Additive, epistatic, and environmental effects through the lens of expression variability QTLs in a twin cohort

Gang Wang¹,§, Ence Yang¹,§, Candice L. Brinkmeyer-Langford¹, James J. Cai¹,²,*

¹ Department of Veterinary Integrative Biosciences, Texas A&M University, College Station, TX 77843, USA.

² Interdisciplinary Program in Genetics, Texas A&M University, College Station, TX 77843, USA.

§ Equally contributing authors

*Corresponding author – jcai@tamu.edu

Keywords: expression variability, evQTL, additive genetic effect, genetic interaction, epistasis, twin cohort

Running title: Dissecting evQTLs in a twin cohort
Abstract

The expression of a gene can vary across individuals in the general population, as well as between monozygotic twins. This variable expression is assumed to be due to the influence of both genetic and nongenetic factors. Yet little evidence supporting this assumption has been obtained from empirical data. In this study, we used expression data from a large twin cohort to investigate the influences of genetic and nongenetic factors on variable gene expression. We focused on a set of expression variability QTLs (evQTLs)—i.e., genetic loci associated with the variance, as opposed to the mean, of gene expression. We identified evQTLs for 99, 56, and 79 genes in lymphoblastoid cell lines, skin, and fat, respectively. The differences in gene expression, measured by the relative mean difference (RMD), tended to be larger between pairs of dizygotic (DZ) twins than between pairs of monozygotic (MZ) twins, showing that genetic background influenced the expression variability. Furthermore, a more profound RMD was observed between pairs of MZ twins whose genotypes were associated with greater expression variability than the RMD found between pairs of MZ twins whose genotypes were associated with smaller expression variability. This suggests that nongenetic (e.g., environmental) factor contributes to the variable expression. Lastly, we demonstrated that the formation of evQTLs is likely due to partial linkages between eQTL SNPs that are additively associated with the mean of gene expression; in most cases, no epistatic effect is involved. Our findings have implications for understanding divergent sources of gene expression variability.
Introduction
Variation and variability are central concepts in biology (Hallgrímsson and Hall 2005). Although often used interchangeably in the scientific literature, the two are not synonymous. Variation refers to the differences among individuals, whereas variability refers to the potential of a population to vary (Wagner 1995; Wagner and Altenberg 1996). In many cases, greater phenotypic variability (e.g., transcriptional noise) is disadvantageous (Kemkemer et al. 2002; Bahar et al. 2006; Wang and Zhang 2011) unless it gives rise to greater organismal plasticity—first at the level of an individual organism and eventually at the population level. Genetic factors resulting in more variable phenotypes become favored when they enable a population to more effectively respond to environmental changes (Hill and Zhang 2004; Kaern et al. 2005; Acar et al. 2008; Zhang et al. 2009). Thus, understanding to what extent and in what ways genotypes influence phenotypic variability is of fundamental importance.

Much effort has been focused on identifying genetic loci such as expression quantitative trait loci, or eQTL (Stranger et al. 2005; Stranger et al. 2007; Choy et al. 2008; Montgomery et al. 2010; Pickrell et al. 2010; Montgomery and Dermitzakis 2011), that affect the average value of a phenotype, while ignoring those that affect the variance of a phenotype. However, there is increasing evidence across species for genetic loci that affect the variance of phenotype (Queitsch et al. 2002; Jimenez-Gomez et al. 2011; Ronnegard and Valdar 2011; Perry et al. 2012; Shen et al. 2012; Yang et al. 2012). Recently we introduced the concept of expression variability QTL, or evQTL (Hulse and Cai 2013). By definition, an evQTL is a genetic locus linked to or associated with genetic variation influencing the variance of gene expression in a population. To identify evQTLs, we previously adapted the method developed by Ronnegard and Valdar (2011) for detecting vQTLs based on the double generalized linear model (dglm) (Verbyla and Smyth 1998). The dglm method tests for expression variances and measures the contribution of genetic variants to the expression heteroscedasticity. It compares the fit of a full model, which takes into account the contribution of genotype to both the mean and the variance of gene expression simultaneously, and a mean model, which only takes into account the contribution of genotype to the mean and ignores the contribution to the variance. A significant result of dglm test shows the nonrandom
association between genotypes and gene expression variances. Using this method, we have conducted a genome-wide scan for evQTLs in the human genome (Hulse and Cai 2013).

