Chapter 3A

Frederick Boehm

February 5, 2019

Todo list

Contents

1 Introduction

A central goal of systems genetics studies is to identify causal relationships between genetic variants and quantitative traits. Recent work by Chick et al. (2016) has popularized linear regression-based methods for causal inference in systems genetics. The great successes of these methods in identifying causal relationships among genetic variants and quantitative biomolecule concentrations motivates our goal of clarifying a role for our test of pleiotropy vs. separate QTL in systems genetics studies. We argue that our test of pleiotropy vs. separate QTL complements mediation analysis in two ways. First, our test limits the set of candidate mediators by ruling out traits that don't share a single pleiotropic QTL. Second, when regression-based mediation analysis fails to identify a mediator, our test of pleiotropy vs. separate QTL still provides information on the number of QTL. The number of separate QTL may provide clues to aid biological understanding.

We begin below with an overview of a causal mediation analysis framework that VanderWeele (2015) summarizes. We then discuss applications of parts of this framework to the systems genetics context. We describe in our methods section the study design and statistical analyses in which we use mediation analysis and our pleiotropy vs. separate QTL test to dissect an expression trait hotspot with data from 378 Diversity Outbred mice.

Specifically, one may formulate a mediation analysis by identifying a marker (and its founder allele probabilities) as the "driver" of a causal diagram. In recalling the central dogma of molecular biology, we want to consider a marker that affects transcript levels of a nearby gene (Crick, 1958). For example, if a gene is located on Chromosome 2 and centered at 100 Mb, then we would want to consider markers that are also on Chromosome 2 and near 100 Mb, perhaps between 98 Mb and 102 Mb. If a second expression trait - typically transcripts from a gene that is not "local" - maps near the marker that we've identified, then we can ask whether trait1 "mediates" the relationship between trait2 and the marker.

To determine whether trait1 mediates the relationship between trait2 and the marker genotypes, *i.e.*, the driver, we perform a series of regression analyses, which we detail below (Equations 1, 2, 3, 4). In brief, we perform regression-based QTL mapping between trait2 and the marker, with and without conditioning on the candidate mediator. If the LOD score diminishes sufficiently upon conditioning on a candidate mediator, then we declare the candidate mediator a mediator.

The rationale behind this strategy is as follows. If the driver affects trait2 solely by way of trait1, then conditioning on trait1 would nullify the relationship between driver and trait2. At the other extreme, if the driver affects trait2 solely through mechanisms that don't involve trait1, then conditioning on trait1 values would not affect the association between the driver and trait2.

One role for our test of pleiotropy vs separate QTL is to potentially limit the set of candidate mediators

that require study. For example, if trait1 and trait2 map to separate QTL, then it would be unlikely that trait1 affects trait2 (or vice versa). On the other hand, if we believe that a single pleiotropic locus affects both trait1 and trait2, then a mediation analysis is particularly useful in efforts to clarify the relationship between trait1 and trait2. In this setting, where a single pleiotropic locus affects both trait1 and trait2, and we seek to identify whether the trait1 or trait2 is the intermediate, we would perform two mediation analyses, one with trait1 as candidate intermediate between driver and trait2, and a second analysis with trait2 as a candidate intermediate between driver and trait1.

1.1 Regression-based mediation methods

Frequently one uses regression-based methods to identify causal intermediates from among multiple measured variables (Baron and Kenny, 1986). For example, Chick et al. (2016) measured 16,921 gene transcript levels and 6,756 protein concentrations in liver tissue from 192 mice. After identifying QTL for gene transcript levels and protein concentrations, they performed regression-based mediation analyses. For each candidate mediator, they fitted four linear models:

$$Y = b1 + WC + E \tag{1}$$

1: Linear model with intercept and covariates only.

$$Y = XB + WC + E \tag{2}$$

2: Linear model with founder allele dosages and covariates.

$$Y = b1 + WC + M\beta + E \tag{3}$$

3: Linear model with intercept, covariates, and candidate mediator.

$$Y = XB + WC + M\beta + E \tag{4}$$

4: Linear model with founder allele dosages, covariates, and candidate mediator.

In the above four models, X is a n by 8 matrix of founder allele probabilities at a single marker, B is a 8 by 1 matrix of founder allele effects, E is a n by 1 matrix of random errors, b is an number, 1 is a n by 1 matrix with all entries set to 1, Y is a n by 1 matrix of phenotype values (for a single trait), and M is a n by 1 matrix of values for a putative mediator. C is a matrix of covariate effects, and W is a matrix of covariates. We denote the coefficient of the mediator by β.

We assume that the vector E is (multivariate) normally distributed with zero vector as mean and covariance matrix $\Sigma = \sigma^2 I_n$, where I_n is the n by n identity matrix.

With the above models, the log-likelihoods are easily calculated. For example, in Equation 1, the vector Y follows a multivariate normal distribution with mean $(b1 + WC)$ and covariance $\Sigma = \sigma^2 I$. Thus, we can write the likelihood for Model 1 as:

$$L(b, C, \sigma^2 | Y, W) = (2\pi)^{-\frac{n}{2}} \exp\left(-\frac{1}{2}(Y - b1 - WC)^T \Sigma^{-1} (Y - b1 - WC)\right) \tag{5}$$

We thus have the following equation for the log-likelihood for model 1:

$$\log L(b, C, \sigma^2 | Y, W) = -\frac{n}{2}\log(2\pi) - \frac{1}{2}(Y - b1 - WC)^T \Sigma^{-1} (Y - b1 - WC) \tag{6}$$

Chick et al. (2016) calculated the $\log_{10}$ likelihoods for all four models before determining two LOD scores:

$$LOD_1 = log_{10}(\text{model 2 likelihood}) - log_{10}(\text{model 1 likelihood}) \tag{7}$$

$$LOD_2 = log_{10}(\text{model 4 likelihood}) - log_{10}(\text{model 3 likelihood}) \tag{8}$$

And, finally, Keller et al. (2018) calculated the statistic:

$$\text{LOD difference} = LOD_1 - LOD_2 \tag{9}$$

LOD difference need not be positive. For example, in the setting where the putative mediator is not a true mediator, LOD_1 and LOD_2 values may lead to a negative LOD difference statistic. In our analyses in this section, we truncated these negative LOD difference values to zero.

