PLOS | GENETICS

A Statistical Framework for Joint eQTL Analysis in Multiple Tissues

Timothée Flutre[1,2]⁹, Xiaoquan Wen[3]⁹, Jonathan Pritchard[1,4], Matthew Stephens[1,5]*

1 Department of Human Genetics, University of Chicago, Chicago, Illinois, United States of America, 2 Department of Plant Genetics, Institut National de la Recherche Agronomique, Paris, France, 3 Department of Biostatistics, University of Michigan, Ann Harbor, Michigan, United States of America, 4 Howard Hughes Medical Institute, Chevy Chase, Maryland, United States of America, 5 Department of Statistics, University of Chicago, Chicago, Illinois, United States of America

Abstract

Mapping expression Quantitative Trait Loci (eQTLs) represents a powerful and widely adopted approach to identifying putative regulatory variants and linking them to specific genes. Up to now eQTL studies have been conducted in a relatively narrow range of tissues or cell types. However, understanding the biology of organismal phenotypes will involve understanding regulation in multiple tissues, and ongoing studies are collecting eQTL data in dozens of cell types. Here we present a statistical framework for powerfully detecting eQTLs in multiple tissues or cell types (or, more generally, multiple subgroups). The framework explicitly models the potential for each eQTL to be active in some tissues and inactive in others. By modeling the sharing of active eQTLs among tissues, this framework increases power to detect eQTLs that are present in more than one tissue compared with "tissue-by-tissue" analyses that examine each tissue separately. Conversely, by modeling the inactivity of eQTLs in some tissues, the framework allows the proportion of eQTLs shared across different tissues to be formally estimated as parameters of a model, addressing the difficulties of accounting for incomplete power when comparing overlaps of eQTLs identified by tissue-by-tissue analyses. Applying our framework to re-analyze data from transformed B cells, T cells, and fibroblasts, we find that it substantially increases power compared with tissue-by-tissue analysis, identifying 63% more genes with eQTLs (at FDR = 0.05). Further, the results suggest that, in contrast to previous analyses of the same data, the majority of eQTLs detectable in these data are shared among all three tissues.

Introduction

Regulatory variation plays an essential role in the genetics of disease and other phenotypes as well as in evolutionary change [1–3]. However, in sharp contrast to nonsynonymous variants in the human genome, which can now be identified with great accuracy, it remains extremely difficult to know which variants in the genome may impact gene regulation in any given tissue or cell type. [Henceforth we use "tissue" for brevity, but everything applies equally to cell types.] Expression QTL mapping (e.g. [4–6]) represents a powerful approach for bridging this gap, by allowing regulatory variants to be identified, and linked to specific genes. Indeed, numerous studies (e.g. [7,8]) have shown highly significant overlaps between eQTLs and SNPs associated with organismal-level phenotypes in genome-wide association studies (GWAS), suggesting that a large fraction of GWAS associations may be due to variants that affect gene expression.

Ultimately, understanding the biology of organismal pheno-types, such as diseases, is likely to require understanding regulatory variation in many different tissues ([9,10]). For example, if regulatory variants differ across tissues, then, in understanding GWAS hits, and using them to understand the biology of disease, we would like to know which variants are affecting which tissues.

At a more fundamental level, identifying differential genetic regulation in different tissues could yield insights into the basic biological processes that influence tissue differentiation. To date, eQTL studies have been performed in a relatively narrow range of tissue types. However, this is changing quickly: for example, the NIH "Genotype-Tissue Expression" (GTEx) project aims to collect expression and genotype data in 30 tissues across 900 individuals. Motivated by this, here we describe and illustrate a statistical framework for mapping eQTLs in expression data on multiple tissues.

While statistical methods for identifying eQTLs in a single tissue or cell type are now relatively mature (e.g. [11]) current analytic tools are limited in their ability to fully exploit the richness of data across multiple tissues. In particular, available methods fall short in their ability to *jointly analyze data on all tissues* to maximize power, while simultaneously *allowing for differences among eQTLs* present in each tissue. Indeed relatively few papers have considered the problem. The simplest approach (e.g. [12,13]) is to analyze data on each tissue separately ("tissue-by-tissue" analysis), and then to examine overlap of results among tissues. However, this fails to leverage commonalities among tissues to improve power to detect shared eQTLs. Furthermore, although examining overlap of eQTLs among tissues may appear a natural approach to

Author Summary

Genetic variants that are associated with gene expression are known as expression Quantitative Trait Loci, or eQTLs. Many studies have been conducted to identify eQTLs, and they have proven an effective tool for identifying putative regulatory variants and linking them to specific genes. Up to now most studies have been conducted in a single tissue or cell type, but moving forward this is changing, and ongoing studies are collecting data aimed at mapping eQTLs in dozens of tissues. Current statistical methods are not able to fully exploit the richness of these kinds of data, taking account of both the sharing and differences in eQTLs among tissues. In this paper we develop a statistical framework to address this problem, to improve power to detect eQTLs when they are shared among multiple tissues, and to allow for differences among tissues to be estimated. Applying these methods to data from three tissues suggests that sharing of eQTLs among tissues may be substantially more common than it appeared in previous analyses of the same data.

examining heterogeneity, in practice interpretation of results is complicated by the difficulty of accounting for incomplete power. Both [13] and [14] provide approaches to address this, but only for pairwise comparisons of tissues.

Compared with tissue-by-tissue analysis, joint analysis of multiple tissues has the potential to increase power to identify eQTLs that have similar effects across tissues. Both [15] and [16] conduct such joint analyses – the first using ANOVA, and the second using a weighted Z-score meta-analysis – and [16] confirm that their joint analysis has greater power than tissue-by-tissue analysis. The ANOVA and Z-score methods each have different advantages. The ANOVA framework has the advantage that, by including interaction terms, it can be used to investigate heterogeneity in eQTL effects among tissues. Gerrits *et al.* ([15]) use this to identify eQTLs that show significant heterogeneity, and then classify these eQTLs, post-hoc, into different types based on estimated effect sizes. The weighted Z-score method has the advantage that, unlike ANOVA, it allows for different variances of expression levels in different tissues (which are likely to occur in practice). However, it does not so easily allow for investigation of heterogeneity; Fu *et al.* ([16]) hence assess heterogeneity for pairs of tissues by using a resampling-based procedure to assess the significance of observed differences in Z scores. Other papers, including [17] and [18], also show that joint analyses provide more power. Our work goes beyond these papers in its modeling of heterogeneity, and in its use of a hierarchical model to borrow information across genes to estimate weights associated with different types of heterogeneity.

Here we introduce a statistical framework for the joint analysis of eQTLs among multiple tissue types, that combines advantages of some of the methods above, as well as introducing some new ones. In brief, our framework integrates recently-developed GWAS meta-analysis methods that allow for heterogeneity of effects among groups [19–23], into a hierarchical model (e.g. [24,25]) that combines information across genes to estimate the relative frequency of patterns of eQTL sharing among tissues. Like ANOVA, our approach allows investigation of heterogeneity among several tissues, not just pairs of tissues. However, in contrast to ANOVA, our framework allows for different variances in different tissues. Moreover, unlike any of the methods described above, our framework explicitly models the fact that some tissues may share eQTLs more than others, and estimates these patterns

of sharing from the data (a similar idea was applied to ChIP-Seq data by [26]). Our methods also allow for intra-individual correlations when samples are obtained from a common set of individuals. While we focus here on comparing and combining information across different tissue types, our framework could be applied equally to comparing and combining across other units, e.g. different experimental platforms, multiple datasets on the same tissue types, or data on individuals from different populations.

The remainder of the paper is as follows. After providing a brief overview of our framework, we use simulations to illustrate its power compared to other methods, and then apply it to map eQTLs, and assess heterogeneity among tissues, using data from Fibroblasts, LCLs and T-cells ([12]). Consistent with results from [16], we show that our joint analysis framework provides a large gain in power compared with a tissue-by-tissue analysis. Furthermore, compared with previous analyses of these data, we find a much higher rate of tissue-consistent eQTLs.

Results

Methods Overview

Consider mapping eQTLs in S tissues. In our applications here the expression data are from micro-arrays, and so we assume a normal model for the expression levels, suitably-transformed. (These methods can also be applied to RNA-seq data after suitable transformation; see Discussion). That is, in each tissue, $s = 1, \ldots, S$, we model the potential association between a candidate SNP and a target gene by a linear regression:

$$y_{si} = \mu_s + \beta_s g_i + \varepsilon_{si} \text{ with } \varepsilon_{si} \sim \mathcal{N}(0, \sigma_s^2), \quad (1)$$

where y_{si} denotes the observed expression level of the target gene in tissue s for the ith individual, μ_s the mean expression level of this gene in tissue s, β_s the effect of a candidate SNP on this gene expression in tissue s, g_i the genotype of the ith individual at the SNP (coded as 0,1 or 2 copies of a reference allele) and ε_{si} the residual error for tissue s and individual i. Note that the subscript s on residual variance σ_s^2 indicates that we allow the residual variance to be different in each tissue. In addition, when tissues are sampled from the same set of individuals, we allow that the residual errors $\varepsilon_{1i}, \ldots, \varepsilon_{Si}$ may be correlated (with the correlation matrix to be estimated from the data).

The primary questions of interest are whether the SNP is an eQTL in any tissue, and, if so, in which tissues. To address these questions we use the idea of a "configuration" from [21,23]. A configuration is a binary vector $\gamma = (\gamma_1, \ldots, \gamma_S)$ where $\gamma_s \in \{0,1\}$ indicates whether the SNP is an eQTL in tissue s. If $\gamma_s = 1$ then we say the eQTL is "active" in tissue s. The "global null hypothesis", H_0, that the SNP is not an eQTL in any tissue, is therefore $\gamma = (0, \ldots, 0)$. Every other possible value of γ can be thought of as representing a particular alternative hypothesis. For example, $\gamma = (1, \ldots, 1)$ represents the alternative hypothesis that the SNP is an eQTL in all S tissues, and $\gamma = (1,0, \ldots, 0)$ represents the alternative hypothesis that the SNP is an eQTL in just the first tissue.

