Eve Syrkin Wurtele1, Ling Li1, Dan Berleant2, Dianne Cook2, Julie A. Dickerson3, Jing Ding2, Heike Hofmann2, Michael Lawrence2, Eun-kyung Lee2, Jie Li1, Wieslawa Mentzen1, Leslie Miller4, Basil J. Nikolau4, Nick Ransom1, Yingjun Wang1,

Departments of: 1. Genetics, Development and Cell Biology , 2 Statistics, 3 Electrical and Computer Engineering 4 Computer Science 5 Biochemistry Biophysics and Molecular Biology.

METNET: SYSTEMS BIOLOGY TOOLS FOR ARABIDOPSIS

Abstract. MetNet (http://metnet.vrac.iastate.edu/) is an emerging open-source software platform for exploration of disparate experimental data types and regulatory and metabolic networks in the context of Arabidopsis systems biology. The MetNet platform features graph visualization, interactive displays, graph theoretic computations for determining biological distances, a unique multivariate display and statistical analysis tool, graph modeling using the open source statistical analysis language, R, and versatile text mining. The use of these tools is illustrated with data from the *bio1* mutant of Arabidopsis.

1. INTRODUCTION

Plant composition, form, and function are the ultimate consequence of gene expression. High-throughput detection and measurement of changes in the accumulation of tens of thousands of cellular components - RNAs, proteins, and metabolites, and metabolic flux information, lead to complex, valuable datasets (Oliver et al., 2002; Sriram et al., 2004; Fermi et al., 2005; Nikirofrova et al., 2005). Each dataset has the potential to contribute to our understanding of cellular function, and combined experimental datasets impart an added potential to understand and predict the behaviour of a cell. Comparative analysis of mRNA and proteins can provide insights into the processes that affect mRNA accumulation (gene transcription and/or mRNA stability) and protein accumulation (mRNA translation and/or protein stability), but do not give direct information on metabolism. Metabolite profiling gives information about the accumulation of metabolites, but does not reveal which pathways produced those metabolites; however, in combination with microarray and proteomics pathways may be surmised. Techniques for metabolomic flux analysis in plants are becoming more sophisticated (Sriram et al., 2004; Ratcliffe RG, Shachar-Hill, 2005), and these data can contribute information on the flow through specific metabolic pathways and when combined with ‘omics data can provide clues about regulatory mechanisms. Other datasets for

T.H.E. Editor(s) (ed.), Book title, 1—6.

plants that could provide additional information for analysis of cellular systems, such as protein-protein or protein-DNA interactions, are on the horizon.

Due to the complexity of each dataset, a human mind cannot comprehend data of a single type, let alone the datasets *en toto*. Also, the datasets are flawed. Even for the model plant species Arabidopsis, the majority of genes are not yet well annotated, and current technologies to identify metabolites and proteins yield incomplete datasets. Furthermore, most interactions between the biomolecules, as well as most of the kinetics of the established interactions, are not yet known. Even given the availability of comprehensive 'omics datasets, and a full understanding of the interactions and kinetics of a cell, there are not yet modeling methods capable of predicting the behaviour of such a complex system (Du et al, 2004; Ma'ayan et al., 2005; Lee et al., 2005; Xiong et al., 2005)

Thus, the challenge in prediction of a biological network is complex, and requires consideration of a variety of factors: 1) How to represent a biological network; 2) How to evaluate datasets that have only part of their constituents determined and a subset of the possible interactions elucidated; 3) How to model processes that have wide ranging kinetics parameters, most of which are not yet determined.

MetNet is being designed to provide an integrated, open-source platform to develop hypotheses about which genes and proteins might be involved in a process, which pathways and interactions might be important under particular conditions, and ultimately how the biological system functions. We discuss MetNet, and illustrate its use with data from an experiment designed to analyse the biotin metabolic network. Biotin is required as a cofactor by all living organisms. It is synthesised almost exclusively by photosynthetic organisms, is an essential cofactor for several key enzymes in plants (Nikolau et al., 2003). It is also a potential metabolic regulator (Che et al., 2002, 2003). Understanding the multiple functions of this metabolite presents a formidable challenge in systems biology.

