



APPLICATION NOTE

ezQTL: A Web Platform for Interactive Visualization and Colocalization of QTLs and GWAS Loci



Tongwu Zhang, Alyssa Klein, Jian Sang, Jiyeon Choi^{*,#}, Kevin M. Brown^{*,#}

Division of Cancer Epidemiology and Genetics, National Cancer Institute, Bethesda, MD 20892, USA

Received 17 September 2021; revised 12 March 2022; accepted 8 May 2022

Available online 25 May 2022

Handled by Zhongming Zhao

KEYWORDS

Genome-wide association study;
Expression quantitative trait locus;
Colocalization;
Linkage disequilibrium;
Visualization

Abstract Genome-wide association studies (GWAS) have identified thousands of genomic loci associated with complex diseases and traits, including cancer. The vast majority of common trait-associated variants identified via GWAS fall in non-coding regions of the genome, posing a challenge in elucidating the causal variants, genes, and mechanisms involved. **Expression quantitative trait locus** (eQTL) and other molecular QTL studies have been valuable resources in identifying candidate causal genes from GWAS loci through statistical **colocalization** methods. While QTL colocalization is becoming a standard analysis in post-GWAS investigation, an easy web tool for users to perform formal colocalization analyses with either user-provided or public GWAS and eQTL datasets has been lacking. Here, we present ezQTL, a web-based bioinformatic application to interactively visualize and analyze genetic association data such as GWAS loci and molecular QTLs under different **linkage disequilibrium** (LD) patterns (1000 Genomes Project, UK Biobank, or user-provided data). This application allows users to perform data quality control for variants matched between different datasets, LD **visualization**, and two-trait colocalization analyses using two state-of-the-art methodologies (eCAVIAR and HyPrColoc), including batch processing. ezQTL is a free and publicly available cross-platform web tool, which can be accessed online at <https://analysistools.cancer.gov/ezqtl>.

Introduction

Genome-wide association studies (GWAS) have proven powerful in identifying common genetic variants and loci associated with many complex diseases and traits, including cancer susceptibility [1]. Lead single nucleotide polymorphisms (SNPs) identified via GWAS are not necessarily causal variants themselves; rather these variants often tag a larger set of many

* Corresponding authors.

E-mail: kevin.brown3@nih.gov (Brown KM), jiyeon.choi2@nih.gov (Choi J).

Equal contribution.

Peer review under responsibility of Beijing Institute of Genomics, Chinese Academy of Sciences / China National Center for Bioinformation and Genetics Society of China.

<https://doi.org/10.1016/j.gpb.2022.05.004>

1672-0229 © 2022 The Authors. Published by Elsevier B.V. and Science Press on behalf of Beijing Institute of Genomics, Chinese Academy of Sciences / China National Center for Bioinformation and Genetics Society of China.

This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

variants in high linkage disequilibrium (LD), one or more of which may be causal variants underlying the trait [2]. Most of these trait-associated variants are non-coding, and for a large proportion of loci, causal variants are thought to be *cis*-regulatory. Expression quantitative trait locus (eQTL) studies have shown that common *cis*-regulatory variants often regulate many genes other than the nearest one [3], meaning that many distant genes must be considered as plausible candidate genes. Sifting through and functionally characterizing a large number of statistically-correlated candidate variants, as well as in many cases identifying the likely causal gene(s) affecting disease risk at GWAS loci, remain significant challenges in the field. These challenges hinder the translation of GWAS findings to a better understanding of the biological processes underlying complex traits including cancer susceptibility.

Most common causal risk variants are hypothesized to function by affecting gene expression, or alternatively, via potentially affecting patterns of splicing, DNA methylation, and/or expression levels of microRNAs or other non-coding RNAs. Genetic association tests of these various molecular quantitative trait loci (QTLs) are commonly used for prioritization of common variants within GWAS loci for functional study, as well as identification of the gene(s) likely to underlie risk at these loci [4–6]. Numerous eQTL datasets derived from normal human tissues have been made publicly available, such as the Genotype-Tissue Expression (GTEx) project [7], which has become a key resource for investigating tissue-specific QTLs underlying complex traits. Colocalization analysis has been proposed as an important step to test whether complex trait associations (*e.g.*, cancer risk) and molecular QTLs (such as those for gene expression, eQTLs) share common causal genetic variants for both signals, which will aid biological insight following GWAS. Statistical colocalization methods have been developed to assess the overlap of genetic association signals across multiple related traits (such as molecular QTLs and GWAS loci) [8]. Rather than focusing on lead variants, a colocalization analysis compares the distribution of summary statistics from two association signals and accounts for LD, which reduces the false positive discoveries by using multiple variants [9–11]. However, current web-based applications, such as LocusZoom [12], do not provide visualization of GWAS locus–QTL colocalization, whereas LocusCompare [13] provides only limited colocalization statistics. Although there is an abundance of implementation of colocalization algorithms in R and tools for visualization, running these various programs may require additional quality control (QC) of user-supplied data and different data formatting, which can be challenging for a wet-lab researcher with little background in bioinformatics. There is also a critical need for a colocalization platform with large available QTL datasets including different types and tissues, and LD datasets including different populations.

