QTLViewer: an interactive webtool for genetic analysis in the Collaborative Cross and Diversity Outbred mouse populations

Matthew Vincent 1,1,† Isabela Gerdes Gyuricza 1,1,‡ Gregory R. Keele 1,‡ Daniel M. Gatti 1, Mark P. Keller 2, Karl W. Broman 3,*, Gary A. Churchill 1,1,‡

1The Jackson Laboratory, Bar Harbor, ME 04609, USA
2Department of Biochemistry, University of Wisconsin–Madison, Madison, WI 53706-1544, USA
3Department of Biostatistics and Medical Informatics, University of Wisconsin–Madison, Madison, WI 53706-1544, USA

†Corresponding author: The Jackson Laboratory, Bar Harbor, ME 04609, USA. Email: gary.churchill@jax.org
‡These authors contributed as first authors.

Abstract

The Collaborative Cross and the Diversity Outbred mouse populations are related multiparental populations, derived from the same 8 isogenic founder strains. They carry >50 M known genetic variants, which makes them ideal tools for mapping genetic loci that regulate phenotypes, including physiological and molecular traits. Mapping quantitative trait loci (QTLs), which can be powerful in recombinant populations, the CC and DO are well-suited for mapping quantitative trait loci mapping and related analyses performed through the R/qtl2 package. Additionally, the QTLViewer website serves as a repository for published Collaborative Cross and Diversity Outbred studies, increasing the accessibility of these genetic resources to the broader scientific community.

Keywords: multiparental populations; MPP; genome-wide association studies; GWAS; quantitative trait loci mapping; QTL mapping

Introduction

Multiparental populations (MPPs; de Koning and McIntyre 2017) improve upon traditional experimental crosses by interbreeding more than 2 isogenic strains capturing greater genetic diversity. The Collaborative Cross (CC) and Diversity Outbred (DO) mouse populations are related MPPs that have been used to genetically dissect complex traits, including obesity (Svenson et al. 2012), bone strength (Al-Barghouthi et al. 2021), short-term memory (Hsiao et al. 2020), and benzene response (French et al. 2015). As recombinant populations, the CC and DO are well-suited for mapping quantitative trait loci (QTLs), which can be powerful in studies of -omic traits, including gene expression, protein abundance, and chromatin accessibility (Aylor et al. 2011; Chick et al. 2016; Abu Toamih Atamni et al. 2018; Keller et al. 2018; Keele et al. 2020, 2021; Takemon et al. 2021; Gerdes Gyuricza et al. 2022).

The CC and the DO were bred from the same 8 strains: A/J, C57BL/6J, 129S1/SvImJ, NOD/ShiLtJ, NZO/HILJ, CAST/EiJ, PWK/PfJ, and WSB/EiJ (Chickill et al. 2004, 2012; Iraqi et al. 2012). The use of classical and wild-derived founder strains drives high levels of genetic variability (Yang et al. 2007, 2011). The CC represents ~60 recombiant strains that are fully inbred (>99%) whereas DO mice are outbred, thus each mouse is genetically unique. Mapping resolution is finer for the CC and the DO compared to intercrosses or backcrosses due to additional generations of meiosis, which facilitates the identification of candidate genes (Solberg Woods 2014). Mouse experiments support the collection of multiple types of data on the same individuals, including physiological traits and molecular assays (e.g. gene expression) across multiple tissues. Integrative approaches like mediation analysis can be used to delineate the relationships among traits with shared QTLs (Chick et al. 2016; Keele et al. 2020). This series of analyses enable richer findings from CC and DO experiments but can also obstruct researchers without statistical and computational expertise.