2. RESULTS

2.1. MetNetDB contains an integrated metabolic and regulatory map of Arabidopsis interactions.

The MetNetDB database contains a repository of curated expert-created regulatory and metabolic pathways, as well as processed information from repositories of metabolic-only pathways for Arabidopsis: AraCyc [Mueller et al., 2003], and in the near future, BioPathAt [Lange and Ghassemian, 2005], and MapMan [Thimm et al., 2004]. Expansion of the MetNetDB database is ongoing. Biomolecules that can be represented in MetNetDB, include metabolites, genes, RNAs, polypeptides, and protein complexes; interactions that can be represented include catalysis, conversion, transport, and a wide variety of regulatory interactions (eg., allosteric inhibition, transcriptional inhibition, covalent modification). Because the concentration of each biomolecule, as well as the interactions it is able to participate

in, vary across subcellular compartments, MetNetDB includes subcellular location information. Thus, multiple entries are permitted for each biomolecule (eg., a metabolite can participate in more than one reaction, and can be located in more than one subcellular compartment). The MetNetDB curator interface is designed for curation of biomolecules, interactions, and associated information about subcellular location, synonyms, and references. The interface includes a simple graphic representation of the pathways in which biological interactions and complexes can be viewed, created, or modified.

The network is stored in a MYSQL (www.mysql.org) relational database. We have constructed an XML file format that accurately encodes the network topology information from MetNetDB. The network itself is designed for analysis with experimental data, using tools such as MetNet in Cytoscape and ExploRase, which currently receive network information in XML format. A versatile XML file-builder (http://metnetdb.gdcb.iastate.edu/MetNet/MapBuilder.html) can be used to export current data from the MetNetDB database network.

22. Statistical and Visualization Software Tools

ExploRase provides a multivariate approach to detect patterns in gene expression, and to explore connections between 'omics data sets and the known and hypothesized regulatory and metabolic network of Arabidopsis. ExploRase is built on the open-source statistical analysis software R (*http://*www.R-project.org), and the open-source data visualization software, GGobi (*http://www.ggobi.org*), and includes a user-friendly interface for both. ExploRase also adds a spreadsheet with TAIR annotations about each gene, links to literature, menus of analysis and visualization options, and an interface to lists of genes in MetNetDB pathways. Common statistical analyses are provided through GUIs. Alternatively, code for new functionality can be written using R commands. Thus, the GUIs in ExploRase make the R functionality transparent for the novice, but allow a more advanced user to do more sophisticated analysis.

ExploRase has a highly interactive graphics system, designed specifically for exploratory mining of high-dimensional data. It has multivariate graphics including parallel coordinate plots and tours (rotations of high-dimensional scatterclouds). Users can label elements of the plots by clicking on genes, proteins, and/or metabolites of interest. Metabolic and regulatory networks can be displayed using the add-on package GGVis. Users can layout a network (in 2, 3 or higher dimensions), or read in a layout from another package such as MetNet in Cytoscape.

To elucidate the biotin network of plants from a systems biology viewpoint, we have been analysing mutants blocked, overexpressed, or underexpressed in steps of this network. One such step is encoded by the *bio1* gene, which encodes 7,8-diaminopelargonic acid aminotransferase, the third step in the synthesis of biotin from pimelic acid (Patton et al., 1996)**.** A homozygous mutation in *bio1* is lethal without addition of exogenous biotin, however the seedlings appear normal for several days, due to a residue of biotin originally supplied to the parent plants (Weaver et al., 1996; Patton et al., 1996; Che et al., 2000).

One aspect of exploRase is illustrated with an example of microarray data from a portion of a larger experiment (Figure 1). In this experiment, seeds of homozygous mutants for the *bio1* gene are grown in medium with and without biotin. The upper part of Figure 1A shows a dialog window with information on each mRNA. This includes Affy8k ID, Locus ID, TAIR annotations and otherdescriptions. On the left, is a list of available chips (biotin.bio101, biotin.bio102, biotin.bio1B1,... WT1, WT2). Prior to this visualization, the chips were normalized using a quantiles' normalization (Bolstad et al., 2003), and a robust median is used for the expression value. Below the dialogs are two plots: a scatterplot and a parallel coordinate plot. The scatter plot shows a comparison of transcript accumulation between two replicates of WT seedling grown without biotin. The two RNAs marked by the user in yellow, both chloroplast encoded transcripts, are seen to accumulate at a much higher level in the second replication than in the first. The yellow highlight marking is automatically shown on the parallel coordinate plot on the right, and on the annotation list. Both genes exhibit a similar pattern: they are expressed at a high but fairly stable level for all of the chips except for biotin.WT.02, indicating, that a data error might have occurred on the chip biotin.WT.02.