For each mediation analysis, we consider the LOD difference proportion, which we define below (Equation 10).

$$\text{LOD difference proportion} = \frac{(LOD_1 - LOD_2)}{LOD_1} \tag{10}$$

In other words, we consider what proportion of the association strength, on the LOD scale, is diminished by conditioning on a putative mediator. This statistic differs from the LOD difference statistic in that it scales the mediation LOD difference by LOD_1 value. We do this in efforts to accommodate the diversity of LOD peak heights in our data. While a value of 5 for LOD difference may be important for a trait with a

LOD_1 of 12, another trait with a LOD_1 of 120 that has a LOD_2 of 115, suggests that nearly all of the signal remains transmitted when conditioning on the putative mediator.

To perform a mediation analysis, we must specify three components of a possible directed acyclic graph. For our purposes, we always begin with a marker's founder allele probabilities. Since we are working in the context of dissecting an expression QTL hotspot, we choose the "driver" to be the founder allele probabilities at the peak position for a local expression trait. We then choose the terminus of the putative graph, i.e., the "target", to be a nonlocal expression trait.

1.2 Four assumptions for causal inference

To claim that the LOD difference statistic reflects a causal relationship, four assumptions about confounding are needed (VanderWeele, 2015).

1. No unmeasured confounding of the treatment-outcome relationship
2. No unmeasured confounding of the mediator-outcome relationship
3. No unmeasured confounding of the treatment-mediator relationship
4. No mediator-outcome confounder that is affected by the treatment

In the above terminology, treatment refers to the founder allele dosages matrix, while outcome is the nonlocal gene expression trait, and mediator is the local gene expression trait.

1.3 Declaring mediators

Multiple research teams have proposed strategies for assessing statistical significance of the "mediation LOD difference" statistic. Chick et al. (2016) used individual transcript levels as sham mediators and tabulated their LOD difference statistics. They then compared the observed LOD difference statistics for putative mediators to the empirical distribution of LOD difference statistics obtained from the collection of sham mediators. Keller et al. (2018), on the other hand, in their study of pancreatic islet cell biology, declared mediators those local transcripts that diminished the LOD score of nonlocal transcripts by at least 1.5 (Keller et al., 2018). While significance threshold determination remains an active area of research, we proceed below by examining all 147 nonlocal transcript levels that Keller et al. (2018) identified as mapping to the Chromosome 2 hotspot.

2 Methods

We examined the potential that the two methods - 1. pleiotropy vs. separate QTL testing and 2. mediation analysis - could play complementary roles in efforts to dissect gene expression trait hotspots. After describing our data, we explain our statistical analyses involving 13 local gene expression traits and 147 nonlocal gene expression traits, all of which map to a 4-Mb hotspot on Chromosome 2.

2.1 Data description

We analyzed data from 378 Diversity Outbred mice (Keller et al., 2018). Keller et al. (2018) genotyped tail biopsies with the GigaMUGA microarray (Morgan et al., 2015). They also used RNA sequencing to measure genome-wide pancreatic islet cell gene expression for each mouse at the time of sacrifice (Keller et al., 2018). They shared these data, together with inferred founder allele probabilities, on the Data Dryad site (https://datadryad.org/resource/doi:10.5061/dryad.pj105).

We downloaded the file "Attie_DO378_eQTL_viewer_v1.Rdata" from Data Dryad (Keller et al., 2018) and analyzed the data in the R statistical computing environment (R Core Team, 2018).

We examined the Chromosome 2 pancreatic islet cell expression trait hotspot that Keller et al. (2018) identified with these data. Keller et al. (2018) identified 147 nonlocal traits that map to the 4-Mb region centered at 165.5 Mb on chromosome 2. The 147 nonlocal traits all exceeded the genome-wide LOD significance threshold, 7.18 (Keller et al., 2018). With regression-based mediation analyses, they identified expression levels of local gene *Hnf4a* as a mediator of 88 of these 147 nonlocal traits.

Because Keller et al. (2018) reported that some nonlocal traits that map to the Chromosome 2 hotspot did not demonstrate evidence of mediation by *Hnf4a* expression levels, we elected to study a collection of local gene expression traits, rather than *Hnf4a* alone. This strategy enabled us to ask whether one of twelve other local traits mediates those nonlocal hotspot traits that are not mediated by *Hnf4a*. Our set of local gene expression traits includes *Hnf4a* and 12 other local genes (Table 1). Our 13 local genes are the only genes that met all of the following criteria:

1. QTL peak with LOD > 40
2. QTL peak position within 2 Mb of hotspot center (165.5 Mb)
3. Gene midpoint within 2 Mb of hotspot center (165.5 Mb)

The 147 nonlocal traits that we studied all had LOD peak heights above 7.18 and LOD peak positions within 2 Mb of the center of the hotspot (at 165.5 Mb).

2.2 Statistical analyses

We used both univariate and bivariate QTL scans. In the univariate scans, we identified LOD peaks for each of 160 expression traits that mapped to within 2 Mb of the center of the Chromosome 2 expression trait hotspot. Thirteen of these 160 gene expression traits arise from local genes, *i.e.*, genes on Chromosome 2. The remaining 147 expression traits are nonlocal traits, meaning that they arise from genes that either are not on Chromosome 2 or are on Chromosome 2 but more than 2 Mb from the center of the expression trait hotspot (at 165.5 Mb).

