Published in "Trends in Plant Science 11(7): 320-322, 2006"
which should be cited to refer to this work.

http://doc.rero.ch

FiRe and microarrays: a fast answer to burning questions

Christophe Garcion[1], ApplimathFRI[2,3] and Jean-Pierre Métraux[1]

[1]Department of Biology, ch. du Musée 10, University of Fribourg CH-1700 Fribourg, Switzerland
[2]Department of Mathematics, ch. du Musée 10, University of Fribourg CH-1700 Fribourg, Switzerland
[3]Ecole d'ingénieurs, Bd de Pérolles 80, 1705 Fribourg, Switzerland

FiRe is a user-friendly Excel® macro designed to survey microarray data rapidly. This software interactively assembles data from different experiments and produces lists of candidate genes according to patterns of gene expression. Furthermore, macros bundled with FiRe can compare lists of genes, merge information from different spreadsheets, link candidates to information available from web-based databases, and produce heatmaps for easy visualization of microarray data. FiRe is freely available at http://www.unifr.ch/plantbio/FiRe/main.html.

Microarray data: joy and awe

DNA microarray expression studies generate an unprecedented wealth of data and place the researcher in the privileged position of being able to uncover hitherto unseen relationships between physiological situations and gene regulation. They offer the opportunity to single out genes for further analysis, based on the expectation that observed changes in the levels of accumulation of mRNA will be relevant to the process under study. Several algorithms have been developed to identify differentially expressed genes from microarray data [1–4], and some of them have been used successfully to isolate candidate genes (e.g. Ref. [5]). These programmes rely on complex statistical evaluations and sophisticated user interfaces, and might require a lot of effort from the untrained and unassisted biologist needing an answer to simple questions. Another widespread approach to mine microarray data relies on the clustering of genes according to their pattern of accumulation of mRNA. This method might be appealing for the user to retrieve candidate genes displaying a desired pattern of expression, and is also advocated when some marker genes have already been shown to be involved in the process under study [6–8]. However, different results can be obtained depending on the clustering algorithms used [9], making the task difficult for biologists not familiar with the complexity of dedicated software and the subsequent analysis of data.

All FiRed up

Microarray data from multiple experiments can also be sorted using the ratio of expression values of 'treated' over 'controls' (fold-change ratio). This straightforward criterion has been reported to perform well for many biological datasets [1], bringing candidate genes for further experimental assessment to the attention of the user. To perform a quick and intuitive analysis of microarray data, requiring little technical knowledge and no specialized software, we created a macro named FiRe (for *Fi*nd *Re*gulons) that operates within the widely distributed Excel® software and selects candidate genes based upon their fold-change ratios. FiRe does not address the difficult problem of estimating the statistical significance of the fold-change ratios from replicate datasets, but instead identifies candidate genes that show a specific pattern of mRNA accumulation under different experimental conditions. When FiRe is activated, normalized microarray data are first assembled interactively into a single table (Figure 1). A filter button is provided to remove signals below a cut-off value. Fold-change ratios of the 'treated' compared with the 'control' conditions are then computed, and genes are selected depending on user-defined upper or lower thresholds. An easy formulation of the desired expression pattern is made possible by the assignment of numerical qualifiers (NQs). A NQ of 1 or −1 is attributed if the expression ratio of 'treated' over 'control' is greater or smaller than the upper or lower cut-offs. A NQ of 0 is attributed if the expression ratio lies between the two thresholds. The user can then activate NQ selection buttons to extract subgroups of genes corresponding to a defined expression pattern (regulons). For example, in an experiment with four treatments, the regulon with the NQ formula 1010 will correspond to genes induced by treatments 1 and 3 but unchanged in the others. The upper and lower threshold values can be modified independently for each treatment, and the NQs will be recalculated in real time. It is up to the user to check the relevance of the genes displayed as candidates, or to define appropriate, biologically significant cut-off values. Interesting regulons can be copied on a separate sheet equipped for further sorting.