To identify patterns of co-accumulation of RNAs associated with biotin, the user selected a parallel coordinate plot, gene annotation list, and a scatter plot, and displayed data from the *bio1* genotype grown with and without biotin (Fig 1B). There are numerous RNAs, visible in the scatter plot that accumulate to similar levels. Clicking on outliers in the scatter plot (those accumulating at higher levels in the *bio1* mutant with added biotin were colored in blue; those with decreased accumulation when biotin is absent are colored orange) links to the same genes in the parallel coordinate plot and the highlighted genes are also displayed in the gene annotation list. For comparison, two genes that are not outliers in the scatter plot were highlighted in yellow. Using this approach identified At2g02500 (encoding 4-diphosphocytidyl-2C-methyl-D-erythritol synthase (ISPD)) and At4g15560 (encoding putative 1-deoxy-D-xylulose 5-phosphate synthase (DXPS), both genes of isoprenoid synthesis. At2g02500 and At4g15560 were up-regulated 7- and 2-fold, respectively in the *bio1* mutant plus biotin as compared to the *bio1* mutant without biotin.

2.3. Metabolic Network Display and Modeling (MetNet in Cytoscape)

MetNet in Cytoscape (Figure 2A) uses the Cytoscape (http://www.cytoscape.org/) Java program, together with plug-ins specialized for MetNet and its database, to dynamically display complex biological networks and analyse their structure (Wurtele et al., 2003; Du et al., 2004). Data from experiments (i.e., microarray, proteomics, or metabolomics) can be directly overlayed on the network. An interface to R allows the user to analyse 'omics data in R, cluster biomolecules that behave similarly, search for biomolecules with significant changes, and to custom-write R scripts and apply them to experimental data.

MetNet in Cytoscape uses graph theoretic methods to display and analyse biological networks, such as those in MetNetDB. Graphs can be visualized by employing the P-neighborhood function around nodes or reactions of interest; in this mode, the user selects any group of biomolecules or pathways in the MetNetDB network, and extends the network in all directions by a user-designated number of steps. Graphs also can be dynamically displayed as pathways and cycles. (A simple cycle could include a gene transcribed to a protein which, when that protein was over-accumulated, would inhibit the gene's transcription.) For example, a user could display the network that includes all genes that are differentially expressed between *bio1* seedlings grown with and without biotin, and find pathways in that network. Different pathways might indicate multiple mechanisms for control of a process. Common steps among pathways may reflect critical paths in the network.

By displaying all pathways containing genes identified as differentially expressed using ExploRase, the user obtained a very complex network (the insert at upper left of in Figure 2A shows a portion of this network); the network was pared down so that only steps connecting biotin with the At2g02500 and At4g15560 proteins remained (Figure 2A). Both these encode enzymes that are early in the plastidic methylerythritol 4-phosphate (MEP) isoprenoid pathway. Pyruvate is a common substrate for both plastidic fatty acid synthesis and the MEP pathway. Biotin is required for acetyl-CoA carboxylase activity; therefore the carbon needed for formation of plastidic malonyl-CoA (indirectly from plastidic pyruvate via acetyl-CoA) could be limited in the *bio1* mutant when grown without added biotin. A decrease in the flux through the fatty acid biosynthetic pathway due to decreased acetyl-CoA carboxylase activity might influence the flux of pyruvate towards MEP, and provide a signal that alters gene expression in the MEP pathway. This potential interconnection between the fatty acid and MEP pathway could be explored further by experimentation and by modelling (eg, Du et al., 2004).

2.4. MetachipGraph

MetachipGraph is a JAVA program designed to analyse co-expressed genes across large datasets. Unlike other analysis programs, which store all the data in memory all the time, MetachipGraph uses the RandomAccessFile class to read and store data only as it's needed. This allows the program to work with extremely large sets of data while requiring relatively little memory. The program comes with a set of Arabidopsis data (both experimental data and metadata) from NASCArrays (http://arabidopsis.info/) that we have selected as being high quality and normalized. It is also simple to analyse a microarray dataset from any species (or indeed any other type of dataset) using MetachipGraph. Graphs can be sorted according to the expression value and metadata information.

MetachipGraph was used to determine the Pearson correlations of the differentially expressed genes At2g02500 and At4g15560, across 1000 chips from the NASCArrays database. At2g02500 and At4g15560 have a 63% correlation with each other across all the chips (not shown). This corresponds to a p-value below 1.4e-45. Among the 22,746 genes on the Affymetrix ATH1chip, the most similar expression profiles to that of At2g02500 are those of At5g45930 and At1g32990

(87% and 86% correlation respectively) (Fig 2B). At5g45930 encodes a magnesium-chelatase subunit, ChlI, which is required for chlorophyll biosynthesis. At1g32990 encodes plastidic ribosomal protein L11. These results suggest a possible relationship between the plastidic biosynthetic processes of isoprenoid synthesis and photosynthesis.