The R/qtlcharts software (Broman 2015) provided a set of interactive data visualization tools for QTL analysis, built on the JavaScript library D3 (Bostock et al. 2011) and accessible from R (R Core Team 2021), but it does not provide a comprehensive or user-friendly graphical user interface. Konganti et al. (2018) developed the gQTL webtool, a graphical interface QTL mapping tool for CC data based on the R package DOQTL (Gatti et al. 2014). However, DOQTL is no longer maintained and instead, the R/qtl2 package (Broman et al. 2019) should be used. Here, we present the QTLViewer software (https://qtlviewer.jax.org), an interactive QTL mapping webtool built for both CC and DO that utilizes the R/qtl2 package to perform QTL mapping as well as variant association mapping based on the founder strain genotypes. QTLViewer can perform mediation analysis for QTL through -omic traits to identify candidate causal mediators of physiological trait QTL. The primary goal of QTLViewer is to allow users to interactively perform and visualize QTL mapping for publicly...
available CC and DO datasets, and so, QTLViewer represents a data repository that can be also downloaded for further investigation. This software will empower researchers to analyze and explore CC and DO data while facilitating access to relevant data and findings from others across the community.

**Methods**

The QTLViewer is an interactive web application designed to interactively perform and visualize QTL mapping and related analyses in experimental CC or DO data. We make use of modern computing tools, such as application programming interfaces (APIs) and Docker containers, to make QTLViewer efficient, portable, and extendable.

**Implementation**

The QTLViewer is comprised of 3 Docker (https://www.docker.com/—v18.0.3 last accessed 06/13/2022) images that are managed with Docker-Compose (https://docs.docker.com/compose/—v1.23.2 last accessed 06/13/2022). The user interface is contained in the churchilllab/qtl2web container (v1.0.0), which is a Python (v3.6.9) web application that processes all requests made by the user. Each request is parsed and analyzed to determine which analyses need to be performed.

All computational requests are performed by the churchilllab/qtl2rest container (v0.1.0). This container loads data upon startup and listens to Web API calls via RestReserve (https://restreserve.org/—v0.4.1 last accessed 06/13/2022). All data returned from these APIs are in JSON format. Since some requests require longer time to complete than others, we use another instance of the churchilllab/qtl2web container to manage them. We utilize Celery (https://docs.celeryproject.org/—v4.7 last accessed 06/13/2022) with a Redis (https://redis.io/—v6.2 last accessed 06/13/2022) backend as a task queue.

The containers work together to handle all requests. Requests generated from the web page are submitted via AJAX (Asynchronous JavaScript) and processed by the churchilllab/qtl2web container. A dialog box with a spinning wheel is displayed to the end-user to show that the request is being processed. The request is parsed to determine which APIs need to be called. There could be one to many API calls per request. Those API calls are bundled together and submitted to the task queue for processing. Once the task queue receives the request, a unique ID is generated and sent back to the web page. The web page parses the unique ID and begins polling the task queue until the task is complete or an error occurs. While the web page is polling the task queue for a status, the API calls are submitted to the churchilllab/qtl2rest container for computational analyses. When all API calls are complete, the task queue bundles the data together and marks the job as complete. The background polling that has been happening on the web page see that the job is complete, removes the dialog box, and processes the JSON data to display the appropriate graphs and data.

A diagram illustrating the implementation process is shown in Fig. 1.

**Supporting tools**

QTLViewer relies on a set of tools that facilitate the analyses of MPPs. Ensimpl (v 1.0.0) is a customized small version of Ensembl (https://www.ensembl.org last accessed 06/13/2022) that provides a web API to retrieve genomic information. Ensimpl extracts the necessary data from Ensembl and creates efficient SQLite (https://www.sqlite.org/ last accessed 06/13/2022) databases for querying purposes. The databases are separated by Ensembl release and species (mouse and human). The web API is written in Python utilizing FastAPI (https://fastapi.tiangolo.com/ last accessed 06/13/2022) framework. The API functionality provides search, genomic location lookup, gene lookup, and gene history.