After the harvest

FiRe is a handy tool to extract subsets of genes according to a desired pattern of accumulation of mRNA in relation to various stimuli or genetic lesions. The results of this analysis can be used, if desired, for further statistical evaluation with other published algorithms.

Corresponding author: Métraux, J-P. (jean-pierre.metraux@unifr.ch).

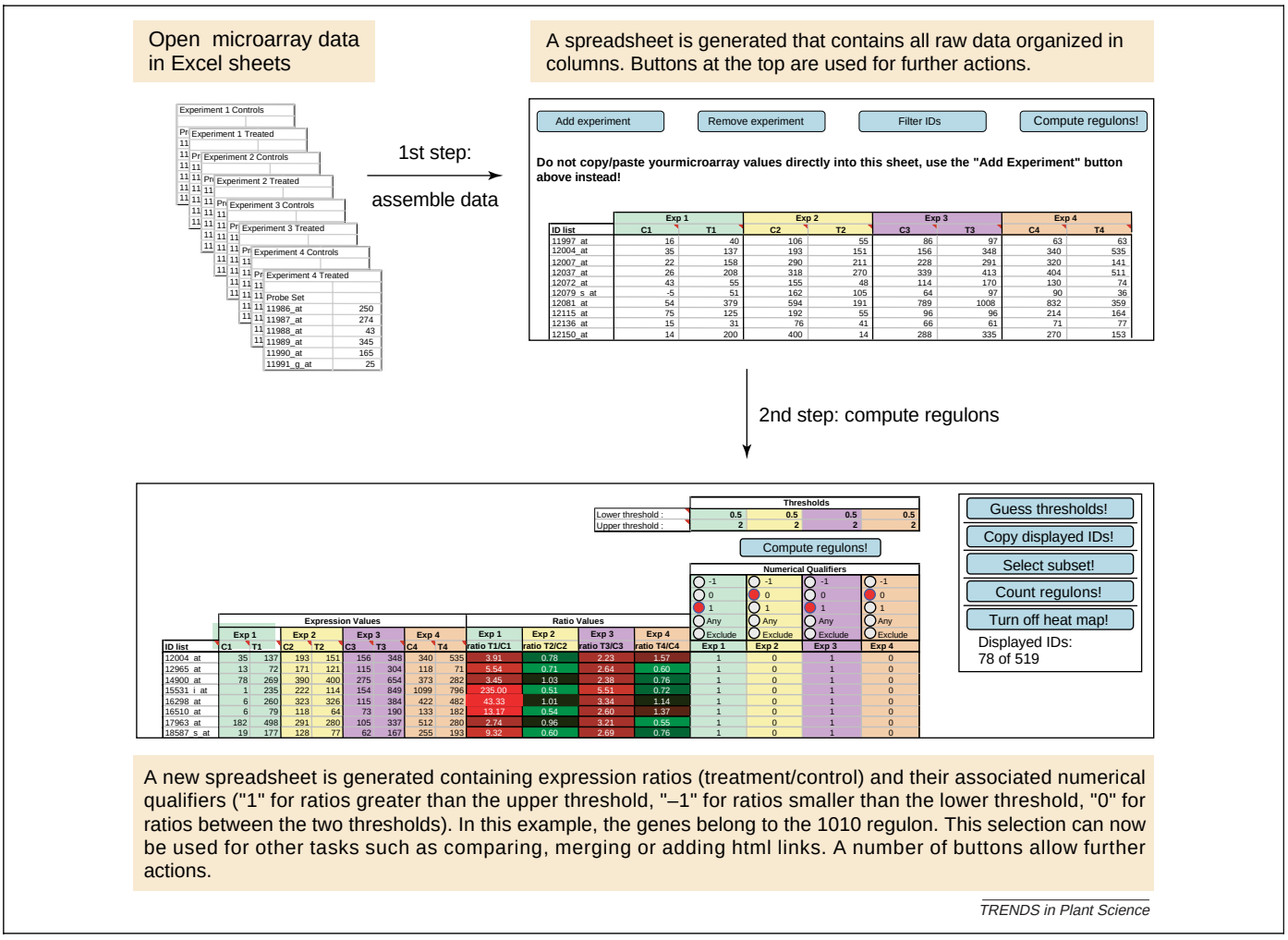

Figure 1. An interactive search using FiRe to search for candidate genes using a set of raw data available on Excel® spreadsheets. FiRe calculates the expression ratios, assigns numerical qualifiers (NQs) for each treatment and displays the results in a new sheet that also contains several action buttons. From this initial result sheet the user can then redefine upper and lower thresholds or select new NQ values. FiRe then displays genes corresponding to these search criteria. The results of such a search can be stored on a separate sheet for further sorting and analysis.

http://doc.rero.ch

The FiRe package also includes macros for grouping, accessing, comparing or visualizing supplementary data available for candidate genes. For instance one macro can merge data from a source table containing annotations directly into the target table containing the candidate genes. Another macro can link each candidate to the associated entry of a web-based database, thus enabling the Internet browser to display specific information about the candidate with a single click. Furthermore, different lists of candidates can be compared together, or with members of a gene family for instance, allowing the immediate identification of items that are shared between the lists. The candidates can also be further characterized by their mRNA accumulation patterns in other, non-selecting conditions. In that case, adding color backgrounds to create a heatmap using a dedicated macro available together with FiRe can emphasize the visual impact of the ratios calculated. Importantly, these stand-alone tools can also be used separately from FiRe and outside of the context of microarray analysis.