2.4. Textmining

PathBinder (http://www.public.iastate.edu/~mash/MetNet/MetNet_PathBinder.htm) is a text-mining tool designed specifically to explore metabolic and regulatory interactions in plants. The tool queries the MEDLINE database and retrieves sentences that contain two terms of interest; each sentence is a clickable link pointing to the original on-line PubMed citation. PathBinder contains an extensive set of synonyms from MetNetDB, many tailored to plant biology, which are co-searched when a term is selected. An API (applications programming interface) is provided so that the PathBinder text-mining tool can be integrated into other analysis tools. The API has been used to incorporate PathBinder into the MetNetDB database, and in the future could automatically extract references for interactions, which could then be manually curated. We have created a novel “hidden links” tool to identify and explore potential intermediate links in networks. Given biomolecules A and C that do not co-occur in any sentence, the tool will find biomolecules B that co-occur in sentences both with A and with C. In Figure 2C, the user explored possible literature connections between isoprenoids and biotin. These two terms did not co-occur in any sentence. The user chose isopentenyl diphosphate as biomolecule A, and biotin [an automatically-selected synonym was vitamin H] as biomolecule C, and selected the Hidden Links algorithm. Two hundred and eleven biomolecule B terms co-occurred independently in sentences with both isopentenyl diphosphate and with biotin. The user clicked on LEUCINE; a single sentence containing isopentenyl diphosphate and LEUCINE, and 34 sentences containing biotin and LEUCINE were retrieved. Here, a second possible connection between biotin and isoprenoids is suggested, as leucine catabolism requires the biotin-enzyme methylcrotonyl-CoA carboxylase.

2.6 Major venues for improvement and expansion of the MetNet platform.

Expansion of the MetNet platform is in progress. The software tools will be further integrated such that the users will be able to access all tools from a single platform. A node- and edge-labelled graph model for the database will be implemented. This model will address a major database challenge: tracking changes in biological network data, as such data are being continuously revised and expanded. By broadening the current MetNetDB relational database to a node- and edge-labelled graph model format, information about the date and person (or web source) for each data entry must be captured, to track data revisions and new biological knowledge, as well as to provide automated methods for addition of large-scale data-dumps from on-line resources such as AraCyc. This model would enable addition of several features not present in other databases. In particular it would provide a flexible method for tracking changes. Such a model would also enable researchers to create

their own version of networks to test, model and compare with other networks. In addition, the database would be able to model, as well as store, the data.
MetNet can be modified for analysis of species other than Arabidopsis; in particular, we are beginning to expand the MetNet platform to soybean.

3. CONCLUSION

The MetNet platform is designed for exploration of diverse data sets, and formulation of hypotheses based on this data in the context of known Arabidopsis regulatory and metabolic interactions.

4. REFERENCES

Che P, Weaver LM, Wurtele ES, Nikolau BJ. 2003. The role of biotin in regulating 3-methylcrotonyl-CoA carboxylase expression in Arabidopsis. Plant Physiol 131: 1479-1486;

Che P, Wurtele ES, Nikolau BJ. 2002 Metabolic and Environmental Regulation of 3-Methylcrotonyl-CoA Carboxylase Expression in Arabidopsis. Plant Physiol 129:625-637

Cook D, Hofmann H, Lee E-K, Yang H, Nikolau B, Wurtele ES. 2004 "Visual Methods for Data from Two Factor Single Replicate Gene Expression Studies", Technical Report 04-06:http://www.stat.iastate.edu/preprint/articles/2004-04.pdf, Department of Statistics, Iowa State University.

Dickerson' JA, Berleant, D, Cox Z, Qi, W, Ashlock D, Fulmer AW, Wurtele ES. 2002 "Creating and Modeling Metabolic and Regulatory Networks Using Text Mining and Expert Knowledge," in Computational Biology and Genome Informatics, edited by Cathy H. Wu, Paul Wang, Jason T. L. Wang, World Scientific

Ding J, Berleant D, Nettleton D, E. Wurtele ES. 2002. Mining Medline: Abstracts, sentences, or phrases? Pacific Symposium on Biocomputing, 1-12.

Du P, Gong J, Wurtele ES, Dickerson JA, 2005 Modeling Gene Expression Networks using Fuzzy Logic. *IEEE* Trans. On Systems, Man and Cybernetics, Part B. In press.