2.2.1 Univariate QTL scans

Keller et al. (2018) provided univariate QTL mapping results in their Data Dryad file, so we didn't repeat the univariate QTL scans. We briefly summarize their analyses below.

For a given univariate phenotype, Keller et al. (2018) fitted a linear mixed effects model at every marker across the genome.

$$Y = XB + WC + G + E \tag{11}$$

where Y is a n by 1 matrix of phenotype values, X is a n by 8 matrix of founder allele probabilities at a single marker, B is a 8 by 1 matrix of founder allele effects, W is a n by 4 matrix of four covariates (sex and three binary indicators for wave membership), C is a 4 by 1 matrix of covariate effects, G is a n by 1 matrix of polygenic random effects, and E is a n by 1 matrix of random errors.

G and E are assumed independent and to follow the distributions below.

$$G \sim N(0, \sigma_g^2 K) \tag{12}$$

$$E \sim N(0, \sigma_e^2 I) \tag{13}$$

where K is a leave-one-chromosome-out kinship matrix, and I is the n by n identity matrix. Keller et al. (2018) fitted the models by first estimating the variance components σ_g^2 and σ_e^2 from the null model, *i.e.*, a model, for a given phenotype, that omits founder allele probabilities. With the estimated variance components, they then performed the univariate QTL scan with model fitting by generalized least squares. Keller et al. (2018) performed calculations with the `qtl2` R package (Karl W. Broman et al., 2018).

2.2.2 Bivariate QTL scans

Our analyses involved 13 local expression traits and 147 nonlocal expression traits. We described above the criteria for choosing these expression traits.

We performed a series of two-dimensional QTL scans in which we paired each local gene's transcript levels with each nonlocal gene's transcript levels, for a total of 13 * 147 = 1911 two-dimensional scans. Each scan examined the same set of 180 markers, which spanned the interval from 163.1 Mb to 167.8 Mb and included univariate peaks for all 13 + 147 = 160 expression traits. We performed these analyses with the R package `qtl2pleio` (Boehm, 2018).

For each QTL scan, we fitted a collection of bivariate models for all 180 * 180 = 32,400 ordered pairs of markers. Throughout our analyses, we treated the founder allele probabilities as known, despite the fact that they are inferred through hidden Markov model methods (Karl W Broman, 2012a; Karl W Broman, 2012b; Karl W. Broman et al., 2018).

For each ordered pair of markers, we fitted the model:

$$vec(Y) = Xvec(B) + vec(G) + vec(E) \tag{14}$$

where Y is a n by 2 matrix of phenotypes, B is a 12 by 2 matrix of founder allele and covariate effects, G is a n by 2 matrix of polygenic random effects, and E is a n by 2 matrix of random errors. We impose distributional assumptions on G and E, which are assumed to be independent.

$$G \sim MN(0, K, V_g) \tag{15}$$

$$E \sim MN(0, I, V_e) \tag{16}$$

K is a leave-one-chromosome-out n by n kinship matrix. We calculated its entries by using all markers except those on Chromosome 2. The i, j entry of K is calculated with Equation 17.

$$K[i, j] = \sum_{k,l} (p_{ikl} p_{jkl}) \tag{17}$$

where k indexes marker positions and includes all markers except those on Chromosome 2, l indexes the eight founder alleles, and i and j represent two subjects.

In equation 14, X is a $2n$ by 24 matrix of founder allele probabilities and covariates. X has a block-diagonal structure, with the two n by 12 blocks on the diagonal containing nonzero entries, while the

off-diagonal n by 12 blocks contain only zeros. Each n by 12 block on the diagonal contains eight founder allele probabilities for one marker and values for four covariates. We used as covariates sex and three binary indicators for wave number. As Keller et al. (2018) describe, the 378 mice represent four distinct waves, or batches.

2.2.3 Mediation analyses

We performed linear regression-based mediation analyses for all 1,911 local-nonlocal trait pairs. We probed the extent to which each nonlocal trait's association strength diminished upon conditioning on transcript levels of a putative mediator. Each of the 13 local expression traits, considered one at a time, served as putative mediators. We thus fitted the four linear regression models that we describe above (Equations 1, 2, 3, 4).

One question that needs clarification is the choice of "driver" for each mediation analysis. We elected to use the founder allele probabilities at the marker that demonstrated the univariate LOD peak for the putative mediator. Alternative analyses, in which one chooses as driver the founder allele probabilities at which the nonlocal trait has its univariate peak, are also possible.

2.2.4 Local gene analyses

To visualize the summary statistics for our two methods, we plotted, for each local gene, a scatterplot of LOD difference proportion values against pleiotropy vs. separate QTL test statistics.

symbol	start	end	LOD_peak_position	LOD_peak_height
Pkig	163.66	163.73	163.52	51.68
Serinc3	163.62	163.65	163.58	126.93
Hnf4a	163.51	163.57	164.02	48.98
Stk4	164.07	164.16	164.03	60.39
Pabpc1l	164.03	164.05	164.03	52.50
Slpi	164.35	164.39	164.61	40.50
Neurl2	164.83	164.83	164.64	64.58
Cdh22	165.11	165.23	165.05	53.84
2810408M09Rik	165.49	165.49	165.57	67.34
Eya2	165.60	165.77	165.72	98.89
Prex1	166.57	166.71	166.75	46.91
Ptgis	167.19	167.24	167.27	56.25
Gm14291	167.20	167.20	167.27	73.72

Table 1: Local gene annotations for analysis of Chromosome 2 expression trait hotspot. All positions are in units of Mb on Chromosome 2. LOD peak position and LOD peak height refer to those obtained from univariate analyses. "Start" and "end" refer to the local gene's DNA start and end positions, as annotated by Ensembl version 75.