In addition to Ensimpl, QTLViewer also utilizes a Founder SNP Database (https://churchilllab.jax.org/foundersnp/ last accessed 06/13/2022). The Founder SNP database is a custom version of the Sanger SNP VCF files that have been processed via a Python script and stored in an SQLite (https://www.sqlite.org/ last accessed 06/13/2022) database. The SNP genomic location and allele calls for the 8 founder strain alleles are stored. A custom Strain Distribution Pattern is stored for easy query and dissemination of data. Currently, GRCm38 (https://www.ncbi.nlm.nih.gov/assembly/GCF_000001635.20/ last accessed 06/13/2022) release 1410 and 1505 are supported.

Data visualizations make use of the JavaScript library Highcharts (https://highcharts.com last accessed 06/13/2022) which is built on D3 (Bostock et al. 2011).

**Statistical methods**

The QTLViewer carries out several types of analysis to map and characterize QTL. QTL mapping analysis is run using the R/qtl2 package (Broman et al. 2019—v0.28). The QTLViewer maps QTL by testing for an additive locus effect (additive QTL) or a locus-by-factor interaction effect (interactive QTL). For the additive model, the following linear mixed model is used to test an additive locus effect at loci spanning the genome:

\[
\text{trait}_i = \text{QTL}_i + \text{covariates}_i + \text{kinship}_i + \text{error}_i, \quad \text{Equation 1}
\]

where \(\text{trait}_i\) is the phenotype value of mouse \(i\), \(\text{QTL}_i\) is the effect of locus \(m\) on mouse \(i\) being tested, \(\text{covariates}_i\) is the cumulative effect of all covariates on mouse \(i\), \(\text{kinship}_i\) is a random term that captures noise variation for mouse \(i\) due to population structure, and \(\text{error}_i\) is the independent random noise for mouse \(i\). The structure of the kinship term is encoded in a genetic relationship matrix (K) estimated from the genotypes. We use the “leave one chromosome out” (LOCO) approach in which the K used for each locus \(m\) fit by \text{Equation 1} excludes all markers from the chromosome of locus \(m\), which improves QTL mapping power (Wei and Xu 2016). The QTLViewer uses the same statistical model for mapping on all chromosomes including the X chromosome. The precomputed genotype probabilities take account of sex-specific differences (Broman et al. 2006).

For additive QTL, the \(\text{QTL}_i\) term represents allele dosages founder haplotypes at the locus. The QTLViewer will also plot the regression coefficients from the \(\text{QTL}_i\) term, i.e. the founder allele effects, when doing haplotype-based analysis. These effects can also be re-estimated as best linear unbiased predictors (BLUPs), which reduce the impact of rare alleles and can make signals clearer.

For interactive QTL, a similar model to \text{Equation 1} is used to test a locus-by-factor interaction effect at loci spanning the genome:

\[
\text{trait}_i = \text{QTL}_i + \text{factor}_i + (\text{QTL}_i \times \text{factor}_i)_m + \text{covariates}_i + \text{kinship}_i + \text{error}_i, \quad \text{Equation 2}
\]

where \text{factor}_i is a covariate of interest for mouse \(i\) that may have an interaction effect with genotype at the locus \(m\), \((\text{QTL}_i \times \text{factor}_i)_m\) is the QTL-by-factor interaction term at locus

\(m\), \(\text{covariates}_i\) is a cumulative effect of all covariates on mouse \(i\), \(\text{kinship}_i\) is a random term that captures noise variation for mouse \(i\) due to population structure, and \(\text{error}_i\) is the independent random noise for mouse \(i\).
m being tested, and all other terms as defined before. Possible factors include sex and age.

The QTLViewer can also perform association mapping on biallelic variants using both additive and interactive models. Rather than encoding genetic effects based on doses of founder haplotypes, the $QTL_m$ and $(QTL_m \times \text{factor})_m$ terms in Equations (1) and (2) can be fit to doses of alleles of specific variants, imputed from dense genotypes of the founder strains. Although variant association mapping sacrifices some of the information present in the founder haplotypes, it does enable researchers to potentially identify specific variants of interests and prioritize candidate genes.