Fernie AR, P Geigenberger, M Stitt. (2005) Flux an important, but neglected component of functional genomics. Current Opinion in Plant Biology 8(174-182).

Lange BM, M Ghassemian 2005 Comprehensive post-genomic data analysis approaches integrating biochemical pathway map Phytochemistry, 66:413-451.

Lee E-K, Cook D, Hofmann H, Eve Wurtele, Dongshin Kim, Jihong Kim, and Hogeun An. 2004 GENEGOBI : VISUAL DATA ANALYSIS TOOLS FOR MICROARRAY DATA COMPSTAT'2004 Symposium □c Physica-Verlag/Springer

Lee SG, Kim CM, Hwang KS. 2005. Development of a software tool for in silico simulation of Escherichia coli using a visual programming environment. J Biotechnol (Epub)

Ma'ayan A, Blitzer RD, Iyengar R. 2005. Toward predictive models of mammalian cells.Annu Rev Biophys Biomol Struct. 34:319-349

Mueller LA, Peifen Zhang and Seung Y. Rhee 2003. AraCyc: A Biochemical Pathway Database for Arabidopsis. Plant Physiology, June 2003, Vol. 132, pp. 453-460

Nikiforova VJ, Daub CO, Hesse H, Willmitzer L, Hoefgen R. 2005. Integrative gene-metabolite network with implemented causality deciphers informational fluxes of sulphur stress response. J Exp Bot. 56:1887-1896.

Nikolau BJ, Ohlrogge JB, Wurtele ES 2003 Plant Biotin-Containing Carboxylases. Arch Biochem Biophys. 414:211-222

Oliver DJ, Nikolau B, Wurtele ES. 2002 Functional genomics: high-throughput mRNA, protein, and metabolite analyses. Metab Eng. 4:98-106

Patton DA, Sandy Volrath, Eric R. Ward 1996. Complementation of an Arabidopsis thaliana biotin auxotroph with an Escherichia coli biotin biosynthetic gene. Molecular Genetics and Genomics 251:261 - 266

Ratcliffe RG, Shachar-Hill Y. 2005 Revealing metabolic phenotypes in plants: inputs from NMR analysis. Biol Rev Camb Philos Soc. 80:27-43.

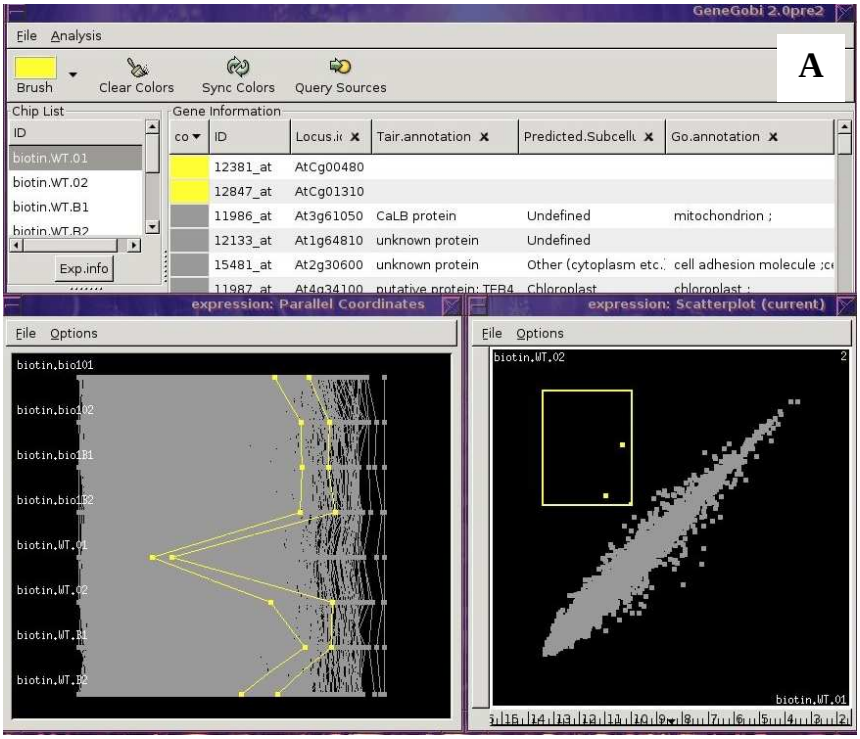
Sriram G, DB Fulton, Vidya V. Iyer, Joan Marie Peterson, Ruilian Zhou, Mark E. Westgate, Martin H. Spalding and Jacqueline V. Shanks . 2004. Quantification of Compartmented Metabolic Fluxes in Developing Soybean Embryos by Employing Biosynthetically Directed Fractional 13C Labeling, Two-Dimensional [13C, 1H] Nuclear Magnetic Resonance, and Comprehensive Isotopomer Balancing. Plant Physiol. 136: 3043-3057.

