

Software

Open Access

## PageMan: An interactive ontology tool to generate, display, and annotate overview graphs for profiling experiments

Björn Usadel\*<sup>1</sup>, Axel Nagel<sup>2</sup>, Dirk Steinhauser<sup>1</sup>, Yves Gibon<sup>1</sup>,  
Oliver E Bläsing<sup>1</sup>, Henning Redestig<sup>1</sup>, Nese Sreenivasulu<sup>3</sup>, Leonard Krall<sup>1</sup>,  
Matthew A Hannah<sup>1</sup>, Fabien Poree<sup>1</sup>, Alisdair R Fernie<sup>1</sup> and Mark Stitt

Address: <sup>1</sup>Max Planck Institute of Molecular Plant Physiology, Am Mühlenberg 1, 14476 Potsdam, Germany, <sup>2</sup>RZPD: Deutsches Ressourcenzentrum für Genomforschung GmbH, Heubnerweg 6, 14059 Berlin, Germany and <sup>3</sup>Leibniz Institute of Plant Genetics and Crop Plant Research (IPK), Corrensstr. 3, 06466 Gatersleben, Germany

Email: Björn Usadel\* - [usadel@mpimp-golm.mpg.de](mailto:usadel@mpimp-golm.mpg.de); Axel Nagel - [nagel@rzpd.de](mailto:nagel@rzpd.de); Dirk Steinhauser - [steinhauser@mpimp-golm.mpg.de](mailto:steinhauser@mpimp-golm.mpg.de); Yves Gibon - [gibon@mpimp-golm.mpg.de](mailto:gibon@mpimp-golm.mpg.de); Oliver E Bläsing - [oliver.blaesing@metanomics.de](mailto:oliver.blaesing@metanomics.de); Henning Redestig - [redestig@mpimp-golm.mpg.de](mailto:redestig@mpimp-golm.mpg.de); Nese Sreenivasulu - [srinivas@ipk-gatersleben.de](mailto:srinivas@ipk-gatersleben.de); Leonard Krall - [krall@mpimp-golm.mpg.de](mailto:krall@mpimp-golm.mpg.de); Matthew A Hannah - [hannah@mpimp-golm.mpg.de](mailto:hannah@mpimp-golm.mpg.de); Fabien Poree - [poree@mpimp-golm.mpg.de](mailto:poree@mpimp-golm.mpg.de); Alisdair R Fernie - [fernie@mpimp-golm.mpg.de](mailto:fernie@mpimp-golm.mpg.de); Mark Stitt - [mstitt@mpimp-golm.mpg.de](mailto:mstitt@mpimp-golm.mpg.de)

\* Corresponding author

Published: 18 December 2006

Received: 05 September 2006

*BMC Bioinformatics* 2006, **7**:535 doi:10.1186/1471-2105-7-535

Accepted: 18 December 2006

This article is available from: <http://www.biomedcentral.com/1471-2105/7/535>

© 2006 Usadel et al; licensee BioMed Central Ltd.

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/2.0>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

### Abstract

**Background:** Microarray technology has become a widely accepted and standardized tool in biology. The first microarray data analysis programs were developed to support pair-wise comparison. However, as microarray experiments have become more routine, large scale experiments have become more common, which investigate multiple time points or sets of mutants or transgenics. To extract biological information from such high-throughput expression data, it is necessary to develop efficient analytical platforms, which combine manually curated gene ontologies with efficient visualization and navigation tools. Currently, most tools focus on a few limited biological aspects, rather than offering a holistic, integrated analysis.

**Results:** Here we introduce PageMan, a multiplatform, user-friendly, and stand-alone software tool that annotates, investigates, and condenses high-throughput microarray data in the context of functional ontologies. It includes a GUI tool to transform different ontologies into a suitable format, enabling the user to compare and choose between different ontologies. It is equipped with several statistical modules for data analysis, including over-representation analysis and Wilcoxon statistical testing. Results are exported in a graphical format for direct use, or for further editing in graphics programs.

PageMan provides a fast overview of single treatments, allows genome-level responses to be compared across several microarray experiments covering, for example, stress responses at multiple time points. This aids in searching for trait-specific changes in pathways using mutants or transgenics, analyzing development time-courses, and comparison between species. In a case study, we analyze the results of publicly available microarrays of multiple cold stress experiments using PageMan, and compare the results to a previously published meta-analysis.

PageMan offers a complete user's guide, a web-based over-representation analysis as well as a tutorial, and is freely available at <http://mapman.mpimp-golm.mpg.de/pageman/>.