2.2.5 Nonlocal gene analyses

We also examined the 1911 pairs from the per-nonlocal gene perspective. For each of the 147 nonlocal genes, we plotted LOD difference proportion against pleiotropy vs. separate QTL test statistic values. We present below (Figure 4) examples that illustrate some of the observed patterns between pleiotropy vs. separate QTL test statistics and LOD difference proportion values.

3 Results

3.1 Scatter plot for all 1911 pairs

We present below our scatter plot for all 147 * 13 = 1911 pairs of traits (Figure 1). Each pair contains one local expression trait and one nonlocal expression trait. Each point in the figure represents a single pair. We see that points with high values of LOD difference proportion tend to have small values of pleiotropy vs. separate QTL test statistic, and those points with high values of the pleiotropy vs. separate QTL test statistic tend to have small values of LOD difference proportion.

We colored blue the points that involve *Hnf4a* as the local gene; all other local genes have red points. The most striking feature of the coloring is that many blue points have small values of the pleiotropy vs. separate QTL test statistics and very high values of LOD difference proportion.

3.2 Local gene analyses

To more thoroughly examine the relationships across the 13 local genes, we created 13 plots of LOD difference proportion (from the mediation analyses) against pleiotropy vs. separate QTL test statistic. They reveal common patterns. First, we see no points in the upper right quadrant of each plot. This tells us that those nonlocal genes with high values of pleiotropy vs. separate QTL test statistic have low values of LOD difference proportion. Similarly, those nonlocal genes with high values of LOD difference proportion tend to have small values of the pleiotropy vs. separate QTL test statistic. Finally, some trait pairs demonstrate low values of both the LOD difference proportion and pleiotropy vs. separate QTL test statistic. This observation suggests that, for a given local expression trait, these pairs are not mediated by the local expression trait yet they arise from a shared pleiotropic locus.

In comparing the *Hnf4a* plot (Figure 2) with the other 12 plots (Figure 3), we see that none of the 12 plots in Figure 3 closely resembles Figure 2. *Serinc3*, *Stk4*, *Neurl2*, and *Cdh22* are closest in appearance to the plot of *Hnf4a*. However, each of *Serinc3*, *Stk4*, *Neurl2*, and *Cdh22* has very few points with LOD difference proportion above 0.5, while *Hnf4a* has many points with LOD difference proportion above 0.5.

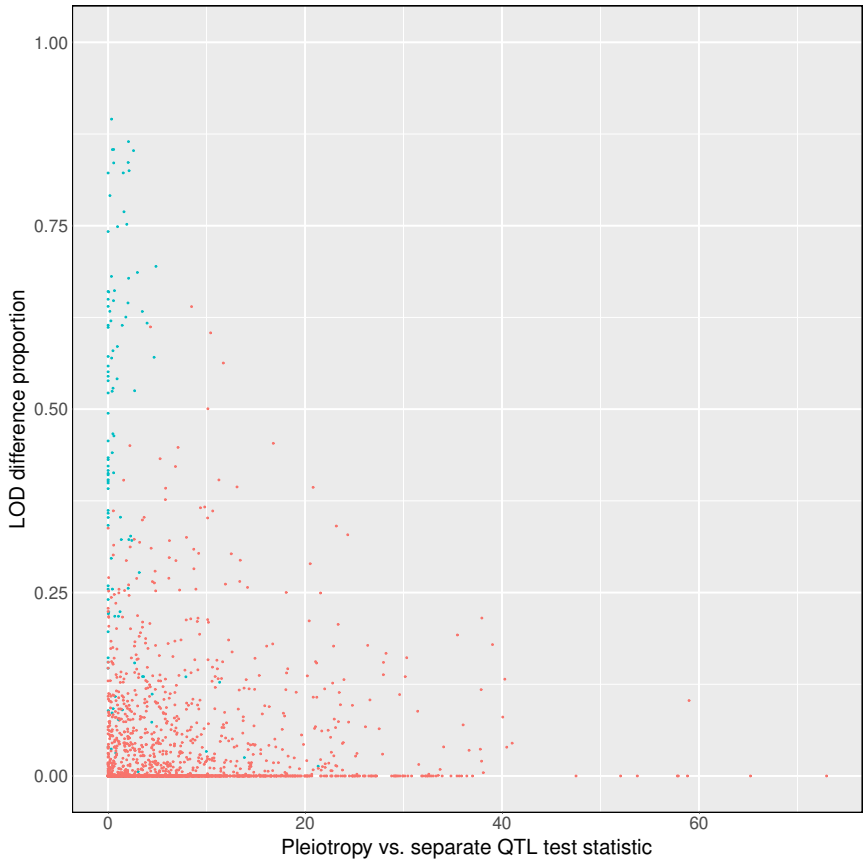


Figure 1: LOD difference proportion against pleiotropy vs. separate QTL test statistics for 1911 pairs of traits. Each point represents a single pair. Pairs that involve local expression trait Hnf4a are colored blue, while all others are colored red. Note the high prevalence of blue points in the upper left quadrant of the figure. These points, with low values of the pleiotropy vs. separate QTL test statistic and high values of LOD difference proportion, are consistent with Hnf4a transcript levels mediating the effect of Hnf4a genetic variants affecting nonlocal transcript abundances.