Our aim is to perform inference for γ. A natural approach is to specify a probability model, $P(\text{data} | \gamma)$, being the probability of obtaining the observed data if the true configuration were γ, and then perform likelihood-based inference for γ. The support in the data for each possible value of γ, relative to the null H_0, is quantified by the

likelihood ratio, or *Bayes Factor* (BF, [27]):

$$\text{BF}_\gamma = \frac{\text{P(data|true configuration is } \gamma)}{\text{P(data|}H_0)}. \tag{2}$$

Specifying these likelihoods requires assumptions about $\text{P}(\beta|\gamma)$, the distribution of the effect sizes β for each possible configuration γ (as well as less crucial assumptions about nuisance parameters such as μ and σ_s). Of course, if $\gamma_s = 0$ then $\beta_s = 0$ by definition, but for the tissues where $\gamma_s = 1$ various assumptions are possible – for example, one could assume that the effect β_s is the same in all these tissues, or allow it to vary among tissues. Here we use a flexible family of distributions, $\text{P}(\beta|\gamma,\theta)$ (see Methods), where the hyper-parameters θ can be varied to control both the typical effect size, and the heterogeneity of effects across tissues (see below).

The value of BF_γ measures the support in the data for one specific alternative configuration γ, compared against the null hypothesis H_0. To account for the fact that there are many possible alternatives, the *overall* strength of evidence against H_0 at the candidate SNP is obtained by "Bayesian Model Averaging" (BMA), which involves averaging BF_γ over the possible alternative configurations γ, weighting each by its prior probability, $\eta_\gamma = \text{P}(\gamma|H_0 \text{ false})$:

$$\text{BF}_\text{BMA} = \frac{\text{P(data|}H_0 \text{ false)}}{\text{P(data|}H_0 \text{ true)}} = \sum_{\gamma \neq (0,\ldots,0)} \eta_\gamma \, \text{BF}_\gamma. \tag{3}$$

Further, under an assumption of at most one eQTL per gene, the overall evidence against H_0 for the entire gene (i.e. that the gene contains no eQTL in any tissue) is given by averaging BF_BMA across all candidate SNPs [28]. In either case, at either the SNP or gene level, large values of BF_BMA constitute strong evidence against H_0. BF_BMA has a direct Bayesian interpretation as the strength of the evidence against H_0, but here we also use it as a frequentist test statistic ([28,29]), assessing significance by permutation or simulation. The latter has the advantage that $p-$values and $q-$values obtained in this way are "valid" even if not all the prior assumptions are exactly correct.

Note that BF_BMA depends on the choice of (θ,η), and the power of BF_BMA as a test statistic is expected to depend on how well this choice of these hyper-parameters captures the range of alternative scenarios present in the data. Here we make use of three different choices:

- A "data-driven" choice, where the hyper-parameters are estimated from the data using a hierarchical model (HM, [30]) that combines information across all genes. We use $\text{BF}_\text{BMA}^\text{HM}$ to denote this choice.
- A "default" choice, which chooses η to cover a wide range of different possible alternative configurations, and θ is set to allow modest heterogeneity. We use BF_BMA to denote this choice.
- A "lite" choice, which puts weight only on the most extreme configurations (where the eQTL is active in only one tissue, or in all tissues), but compensates by setting θ to allow for more heterogeneity. We use BF_BMAlite to denote this choice.

Each of these choices has something to recommend it. The first, being data driven, is the most attractive in principle, but also the most complex to implement. The default choice is simpler to implement, and is included partly to demonstrate that one does not have to get the hyper-parameter values exactly "right" for

BF_BMAlite to be a powerful test statistic. Finally, BF_BMAlite has the advantage that it is easily applied to large numbers of tissues; neither of the other methods scales well, either computationally or statistically, with the number of tissues, because the number of terms in the sum in equation (3) is $2^S - 1$.

When there is strong evidence against H_0, the Bayes Factors can also be used to assess which alternative configurations γ are consistent with the data. Specifically the posterior probability on each configuration is:

$$\text{P(true configuration is } \gamma \text{ |data, } H_0 \text{ false)} = \frac{\eta_\gamma \, \text{BF}_\gamma}{\sum\limits_{\gamma \neq (0,\ldots,0)} \eta_\gamma \, \text{BF}_\gamma}, \tag{4}$$

and the posterior probability that the SNP is an eQTL in tissue s is obtained by summing the probabilities over configurations in which $\gamma_s = 1$:

$$\text{P(eQTL in tissue } s|\text{data, } H_0 \text{ false)} =$$
$$\sum_{\gamma:\gamma_s = 1} \text{P(true configuration is } \gamma \text{ |data, } H_0 \text{ false)}. \tag{5}$$

The second of these is particularly helpful when the data are informative for an eQTL in tissue s, but ambiguous in other tissues: in such a case the probability () will be close to 1, even though the "true" configuration will be uncertain (so none of the probabilities (4) will be close to 1). Because both (4) and (5) are sensitive to choice of hyper-parameters, we compute them using $\text{BF}_\text{BMA}^\text{HM}$ (where the hyper-parameters are estimated from the data).

Further details of methods used are provided in the Methods section.

Simulations

Power to detect eQTLs. We begin by comparing the ability of different methods to reject the global null hypothesis H_0; i.e. to detect eQTLs that occur in *any* tissues. We expect that a tissue-by-tissue analysis, which analyzes each tissue separately, will perform well for detecting eQTLs that are present in a single tissue. Conversely, we expect joint analysis of all tissues to perform well for detecting eQTLs that are present in all tissues. Our Bayesian model averaging (BMA) approach attempts, by averaging over different possible eQTL configurations, to combine the advantages of both types of analysis, and thus aims to perform well across all scenarios.

To assess this we performed a series of simulations, with five tissues measured in 100 individuals (and no intra-individual correlations). Each simulation consisted of 2,000 gene-SNP pairs (i.e. one candidate SNP per gene), half of which were "null" (i.e. the SNP was not an eQTL in any tissue), and the other half following an alternative hypothesis where the SNP was an eQTL in exactly k tissues, with k varying from 1 to 5. Thus, for example, the simulations with $k = 1$ assess power to detect eQTLs that are active in just one tissue, whereas the simulations with $k = 5$ assess power to detect eQTLs that are active in all five tissues. When simulating eQTLs that are active in multiple tissues we assumed their effects to be similar, but not identical, across tissues (see Methods). We applied four analysis methods to these data: 1) BF_BMA, being our Bayesian Model Averaging approach with default weights described above; 2) BF_BMAlite, being the computationally-scalable version of BMA described above; 3) ANOVA/ linear regression (ANOVA/LR) (c.f. [15] and see Methods), which jointly analyzes all tissues in a regression model, and compares the general alternative model (which allows a different genetic effect in

each tissue) with the null model (no effect in all tissues); and 4) a "tissue-by-tissue" analysis (c.f. [12]), where we use linear regression to test for an eQTL separately in each tissue, and take the minimum $p-$value across tissues as a test statistic. For simplicity we defer consideration of the more sophisticated of our approaches, BF_{BMA}^{HM}, to slightly more complex simulations described later. Each of these methods produces a test statistic for each SNP-gene pair, testing the global null hypothesis H_0. For each test statistic, we found the threshold that yielded a False Discovery Rate of 0.05 (based on the known null/alternative status of each SNP-gene pair), and assessed the effectiveness of each method by the number of discoveries it made at that FDR.

The results of these comparisons are shown in Figure 1A. As expected, for eQTLs that occur in just one tissue, the tissue-by-tissue analysis performs best. However, it is only slightly more effective than the joint analysis approaches in this setting. Conversely, the joint analysis approaches outperform the tissue-by-tissue analysis for eQTLs that occur in more than one tissue, with the gains becoming larger as the number of tissues sharing the eQTL increases. The BMA analyses generally perform similarly to one another, and outperform ANOVA/LR. This is presumably because our simulations involved eQTLs that have similar effects in each tissue, and our prior distribution $p(\beta|\gamma,\theta)$ explicitly up-weights eQTLs with this feature.

This first set of simulations assumed error variances to be equal among tissues. This assumption is made by ANOVA/LR, but not by the other methods, and is likely often to be violated in practice. To assess the effects of this we repeated the simulations, but with error variances differing among tissues. The results (Figure 1B) confirm that, relative to other methods, ANOVA/LR performs less well when error variances vary among tissues.

To assess performance in larger numbers of tissues we repeated the simulations above, but with 20 tissues (so $k=1,\ldots,20$). For this many tissues computing BF_{BMA} involves averaging over all $2^{20} > 10^6$ possible eQTL configurations, which is computationally inconvenient, so we omitted BF_{BMA} from this comparison. The results (Figure 1C) show that $BF_{BMAlite}$ performs similarly to the tissue-by-tissue analysis for eQTLs that occur in just one or two tissues, and outperforms it substantially for eQTLs occurring in many tissues. As expected, ANOVA/LR outperforms tissue-by-tissue analysis for eQTLs occurring in many tissues, but is noticeably less effective for eQTLs occurring in only one tissue, and performs consistently less well than $BF_{BMAlite}$.

In summary, these simulations illustrate the benefits of Bayesian Model Averaging as a general strategy for producing powerful test statistics: by explicitly averaging over a range of alternative models, these test statistics are able to achieve good power to detect a wide range of different types of eQTL.

Identifying eQTLs in particular tissues: borrowing information among tissues. Next we consider the benefits of jointly analyzing multiple tissues, even when the main goal is to identify eQTLs in a particular tissue of interest. For intuition, suppose for a moment that every eQTL is shared among all tissues. Then, from the simulation results above, we know that a joint analysis will identify more eQTLs overall, and hence more eQTLs in the tissue of interest. Of course, not all eQTLs are shared among all tissues, but some are, and some tissues may share eQTLs more than others. To allow for this, our hierarchical model attempts to infer the extent of such sharing (by estimating the configuration weights η_γ), and exploits any sharing that does occur to increase power to detect eQTLs in each tissue. By estimating sharing from the data, rather than assuming that all tissues share equally with one another (as do the simpler test statistics BF_{BMA} and $BF_{BMAlite}$ used above), we expect BF_{BMA}^{HM} to make more

effective use of sharing in the data to further improve power to identify eQTLs.