**Conclusion:** PageMan allows multiple microarray experiments to be efficiently condensed into a single page graphical display. The flexible interface allows data to be quickly and easily visualized, facilitating comparisons within experiments and to published experiments, thus enabling researchers to gain a rapid overview of the biological responses in the experiments.

## Background

Recent advances in microarray technologies have led to an avalanche of gene expression data from a variety of organisms. Many of these microarray experiments have been deposited in array databases and are available for public scrutinizing and data mining purposes [1-3]. Initially, microarray experiments involved comparison of one or a small number of treatments to a control. There is now a trend to more complex experiments, consisting of time-course (e.g. [4,5]) or dose-response studies. As an increasing number of data sets are deposited in the public domain, it is also becoming important to compare large numbers of treatments that are putatively similar, or that may share common components.

Given the initial limitation of only a few arrays per experiment, tools were first developed to visualize data on a one-experiment-at-a-time basis ([6-8]). Tools have also been developed that focus on the response of one (or several related) genes through multiple time points/experiments with a limited ontology structure [9]. There is still a need for visualization tools that allow data sets from multiple experimental conditions to be integrated and condensed into a single graphic display.

One common way to get an overview of a given microarray experiment is by using biological ontology structures. For example, many tools exist that utilize GO [10], KEGG, or MIPS functional categories to provide an overview of changes in expression using overrepresentation (ORA) Reviewed by [11] or other approaches [6,12]. As a complement to pre-defined annotations, *ad hoc* manual annotations or groupings of genes that are differentially expressed can be used for presentation and/or interpretation. In these approaches the functional ontology is often displayed as a tree-like graph, and the individual results are usually represented in tabular format.

In this paper we describe PageMan, a software tool which facilitates an ontologically-defined overview of the global response of the transcriptome. It provides a statistically-based overview of the enriched functional categories from global transcriptome responses, and can be viewed either in tabular form, or via a false color heat-map like display. This data-condensation allows the main global features of a single treatment to be rapidly identified, and facilitates the comparison of large numbers of treatments. PageMan can be used with various functional ontologies. It implements a Wilcoxon analysis to directly infer the contribution of individual categories to the response of the whole experiment [6]. It also implements an over/under representation analysis using either Fisher's exact test, or a  $\chi^2$  test combined with thresholds that the user can set individually. Furthermore, for convenience, a web based ORA analysis is offered on the PageMan website. PageMan gen-

erates direct-to-use, editable figures, which display both the hierarchy tree, as well as the changes within the individual experiments.

## Implementation

### Overview

PageMan is a standalone desktop application with Graphical User Interface (GUI) implemented in Java using the java swing libraries and parts of the MapMan source tree. Thus, PageMan should run on any Java enabled platform. It has been tested on Microsoft Windows XP, various Linux systems and on Apple's OS X operating system. Using a standard installation, PageMan can handle between 30-40 experiments at one time. However, by increasing the available java heap space to 1 GiByte enables PageMan to deal with hundreds of arrays.

The analysis algorithm for the Wilcoxon test has been described earlier [6]. For the ORA analysis a Fisher's exact or a  $\chi^2$  test is performed using a newly written Java class. The output of this class has been tested by performing the same tests in R [13]. This class can also be used independently of PageMan. In addition to the standard Java libraries, PageMan uses a number of third party libraries as support for specific operations. It relies on FreeHep libraries for graphical export, on JexcelApi library for import of excel files, which are one of the most widely used file formats of data representation in the biological sciences, and makes use of the Dom4j libraries for parsing XML files.

### Inputs

PageMan requires a mapping file, which assigns each probe identifier on the array into at least one functional category. In the examples presented here, the mapping file is based on the MapMan ontology for Arabidopsis genes described in the user's manual and elsewhere [6,14]. We provide a GUI tool to translate MIPS, KEGG, or GO hierarchies into this format thus enabling the user to choose the most appropriate ontology or compare the different ontologies.

PageMan can use several different types of input for the experimental data, depending on the operation desired e.g.  $\log_2$  fold change or p-values of differential expression obtained by freely available, standard array handling software such as BioConductor [15]. Alternatively, if the user requires only data visualization it is possible to generate a PageMan native file (in tab separated text format, see user's manual for details). Using this format it is possible to display any kind of data as false-color boxes, thus enabling the use of PageMan for a multitude of other applications.