To address this problem, we developed ezQTL to interactively visualize the genetic associations and perform colocalization between two traits. The web-based modules (Locus QC, Locus LD, Locus Alignment, Locus Colocalization, Locus Table, Locus Quantification, and Locus Download) utilize GWAS summary statistics, QTL association data, and LD matrix data to perform a comprehensive colocalization analy-

sis, culminating in a QC report and interactive visualization. ezQTL bridges current gaps of existing colocalization analysis tools by integrating user-supplied as well as publicly available GWAS and QTL resources, LD reference datasets, different colocalization algorithms, and interactive visualization, through a user-friendly web interface.

Implementation

The ezQTL web application includes seven modules (**Figure 1**): Locus QC, Locus LD, Locus Alignment, Locus Colocalization, Locus Table, Locus Quantification, and Locus Download. All seven ezQTL modules are written in R and run on a virtual machine using the UNIX operating system. The web tool is primarily written in Javascript and R. All web content is programmed in HTML5 for cross platform versatility. ezQTL is built using the following packages for colocalization analyses: eCAVIAR (<https://github.com/fhormoz/caviar>) and HyPrColoc (<https://github.com/jrs95/hyprcoloc>). It also uses the IntAssoPlot for LD visualization (<https://github.com/whweve/IntAssoPlot>). The Plotly R graphing library is used to generate interactive plots. All datasets, QC files, and results, including plots and tables in each analysis module, can be downloaded using the Locus Download module. For genomic data privacy, the input data in ezQTL is encrypted and secured during the data transition, calculation, and temporary storage. All input data and results for each query or submission will be deleted automatically after seven days.

User- and publicly-supplied input data

Currently, for colocalization, ezQTL requires QTL, GWAS, and LD matrix data that can be supplied by the user. However, ezQTL also hosts a large collection of publicly available QTL, GWAS, and LD datasets; currently 594 datasets are collectively listed on the searchable Public Data Source tab. We provide public datasets that are pre-processed into the standard ezQTL input format and example scripts to assist users in preparing user-supplied datasets into the correct input format. GWAS summary statistics included in ezQTL can be downloaded at <https://github.com/mikegloudemans/gwas-download> and <https://www.ebi.ac.uk/gwas/home>. The GTEx datasets for both eQTLs and splicing QTLs (sQTLs) can be accessed via <https://gtexportal.org/home/datasets>. The link used to download other GWAS or QTL datasets from individual studies can be accessed through the ezQTL website. 1000 Genomes data used to generate the LD matrix are available at <https://ftp.1000genomes.ebi.ac.uk/vol1/ftp/release/20130502/>. The LD matrix from the UK Biobank dataset can be accessed at https://alkesgroup.broadinstitute.org/UKBB_LD/.

User-supplied QTL summary statistics can be provided for any given locus to be tested (*e.g.*, by default variants of ± 75 kb of the lead SNP can be used) in the same format as the output from commonly used QTL data analysis tools (*e.g.*, FastQTL [14] and Matrix eQTL [15]); the fields required by ezQTL are matched by the column header names of the input file, and a warning is provided when no match is found. ezQTL also provides commonly used QTL datasets, including both sQTLs and eQTLs from GTEx in two genome builds

