A Bayesian model selection approach to mediation analysis

Wesley L Crouse[1†], Gregory R Keele[2†], Madeleine S Gastonguay[2], Gary A Churchill[2*^] and William Valdar[1,3^]

[*]Correspondence:
gary.churchill@jax.org,
william.valdar@unc.edu
[1]Department of Genetics, University of North Carolina, 27599 Chapel Hill, NC, US
[2]The Jackson Laboratory, 600 Main Street, 04609 Bar Harbor, ME, US
Full list of author information is available at the end of the article
[†]Equal contributor [^]Also equal contributor

Abstract

Mediation analysis is a powerful tool for discovery of causal relationships. We describe a Bayesian model selection approach to mediation analysis that is implemented in our bmediatR software. Using simulations, we show that bmediatR performs as well or better than established methods including the Sobel test, while allowing greater flexibility in both model specification and in the types of inference that are possible. We applied bmediatR to genetic data from mice and human cell lines to demonstrate its ability to derive biologically meaningful findings. The Bayesian model selection framework is extensible to support a wide variety of mediation models.

Keywords: Mediation; Causal inference; Sobel test; Quantitative trait loci; QTL; Genome-wide association studies; GWAS

Introduction

Mediation analysis seeks to understand a causal process by determining whether an intermediate variable (M) explains (at least partially) the response of a dependent variable (Y) to changes in an independent variable (X). Though the causal interpretations of mediation analysis are subject to a number of assumptions [1], it has been widely used in both the social and natural sciences [2]. In a biomedical context, mediation analysis has been used to investigate how gene expression mediates the effects of genetic variants on complex phenotypes and disease [3, 4]. It has been used to infer causal relationships between biomolecular phenotypes, such as transcripts, chromatin states, and proteins [5, 6]. Here we focus on mediating genetic associations on biomolecular phenotypes, including gene expression, protein abundance, and chromatin accessibility, in samples of model organisms and human cell lines, but the mediation approach we introduce is broadly applicable.

Mediation analysis requires that X, M, and Y are measured in the same individuals, and its causal interpretation relies on assumptions that are not verifiable based solely on the data [7, 1, 8]. In particular, it assumes that the direction of causal effects is from X to Y (X→Y), referred to as the direct effect of X on Y, and from X to M to Y (X→M→Y), referred to as the indirect effect of X on Y. Also assumed is that these relationships are not subject to unobserved confounding, that observed confounders are conditioned on, and that observed confounders of M to Y do not depend on X. If these assumptions are satisfied, evidence for causal mediation lies in the magnitude of the indirect effect. If this indirect effect is non-zero, then M is a mediator of X on Y. Further, if M is a mediator and the direct effect is zero, then M is a complete mediator of X on Y, whereas if the direct effect is non-zero, then M is a partial mediator of X on Y. Our objective is to assess evidence of complete or partial mediation including the non-standard case where X comprises more than one independent variable.

Our motivating context is quantitative trait locus (QTL) mapping, where a univariate trait of interest Y (*e.g.*, protein abundance) is associated with genetic variation represented in a matrix X, which could encode multiple variants or their combinations (*i.e.*, haplotypes). In particular, we are interested in assessing whether one or more univariate candidate variables M may mediate the relationship between the genetic matrix and the trait. For example, a candidate mediator M could be the protein abundance of a gene encoded nearby a QTL for Y. In this case, it is reasonable to assume that the direction of causal effects is $X \rightarrow M \rightarrow Y$, and we further assume that the there is no unexplained confounding of these relationships. With the assumptions of mediation satisfied, our objectives are two-fold: 1) assess the evidence in favor of M being a causal mediator, 2) and determine if the mediation is partial or complete, given that X contains complex genetic information.

Traditional methods for mediation analysis are poorly suited to the dual objectives of detecting mediation and distinguishing complete from partial mediation when there are many candidate mediators and when X is a matrix. A classic approach for establishing mediation was introduced by Baron & Kenny [9]. This approach, termed the causal steps (CS) method, establishes evidence for partial or complete mediation by sequentially testing the relationships between X, M, and Y. Specifically, CS uses linear regression models to establish the following four conditions: 1) X has a marginal effect on Y [$X \rightarrow Y$]; 2) X has an effect on M [$X \rightarrow M$]; 3) M is a partial mediator of the effect of X on Y [$M \rightarrow Y|X$]; and 4) M is a complete mediator of the effect of X on Y [$X \perp\!\!\!\perp Y|M$]. The CS method can accommodate a matrix of independent variables by using a likelihood ratio test for grouped predictors. Although the CS method is useful due to its conceptual accessibility, its implementation in a genomics setting with many candidate mediators can be awkward. In particular, it is not straightforward to combine statistics across the steps while also accounting for multiple testing, particularly for step (4), which requires a failure to reject a null hypothesis. This makes it difficult to succinctly summarize evidence for complete or partial mediation for many candidate M.

Other common tests for mediation analysis address the problems of the CS method by ignoring complete mediation and instead providing a single test statistic for the significance of the indirect effect. The indirect effect is formally given as the product of regression coefficients from $X \rightarrow M$ and $M \rightarrow Y \mid X$, that is, the effect of M on Y controlling for the effect of X. Establishing that this coefficient product is non-zero gives evidence for (at least) partial mediation, but it does not provide information about complete mediation. The most popular methods for testing the indirect effect include the Sobel test [10], which is based on an approximation to the asymptotic distribution of the indirect effect; alternatively, bootstrapping can be used to assess significance [11], which does not make distributional assumptions but is computationally expensive. In addition to not providing information about complete mediation, the Sobel test does not generalize when X is a matrix rather than a vector.

Applications of mediation analysis to large scale genetic and molecular profiling data have used modified versions of the traditional tests described above. Approximations to CS have been used in the multi-parent Collaborative Cross (CC [12, 6, 13]) and Diversity Outbred mouse populations (DO [14, 5, 15, 16]). These

studies identified QTL for gene expression (eQTL), protein abundance (pQTL), and chromatin accessibility (cQTL). Detection of a significant QTL for the target phenotype (Y) satisfies step (1) and detection of a QTL local to the molecular trait (M), *i.e.*, near the genomic position of M, satisfies step (2). For a given phenotype QTL, a mediation scan is performed by testing the effect of X on Y as being mediated through each M (*e.g.*, each observed gene transcript). For the approximation to CS, significant mediators are determined based on the reduction in log-odds (LOD) score before and after accounting for the effects of M (hereafter referred to as LOD drop). This approximates step (4) without requiring complete independence between X and Y. Notably, the LOD drop method does not directly check step (3) and, as a result, it may detect candidates that are correlated with the true mediator but are not mediating the effect of X on Y. Thus, care is needed when interpreting the LOD drop mediation scan.

More recent methodological developments in large-scale mediation of genetic and molecular profiling data include the multi-SNP intersection union test [17], an extension of the CS method that simultaneously models the overall effect of multiple genetic predictors by representing them as a similarity matrix-based (*i.e.*, kernel-based) random effect, and the divide-aggregate composite null test [18], an extension of the joint significance test [19] that improves power relative to the Sobel and joint significance tests by utilizing an empirical null distribution. As with other methods based on the indirect effect, neither of these methods provide inference on distinguishing partial and complete mediation.

All of the approaches described rely on hypothesis testing, in which significance criteria are used to choose between nested alternative models. In our view, a more natural perspective for mediation analysis is provided by Bayesian model selection. Specifically, the goal of mediation analysis is to classify the relationship between X, M, and Y as a particular causal model. We note that the full space of potential causal models is not nested; the classification of the observed data into a particular causal model is, with finite data, necessarily uncertain; estimation of parameters when the model is uncertain ideally requires incorporation of model uncertainty into the estimate. These properties are well suited to the Bayesian model selection paradigm, which considers a set of potential models (nested or otherwise) and assigns to each a posterior probability.

Bayesian methods have been previously used in mediation analysis to estimate the posterior distribution of the indirect effect [20, 21, 22], and Bayesian model selection has been used to test for the presence of an indirect effect [23]. Here we develop a more expansive Bayesian model selection approach that considers the full space of causal models, distinguishing between—among others—complete mediation, partial mediation, and independence (where X affects M and Y independently, e.g. co-localized but distinct eQTL and pQTL). Bayesian approaches are often computationally intensive, and even infeasible for many applications. We avoid this problem by employing conjugate priors that allow us to avoid costly sampling when calculating posterior summaries. Our approach is computationally efficient, generalizes to multiple independent variables, and provides unique posterior summaries in the form of causal model probabilities.

Results

We developed a Bayesian model selection approach, implemented in the bmediatR R package, to evaluate different causal models that define relationships between a continuous dependent variable Y, an independent (or exogenous) variable X, and a continuous mediator variable M. The causal models can be described as directed acyclic graphs (DAGs) (Figure 1). There are three possible edges (a, b, c) and we define each to be either present or absent using an indicator vector $\boldsymbol{\theta} = (\theta_a, \theta_b, \theta_c)$, where for example $\boldsymbol{\theta} = (1, 0, 0)$ denotes presence of a only. This leads to eight possible combinations of edges, each defining a different causal relationship (causal model) (ML1-8 in Figure 1). As an extension, the Bayesian model selection approach can also consider four additional causal models in which the direction of edge b (between M and Y) is reversed (ML9-12 in Figure 1, termed "reactive" models.), though some of these are are not causally identifiable (see Methods).

We are primarily interested in models that describe mediation, defined as any causal relationship where edges a and b are both present, *i.e.*, $\boldsymbol{\theta} = (1, 1, \theta_c)$, indicating a causal path from X to Y through M. Complete mediation describes models where X acts entirely through M and thus X conveys no additional information about Y beyond that provided by M ($\theta_c = 0$); partial mediation describes models where X conveys additional information beyond that contained in M ($\theta_c = 1$); and there is no mediation when $\theta_a = 0$ or $\theta_b = 0$, which includes the co-local model in which X affects both Y and M but independently, and also includes the model in which X and M affect Y but are independent of each other. Our approach calculates a posterior probability for each causal model $\boldsymbol{\theta}$ given data X, M, and Y.

Bayesian model selection requires specifying prior probabilities for the models under comparison, as well as prior distributions for the parameters of those models (model priors). Priors can be used to incorporate external information into mediation analysis, which can improve power and reduce false positive rates. Since the results of Bayesian model selection analysis can be sensitive to the choice of priors, it is important to understand the role and interpretation of different prior specifications. The model priors are specified by assigning probabilities to each of the twelve possible configurations of $\boldsymbol{\theta}$ (Figure 1). These probabilities can be viewed as weights and modified to define informative prior expectations on the relative frequency of different causal models. Setting a model prior to zero removes it from consideration, and this allows a user to define a set of allowable models. By default, we assign equal prior probability to the models that assume no reverse causality from Y to M for edge b (*i.e.*, probability 1/8 for each model ML1-8 but probability zero for ML9-12).

Prior distributions on the effect sizes of the edges connecting X, M, and Y (effect priors) are given conjugate prior distributions, which yields a closed form for the joint likelihood of M and Y. The size of the effects is controlled by specifying the hyperparameters $\boldsymbol{\phi}$, which are ratios of the edge effect sizes to the variability of the errors (Methods). The default specification $\boldsymbol{\phi} = (1, 1, 1)$ assumes that all edges have *a priori* equal effect sizes, also equal in size to the error variances of M and Y. Using conjugate priors and specifying the edge effect sizes make computing the posterior fast and exact.

Bayesian model selection performance in simulated data

We evaluated our Bayesian model selection against two alternative mediation analysis methods, the Sobel test and LOD drop, using simulated data. Data was simulated by first specifying X and then simulating M and Y according to linear models with normal error and effect sizes, expressed as proportion of variation explained, that ranged from 0.05 to 0.95. This approach allows X to represent a non-normal (and non-scalar) quantity. In particular, for genetics applications, X may represent minor allele counts, multi-state founder haplotype probabilities, or multiple genetic variants. Data were simulated for five scenarios: co-local, where X drives M and Y independently; partial mediation, where X drives Y both directly and indirectly through M; complete mediation, where X drives Y only through its effect on M; and reactive versions of partial and complete mediation, where X drives M through Y instead. For each scenario, we simulated 100 data sets with 200 observations each. For each simulated data set, we applied bmediatR's default model priors (*i.e.*, uniform over ML1-8 in Figure 1). We also considered two other types of model priors: reduced model priors, which sets a uniform prior probability over models that assume the direct effect c (ML4-8), and which is analogous to testing the indirect effect like the Sobel test; and expanded model priors, which sets a uniform prior over all models (ML1-12), including reactive ones (ML9-12) (Figures S1 and S2).

It should be noted that the Sobel test, LOD drop, and bmediatR provide different inferences. The Sobel test does not distinguish between partial and complete mediation, or their reactive forms. The LOD drop method does not distinguish between co-local, complete, partial mediation, or the reactive forms of the mediation models. Which models can be distinguished by the Bayesian model selection will depend on which prior model probabilities are non-zero, resulting in less or more complicated inference. If a direct effect from X to Y is assumed (edge c fixed), Bayesian model selection no longer distinguishes partial and complete mediation, similar to the Sobel test. Alternatively, more complex inference is possible by including reactive models as non-zero prior models in situations when the effect from Y to M is plausible, though inference between some models is not causally identifiable (see Discussion).

Bayesian model selection performs well when genotypes are correctly specified

We evaluated the performance of the three mediation methods, with bmediatR set to default model priors, on data with a bi-allelic QTL with an allele frequency of 0.5, simulated under the co-local, partial mediation, and complete mediation models. Results from bmediatR with default model priors are shown in Figure 2; results from all prior settings and all simulation settings, including reactive, are shown in Figures S1 and S2.

The left panel of Figure 2a shows results for when the SNP affects both M and Y but independently, as would be the case for an eQTL and a pQTL that are co-local but otherwise unrelated. The top four boxes show the posterior probability given by bmediatR for four types of models: complete mediation (ML4), partial mediation (ML8), co-local (ML7), and other non-mediation models (ML1-3 and ML5-6). Each box is a heatmap of the posterior probability under all possible simulated effect sizes of X→M (x-axis) and X→Y (y-axis). Together, the four boxes show that, for all but the most extreme combinations of effect sizes, bmediatR overwhelmingly favors

the (correct) co-local model. Below these are heatmaps for the Sobel test and the LOD drop. The Sobel test aims to identify mediation, and so any detections for this co-local scenario will be false positives. Its heatmap shows that the false positive rate is low in all cases. The LOD drop heatmap shows the magnitude of the signal for mediation by that method. LOD drop is seen to correctly register a minimal mediation signal when the effect size of the QTL for Y is large and that for M is small; but it incorrectly detects mediation when the QTL for M is large (>80%), a situation that is common in eQTL data.