### **Computation within PageMan**

#### *Over/Under-representation analysis*

For over-representation analysis, PageMan uses Fisher's exact or a  $\chi^2$  test to calculate the likelihood for each category to contain the number of objects exceeding a user-definable threshold that is actually observed, given the total number of objects and the total number of objects exceeding this threshold. This threshold depends on an analysis previously performed. For example, one could analyze which genes surpass a certain fold-change value, or which genes are below a certain p-value, if differential expression has already been calculated. To facilitate interpretation, PageMan applies the same procedure for objects below the negative value of the threshold and adds the tags "up" or "down" to the experiment names respectively.

Unlike the  $\chi^2$  test, Fisher's exact test is also applicable for extreme cases of test situations, such as only observing a small number of objects per class or small classes. However, a  $\chi^2$  test with Yates continuity correction is offered as an alternative for testing; in this case, ontological groups with too few items are omitted. The calculations are based upon the approximation of the Gamma function by Lanczos [16] as implemented in the Gnu Scientific Library which has been ported to Java [17].

#### *Wilcoxon Test Statistic*

PageMan uses an internal routine to compute an unpaired Wilcoxon rank sum test statistic (equivalent to Mann-Whitney's U test). If a table of fold-change values was given as an input, this feature would test whether the median fold-change within a particular ontological group was the same as the median fold-change of all genes not in that ontological group. Unlike ORA based tests, the Wilcoxon test does not require setting a sometimes subjective threshold.

#### *Multiple hypothesis testing correction*

PageMan allows the user to not only test one hypothesis (e.g. is glycolysis up-regulated/over-represented?) but to test up to hundreds of hypotheses at once (are any of the functional categories changed?). It is therefore necessary to implement multiple-hypothesis-testing correction methods. This is achieved using three different methods: the conservative Bonferroni correction that controls the family wise error rate, and the false discovery rate control methods by Benjamini, Hochberg [18] and Benjamini, Yekutieli [19]. After correcting for multiple testing, "adjusted" p-values are computed according to the correction method specified. In the case of the false discovery rate controlling corrections, these new values actually represent the false discovery rate level e.g. using a value of 0.05 as a cut-off would mean accepting a false discovery rate level of 5%. None of these testing corrections takes

the nested hierarchy into account and may therefore lead to slightly biased results.

#### *Conversion for display*

In order to display (adjusted) p-values in PageMan, they are transformed into their respective z-values. All p-values above 0.05 are set to a z-value of 0 to avoid misinterpretation. The resulting values are then false color coded in a user-adjustable two color scale. Here, a highly saturated color indicates a high absolute value, whereas smaller values are indicated by a lower color saturation. For the Wilcoxon's test p-values, two different colors (e.g., blue and red) can be selected to distinguish between categories where the average of the signals for all the genes in a category increases or decreases.