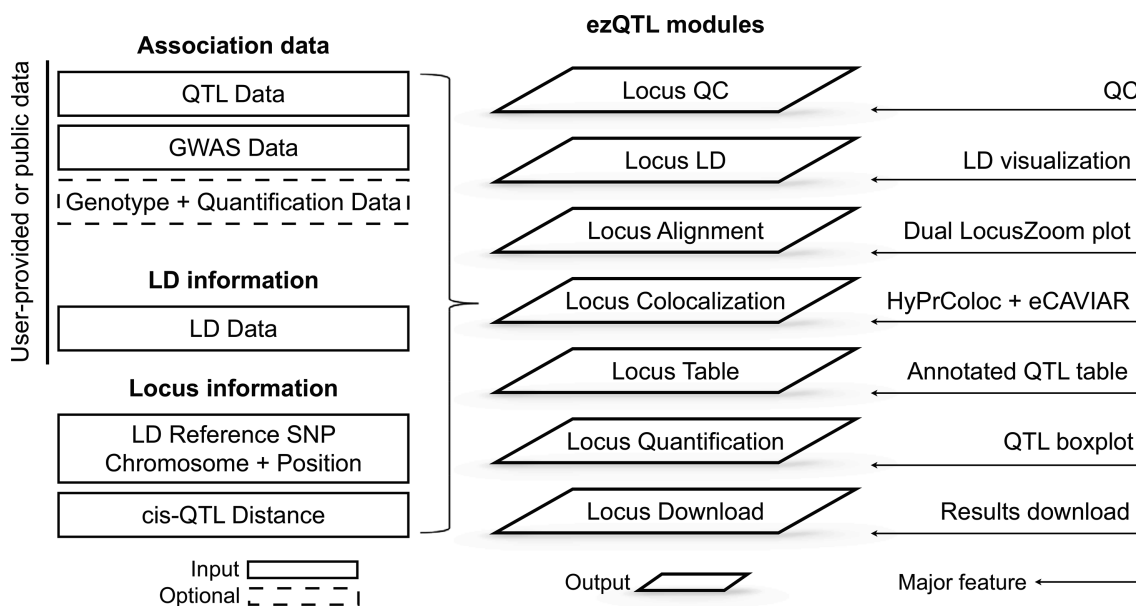


Figure 1 The architecture of the ezQTL analysis platform

The overall structure of the ezQTL analysis platform including the input data types, analysis modules, and major features is shown. QTL, quantitative trait locus; SNP, single nucleotide polymorphism; QC, quality control; LD, linkage disequilibrium; GWAS, genome-wide association study.

(GRCh37 and GRCh38) [16] and QTLs from additional tissue types [4,17–23].

For GWAS summary data, ezQTL similarly accepts user-supplied summary data in a general format for each locus (*e.g.*, variants of ± 75 kb of the lead SNP), matched by column headers for the required fields. We also provide an example script to transfer GWAS summary statistics from National Human Genome Research Institute (NHGRI) GWAS Catalog [24] into ezQTL GWAS data format through ezQTL GitHub page. Currently, ezQTL has pre-loaded GWAS summary statistics previously collected by LocusCompare [13] and most of the publicly available cancer-related GWAS summary data downloaded from the NHGRI GWAS Catalog [24] or the study websites. At this time, there are more than 420 phenotypes included in the set of pre-supplied GWAS summary data.

For LD reference data, ezQTL will instantly calculate the LD matrix for any locus from any combination of 1000 Genomes (Phase 3) populations [25] for both GRCh37 and GRCh38 genome builds, or alternatively provides a pre-calculated LD matrix from the UK Biobank European populations [26] (GRCh37 only). Users may also supply their own LD matrix data generated by emeraLD [27].

In addition to colocalization analyses using these three data types, users can provide raw QTL data (*e.g.*, matched gene expression or other QTL trait levels and individual genotypes) for extended QTL visualization functions. We anticipate future implementations of this visualization feature extended to pre-loaded public eQTL data such as those from GTEx.