Finally, the QTLViewer can perform mediation analysis for additive QTL to identify candidate causal mediators (Chick et al. 2017).
interaction (e.g. sex-interactive QTL) LOD peaks can be viewed by
Briefly, Equation (1) is reused (with the kinship term excluded for
gene expression, have been collected on the same mice or strains.
A strong candidate mediator will localize near the QTL and strongly re-
now conditioning one-by-one on candidate mediators. A strong
candidate mediator may be viewed by changing the Plot type select box. Additional information on
specific QTLs can be interactively accessed by hovering and clicking on the points in the grid. The plot supports pan and zoom features.

Step 4: Generating genome-wide LOD plots
(i.e. genome scans) for individual traits
To visualize a genome-wide LOD plot for a specific trait, users
should click on a point from the LOD Peaks grid or select a trait in
the text search. Genome-wide LOD plots are useful to visualize
peaks on the genome associated with the phenotype of interest.
For example, the LOD plot for the gene Sfi1 reveals a clear expres-
sion QTL (eQTL) on chromosome 13 at 65 Mb (Fig. 3a), which
is considered a distal eQTL because Sfi1 is located on chromosome
11. Information about each LOD score may be accessed by hover-
ing. The plots select box will switch between additive and factor-
itive QTL models. Clicking on a locus of interest (i.e. the
peak LOD score) is used to drive additional analyses on the locus,
which are visualized as plots in the lower panel. These analyses
include estimation of founder allele effects, mediation of the loc-
us effect through -omic traits, and SNP association mapping in the
locus region.

Step 5: Generating founder allele effects plot
The estimation of founder allele effects is used to obtain the con-
tributions of each founder haplotype to the QTL. To perform this
analysis, users should click on a specific peak on the genomewide
LOD plot for a trait and, by default, the allele effects estima-
tion will be the first analysis to be performed (Fig. 3b). The allele
effects plot shows the estimated founder allele effects across the
chromosome that contains the selected peak. The allele effects
can be reported as fixed effects coefficients (default) or as con-
strained BLUPs by checking the bar on top of the effect plot. BLUP
estimates are generally preferable because they shrink extreme
effects from rarely observed alleles, but they are computationally
more intensive to calculate. These effects can be used to distin-
guish which founder strains likely possess the causal genetic var-
ants. They also highlight QTL that is multiallelic (Crouse et al.
2020)—a unique feature of MPPs. The allele effects of the distal
eQTL on chromosome 13 for the gene Sfi1 show a biallelic pattern
driven by lower expression from the CAST/EiJ and C57BL/6j
alleles compared to higher expression from the other alleles
(Fig. 3b).

Step 6: Performing mediation of QTLs
through -omic traits
Mediation analysis can be used to identify candidate mediators of
a QTL and thus reveals relationships between coregulated
traits. For example, genetic variation local to a gene encoding a
transcription factor can alter its expression levels, which then
mediate changes in expression on its downstream targets.
Another example of coregulated traits is a gene that possesses
comapping eQTL for its transcript and pQTL for its protein, sug-
gesting that its protein levels are transcriptionally regulated.
Support for these relationships in the data can be assessed
through mediation analysis. To generate mediation plots users
should click on the Mediation tab adjacent to the Effect tab, and
then click back on the locus of interest on the LOD plot
(Fig. 4a). Doing so runs a mediation analysis which involves
restesting a QTL effect at the locus of interest, iteratively condi-
tioning on candidate mediators. Promising candidates will sig-
ificantly reduce or drop the initial QTL LOD score and be
encoded on the genome near the QTL. When performing a
mediation analysis of the gene Sfi1 against the transcriptome heart data, we see an LOD drop on chromosome 13 (where the distal eQTL is located; Fig. 4a). The gene with the lowest LOD score on chromosome 13 is Rsl1, the candidate mediator. We also see an LOD drop on chromosome 11 at the position of the Sfi1 gene (which corresponds to a local eQTL; Fig. 4a). This “LOD-drop” method of mediation is a powerful tool to identify candidate mediators, but it is also susceptible to false-positive detection of linked independent effects (Chick et al. 2016; Crouse et al. 2020). Mediation can be performed against different datasets by changing the selection in the Mediate Against box.