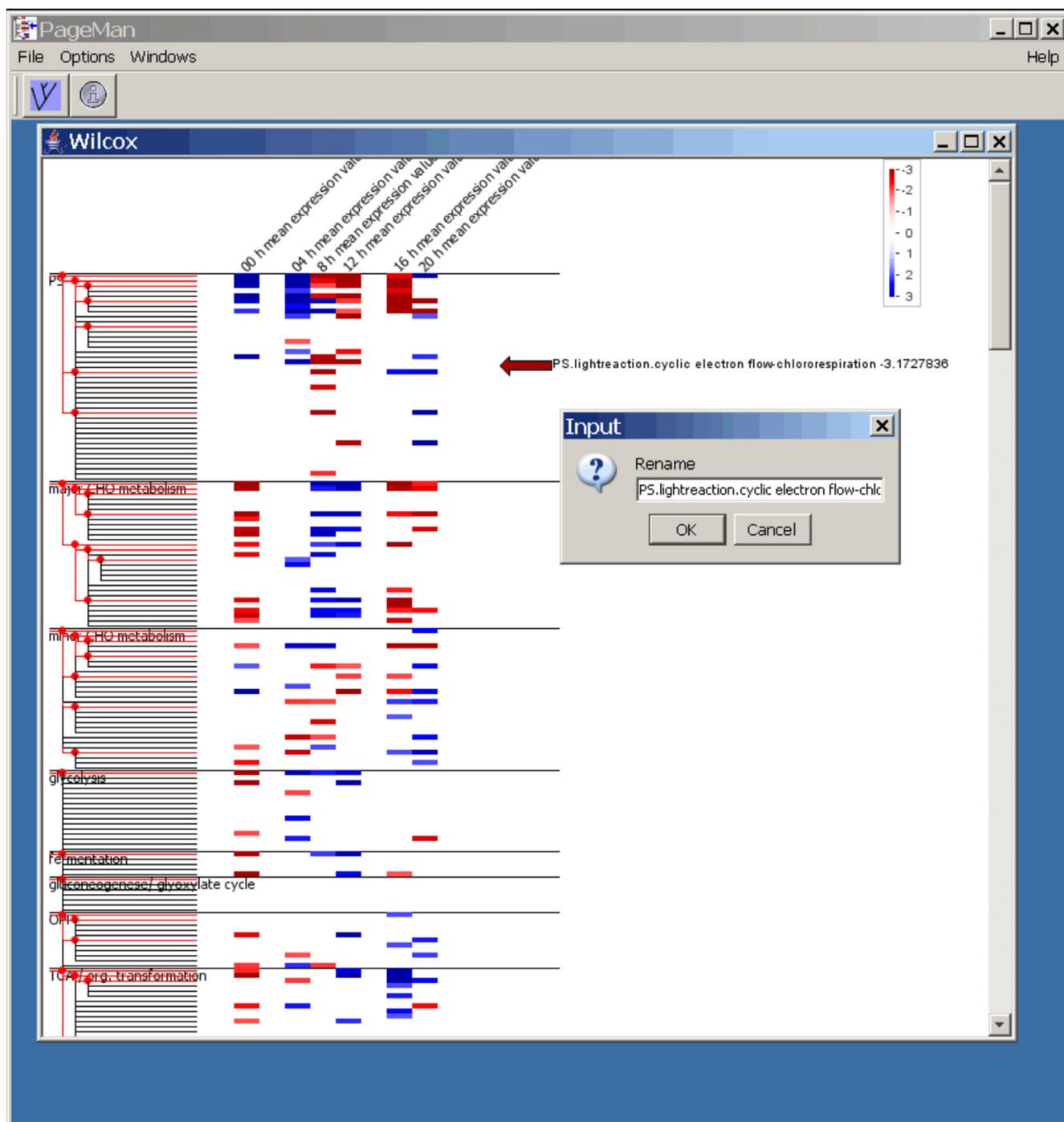
### **The PageMan GUI**

PageMan was designed with ease of use in mind: for this reason the user is guided through the analysis by a wizard. Once the analysis has finished, the user is presented with a heat map (overview transcript map) with representation of the differently enriched/differently behaving functional categories within the various experiments by false color coded boxes. This view can be overlaid with a tree representing the hierarchical information among different functional categories (Figure 1, left hand side). For flexible visualization, individual nodes of the hierarchy tree can be collapsed to remove areas of the tree that are uninteresting (e.g. because there are non-significant changes). Alternatively, all parents having only non-significant nodes can be collapsed or all non-significant nodes can be hidden. The boxes of the false color display can be identified and annotated by clicking or by using the command to "annotate all significant nodes"; this allows an editable, moveable annotation arrow to be added directly opposite the heat map feature. Annotations that are not required can be manually removed. Finally, experiments can be deleted from the display, and spacers can be added to separate groups of experiments to optimize the visual appearance.

For the layout, several different options are accessible via the options menu. For example, the dimensions of the boxes as well as the color intensities for the boxes can be set according to the user's choice. For depicting differential expression, several different color schemes (red-white-blue, red-black-green etc.) are available. Finally, sub-categories can be opened in a separate window for closer inspection.

#### **Graphical Output**

PageMan comes with various graphical export capabilities, which support the production of suitable graphics for viewing or even for pre-publication stages. The visualization display can be exported in standard bitmapped for-



**Figure 1**  
**PageMan Analysis and Annotation Session.** An ongoing annotation session. The user is presented with the results of the statistical analysis using a false color scale. On the left hand side, the user can access an ontology tree which can be collapsed on demand. On the right hand side a color scale is shown.

mats (such as png or jpeg), and in vector formats such as svg, ps, pdf and the windows specific emf format. This allows the visualization to be imported into various downstream applications such as Microsoft PowerPoint

or Corel Draw, where the individual elements can be further edited without loss of quality for final manuscript preparation and/or presentation. As indicated above, it is possible in advance to collapse or expand nodes while

preparing the visualization display, in order to focus on selected features of the response.

### **Documentation**

Help is available directly from the program itself by simply selecting the help menu item from within PageMan. Moreover, on our website we offer a step-by-step tutorial that guides one through the use of PageMan.