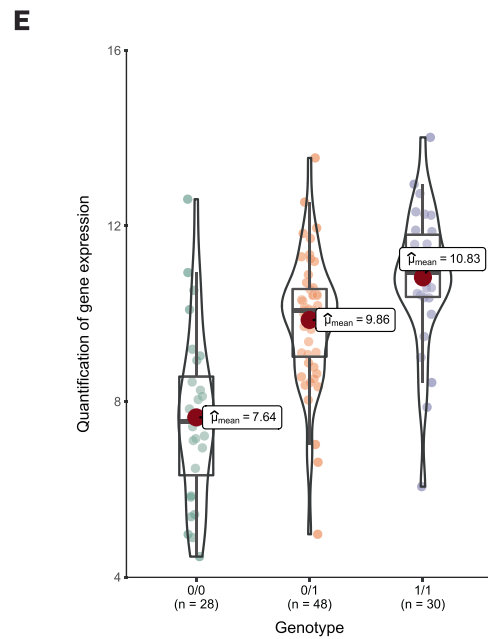
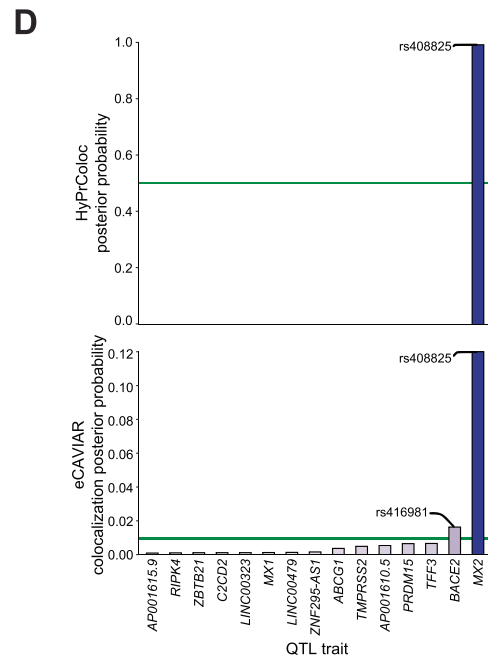
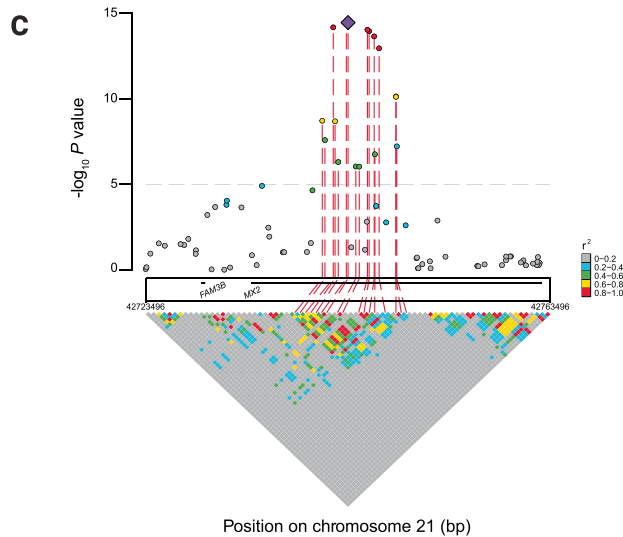
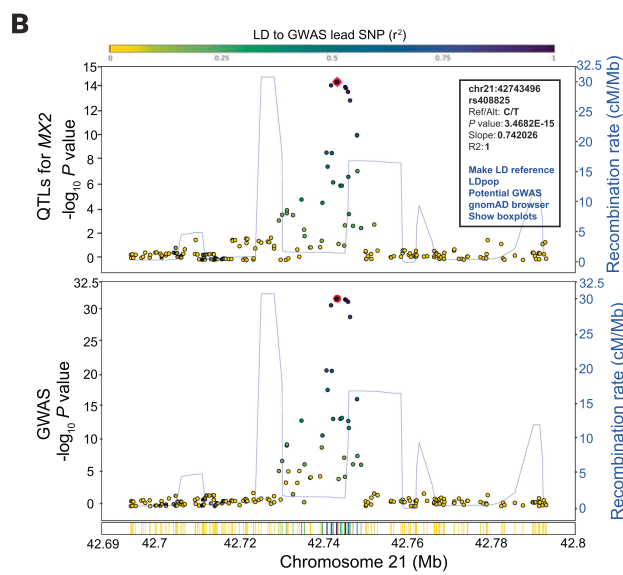
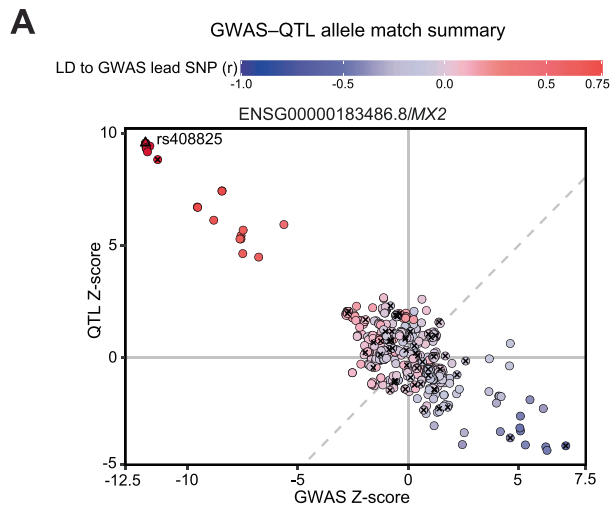
Features of ezQTL modules

ezQTL takes as minimal input both GWAS and QTL data for colocalization functions and any one of GWAS, QTL, or LD for visualization functions. The detailed relationships between

module functions and input datasets are included in the documentation of ezQTL. The Locus QC module performs data QC for user-supplied GWAS, QTL, and/or LD data (such as for data formatting, allele-matching, variant filtering, and suggesting a reference variant), generates a variant level summary report for each dataset, and reports overlapping SNPs. This module also provides visualizations for each input dataset and highlights potential issues (such as insufficient overlapping variants) for subsequent colocalization analyses. The post-QC data generated by Locus QC will be used for all other modules.

Locus LD and Locus Alignment modules both visualize QTL and GWAS data together with LD information for the locus to provide a quick visual overview of the colocalization pattern. Locus LD generates Manhattan plots for either QTL or GWAS P values along with local LD patterns using the IntAssoPlot R package [28]. The Locus Alignment module simultaneously and interactively visualizes association P values and LD patterns (recombination rate) for both GWAS and QTL datasets using two LocusZoom plots for a given gene or probe in a locus of interest. An example is provided for the melanoma GWAS locus at Chr21q22.3 and melanocyte-specific eQTL data for *MX2* [4,29] (Figure 2). In addition, clicking on any given variant on the LocusZoom plot can directly connect to other functions, such as setting that variant as the LD reference, creating a QTL boxplot (if users provide individual-level QTL data), or linking to multiple external databases (LDpop [30], GWAS Catalog [24], gnomAD [31]). A P value correlation plot between these two LocusZoom plots, similar to that provided by LocusCompare, is generated to visualize colocalization.