How an evQTL is created in the first place is not clear. We consider two possible scenarios, emphasizing either the role of environmental or genetic factors. The first possibility is that specific genetic variants disrupt the stabilizing genetic architecture that buffers stochastic variation in phenotype. As a result of such an effect of decanalization, along with the sensitization of the stabilizer (e.g., heat-shock protein 90), the phenotype becomes more sensitive to external environment and varies more greatly between individuals (Ronnegard and Valdar 2011; Hulse and Cai 2013).

The other possible scenario concerns the role of genetic interactions in the formation of evQTLs. Through either epistatic or non-epistatic (e.g., additive and dominance) effects within or between loci, genetic interactions contribute to genotypic variance. Epistasis may increase the variance of a quantitative trait (Pare et al. 2010; Ronnegard and Valdar 2012). However, it is extremely difficult to distinguish the contributions of epistatic or non-epistatic effects to variable expression of genes. Epistasis, in particular, is known to produce predominantly additive and dominance genetic variance when the low-frequency alleles of some SNPs are involved (Cheverud and Routman 1995).

Here we investigated the distribution and formation of evQTLs by leveraging the existing dataset (Grundberg et al. 2012a) derived from a population-based cohort of twin studies (Moayyeri et al. 2013b). We interrogated this dataset for evQTLs, and investigated the roles of genetic and nongenetic factors in the formation of the evQTLs we identified. The twin cohort offered a unique advantage for studying the relative contributions of various factors that influence expression variability. Importantly, comparing expression data of monozygotic and dizygotic twins allowed us to distinguish between genetic and nongenetic effects. In the following sections, we present the descriptive statistics for expression variability in the twin cohort, describe the detection of evQTLs, and finally estimate the relative contributions of genetic and nongenetic factors, as well as epistatic and non-epistatic effects, to the creation of evQTLs.
Materials & Methods

The TwinsUK dataset

We obtained the TwinsUK dataset including both genotype and expression data, which had been used in the eQTL study of (GRUNDBERG et al. 2012b). Here, we briefly describe the cohort and data processing performed in this previous study (GRUNDBERG et al. 2012b). The TwinsUK cohort includes 856 female individuals of European descent recruited from the TwinsUK Adult twin registry (SPECTOR and WILLIAMS 2006; MOAYYERI et al. 2013a). Subcutaneous adipose tissue, skin tissue, and LCLs were collected from each individual. Genotyping was performed with a combination of Illumina HumanHap300, HumanHap610Q, 1M-Duo and 1.2MDuo 1M chips. Genotypes were called with the Illuminus calling algorithm (TEO et al. 2007), and SNPs were filtered for minor allele frequency (MAF) of >5%. Gene expression levels were measured in LCLs, skin, and adipose (GRUNDBERG et al. 2012b). Expression profiling of the samples, each with either two or three technical replicates, was performed using Illumina Human HT-12 V3 BeadChips (Illumina). All samples were randomized before array hybridization, and replicates were hybridized on different BeadChips. Raw data were imported to Illumina BeadStudio software, and probes with less than three beads present were excluded. Log₂-transformed expression signals were normalized separately per tissue, with quantile normalization of the replicates of each individual followed by quantile normalization across all individuals (NICA et al. 2011).

In this study, we used available gene expression data for both individuals of a twin pair. All 48,804 probe sequences were mapped by BLAST to the reference genome (hg18), and probes found to map to more than one location were not used. Polymorphisms in the target mRNA sequence can greatly affect the binding affinity of microarray probe sequences, leading to false-positive and false-negative signals with any other polymorphisms in LD (RAMASAMY et al. 2013). In order to control for this, we used a comprehensive compendium of SNPs in CEU of the 1,000 Genomes Project (THE-1000-GENOMES-PROJECT-CONSORTIUM et al. 2012) to remove an additional 13,600 probes found to anneal in regions with SNPs present at a MAF of 5% or greater. Similarly,
probes mapping to non-autosomal locations were excluded from further analysis. Finally, 35,078 probes were left for our analysis.

The coefficient of variation (CV) is used as a normalized measure of the dispersion of expression distribution (Maheshri and O'Shea 2007; Ansel et al. 2008; Ronnegard and Valdar 2011). The CV was computed as \( CV = \frac{\sigma}{\mu} \), where \( \sigma \) and \( \mu \) are the standard deviation and the mean of gene expression levels, respectively. LD block plots were obtained by using HaploView (Barrett et al. 2005).

