transcript levels of dry seeds are compared to seeds that have been imbibed for 24 h many important molecular processes can be recognized (Fig. 2A). Genes in the Calvin cycle, glycolysis and TCA cycle seem to be up-regulated, as well as redox modification, cell cycle, cell-wall modification and degradation, and protein activation and folding. Contrarily, transcripts for seed storage proteins, LEA proteins and TAG synthesis are severely decreased in imbibed seeds. It is noteworthy that expression of genes involved in electron transport and respiration seems to be lower in 24-h-imbibed seeds as compared to dry seeds. Our analysis indicated that the major differences between

dormant and non-dormant seeds were in the 'Hormone' and 'Cell wall' clusters (Fig. 3). Interestingly, only minor differences were observed in other processes, such as seed storage proteins and LEA proteins, between dormant and non-dormant seeds. This is due to the decrease in the transcript levels of these proteins to a similar extent in both dormant and non-dormant seeds upon imbibition. It is possible that seed storage proteins and *LEA* transcripts are remnants from the seed developmental stages, which are probably no longer necessary during imbibition. These MapMan diagrams allow users not only to observe global changes but also to retrieve detailed information about gene annotation and expression, because the users can click on an individual process or gene in the interactive MapMan tools.

By plotting the transcriptome differences between the *comatose-1* mutant and wild-type we showed that our selection of molecular processes has the potential to clearly visualize the affected genes and pathways as they were previously described (Fig. 4) (Theodoulou *et al.*, 2005; Carrera *et al.*, 2007). For a more detailed view on a certain process or metabolic pathways one can make a customized diagram (as we showed for cell wall changes) or use already available diagrams in the MapMan package.

We have explored the possibility of using MapMan for multi-level data by combining transcriptome
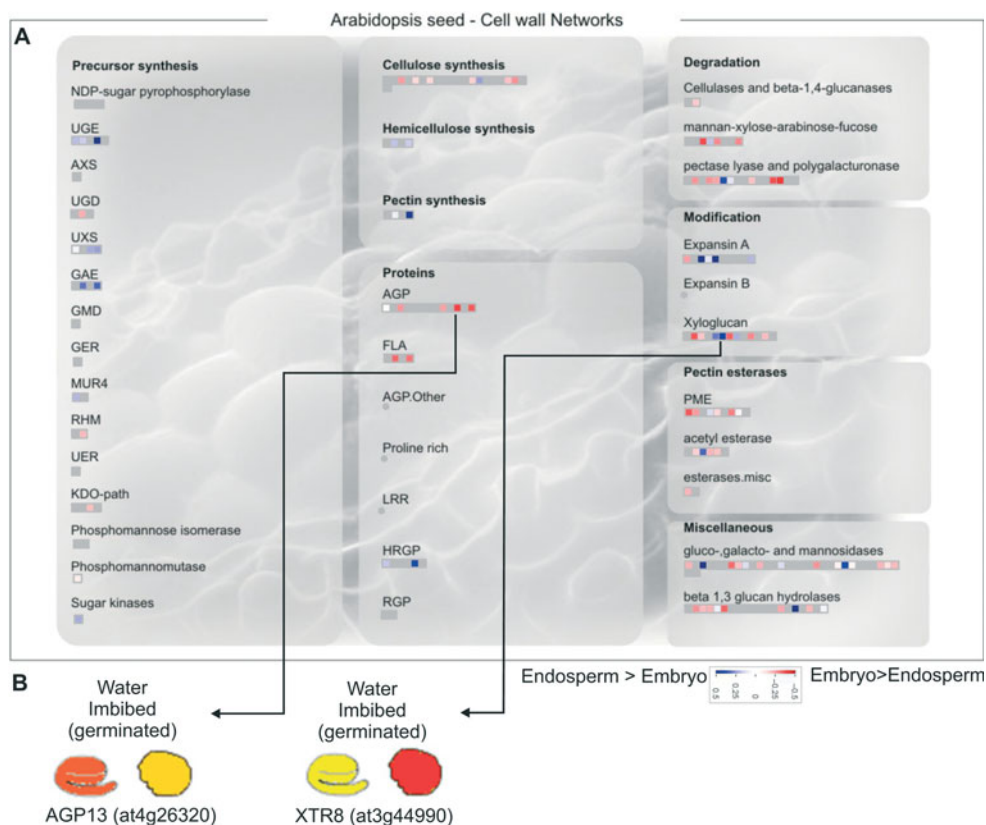
Figure 5. (A) MapMan Cell wall map showing differences in transcript levels between the embryo and endosperm. Log2−1 ratios are used to express relative transcript levels of the endosperm versus the embryo shortly after radicle protrusion. Red, higher levels in the embryo; blue, higher levels in the endosperm (A). Only ratios with a *P* value < 0.05 are presented. (B) Images of the eFP browser (BAR website: http://bar.utoronto.ca; Winter *et al.*, 2007; Bassel *et al.*, 2008) are included as examples of tissue-specific expression. Left and right, images of the enhanced expression of *AGP13* (at4g26320) in the embryo and the expression of *XTR8* (at3g44990) in the endosperm, respectively.