The Locus Colocalization function performs a colocalization analysis between GWAS loci and molecular QTLs (or any two traits) within a locus using two different algorithms, eCAVIAR [9] and HyPrColoc [8] (Figure 2). By default, data



used for colocalization are selected from a genomic window centered on the variant with the smallest GWAS P value within the region uploaded. Re-centering on a different LD reference SNP can be specified either at the time of uploading the data, or alternatively by entering or selecting a specific variant on the Locus Alignment page following data upload. HyPrColoc [8] is a computationally efficient method based on the Bayesian algorithm to estimate the posterior probability that a variant is causal in both GWAS and QTL studies. This approach extends the underlying methodology of COLOC [32] and MOLOC [33]. As implemented in ezQTL, HyPrColoc uses all genetic variants within a user-specified window (specified on the data input panel, the default is ± 75 kb centered over the lead GWAS or a user-specified reference variant). For eCAVIAR, the colocalization test is performed by multiplying posterior causal probabilities from the QTL and GWAS studies, to generate a colocalization posterior probability (CLPP), as described by Hormozdiari and his colleagues [9]. eCAVIAR colocalization as implemented in ezQTL is performed using all genetic variants within the user-specified window used by HyPrColoc, as well as the set of 100 variants surrounding the lead GWAS variant (50 variants on each side as well as the lead variant itself). Both eCAVIAR and HyPrColoc results are presented and compared as exportable barplots and tables including all the molecular traits (such as eQTL genes) tested in each locus.

We recognize that the colocalization results for two traits in the same locus might be different based on different colocalization approaches. HyPrColoc is a deterministic Bayesian colocalization algorithm that makes the assumption of a single causal variant in a given trait, similar to other existing methods, such as COLOC [32] and MOLOC [33]. Although HyPrColoc can be used for multiple-trait colocalization, the HyPrColoc analysis implemented in ezQTL is set for two-trait analysis at a time and does not use LD information (except for two traits with overlapping samples) because of the single causal variant assumption. eCAVIAR is a probabilistic model that could consider multiple causal variants in a locus, and the eCAVIAR analysis provided by ezQTL allows up to two casuals using the LD information that users choose for the analysis. Some loci may harbor more than one causal variant for any given trait, and the colocalization results under the single causal variant assumption would not hold when secondary causal variants explain as much trait variation as the identified common causal variant [8]. On the other hand, eCAVIAR was designed to account for allelic heterogeneity and LD, under the assumption that the number of individuals and the LD are the same between GWAS and QTL popula-

tions [9]. However, misspecification of the LD structure when aiming to fine-map more than a single causal variant can lead to major biases in the results. If an accurate LD reference panel of a large sample size that matches both GWAS and QTL populations is not available, a single causal variant model without using LD could be more reliable [8]. In addition, the number of variants selected for a colocalization analysis is different between eCAVIAR and HyPrColoc, as described in the Locus Colocalization module. Thus, we recommend users to consider these differences in the statistical models and assumptions as described in the original publications of eCAVIAR [9] and HyPrColoc [8] when reporting and interpreting the results.

Locus Quantification provides visualizations (violin, distribution, and pairwise correlation plots) of optional quantification data used in the QTL testing of the locus of interest. The module requires user-provided individual-level genotypes and molecular trait data that have been used for calculating QTLs (e.g., gene expression or DNA methylation); in the current version, QTLs provided by ezQTL are pre-computed summary statistics and do not use individual-level genotypes or trait data. These data allow generation of violin plots for each trait, e.g., in the Locus Quantification module), as well as boxplots for QTLs that can be accessed through the Locus Alignment module.

Finally, Locus Table provides a sortable table of user-provided QTLs for any given trait in a region. QTL results are annotated with LD to any given index SNP and link to multiple external databases. All datasets, QC files, and results, including plots and tables in each analysis module, can be exported using the Locus Download module.