Fig. 2. Navigating datasets in the QTLViewer. The QTLViewer page for a project provides access to multiple datasets which can be explored by changing the option on “Current Data Set.” In the screenshot, the aging DO heart transcriptomics data is selected (a). Within this specific dataset, it is possible to search for specific traits. For example, in the transcriptome dataset, the genes Ace and Ace2 are available, but not Ace3 (b). Switching to the “Lod Peaks” mode, visualizes a genome-wide transcriptome map with marker IDs on the x-axis and gene position on the y-axis, which is a common plot type to summarize all the eQTLs in the data according to a specified LOD threshold (c).
**Step 7: Performing SNP association scans**

SNP (or variant) association mapping is useful to identify genetic variants of interest in the QTL region. To perform this analysis, users should click the SNP Association tab adjacent to the Mediation tab, and then click on a locus of interest on the genome-wide LOD plot (Fig. 4b). This will generate an SNP association plot, which shows LOD scores against the genomic position for all variants overlaid with annotated genes by their corresponding genomic positions on the bottom. The plot of the chromosome 13 shows multiple variants in strong linkage disequilibrium (LD) with each other (haplotype) and with high LOD scores at the position of the candidate mediator Rsl1 (Fig. 4b). Users can pan and zoom in the SNP association plots. Hovering over a variant will display additional information such as variant consequences, and the allele distribution pattern among the founder strains, which shows the magnitude and direction of the effect of each founder’s haplotypes on the trait.

**Step 8: Exploring the relationships among traits and covariates**

In addition to QTL mapping results, the QTLViewer can be used to explore the relationship between a trait and a covariate of interests (e.g. sex). If users are interested in this type of analysis, they should scroll down to the bottom left corner of the screen, under the Profile Plot tab (Fig. 5a). The profile plot is generated automatically after searching or clicking in a trait of interest. If users wish to modify the plots based upon the covariates observed in the experiment, they should click on the Select your factors selection box, which will turn factors on or off. In addition,
users can assess the relationships among traits based on correlation. This analysis can be useful when, for example, checking the correlation between the trait of interest and its genetic mediator identified in the mediation analysis (Step 6). Correlation tab adjacent to the Profile Plot tab will show the correlation of the element of interest with any other element in the data (Fig. 5b). The tool will display a scatter plot of the focus trait with another selected trait. To do that, users should choose a dataset in the Select Correlation Dataset box, and then select the element of interest from the table (Fig. 5c). The values of the elements can be adjusted by the covariates specified on the Covariate Adjustment box, and points can be colored by specific covariates determined...

**Fig. 4.** Mediation and SNP association plots. a) Mediation analysis can be performed on QTL of interest to identify candidate mediators as long as the QTL’s trait and the mediators are observed for the same mice (for the DO specifically) or for the same strains. The mediation plot shows genomic position across chromosomes on the x-axis and conditional LOD scores on the y-axis. By mediating the distant eQTL of Sfi1 on chromosome 13 through the DO heart transcriptomics data, the gene Sfi1 on chromosome 13 at 67 Mb is identified as a candidate mediator, which matches the eQTL position. We also see candidate mediators on chromosome 11. The one with the lowest LOD score is the gene Sfi1 itself and the other is a gene model (Gm11400) that is likely on LD with Sfi1. b) An SNP association scan can be performed in the QTL region by clicking on the “SNP Association” button. The SNP association plot shows genomic position in bp on the x-axis and LOD scores on the y-axis. Annotated genes are overlayed below the x-axis. Variant association in the CC and DO reveals haplotypes shelves of variants in strong LD with each other. Additional information on variants or genes can be accessed by hovering the cursor over the dot or gene track.
Select a series to color. We can see that Sfi1 has a strong positive correlation with the gene Gm114400 located on chromosome 11, which is likely in LD with Sfi1 (Fig. 5c). The gene Sfi1 also has a negative correlation with the candidate mediator Rsl1 (Fig. 5c). The negative correlation between Sfi1 and Rsl1 makes sense given that Rsl1 encodes a zinc-finger protein that functions as a negative regulator of gene expression (Krebs et al. 2003, 2014).