data from Nakabayashi *et al.* (2005) (dry versus 24-h-imbibed seeds) with metabolome data from Fait *et al.* (2006) (dry versus germinating seeds). In this way, relationships between gene expression and metabolome changes can easily be visualized, as we depicted for the TCA cycle genes and metabolites (Fig. 2B).

While our analysis demonstrated the usefulness of the Seed – Molecular Networks diagram, it does not cover all functional categories in every possible seed experiment. Besides, one should bear in mind that non-annotated genes are rarely selected for visualization, which hampers the discovery of new genes with unknown function. Also, it can be misleading when the original functional annotation is incorrect. Despite the aforementioned issues, the annotation used in Map-Man has attained a high-quality level and its usability will only improve, because knowledge about many genes and biological processes is rapidly increasing.

In conclusion, the MapMan tool allows a quick identification of the molecular processes that are regulated during a developmental programme of interest, for which candidate genes with known annotation can easily be identified. This way of data

visualization and the two pathway files that we have created provide a solid base for the next level of statistical data analysis, and are useful tools for the seed science community.

Acknowledgements

Supplemental data

Figure F1. Seed – Molecular Networks map showing differences in transcript levels between embryo and endosperm. Log2−1 ratios of endosperm versus embryo samples of seeds shortly after radicle

protrusion are used to express level differences with the help of a false colour scale. Red indicates higher levels in embryo, blue indicates higher levels in endosperm. Only ratios with a P value < 0.05 are represented.

Table S1. Transcriptome data of 20 microarray experiments used for PageMan analysis.

Table S2. Log2 ratios with P values used for MapMan analysis.

References

Baerenfaller, K., Grossmann, J., Grobei, M.A., Hull, R., Hirsch-Hoffmann, M., Yalovsky, S., Zimmermann, P., Grossniklaus, U., Gruissem, W. and Baginsky, S. (2008) Genome-scale proteomics reveals *Arabidopsis thaliana* gene models and proteome dynamics. *Science* **320**, 938–941.

Bassel, G.W., Fung, P., Chow, T.F., Foong, J.A., Provart, N.J. and Cutler, S.R. (2008) Elucidating the germination transcriptional program using small molecules. *Plant Physiology* **147**, 143–155.

Bewley, J.D. (1997) Seed germination and dormancy. *Plant Cell* **9**, 1055–1066.

Blasing, O.E., Gibon, Y., Gunther, M., Hohne, M., Morcuende, R., Osuna, D., Thimm, O., Usadel, B., Scheible, W.-R. and Stitt, M. (2005) Sugars and circadian regulation make major contributions to the global regulation of diurnal gene expression in Arabidopsis. *Plant Cell* **17**, 3257–3281.

Brazma, A., Parkinson, H., Sarkans, U., Shojatalab, M., Vilo, J., Abeygunawardena, N., Holloway, E., Kapushesky, M., Kemmeren, P., Lara, G.G., Oezcimen, A., Rocca-Serra, P. and Sansone, S.A. (2003) ArrayExpress – A public repository for microarray gene expression data at the EBI. *Nucleic Acids Research* **31**, 68–71.

Cadman, C.S.C., Toorop, P.E., Hilhorst, H.W.M. and Finch-Savage, W.E. (2006) Gene expression profiles of Arabidopsis Cvi seeds during dormancy cycling indicate a common underlying dormancy control mechanism. *The Plant Journal* **46**, 805–822.

Carrera, E., Holman, T., Medhurst, A., Peer, W., Schmuths, H., Footitt, S., Theodoulou, F.L. and Holdsworth, M.J. (2007) Gene expression profiling reveals defined functions of the ATP-binding cassette transporter COMATOSE late in phase II of germination. *Plant Physiology* **143**, 1669–1679.