Comparison with other tools

In comparison to similar tools (such as LocusZoom [12], LocusCompare [13], LocusFocus [34], GTEx Portal [35], and eQTLplot [36]), ezQTL contains several new features and functions for the integrated analysis of GWAS loci and QTLs (**Table 1**). ezQTL provides interactive dual-LocusZoom plots, which allows users to explore and integrate the association data between GWAS and QTL with availability to recalculate LD information based on the selected reference variant. ezQTL performs variant-level QC before the colocalization analysis for all user-supplied and pre-provided datasets, and also generates reports and plots for future investigation. For example, matching allelic directions between GWAS and QTL datasets is essential for the colocalization analysis, and ezQTL will report any flipped or unmatched variants to users



Figure 2 Example of an ezQTL colocalization analysis between GWAS loci and eQTLs in a melanoma GWAS locus

A. Correlation plot of Z-scores between QTLs and GWAS loci from the Locus QC module for identification of potentially mismatched alleles. Variants marked by “×” are C/G or A/T SNPs. **B.** A dual LocusZoom plot with eQTL association on top and GWAS association at the bottom from Locus Alignment. The pink diamond in the LocusZoom plot indicates the most significant SNP among the eQTL association, and the red circle highlights the current LD reference SNP. The pop-up window after clicking each SNP shows association information, additional actions, and links. **C.** LD matrix visualization from Locus LD. The purple diamond and red dashed line indicate the reference SNP. **D.** Colocalization analysis by both eCAVIAR and HyPrColoc from Locus Colocalization. **E.** QTL boxplot between genotype of reference SNP and gene quantification from Locus Quantification. eQTL, expression quantitative trait locus.

Table 1 Comparisons of QTL visualization and analysis tools

	ezQTL	LocusZoom	LocusCompare	LocusFocus	GTEx Portal	eQTLpLot
LocusZoom plot	✔	✔	✔	✔	✔	✔
Interactive visualization of variants within plots	✔	×	×	×	✔	×
Data QC (e.g., allele/strand matching)	✔	×	×	×	✔	×
Colocalization tools	eCAVIAR; HyPerColoc; correlation	×	FINEMAP; eCAVIAR	COLOC2; simple sum test	×	Correlation; Fisher's exact test
Provide LD source	✔	×	×	✔	×	✔
LD matrix visualization	User + public	×	×	×	×	User only
Support user dataset for colocalization analyses	✔	×	×	×	×	×
Host public datasets	✔	×	×	GWAS only	×	×
Genome build	✔	GRCh37 and GRCh38	✔	GRCh37 and GRCh38	GTEx QTL only GRCh37 and GRCh38	GRCh37 and GRCh38
Visualization of QTL association	✔	×	GRCh37	×	✔	×
Multi-locus query	✔	×	×	×	×	×
Reference	This study	[12]	[13]	[34]	[7]	[36]

Note: QTL, quantitative trait locus; GTEx, Genotype-Tissue Expression; QC, quality control; LD, linkage disequilibrium; GWAS, genome-wide association study.

for recalibration. ezQTL hosts a large number of public datasets (including those from different genome builds, different tissue types, and different populations) for each data type and allows users to perform a colocalization analysis using mixed data sources. These public datasets in ezQTL also allow users to perform an exploratory data analysis in any given locus. ezQTL also implements two state-of-the-art methodologies (eCAVIAR and HyPrColoc) for formal colocalization analyses, which represent two distinct statistical approaches with different assumptions of the number of causal variants. In addition, serial visualizations have been implemented in ezQTL, including multiple QC-related plots, *P* value correlation plots, LD matrix visualization, colocalization summary plots, and QTL association plots. For the same dataset, ezQTL is designed to allow users to perform an analysis for multiple loci by inputting different locus information. In addition, ezQTL allows users to submit a job to the ezQTL server and an email with a link to access all of the results will sent to the user.

Conclusion

Here, we present ezQTL as a web tool to interactively visualize and analyze genetic association data and perform colocalization between two traits without much bioinformatics skill. ezQTL is competitive with two state-of-the-art colocalization approaches as well as large public datasets. ezQTL enhances existing colocalization analysis tools for GWAS by integrating user-supplied as well as publicly available GWAS and QTL resources, LD reference datasets, different colocalization algorithms, and interactive visualization through a user-friendly web interface. In summary, ezQTL facilitates mapping disease susceptibility regions and assists researchers in characterizing and prioritizing functional genes and variants based on the genotype–phenotype associations.

Code availability

The source code for ezQTL can be found at the Center for Biomedical Informatics and Information Technology (CBIIT) GitHub page: <https://github.com/CBIIT/nci-webtools-dceg-ezQTL>, and at BioCode at the National Genomics Data Center (NGDC): <https://ngdc.cncb.ac.cn/biocode/tools/BT007295>. It is licensed under the GNU General Public License version 3.0 (GPLv3).

Data availability

ezQTL is a free and publicly available cross-platform web tool which can be accessed online at <https://analysistools.cancer.gov/ezqtl>.