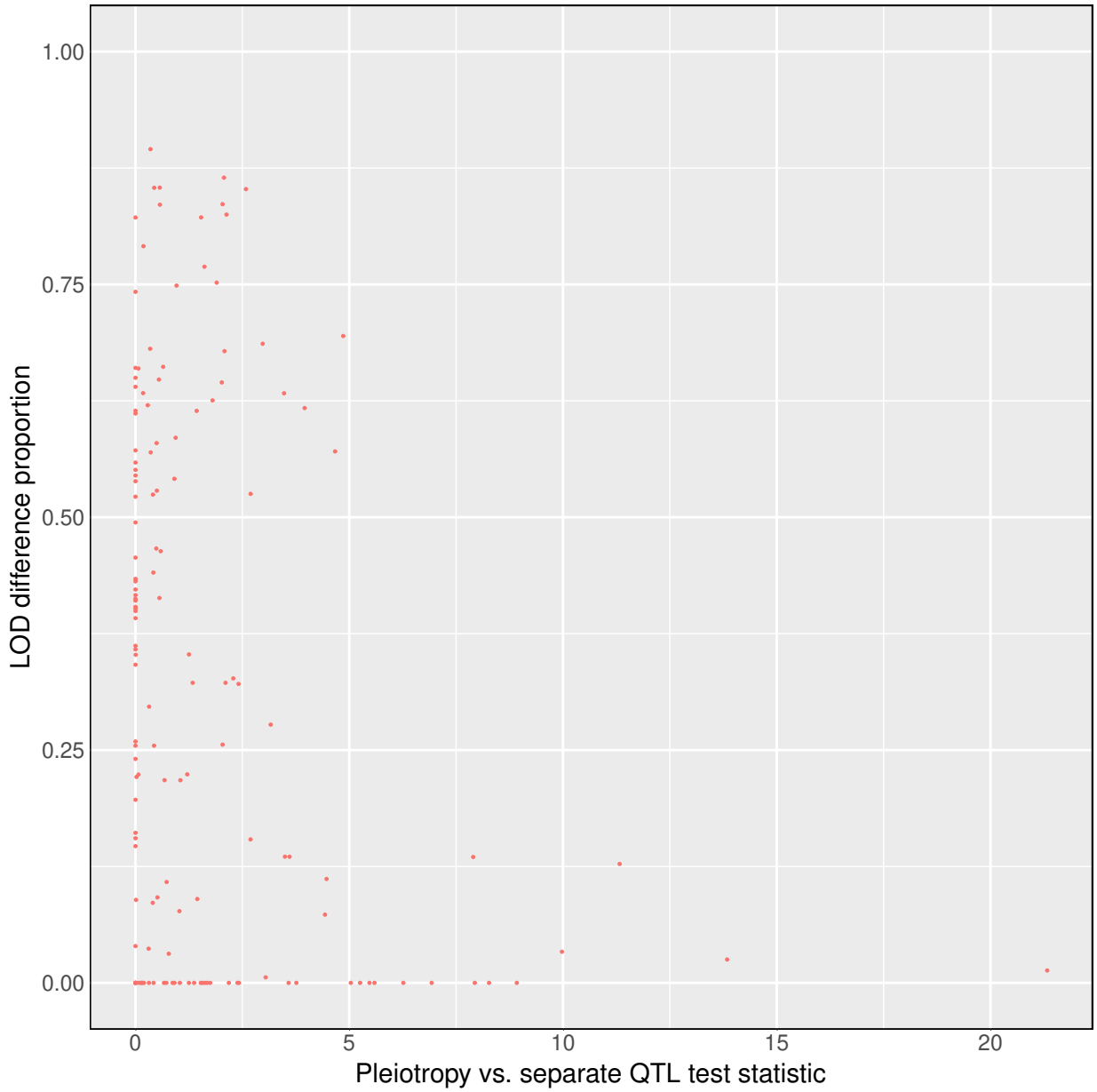


Figure 2: Scatter plot of LOD difference proportion against pleiotropy vs. separate QTL test statistic for 147 pairs of traits. Each pair includes *Hnf4a* and one of the nonlocal gene expression traits that map to the Chromosome 2 hotspot.