FiRe will be attractive to eager biologists discouraged by the complexity or price of specialized software, and who eventually resort to excruciating manual explorations of their microarray data. FiRe was designed to be user-friendly and comes with full instructions at the website http://www.unifr.ch/plantbio/FiRe/main.html. It provides the opening steps to the exploration of novel gene functions and should stimulate the formulation of new working hypotheses.

Acknowledgements

The financial support of the Swiss National Science Foundation (Grant 3100A0-104224) is gratefully acknowledged. ApplimathFRI is an applied mathematics group including Richard Baltensperger, Thomas Fournier, Jérôme Pasquier, Marc-Adrien Schnetzer and Jean-Pierre Gabriel. We thank Antony Buchala for critical comments on the manuscript.

References

1 Breitling, R. *et al.* (2004) Rank products: a simple, yet powerful, new method to detect differentially regulated genes in replicated microarray experiments. *FEBS Lett.* 573, 83–92

2 Hsiao, A. *et al*. (2005) VAMPIRE microarray suite: a web-based platform for the interpretation of gene expression data. *Nucleic Acids Res.* 33, W627–W632
3 Tusher, V.G. *et al*. (2001) Significance analysis of microarrays applied to the ionizing radiation response. *Proc. Natl. Acad. Sci. U. S. A.* 98, 5116–5121
4 Vaquerizas, J.M. *et al*. (2005) GEPAS, an experiment-oriented pipeline for the analysis of microarray gene expression data. *Nucleic Acids Res.* 33, W616–W620
5 Ramonell, K. *et al*. (2005) Loss-of-function mutations in chitin responsive genes show increased susceptibility to the powdery mildew pathogen *Erysiphe cichoracearum*. *Plant Physiol.* 138, 1027–1036
6 Gachon, C.M. *et al*. (2005) Transcriptional co-regulation of secondary metabolism enzymes in *Arabidopsis*: functional and evolutionary implications. *Plant Mol. Biol.* 58, 229–245
7 Brown, D.M. *et al*. (2005) Identification of novel genes in *Arabidopsis* involved in secondary cell wall formation using expression profiling and reverse genetics. *Plant Cell* 17, 2281–2295
8 Persson, S. *et al*. (2005) Identification of genes required for cellulose synthesis by regression analysis of public microarray data sets. *Proc. Natl. Acad. Sci. U. S. A.* 102, 8633–8638
9 Leung, Y.F. and Cavalieri, D. (2003) Fundamentals of cDNA microarray data analysis. *Trends Genet.* 19, 649–659

http://doc.rero.ch