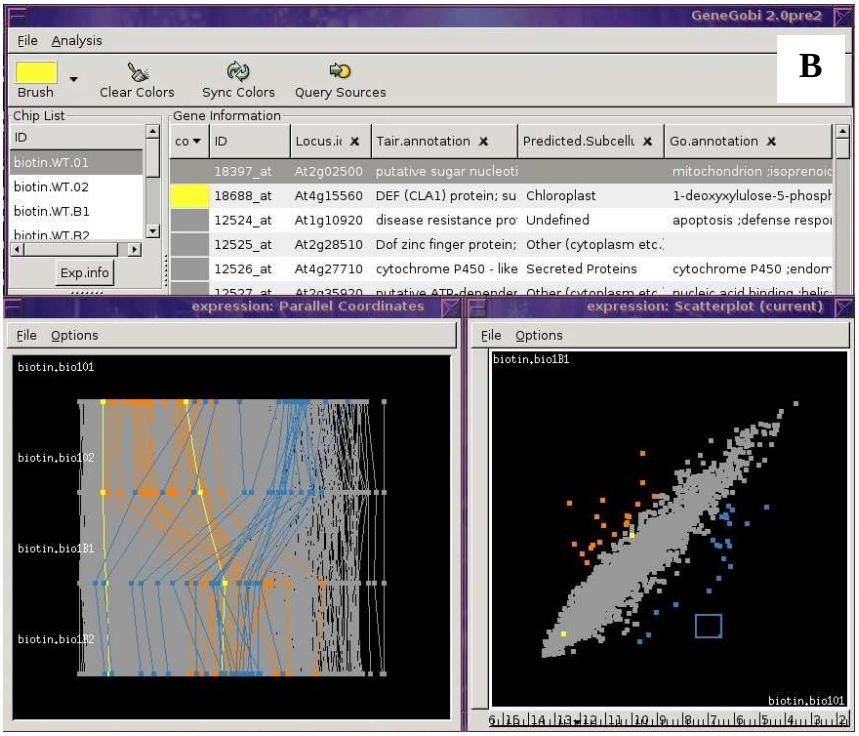
Thimm O. Bläsing O, Gibon Y, Nagel A, Meyer S, Krüger P, Selbig J, Müller L.A, Rhee SY, Stitt M. 2004. Mapman: a user-driven tool to display genomics data sets onto diagrams of metabolic pathways and other biological processes. The Plant Journal: 37:914-939.

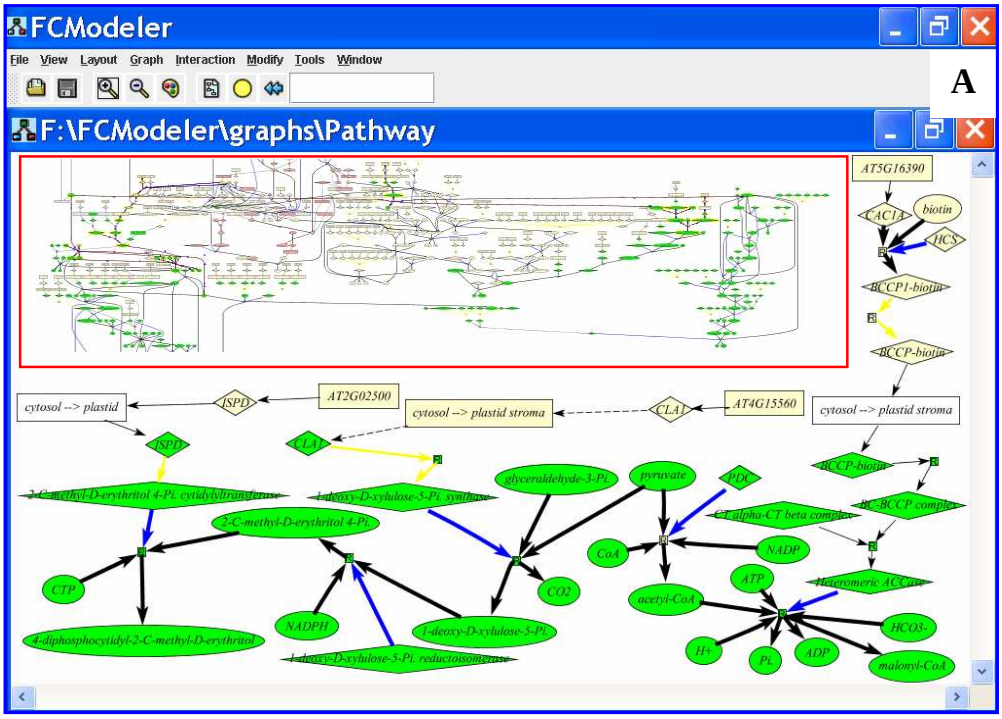
Weaver LM, Yu F, Wurtele ES, Nikolau BJ. 1996. Characterization of the cDNA and gene coding for the biotin synthase of Arabidopsis thaliana. Plant Physiol 110:1021-1028