To illustrate this, we simulated eQTL data for five tissues. Some eQTLs were shared by all tissues, some were specific to each tissue, some were shared by Tissues 1 and 2 only, and some were shared by Tissues 3, 4 and 5. To show how the benefits of sharing can change with sample size, we simulated 60 samples for Tissue 1, and 100 samples for the others. This mimics a setting where Tissue 1 is harder to obtain than the other tissues, with Tissue 2 being the best proxy for Tissue 1.

We applied our Bayesian methods and a tissue-by-tissue analysis to these data, and assessed their ability to identify eQTLs in each tissue. For the tissue-by-tissue analysis the test statistic in each tissue is simply the linear regression $p-$value in that tissue. For our Bayesian methods, the test statistic in each tissue is the posterior probability of the SNP being an active eQTL in that tissue (5). Note that this posterior probability is computed from joint analysis of all tissues, and takes account of sharing of eQTLs among tissues. For example, consider a SNP showing modest association with expression in Tissue 1. If this SNP also shows strong association in the other tissues, then it will be assigned a higher probability of being an active eQTL in Tissue 1 than it would if it showed no association in the other tissues. For each method, separately in each tissue, we identified the threshold of the test statistic value that yields a FDR of 0.05 in that tissue, based on the true active/inactive status of each SNP in that tissue (known since this is simulated data). (A SNP that is an eQTL in some tissues but not others counts as a "false discovery" if it is called as an eQTL in a tissue where it is inactive.) For the Bayesian methods we obtained results both using "default" weights on configurations (BF_{BMA}), and using weights estimated from the data by the hierarchical model (BF_{BMA}^{HM}). The latter is expected to be more effective as it should learn, for example, that Tissue 1 shares more eQTLs with Tissue 2 than with other tissues.

The results (Figure 2) show that, for all tissues, our joint analyses outperform the tissue-by-tissue analysis. Further, BF_{BMA}^{HM} outperforms BF_{BMA}, demonstrating the benefits of learning patterns of sharing from the data. Finally, the gain of BF_{BMA}^{HM} is greater for Tissue 1 than for Tissue 2, illustrating that benefits of sharing information are greater for tissues with small sample sizes.

Furthermore, using the hierarchical model which pools all genes together, we can estimate the configuration proportions, i.e. η_γ. In the setting described above, we simulated one thousand eQTLs in each of 8 different configurations, as well as one thousand genes with no eQTLs. Averaged over replicates, the proportions are estimated to be in $[0.124-0.127]$ for each of the 8 active configurations (negligible differences between replicates). These estimates are fairly accurate knowing that the true proportion is $1/8 = 0.125$ for each configuration.

Analysis of eQTL Data in Three Cell Types from Dimas *et al.*

We now analyze data from [12], consisting of gene expression levels measured in fibroblasts, LCLs and T-cells from 75 unrelated individuals genotyped at approximately 400,000 SNPs. The data were pre-processed similarly to the original publication, as described in the Methods section. Throughout we focus on testing SNPs that lie within 1 Mb of the transcription start site of each gene (the "*cis* candidate region"), and on a subset of 5012 genes robustly expressed in all three cell-types.

Gain in power from joint analysis. First we assess the gain in power from mapping eQTLs in all three cell types jointly, using BF_{BMA}, compared with a "tissue-by-tissue" analysis similar to that
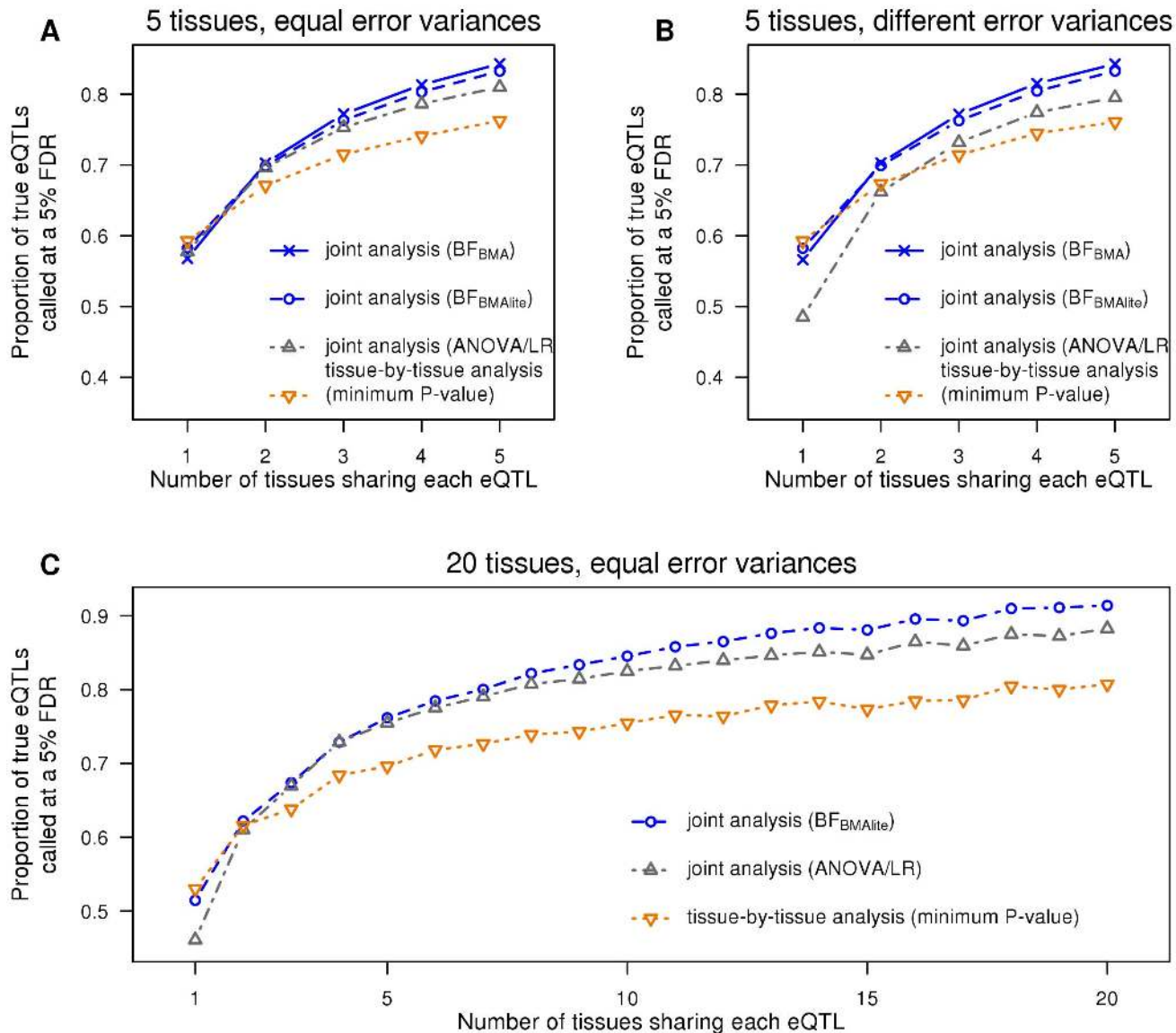
Figure 1. The BF_{BMA} **joint analysis has more power across a range of alternatives.** A. Five tissues are simulated, each with the error variance equal to 1. B. Five tissues are simulated, with error variances being 1, 1.5 or 2. C. Twenty tissues are simulated, each with the error variance equal to 1.
doi:10.1371/journal.pgen.1003486.g001

in [12]. For each method we compute a test statistic for each gene, combining information across SNPs, to assess the overall support for any eQTL in that gene in any tissue. For our Bayesian approach the test statistic is the average value of BF_{BMA} over all SNPs in the cis candidate region; for the tissue-by-tissue analysis the test statistic is the minimum $p-$value from linear regressions performed separately in each tissue for each SNP in the cis candidate region. We translate each of these test statistics into a $p-$value for each gene by comparing observed values with simulated values obtained under H_0 (by permutation of individual labels). Finally we computed, for each method, the number of genes identified as having an eQTL at an FDR of 0.05 (using the qvalue package [31]).

Joint mapping, via BF_{BMA}, substantially increased power to identify eQTLs compared with tissue-by-tissue analysis. For example, BF_{BMA} identified 1022 eQTLs at FDR = 0.05, which is 63% more than the 627 eQTLs identified by the tissue-by-tissue

analysis at the same FDR (Figure 3A, 3B, and 3C). Further, the vast majority of eQTLs identified by the tissue-by-tissue analysis (94%) are also detected by BF_{BMA} (Figure 3C).