De Vos, R.C.H., Moco, S., Lommen, A., Keurentjes, J.J.B., Bino, R.J. and Hall, R.D. (2007) Untargeted large-scale plant metabolomics using liquid chromatography coupled to mass spectrometry. *Nature Protocols* **2**, 778–791.

Fait, A., Angelovici, R., Less, H., Ohad, I., Urbanczyk-Wochniak, E., Fernie, A.R. and Galili, G. (2006) Arabidopsis seed development and germination is associated with temporally distinct metabolic switches. *Plant Physiology* **142**, 839–854.

Finch-Savage, W.E., Cadman, C.S.C., Toorop, P.E., Lynn, J.R. and Hilhorst, H.W.M. (2007) Seed dormancy release in Arabidopsis Cvi by dry after-ripening, low temperature, nitrate and light shows common quantitative patterns of gene expression directed by environmentally specific sensing. *The Plant Journal* **51**, 60–78.

Finkelstein, R.R., Gampala, S.S. and Rock, C.D. (2002) Abscisic acid signaling in seeds and seedlings. *Plant Cell* **14 (Suppl.)**, S15–S45.

Finkelstein, R., Reeves, W., Ariizumi, T. and Steber, C. (2008) Molecular aspects of seed dormancy. *Annual Review of Plant Biology* **59**, 387–415.

Franklin, K.A. and Quail, P.H. (2010) Phytochrome functions in Arabidopsis development. *Journal of Experimental Botany* **61**, 11–24.

Hennig, L., Menges, M., Murray, J.A. and Gruissem, W. (2003) Arabidopsis transcript profiling on Affymetrix GeneChip arrays. *Plant Molecular Biology* **53**, 457–465.

Holdsworth, M.J., Bentsink, L. and Soppe, W.J.J. (2008a) Molecular networks regulating Arabidopsis seed maturation, after-ripening, dormancy and germination. *New Phytologist* **179**, 33–54.

Holdsworth, M.J., Finch-Savage, W.E., Grappin, P. and Job, D. (2008b) Post-genomics dissection of seed dormancy and germination. *Trends in Plant Science* **13**, 7–13.

Huang, X., Schmitt, J., Dorn, L., Griffith, C., Effgen, S., Takao, S., Koornneef, M. and Donohue, K. (2010) The earliest stages of adaptation in an experimental plant population: Strong selection on QTLS for seed dormancy. *Molecular Ecology* **19**, 1335–1351.

Joosen, R.V., Ligterink, W., Hilhorst, H.W. and Keurentjes, J.J. (2009) Advances in genetical genomics of plants. *Current Genomics* **10**, 540–549.

Kucera, B., Cohn, M.A. and Leubner-Metzger, G. (2005) Plant hormone interactions during seed dormancy release and germination. *Seed Science Research* **15**, 281–307.

Liu, P.P., Koizuka, N., Homrichhausen, T.M., Hewitt, J.R., Martin, R.C. and Nonogaki, H. (2005) Large-scale screening of Arabidopsis enhancer-trap lines for seed germination-associated genes. *The Plant Journal* **41**, 936–944.

Meyer, R.C., Steinfath, M., Lisec, J., Becher, M., Witucka-Wall, H., Törjék, O., Fiehn, O., Eckardt, Ä., Willmitzer, L., Selbig, J. and Altmann, T. (2007) The metabolic signature related to high plant growth rate in *Arabidopsis thaliana*. *Proceedings of the National Academy of Sciences of the United States of America* **104**, 4759–4764.

Moyers, B.T. and Kane, N.C. (2010) The genetics of adaptation to novel environments: Selection on germination timing in *Arabidopsis thaliana*. *Molecular Ecology* **19**, 1270–1272.

Müller, K., Tintelnot, S. and Leubner-Metzger, G. (2006) Endosperm-limited Brassicaceae seed germination: abscisic acid inhibits embryo-induced endosperm weakening of *Lepidium sativum* (cress) and endosperm rupture of cress and *Arabidopsis thaliana*. *Plant Cell Physioogy* **47**, 864–877.

Nakabayashi, K., Okamoto, M., Koshiba, T., Kamiya, Y. and Nambara, E. (2005) Genome-wide profiling of stored mRNA in *Arabidopsis thaliana* seed germination: epigenetic and genetic regulation of transcription in seed. *The Plant Journal* **41**, 697–709.