The left panels of Figure 2b,c show results for the same set up but where mediation is present, either partial or complete. For partial (Figure 2b), bmediatR correctly identifies the type of mediation in most cases except for when the effect size of the QTL for M much is disproportionately large, in which case it is misclassified as complete or co-local. The Sobel test correctly identifies the presence of mediation (type unspecified) with a similar error pattern as for bmediatR. The LOD drop performs a little better than other methods, albeit with different error characteristics. For complete mediation (Figure 2c), the Sobel test and LOD drop perform slightly better than for partial mediation, while bmediatR accurately identifies the mediation type across all but the most extreme effect size settings.

Misspecification of bi-allelic genotypes induces false mediation

Mediation was then evaluated when the bi-allelic X is misspecified. This can occur, for example, when X represents a variant that is imperfectly correlated with the true causal variant through linkage disequilibrium (LD) (Figure 2 right panels). In this setting, Bayesian model selection begins to favor the partial mediation model and the Sobel test begins to detect mediation when it is not present. LOD drop also suffers with misspecification, exacerbating its issues when the QTL on M is large. When data were simulated with mediation present (partial and complete), misspecification of X was less problematic (Figure 2b-c right), although Bayesian model selection was less accurate at distinguishing partial and complete mediation.

Misspecification of multi-allelic genotypes as bi-allelic induces false mediation

Next we evaluated how mediation would be affected by misspecifying a multi-state haplotype effect. Data were simulated for 200 individuals with equal frequency of four alleles with distinct effects. Fitting X based on a bi-allelic SNP that tags the low and high haplotype groups worked well for complete mediation, but resulted in false mediation signals for co-local data, only correctly preferring the co-local model when at least one of the QTL on M and Y were small ($< 50\%$) (Figure 3a). This issue was exacerbated when the SNP was more imbalanced, resulting in false mediation signal for co-local data across a wider range of effect sizes (Figure 3b). We then looked at simulated bi-allelic data but modeled the genetic effect as eight haplotypes. Bayesian model selection performed well for co-local and complete mediation data, aside from some edge cases with extreme effect sizes, most notably at the corners of large QTL on M with small QTL on Y (prefers complete mediation) and small QTL on M and large QTL on Y (prefers partial mediation) (Figure 3c).

Bayesian model selection can detect mediation of multi-allelic QTL

To demonstrate mediation analysis in the context of multi-allelic QTL analysis, we simulated data based on the genomes of 192 DO mice [5]. A genetic locus X was randomly selected and M and Y were simulated assuming no mediation (QTL for Y, no QTL for M, M and Y uncorrelated; ML5), co-local (ML7), and complete mediation (ML4). QTL mapping was then performed for both M and Y to obtain LOD scores and estimated haplotype effects (Figure 4). The simulations demonstrate characteristic features of co-mapping QTL with correlated effects for co-local and complete mediation simulations, as they would appear in a multi-allelic QTL analysis. Data were simulated assuming a bi-allelic genetic variant, but the QTL mapping and Bayesian model selection were performed based on 8-state founder haplotypes.

Bayesian model selection applied to DO mice

In order to illustrate how bmediatR works in QTL mapping applications, we analyzed previously reported liver proteomics data from 192 DO mice [5, 13]. We first illustrate how multi-state haplotypes help to identify the causal driver of a distal pQTL for *Snx4*. Then we look at a case in which bmediatR identifies a biologically plausible mediator for a distal pQTL of *Tubg1* where LOD drop favored a less plausible candidate.

Multi-state haplotypes improves mediation inference for SNX4 distal pQTL

The *Snx4* gene is located on chromosome 16 and has a distal pQTL on chromosome 3 that co-maps with a local pQTL for *Snx7* (Figure 5). The proteins for both genes are sorting nexins that bind phospholipids, form protein-protein interactions, and play a role in membrane trafficking and protein sorting [24]. The haplotype effects of the *Snx7* local pQTL and the *Snx4* distal pQTL are highly correlated ($r = 0.99$) and reveal a complex haplotype effects pattern that cannot be explained by a single bi-allelic variant. We used the TIMBR software [25] to determine that the pQTL has as many as ≈5 distinct functional alleles (Figure S3a). Bi-allelic variants in the pQTL region with alleles shared by the B6 and 129 strains partially match the haplotype effects and thus have strong associations with both proteins. In order to evaluate SNX7 as a mediator of the pQTL effect, we first applied Bayesian mediation using the peak SNP (with B6 + 129 allele) and find that this analysis assigns most of the posterior probability to the partial mediation model. In contrast, Bayesian mediation using the 8-state haplotype information finds strong support for complete mediation. We then evaluated all proteins (genome-wide) as potential mediators of the *Snx4* distal pQTL, and only SNX7 is identified as a likely complete mediator based on the log posterior odds (Figure 5e). The co-local model is highly unlikely for SNX7. Summing the posterior probabilities for complete and partial mediation we see clearly that SNX7 is the most likely candidate mediator of the distal QTL for *Snx4*.

Bayesian model selection favors a biologically plausible driver for TUBG1

The gene *Tubg1* is located on chromosome 11 and has a distal pQTL on chromosome 8. In our previous work, genome-wide mediation analysis by the LOD drop method

identified NAXD as the best candidate mediator, but also revealed a significant LOD drop score for TUBGCP3 (Figure 6). Both mediation candidates have local pQTL. The *Naxd* local pQTL is stronger (LOD score ≈40) than both the *Tubg1* distal and *Tubgcp3* local pQTL (LOD scores ≈10). The pQTL haplotype effects for *Tubgcp3* and *Naxd* are highly correlated with the distal pQTL effects on TUBG1 (although flipped for NAXD) (Figure 6c). The allelic series could be either bi- or tri-allelic (k of 2-3; Figure S3b). We applied Bayesian model selection to both candidate mediators using the haplotype effects. The posterior probabilities at the two candidate mediators show that (Figure 6f) TUBGCP3 has 98% probability as a partial mediator, with the remaining probability for complete mediator. In contrast, NAXD, which is the best candidate using the LOD drop method, has 79% probability of being a complete mediator, 20% partial mediator, and 1% co-local probability. Due to the non-zero co-local probability of NAXD, TUBGCP3 is preferred based on the combined posterior odds of partial and complete mediation. A related tubulin gene, *Tubgcp2*, encoded on chromosome 7, has a distal pQTL at this locus on chromosome 8 (Figure 6g). Bayesian model selection analysis of the *Tubgcp2* distal pQTL also supports TUBGCP3 as a candidate mediator. These findings support a model for stoichiometric co-regulation of the protein constituents of the tubulin small complex [26, 27] that is mediated by the abundance of TUBGCP3.

Mediation of genetic effects on gene expression through chromatin accessibility in human cell lines

We applied Bayesian model selection to data from 63 human lymphoblastoid cell lines (LCLs) using genotype [28], RNA-seq [29], and DNase-seq [30, 31] data. We used both Bayesian model selection and the Sobel test to identify chromatin regions near transcription start sites that may act as mediators of gene expression. Previously, genetic variants that affect chromatin accessibility and gene expression were mapped in the LCLs [30]. Variants associated with chromatin accessibility were also likely to be associated with the expression of nearby genes, indicating that chromatin accessibility may be a common mechanism by which transcription is regulated.

To illustrate, we looked at the expression of gene SLFN5 which was strongly associated with local genetic variation (eQTL). These genetic variants were also strongly associated with chromatin accessibility (cQTL) (Figure 7a). The SNP most strongly associated with SLFN5 expression and chromatin accessibility is located within an interferon-stimulated response element (ISRE) in the first intron of SLFN5 [30], and the co-mapping chromatin site is directly above the SNP. The juxtaposition of chromatin site to gene makes it a likely mediator for the expression of SLFN5, an interferon-regulated gene, because it controls the accessibility of the ISRE to transcription factors [32]. Bayesian model selection and the Sobel test both detect partial mediation.

Co-mapping eQTL and cQTL can also represent independent signals. A SNP within 10 Kbp of the start of GPR63 is a local eQTL and a local cQTL for a nearby chromatin site (Figure 7b). The Sobel test does not support the chromatin site as a mediator of the genetic effect on GPR63 expression. Bayesian model selection results are consistent with the Sobel test in ruling out mediation but, unlike the Sobel test, it clearly identifies the relationship as co-local.

Discussion

Results from our Bayesian model selection analysis can be sensitive to the choice of model prior; therefore, it is important to understand the role and interpretation of different prior specifications. In our framework, is is possible to define any configuration and relative weight of causal models using the model prior. For example, if one believes that X is not a QTL for most candidate mediators, we could assign a smaller prior probability to models that include the a edge; or, if we believe that reactive models are unlikely but do not want to exclude them completely, we could assign these models a small prior probability (with the caveat that some reactive models are not identifiable). The appropriateness of such choices depends on the context of the analysis and the set of candidate mediators under consideration (*e.g.*, considering only genes nearby X, versus all genes in the genome).

A straightforward elaboration of our approach is to use a set of candidate mediators to learn an empirical prior distribution over the causal models. If suitably constructed, an empirical prior could improve power within a given dataset. An empirical prior could be implemented, for example, by computing the model posterior for all candidate mediators using the uniform model prior, and then taking the average over all posteriors as the empirical prior distribution for downstream analyses. Conceptually, this approach is similar to the empirical null employed in Liu *et al* [18]. As an illustration, in the case where X is a QTL for Y, if most of the candidates are not actually mediators, an empirical prior would put a high prior probability on models that include the c edge and do not represent mediation. Our simulations indicate that the default uniform model prior led to reasonable model selection inference over a range of true effect sizes and causal models. That said, exploring alternative prior distributions over the causal models is an intriguing avenue for future research, and many strategies could be implemented within our flexible framework.

There may also be opportunities to improve inference by adjusting the effect size hyperparameters, $\boldsymbol{\phi}$. By default, we assume *a priori* equal effect sizes for each edge, which are also equal to the size of the error variances. Tuning these hyperparameters to reflect prior beliefs about relative effect sizes may improve posterior model inference, provided these prior beliefs are closer to the truth. It is possible to put prior distributions on the effect size hyperparameters, but doing so would complicate posterior inference, as conditioning on these hyperparameters makes inference fast and exact. Other possibilities for setting the effect size hyperparameters include maximum *a posteriori* estimation (*e.g.*, via grid search) or defining an empirical prior based on the data, though we have not explored these ideas here.

Mediation analysis critically assumes that there is no confounding between M and Y, *i.e.*, $\mathrm{M} \leftarrow \mathrm{U} \rightarrow \mathrm{Y}$. In the presence of confounding, M and Y may be associated even if there is no direct effect from M to Y. Known confounders can be adjusted for as covariates ($\mathbf{Z}$ in the Methods). Unobserved confounders present a more fundamental problem. Approaches that account for unobserved confounders [33] could be implemented, but only as part of a two-stage approach, where latent factors are inferred first, naive to the mediation model, and then either removed or included as covariates in the mediation analysis. In our real data applications, we analyzed experimental data from mice and human cell lines, which we assumed was unconfounded after controlling for sex, batch, and other experimental variables.

There are other non-mediation approaches that are commonly used to investigate the relationship between X, M, and Y. These include Mendelian randomization (MR) [34, 35, 36, 37], which is an instrumental variables approach for estimating the causal effect of M on Y in the presence of confounding. MR involves building a predicted M (using X $\rightarrow$ M), and regressing Y on the predicted (rather than observed) M. This approach is robust to confounding, but it requires the additional strong assumption of no direct effect of X on Y (*i.e.*, no partial mediation). This means that classical MR cannot be used to distinguish between partial and complete mediation. Emerging MR methods relax the assumption of no direct effect and may be useful in this context, but this field is rapidly developing and a full review of these methods is beyond the scope of this paper. A different, non-causal method for evaluating the relationship between M and Y is colocalization analysis [38, 39, 40, 41, 42], which evaluates if two variables (M and Y) share a common X, when only one of many independent variables are under consideration. This approach is typically motivated by fine-mapping genetic variants (*i.e.*, evaluating candidate X, rather than candidate M), and it does not establish a causal relationship between M and Y in the way that mediation analysis can.

Mediation analysis commonly assumes that the direction of causal effect is from M to Y. This assumption must be justified based on the context of the analysis. For example, the mediator of a distal pQTL is expected to possess a corresponding local QTL. In the presence of reverse causality from Y to M, however, standard mediation analysis (*i.e.*, inference that excludes the reactive models) will still detect an edge from M to Y, yielding spurious evidence for mediation. In our framework, is it straightforward to include the reactive models in settings where reverse causality is a possibility. We examined inference in the presence of reverse causality via simulation (Figures S1 and S2). These results demonstrate that it is possible to distinguish complete mediation from complete reactive mediation, but also indicate that it is not possible to distinguish partial mediation from reactive partial mediation. This is not unexpected, as these models are not identifiable. Our simulation results also show that, for the priors we specified, true partial mediation can appear more consistent with reactive complete mediation than partial mediation, for a particular space of effect sizes (Figure S2b bottom row). These findings emphasize the importance of the assumption that M $\rightarrow$ Y in mediation analysis. By default, our bmediatR software excludes reactive models, and we recommend interpreting inferences made using the reactive models with caution.

Bayesian model selection provides an opportunity for more flexible model specifications than are possible for other methods, such as the Sobel test. In particular, the exogenous variable X can be categorical (with more than two groups) or multivariable. This is useful for modeling genetic effects in some contexts, as we demonstrated by the improvement in mediation using 8-state haplotypes in the DO mouse data. Mediation analysis with multiple exogenous variables [43] is easily implemented in bmediatR. The bmediatR software offers a framework that could be generalized to account for moderated mediation [44, 9, 45], in which the effect of the mediator is moderated by another factor. Increasing the complexity of the model space can increase computation time and may require more complex posterior summaries. Despite these hurdles, compared to methods like the Sobel test, implementing more

general mediation models settings is straightforward within the Bayesian model selection framework.

Lastly, we caution that mediation analysis is sensitive to misspecification. In our simulations, we showed that when a genotype X is misspecified (*i.e.*, a haplotype is incompletely described by a single variant), a true co-local model can be inferred as mediation. Simulations in other studies have demonstrated that greater precision in the measurement of Y relative to M can lead to true complete mediation appearing as partial mediation [46]. These issues highlight the need for careful examination and independent validation of inferences based solely on mediation analysis.

Conclusions

Mediation analysis is a powerful tool for discovery and integrative analyses of high dimensional biological data, with the caveat that proper caution and awareness of the potential pitfalls are needed. Here we describe a flexible Bayesian model selection approach to mediation analysis that is implemented as the R package bmediatR. Using simulations, we show that bmediatR performs as well or better than established methods including the Sobel test, while allowing greater flexibility in both model specification and in the types of inference that are possible. We applied bmediatR to genetic data from mice and from human cell lines, demonstrating its ability to derive biologically meaningful findings. The Bayesian model selection approach provides a flexible framework to support further advances in mediation analysis methods.