## **Results and discussion**

### **Exemplary comparison of multiple experiments**

To demonstrate the use of PageMan for multiple experiments, Arabidopsis cold stress experiments were downloaded from NASC Arrays [3] and evaluated. The RMA expression values for the samples were calculated [20] and a linear model was fitted using BioConductor [15,21]. The resulting  $\log_2$  fold change values at each time point were calculated and used for PageMan. The data was processed in PageMan using ORA analysis with Fisher's exact test, setting a threshold of 1 (at least a two fold change). All categories that have more/less genes than expected that exceed this threshold are colored with increasing intensity. An example from the PageMan visualization is shown in figure 2, where categories for transcription factors have been magnified using PageMan's "extract and enlarge" function. AP2/EREBP and Constans-like transcription factors are consistently over-represented amongst the up-regulated genes. Over-representation of MYB related genes amongst the down-regulated genes can be seen in most experiments. These responses are in accordance with earlier meta-analysis of cold acclimation using MapMan ontologies performed by Hannah et al. 2005 (see Table S6 from Hannah et al. 2005[22]). However, the earlier analysis was time-consuming, requiring either manual bin counts or scripting based on customized mapping files. PageMan performs this type of analysis in a few minutes, including annotation and layout, resulting in a graph like that shown in figure 1. It also allows equally rapid analysis using the other three enrichment-based statistical tests included in the package. Thus, PageMan provides a quick integration at the ontology level across multiple similar experiments, and allows comparison of their similarities and differences. As exemplified, the PageMan graphical interface provides an intuitive visualization overview representing "hot-spot pathways" activated during Arabidopsis cold stress across experiments performed by different labs.

### **Comparison of PageMan with related tools**

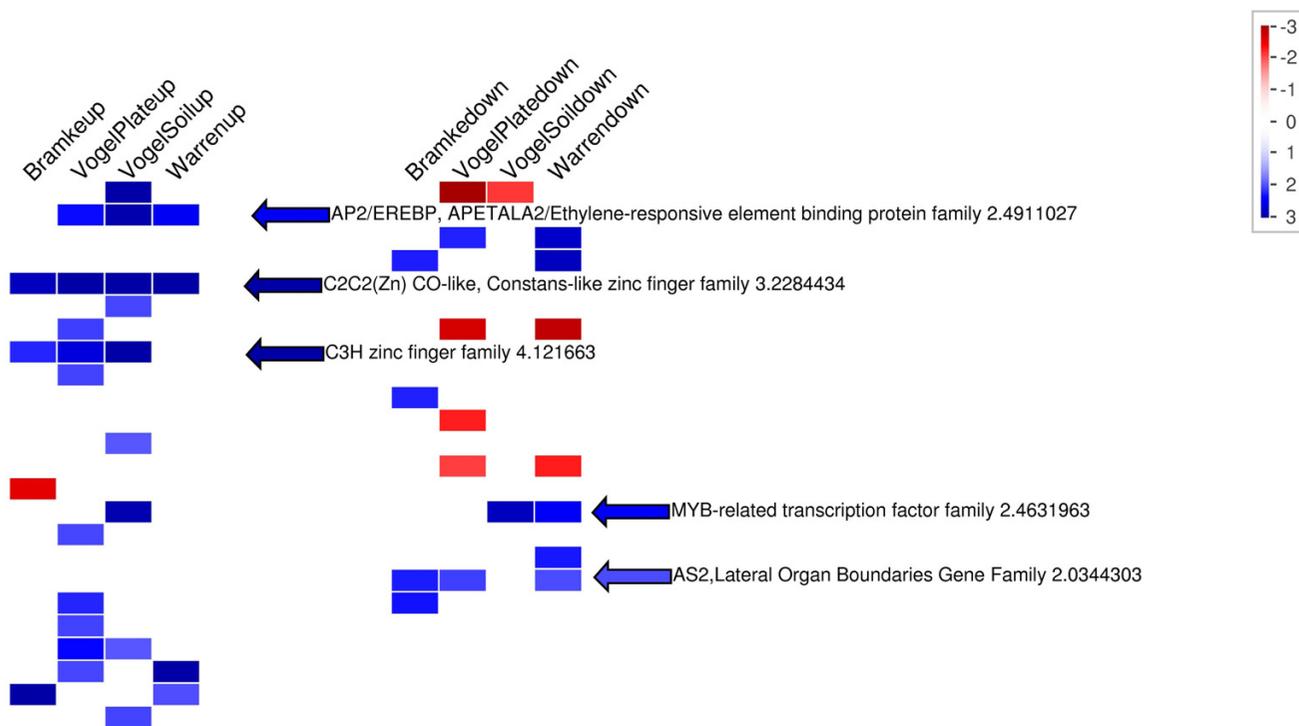
Most currently available tools are limited to a few (usually enrichment-based) statistical models such as either the hypergeometric or the binomial distribution. Within the 14 tools reviewed by Khatri and Draghici, only one, namely the Onto-Express tool, supported four different enrichment-based statistical tests. PageMan supports the