CRedit author statement

Tongwu Zhang: Conceptualization, Software, Resources, Data curation, Writing - original draft, Visualization, Funding acquisition. **Alyssa Klein:** Software, Validation, Data curation. **Jian Sang:** Data curation, Validation. **Jiyeon Choi:** Conceptualization, Validation, Resources, Data curation,

Writing - original draft, Supervision, Funding acquisition. **Kevin M. Brown:** Conceptualization, Validation, Resources, Data curation, Writing - original draft, Supervision, Funding acquisition. All authors have read and approved the final manuscript.

Competing interests

The authors have declared no competing interests.

Acknowledgments

The authors thank Kailing Chen, Kevin Jiang, Phyllip Cho, and Ben Chen from the CBIIT of National Cancer Institute (NCI) for technical assistance and web development, as well as the Laboratory of Translational Genomics and Division of Cancer Epidemiology and Genetics of NCI for valuable input and testing. Support comes from the Intramural Research Program (Grant No. 1ZIACP010201) and the Division of Cancer Epidemiology and Genetics Informatics Tool Challenge of NCI. The content of this publication does not necessarily reflect the views or policies of the US Department of Health and Human Services, nor does mention of trade names, commercial products, or organizations imply endorsement by the US government.

ORCID

ORCID 0000-0003-2124-2706 (Tongwu Zhang)

ORCID 0000-0003-3763-5731 (Alyssa Klein)

ORCID 0000-0003-4953-3417 (Jian Sang)

ORCID 0000-0002-0955-2384 (Jiyeon Choi)

ORCID 0000-0002-8558-6711 (Kevin M. Brown)

References

- [1] Visscher PM, Wray NR, Zhang Q, Sklar P, McCarthy MI, Brown MA, et al. 10 years of GWAS discovery: biology, function, and translation. *Am J Hum Genet* 2017;101:5–22.
- [2] Schaid DJ, Chen W, Larson NB. From genome-wide associations to candidate causal variants by statistical fine-mapping. *Nat Rev Genet* 2018;19:491–504.
- [3] GTEx Consortium. Genetic effects on gene expression across human tissues. *Nature* 2017;550:204–13.
- [4] Zhang T, Choi J, Kovacs MA, Shi J, Xu M, NISC Comparative Sequencing Program, et al. Cell-type-specific eQTL of primary melanocytes facilitates identification of melanoma susceptibility genes. *Genome Res* 2018;28:1621–35.
- [5] Zhang T, Choi J, Dilshat R, Einarsdóttir BÓ, Kovacs MA, Xu M, et al. Cell-type-specific meQTLs extend melanoma GWAS annotation beyond eQTLs and inform melanocyte gene-regulatory mechanisms. *Am J Hum Genet* 2021;108:1631–46.
- [6] Do C, Shearer A, Suzuki M, Terry MB, Gelernter J, Grealley JM, et al. Genetic-epigenetic interactions in *cis*: a major focus in the post-GWAS era. *Genome Biol* 2017;18:120.
- [7] GTEx Consortium. The Genotype-Tissue Expression (GTEx) pilot analysis: multitissue gene regulation in humans. *Science* 2015;348:648–60.
- [8] Foley CN, Staley JR, Breen PG, Sun BB, Kirk PDW, Burgess S, et al. A fast and efficient colocalization algorithm for identifying shared genetic risk factors across multiple traits. *Nat Commun* 2021;12:764.
- [9] Hormozdiari F, van de Bunt M, Segrè AV, Li X, Joo JWJ, Bilow M, et al. Colocalization of GWAS and eQTL signals detects target genes. *Am J Hum Genet* 2016;99:1245–60.
- [10] Nica AC, Montgomery SB, Dimas AS, Stranger BE, Beazley C, Barroso I, et al. Candidate causal regulatory effects by integration of expression QTLs with complex trait genetic associations. *PLoS Genet* 2010;6:e1000895.
- [11] Zhu Z, Zhang F, Hu H, Bakshi A, Robinson MR, Powell JE, et al. Integration of summary data from GWAS and eQTL studies predicts complex trait gene targets. *Nat Genet* 2016;48:481–7.
- [12] Pruim RJ, Welch RP, Sanna S, Teslovich TM, Chines PS, Gliedt TP, et al. LocusZoom: regional visualization of genome-wide association scan results. *Bioinformatics* 2010;26:2336–7.
- [13] Liu B, Gloudeans MJ, Rao AS, Ingelsson E, Montgomery SB. Abundant associations with gene expression complicate GWAS follow-up. *Nat Genet* 2019;51:768–9.
- [14] Ongen H, Buil A, Brown AA, Dermizakis ET, Delaneau O. Fast and efficient QTL mapper for thousands of molecular phenotypes. *Bioinformatics* 2016;32:1479–85.
- [15] Shabalin AA. Matrix eQTL: ultra fast eQTL analysis via large matrix operations. *Bioinformatics* 2012;28:1353–8.
- [16] GTEx Consortium. The GTEx Consortium atlas of genetic regulatory effects across human tissues. *Science* 2020;369:1318–30.
- [17] McRae AF, Marioni RE, Shah S, Yang J, Powell JE, Harris SE, et al. Identification of 55,000 replicated DNA methylation QTL. *Sci Rep* 2018;8:17605.
- [18] Hannon E, Spiers H, Viana J, Pidsley R, Burrage J, Murphy TM, et al. Methylation QTLs in the developing brain and their enrichment in schizophrenia risk loci. *Nat Neurosci* 2016;19:48–54.
- [19] Hannon E, Dempster E, Viana J, Burrage J, Smith AR, Macdonald R, et al. An integrated genetic-epigenetic analysis of schizophrenia: evidence for co-localization of genetic associations and differential DNA methylation. *Genome Biol* 2016;17:176.
- [20] Hannon E, Gorrie-Stone TJ, Smart MC, Burrage J, Hughes A, Bao Y, et al. Leveraging DNA-methylation quantitative-trait loci to characterize the relationship between methylomic variation, gene expression, and complex traits. *Am J Hum Genet* 2018;103:654–65.
- [21] Lappalainen T, Sammeth M, Friedländer MR, 't Hoen PAC, Monlong J, Rivas MA, et al. Transcriptome and genome sequencing uncovers functional variation in humans. *Nature* 2013;501:506–11.
- [22] Qi T, Wu Y, Zeng J, Zhang F, Xue A, Jiang L, et al. Identifying gene targets for brain-related traits using transcriptomic and methylomic data from blood. *Nat Commun* 2018;9:2282.
- [23] Bryois J, Garrett ME, Song L, Safi A, Giusti-Rodríguez P, Johnson GD, et al. Evaluation of chromatin accessibility in prefrontal cortex of individuals with schizophrenia. *Nat Commun* 2018;9:3121.
- [24] Buniello A, MacArthur JAL, Cerezo M, Harris LW, Hayhurst J, Malangone C, et al. The NHGRI-EBI GWAS Catalog of published genome-wide association studies, targeted arrays and summary statistics 2019. *Nucleic Acids Res* 2019;47:D1005–12.
- [25] 1000 Genomes Project Consortium, Auton A, Brooks LD, Durbin RM, Garrison EP, Kang HM, et al. A global reference for human genetic variation. *Nature* 2015;526:68–74.
- [26] Weissbrod O, Hormozdiari F, Benner C, Cui R, Ulirsch J, Gazal S, et al. Functionally informed fine-mapping and polygenic localization of complex trait heritability. *Nat Genet* 2020;52:1355–63.
- [27] Quick C, Fuchsberger C, Taliun D, Abecasis G, Boehnke M, Kang HM. emeraLD: rapid linkage disequilibrium estimation with massive datasets. *Bioinformatics* 2019;35:164–6.
- [28] He F, Ding S, Wang H, Qin F. IntAssoPlot: an R package for integrated visualization of genome-wide association study results with gene structure and linkage disequilibrium matrix. *Front Genet* 2020;11:260.