Figures

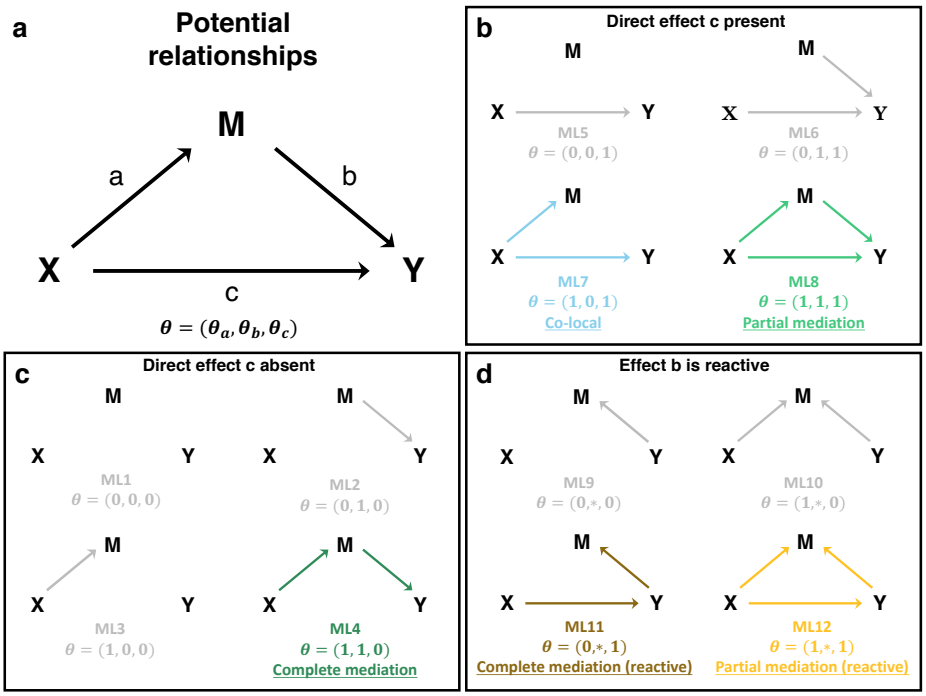

Figure 1 Possible relationships among X, M, and Y. X is assumed to be exogenous, and thus M and Y have no effects on X. A model and corresponding marginal likelihood (ML) are defined by the presence or absence of any the three edges a, b, and c according to an indicator variable $\boldsymbol{\theta}$. In this work and by default in bmediatR, the direction of edge b is assumed to be from M to Y (M $\rightarrow$ Y), but a set of reactive models can also be accommodated in which the direction of edge b is reversed (M $\leftarrow$ Y), indicated with $\boldsymbol{\theta} = (\theta_a, *, \theta_b)$. Models can be favored or even excluded by adjusting the model priors. By default, there are five models (ML1-3 and ML5-6) that represent non-mediation, *i.e.*, the effect of X on Y, if present, is not mediated through M. The co-local model (ML7) represents a special case where there is no mediation between X and Y, but X independently affects M and Y. The complete mediation model (ML4) and the partial mediation model (ML8) represent cases where the effect of X on Y is explained, completely or partially, by the effect of X on M.

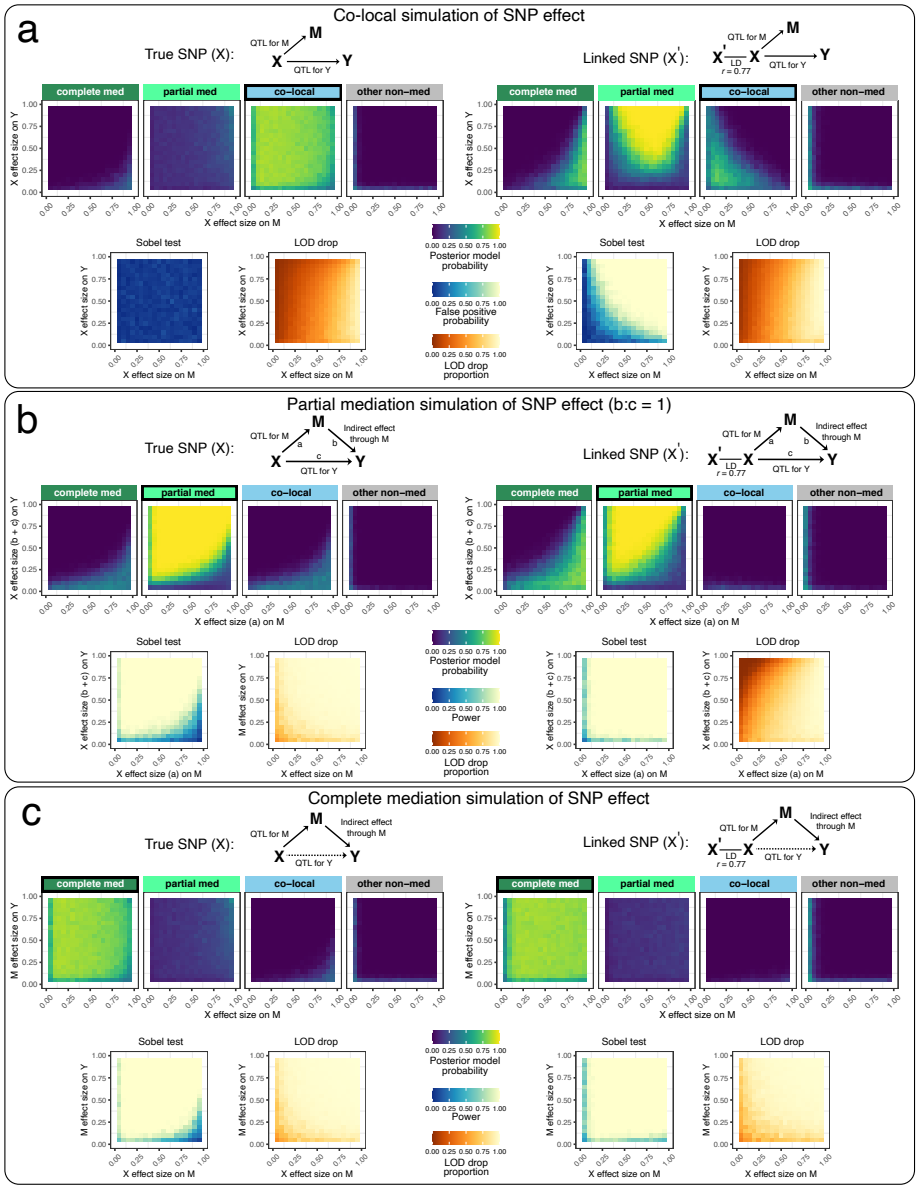

Figure 2 Performance of Bayesian model selection, Sobel test, and LOD drop in simulated data with a binary exogenous variable. Data for 200 individuals were simulated according to (a) co-local, (b) partial mediation, and (c) complete mediation models based on a balanced bi-allelic variant X. We applied mediation analysis with X as the true variant (left) and as a variant in linkage disequilibrium ($r = 0.77$) with the true variant (right). DAGs indicate the model used to simulate the data. Heat maps for Bayesian model selection represent the mean posterior probability associated with each inferred model for a range of fixed settings of the model parameters as indicated on x- and y-axes. Heat maps for the Sobel test represent false positive probability for co-local simulations and power for mediation simulations. Heatmaps for LOD drop represent mean LOD drop, scaled to the proportion of the simulated QTL's LOD score. See Figures S1 and S2 for results from non-default model priors and all simulated models, including reactive.

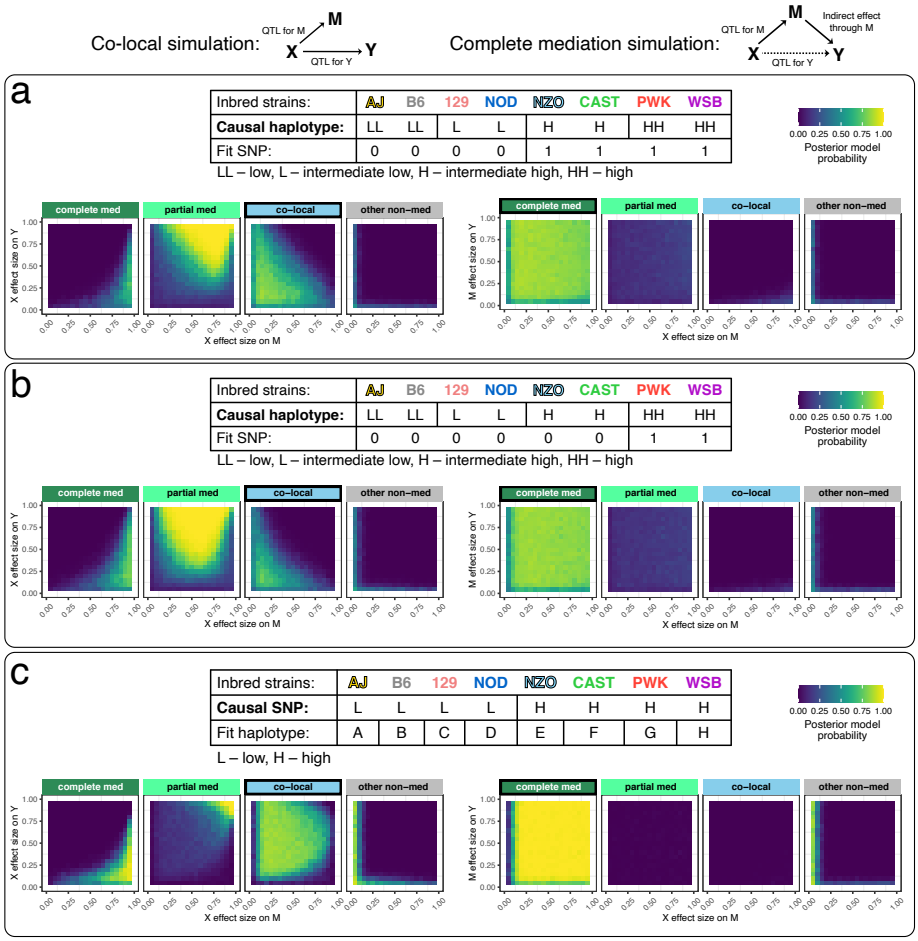

Figure 3 Performance of Bayesian model selection in simulated data with a multi-state exogenous variable. Data for 200 individuals were simulated according to co-local (left) and complete mediation (right) models. DAGs indicate the model used to simulate the data. (a-b) The genetic effect assumes four functional alleles with balanced allele frequencies (25%). Mediation analysis was performed using (a) a variant that tags the two higher functional alleles and (b) a variant that tags only the highest functional allele. (c) Data from a bi-allelic variant with allele frequency 50% were simulated, and mediation analysis performed using 8 founder haplotype states. Tables describe the relationships of the structure of X used to simulate the data versus the structure of X used in the mediation analysis. Heat maps for Bayesian model selection represent the mean posterior probability associated with each inferred model for a range of fixed settings of the model parameters as indicated on x- and y-axes.

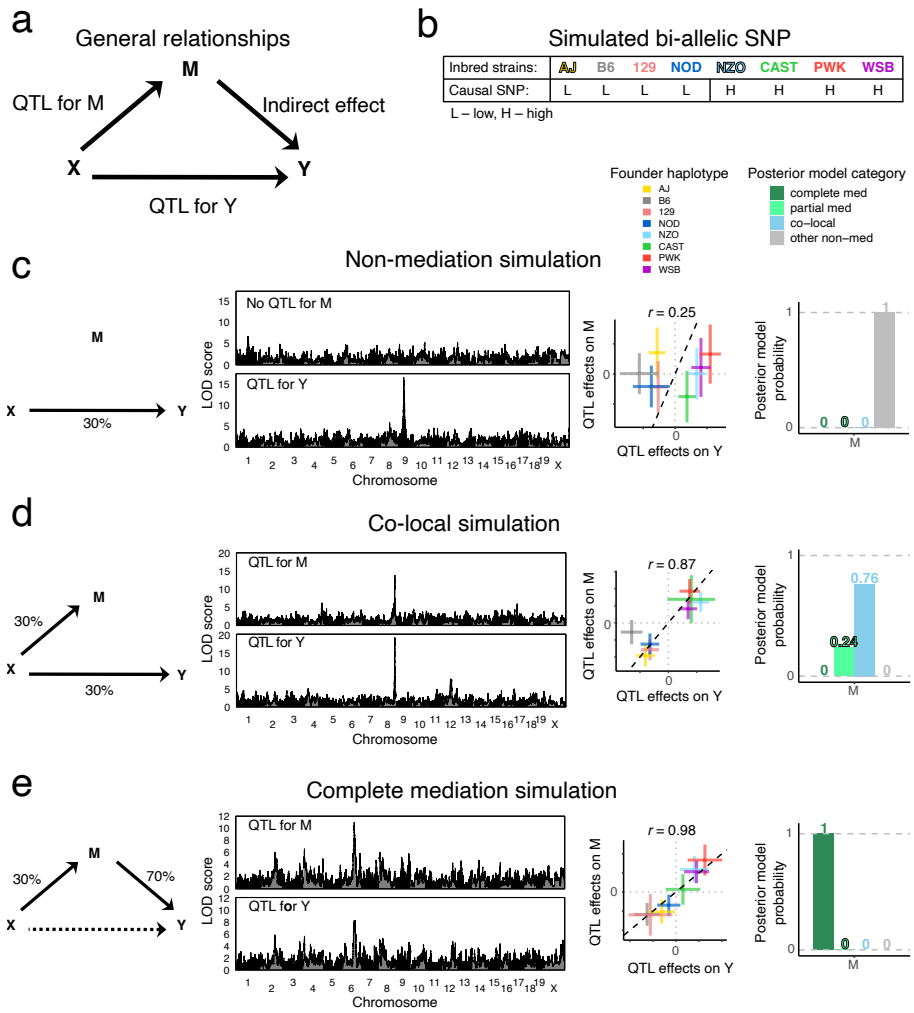

Figure 4 Illustration of Bayesian model selection applied to QTL mapping with simulated DO mouse data. (a) The DAG is labeled to indicate how each arm in the mediation model is interpreted in the QTL mapping setting. Y and M were simulated based on a bi-allelic QTL X at a randomly selected locus, with (b) each allele distributed to four founder strains. Genome-wide genotype data were obtained from 192 Diversity Outbred mice, according to one of three models: (c) M is a non-mediator of X on Y, (d) M and Y are independently driven by X (co-local), and (e) M is a complete mediator of X on Y, as illustrated with the corresponding model DAG (left). Genome-wide LOD scores for QTL mapping of M and Y, a scatter plot of the founder haplotype effects at the QTL for M and Y, and the Bayesian model selection posterior model probabilities are shown (from left to right).