use of Fisher's exact test and  $\chi^2$  statistics as well as Wilcoxon's test. Unlike the web-based tool JProGO [23] it allows use of the non-enrichment based Wilcoxon test, without the web-based limitations. Also, many tools are limited to a few experiments at a time, whereas using PageMan evaluating hundreds of experiments at once is possible.

Most available tools only support the GO ontology, or GO and KEGG in the case of AMDA. To the best of our knowledge, PageMan is the only tool supporting the use of MapMan, KEGG, MIPS, and GO ontologies. By providing a parser to automatically format these ontologies, PageMan offers the user unprecedented flexibility to use whatever ontology is strongest or most advanced in their particular field of study. Thus using PageMan it is also possible to classify metabolite data, which is not possible based on the GO ontology. As discussed by Khatri and Draghici [11], most tools that use the GO ontology are not able to use a higher level of abstraction because they can only use the lowest level of the hierarchy. PageMan allows the user to flexibly collapse nodes that are of no interest for the user, and by default analyze all levels of abstraction.

Also, many tools are limited to a few experiments at a time, whereas using PageMan evaluating hundreds of experiments at once is possible. Further, PageMan supports the subsequent introduction of more array data (including that from a different organism or a different array platform) for comparison. Among the tools having a user interface, this represents a rather novel feature. Although this is also possible by using R/Bioconductor, substantial programming skills are required. As Manoli et al. point out, group testing helps in comparing different datasets [24].

In terms of graphical capabilities and the ability to upload multiple experiments, High-Throughput GO-Miner [25] and AMDA [26] are most similar to PageMan, as they also offer heat-map like graphics. However, unlike PageMan, these tools are not interactive. High-Throughput Go-Miner sometimes requires manual editing of configuration files, and the installation requires connecting to an SQL server or for the user to install their own SQL server. This is typically beyond the means of most users. Further, although a database approach offers more flexibility, the use of dumped files for classification (as in PageMan) offers a significant speed improvement because the time intensive step of connecting to a remote database or a web-service only needs to happen once. AMDA, while offering a widget based interface still requires the installation of Tcl/Tk on top of R and BioConductor, which can be cumbersome on Windows. PageMan, on the other hand, is packaged with an installer and the user can down-



**Figure 2**  
**PageMan Analysis of a Cold Stress Time Series.** Here, different cold stress experiments obtained from NASC were supplied as log<sub>2</sub> fold change values. An over-representation analysis was performed using Fisher's exact test and 1 as a cut-off. Using PageMan's "zoom in" function, only potential transcription factors were displayed.

load necessary files from the internet and thus remain totally anonymous and independent of internet connections for analyses. Further, unlike PageMan, the heatmaps generated by High-Through Put GO Miner or AMDA are static and annotations can not be easily edited.

**Conclusion**

We have developed a novel, platform-independent tool, PageMan, which is available free of charge. It aids in interpreting individual microarray experiments and in exploring large sets of microarray experiments by analyzing and summarizing the data, and then visualizing it in an ontological context. With this tool it is possible to quickly compare given data to published results and/or to pinpoint special biological processes or pathways that may need to be investigated more thoroughly. PageMan also allows comparison of the global response to analogous treatments in different species, provided that a comparable ontology is possible (see e.g. [27] and [28], for an extension of the MapMan ontology to tomato and *Medicago*). It is planned to extend the MapMan ontology to fur-

ther crop plants and wild plant species, as large scale array sets become available for them. Moreover, PageMan will be adapted to include p-value corrections that take nested hierarchies into account, once these become widely accepted.

Even though many tools have been generated over the years that use ontological categories to statistically assess and summarize data, PageMan offers the unique possibility to layout, visualize, and annotate information from large transcriptome series experiments in an integrated manner using a single tool. Furthermore, it is generic, and can be applied to other large quantitative data sets obtained from enzymatic, metabolomics, or proteomic approaches. This offers the research community a tool to both globally analyze and identify "hot-spot regulated pathways" and immediately export publication ready pictures.