In many cases, the eQTLs detected by BF_{BMA} but not by the tissue-by-tissue analysis have modest effects that are consistent across tissues. Because their effects are modest in each tissue, they fail to reach the threshold for statistical significance in any single tissue, and so the tissue-by-tissue analysis misses them. But because their effects are consistent across tissues, the joint analysis is able to detect them. Figure 4 shows an example illustrating this (gene *ASCC1*, Ensembl id ENSG00000138303, with SNP rs1678614). The PC-corrected phenotypes already indicate that this gene-SNP pair looks like a consistent eQTL (Figure 4A), and its effect size estimates are highly concordant across tissues (Figure 4B). However, as indicated by the $q-$values on the forest plot, this eQTL is not called by the tissue-by-tissue analysis in any tissue (all the $q-$values exceed .14). In contrast, the joint analysis pools
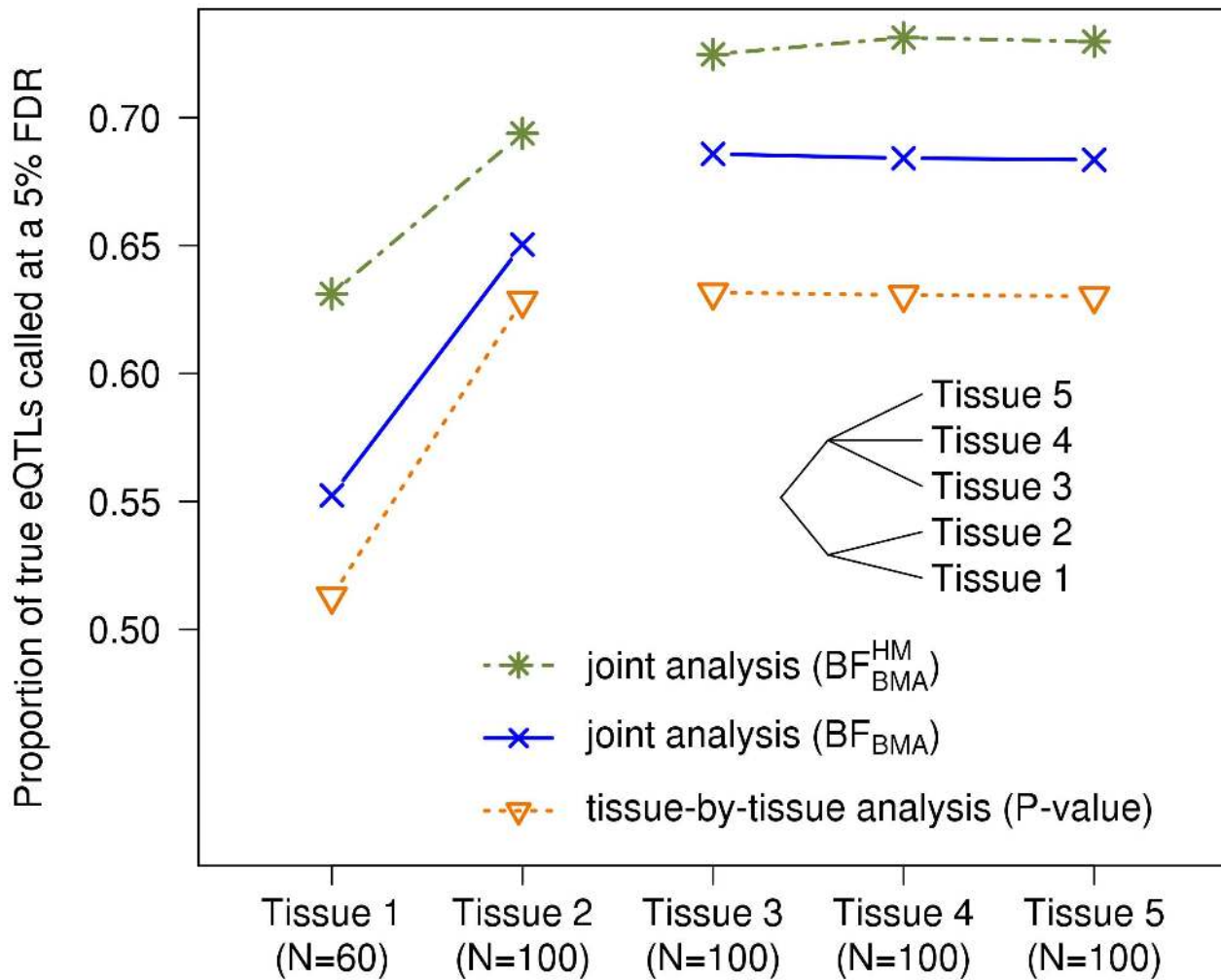
Figure 2. The $\mathrm{BF_{BMA}^{HM}}$ joint analysis efficiently borrows information across genes. Five tissues are simulated. Some eQTLs were shared by all tissues, some were specific to each tissue, and, as depicted by the cladogram, some were shared by Tissues 1 and 2 only, while others were shared by Tissues 3, 4 and 5. Each tissue has 100 samples, except tissue 1 which has only 60.
doi:10.1371/journal.pgen.1003486.g002

information across tissues to conclude that there is strong evidence for an eQTL ($q = 0.001$), and that it is likely an eQTL in all three tissues (probability 1 assigned to the consistent configuration $\gamma = (1,1,1)$, conditional on it being an eQTL).

Many eQTLs are consistent among tissues. The original analyses of these data concluded that 69% to 80% of eQTLs operate in a cell-type specific manner ([12]). These results were obtained by mapping eQTLs separately in each tissue, and then examining which of the eQTLs identified in each tissue also showed some signal (e.g. at a relaxed significance threshold of $p = 0.05$), in another tissue. However, as noted by [14], due to incomplete power, eQTLs that are actually shared between tissues may appear "tissue-specific" in this type of analysis. Our hierarchical model has the potential to help overcome this difficulty by estimating the proportion of eQTLs that follow each configuration type as a parameter of the model, combining information across all genes, and without setting specific significance thresholds (thus sidestepping the problems of incomplete power).

Applying the hierarchical model to these data produced an estimate of just 8% of eQTLs being specific to a single tissue, with an estimated 88% of eQTLs being common to all three tissues

(95% CI = 84%–93%; Table 1). Among eQTLs shared between just two tissues, many more are shared between LCLs and T-cells (Figure S1 shows such an example), than between these cell types and fibroblasts. This is consistent with results from [12], and perhaps expected since LCLs and T-cells are more similar to one another than to fibroblasts.

We obtained qualitatively similar patterns when we varied some of the assumptions in the hierarchical model - specifically, whether or not we allow for intra-individual correlations, whether or not we assume at most one eQTL per gene, whether or not we remove PCs to account for confounders, and whether or not we analyze all genes or only those genes robustly expressed in all tissues (Text S1). Nonetheless, we caution against putting too much weight on any particular number to quantify tissue specificity, not least because the definition of a tissue-specific eQTL is somewhat delicate: for example, it is unclear how to classify a SNP that is very strong eQTL in one tissue, and much weaker in the others. Further, our estimates necessarily reflect patterns of sharing only for moderately strong eQTLs, strong enough to be detected in the modest sample sizes available here: patterns of sharing could be different among weaker eQTLs. Nonetheless, these results do suggest that there is substantial sharing of
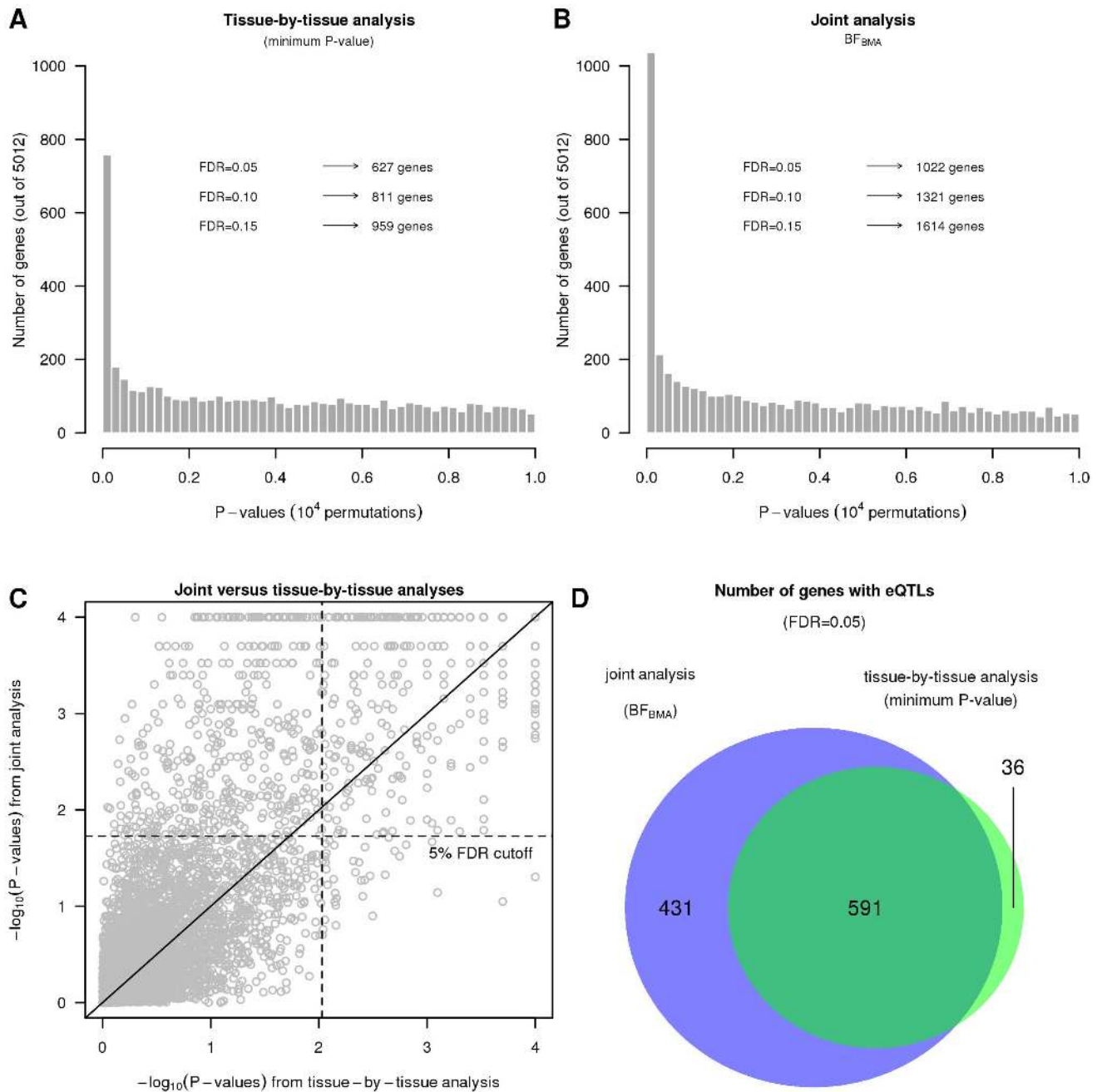
Figure 3. The $\mathrm{BF_{BMA}}$ joint analysis is more powerful on the data set from Dimas *et al*. A and B. Histograms of gene $p-$values obtained by the tissue-by-tissue analysis and the joint analysis. C. Scatter plot of the $p-$values from the joint analysis versus the $p-$values of the tissue-by-tissue analysis. D. Numbers of eQTLs called by both methods or either one of them.
doi:10.1371/journal.pgen.1003486.g003

eQTLs among these three tissue types, considerably higher than the original analysis suggested.