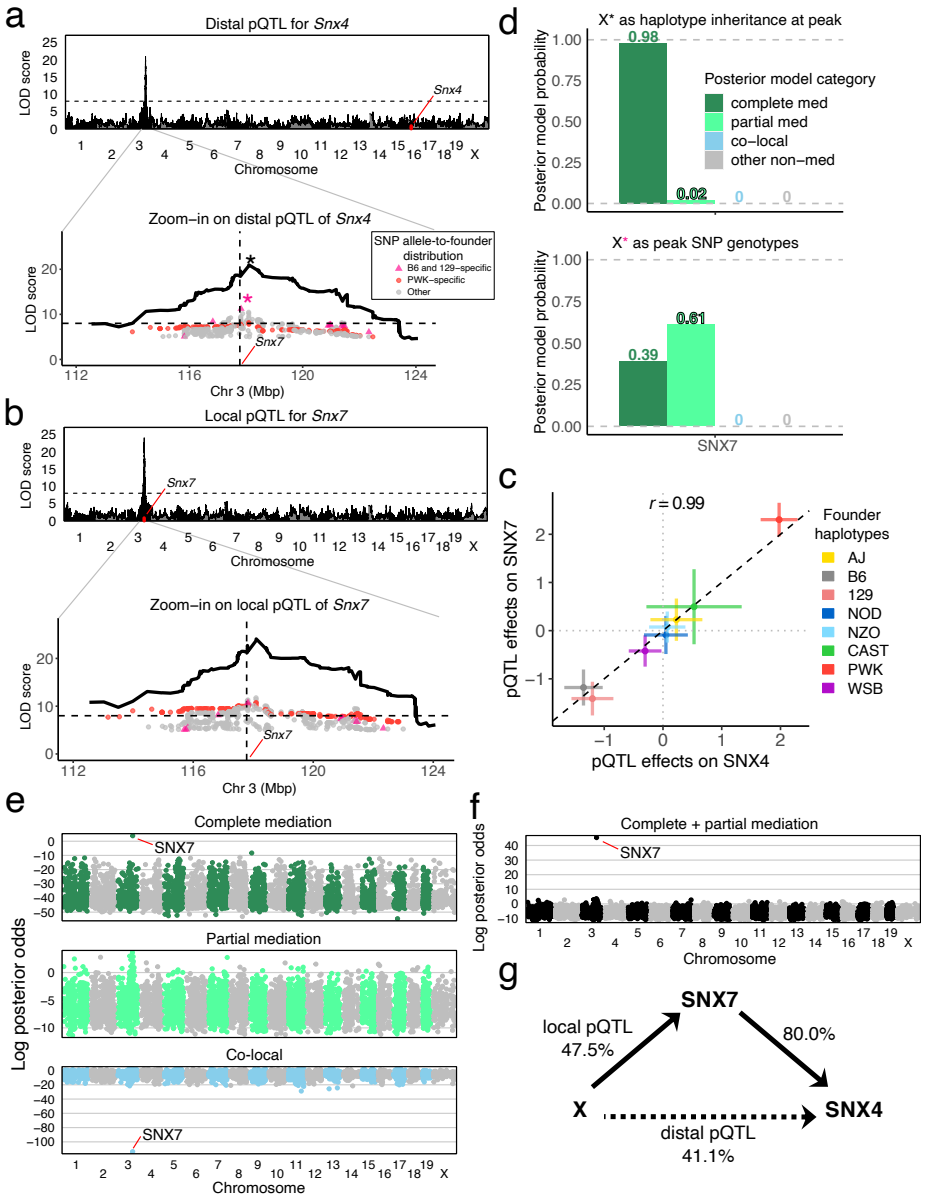

Figure 5 Mediation analysis of a distal pQTL for Snx4 in DO mice. Genome-wide LOD scores for associations of (a) SNX4 and (b) SNX7 abundance were performed using founder haploptye linkage mapping. Zooming into the QTL region, LOD scores for variant association within the pQTL region (peak $\pm$ 5 Mbp) for bi-allelic vartiants with LOD scores $>$ 5 are overlaid on the haplotype association LOD curve. Variants with alleles specific to B6 and 129 (pink) and PWK (red) are highlighted. (c) The founder haplotype effects at the pQTL are multi-allelic and highly similar for the two proteins. (d) Posterior probabilities of mediation models for the pQTL (top) using founder haplotypes and (bottom) using the peak bi-allelic variant. (e) Genome-wide mediation scan where all observed proteins are individually evaluated as mediators of the *Snx4* distal pQTL highlights SNX7 as a complete mediator and strongly indicates that the co-local model is unlikely. Each point represents the log posterior odds for a candidate mediator for the specified mediation model. (f) Summing complete and partial mediation posterior probabilties distinguishes SNX7 as the single best candidate mediator. (g) The complete mediation model with SNX7 as mediator of the *Snx4* distal pQTL is shown as a DAG with estimated effect sizes in units of percent variance explained. The dashed line indicates the strength of the distal pQTL that is not included in the model because it is completely mediated.

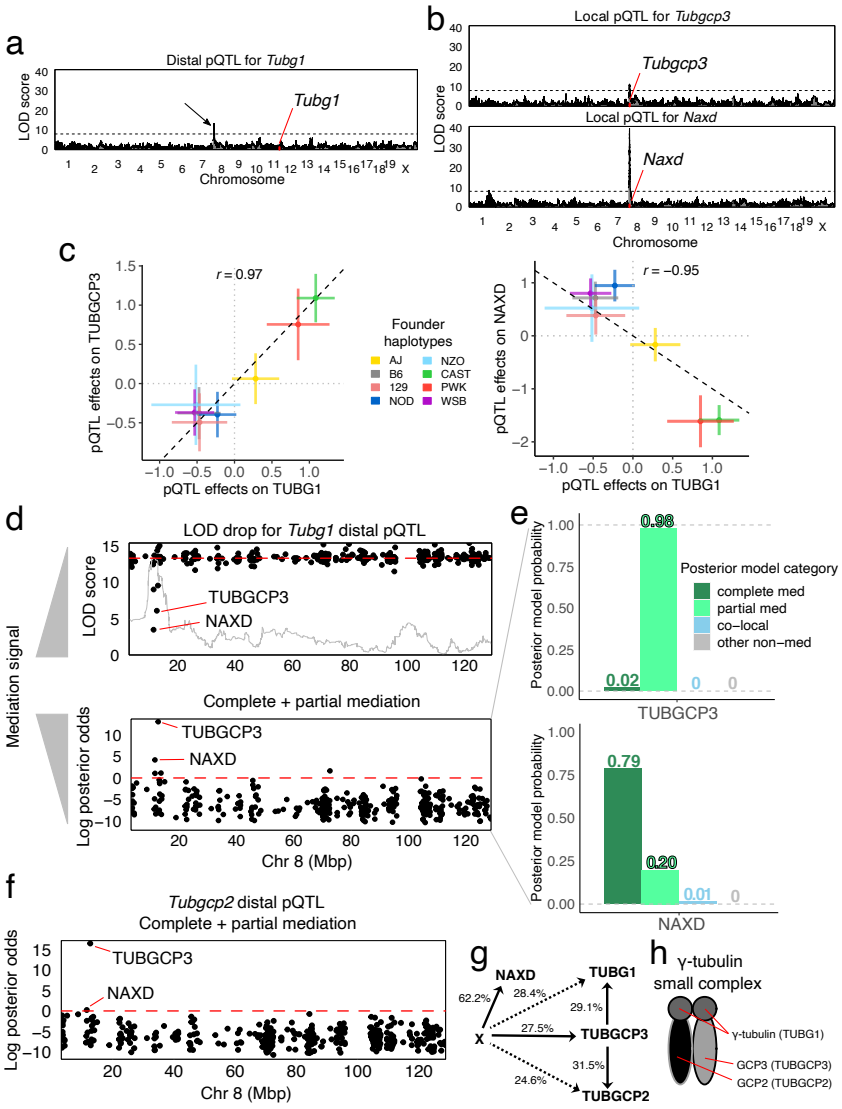

Figure 6 Mediation analysis of a distal pQTL for Tubg1 in DO mice. (a) Genome-wide LOD scores for TUBG1 abundance. Black arrow indicates distal pQTL on chromosome 8. (b) Genome-wide LOD scores for two genes *Tubgcp3* (top) and *Naxd* (bottom), with co-mapping local pQTL. (c) Comparison of the founder haplotype effects of the *Tubg1* pQTL with *Tubgcp3* (left) and *Naxd* (right) pQTL. (d) Mediation scans of all observed proteins on chromosome 8 by LOD drop with an overlay of the pQTL LOD scores in gray (top) and Bayesian model selection log posterior odds for mediation (bottom) show different prioritization for candidate mediators NAXD and TUBCP3. Note that low LOD drop scores indicate stronger mediation signal. (e) Posterior model probabilities for the *Tubg1* distal pQTL for candidate mediators (top) TUBGCP3 and (bottom) NAXD. (f) Mediation scan for the *Tubgcp2* distal pQTL identifies TUBGCP3 as the best candidate mediator. (g) The DAG summarizes the mediation analysis results with effect size estimates shown as percent variance explained. Dashed lines indicate the strength of distal pQTL effects that are not part of the model assuming complete mediation through TUBGCP3. (h) TUBG1, TUBGCP2, and TUBGCP3 comprise the γ-tubulin small complex.

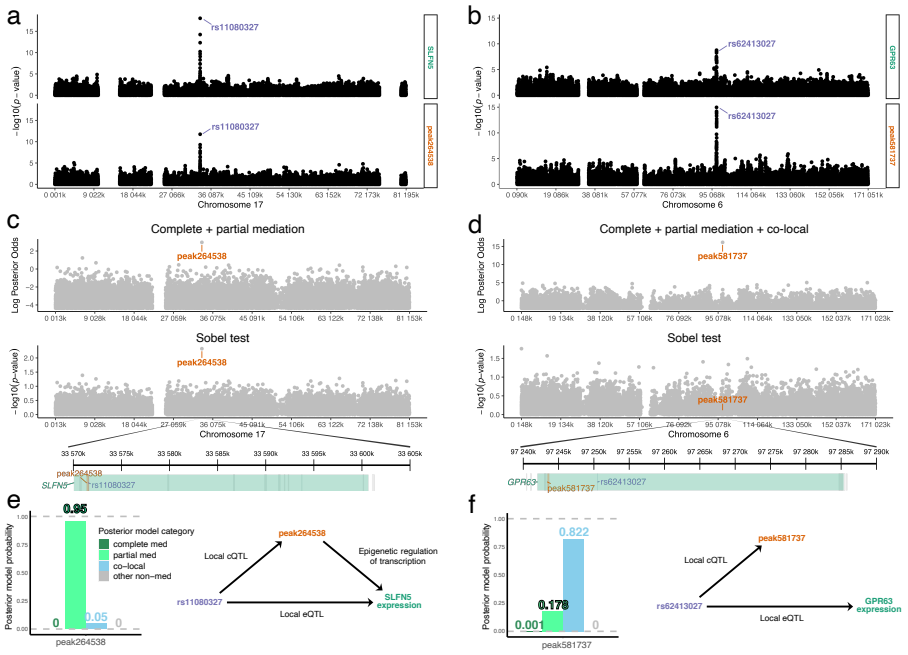

Figure 7 Mediation analysis of local chromatin state and gene expression data in human cell lines. SNP associations with (a) SLFN5 and (b) GPR63 expression (top) and nearby chromatin accessibility (bottom) for variants on the genes' chromosome. Peak SNPs are labeled. Mediation results for the (c) SLFN5 eQTL and (d) GPR63 eQTL. Log posterior odds from Bayesian model selection (top), -log10 p-values from Sobel test (middle), and zoomed-in window highlighting gene start, peak SNP, and peak mediator (bottom). Peak mediator or co-local chromatin peak is labeled. Each gray point represents a chromatin peak candidate mediator located near the gene of interest. For SLFN5 expression, complete and partial mediation models were summed in the posterior summary from Bayesian model selection. For GPR63 expression, the co-local model was also summed with the mediation models. Posterior model probabilities from Bayesian model selection for the peak mediator and co-local chromatin peaks and the implied DAG for the (e) SLFN5 eQTL and (f) GPR63 eQTL.

Methods

The standard mediation model, the Sobel test, and LOD drop

The standard mediation model can be described using two linked linear models. For individual $i = 1, \ldots, N$, let y_i be the value of the primary outcome or dependent variable, let m_i be the value of the mediator variable, and let x_i be a (scalar) independent variable. Mediation analysis frames the relationships between these variables in terms of the following regressions:

$$
\begin{aligned}
m_i &= \mu_m + x_i \boldsymbol{\beta}_a + \epsilon_i \\
y_i &= \mu_y + x_i \boldsymbol{\beta}_c + m_i \beta_b + \varepsilon_i \,,
\end{aligned}
$$

where μ_m and μ_y are intercepts, ϵ_i and ε_i are normally distributed, independent random noise variables, β_a is the effect of x_i on m_i, β_b is the effect of m_i on y_i, and β_c is the effect of x_i on y_i. The effects $\{\beta_a, \beta_b, \beta_c\}$ correspond to edges $\{a, b, c\}$ in the DAG in Figure 1. In the language of mediation, β_c describes the direct effect of x_i on y_i, whereas the combination of β_a and β_b describe the indirect effect of x_i on y_i via m_i.

The Sobel test for mediation focuses on the estimation of the product $\beta_a \beta_b$. This product, which provides a single number description of the indirect effect, can take any value from $-\infty$ to ∞ but is zero only when either $\beta_a = 0$ or $\beta_b = 0$ or both. The traditional implementation of the Sobel test uses classical estimates of $\beta_a \beta_b$, and bootstrapping or other approximations [47], to estimate confidence interval for $\beta_a \beta_b$ and thereby a p-value for $\beta_a \beta_b \neq 0$. Bayesian implementations of the Sobel test analogously determine a posterior for $\beta_a \beta_b$ and thereby a suitable tail probability [20] or Bayes factor [23]. In either implementation, the Sobel test provides a succinct way to quantify the evidence for mediation and uncertainty about the strength and direction of the indirect effect. By contrast, the CS approach [9], and thereby its approximation, the LOD drop method (both defined in the Introduction), is awkwardly constructed [48], lacks power [19, 49], and does not quantify uncertainty; it does, however, take into account the X $\rightarrow$ Y relationship, potentially enabling discrimination between partial and complete mediation. Note that distinguishing partial and complete mediation is sensitive to misspecification [46] and its value is subject to debate [50].

Mediation with multiple independent variables

The standard mediation model is trivially adapted to cases where the independent variable is multivariable, that is, of dimension $D > 1$: the scalar predictor x_i and scalar effects β_a and β_c are simply replaced by their D-vector counterparts, $\mathbf{x}_i$, $\boldsymbol{\beta}_a$, $\boldsymbol{\beta}_c$. In this case, the CS (or LOD drop) approach to testing mediation may still be applied but, because the product $\boldsymbol{\beta}_a \beta_b$ is a vector, the Sobel test cannot. Sobel test-like procedures have been developed for specific cases, including a bootstrap-based test for when X is multicategorical [43] and a bespoke Bayesian model for when X is a scalar-by-multi-category interaction [21], but these are computationally intensive and do not generalize easily to multivariable X.