**Identification of evQTLs using the dglm method**

First we used the F-K test filter to greatly reduce the number of SNPs for computationally intensive model fitting. We then adapted the dglm method (Verbyla and Smyth 1998) to test for inequality in expression variances and measure the contribution of genetic variants to the expression heteroscedasticity. We considered the following model: 

\[
y_i = \mu + x_i \beta + g_i \alpha + \epsilon_i, \quad \epsilon_i \sim N(0, \sigma^2 \exp(g_i \theta)),
\]

where \( y_i \) indicates a gene expression trait of individual \( i \), \( g_i \) is the genotype at the given SNP (encoded as 0, 1, or 2 for homozygous rare, heterozygous and homozygous common alleles, respectively), \( \epsilon_i \) is the residual with variance \( \sigma^2 \), and \( \theta \) is the corresponding vector of coefficients of genotype \( g_i \) on the residual variance. Age of subjects and the batch of data collection were modeled as covariates \( x_i \). With this full model, both mean and variance of expression \( y_i \) were controlled by SNP genotype \( g_i \). We coded the fitting procedure using the dglm package in R. A snippet of R code for the dglm analysis is available in the supporting information of (Ronnegard and Valdar 2011). We assumed that the input gene expression data were approximately normally distributed, conditional on the evQTL and covariates, and set family=gaussian in the dglm R code to specify the error distribution and link function used. We tested for each input probe-SNP pair and obtained two \( P \)-values: \( P_{\text{dispersion}} \) and \( P_{\text{mean}} \), for the effects of genotypes on the variance and the mean of expression levels, respectively (Ronnegard and Valdar 2011). Probe-gene pairs that did not make the algorithm converge during computation were discarded. To control for the effect of outlier expression data points, permutation tests (Stranger et al. 2005) were conducted for all \( P_{\text{dispersion}} \) significant pairs. Specifically, for each
probe-SNP pair, we performed 10,000 permutations of expression phenotype relative to SNP genotypes. An association was considered significant if the $P$-value from the analysis of the observed $P_{\text{dispersion}}$ was lower than the threshold of the 0.001 tail of the distribution of the $P_{\text{dispersion}}$ from the 10,000 permutations ($P_{\text{permutation}} < 0.001$).

**Single-cell expression and mRNA decay rate**

Expression level of 96 genes were measured in 1,440 single lymphoblastoid single cells by qPCR assays in another study (LIVAK et al. 2013). We used this data to compute the CV of expression of the same gene in different cells. The mRNA decay rates of 16,823 genes were estimated in 70 human LCLs (PAI et al. 2012). We obtained the mRNA decay rate data to compute the average mRNA decay rate for each gene among these LCL samples.

**Identification of interacting SNPs**

We used a two-step procedure to identify SNPs that may “interact” with evSNPs. Assuming an additive interaction between the SNP to be identified and an evSNP, we first partitioned individuals into L and S groups according to genotypes of the evSNP, which were associated with large (L) and small (S) variances of gene expression. Next we scanned genome-wide SNPs. For each SNP, we computed heterozygosity of the polymorphic site among individuals in L and S groups as $Het_L = 1 - \sum_{i=1}^{n} p_i^2$ and $Het_S = 1 - \sum_{i=1}^{n} p_i^2$, respectively, where $p_i$ are allele frequencies of SNPs and, for diallelic SNPs, $n = 2$. All SNPs were then ranked by the value of $Het_L - Het_S$ and the top 100 SNPs with largest values were selected for further analysis. In the next step, a typical eQTL (not evQTL) analysis (STRANGER et al. 2005) was conducted among individuals of the L group. In other words, for each top SNP with high genotype heterozygosity difference, a simple linear regression was performed between the SNP’s genotypes and gene expression. The most significant SNPs were retained after applying a Bonferroni adjusted $P$-value cutoff = 0.05 and were reported as candidate interacting SNPs. To maintain sample independence, only one group of the twin sets was used in this analysis.
Results