To illustrate the potential pitfalls of investigating heterogeneity in a tissue-by-tissue analyses, we also ran a tissue-by-tissue analysis on these data. Specifically, we called eQTLs separately in each tissue (at an FDR of 0.05), and then examined the overlap in the genes identified in each tissue. Using this procedure, in strong contrast with results from the joint analysis, 65% of eQTLs are called in only one tissue, with fewer than 15% called in all three tissues (Table 1). Qualitatively similar results are obtained for different FDR thresholds. However, these results cannot be taken

as reliable indications of tissue specificity, because the procedure fails to take account of incomplete power to detect eQTLs at any given threshold, and therefore tends to over-estimate tissue specificity. Figure 5 shows an eQTL that illustrates this behavior (gene *CHPT1*, Ensembl id ENSG00000111666, with SNP rs10860794). Visual examination of the expression levels in each genotype class (Figure 5A), suggest that this SNP is an eQTL in all three tissues, with similar effects in each tissue (Figure 5B). This is supported by the joint analysis, which shows strong evidence for an eQTL $q=0.001$, and assigns probability effectively 1 to the consistent configuration $\gamma=(1,1,1)$. However, as shown by the
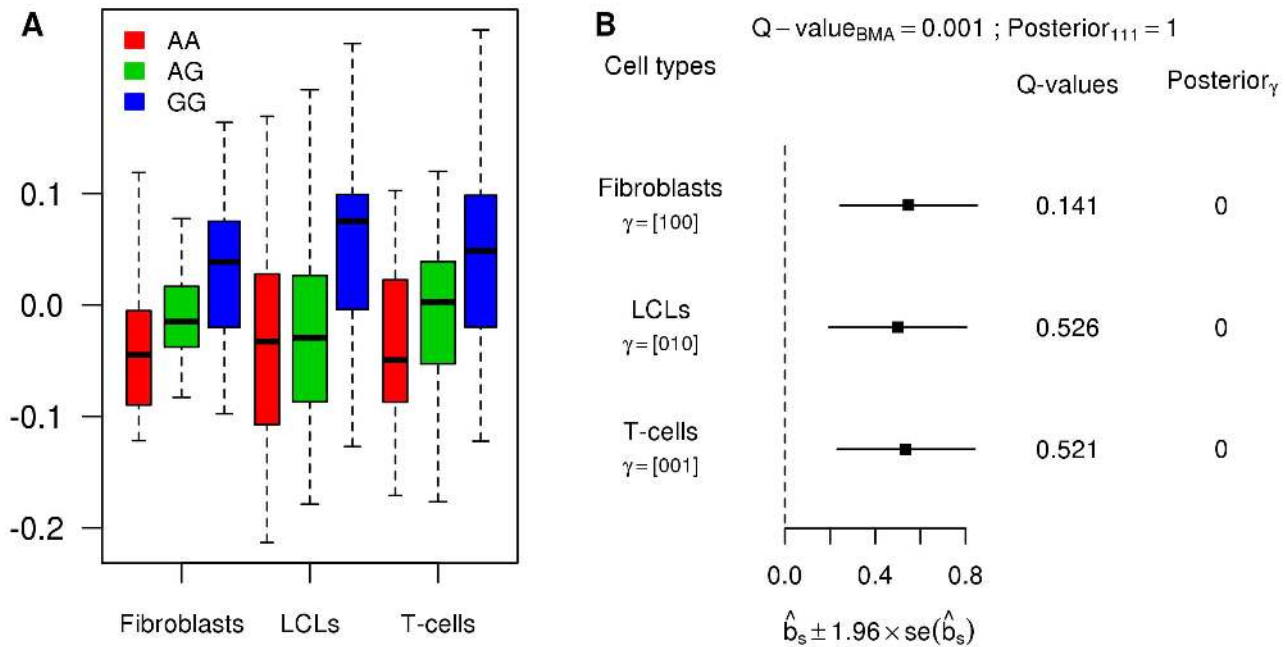
Figure 4. Example of an eQTL with weak, yet consistent effects. A. Boxplots of the PC-corrected expression levels from gene *ASCC1* (Ensembl id ENSG00000138303) in all three cell types, color-coded by genotype class at SNP rs1678614. B. Forest plot of estimated standardized effect sizes of this eQTL. Note that none of the $q-$values from the tissue-by-tissue analysis are significant at FDR = 0.05.
doi:10.1371/journal.pgen.1003486.g004

$q-$values, at an FDR of 0.05, the tissue-by-tissue analysis calls the eQTL only in fibroblasts.

Given the disagreement between the results from our novel framework and the original analyses of these data, we checked the plausibility of our results by applying a previously-used method for examining pairs of tissues to these data ([13]). This analysis takes the best eQTL in each gene identified in one tissue, and then estimates the proportion of these (π_1) that are also eQTLs in a second tissue (by applying Storey's method [31] to their nominal $p-$values in that second tissue, uncorrected for multiple comparisons). Unlike the tissue-by-tissue analysis above, this approach avoids thresholding of $p-$values, and makes some allowance for incomplete power. However, unlike our framework, this approach can only be applied to compare pairs of tissues.

Applying this approach to each pair yielded a mean estimate of $\pi_1 \approx 88\%$ (range 77% to 94%), broadly consistent with our qualitative conclusion that a substantial proportion of eQTLs are shared among tissues.

Discussion

In this work, we have presented a statistical framework for analyzing and identifying eQTLs, combining data from multiple tissues. Our approach considers a range of alternative models, one for each possible configuration of eQTL sharing among tissues. We compute Bayes Factors that quantify the support in the data for each possible configuration, and these are used both to develop powerful test statistics for detecting genes that have an eQTL in at least one tissue (by Bayesian model averaging across configurations), and to identify the tissue(s) in which these eQTLs are active (by comparing the Bayes factors for different configurations against one another). Our framework allows for heterogeneity of eQTL effects among tissues in which the eQTL is active, for different variances of gene expression measurements in each tissue, and for intra-individual correlations that may exist due to samples being obtained from the same individuals. For eQTL detection, our framework provides consistent, and sometimes substantial, gains in power compared to a tissue-by-tissue analysis and ANOVA or simple linear regression. Concerning the tissue specificity of eQTLs, our framework efficiently borrows information across genes to estimate configuration proportions, and then uses these estimates to assess the evidence for each possible configuration. When re-analyzing the gene expression levels in three cell types from 75 individuals ([12]), we found that there appears to be a substantial amount of sharing of eQTLs among tissues, substantially more than suggested by the original analysis.

In the next few years, we expect that expression data will be available on large numbers of diverse tissue types in sufficient

Table 1. Inference of the proportion of tissue specificity.

Configuration	Hierarchical model	Tissue-by-tissue
F-L-T	0.882 [0.840, 0.925]	0.187
L-T	0.051 [0.025, 0.085]	0.080
F-L	0.005 [0.000, 0.018]	0.050
F-T	0.002 [0.000, 0.011]	0.047
F	0.033 [0.014, 0.065]	0.246
L	0.015 [0.000, 0.039]	0.165
T	0.011 [0.000, 0.033]	0.224

The configurations are denoted here using the first letter of each tissue, e.g. "F-L-T" corresponds to the consistent configuration $\gamma = (1,1,1)$. The results for the hierarchical model were obtained with the multivariate Bayes Factors allowing correlated residuals and the EM algorithm. The results for the tissue-by-tissue analysis were obtained by calling eQTLs at an FDR of 0.05 after performing permutations in each tissue separately, and calculating the overlaps among tissues.
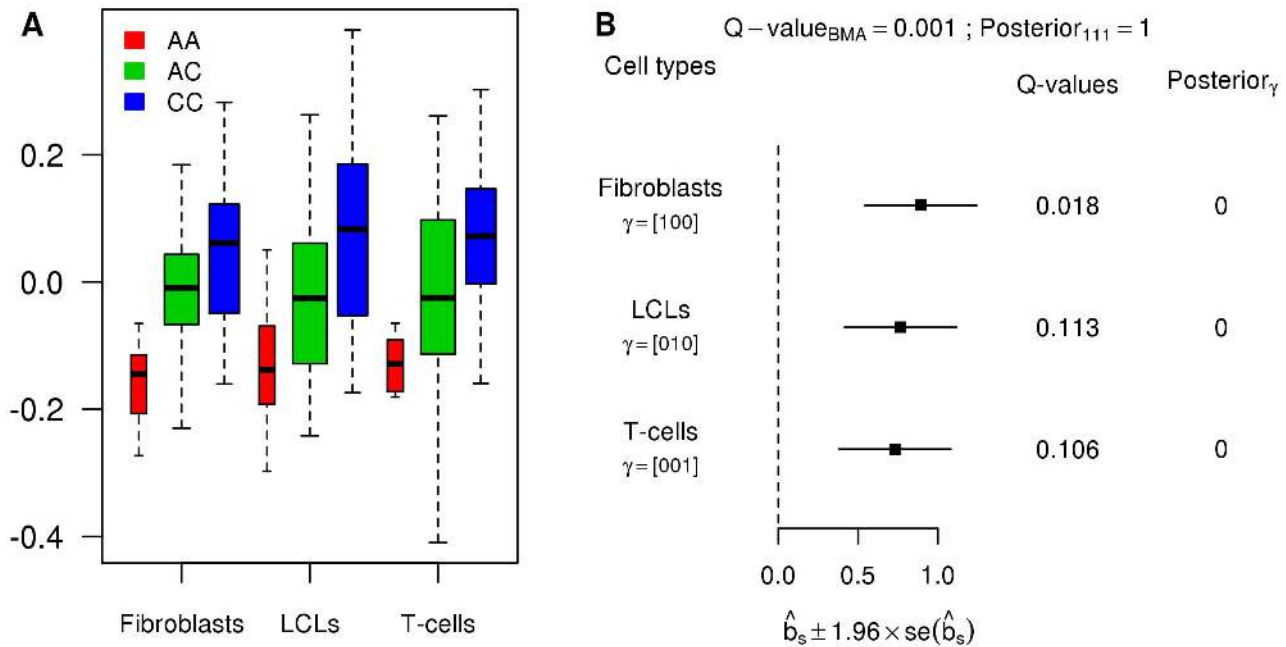doi:10.1371/journal.pgen.1003486.t001