A Bayesian model selection approach to mediation analysis

We propose a Bayesian model selection approach that combines the generality of the CS method with the inferential coherence of the Sobel test. Our approach is most similar to the one in Nuijten et al. 2015 [23], which is also a Bayesian formulation of mediation, but only considers a single independent variable in X. To describe our approach, we first elaborate the standard mediation model to

$$
\begin{aligned}
m_i &= \mu_{\mathbf{m}} + \mathbf{z}_{m,i}^\top \boldsymbol{\alpha}_m + \theta_a \mathbf{x}_i^\top \boldsymbol{\beta}_a + \epsilon_i \\
y_i &= \mu_{\mathbf{y}} + \mathbf{z}_{y,i}^\top \boldsymbol{\alpha}_m + \theta_c \mathbf{x}_i^\top \boldsymbol{\beta}_c + \theta_b m_i \beta_b + \varepsilon_i \,,
\end{aligned}
$$

where $\mathbf{x}_i$ is a D-length column vector, $\boldsymbol{\beta}_a$ and $\boldsymbol{\beta}_c$ are D-length effects vectors, $\mathbf{z}_{m,i}$ and $\mathbf{z}_{y,i}$ are column vectors encoding any covariates for m_i and y_i, $\boldsymbol{\alpha}_m$ and $\boldsymbol{\alpha}_{\mathbf{y}}$ are the corresponding covariate effects, $\epsilon_i \sim \mathrm{N}(0, \sigma^2_{\mathbf{m}}/w_{\mathbf{m},i})$ and $\varepsilon_i \sim \mathrm{N}(0, \sigma^2_{\mathbf{y}}/w_{\mathbf{y},i})$ are noise variables with individual-specific weights $w_{\mathbf{m},i}$ and $w_{\mathbf{y},i}$, and $\boldsymbol{\theta} = \{\theta_a, \theta_b, \theta_c\} \in \{0,1\}^3$ are indicator variables denoting the presence or absence of edges $\{a, b, c\}$ in the DAG in Figure 1, such that, for example, $\boldsymbol{\theta} = \{1, 1, 0\}$ denotes complete mediation (a and b active) with no effect of edge c. The parameter space of $\boldsymbol{\theta}$ contains $2^3 = 8$ possible combinations of edges, each of which corresponds to a particular causal model. From a Bayesian perspective, making inferences about the identity of the true causal model is equivalent to calulating the posterior distribution of $\boldsymbol{\theta}$, that is,

$$
\mathrm{p}(\boldsymbol{\theta}|\mathbf{y}, \mathbf{m}) \propto \mathrm{p}(\mathbf{y}, \mathbf{m}|\boldsymbol{\theta})\mathrm{p}(\boldsymbol{\theta}),
$$

where $\mathbf{y} = \{y_i\}_{i=1}^N$ and $\mathbf{m} = \{m_i\}_{i=1}^N$. This involves calculating a joint likelihood for Y and M given a causal model, $\mathrm{p}(\mathbf{y}, \mathbf{m}|\boldsymbol{\theta})$, and a prior distribution over causal models, $\mathrm{p}(\boldsymbol{\theta})$. These are described separately below.

Likelihood and priors conditional on a causal model

The conditional joint likelihood function is given by

$$
\begin{aligned}
\mathbf{m}|\theta_a, \mu_{\mathbf{m}}, \boldsymbol{\alpha}_{\mathbf{m}}, \boldsymbol{\beta}_a, \sigma_{\mathbf{m}} &\sim \mathrm{N}(\mu_{\mathbf{m}}\mathbf{1} + \mathbf{Z}_{\mathbf{m}}\boldsymbol{\alpha}_{\mathbf{m}} + \theta_a \mathbf{X}\boldsymbol{\beta}_a, \sigma^2_{\mathbf{m}}\mathbf{W}_{\mathbf{m}}^{-1}), \\
\mathbf{y}|\mathbf{m}, \theta_b, \theta_c, \mu_{\mathbf{y}}, \boldsymbol{\alpha}_{\mathbf{y}}, \beta_b, \boldsymbol{\beta}_c, \sigma_{\mathbf{y}} &\sim \mathrm{N}(\mu_{\mathbf{y}}\mathbf{1} + \mathbf{Z}_{\mathbf{y}}\boldsymbol{\alpha}_{\mathbf{y}} + \theta_c \mathbf{X}\boldsymbol{\beta}_c + \theta_b \mathbf{m}\beta_b, \sigma^2_{\mathbf{y}}\mathbf{W}_{\mathbf{y}}^{-1}),
\end{aligned}
$$

where $\mathbf{Z}_{\mathbf{m}}$ and $\mathbf{Z}_{\mathbf{y}}$ are matrices for the covariates, $\mathbf{W}_{\mathbf{m}} = \mathrm{diag}(\{w_{\mathbf{m},i}\}_{i=1}^N)$ and $\mathbf{W}_{\mathbf{y}} = \mathrm{diag}(\{w_{\mathbf{y},i}\}_{i=1}^N)$ are diagonal matrices of observation weights. By default, $\mathbf{Z}_{\mathbf{m}} = \mathbf{Z}_{\mathbf{y}} = \mathbf{0}$, and $\mathbf{W}_{\mathbf{m}} = \mathbf{W}_{\mathbf{y}} = \mathbf{I}$.

Conjugate priors are given for the following variables:

$$
\begin{aligned}
\mu_{\mathbf{m}}|\sigma_{\mathbf{m}} &\sim \mathrm{N}(0, \sigma^2_{\mathbf{m}}\tau^2_{\mu_{\mathbf{m}}}) & \mu_{\mathbf{y}}|\sigma_{\mathbf{y}} &\sim \mathrm{N}(0, \sigma^2_{\mathbf{y}}\tau^2_{\mu_{\mathbf{y}}}) \\
\boldsymbol{\alpha}_{\mathbf{m}}|\sigma_{\mathbf{m}} &\sim \mathrm{N}(\mathbf{0}, \sigma^2_{\mathbf{m}}\tau^2_{\mathbf{Z}_{\mathbf{m}}}\mathbf{I}) & \boldsymbol{\alpha}_{\mathbf{y}}|\sigma_{\mathbf{y}} &\sim \mathrm{N}(\mathbf{0}, \sigma^2_{\mathbf{y}}\tau^2_{\mathbf{Z}_{\mathbf{y}}}\mathbf{I}) \\
\boldsymbol{\beta}_a|\sigma_{\mathbf{m}} &\sim \mathrm{N}(\mathbf{0}, \sigma^2_{\mathbf{m}}\phi^2_{\mathbf{a}}\mathbf{I}) & \boldsymbol{\beta}_c|\sigma_{\mathbf{y}} &\sim \mathrm{N}(\mathbf{0}, \sigma^2_{\mathbf{y}}\phi^2_{\mathbf{c}}\mathbf{I}) \\
& & \beta_b|\sigma_{\mathbf{y}} &\sim \mathrm{N}(0, \sigma^2_{\mathbf{y}}\phi^2_{\mathbf{b}}) \\
\sigma_{\mathbf{m}}^{-2} &\sim \mathrm{Ga}(0.5\kappa_{\mathbf{m}}, 0.5\lambda_{\mathbf{m}}) & \sigma_{\mathbf{y}}^{-2} &\sim \mathrm{Ga}(0.5\kappa_{\mathbf{y}}, 0.5\lambda_{\mathbf{y}})
\end{aligned}
$$

These variables can be integrated from the likelihood due to conjugacy, giving a closed form expression for the marginal joint likelihood function:

$$\begin{aligned}\mathbf{m}|\theta_a &\sim \mathrm{t}_{\kappa_\mathbf{m}}(\mathbf{0}, \lambda_\mathbf{m}[\mathbf{W}^{-1} + \mathcal{X}_\mathbf{m}\mathbf{V}_\mathbf{m}\mathcal{X}_\mathbf{m}^T]) \\ \mathbf{y}|\mathbf{m}, \theta_b, \theta_c &\sim \mathrm{t}_{\kappa_\mathbf{y}}(\mathbf{0}, \lambda_\mathbf{y}[\mathbf{W}^{-1} + \mathcal{X}_\mathbf{y}\mathbf{V}_\mathbf{y}\mathcal{X}_\mathbf{y}^T])\end{aligned}$$

where $\mathcal{X}_m$ and $\mathcal{X}_\mathbf{y}$ are concatenated design matrices and $\mathbf{V}_\mathbf{m}$ and $\mathbf{V}_\mathbf{y}$ are prior covariance matrices for the effect variables, specifically,

$$\mathcal{X}_\mathbf{m} = \begin{bmatrix}\mathbf{1} & \mathbf{Z}_\mathbf{m} & \theta_a\mathbf{X}\end{bmatrix} \qquad \mathbf{V}_\mathbf{m} = \begin{bmatrix}\tau_\mu^2 & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \tau_\mathbf{Z}^2\mathbf{I} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \phi_a^2\mathbf{I}\end{bmatrix}$$

$$\mathcal{X}_\mathbf{y} = \begin{bmatrix}\mathbf{1} & \mathbf{Z}_\mathbf{y} & \theta_c\mathbf{X} & \theta_b\mathbf{m}\end{bmatrix} \qquad \mathbf{V}_\mathbf{y} = \begin{bmatrix}\tau_\mu^2 & \mathbf{0} & \mathbf{0} & 0 \\ \mathbf{0} & \tau_\mathbf{Z}^2\mathbf{I} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \phi_c^2\mathbf{I} & \mathbf{0} \\ 0 & \mathbf{0} & \mathbf{0} & \phi_b^2\end{bmatrix}$$

This marginal joint likelihood is evaluated for all causal models $\boldsymbol{\theta} = (\theta_a, \theta_b, \theta_c)$, given prior hyperparameters $\boldsymbol{\kappa} = (\kappa_\mathbf{m}, \kappa_\mathbf{y})$, $\boldsymbol{\lambda} = (\lambda_\mathbf{m}, \lambda_\mathbf{y})$, $\boldsymbol{\tau}_\mu^2 = (\tau_{\mu_\mathbf{m}}^2, \tau_{\mu_\mathbf{y}}^2)$, $\boldsymbol{\tau}_\mathbf{Z}^2 = (\tau_{\mathbf{Z}_\mathbf{m}}^2, \tau_{\mathbf{Z}_\mathbf{y}}^2)$, and $\boldsymbol{\phi}^2 = (\phi_a^2, \phi_b^2, \phi_c^2)$. Non-informative priors are used for the scale of the data [$\boldsymbol{\kappa} = \boldsymbol{\lambda} = (0.001, 0.001)$], the location of the intercept [$\boldsymbol{\tau}_\mu = (1000, 1000)$], and covariate effects [$\boldsymbol{\tau}_\mathbf{Z} = (1000, 1000)$]. The hyperparameter $\boldsymbol{\phi}^2$ controls the prior effect size of each edge, relative to error, on $\mathbf{m}$ or $\mathbf{y}$. We set $\boldsymbol{\phi}^2 = (1, 1, 1)$ by default, such that effect sizes for all edges are equal and relatively large *a priori*.

Priors for the causal models

In our framework, a prior probability is assigned to each possible causal model. The default model prior places uniform prior probability across the eight models for which edge b is either absent or present as M→Y (no reactive models). That is,

$$\text{Default:} \quad \mathrm{p}(\boldsymbol{\theta}) \propto 1 \quad \text{for} \quad \boldsymbol{\theta} \in \{0, 1\}^3\,,$$

such that the prior probability for each of ML1-8 in Figure 1 is $\frac{1}{8}$. We also consider two other prior specifications: the reduced model prior, which additionally assumes that edge c is present, that is,

$$\text{Reduced:} \quad \mathrm{p}(\boldsymbol{\theta}) \propto 1 \quad \text{for} \quad (\theta_a, \theta_b) \in \{0, 1\}^2 \quad \text{and} \quad \theta_c = 1\,,$$

such that the prior probability for each of ML4-8 is $\frac{1}{4}$; and the expanded model prior, which allows models with a reversed edge b (*i.e.*, Y→M) and is described in more detail below. In our framework and using our bmediatR software, it is possible to specify priors as any distinct set of allowable causal models, or, with greater granularity, as relative weights across the set of models. This feature provides a foundation that could accommodate empirically-determined prior weights out-of-the-box.

Allowing reverse causality between M and Y

The bmediatR software can also model reverse causality between the candidate mediator and trait (*i.e.*, Y→M). This can be described using a third state for θ_b, denoted by $\theta_b = *$, giving a total of twelve possible causal models encoded by $\boldsymbol{\theta}$. These include the eight causal models previously described, as well as four additional reverse causality (or "reactive") cases, given by $(\theta_a, \theta_b = *, \theta_c)$, where models with $\theta_b = *$ are specified with the roles of all Y- and M-related variable above interchanged (explicit formulae given in Supplemental Methods). Although any set of relative prior probabilities could be used, we here consider a simple prior (the expanded prior) that is uniform over the 12 causal models, that is,

$$\text{Expanded:} \quad \mathrm{p}(\boldsymbol{\theta}) \propto 1 \quad \text{for} \quad (\theta_a, \theta_c) \in \{0,1\}^2 \quad \text{and} \quad \theta_b \in \{0,1,*\}\,,$$

such that the prior probability for each of ML1-12 in Figure 1 is $\frac{1}{12}$.

Note that two pairs of these twelve causal models are not identifiable from one another: $\boldsymbol{\theta} = (0,1,0)$ is indistinguishable from $\boldsymbol{\theta} = (0,*,0)$; and $\boldsymbol{\theta} = (1,1,1)$ (partial mediation) is indistinguishable from $\boldsymbol{\theta} = (1,*,1)$ (reactive partial mediation). The marginal joint likelihood functions for these pairs of models are identical, and these causal relationships cannot be distinguished without additional information. Despite this limitation, we include all twelve causal relationships as possibilities in our expanded model prior, and emphasize that inference about the direction of causality in these cases depends entirely on prior information and assumptions.

Simulation of QTL data

To simulate data with an exogenous variable representing genetic variation (*e.g.*, SNPs, founder haplotypes), we expanded our previous approach for simulating QTL in the Collaborative Cross mouse population [51]. The model describes relationships between two traits with shared genetic drivers, representing either co-local or mediation.