Nonogaki, H. (2006) Seed germination – the biochemical and molecular mechanisms. *Breeding Science* **56**, 93–105.

Nonogaki, H., Chen, F. and Bradford, K.J. (2007) Mechanisms and genes involved in germination *sensu stricto*. pp. 264–304 in Bradford K.J.; Nonogaki, H. (Eds) *Seed development, dormancy and germination. Annual Plant Reviews*, Vol. 27. Oxford, Wiley-Blackwell.

Penfield, S., **Li, Y.**, **Gilday, A.D.**, **Graham, S. and Graham, I.A.** (2006) Arabidopsis ABA INSENSITIVE4 regulates lipid mobilization in the embryo and reveals repression of seed germination by the endosperm. *Plant Cell* **18**, 1887–1899.

Penfield, S., **Pinfield-Wells, H. and Graham, I.A.** (2007) Lipid metabolism in seed dormancy. pp. 133–152 *in* Bradford K.J.; Nonogaki, H. (Eds) *Seed development, dormancy and germination. Annual Plant Reviews*, Vol. 27. Oxford, Wiley-Blackwell.

Rajjou, L., **Gallardo, K.**, **Debeaujon, I.**, **Vandekerckhove, J.**, **Job, C. and Job, D.** (2004) The effect of alpha-amanitin on the Arabidopsis seed proteome highlights the distinct roles of stored and neosynthesized mRNAs during germination. *Plant Physiology* **134**, 1598–1613.

Rensink, W.A. and Hazen, S.P. (2006) Statistical issues in microarray data analysis. *Methods Molecular Biology* **323**, 359–366.

Seifert, G.J. and Roberts, K. (2007) The biology of arabinogalactan proteins. *Annual Review Plant Biology* **58**, 137–161.

Sreenivasulu, N., **Usadel, B.**, **Winter, A.**, **Radchuk, V.**, **Scholz, U.**, **Stein, N.**, **Weschke, W.**, **Strickert, M.**, **Close, T.J.**, **Stitt, M.**, **Graner, A. and Wobus, U.** (2008) Barley grain maturation and germination: metabolic pathway and regulatory network commonalities and differences highlighted by new MapMan/PageMan profiling tools. *Plant Physiology* **146**, 1738–1758.

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

Theodoulou, F.L., **Job, K.**, **Slocombe, S.P.**, **Footitt, S.**, **Holdsworth, M.**, **Baker, A.**, **Larson, T.R. and Graham, I.A.** (2005) Jasmonic acid levels are reduced in COMATOSE ATP-binding cassette transporter mutants. Implications for transport of jasmonate precursors into peroxisomes. *Plant Physiology* **137**, 835–840.

Thimm, O., **Bläsing, O.**, **Gibon, Y.**, **Nagel, A.**, **Meyer, S.**, **Krüger, P.**, **Selbig, J.**, **Müller, L.A.**, **Rhee, S.Y. and Stitt, M.** (2004) Mapman: a user-driven tool to display genomics data sets onto diagrams of metabolic pathways and other biological processes. *The Plant Journal* **37**, 914–939.

Toufighi, K., **Brady, S.M.**, **Austin, R.**, **Ly, E. and Provart, N.J.** (2005) The botany array resource: e-Northerns, expression angling, and promoter analyses. *The Plant Journal* **43**, 153–163.

Usadel, B., **Nagel, A.**, **Steinhauser, D.**, **Gibon, Y.**, **Blasing, O.E.**, **Redestig, H.**, **Sreenivasulu, N.**, **Krall, L.**, **Hannah, M.A.**, **Poree, F.**, **Fernie, A.R. and Stitt, M.** (2006) PageMan: an interactive ontology tool to generate, display, and annotate overview graphs for profiling experiments. *BMC Bioinformatics* **7**, 535.

Winter, D., **Vinegar, B.**, **Nahal, H.**, **Ammar, R.**, **Wilson, G.V. and Provart, N.J.** (2007) An 'Electronic Fluorescent Pictograph' browser for exploring and analyzing large-scale biological data sets. *PLoS ONE* **2**, e718.

Zhou, X. and Su, Z. (2007) EasyGO: Gene Ontology-based annotation and functional enrichment analysis tool for agronomical species. *BMC Genomics* **8**, 246.

Zimmermann, P., **Hennig, L. and Gruissem, W.** (2005) Gene-expression analysis and network discovery using Genevestigator. *Trends in Plant Science* **10**, 407–409.