**Availability and requirements**

- Project name: PageMan

- Project home page: <http://mapman.mpimp-golm.mpg.de/pageman/index.html>
- Operating system(s): Platform independent
- Programming language: Java
- Other requirements: Java 1.4 or higher
- License: freely available. The software uses libraries covered by the LGPL (freehep for graphics import) and others (dom4j for xml import).
- Any restrictions to use by non-academics: none

### Authors' contributions

BU designed the software, implemented the visualization, and tested and improved the statistics. AN implemented data loading, foreign ontology parsing, document handling as well as parts of the user interfaces, AN streamlined the whole code, and fixed bugs. MS presented the problem and came up with an initial concept, discussed additional users' needs and presented UI improvements. HR identified the Wilcoxon test as a possible tool and MH tested and discussed ORA analysis. YG, OB, and MS brought up the concept of data condensation. FP discussed UI improvements. ARF contributed tomato ontologies, NS contributed barley ontologies. YG, OB, ARF, NS, DS, HR, LK, FP and MS located mistakes in the software. All authors have read and approved the final manuscript.

### Acknowledgements

BU and AN were supported by the BMBF GABI grant no. 0313110/0313112.

### References

- Parkinson H, Sarkans U, Shojatalab M, Abeygunawardena N, Contrino S, Coulson R, Farnie A, Lara GG, Holloway E, Kapushesky M, Lilja P, Mukherjee G, Oezcimen A, Rayner T, Rocca-Serra P, Sharma A, Sansone S, Brazma A: **ArrayExpress – a public repository for microarray gene expression data at the EBI.** *Nucleic Acids Res* :D553-5. 2005 Jan 1
- Barrett T, Suzek TO, Troup DB, Wilhite SE, Ngau WC, Ledoux P, Rudnev D, Lash AE, Fujibuchi W, Edgar R: **NCBI GEO: mining millions of expression profiles – database and tools.** *Nucleic Acids Res* :D562-6. 2005 Jan 1
- Craigon DJ, James N, Okyere J, Higgins J, Jotham J, May S: **NASCARrays: a repository for microarray data generated by NASC's transcriptomics service.** *Nucleic Acids Res* 2004, **32**:D575-7.
- Blasing OE, Gibon Y, Gunther M, Hohne M, Morcuende R, Osuna D, Thimm O, Usadel B, Scheible WR, Stitt M: **Sugars and circadian regulation make major contributions to the global regulation of diurnal gene expression in Arabidopsis.** *Plant Cell* 2005, **17**:3257-81.
- [<http://web.uni-frankfurt.de/fb15/botanik/mcb/AFGN/atgenex.htm>].
- Usadel B, Nagel A, Thimm O, Redestig H, Blasing OE, Palacios-Rojas N, Selbig J, Hannemann J, Piques MC, Steinhauser D, Scheible WR, Gibon Y, Morcuende R, Weicht D, Meyer S, Stitt M: **Extension of the visualization tool MapMan to allow statistical analysis of arrays, display of corresponding genes, and comparison with known responses.** *Plant Physiol* 2005, **138**:1195-204.
- Tokimatsu T, Sakurai N, Suzuki H, Ohta H, Nishitani K, Koyama T, Umezawa T, Misawa N, Saito K, Shibata : **KaPPA-view: a web-based analysis tool for integration of transcript and metabolite data on plant metabolic pathway maps.** *Plant Physiol* 2005, **138**(3):1289-300.
- Mueller LA, Zhang P, Rhee SY: **AraCyc: a biochemical pathway database for Arabidopsis.** *Plant Physiol* 2003, **132**(2):453-60.
- Junker BH, Klukas C, Schreiber F: **VANTED: a system for advanced data analysis and visualization in the context of biological networks.** *BMC Bioinformatics* 2006, **7**:109.
- The Gene Ontology Consortium: **Gene Ontology: tool for the unification of biology.** *Nature Genet* 2000, **25**:25-29.
- Khatri P, Draghici S: **Ontological analysis of gene expression data: current tools, limitations, and open problems.** *Bioinformatics* 2005, **21**:3587-3595.
- Lee HK, Braynen W, Keshav K, Pavlidis P: **ErmineJ: tool for functional analysis of gene expression data sets.** *BMC Bioinformatics* 2005, **6**:269.
- R Development Core Team R: **A language and environment for statistical computing R Foundation for Statistical Computing, Vienna, Austria.** . ISBN 3-900051-07-0
