BIOINFORMATICS APPLICATIONS NOTE Vol. 22 no. 4 2006, pages 509–511
doi:10.1093/bioinformatics/btk007

Gene expression

eQTL Explorer: integrated mining of combined genetic linkage and expression experiments

Michael Mueller, Anuj Goel, Manjula Thimma, Nicholas J Dickens, Timothy J. Aitman* and Jonathan Mangion*

MRC Clinical Sciences Centre, Imperial College, London W12 0NN, UK

Received on September 30, 2005; revised on November 23, 2005; accepted on December 12, 2005
Advance Access publication December 15 2005
Associate Editor: Thomas Lengauer

ABSTRACT

Summary: The development of computational resources to visualize and explore data from combined genome-wide expression and linkage studies is essential for the development of testable hypotheses. eQTL Explorer stores expression profiles, linkage data and information from external sources in a relational database and enables simultaneous visualization and intuitive interpretation of the combined data via a Java graphical interface. eQTL Explorer provides a new and powerful tool to interrogate these very large and complex datasets.

Availability: The application is freely available for non-commercial research.

Contact: jon.mangion@csc.mrc.ac.uk

Supplementary Information: Documentation, source code and a demonstration version are available at www.bioinformatics.ic.ac.uk/eqtl/

INTRODUCTION

The combined application of genome-wide expression profiling from microarray experiments with genetic linkage analysis has been used in a variety of species, including humans and rodents, to help elucidate the molecular basis of complex disease (reviewed in Gibson and Weir, 2005). This approach enables the mapping of hundreds of expression quantitative trait loci (eQTLs), which are primary genomic control points that influence levels of gene expression.

Cis-acting eQTLs are caused by genomic sequence variants that reside within or close to the gene being regulated and are attractive candidate genes for (patho)physiological QTLs (pQTLs) mapped to the same location (Hubner *et al*., 2005). *Trans*-regulated eQTLs reflect differences in remotely regulated gene expression and often occur in clusters, suggesting co-ordinated regulation of many genes by a single 'master regulator'. It is probable that master regulators of gene expression are key control points in gene networks whose dysregulation leads to complex whole-body phenotypes including disease states (Yvert *et al*., 2003).

Here we present eQTL Explorer, an application that enables integrated visualization and mining of results from genome-wide linkage analyses and expression profiling. eQTL Explorer complements already available tools such as WebQTL (Wang *et al*., 2003) or QTL Express (Seaton *et al*., 2002), but provides further capabilities by focusing on the visualization of genetic mapping results and their integration with external information sources. eQTL Explorer is unique in that it allows eQTL results across the whole genome from multiple array experiments to be displayed alongside pQTLs mapped to the genome (Fig. 1A).

*To whom correspondence should be addressed.

IMPLEMENTATION AND VISUALIZATION

eQTL Explorer comprises a relational database, to store and manage expression, linkage and external data, and a Java interface. The application was designed to interrogate linkage analysis results carried out with QTL reaper, the batch-oriented version of WebQTL (Wang *et al*., 2003), for a study (Hubner *et al*., 2005) carried out in the BXH/HXB panel of recombinant inbred rat strains (Pravenec *et al*., 1989) using the Affymetrix microarray platform. However, it is generically applicable to any similar study.

Upon import into the database eQTLs are generated by amalgamating linkages of expression phenotypes to multiple tightly linked markers, using an algorithm described elsewhere (Hubner *et al*., 2005). The software also determines and indicates whether the eQTLs are *cis*- or *trans*-acting.

Exploration of the data starts with a 'birds-eye' view (Fig. 1A) showing the distribution of eQTLs across the genome alongside the location of known physiological QTLs (pQTLs). By default, eQTLs are plotted according to the location of the genetic marker at the peak of linkage and can be filtered according to whether they are *cis*- or *trans*-acting. Displayed eQTLs can be filtered by level of significance by altering the *P*-value. Less stringent *P*-values result in many datapoints being plotted and putative master regulators of expression are readily identifiable as clusters of co-localized eQTLs.

INTEGRATION WITH EXTERNAL DATA SOURCES

Individual chromosomes can be displayed in a separate view (Fig. 1B), which provides options to browse, zoom and export data for detailed analysis. Upon 'mousing-over' an eQTL or pQTL information relating to its genetic markers or probeset is displayed as tool tip text. A pop-up menu provides access to annotations and cross-references to external data sources including Ensembl (http://www.ensembl.org/), the Rat Genome Database (http://rgd.mcw.edu/) and NCBI Entrez (http://www.ncbi.nih.gov/Entrez/).