Step 9: Downloading QTLViewer data objects

The QTLViewer contains multiple R data objects that supply all the necessary inputs for the analyses, including traits, covariates, and genotype probabilities. In addition to exploring the data interactively on the QTLViewer webpage, users can download the corresponding R data objects by clicking on Download Data on the top of the page (Fig. 6a). All plots and analyses generated above can be downloaded as figures or data tables by clicking on the top.
right corner button on each plot (Fig. 6a). When clicking on Download Data, users will be redirected to a new page displaying all the downloadable RData files (Fig. 6b). There, users will find a "core" RData file containing all the input needed for mapping, such as genotype probabilities and marker information. Users can also download the "dataset" remote desktop services (RDS)
files that contain the trait data, sample and assay metadata annotations, and a summary of the QTL mapping results (Fig. 6c). This functionality enables users to quickly gain access to processed data files to run further analyses.

**Building your own QTLViewer**

The main goal of the QTLViewer is to allow users to interactively perform QTL mapping on publicly available data and to graphically explore results independent of their computational skills. However, if users wish to analyze their own data it is also possible to build a new QTLViewer object, which will require some coding experience. Instructions on how to set up a new QTLViewer can be found in the Supplementary File.

**Future directions**

The R/qtl2 software at the core of QTLViewer is very general and can accommodate a wide range of cross designs and data from model organisms other than the mouse. We plan to extend the QTLViewer to include mouse backcrosses, intercrosses, and 2-way recombinant inbred panels. Adapting QTLViewer to MPPs from species other than the mouse will require modification of settings in the configuration files and the gene-search features for -omics data, and creation of a variant database.

We are in the process of adding a new method for mediation analysis based on Bayesian model selection (Crouse et al. 2022) that can provide richer inference for mediators of interest. We continue to expand the QTLViewer for new types of genomic data, such as chromatin profiling data, which will be incorporated and layered onto the genome browser tracks.

We continue to add new datasets to the QTLViewer website, which already represents a powerful resource for the community. We welcome contributions from outside investigators and will assist in the process of preparing their data for the QTLViewer.

**Web resources**

Users can access the QTLViewer webpage at https://qtlviewer.jax.org/ (last accessed 06/13/2022). The QTLViewer software is version controlled and available from GitHub (https://github.com/churchill-lab; last accessed 06/13/2022).

The supporting tools Ensimpl and SNPDB are available online at https://churchilllab.jax.org/ensimpl (last accessed 06/13/2022) and https://churchilllab.jax.org/foundersnps (last accessed 06/13/2022), respectively. The Docker containers code can be found at https://github.com/churchill-lab/qtl2web (last accessed 06/13/2022), https://github.com/churchill-lab/qtl2rest (last accessed 06/13/2022), and https://github.com/churchill-lab/ensimpl (last accessed 06/13/2022). The Docker images are available at https://hub.docker.com/r/churchilllab/qtl2web/tags (last accessed 06/13/2022), https://hub.docker.com/r/churchilllab/qtl2rest/tags (last accessed 06/13/2022), and https://hub.docker.com/r/churchilllab/ensimpl/tags (last accessed 06/13/2022). The source code for QTL mapping analysis through R/qtl2 package can be found at https://github.com/rqtl/qtl2 (last accessed 06/13/2022).