We simulate a single trait with a single QTL based on a simple linear model:

$$\mathbf{y}_{\text{sim}} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon} \tag{1}$$

where $\mathbf{y}$ is the trait vector, $\mathbf{X}$ is the design matrix of the genetic information (*e.g.*, SNP allele count, founder haplotype count), $\boldsymbol{\beta}$ is the genetic effects vector (*e.g.*, SNP allele effect, haplotype effects), and $\boldsymbol{\varepsilon}$ is a random noise vector. We scale $\boldsymbol{\beta}$ and $\boldsymbol{\varepsilon}$ such that their relative contributions to $\mathbf{y}_{\text{sim}}$ match a specified proportion of variation of $\mathbf{y}_{\text{sim}}$ explained by $\mathbf{X}\boldsymbol{\beta}$, *i.e.*, the effect size of X on Y. If an initial genetic effect vector ($\boldsymbol{\beta}_{\text{raw}}$) is not specified, we sample one according to $\boldsymbol{\beta}_{\text{raw}} \sim \mathrm{N}(\mathbf{0}, \mathbf{I})$, where $\mathbf{I}$ is the identity matrix of rank equal to length of $\boldsymbol{\beta}$. An initial noise vector is also sampled as $\boldsymbol{\varepsilon}_{\text{raw}} \sim \mathrm{N}(\mathbf{0}, \mathbf{I}_n)$ where $\mathbf{I}_n$ is the identity matrix of rank n, the number of rows of $\mathbf{X}$. For a specified effect size ϕ^2 (ranging from 0 to 1), we define scaling factors for $\boldsymbol{\beta}_{\text{raw}}$ and $\boldsymbol{\varepsilon}_{\text{raw}}$ to approximate the desired effect size of X on Y:

$$s_\beta = \left(\frac{n\phi^2}{(n-1)V(\mathbf{X}\boldsymbol{\beta}_{\text{raw}})}\right)^{\frac{1}{2}} \qquad s_\varepsilon = \left(\frac{n(1-\phi^2)}{(n-1)V(\boldsymbol{\varepsilon}_{\text{raw}})}\right)^{\frac{1}{2}}$$

where $V()$ returns the population variance of its argument. The effects and noise vectors are then scaled: $\boldsymbol{\beta} = s_\beta \boldsymbol{\beta}_{\text{raw}}$ and $\boldsymbol{\varepsilon} = s_\varepsilon \boldsymbol{\varepsilon}_{\text{raw}}$.

For the co-local simulations, both the dependent variable $\mathbf{y}_{\text{sim}}$ and the mediator variable $\mathbf{m}_{\text{sim}}$ are simulated as in Equation 1 with the same $\boldsymbol{\beta}_{\text{raw}}$ but independent draws of $\boldsymbol{\varepsilon}_{\text{raw}}$. For the complete mediation simulations, $\mathbf{m}_{\text{sim}}$ is simulated according to Equation 1, and the dependent variable $\mathbf{y}_{\text{sim}}$ is simulated based on a linear model with the mediator as a predictor:

$$\mathbf{y}_{\text{sim}} = \mathbf{m}_{\text{sim}} \boldsymbol{\beta}_{\mathbf{m}} + \boldsymbol{\epsilon}$$

where $\boldsymbol{\beta}_{\mathbf{m}}$ and $\boldsymbol{\epsilon}$ are derived from scaled raw variables just as $\boldsymbol{\beta}$ and $\boldsymbol{\varepsilon}$ were in Equation 1 in order to strictly control how M contributes to variation in Y (*i.e.*, effect size of M on Y).

The partial mediation simulations are necessarily more complicated. The mediator $\mathbf{m}_{\text{sim}}$ is again simulated according to Equation 1. The linear model for $\mathbf{y}_{\text{sim}}$ includes effects from both X and M:

$$\mathbf{y}_{\text{sim}} = \mathbf{X}\boldsymbol{\beta} + \mathbf{m}_{\text{sim}} \boldsymbol{\beta}_{\mathbf{m}} + \boldsymbol{\epsilon}$$

where $\boldsymbol{\beta}$ is the direct effect of X on Y and $\boldsymbol{\beta}_{\mathbf{m}}$ is the indirect effect. $\boldsymbol{\beta}$, $\boldsymbol{\beta}_{\mathbf{m}}$, and $\boldsymbol{\epsilon}$ were derived from scaled raw variables as before in order to strictly control how X and M together contribute to variation in Y (*i.e.*, combined effect size on Y). Reactive partial and reactive complete mediation were simulated exactly as partial and complete mediation, but with $\mathbf{y}_{\text{sim}}$ and $\mathbf{m}_{\text{sim}}$ swapped in the mediation procedures.

For the large-scale simulations of a single local (Figures 2, 3, S1, S2), $\mathbf{X}$ represented balanced functional allele counts (SNP or founder haplotypes) for 200 individuals. One hundred simulations were performed for each combination of effect size on M and Y (ranging from 0.05 to 0.95 at regular intervals of 0.05) for each data-generating model (co-local, partial mediation, complete mediation, reactive complete mediation, and reactive partial mediation). For the partial mediation simulations, the ratio of direct and indirect effect were fixed at 1 (b:c = 1). For each data-generating model, Bayesian model selection, Sobel test (if $\mathbf{X}$ represented a SNP), and LOD drop were summarized. For the simulations demonstrating multi-allelic QTL analysis with mediation in Figure 4, the design matrix $\mathbf{X}$ represents additive effects of founder haplotypes at a randomly sampled genetic locus from a population of 192 DO mice [5]. The underlying simulated genetic effect was from a bi-allelic SNP, with each allele present in four of the founder strains.

Diversity Outbred mouse data

The DO mouse data represent 192 animals [5] with both gene expression and protein abundance from bulk liver tissue, representing a subset of a larger cohort of 850 animals [52]. Approximately equal numbers of males and females are present as well as animals on standard and high fat diets. The data are publicly available for interactive analysis as a QTL Viewer (`https://github.com/churchill-lab/qtlapi`) which also allows a bulk download of the underlying R data files (`https://qtlviewer.jax.org/viewer/SvensonHFD`).

QTL analysis

Samples of liver tissue were collected and processed for quantitative mass-spectrometry as previously described [5]. Estimation and normalization of the protein abundance data from component quantitative peptide data and subsequent QTL analysis have been previously described [13]. Genetic mapping was based on final quantities output from the rank-based inverse normal transformation (RINT) [53].

The examples related to the distal pQTL of *Snx4* and *Tubg1* were previously identified [13]. QTL were mapped using the following linear mixed effect model:

$$\text{trait}_i = \text{intercept} + \text{covariates}_i + \text{QTL}_{i,m} + u_i + \varepsilon_i \tag{2}$$

where trait_i is the phenotype of interest (abuandance of a protein) for individual i, intercept is a shared intercept that represents the mean trait value, covariates_i is the effect of known covariates on individual i, $\text{QTL}_{i,m}$ is the effect of genetic variation at genomic interval m on individual i, u_i is a random error term that accounts for the similarity of individual i to other samples proportional to overall genetic relatedness (kinship effect), and ε_i is the unstructured error on individual i. For the QTL term: $\text{QTL}_{i,m} = \mathbf{d}_i^{\mathrm{T}} \boldsymbol{\beta}_{\text{QTL}}$, where $\mathbf{d}_i$ is the vector of additive dosages of founder haplotypes for the genomic interval m and $\boldsymbol{\beta}_{\text{QTL}}$ are the haplotype effects at the putative QTL, estimated as fixed effects. The random error terms are modeled as $\mathbf{u} \sim \text{N}(\mathbf{0}, \mathbf{K}\tau^2)$ and $\varepsilon_i \sim \text{N}(0, \sigma^2)$, where $\mathbf{K}$ is the realized genetic relationship matrix excluding information from the chromosome of the current genomic interval m ("loco" method), and τ^2 and σ^2 are variance components for u_i and ε_i, respectively. Covariates adjusted for include sex, diet, and DO litter (2 levels). Equation 2 is fit at genomic intervals spanning the entire chromosome, comprising a QTL genome scan. Models were fit using the qtl2 R package [54]. For the variant association performed for the *Snx4* and *Snx7* pQTL, Equation 2 was used, but with the $\text{QTL}_{i,m}$ adjusted. Instead of representing the effects of doses of founder haplotypes, variant allele dosages were imputed based on the founder haplotype dosages and the variant genotype-to-founder strains distribution (SQLite variant database: https://doi.org/10.6084/m9.figshare.5280229.v3).

Haplotype effect estimation

To compare the similarity of the genetic effects on M and Y, the correlation coefficients between haplotype effects of QTL were calculated. To stabilize haplotype effect estimates, instead of the fixed effect estimate from Equation 2, they were estimated as best linear unbiased predictions (BLUPs) in the qtl2 R package.

Modeling the allelic series of QTL with TIMBR

We modeled the allelic series at the pQTL for to SNX4 and TUBG1 using a Bayesian hierarchical model as implemented in TIMBR [25]. The model is roughly equivalent to Equation 2, with the kinship effect excluded in order to make the computation feasible. A Chinese restaurant process prior was used for the allelic series, as well as the prescribed shape and rate parameters of 1 and ≈2.33, respectively, for the concentration parameter, which favors smaller numbers of functional alleles with low variance.

TIMBR models founder haplotype uncertainty at the QTL based on the 36 genotype states. We reconstructed founder haplotype probabilities from the available genotype array data (https://www.jax.org/research-and-faculty/genetic-diversity-initiative/tools-data/diversity-outbred-database), which represented 187 of the 192 mice.

Human cell line data

The genotype data from 119 Yoruba LCLs represent variants from the intersection of HapMap2 and HapMap3, coded as the major allele count for each SNP [28]. The RNA-seq data [29] were adjusted with WASP [55] and normalized [30], representing 69 LCLs [28]. Both genotype and RNA-seq data are publicly available for download (http://eqtl.uchicago.edu/jointLCL/). The DNase-seq data for the 69 cell lines with RNA-seq data were used as previously processed [31]. The overlap of samples across genotypes, RNA-seq, and DNase-seq was 63 LCLs.

Chromatin accessibility and expression QTL analysis

To identify candidates for co-mapping eQTL and cQTL, we correlated the expression of gene with chromatin peaks within 80 Kbp of the gene start. QTL mapping was then performed for genes and chromatin peaks that had correlations > 0.5. Mapping was performed by regressing a trait (gene expression or chromatin accessibility) onto the genotypes of individual SNPs located on the gene's chromosome, and compared to a null model with no genotype term. QTL were called for SNPs that produced a -log10 p-value > 8, representing a stringent threshold for 63 individuals. Gene-chromatin peak pairs were then filtered to those with an eQTL and cQTL co-mapping to the same SNP. For each passing gene, all chromatin peaks on the gene's chromosome were tested as candidate mediators using both Bayesian model selection, without default prior settings, and the Sobel test.

Availability of data and materials

All data and R code used to generate the results are available at figshare (https://doi.org/10.6084/m9.figshare.14912484).

Acknowledgments

We thank Michael I. Love of the University of North Carolina at Chapel Hill for discussions related to this work and for pointing us to publicly available human cell line data.

Authors' contributions

WLC, GRK, and GAC conceived the project. GAC and WV supervised and directed the research. WLC developed and implemented the methodology. WLC and GRK developed the software package. GRK and MSG analyzed data and visualized results for the manuscript. WLC, GRK, and MSG drafted the manuscript. GAC and WV edited the manuscript. All authors read and approved the manuscript before review.

Funding

This work was supported by grants from the National Institutes of Health (NIH): T32 Toxicology Training Grant (5T32ES007126-33) to WLC, F32GM134599 to GRK, R01GM070683 to GAC, and R35GM127000 to WV.

Competing interests

The authors declare that they have no competing interests.

Author details

[1]Department of Genetics, University of North Carolina, 27599 Chapel Hill, NC, US. [2]The Jackson Laboratory, 600 Main Street, 04609 Bar Harbor, ME, US. [3]Lineberger Comprehensive Cancer Center, University of North Carolina, 27599 Chapel Hill, NC, US.