Wurtele ES, Dickerson JD, Cook D, Hofmann H, Li J, Diao L. 2003. MetNet: software to build and model the biogenetic lattice of Arabidopsis. Comp Funct Genomics. 4:239-245.

Xiong M, Zhao J, Xiong H. 2004 Network-based regulatory pathways analysis. Bioinformatics 20:2056-2066.




GeneGobi 2.0pre2
File Analysis
B
Brush
Clear Colors
Sync Colors
Query Sources
Chip List
ID
biotin.WT.01
biotin.WT.02
biotin.WT.B1
biotin.WT.B2
Exp.info
Gene Information
co
ID
Locus.i
Tair.annotation
Predicted.Subcell
Go.annotation
18397_at At2g02500 putative sugar nucleoti mitochondrion ;isoprenoic
18688_at At4g15560 DEF (CLA1) protein; su Chloroplast 1-deoxyxylulose-5-phosph
12524_at At1g10920 disease resistance pro Undefined apoptosis ;defense respo
12525_at At2g28510 Dof zinc finger protein; Other (cytoplasm etc.
12526_at At4g27710 cytochrome P450 - like Secreted Proteins cytochrome P450 ;endom
expression: Parallel Coordinates
File Options
biotin.bio101
biotin.bio1B1
biotin.bio1B2
expression: Scatterplot (current)
File Options
biotin.bio1B1
biotin.bio101


FCModeler
File View Layout Graph Interaction Modify Tools Window
F:\FCModeler\graphs\Pathway
A
AT5G16390
CAC1A
biotin
HCS
BCCP1-biotin
BCCP-biotin
cytosol --> plastid
ISPD
AT2G02500
cytosol --> plastid stroma
CLA1
AT4G15560
cytosol --> plastid stroma
2-C-methyl-D-erythritol 4-Pi cytidylyltransferase
1-deoxy-D-xylulose-5-Pi synthase
glyceraldehyde-3-Pi
pyruvate
PDC
CT alpha-CT beta complex
BC-BCCP complex
2-C-methyl-D-erythritol 4-Pi
CoA
NADP
CTP
CO2
acetyl-CoA
ATP
Heteromeric ACCase
4-diphosphocytidyl-2-C-methyl-D-erythritol
NADPH
1-deoxy-D-xylulose-5-Pi
HCO3-
1-deoxy-D-xylulose-5-Pi reductoisomerase
H+
Pi
ADP
malonyl-CoA


MetachipGraph - supercluster4-1-05.mcg
File Project Help
Chip Name vs Expression Level Plot
B
Name Length:
Annotation:
Expression Level
Chip Name
At5g45930
At1g32990
At2g02500
Series: At2g02500
TAIR
Column:
Info
Y-Value:
Mark

..