Figure 5. Example of an eQTL wrongly called as tissue-specific by the tissue-by-tissue analysis. A. Boxplots of the PC-corrected expression levels from gene *CHPT1* (Ensembl id ENSG00000111666) in all three cell types, color-coded by genotype class at SNP rs10860794. B. Forest plot of estimated standardized effect sizes of this eQTL. Note that, from the $q-$values of the tissue-by-tissue analysis, the eQTL is significant at FDR = 0.05 only in fibroblasts.
doi:10.1371/journal.pgen.1003486.g005

sample sizes to allow eQTLs to be mapped effectively (for example, the NIH GTEx project aims to collect such data). The methods presented here represent a substantive step towards improved analyses that fully exploit the richness of these kinds of data. However, we also see several directions for potential extensions and improvements. First, our current framework can only partially deal with the challenges of large numbers of tissues. Specifically, because with S tissues, there are 2^S possible configurations of eQTL sharing among tissues, some of our current methods, which consider all possible configurations, will become impractical for moderate S (speculatively, above about 10, perhaps). Our test statistic $BF_{BMAlite}$ partially addresses this problem, by allowing for heterogeneity while averaging over only $S+1$ configurations, which is practical for very large S. Our simulation results suggest that $BF_{BMAlite}$ is a powerful test statistic for identifying SNPs that are an eQTL in at least one tissue. However our preferred approach for identifying *which* tissues such SNPs are active in involves a hierarchical model that estimates the frequency of different patterns of sharing from the data, and this hierarchical model scales poorly with S. In particular, having a separate parameter for each possible configuration is unattractive (both statistically and computationally) for large S, and alternative approaches will likely be required. There are several possible ways forward here: for example, one would be to reduce the number of distinct configurations by clustering "similar" configurations together; another would be to focus less on the discrete configurations, and instead to focus on modeling heterogeneity in effect sizes in a continuous way - perhaps using a mixtures of multivariate normal distributions with more complex covariance structures than we allow here. We expect this to remain an area of active research in the coming years, especially since these types of issues will likely arise in many genomics applications involving multiple cell types, and not only in eQTL mapping.

Another important issue to address is that most future expression data sets will likely be collected by RNA-seq, which provides count data that are not normally distributed. Previous eQTL analyses of RNA-seq (e.g. [32]) have nonetheless performed eQTL mapping using a normal model, by first transforming (normalized) count data at each gene to the quantiles of a standard normal distribution. Although this approach would not be attractive in experiments with small sample sizes, with the moderate to large sample sizes typically used in eQTL mapping experiments this approach works well. As a first step, this approach could also be used to apply our methods to count data. However, ultimately it would seem preferable to replace the normal model with a model that is better adapted to count-based data, perhaps a quasi-Poisson generalized linear model ([33]); Bayes Factors under these models could be approximated using Laplace approximations, similar to the approximations used here for the normal model [21]. The quasi-Poisson model has the advantage over the normal transformation approach that it preserves the fact that there is more information about eQTL effects in tissues where a gene is high expressed than in tissues where it is low expressed. This information is lost by normal transformation. In our primary analyses here we addressed this by analyzing only genes that were robustly expressed in all tissues, but this is sub-optimal, and will become increasingly unattractive as the number of tissues grows.

Our analyses here assess (cis) eQTL sharing among tissues by performing association testing at the level of individual SNPs. A different approach to investigating eQTL sharing among tissues is to study the "cross-heritability" of expression levels among tissues (e.g. [34,35]). These methods are based on polygenic models, and attempt to estimate the combined influence of all shared eQTLs; this contrasts with our analysis, where the focus is on sharing of individually-identifiable eQTLs of moderate-to-large effect. Both [34] and [35] estimate cross-tissue heritability to be low. [34],

studying expression in Blood and Adipose tissues from Icelanders, estimated cross-tissue heritability as $\xi^2 \approx 3\%$; [35] obtained an estimate of mean genetic correlation close to zero for Blood and LCLs in monozygotic twins ($\hat{r}_G = -0.031$). These results may appear to conflict with our results (both from our model-based approach, and the less-model-based pairwise analysis approach from [13]), which suggest that most large-effect cis eQTLs are shared among fibroblasts, LCLs and T cells. However, these low estimates of cross-tissue heritability reflect not only the extent of sharing of eQTLs, but also the absolute size of the eQTL effects. If eQTL effects are small, explaining only a small proportion of the total variance in gene expression, then cross-tissue heritability will be also small, even if all eQTLs have exactly the same effect in all tissues. Thus, to assess eQTL sharing in the heritability-based approaches, it is helpful to contrast cross-tissue heritability, ξ^2, with within-tissue heritability, h^2, (which is also affected by eQTL effect size, but not by sharing). Specifically, within the polygenic model it can be shown that the correlation coefficient ρ of the eQTL effects in two tissues a and b is: $\rho := \xi^2/\sqrt{h_a^2 h_b^2}$. Applying this to the cis estimates of ξ^2 and h_a^2, h_b^2 from [34], for adipose and blood, yields $\rho = 0.03/\sqrt{(0.055)(0.057)} = 0.53$. Although this estimate of effect correlation within a polygenic model, is not directly comparable with our estimate of sharing of eQTLs in a decidedly non-polygenic model (and for different cell types!), this result suggests that the two analyses may be less in conflict than they initially appear.

Methods

Software implementing our methods are available on the website http://stephenslab.uchicago.edu/software.html.

Bayesian Methods for Mapping Multiple-Tissue eQTLs

Models for multiple-tissue eQTLs. For each tissue, we model the potential genetic association between a target SNP and the expression levels of a target gene by the simple linear regression model (1). In vector form, this model is represented by

$$\boldsymbol{y}_s = \mu_s 1 + \beta_s \boldsymbol{g}_s + \boldsymbol{e}_s, \, \boldsymbol{e}_s \sim \mathcal{N}(0, \sigma_s^2 I), \qquad (6)$$

where s indexes one of the S tissue types examined and the vectors $\boldsymbol{y}_s, \boldsymbol{g}_s$ and $\boldsymbol{e}_s$ denote the expression levels, the genotypes of the samples and the residual errors respectively for the sth tissue type. The intercept term, μ_s, and the residual error variance, σ_s^2 are allowed to vary with tissue type. The regression coefficient β_s denotes the effect of the eQTL in tissue s, but we follow [21,28] in using the (unitless) standardized regression coefficient $b_s := \beta_s/\sigma_s$, as the main measure of effect size. As a result, inference is invariant to scale transformations of the response variables ($\boldsymbol{y}_s$) within each tissue.

When the tissue samples are taken from the same individuals we allow that the observations on the same individual may be correlated with one another. Specifically, let $E := (_1 \cdots_s)$ denote the $N \times S$ matrix of residual errors, we assume it to follow a matrix-variate normal (MN) distribution, i.e.,

$$E \sim MN(0, I, \Sigma). \qquad (7)$$

That is, the vectors $(\varepsilon_{1i}, \ldots, \varepsilon_{Si})$ are independent and identically distributed as $\mathcal{N}(0, \Sigma)$. The (unknown) $S \times S$ covariance matrix Σ quantifies the correlations between the S tissues; it can vary from gene to gene and is estimated from the data (see below). [When the

tissue samples are collected from different individuals then we assume their error terms are independent; methods for this case are given in [21].

Prior on effect sizes. A key component of our Bayesian model is the distribution $p(|\gamma, \theta)$, where θ denotes hyper-parameters that are to be specified or estimated from the data. (In the main text we used $p(\beta|\gamma, \theta)$ to simplify exposition, but we actually work with the standardized effects $\boldsymbol{b}$.) Of course, if $\gamma_s = 0$ then $b_s = 0$ by definition. So it remains to specify the distribution of the remaining b_s values for which $\gamma_s = 1$.

We use the distribution from [21] (see also [19,20]), which provides a flexible way to model the heterogeneity of genetic effects of an eQTL in multiple tissues. Specifically, [21] consider a distribution $p(|\phi, \omega, \gamma)$, with two hyper-parameters, ϕ, ω, in which the non-zero effects are normally distributed about some mean $\bar{b}$, which itself is normally distributed:

$$b_s | \bar{b}, \gamma_s = 1 \sim \mathcal{N}(\bar{b}, \phi^2), \qquad (8)$$

and

$$\bar{b} \sim \mathcal{N}(0, \omega^2). \qquad (9)$$

Note that $\phi^2 + \omega^2$ controls the variance (and hence the expected absolute size) of b_s, and $\phi^2/(\phi^2 + \omega^2)$ controls the heterogeneity (indeed, $\omega^2/(\phi^2 + \omega^2)$ is the correlation of $b_s, b_{s'}$ for different subgroups $s \neq s'$). If $\phi^2 = 0$ then this model corresponds to the "fixed effects" model in which the effects in all subgroups are equal (e.g. [23]).

To allow for different levels of effect size and heterogeneity, [21] use a fixed grid of values $\{(\phi_i, \omega_i) : i = 1, \ldots, L\}$, with the ith grid point having weight w_i. Thus

$$p(\boldsymbol{b}|\gamma, \theta) = \sum_i w_i p(\boldsymbol{b}|\phi_i, \omega_i, \gamma). \qquad (10)$$

In all our applications here we consider the grid of values fixed, and treat the weights $w_1, \ldots, w_L$ as hyper-parameters (so $\theta = (w_1, \ldots, w_L)$), which can be either fixed or estimated.

Choice of grid for (ϕ, ω). We define a grid of values for (ϕ, ω) by specifying a set A of values for the average effect size, $\omega^2 + \phi^2$, and a set H of values for the heterogeneity $\phi^2/(\phi^2 + \omega^2)$, and then taking the grid to be all $L = |A| \times |H|$ possible combinations of values. For all methods we use $A = \{0.1^2, 0.2^2, 0.4^2, 0.8^2, 1.6^2\}$, which is designed to span a wide range of eQTL effect sizes (see Text S1 as well as Figure S2). For BF_{BMA} and $\text{BF}_{\text{BMA}}^{\text{HM}}$ we allow for only a limited range of heterogeneity: $H = \{0, 0.25\}$. In this way we assume that when the eQTL is present in multiple tissues, it has a similar (but not necessarily identical) effect in each tissue. For $\text{BF}_{\text{BMAlite}}$ we allow a much wider range of heterogeneity: $H = \{0, 0.25, 0.5, 0.75, 1\}$. The rationale here is that the large heterogeneity values will help capture eQTLs that are present in only a subset of tissues, a feature that is not otherwise captured by $\text{BF}_{\text{BMAlite}}$ as it averages over a small number of configurations.

