Genetics and population analysis

xQTL workbench: a scalable web environment for multi-level QTL analysis

Danny Arends1,†, K. Joeri van der Velde1,†, Pjotr Prins1,2, Karl W. Broman3, Steffen Möller4, Ritsert C. Jansen1 and Morris A. Swertz1,5,6,*

1Groningen Bioinformatics Centre, University of Groningen, Groningen, 2Laboratory of Nematology, Wageningen University, Wageningen, 3Department of Biostatistics and Medical Informatics, University of Wisconsin-Madison, Madison, WI, USA, 4Institut für Neuro- und Bioinformatik, Universität zu Lübeck, 5Genomics Coordination Centre, University Medical Centre Groningen, University of Groningen, The Netherlands and 6EMBL-EBI, the European Bioinformatics Institute, Hinxton, UK

Associate Editor: Jeffrey Barrett


Advance Access publication January 3, 2012

BIOINFORMATICS APPLICATIONS NOTE

Vol. 28 no. 7 2012, pages 1042–1044
doi:10.1093/bioinformatics/bts049

ABSTRACT
Summary: xQTL workbench is a scalable web platform for the mapping of quantitative trait loci (QTLs) at multiple levels: for example gene expression (eQTL), protein abundance (pQTL), metabolite abundance (mQTL) and phenotype (phQTL) data. Popular QTL mapping methods for model organism and human populations are accessible via the web user interface. Large calculations scale easily to multi-core computers, clusters and Cloud. All data involved can be uploaded and queried online: markers, genotypes, microarrays, NGS, LC-MS, GC-MS, NMR, etc. When new data types come available, xQTL workbench is quickly customized using the Molgenis software generator.

Availability: xQTL workbench runs on all common platforms, including Linux, Mac OS X and Windows. An online demo system, installation guide, tutorials, software and source code are available under the LGPL3 license from http://www.xqtl.org.

Contact: m.a.swertz@rug.nl

Received on September 30, 2011; revised on December 19, 2011; accepted on January 20, 2012

1 INTRODUCTION

Modern high-throughput technologies generate large amounts of genomic, transcriptomic, proteomic and metabolomic data. However, existing open source web-based tools for QTL analysis, such as webQTL (Wang et al., 2004) and QTLNetwork (Yang et al., 2008), are not easily extendable to different settings and computationally scalable for whole genome analyses. xQTL workbench makes it easy to analyse large and complex datasets using state-of-the-art QTL mapping tools and to apply these methods to millions of phenotypes using parallelized ‘Big Data’ solutions.

To whom correspondence should be addressed.

†The authors wish it to be known that, in their option, the first two authors should be regarded as joint First Authors.

2 FEATURES

xQTL workbench provides visualization of QTL profiles, single and multiple QTL mapping methods, easy addition of new QTL analyses, scalable data management and analysis tracking.

2.1 Explore QTL profiles

Through the web interface, users can explore mapped QTLs, and underlying information by viewing QTL plots in an interactive scrollable and zoomable window. xQTL workbench has support for other common image formats, such as PNG, JPG, SVG and embedded postscript; useful for publishing scientific results online, and on paper. From the output, main database identifiers, such as gene, protein and/or metabolite identifiers are automatically linked-out to matching external web pages of public database such as NCBI, KEGG and Wormbase.

2.2 Single and multiple QTL mapping

xQTL workbench—one kind of QTLNetwork (Arends et al., 2010; Broman et al., 2009) in a web-based analysis framework offering all important QTL mapping routines, such as the EM algorithm, imputation, Haley-Knott regression, the extended Haley-Knott method, marker regression and Multiple QTL mapping. In addition, xQTL workbench includes RQTLbim, a library that provides a Bayesian model selection approach for mapping multiple interacting QTL (Yandell et al., 2009) and Plink, a library for association QTL mapping on single nucleotide polymorphisms (SNP) in natural populations (Purcell et al., 2007).

2.3 Add new analysis tools

xQTL workbench supports flexible adding of more QTL analysis software: any R-based, or command-line tool, can be plugged in. All analysis results are uploaded, stored and tracked in the xQTL workbench database through an R-API. When new tools are added, they can build on the high-level multi-core computer, cluster and Cloud management functions, based on TORQUE/OpenPBS and BioNode (Prins et al., 2011). xQTL workbench can be made part of a larger analysis pipeline using interfaces to R, Excel, REST and SOAP web services and Galaxy (Krocks et al., 2010).

© The Author(s) 2012. Published by Oxford University Press. This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (http://creativecommons.org/licenses/by-nc/3.0), which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.
xQTL workbench

Fig. 1. Screenshot of xQTL workbench with all features enabled; (1) import phenotype, genotype and generic map data, examples are given per import type; (2) search through the whole database, explore and browse your data using molgenis generated web-interfaces; (3) run R/qtl QTL mapping, the general plugin allows users to perform not only QTL mapping but also other analyze; (4) use default (or custom) plugins to explore results (e.g. Heatmaps, QTL profiles); (5) add new tools to the workbench (for Bioinformaticians); (6) user management and access control of the system (Only for admins); (7) expert settings can be altered in the admin tab (Only for admins); (8) connect/share data using generated API’s to R statistics, REST/JSON, SOAP.
2.4 Track analysis and monitor performance

When a new analysis protocol or R script is defined, this protocol can easily be applied to new data. Also, xQTL workbench keeps track of history. Re-use of analysis protocols can be done in an automated fashion. Previous analyses can be rerun without resetting parameters. xQTL workbench provides an online overview of past analyses e.g. which analyses were performed, by who, when and display settings applied.

2.5 Scalable data management

xQTL workbench has a consistency checking database based on XGAP specification (Swertz et al. 2010b), user interfaces to manage and query genotype and phenotype datasets and support for various database back-ends including HSQIL (standalone) and MySQL. Phenotype, genotype and genetic map data can be imported as text (TXT), comma separated (CSV) and Excel files. xQTL workbench handles and stores large data in a new and efficient binary edition of the XGAP format, named XGAPhim (extension .xbin), documented online. Such binary formats are essential when handling, storing and transporting multi-Gigabyte datasets.

2.6 Customizable to research needs


ACKNOWLEDGEMENTS

We thank Konrad Zych for Figure 1.

Funding: National Institutes of Health (GM074244 to KB); Netherlands Organisation for Scientific Research (NWO)/TTI Green Genetics (1CC029RP to F.P.); NWO (Rubicon 825.09.008 to M.A.S), Center for BioSystems Genomics (CBSG), Netherlands Consortium of Systems Biology (NCSB) (to D.A.), Netherlands Bioinformatics Center (NBI) and the European Regional Development Fund and NWO (to M.A.S.); and EU-FP7 Projects PANACEA (222936 to K.J.v.d.v.) and EURATRANS (214504 to R.C.J.).

Conflict of Interest: none declared.

REFERENCES