References

1. Judd, C.M., Kenny, D.A.: Data analysis in social psychology: Recent and recurring issues. In: Fiske, S.T., Gilbert, D.T., Lindzey, G. (eds.) Handbook of Social Psychology. American Cancer Society, Hoboken, N.J. (2010). Chap. 4. https://onlinelibrary.wiley.com/doi/pdf/10.1002/9780470561119.socpsy001004. doi:10.1002/9780470561119.socpsy001004. https://onlinelibrary.wiley.com/doi/abs/10.1002/9780470561119.socpsy001004
2. MacKinnon, D.P., Fairchild, A.J., Fritz, M.S.: Mediation Analysis. Annual Review of Psychology **58**(1), 593–614 (2007). doi:10.1146/annurev.psych.58.110405.085542
3. Raulerson, C.K., Ko, A., Kidd, J.C., Currin, K.W., Brotman, S.M., Cannon, M.E., Wu, Y., Spracklen, C.N., Jackson, A.U., Stringham, H.M., Welch, R.P., Fuchsberger, C., Locke, A.E., Narisu, N., Lusis, A.J., Civelek, M., Furey, T.S., Kuusisto, J., Collins, F.S., Boehnke, M., Scott, L.J., Lin, D.-Y., Love, M.I., Laakso, M., Pajukanta, P., Mohlke, K.L.: Adipose Tissue Gene Expression Associations Reveal Hundreds of Candidate Genes for Cardiometabolic Traits. The American Journal of Human Genetics **105**(4), 773–787 (2019). doi:10.1016/j.ajhg.2019.09.001
4. Yao, D.W., O'Connor, L.J., Price, A.L., Gusev, A.: Quantifying genetic effects on disease mediated by assayed gene expression levels. Nature Genetics **52**(6), 626–633 (2020). doi:10.1038/s41588-020-0625-2
5. Chick, J.M., Munger, S.C., Simecek, P., Huttlin, E.L., Choi, K., Gatti, D.M., Raghupathy, N., Svenson, K.L., Churchill, G.A., Gygi, S.P.: Defining the consequences of genetic variation on a proteome-wide scale. Nature **534**(7608), 500–5 (2016). doi:10.1038/nature18270. 15334406
6. Keele, G.R., Quach, B.C., Israel, J.W., Chappell, G.A., Lewis, L., Safi, A., Simon, J.M., Cotney, P., Crawford, G.E., Valdar, W., Rusyn, I., Furey, T.S.: Integrative QTL analysis of gene expression and chromatin accessibility identifies multi-tissue patterns of genetic regulation. PLOS Genetics **16**(1), 1008537 (2020). doi:10.1371/journal.pgen.1008537
7. Judd, C.M., Kenny, D.A.: Process analysis: Estimating mediation in treatment evaluations. Evaluation Review **5**(5), 602–619 (1981). doi:10.1177/0193841X8100500502. https://doi.org/10.1177/0193841X8100500502
8. Pearl, J.: Interpretation and identification of causal mediation. Psychological Methods **19**(4), 459–481 (2014). doi:10.1037/a0036434
9. Baron, R.M., Kenny, D.A.: The moderator–mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. Journal of Personality and Social Psychology **51**(6), 1173–1182 (1986). doi:10.1037/0022-3514.51.6.1173. 4
10. Sobel, M.E.: Asymptotic Confidence Intervals for Indirect Effects in Structural Equation Models. Sociological Methodology **13**(1982), 290 (1982). doi:10.2307/270723
11. Preacher, K.J., Hayes, A.F.: Asymptotic and resampling strategies for assessing and comparing indirect effects in multiple mediator models. Behavior Research Methods **40**(3), 879–891 (2008). doi:10.3758/BRM.40.3.879
12. Collaborative Cross Consortium: The genome architecture of the Collaborative Cross mouse genetic reference population. Genetics **190**(2), 389–401 (2012). doi:10.1534/genetics.111.132639
13. Keele, G.R., Zhang, T., Pham, D.T., Vincent, M., Bell, T.A., Hock, P., Shaw, G.D., Munger, S.C., de Villena, F.P.-M., Ferris, M.T., Gygi, S.P., Churchill, G.A.: Regulation of protein abundance in genetically diverse mouse populations. bioRxiv (2020). doi:10.1101/2020.09.18.296657. https://www.biorxiv.org/content/early/2020/09/24/2020.09.18.296657.full.pdf
14. Churchill, G., Gatti, D., Munger, S., Svenson, K.: The Diversity outbred mouse population. Mammalian Genome **23**, 713–8 (2012). doi:10.1007/s00335-012-9414-2.The
15. Keller, M.P., Gatti, D.M., Schueler, K.L., Rabaglia, M.E., Stapleton, D.S., Simecek, P., Vincent, M., Allen, S., Broman, A.T., Bacher, R., Kendziorski, C., Broman, K.W., Yandell, B.S., Churchill, G.A., Attie, A.D.: Genetic Drivers of Pancreatic Islet Function. Genetics **209**(1), 335–356 (2018). doi:10.1534/genetics.118.300864
16. Skelly, D.A., Czechanski, A., Byers, C., Aydin, S., Spruce, C., Olivier, C., Choi, K., Gatti, D.M., Raghupathy, N., Keele, G.R., Stanton, A., Vincent, M., Dion, S., Greenstein, I., Pankratz, M., Porter, D.K., Martin, W., O'Connor, C., Qin, W., Harrill, A.H., Choi, T., Churchill, G.A., Munger, S.C., Baker, C.L., Reinholdt, L.G.: Mapping the Effects of Genetic Variation on Chromatin State and Gene Expression Reveals Loci That Control Ground State Pluripotency. Cell Stem Cell, 1–11 (2020). doi:10.1016/j.stem.2020.07.005
17. Zhong, W., Spracklen, C.N., Mohlke, K.L., Zheng, X., Fine, J., Li, Y.: Multi-snp mediation intersection-union test. Bioinformatics **35**(22), 4724–4729 (2019). doi:10.1093/bioinformatics/btz285
18. Liu, Z., Shen, J., Barfield, R., Schwartz, J., Baccarelli, A.A., Lin, X.: Large-scale hypothesis testing for causal mediation effects with applications in genome-wide epigenetic studies. Journal of the American Statistical Association, 1–39 (2021). doi:10.1080/01621459.2021.1914634
19. MacKinnon, D.P., Lockwood, C.M., Hoffman, J.M., West, S.G., Sheets, V.: A comparison of methods to test mediation and other intervening variable effects. Psychological methods **7**(1), 83–104 (2002). doi:10.1037/1082-989x.7.1.83. NIHMS150003
20. Yuan, Y., MacKinnon, D.P.: Bayesian mediation analysis. Psychological Methods **14**(4), 301–322 (2009). doi:10.1037/a0016972. NIHMS150003
21. Oreper, D., Schoenrock, S.A., McMullan, R., Ervin, R., Farrington, J., Miller, D.R., de Villena, F.P.-M., Valdar, W., Tarantino, L.M.: Reciprocal F1 Hybrids of Two Inbred Mouse Strains Reveal Parent-of-Origin and Perinatal Diet Effects on Behavior and Expression. G3: Genes, Genomes, Genetics **8**(11), 3447–3468 (2018). doi:10.1534/g3.118.200135
22. Song, Y., Zhou, X., Zhang, M., Zhao, W., Liu, Y., Kardia, S.L.R., Roux, A.V.D., Needham, B.L., Smith, J.A., Mukherjee, B.: Bayesian shrinkage estimation of high dimensional causal mediation effects in omics studies. Biometrics **76**(3), 700–710 (2020). doi:10.1111/biom.13189
23. Nuijten, M.B., Wetzels, R., Matzke, D., Dolan, C.V., Wagenmakers, E.-J.: A default bayesian hypothesis test for mediation. Behavior Research Methods **47**(1), 85–97 (2015). doi:10.3758/s13428-014-0470-2
24. Worby, C.A., Dixon, J.E.: Sorting out the cellular functions of sorting nexins. Nature Reviews Molecular Cell Biology **3**(12), 919–931 (2002). doi:10.1038/nrm974

25. Crouse, W.L., Kelada, S.N.P., Valdar, W.: Inferring the Allelic Series at QTL in Multiparental Populations. Genetics **216**(4), 957–983 (2020). doi:10.1534/genetics.120.303393
26. Oakley, B.R., Paolillo, V., Zheng, Y.: γ-Tubulin complexes in microtubule nucleation and beyond. Molecular Biology of the Cell **26**(17), 2957–2962 (2015). doi:10.1091/mbc.E14-11-1514
27. Farache, D., Emorine, L., Haren, L., Merdes, A.: Assembly and regulation of γ-tubulin complexes. Open Biology **8**(3), 170266 (2018). doi:10.1098/rsob.170266
28. Li, Y.I., van de Geijn, B., Raj, A., Knowles, D.A., Petti, A.A., Golan, D., Gilad, Y., Pritchard, J.K.: RNA splicing is a primary link between genetic variation and disease. Science **352**(6285), 600–604 (2016). doi:10.1126/science.aad9417
29. Pickrell, J.K., Marioni, J.C., Pai, A.A., Degner, J.F., Engelhardt, B.E., Nkadori, E., Veyrieras, J.-B., Stephens, M., Gilad, Y., Pritchard, J.K.: Understanding mechanisms underlying human gene expression variation with RNA sequencing. Nature **464**(7289), 768–772 (2010). doi:10.1038/nature08872
30. Degner, J.F., Pai, A.A., Pique-Regi, R., Veyrieras, J.-B., Gaffney, D.J., Pickrell, J.K., De Leon, S., Michelini, K., Lewellen, N., Crawford, G.E., Stephens, M., Gilad, Y., Pritchard, J.K.: DNase I sensitivity QTLs are a major determinant of human expression variation. Nature **482**(7385), 390–394 (2012). doi:10.1038/nature10808
31. Grubert, F., Zaugg, J.B., Kasowski, M., Ursu, O., Spacek, D.V., Martin, A.R., Greenside, P., Srivas, R., Phanstiel, D.H., Pekowska, A., Heidari, N., Euskirchen, G., Huber, W., Pritchard, J.K., Bustamante, C.D., Steinmetz, L.M., Kundaje, A., Snyder, M.: Genetic Control of Chromatin States in Humans Involves Local and Distal Chromosomal Interactions. Cell **162**(5), 1051–1065 (2015). doi:10.1016/j.cell.2015.07.048
32. Mavrommatis, E., Fish, E.N., Platanias, L.C.: The schlafen family of proteins and their regulation by interferons. Journal of interferon & cytokine research : the official journal of the International Society for Interferon and Cytokine Research **33**(4), 206–210 (2013). doi:10.1089/jir.2012.0133
33. Leek, J.T., Storey, J.D.: Capturing Heterogeneity in Gene Expression Studies by Surrogate Variable Analysis. PLoS Genetics **3**(9), 161 (2007). doi:10.1371/journal.pgen.0030161
34. Katan, M.: APOUPOPROTEIN E ISOFORMS, SERUM CHOLESTEROL, AND CANCER. The Lancet **327**(8479), 507–508 (1986). doi:10.1016/S0140-6736(86)92972-7
35. Voight, B.F., Peloso, G.M., Orho-Melander, M., Frikke-Schmidt, R., Barbalic, M., Jensen, M.K., Hindy, G., Hólm, H., Ding, E.L., Johnson, T., Schunkert, H., Samani, N.J., Clarke, R., Hopewell, J.C., Thompson, J.F., Li, M., Thorleifsson, G., Newton-Cheh, C., Musunuru, K., Pirruccello, J.P., Saleheen, D., Chen, L., Stewart, A.F.R., Schillert, A., Thorsteinsdottir, U., Thorgeirsson, G., Anand, S., Engert, J.C., Morgan, T., Spertus, J., Stoll, M., Berger, K., Martinelli, N., Girelli, D., McKeown, P.P., Patterson, C.C., Epstein, S.E., Devaney, J., Burnett, M.-S., Mooser, V., Ripatti, S., Surakka, I., Nieminen, M.S., Sinisalo, J., Lokki, M.-L., Perola, M., Havulinna, A., de Faire, U., Gigante, B., Ingelsson, E., Zeller, T., Wild, P., de Bakker, P.I.W., Klungel, O.H., Maitland-van der Zee, A.-H., Peters, B.J.M., de Boer, A., Grobbee, D.E., Kamphuisen, P.W., Deneer, V.H.M., Elbers, C.C., Onland-Moret, N.C., Hofker, M.H., Wijmenga, C., Verschuren, W.M., Boer, J.M.A., van der Schouw, Y.T., Rasheed, A., Frossard, P., Demissie, S., Willer, C., Do, R., Ordovas, J.M., Abecasis, G.R., Boehnke, M., Mohlke, K.L., Daly, M.J., Guiducci, C., Burtt, N.P., Surti, A., Gonzalez, E., Purcell, S., Gabriel, S., Marrugat, J., Peden, J., Erdmann, J., Diemert, P., Willenborg, C., König, I.R., Fischer, M., Hengstenberg, C., Ziegler, A., Buysschaert, I., Lambrechts, D., Van de Werf, F., Fox, K.A., El Mokhtari, N.E., Rubin, D., Schrezenmeir, J., Schreiber, S., Schäfer, A., Danesh, J., Blankenberg, S., Roberts, R., McPherson, R., Watkins, H., Hall, A.S., Overvad, K., Rimm, E., Boerwinkle, E., Tybjaerg-Hansen, A., Cupples, L.A., Reilly, M.P., Melander, O., Mannucci, P.M., Ardissino, D., Siscovick, D., Elosua, R., Stefansson, K., O'Donnell, C.J., Salomaa, V., Rader, D.J., Peltonen, L., Schwartz, S.M., Altshuler, D., Kathiresan, S.: Plasma HDL cholesterol and risk of myocardial infarction: a mendelian randomisation study. The Lancet **380**(9841), 572–580 (2012). doi:10.1016/S0140-6736(12)60312-2
36. Didelez, V., Sheehan, N.: Mendelian randomization as an instrumental variable approach to causal inference. Statistical methods in medical research **16**(4), 309–30 (2007). doi:10.1177/0962280206077743
37. Davies, N.M., Holmes, M.V., Davey Smith, G.: Reading Mendelian randomisation studies: a guide, glossary, and checklist for clinicians. BMJ **362**, 601 (2018). doi:10.1136/bmj.k601
38. Giambartolomei, C., Vukcevic, D., Schadt, E.E., Franke, L., Hingorani, A.D., Wallace, C., Plagnol, V.: Bayesian Test for Colocalisation between Pairs of Genetic Association Studies Using Summary Statistics. PLOS Genetics **10**(5), 1–15 (2014). doi:10.1371/journal.pgen.1004383
39. Pickrell, J.K., Berisa, T., Liu, J.Z., Ségurel, L., Tung, J.Y., Hinds, D.A.: Detection and interpretation of shared genetic influences on 42 human traits. Nature Genetics **48**(7), 709–717 (2016). doi:10.1038/ng.3570
40. Hormozdiari, F., van de Bunt, M., Segrè, A.V., Li, X., Joo, J.W.J., Bilow, M., Sul, J.H., Sankararaman, S., Pasaniuc, B., Eskin, E.: Colocalization of GWAS and eQTL Signals Detects Target Genes. The American Journal of Human Genetics **99**(6) (2016). doi:10.1016/j.ajhg.2016.10.003. Accessed 2021-07-16
41. Foley, C.N., Staley, J.R., Breen, P.G., Sun, B.B., Kirk, P.D.W., Burgess, S., Howson, J.M.M.: A fast and efficient colocalization algorithm for identifying shared genetic risk factors across multiple traits. Nature Communications **12**(1) (2021). doi:10.1038/s41467-020-20885-8. Accessed 2021-07-16
42. Wallace, C.: A more accurate method for colocalisation analysis allowing for multiple causal variants. bioRxiv (2021). doi:10.1101/2021.02.23.432421. https://www.biorxiv.org/content/early/2021/02/23/2021.02.23.432421.full.pdf
43. Hayes, A.F., Preacher, K.J.: Statistical mediation analysis with a multicategorical independent variable. British Journal of Mathematical and Statistical Psychology **67**(3), 451–470 (2014). doi:10.1111/bmsp.12028
44. James, L.R., Brett, J.M., Demaree, R.G., Dugas, J.L., Gustafson, S.B., Jones, A.P.: Mediators, Moderators, and Tests for Mediation Mulaik, and Gernt Wolf for their helpful suggestions and advice Requests for reprints should be sent to. Journal of Applied Psychology **69**(2), 307–321 (1984)
45. Muller, D., Judd, C.M., Yzerbyt, V.Y.: When moderation is mediated and mediation is moderated. Journal of Personality and Social Psychology **89**(6), 852–863 (2005). doi:10.1037/0022-3514.89.6.852
46. Otter, T., Pachali, M.J., Mayer, S., Landwehr, J.: Causal Inference Using Mediation Analysis or Instrumental Variables - Full Mediation in the Absence of Conditional Independence. Marketing ZFP – Journal of Research

and Management **40**(2), 41–57 (2018). doi:10.2139/ssrn.3135313
47. Hayes, A.F., Scharkow, M.: The relative trustworthiness of inferential tests of the indirect effect in statistical mediation analysis. Psychological Science **24**(10), 1918–1927 (2013). doi:10.1177/0956797613480187
48. Hayes, A.F.: Beyond baron and kenny: Statistical mediation analysis in the new millennium. Communication Monographs **76**(4), 408–420 (2009). doi:10.1080/03637750903310360
49. Fritz, M.S., MacKinnon, D.P.: Required Sample Size to Detect the Mediated Effect. Psychological Science **18**(3), 233–239 (2007). doi:10.1111/j.1467-9280.2007.01882.x
50. Rucker, D.D., Preacher, K.J., Tormala, Z.L., Petty, R.E.: Mediation analysis in social psychology: Current practices and new recommendations. Social and Personality Psychology Compass **5**(6), 359–371 (2011). doi:10.1111/j.1751-9004.2011.00355.x
51. Keele, G.R., Crouse, W.L., Kelada, S.N.P., Valdar, W.: Determinants of QTL Mapping Power in the Realized Collaborative Cross. G3 (Bethesda, Md.) **9**(May), 459966 (2019). doi:10.1534/g3.119.400194
52. Gatti, D.M., Simecek, P., Somes, L., Jeffrey, C.T., Vincent, M.J., Choi, K., Chen, X., Churchill, G.A., Svenson, K.L.: The effects of sex and diet on physiology and liver gene expression in diversity outbred mice. bioRxiv (2017). doi:10.1101/098657. https://www.biorxiv.org/content/early/2017/01/05/098657.full.pdf
53. Beasley, T.M., Erickson, S., Allison, D.B.: Rank-Based Inverse Normal Transformations are Increasingly Used, But are They Merited? Behavior Genetics **39**(5), 580–595 (2009). doi:10.1007/s10519-009-9281-0
54. R/qtl2: Software for Mapping Quantitative Trait Loci with High-Dimensional Data and Multiparent Populations. Genetics **211**(2), 495–502 (2019). doi:10.1534/genetics.118.301595
55. van de Geijn, B., McVicker, G., Gilad, Y., Pritchard, J.K.: WASP: allele-specific software for robust molecular quantitative trait locus discovery. Nature Methods **12**(11), 1061–1063 (2015). doi:10.1038/nmeth.3582