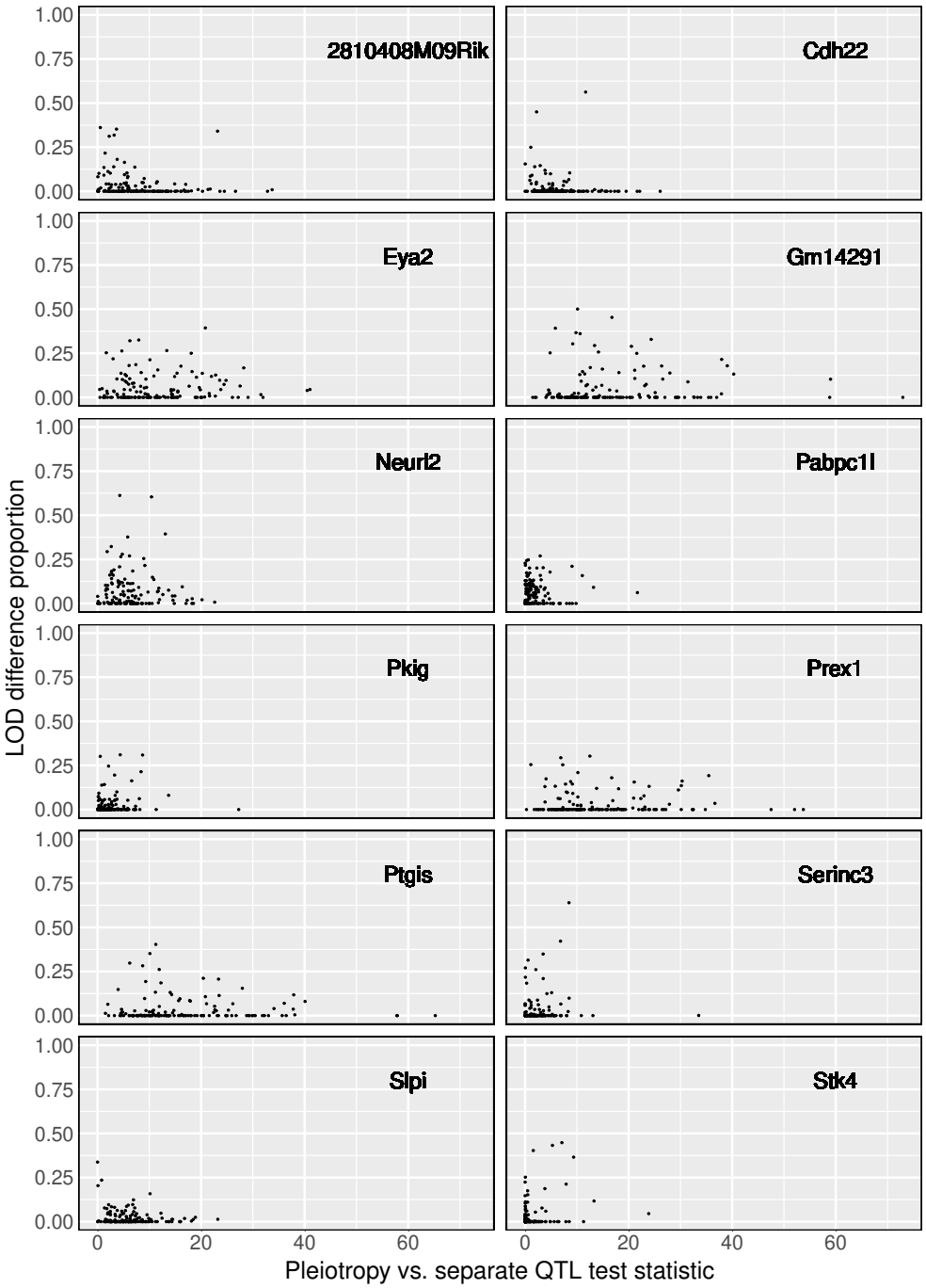


Figure 3: Scatter plots of LOD difference proportion against pleiotropy vs. separate QTL test statistic with 147 pairs of traits per panel. Each panel includes a local gene expression trait (per the label in the upper right quadrant) and one of the 147 nonlocal gene expression traits that map to the Chromosome 2 hotspot.

Local gene	Number of nonlocal gene expression traits
2810408M09Rik	3
Cdh22	4
Eya2	4
Gm14291	5
Hnf4a	89
Neurl2	15
Pabpc1l	4
Pkig	2
Prex1	5
Ptgis	3
Serinc3	7
Slpi	1
Stk4	5

Table 2: Number of nonlocal gene expression traits for which each local gene expression trait is the strongest mediator.

3.3 Nonlocal gene analyses

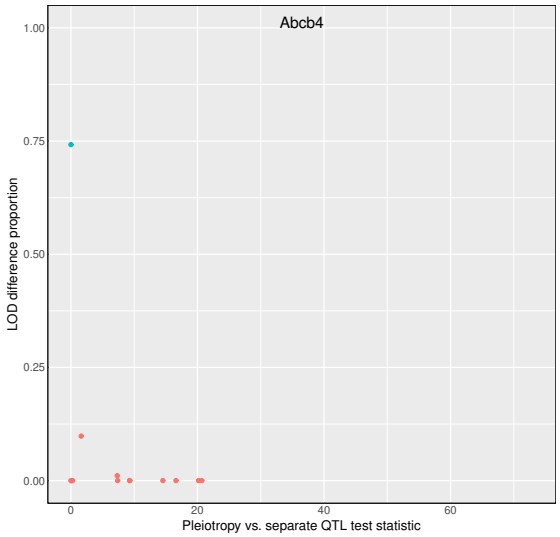
Scatter plots of LOD difference proportion values against pleiotropy vs. separate QTL test statistics for each of the 147 nonlocal expression traits demonstrated multiple patterns. Eighty-nine nonlocal genes' plots showed Hnf4a to have the greatest value, among the 13 putative mediators, of the LOD difference proportion statistic. In many cases, Hnf4a's LOD difference proportion statistic was at least twice that of any of the other 12 local gene expression levels.

4 Discussion

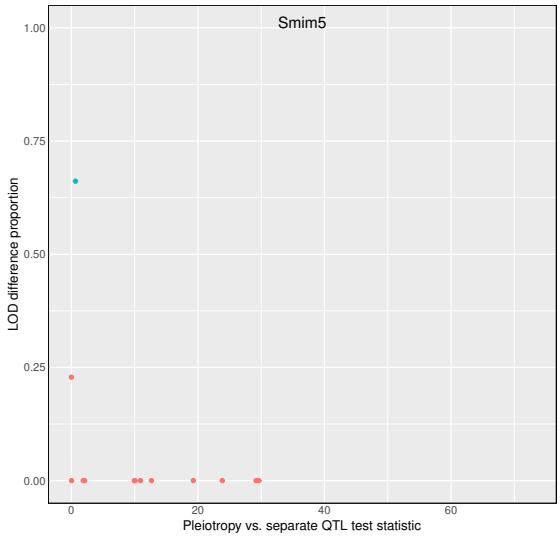
Our pairwise analyses with both mediation analyses and tests of pleiotropy vs. separate QTL provide additional evidence for the importance of Hnf4a in the biology of the Chromosome 2 hotspot in pancreatic islet cells. Our analyses, and, specifically, the test of pleiotropy vs. separate QTL, may be more useful when studying nonlocal traits that map to a hotspot yet don't show strong evidence of mediation by local expression traits. In such a setting, the test of pleiotropy vs. separate QTL can, at least, provide some information about the genetic architecture at the hotspot. Specifically, our test of pleiotropy vs. separate QTL may inform inferences about the number of underlying QTL in a given expression trait hotspot. Additionally, our test may limit the number of expression traits that are potential intermediates between a QTL and a specified nonlocal expression trait. This relies on the assumption that an intermediate expression trait and a target expression trait presumably share a QTL.