Choice of weights w and η. Let $|\gamma| := \sum_s \gamma_s$ denote the number of elements of γ that are equal to 1 (i.e. the number of tissues in which the eQTL is active in configuration γ), and recall that $\eta_\gamma = \text{P}(\gamma|H_0 \text{ false})$.

For BF_{BMA} we fix the weights η so that they put weight $1/S$ on all S possible non-zero values for $|\gamma|$ ($\in\{1, \ldots, S\}$), and, conditional on $|\gamma|$, put equal weight on all $\binom{S}{|\gamma|}$ configurations with that

value for $|\gamma|$. Thus $\eta_\gamma = (1/S)\binom{S}{|\gamma|}^{-1}$. In addition we fix the grid weights w to be equal on all grid values.

For $BF_{BMAlite}$ we put non-zero weights η on only the consistent configuration ($|\gamma|=S$) and configurations with an eQTL in a single tissue ($|\gamma|=1$). We set η so that it puts weight 0.5 on each of $|\gamma|=1$ and $|\gamma|=S$. Conditional on $|\gamma|=1$ we assume all S possibilities are equally likely. Thus $\eta_\gamma = 0.5$ if $\gamma = (1,1,1,\ldots,1)$ and $0.5/S$ if $|\gamma|=1$. Again, we fix the grid weights to be equal on all grid values (but with the larger grid for heterogeneity described above).

For BF_{BMA}^{HM} we estimate the weights, η from the data using a hierarchical model to combine information across genes, as described below.

Bayes Factor computation. To complete model specification, we use (limiting, diffuse) prior distributions for the nuisance parameters μ_s and Σ, as in [36]. Under these priors we can compute the Bayes Factor BF_γ in (2) using

$$BF_\gamma = \sum_{j=1}^{M} w_j BF_\gamma(\phi_j, \omega_j) \qquad (11)$$

where M is the total number of grid points and $BF_\gamma(\phi_j, \omega_j)$ is given by

$$
\begin{aligned}
BF_\gamma(\phi, \omega) &= \frac{p(Y|G, \phi, \omega, \gamma)}{p(Y|G, H_0)} \\
&= \frac{\int p(Y|G, \mu, b, \Sigma)p(\mu, \Sigma)p(b|\gamma, \phi, \omega)db\, d\mu\, d\Sigma}{\int p(Y|G, \mu, b=0, \Sigma)p(\mu, \Sigma)d\mu\, d\Sigma}
\end{aligned} \qquad (12)
$$

where Y and G denote the collection of expression levels and genotypes for a target gene-SNP pair across all tissue types respectively. We use analytic approximations for these Bayes Factors based on Laplace approximation, given in [21,36]. In particular, we use the approximation which in functional forms is connected to Frequentist's score statistic.

Bayesian Hierarchical Model

For BF_{BMA}^{HM} we use a hierarchical model, similar to [24,25], which combines information across genes, to estimate the grid weights η's and configuration weights η's. Following both [25,26] we make the simplifying assumption that each gene has at most one eQTL (which may be active in multiple tissues), and that each SNP is equally likely to be the eQTL. Let m_k be the number of SNPs in the *cis*-region for gene k. Then, if $BF_\gamma^{k,v}(\phi, \omega)$ denotes the Bayes Factor (12) computed for SNP v in gene k, the "overall Bayes Factor" measuring the evidence for an eQTL in gene k, BF^k, is obtained by averaging over the possible eQTL SNPs, the possible configurations γ, and the grid of values for ϕ, ω, weighting by their probabilities:

$$
\begin{aligned}
BF^k(\eta, w) &= \frac{p(\text{data at gene } k|\text{gene contains eQTL})}{p(\text{data at gene } k|\text{gene contains no eQTL})} \\
&= (1/m_k)\sum_{v=1}^{m_k}\sum_{j}\sum_{\gamma} \eta_\gamma w_j BF_\gamma^{k,v}(\phi_j, \omega_j).
\end{aligned} \qquad (13)
$$

Furthermore, if we let π_0 denote the probability that each gene follows the null (i.e. contains no eQTL) then the likelihood for gene k, as a function of π_0, η, w, is given by

$$
\begin{aligned}
L_k(\pi_0, \eta, w) &= (1-\pi_0)p(\text{data at gene } k|\text{gene contains eQTL}) \\
&+ \pi_0 p(\text{data at gene } k|\text{gene contains no eQTL})
\end{aligned} \qquad (14)
$$

$$\propto (1-\pi_0)BF^k + \pi_0 \qquad (15)$$

The overall likelihood for our hierarchical model is obtained by multiplying these likelihoods across genes:

$$L(\pi_0, \eta, w) = \prod_k L_k(\pi_0, \eta, w). \qquad (16)$$

Note that although the expression levels for different genes are not independent, because the SNPs being tested in different genes are mostly independent this independence assumption for the likelihoods across genes is a reasonable starting point. We have developed an EM algorithm to estimate the parameters (π_0, η, w) by maximum likelihood (see Supplementary information).

Relaxation of "one *cis*-eQTL per gene" assumption. To relax the "one *cis*-eQTL per gene" assumption we adopt the following procedure. First we compute the posterior probability of each SNP being the sole eQTL for each gene (i.e. only allowing one *cis*-eQTL per gene) with a set of default parameters, and use these to identify the top SNP for each gene (i.e. the one with the largest posterior probability of being the eQTL).

For each gene, separately in each tissue, we compute the residuals of its expression level after regressing out the effect of the top SNP. If these residuals are strongly associated with a SNP then this is evidence for that SNP being a second independent eQTL for that gene. Therefore, to allow for potentially more than one eQTL per gene we treat these residuals as defining a second set of "artificial" expression data for each gene and each tissue, and fit the hierarchical model using both the original and the artificial expression data.

Simulation Procedures

For our simulations, when simulating SNP-gene pairs, the genotypes at each SNP in each individual were simulated as Binomial(2,0.3): that is, with minor allele frequency 30% and assuming Hardy-Weinberg equilibrium. Phenotypes with eQTLs were simulated, with effect size based on an expected proportion of variance explained (PVE) of 20%; (see Text S1). For Figure 1A and 1B the error variances (one per tissue) were all equal to 1. For Figure 1C the error variances were randomly drawn from $\{1, 1.5, 2\}$, all equally likely.

The ANOVA/LR Method

The ANOVA/LR method uses the same linear model as our Bayesian methods (1), except that the residual errors σ_s are assumed to be equal across tissues s. Within this model we tested the global null hypothesis ($\beta_s = 0$ for all s) using an F test comparing the null model with the unconstrained alternative (β_s unconstrained). See Text S1.

Preprocessing of the Dataset from Dimas *et al.*

The phenotypes from Dimas *et al.* ([12]) were retrieved from the Gene Expression Omnibus (GSE17080). We mapped the 22,651 non-redundant probes to the hg19 human genome reference sequence (only the autosomes) using BWA ([37]), kept 19,965 probes mapping uniquely with at most one mismatch, and removed the probes overlapping several genes from Ensembl. This gave us 12,046 genes overlapped by 16,453 probes. For genes

overlapped by multiple probes, we chose a single probe at random. In our analyses we considered only genes that were robustly expressed in all tissues. A gene was considered robustly expressed in a given tissue if its mean expression level across individuals in this tissue was larger than or equal to the median expression level of all genes across all individuals in this tissue. As a result, we focused on 5012 genes.

Genotypes were obtained from the European Genome-phenome Archive (EGAD00000000027). We extracted the genotypes corresponding to the 85 individuals for which we had phenotypes and converted the SNP coordinates to the hg19 reference using liftOver ([38]). To detect outliers, we performed a PCA of these genotypes using individuals from the CEU, CHB, JPT and YRI populations of the HapMap project using EIGENSOFT ([39]). As in the original study, we identified 10 outliers and removed them from all further analyses, which were therefore performed on 75 individuals.

Gene expression measurements suffer from various confounders, many of which may be unmeasured ([40]), but which can be corrected for using methods such as principal components analysis (PCA). Following [32], we applied PCA in each tissue separately on the 5012×75 matrix of expression levels of each gene in each individual. We sorted principal components (PCs) according to the proportion of variation in the original matrix they explain, and selected PCs so that adding another PC would explain less than 0.0025% of the variation. As a result, this procedure identified 16 PCs in Fibroblasts, 7 in LCLs and 15 in T-cells. We then regressed out these PCs from the original matrix of gene expression levels, and used the residuals as phenotypes for all analyses.

All methods we compared assume that the errors are distributed according to a Normal distribution. Before analysis we therefore rank-transformed the expression levels at each gene to the quantiles of a standard Normal distribution ([28]).

Permutation Procedures

On the data set from Dimas et al., we assessed the performance of two methods, the tissue-by-tissue analysis and the BMA joint analysis, by comparing the number of genes identified as having at least one eQTL in any tissue, at a given FDR. For each method, we defined a test statistic, which was computed for each gene. For the tissue-by-tissue analysis, the test statistic is the minimum $p-\text{value}$ of the linear regressions between the given gene and each cis SNP in each tissue (so the minimum is taken across all SNPs and all tissues). For the BMA joint analysis, the test statistic is the average of the Bayes Factors for the given gene and each cis SNP. (When applying the tissue-by-tissue analysis to test for eQTLs in a single tissue, the test statistic is the minimum $p-\text{value}$ of the linear regressions between the given gene and each cis SNP in that tissue.)