Supplemental Methods

Joint likelihood for causal models in the reactive mode

For causal models describing reverse causality between M and Y, that is, the reactive models $\boldsymbol{\theta} = (\theta_a, \theta_b = *, \theta_c)$, the roles of M and Y are switched relative to the other models, so that the conditional joint likelihood $\mathrm{p}(\mathbf{y}, \mathbf{m}|\boldsymbol{\theta})$ in these case is

$$\begin{aligned}\mathbf{m}|\mathbf{y}, \theta_a, \theta_b = *, \mu_{\mathbf{m}}, \boldsymbol{\alpha}_{\mathbf{m}}, \boldsymbol{\beta}_a, \beta_b, \sigma_{\mathbf{m}} &\sim \mathrm{N}(\mu_{\mathbf{m}}\mathbf{1} + \mathbf{Z}_{\mathbf{m}}\boldsymbol{\alpha}_{\mathbf{m}} + \theta_a\mathbf{X}\boldsymbol{\beta}_a + \mathbf{y}\beta_b, \sigma^2_{\mathbf{m}}\mathbf{W}^{-1}_{\mathbf{m}}),\\ \mathbf{y}|\theta_b = *, \theta_c, \mu_{\mathbf{y}}, \boldsymbol{\alpha}_{\mathbf{y}}, \boldsymbol{\beta}_c, \sigma_{\mathbf{y}} &\sim \mathrm{N}(\mu_{\mathbf{y}}\mathbf{1} + \mathbf{Z}_{\mathbf{y}}\boldsymbol{\alpha}_{\mathbf{y}} + \theta_c\mathbf{X}\boldsymbol{\beta}_c, \sigma^2_{\mathbf{y}}\mathbf{W}^{-1}_{\mathbf{y}}),\end{aligned}$$

where β_b is now the scalar effect of $\mathbf{y}$ on $\mathbf{m}$. Prior distributions for all variables are unchanged (except for $\boldsymbol{\theta}$). When $\theta_b = *$, the marginal joint likelihood function is given by:

$$\begin{aligned}\mathbf{m}|\mathbf{y}, \theta_a, \theta_b = * &\sim \mathrm{t}_{\kappa_{\mathbf{m}}}(\mathbf{0}, \lambda_{\mathbf{m}}[\mathbf{W}^{-1} + \mathcal{X}_{\mathbf{m}}\mathbf{V}_{\mathbf{m}}\mathcal{X}^T_{\mathbf{m}}])\\ \mathbf{y}|\theta_b = *, \theta_c &\sim \mathrm{t}_{\kappa_{\mathbf{y}}}(\mathbf{0}, \lambda_{\mathbf{y}}[\mathbf{W}^{-1} + \mathcal{X}_{\mathbf{y}}\mathbf{V}_{\mathbf{y}}\mathcal{X}^T_{\mathbf{y}}])\end{aligned}$$

$$\mathcal{X}_{\mathbf{m}} = \begin{bmatrix}\mathbf{1} & \mathbf{Z}_{\mathbf{m}} & \theta_a\mathbf{X} & \mathbf{y}\end{bmatrix} \qquad \mathbf{V}_{\mathbf{m}} = \begin{bmatrix}\tau^2_\mu & \mathbf{0} & \mathbf{0} & 0\\ \mathbf{0} & \tau^2_{\mathbf{Z}}\mathbf{I} & \mathbf{0} & \mathbf{0}\\ \mathbf{0} & \mathbf{0} & \phi^2_a\mathbf{I} & \mathbf{0}\\ 0 & \mathbf{0} & \mathbf{0} & \phi^2_b\end{bmatrix}$$

$$\mathcal{X}_{\mathbf{y}} = \begin{bmatrix}\mathbf{1} & \mathbf{Z}_{\mathbf{y}} & \theta_c\mathbf{X}\end{bmatrix} \qquad \mathbf{V}_{\mathbf{y}} = \begin{bmatrix}\tau^2_\mu & \mathbf{0} & \mathbf{0}\\ \mathbf{0} & \tau^2_{\mathbf{Z}}\mathbf{I} & \mathbf{0}\\ \mathbf{0} & \mathbf{0} & \phi^2_c\mathbf{I}\end{bmatrix}$$

Hyperparameters for $\boldsymbol{\kappa}, \boldsymbol{\lambda}, \boldsymbol{\tau}_\mu, \boldsymbol{\tau}_{\mathbf{Z}}$, and $\boldsymbol{\phi}^2$ are unchanged.

Supplemental Figures

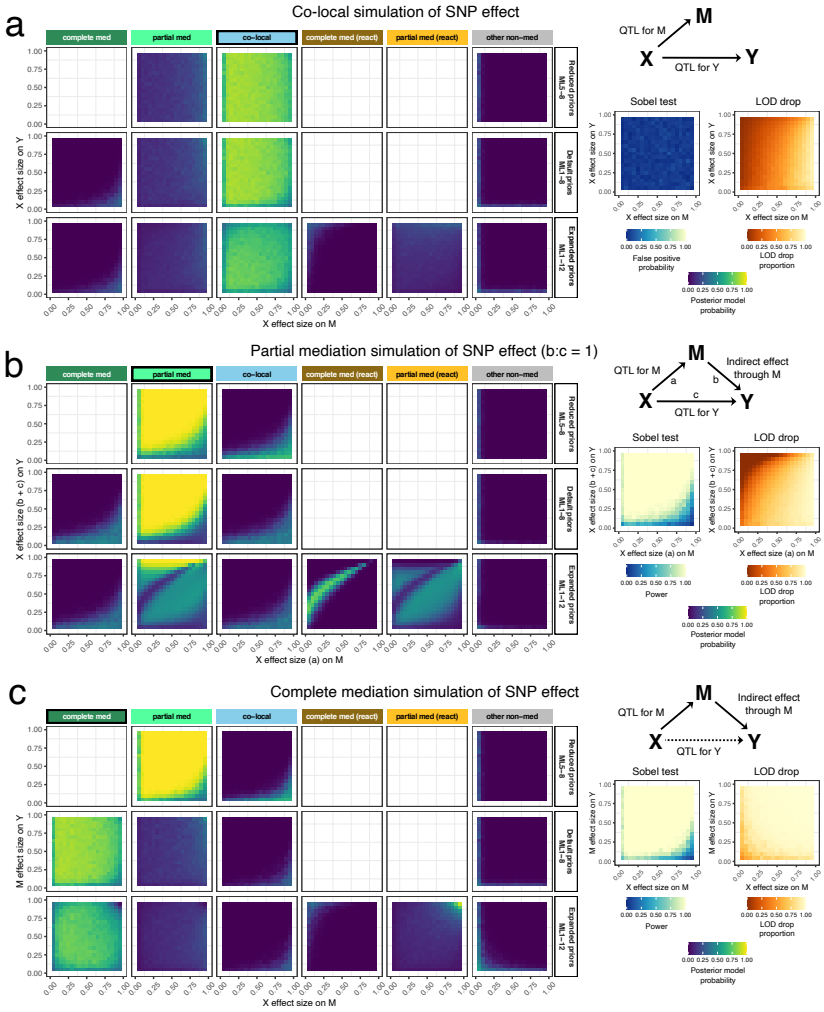

Figure S1 Performance of Bayesian model selection, Sobel test, and LOD drop in simulated QTL data, including reactive models. Data for 200 individuals were simulated according to (a) co-local, (b) partial mediation, and (c) complete mediation models from a balanced bi-allelic SNP. There is no misspecification of X. The DAGs represent the underlying models of each set of simulations. Heat maps for Bayesian model selection represent the mean posterior probability associated with each inferred model for a range of fixed settings of the model parameters as indicated on x- and y-axes. Heat maps for the Sobel test represent false positive probability for co-local simulations and power for mediation simulations. Heat maps for LOD drop represent mean LOD drop, scaled to the proportion of the simulated QTL's LOD score. Empty squares represent posterior model categories not evaluated based on the set of admissible models encoded in the model priors.

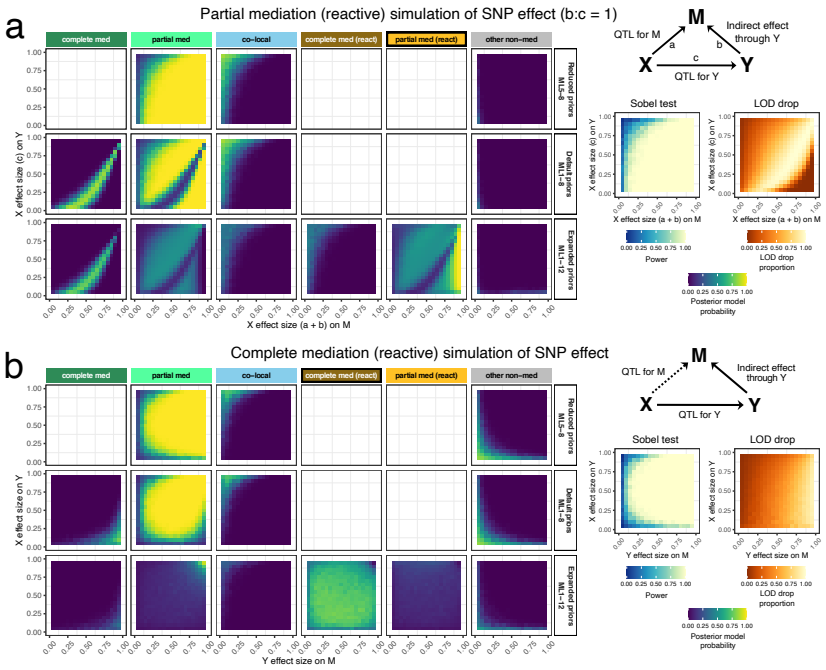

Figure S2 Performance of Bayesian model selection, Sobel test, and LOD drop in simulated reactive QTL data, including reactive models. Data for 200 individuals were simulated according to (a) reactive partial mediation and (b) reactive complete mediation models from a balanced bi-allelic SNP. There is no misspecification of X. The DAGs represent the underlying models of each set of simulations. Heat maps for Bayesian model selection represent the mean posterior probability associated with each inferred model for a range of fixed settings of the model parameters as indicated on x- and y-axes. Heat maps for the Sobel test represent false positive probability for co-local simulations and power for mediation simulations. Heat maps for LOD drop represent mean LOD drop, scaled to the proportion of the simulated QTL's LOD score. Empty squares represent posterior model categories not evaluated based on the set of admissible models encoded in the model priors.

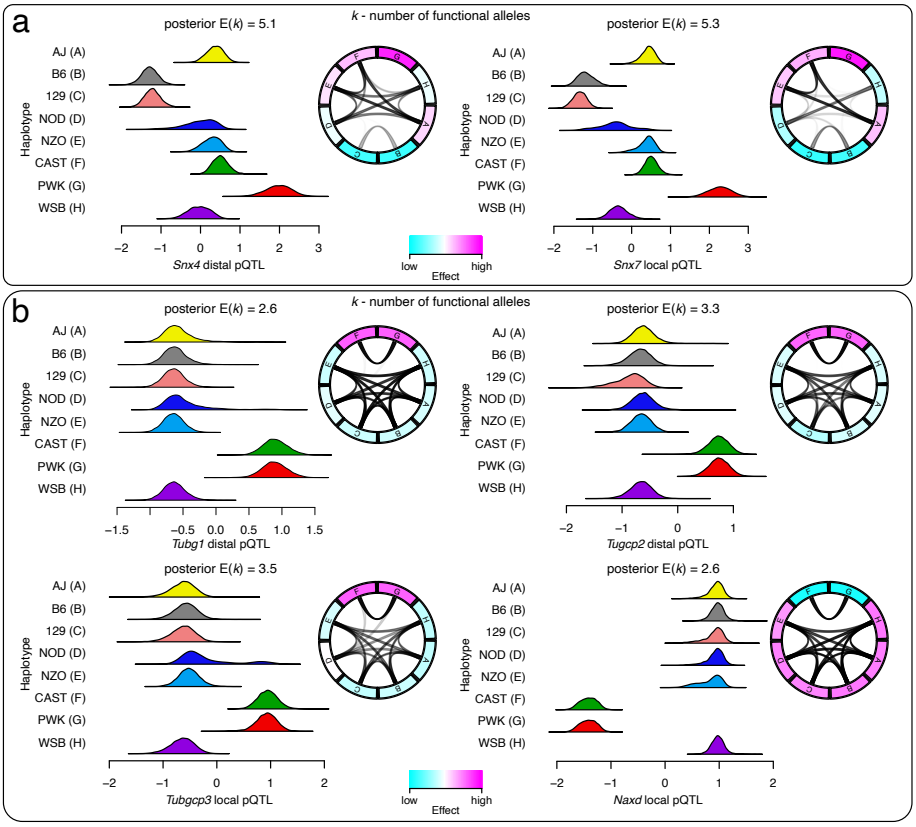

Figure S3 Modeling the allelic series of pQTL for Snx4, Tubg1, and related genes identified through mediation. The allelic series were modeled with TIMBR. (a) Posterior haplotype effects and allelic series for the *Snx4* distal pQTL and the local pQTL for its mediator, *Snx7*. Haploytpe effects are represented as histograms of the posterior samples. Allelic series are represented as circos plots where the opacity of the links represent how often TIMBR assigned founder haplotypes to the same functional allele. (b) Posterior haplotype effects and allelic series for the *Tubg1* and *Tubgcp2* distal pQTL and the local pQTL for their candidate mediators, *Tubgcp3* and *Naxd*. The posterior expected number of functional alleles, k, is included for each pQTL.