On the other hand, mediation analyses, when they provide evidence for mediation of a nonlocal trait by a local expression trait, are more informative than the test of pleiotropy vs. separate QTL, since the

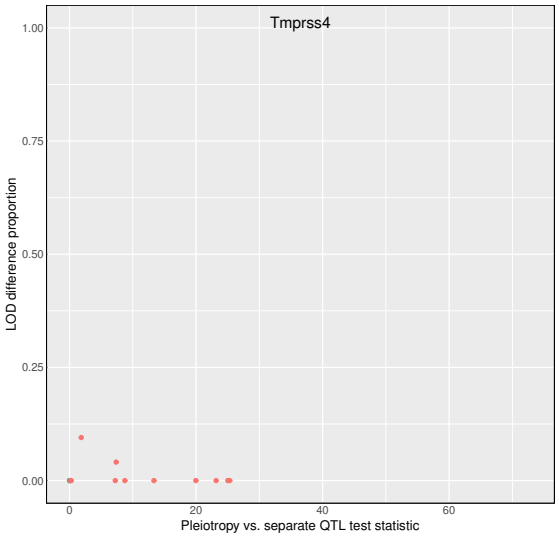
Figure 4: Scatter plots for four nonlocal expression traits. Each plot features 13 points, one for each local gene expression trait. The y axis denotes LOD difference proportion values, while the x axis corresponds to pleiotropy vs. separate QTL test statistics. Blue points represent the pairing with local gene expression trait Hnf4a. Red points represent the other 12 local gene expression traits.



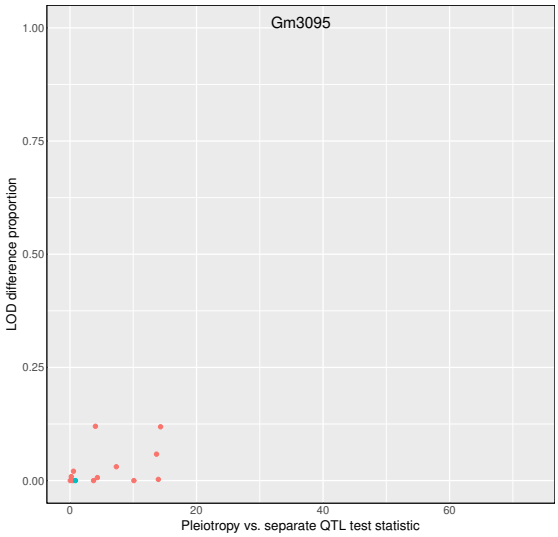
(a) Hnf4a transcript levels mediate Abcb4 transcript levels. The high LOD difference proportion and the very small pleiotropy vs. separate QTL test statistic together provide evidence that Hnf4a mediates Abcb4 transcript levels.



(b) Hnf4a transcript levels mediate Smim5 transcript levels. The high LOD difference proportion and the very small pleiotropy vs. separate QTL test statistic together provide evidence that Hnf4a mediates Smim5 transcript levels.



(c) Tmprww4 transcript levels are not mediated by any of the 13 local gene expression traits. However, several local gene expression traits arise from separate QTL, as evidenced by their large (greater than 5) values of the pleiotropy vs. separate QTL test statistic.



(d) Gm3095 transcript levels are not mediated by any of the 13 local genes. The tests of pleiotropy vs. separate QTL demonstrate evidence of separate QTL for some local expression traits when paired with Gm3095.

mediation analyses identify precisely the intermediate expression trait.

We recommend using both tests of pleiotropy vs separate QTL and mediation analyses when dissecting an expression trait QTL hotspot. In practice, mediation analyses, which involve a collection of univariate regressions at a small set of pre-specified markers, are less computationally intense than tests of pleiotropy vs. separate QTL, since the latter requires a two-dimensional QTL scan over the region of interest. This two-dimensional QTL scan often involves fitting more than ten thousand bivariate regression models. In light of the computational cost of testing pleiotropy vs separate QTL, future researchers may wish to use it as a follow-up to mediation analyses when examining expression trait hotspots.

Future research may investigate the use of polygenic random effects in the statistical models for mediation analysis. Additional methodological questions include approaches for declaring significant a mediation LOD difference proportion and consideration of other possible measures and scales of extent of mediation. Additionally, future researchers may wish to consider biological models that contain two mediators.

The social and health sciences have witnessed much methods research in mediation analysis. The field of statistical genetics has not fully adopted these strategies yet, but, given the nature of current and future data, many opportunities exist for translation of approaches from epidemiology to systems genetics. For example, the 2015 text from Vander Weele contains detailed discussions of many methods issues that arose in mediation analyses in epidemiology studies.

References

Baron, Reuben M and David A Kenny (1986). "The moderator–mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations." In: *Journal of personality and social psychology* 51.6, p. 1173.

Boehm, Frederick (2018). *qtl2pleio: Hypothesis test of close linkage vs pleiotropy in multiparental populations*. R package version 0.1.0.

Broman, Karl W (2012a). "Genotype probabilities at intermediate generations in the construction of recombinant inbred lines". In: *Genetics* 190.2, pp. 403–412.

— (2012b). "Haplotype probabilities in advanced intercross populations". In: *G3: Genes, Genomes, Genetics* 2.2, pp. 199–202.

Broman, Karl W. et al. (2018). "R/qtl2: software for mapping quantitative trait loci with high-dimensional data and multi-parent populations." In: *bioRxiv*. DOI: 10.1101/414748. eprint: https://www.biorxiv.org/content/biorxiv/early/2018/09/12/414748.full.pdf. URL: https://www.biorxiv.org/content/early/2018/09/12/414748.

Chick, Joel M et al. (2016). "Defining the consequences of genetic variation on a proteome-wide scale". In: *Nature* 534.7608, p. 500.

Crick, Francis HC (1958). "On protein synthesis". In: *Symp Soc Exp Biol.* Vol. 12. 138-63, p. 8.