In each case we converted the test statistic to a $p-\text{value}$ for each gene, testing the null hypothesis that the gene contains no eQTL in any tissue, by comparing the observed test statistic with the value of the test statistic obtained on permuted data obtained

by permuting the individuals labels (using the same permutations in each tissue to preserve any intra-individual correlations between gene expression in different tissues). Specifically, let P denote the total number of permutations (we used $P = 10^4$), T_g the value of the test statistic for gene g on the non-permuted data, and $T_g^{(i)}$ the value of the test statistic on the ith-permuted data. The $p-\text{value}$ for gene g from the tissue-by-tissue analysis is: $(1 + \sum_{i=1}^{P} 1_{T_g^{(i)} \leq T_g})/(1 + P)$. For the BMA joint analysis, the $p-\text{value}$ is: $(1 + \sum_{i=1}^{P} 1_{T_g^{(i)} \geq T_g})/(1 + P)$. Note that permutations were performed for each gene, since the null distribution of the test statistic will vary across genes (not least because the genes have different numbers of SNPs in their cis candidate region; see Figure S3).

From the $p-\text{value}$ calculated for each gene we estimate $q-\text{values}$ using the qvalue package ([31]), and determine the number of genes having at least one eQTL in any tissue at an FDR of α by computing the number of genes with $q \leq \alpha$.

When performing the tissue-by-tissue analysis on a single tissue, we performed the permutations in each tissue separately.

Supporting Information

Figure S1 Example of a strong, tissue-specific eQTL. A. Boxplots of the PC-corrected expression levels from gene *ANKDD1A* (Ensembl id ENSG00000166839) in all three cell types, color-coded by genotype class at SNP rs1628955. B. Forest plot of estimated standardized effect sizes of this eQTL. The posterior probability for configuration (0,1,1) is above 0.95, indicating that this eQTL is very likely to be active in LCLs and T-cells but not in fibroblasts.
(PDF)

Figure S2 Histogram of effect sizes when simulated according to the grid A.
(PDF)

Figure S3 Histogram of the number of SNPs in the cis region of each gene for the data set from Dimas et al.
(PDF)

Text S1 Supplementary methods and results.
(PDF)

Author Contributions

Conceived and designed the experiments: TF XW JP MS. Performed the experiments: TF XW. Analyzed the data: TF XW JP MS. Contributed reagents/materials/analysis tools: TF XW. Wrote the paper: TF XW JP MS.

References

1. Frazer KA, Murray SS, Schork NJ, Topol EJ (2009) Human genetic variation and its contribution to complex traits. Nat Rev Genet 10: 241–251.
2. Montgomery SB, Dermitzakis ET (2011) From expression QTLs to personalized transcriptomics. Nature reviews Genetics 12: 277–282.
3. Wray GA (2007) The evolutionary significance of cis-regulatory mutations. Nature Reviews Genetics 8: 206–216.
4. Cheung VG, Conlin LK, Weber TM, Arcaro M, Jen KY, et al. (2003) Natural variation in human gene expression assessed in lymphoblastoid cells. Nature Genetics 33: 422–425.
5. Stranger BE, Nica AC, Forrest MS, Dimas A, Bird CP, et al. (2007) Population genomics of human gene expression. Nature Genetics 39: 1217–1224.
6. Gilad Y, Rifkin SA, Pritchard JK (2008) Revealing the architecture of gene regulation: the promise of eQTL studies. Trends in genetics: TIG 24: 408–415.
7. Nica AC, Montgomery SB, Dimas AS, Stranger BE, Beazley C, et al. (2010) Candidate causal regulatory effects by integration of expression QTLs with complex trait genetic associations. PLoS Genet 6: e1000895. doi:10.1371/journal.pgen.1000895
8. Nicolae DL, Gamazon E, Zhang W, Duan S, Dolan ME, et al. (2010) Trait-associated SNPs are more likely to be eQTLs: annotation to enhance discovery from GWAS. PLoS Genet 6: e1000888. doi:10.1371/journal.pgen.1000888
9. Cookson W, Liang L, Abecasis G, Moffatt M, Lathrop M (2009) Mapping complex disease traits with global gene expression. Nat Rev Genet 10: 184–194.

10. Greenawalt DM, Dobrin R, Chudin E, Hatoum IJ, Suver C, et al. (2011) A survey of the genetics of stomach, liver, and adipose gene expression from a morbidly obese cohort. Genome Research 21: 1008–1016.

11. Stegle O, Parts L, Piipari M, Winn J, Durbin R (2012) Using probabilistic estimation of expression residuals (PEER) to obtain increased power and interpretability of gene expression analyses. Nat Protocols 7: 500–507.

12. Dimas AS, Deutsch S, Stranger BE, Montgomery SB, Borel C, et al. (2009) Common regulatory variation impacts gene expression in a cell type-dependent manner. Science 325: 1246–1250.

13. Nica AC, Parts L, Glass D, Nisbet J, Barrett A, et al. (2011) The architecture of gene regulatory variation across multiple human tissues: the MuTHER study. PLoS Genet 7: e1002003. doi:10.1371/journal.pgen.1002003

14. Ding J, Gudjonsson JE, Liang L, Stuart PE, Li Y, et al. (2010) Gene expression in skin and lymphoblastoid cells: Refined statistical method reveals extensive overlap in cis-eQTL signals. American Journal of Human Genetics 87: 779–789.

15. Gerrits A, Li Y, Tesson BM, Bystrykh LV, Weersing E, et al. (2009) Expression Quantitative Trait Loci Are Highly Sensitive to Cellular Differentiation State. PLoS Genet 5: e1000692. doi:10.1371/journal.pgen.1000692

16. Fu J, Wolfs MGM, Deelen P, Westra HJ, Fehrmann RSN, et al. (2012) Unraveling the Regulatory Mechanisms Underlying Tissue-Dependent Genetic Variation of Gene Expression. PLoS Genet 8: e1002431. doi:10.1371/journal.pgen.1002431

17. Petretto E, Bottolo L, Langley SR, Heinig M, McDermott-Roe C, et al. (2010) New insights into the genetic control of gene expression using a bayesian multi-tissue approach. PLoS Comput Biol 6: e1000737. doi:10.1371/journal.-pcbi.1000737

18. Li J, Tseng GC (2011) An adaptively weighted statistic for detecting differential gene expression when combining multiple transcriptomic studies. The Annals of Applied Statistics 5: 994–1019.

19. Lebrec JJ, Stijnen T, van Houwelingen HC (2010) Dealing with heterogeneity between cohorts in genomewide SNP association studies. Statistical Applications in Genetics and Molecular Biology9.

20. Han B, Eskin E (2011) Random-Effects Model Aimed at Discovering Associations in Meta-Analysis of Genome-wide Association Studies. The American Journal of Human Genetics 88: 586–598.

21. Wen X, Stephens M (2011). Bayesian Methods for Genetic Association Analysis with Heterogeneous Subgroups: from Meta-Analyses to Gene-Environment Interactions. URL http://arxiv.org/abs/1111.1210. 1111.1210.

22. Bhattacharjee S, Rajaraman P, Jacobs KB, Wheeler WA, Melin BS, et al. (2012) A subset-based approach improves power and interpretation for the combined analysis of genetic association studies of heterogeneous traits. American journal of human genetics 90: 821–835.

23. Han B, Eskin E (2012) Interpreting Meta-Analyses of Genome-Wide association studies. PLoS Genet 8: e1002555. doi:10.1371/journal.pgen.1002555

24. Veyrieras JB, Kudaravalli S, Kim SY, Dermitzakis ET, Gilad Y, et al. (2008) High-resolution mapping of expression-QTLs yields insight into human gene regulation. PLoS Genet 4: e1000214. doi:10.1371/journal.pgen.1000214

25. Maranville JC, Luca F, Richards AL, Wen X, Witonsky DB, et al. (2011) Interactions between glucocorticoid treatment and cis-regulatory polymorphisms contribute to cellular response phenotypes. PLoS Genet 7: e1002162. doi:10.1371/journal.pgen.1002162

26. Ferguson JP, Cho JH, Zhao H (2012) A new approach for the joint analysis of multiple ChIP-seq libraries with application to histone modification. Statistical applications in genetics and molecular biology 11.

27. Kass RE, Raftery AE (1995) Bayes Factors. Journal of the American Statistical Association 90: 773–795.

28. Servin B, Stephens M (2007) Imputation-based analysis of association studies: candidate regions and quantitative traits. PLoS Genet 3: e114. doi:10.1371/journal.pgen.0030114

29. Good IJ (1992) The Bayes/Non-bayes compromise: A brief review. Journal of the American Statistical Association 87: 597–606.

30. Gelman A, Hill J (2006) Data Analysis Using Regression and Multilevel/Hierarchical Models. Cambridge University Press, 1 edition. URL http://www.worldcat.org/isbn/0521867061.

31. Storey JD, Tibshirani R (2003) Statistical significance for genomewide studies. Proceedings of the National Academy of Sciences of the United States of America 100: 9440–9445.

32. Pickrell JK, Marioni JC, Pai AA, Degner JF, Engelhardt BE, et al. (2010) Understanding mechanisms underlying human gene expression variation with RNA sequencing. Nature 464: 768–772.

33. Sun W (2012) A statistical framework for eQTL mapping using RNA-seq data. Biometrics 68: 1–11.

34. Price AL, Helgason A, Thorleifsson G, McCarroll SA, Kong A, et al. (2011) Single-tissue and crosstissue heritability of gene expression via identity-by-descent in related or unrelated individuals. PLoS Genet 7: e1001317. doi:10.1371/journal.pgen.1001317

35. Powell JE, Henders AK, McRae AF, Wright MJ, Martin NG, et al. (2012) Genetic control of gene expression in whole blood and lymphoblastoid cell lines is largely independent. Genome Research 22: 456–466.

36. Wen X (2012). Bayesian Analysis of Multiway Tables in Association Studies: A Model Comparison Approach. URL http://arxiv.org/abs/1208.4621. 1208.4621.

37. Li H, Durbin R (2009) Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics (Oxford, England) 25: 1754–1760.

38. Hinrichs AS, Karolchik D, Baertsch R, Barber GP, Bejerano G, et al. (2006) The UCSC Genome Browser Database: update 2006. Nucleic acids research 34: D590–D598.

39. Price AL, Patterson NJ, Plenge RM,Weinblatt ME, Shadick NA, et al. (2006) Principal components analysis corrects for stratification in genome-wide association studies. Nature genetics 38: 904–909.

40. Leek JT, Scharpf RB, Bravo HC, Simcha D, Langmead B, et al. (2010) Tackling the widespread and critical impact of batch effects in high-throughput data. Nature Reviews Genetics 11: 733–739.