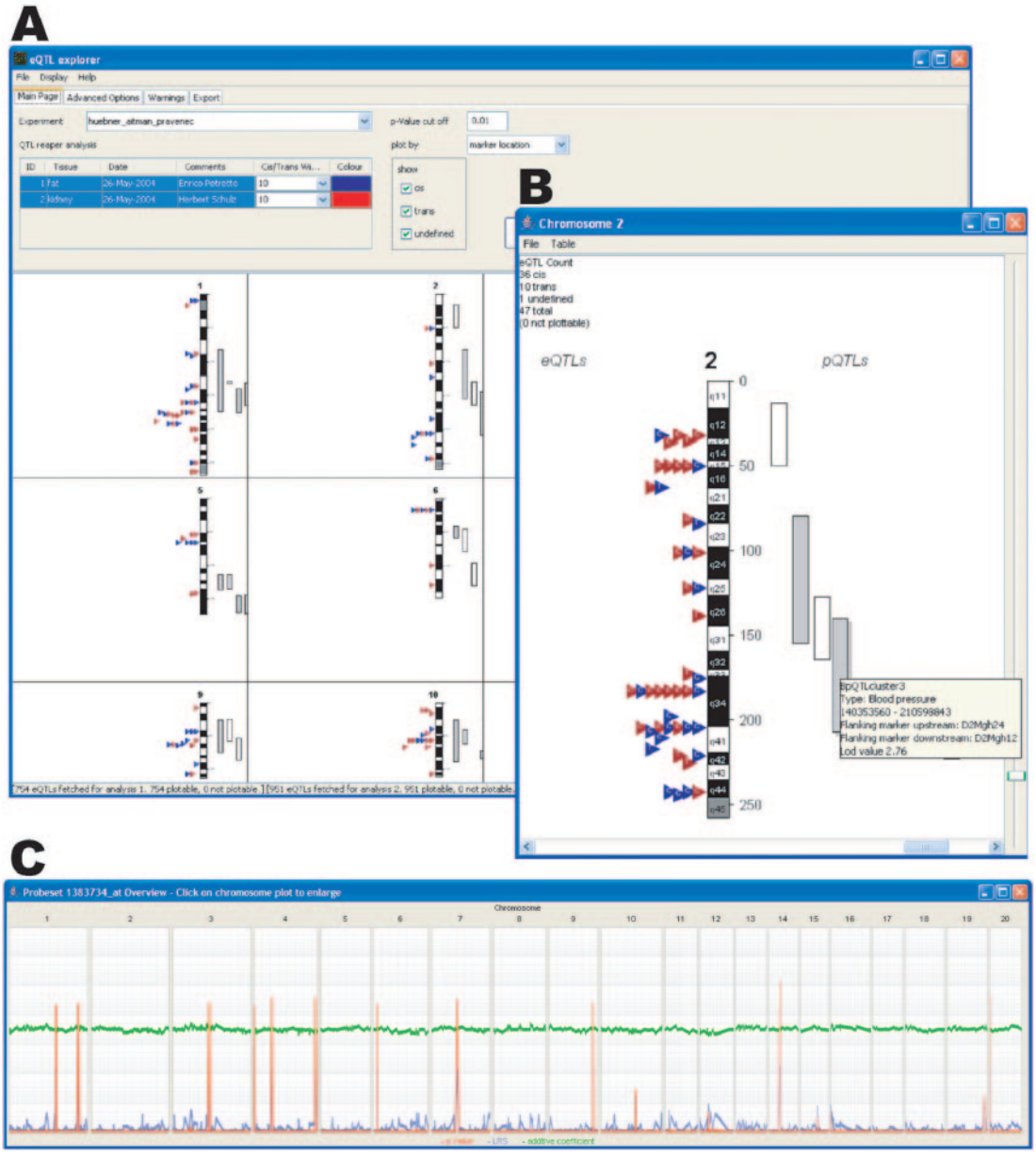

Fig. 1. The eQTL Explorer graphical interface. (**A**) The main window provides a view of the selected exemplar datasets, from two tissues, showing the distribution of eQTLs across the genome (coloured arrow heads) together with the location of pQTLs (vertical bars). The genome display can be adjusted from one to five chromosomes in width. (**B**) The chromosome view provides access to eQTL and pQTL annotations and export options. (**C**) A genome-wide plot of the linkage for expression levels of one probeset.

Advanced options include the ability to select or highlight eQTLs associated with particular probesets. Users can filter the data to display only eQTLs that are located within pQTLs. Information relating to the displayed eQTLs and database cross-references can be exported in tabular format. The genome-wide QTL reaper results for a particular probeset including *P*-value, likelihood ratio statistics and additive coefficient can be viewed as an *XY* plot (Fig. 1C).

In conclusion, we have developed a tool that will greatly aid researchers to interrogate and interpret the massive datasets that are produced by combined microarray and linkage experiments. eQTL Explorer unlocks the data and allows the generation of biological hypotheses about the relationship between eQTLs and pQTLs at the genome level and assists in prioritizing further investigations.

ACKNOWLEDGEMENTS

We thank Fiona Maciver, Michal Pravenec, Norbert Hübner and the Bioinformatics Support Service, Imperial College for their contributions. This work was funded by the MRC Clinical Sciences Centre

and the Wellcome Trust programme for Cardiovascular Functional Genomics.

Conflict of Interest: none declared.

REFERENCES

Gibson,G. and Weir,B. (2005) The quantitative genetics of transcription. *Trends Genet*., **21**, 616–623.

Hubner,N. *et al*. (2005) Integrated transcriptional profiling and linkage analysis for identification of genes underlying disease. *Nat. Genet*., **37**, 243–253.

Pravenec,M. *et al*. (1989) An analysis of spontaneous hypertension in spontaneously hypertensive rats by means of new recombinant inbred strains. *J. Hypertens*., **7**, 217–221.

Seaton,G. *et al*. (2002) QTL Express: mapping quantitative trait loci in simple and complex pedigrees. *Bioinformatics*, **18**, 339–340.

Wang,J. *et al*. (2003) WebQTL: web-based complex trait analysis. *Neuroinformatics*, **1**, 299–308.

Yvert,G. *et al*. (2003) *Trans*-acting regulatory variation in *Saccharomyces Cerevisiae* and the role of transcription factors. *Nat. Genet*., **35**, 57–64.