Keller, Mark P et al. (2018). "Genetic Drivers of Pancreatic Islet Function". In: *Genetics*, genetics–300864.

Morgan, Andrew P et al. (2015). "The mouse universal genotyping array: from substrains to subspecies". In: *G3: Genes— Genomes— Genetics*, g3–115.

R Core Team (2018). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing. Vienna, Austria. URL: https://www.R-project.org/.

VanderWeele, Tyler (2015). *Explanation in causal inference*.

Gene symbol	Peak position	LOD
Mtfp1	163.52	17.79
Slc12a7	163.58	8.30
Gpa33	163.58	26.38
Tmem25	163.58	11.26
Klhl29	163.58	11.03
Slc9a3r1	163.58	8.40
Kif1c	163.58	11.12
Gata4	163.58	8.72
Ppp2r5b	163.58	7.80
Vil1	163.58	26.48
Aldh4a1	163.58	9.38
Cotl1	163.58	16.25
Tmprss4	163.58	18.30
Fhod3	163.58	17.62
Zfp750	163.58	7.83
Svop	163.58	10.12
Abcb4	163.58	16.59
Ccnjl	163.58	6.70
Sephs2	163.58	27.21
Pcdh1	163.58	12.48
Fat1	163.58	7.49
Sox4	163.58	14.71
Gm8206	163.58	11.33
Gm8492	163.58	10.47
Clcn5	164.02	10.49
Slc6a8	164.02	8.35
Bcmo1	164.02	47.31
Arrdc4	164.02	8.75
Cacnb3	164.02	47.06
Sec14l2	164.02	12.00

Pgrmc1	164.02	15.88
Baiap2l2	164.02	20.33
Recql5	164.02	33.97
Cpd	164.02	13.70
Degs2	164.02	15.20
Muc13	164.02	18.91
Clic5	164.02	10.39
Tff3	164.02	16.38
Myo7b	164.02	13.70
Afg3l2	164.02	19.12
Sema4g	164.02	23.39
Agap2	164.02	33.99
Plxna2	164.02	8.61
Aldob	164.02	20.99
Epb4.1l4b	164.02	9.21
Sel1l3	164.02	17.92
Sult1b1	164.02	17.63
Hpgds	164.02	31.14
Ush1c	164.02	21.90
Calml4	164.02	12.62
Fam83b	164.02	16.12
Myo15b	164.02	71.57
Inpp5j	164.02	11.57
Ttyh2	164.02	15.20
Cdhr2	164.02	30.00
Myrf	164.02	37.55
Sh3bp4	164.02	9.04
Vgf	164.02	15.05
Grtp1	164.02	23.88
B4galnt3	164.02	20.43
Gucy2c	164.02	12.02
Smim5	164.02	18.20

Nrip1	164.02	9.50
Clrn3	164.02	20.13
Acot4	164.02	12.17
Hunk	164.02	18.50
Zbtb16	164.02	6.69
Osgin1	164.02	13.51
Zfp541	164.02	25.28
2610042L04Rik	164.02	8.14
Gm9429	164.02	9.13
Agxt2	164.02	14.54
Gm17147	164.02	26.94
Gm8281	164.02	8.40
Ddx23	164.03	9.29
Map2k6	164.03	18.54
Npr1	164.03	9.68
Hdac6	164.03	11.28
Vav3	164.03	13.51
Atp11b	164.03	11.77
Aadat	164.03	8.08
Tmem19	164.03	20.60
Gbp4	164.03	7.14
Papola	164.03	12.14
Als2	164.03	18.35
Sult1d1	164.03	11.23
Myo6	164.03	10.27
Dnajc22	164.03	13.72
Unc5d	164.03	6.96
Gm3095	164.03	10.90
Gm26886	164.03	8.76
Eps8	164.06	12.49
Ddc	164.06	13.61
Fras1	164.06	10.06

Card11	164.06	7.20
Glyat	164.06	10.79
Pipox	164.08	15.90
Iyd	164.08	23.23
Man1a	164.26	8.77
Cdc42ep4	164.26	8.11
Ppara	164.26	8.84
Galr1	164.26	10.10
Ctdsp1	164.26	7.13
Eif2ak3	164.26	8.93
Misp	164.26	9.24
Sun1	164.26	7.35
Kctd8	164.26	9.76
Dcaf12l1	164.26	8.95
4930539E08Rik	164.26	8.25
Slc29a4	164.26	7.94
Gm3239	164.26	9.21
Gm3629	164.26	9.77
Gm3252	164.26	9.87
Gm3002	164.26	7.62
Ctsh	164.29	9.76
Dao	164.29	7.06
Ak7	164.31	10.07
Pcp4l1	164.35	9.11
Gm12929	164.62	37.80
Mgat1	164.63	10.10
Atg7	164.76	7.82
Gm11549	165.05	10.23
Ccdc111	165.12	7.77
Fam20a	165.15	7.59
Oscp1	165.16	7.99
Dsc2	165.28	8.06

Adam10	165.28	9.73
Plb1	165.28	8.85
Ccdc89	165.28	7.84
9930013L23Rik	165.28	7.93
Gm13648	165.28	7.08
Igfbp4	165.45	8.68
Ldlrap1	165.45	7.18
Cdh18	165.45	7.96
Arhgef10l	165.46	7.57
Cib3	165.48	8.70
Macf1	165.57	9.10
Fam63b	165.58	8.33
1190002N15Rik	165.63	9.02
Bcl2l14	165.71	7.54
Hist2h2be	165.88	8.96
Gm6428	165.89	7.32
Dennd5b	166.18	9.42
Gm12168	166.18	7.77
Gm12230	166.42	7.46
Acat3	166.61	7.17
Gpr20	166.84	8.41

Table 3: LOD peak positions and peak heights for 147 expression traits that map to the Chromosome 2 expression trait hotspot.