- Thimm O, Blasing O, Gibon Y, Nagel A, Meyer S, Kruger P, Selbig J, Muller LA, Rhee SY, Stitt M: **MAPMAN: a user-driven tool to display genomics data sets onto diagrams of metabolic pathways and other biological processes.** *Plant J* 2004, **37**(6):914-39.
- Gentleman RC, Carey VJ, Bates DM, Bolstad B, Dettling M, Dudoit S, Ellis B, Gautier L, Ge Y, Gentry J, Hornik K, Hothorn T, Huber W, Iacus S, Irizarry R, Leisch F, Li C, Maechler M, Rossini AJ, Sawitzki G, Smith C, Smyth G, Tierney L, Yang JY, Zhang J: **Bioconductor: open software development for computational biology and bioinformatics.** *Genome Biol* 2004, **5**:R80.
- Lanczos C: **A Precision Approximation of the Gamma Function.** *SIAM Journal on Numerical Analysis series B* 1964, **1**:86-96.
- [<http://www.gnu.org/software/gsl/>].
- Benjamini Y, Hochberg Y: **Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing.** *J Roy Stat Soc Ser B* 1995, **57**:289-300.
- Benjamini Y, Yekutieli D: **The control of the false discovery rate in multiple testing under dependency.** *Annals of Statistics* 2001, **29**:1165-1188.
- Irizarry RA, Hobbs B, Collin F, Beazer-Barclay YD, Antonellis KJ, Scherf U, Speed TP: **Exploration, normalization, and summaries of high density oligonucleotide array probe level data.** *Journal Biostatistics* 2003, **4**:249-264.
- Smyth GK: **Limma: linear models for microarray data.** In *Bioinformatics and Computational Biology Solutions using R and Bioconductor* Edited by: Gentleman R, Carey V, Dudoit S, Irizarry R, Huber W. Springer, New York:397-420.
- Hannah MA, Heyer AG, Hinch DK: **A global survey of gene regulation during cold acclimation in Arabidopsis thaliana.** *PLoS Genet* 2005, **1**(2):e26.
- Scheer M, Klawonn F, Munch R, Grote A, Hiller K, Choi C, Koch I, Schobert M, Hartig E, Klages U, Jahn D: **JProGO: a novel tool for the functional interpretation of prokaryotic microarray data using Gene Ontology information.** *Nucleic Acids Res* 2006, **34**:W510-5.
- Manoli T, Gretz N, Grone HJ, Kenzelmann M, Eils R, Brors B: **Group testing for pathway analysis improves comparability of different microarray datasets.** *Bioinformatics* 2006, **22**:2500-6.
- Pelizzola M, Pavelka N, Foti M, Ricciardi-Castagnoli P: **AMDA: an R package for the automated microarray data analysis.** *BMC Bioinformatics* 2006, **7**:335.
- Zeeberg BR, Qin H, Narasimhan S, Sunshine M, Cao H, Kane DW, Reimers M, Stephens RM, Bryant D, Burt SK, Elnekave E, Hari DM, Wynn TA, Cunningham-Rundles C, Stewart DM, Nelson D, Weinstein JN: **High-Throughput GoMiner, an 'industrial-strength' integrative gene ontology tool for interpretation of multiple-microarray experiments, with application to studies of Common Variable Immune Deficiency (CVID).** *BMC Bioinformatics* 2005, **6**:168.
- Urbanczyk-Wochniak E, Usadel B, Thimm O, Nunes-Nesi A, Carrari F, Davy M, Blasing O, Kowalczyk M, Weicht D, Polinceusz A, Meyer S, Stitt M, Fernie AR: **Conversion of MapMan to allow the analysis of transcript data from Solanaceous species: effects of genetic and environmental alterations in energy metabolism in the leaf.** *Plant Mol Biol* 2006, **60**:773-92.
- Tellström V, Usadel O, Stitt M, Küster H, Niehaus K: **The lipopolysaccharide of Sinorhizobium meliloti suppresses**

defense-associated gene expression in cell cultures of the host plant *Medicago truncatula*. *Plant Physiol* 2007 in press.

Publish with **BioMed Central** and every scientist can read your work free of charge

*"BioMed Central will be the most significant development for disseminating the results of biomedical research in our lifetime."*

Sir Paul Nurse, Cancer Research UK

Your research papers will be:

- available free of charge to the entire biomedical community
- peer reviewed and published immediately upon acceptance
- cited in PubMed and archived on PubMed Central
- yours — you keep the copyright

Submit your manuscript here:  
[http://www.biomedcentral.com/info/publishing\\_adv.asp](http://www.biomedcentral.com/info/publishing_adv.asp)