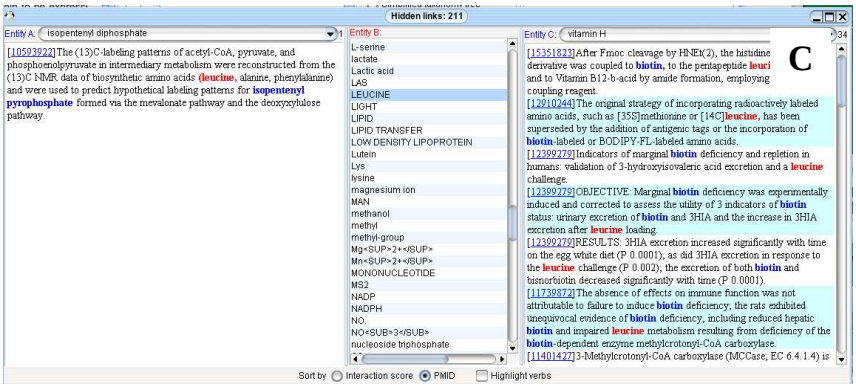

Hidden links: 211
Entity A: isopentenyl diphosphate
[10593922]The (13)C-labeling patterns of acetyl-CoA, pyruvate, and phosphoenolpyruvate in intermediary metabolism were reconstructed from the (13)C NMR data of biosynthetic amino acids (leucine, alanine, phenylalanine) and were used to predict hypothetical labeling patterns for isopentenyl pyrophosphate formed via the mevalonate pathway and the deoxyxylulose pathway.
Entity B:
L-serine
lactate
Lactic acid
LAS
LEUCINE
LIGHT
LIPID
LIPID TRANSFER
LOW DENSITY LIPOPROTEIN
Lutein
Lys
lysine
magnesium ion
MAN
methanol
methyl
methyl-group
Mg²⁺
Mn²⁺
MONONUCLEOTIDE
MS2
NADP
NADPH
NO.
NO₃
nucleoside triphosphate
Entity C: vitamin H
C
[15351823]After Fmoc cleavage by HNEt(2), the histidine derivative was coupled to biotin, to the pentapeptide leuci and to Vitamin B12-b-acid by amide formation, employing coupling reagent.
[12910244]The original strategy of incorporating radioactively labeled amino acids, such as [35S]methionine or [14C]leucine, has been superseded by the addition of antigenic tags or the incorporation of biotin-labeled or BODIPY-FL-labeled amino acids.
[12399279]Indicators of marginal biotin deficiency and repletion in humans: validation of 3-hydroxyisovaleric acid excretion and a leucine challenge.
[12399279]OBJECTIVE: Marginal biotin deficiency was experimentally induced and corrected to assess the utility of 3 indicators of biotin status: urinary excretion of biotin and 3HIA and the increase in 3HIA excretion after leucine loading.
[12399279]RESULTS: 3HIA excretion increased significantly with time on the egg white diet (P 0.0001), as did 3HIA excretion in response to the leucine challenge (P 0.002); the excretion of both biotin and bisnorbiotin decreased significantly with time (P 0.0001).
[11739872]The absence of effects on immune function was not attributable to failure to induce biotin deficiency; the rats exhibited unequivocal evidence of biotin deficiency, including reduced hepatic biotin and impaired leucine metabolism resulting from deficiency of the biotin-dependent enzyme methylcrotonyl-CoA carboxylase.
[11401427]3-Methylcrotonyl-CoA carboxylase (MCCase; EC 6.4.1.4) is
Sort by Interaction score PMID Highlight verbs

Figure 1. Visualization of microarray data by ExploRase.
(A) Examining data quality. A scatter plot analysis of data from microarray replicates of two biological samples quickly reveals a problem with values for two plastid-encoded genes inherent in one of the replicates; these RNAs are simultaneously highlighted in a parallel coordinate plot view of the data. The raw data had previously been normalized, using R functions in ExploRase (not shown). (B) Differentially expressed genes. Scatterplots show data from the *bio1* mutant with added biotin compared to without added biotin. Several genes appear differentially expressed in the scatter plot, and were selected by the user (blue or pink highlights); the corresponding parts of the parallel coordinate plot are simultaneously highlighted; the annotation for these genes is also displayed. Results of statistical analyses can also be superimposed on the visualization results (not shown). The y-axis of the parallel coordinate plot has a log scale.

Fig 2. Exploration of genes whose expression increases in response to biotin in the *bio1* mutants.
(A) MetNet in Cytoscape was used to explore possible interrelationships between a suboptimal level of biotin and changes in gene expression as revealed by global microarray analysis. This example focuses on the increase in accumulation of two RNAs in the methylerythritol phosphate (MEP) pathway of plastidic isoprenoid synthesis. A graph containing pathways of central metabolism, including isoprenoid metabolism and starch metabolism, and a subset of the upregulated genes was selected (insert box in upper left). Clicking within this graph identified a subgraph that includes both biotin and the MEP differentially expressed genes. (B) MetachipGraph was used to determine the Pearson correlations across 1000 Arabidopsis ATH1 chips in the NASCArrays database, comparing the expression pattern of At2g02500 to that of the other 22,746 genes on the chip. The genes most similar to At2g02500 are At5g45930 and At1g32990 (87% and 86% correlation, respectively). Both of these genes are involved in plastid function. At5g45930 encodes a magnesium-chelatase subunit, ChlI, which is required for biosynthesis of the isoprenoid-porferin hybrid molecule, chlorophyll. (C) PathBinder was used to explore interconnections in the literature between biotin and isoprenoids.