**Data availability**

The aging heart data analyzed here to demonstrate the QTLViewer functions can be found online at https://churchilllab.jax.org/qtlviewer/JAC/DOHeart (last accessed 06/13/2022) and at figshare under accession number https://doi.org/10.6084/m9.figshare.12378077. Supplemental material is available at G3 online.

**Funding**

This work was supported by the Jackson Laboratory Cube Initiative and grant funding from the National Institute of Health (NIH): R01GM070683 (GAC and KWB), F32GM134599 (GRK), R01DK101573 (MPK), and RC2DK125961 (MPK). This work was also supported by the University of Wisconsin–Madison, Department of Biochemistry and Office of the Vice Chancellor for Research and Graduate Education with funding from the Wisconsin Alumni Research Foundation (MPK).

**Conflicts of interest**

None declared.

**Literature cited**

Abu Toamih Atammi H, Nashef A, Iraqi FA. The Collaborative Cross mouse model for dissecting genetic susceptibility to infectious diseases. Mamm Genome. 2018;29(7–8):471–487.

Al-Barghouthi BM, Mesner LD, Calabrese GM, Brooks D, Tommasini SM, Bouxein ML, Horowitz MC, Rosen CJ, Nguyen K, Haddox S, et al. Systems genetics in diversity outbred mice inform BMD GWAS and identify determinants of bone strength. Nat Commun. 2021;12(1):3408.

Aylor DL, Valdar W, Foulds-Mathes W, Buus RJ, Verdugo RA, Baric RS, Ferris MT, Freilinger JA, Heise M, Freeman MB, et al. Genetic analysis of complex traits in the emerging Collaborative Cross. Genome Res. 2011;21(8):1213–1222.

Bostock M, Ogievetsky V, Heer J. D3 data-driven documents. IEEE Trans Vis Comput Graph. 2011;17(12):2301–2309.

Broman KW, Gatti DM, Simecek P, Furlotte NA, Prins P, Sen Š, Yandell BS, Churchill GA. R/qtl2: software for mapping quantitative trait loci with high-dimensional data and multiparent populations. Genetics. 2019;211(2):495–502.

Broman KW, Sen S, Owens SE, Manichaikul A, Southard-Smith EM, Churchill GA. The X chromosome in quantitative trait locus mapping. Genetics. 2006;174(4):2151–2158.

Broman KW. R/qtlcharts: interactive graphics for quantitative trait locus mapping. Genetics. 2015;199(2):359–361.

Chick JM, Munger SC, Simecek P, Huttlin EI, Choi K, Daniel M. Defining the consequences of genetic variation on a proteome-wide scale. Nature. 2016;534(7608):500–5.

Churchill GA, Airey DC, Allayee H, Angel JM, Attie AD, Beatty J, Beavis WD, Belknap JK, Bennett B, Berrettini W, et al.; Complex Trait Consortium. The Collaborative Cross, a community resource for the genetic analysis of complex traits. Nat Genet. 2004;36(11):1133–37.

Churchill GA, Gatti DM, Munger SC, Svenson KL. The Diversity Outbred mouse population. Mamm Genome. 2012;23(9–10):713–718.

Crouse WL, Keele GR, Gastonguay MS, Churchill GA, Valdar W. A Bayesian model selection approach to mediation analysis. PLoS Genet. 2022;18(5):e1010184.

Crouse WL, Kelada SNP, Valdar W. Inferring the allelic series at QTL in multiparental populations. Genetics. 2020;216(4):957–983.

de Koning DJ, McIntyre LM. Back to the future: multiparent populations provide the key to unlocking the genetic basis of complex traits. Genetics. 2017;206(7):717–718.