- [29] Choi J, Zhang T, Vu A, Ablain J, Makowski MM, Colli LM, et al. Massively parallel reporter assays of melanoma risk variants identify *MX2* as a gene promoting melanoma. *Nat Commun* 2020;11:2718.
- [30] Alexander TA, Machiela MJ. LDpop: an interactive online tool to calculate and visualize geographic LD patterns. *BMC Bioinformatics* 2020;21:14.
- [31] Karczewski KJ, Francioli LC, Tiao G, Cummings BB, Alföldi J, Wang Q, et al. The mutational constraint spectrum quantified from variation in 141,456 humans. *Nature* 2020;581:434–43.
- [32] Giambartolomei C, Vukcevic D, Schadt EE, Franke L, Hingorani AD, Wallace C, et al. Bayesian test for colocalisation between pairs of genetic association studies using summary statistics. *PLoS Genet* 2014;10:e1004383.
- [33] Giambartolomei C, Zhenli Liu J, Zhang W, Hauberg M, Shi H, Boocock J, et al. A Bayesian framework for multiple trait colocalization from summary association statistics. *Bioinformatics* 2018;34:2538–45.
- [34] Panjwani N, Wang F, Mastromatteo S, Bao A, Wang C, He G, et al. LocusFocus: web-based colocalization for the annotation and functional follow-up of GWAS. *PLoS Comput Biol* 2020;16:e1008336.
- [35] GTEx Consortium. The Genotype-Tissue Expression (GTEx) project. *Nat Genet* 2013;45:580–5.
- [36] Drivas TG, Lucas A, Ritchie MD. eQTPlot: a user-friendly R package for the visualization of colocalization between eQTL and GWAS signals. *BioData Min* 2021;14:32.