French JE, Gatti DM, Morgan DI, Kissling GE, Shockley KR, Knudsen GA, Shepard KG, Price HC, King D, Witt KL, et al. Diversity outbred mice identify population-based exposure thresholds and genetic effects of common complex traits. G3, 2022;12, No. 8


factors that influence benzene-induced genotoxicity. Environ Health Perspect. 2015;123(3):237–45.
Gatti DM, Svenson KL, Shabalin A, Wu LY, Valdar WW, Simecek P, Goodwin N, Cheng R, Pomp D, Palmer A. Quantitative trait locus mapping methods for diversity outbred mice. G3 (Bethesda). 2014;4(9):1623–1633.
Gerdes Gyuricza I, Chick JM, Keele GR, Deighan AG, Munger SC, Korstanje RC, Gygi SP, Churchill GA. Genome-wide transcript and protein analysis highlights the role of protein homeostasis in the aging mouse heart. Genome Res. 2020;32(5):838–852.
Iraqi FA, Noble C, Pitman W, Yadav N, Kumar S, Keele GR, Terceros A, Kanke M, Conniff T, Cheleuitte-Nieves C, et al. A thalamic orphan receptor drives variability in short-term memory. Cell. 2020;183(2):522–536.e19.
Keele GR, Quach BC, Israel JW, Chappell GA, Lewis L, Safi A, Simon JM, Cotney P, Crawford GE, Valdar W, et al. Integrative QTL analysis of gene expression and chromatin accessibility identifies multi-tissue patterns of genetic regulation. PLoS Genet. 2020;16(1):e1008537.
Keller MP, Gatti DM, Schueler KL, Rabaglia ME, Stapleton DS, Simecek P, Vincent M, Allen S, Broman AT, Bacher R, et al. Genetic drivers of pancreatic islet function. Genetics. 2018;209(1):335–356.
Konganti K, Ehrlich A, Ivan R, Threadgill DW. gQTL: a web application for QTL analysis using the Collaborative Cross mouse genetic reference population. G3 (Bethesda). 2018;8:2559–2562.
Krebs CJ, Larkins LK, Price R, Tullis KM, Miller RD, Robins DM. Regulator of sex-limitation (Rsl) encodes a pair of KRB zinc-finger genes that control sexually dimorphic liver gene expression. Genes Dev. 2003;17(21):2664–2674.
Krebs CJ, Zhang D, Yin L, Robins DM. The KRB zinc finger protein RSL1 modulates sex-biased gene expression in liver and adipose tissue to maintain metabolic homeostasis. Mol Cell Biol. 2014;34(2):221–232.
R Core Team. R: a language and environment for statistical computing. 2021 [accessed 2021 Dec 22]. https://www.r-project.org/.
Solberg Woods LC. QTL mapping in outbred populations: successes and challenges. Physiol Genomics. 2014;46(3):81–90.
Svenson KL, Gatti DM, Valdar W, Welsh CE, Cheng R, Chesler EJ, Palmer AA, McMillan L, Churchill GA. High-resolution genetic mapping using the mouse Diversity Outbred population. Genetics. 2012;190(2):437–447.
Takemon Y, Chick JM, Gerdes Gyuricza I, Skelly DA, Devuyst O, Gygi SP, Churchill GA, Korstanje R. Proteomic and transcriptomic profiling reveal different aspects of aging in the kidney. eLife. 2021;10:e62585.
Wei J, Xu S. A random-model approach to QTL Mapping in Multiparent Advanced Generation Intercross (MAGIC) populations. Genetics. 2016;202(2):471–486.
Yang H, Bell TA, Churchill GA, Pardo-Manuel De Villena F. On the subspecific origin of the laboratory mouse. Nat Genet. 2007;39(9):1100–1107.
Yang H, Wang JR, Didion JP, Buus RJ, Bell TA, Welsh CE, Bonhomme F, Yu AH-T, Nachman MW, Pialek J, et al. Subspecific origin and haplotype diversity in the laboratory mouse. Nat Genet. 2011;43(7):648–655.
Yates AD, Achuthan P, Akanni W, Allen J, Allen J, Alvarez-Jarreta J, Amode MR, Armean IM, Azov AG, Bennett R, et al. Ensembl 2020. Nucleic Acids Res. 2020;48(D1):D682–D688.

Communicating editor: S. Macdonald