Expression and genotype data
To investigate the genetic influences underlying variable gene expression, we revisited the published expression data (GRUNDBERG et al. 2012a) of the MuTHER (Multiple Tissue Human Expression Resource) project (NICA et al. 2011). In that study, gene expression was measured for lymphoblastoid cell line (LCL), adipose tissue (subcutaneous fat), and skin (tissue biopsies) using Illumina Human HT-12 V3 BeadChips. These tissues were sampled from a cohort of 856 female twins from the TwinsUK adult registry, including 154 monozygotic (MZ) twin pairs, 232 dizygotic (DZ) twin pairs and 84 singletons (MOAYYERI et al. 2013b). After quality control, expression data for 825 (adipose and LCL) and 705 (skin) individuals were retained (GRUNDBERG et al. 2012a). For each tissue, we downloaded the processed MuTHER expression data files deposited at ArrayExpression (http://www.ebi.ac.uk/arrayexpress/) using accession E-TABM-1140. The data were the quantile-normalized log$_2$-transformed expression signals. Quantile normalization was performed first across the replicates of a single individual and then across all individuals as described in (GRUNDBERG et al. 2012a). Along with the expression data, we also obtained the genotype data of this cohort (GRUNDBERG et al. 2012a). In our analysis, all available twin pairs with complete expression and genotype information were included, corresponding to 134 MZ and 195 DZ pairs with LCL profiles, 139 MZ and 188 DZ pairs with adipose profiles, and 105 MZ and 148 DZ pairs with skin profiles. Members of the TwinsUK cohort have health and lifestyle characteristics that are comparable to those of population singletons (ANDREW et al. 2001). Because of this, we were able to use this cohort as a representative general population to investigate both genetic and nongenetic factors behind expression variability in this study.

Expression variability in the twin cohort
Here we present basic, descriptive statistics for expression data (independent of genotype information), with particular attention to disparities in gene expression among individuals. We chose to focus on the LCL data for this analysis, due to the availability of additional expression-related statistics (such as single-cell expression data and mRNA decay data).
We used the quantile-normalized log₂-transformed expression data in all analyses throughout the paper unless otherwise indicated. From this data, we first determined that expression values for most probes \((n = 35,078)\) approximately fit the normal distribution: 97% of probes were with a skewness between -0.80 and 0.80 and a kurtosis of \(\sim 3.0\) (Supplementary Figure S1A); less than 7% of probes were rejected by Shapiro-Wilk test of normality with Bonferroni adjustment to the level of \(\alpha = 0.01\). These justified the use of the Gaussian error distribution and link function in our dglm model (Materials and Methods). Retrospectively, we showed that the profile distributions for evQTL probes are approximately normal before and after Box-Cox transformation (Supplementary Figure S1B).

To measure the level of dispersion of gene expression values, we computed the coefficient of variation (CV) for each probe. The CVs ranged from 0.0024 [for ILMN_1765043 (RPL38)] to 0.2115 [for ILMN_1715169 (HLA-DRB1)], with a median of 0.0154. The distributions of CVs measured in sub-cohorts are indistinguishable from one another such as when comparing one group of MZ twin sets with the other (i.e., MZ 1 vs. MZ 2) or comparing a group of MZ twin sets with a group of DZ twin sets (e.g., MZ 1 vs. DZ 1) (Figure 1A). Probe data points are located along or close to the 1-1 diagonal line in the CV-CV scatter plot for majority of probes, regardless of the CV-CV comparison between MZ 1 and MZ 2 or between MZ 1 and DZ 1 (Figure 1B). These results indicate that the extent and overall distribution of expression variability measured between individuals across different MZ and DZ cohorts are highly similar when all genes are taken into account.

Next, we measured expression differences between each pair of twins. For each probe, we computed the relative mean difference (RMD) in expression between MZ twin pairs and between DZ twin pairs, separately. For a pair of MZ twin, for example, the RMD was computed using \[ RMD = \frac{1}{2} \left| \frac{y_{MZ1} - y_{MZ2}}{\bar{y}} \right| \], where \(\bar{y}\) is the arithmetic mean the levels of gene expression for that MZ twin pair (designated as \(y_{MZ1}\) and \(y_{MZ2}\)). For most probes, the median RMD of expression between DZ pairs is larger than it is between MZ pairs,
as indicated by the fact that most genes are located above the 1-1 diagonal line in the scatter plot (Figure 1C). That is to say, the normalized difference in gene expression between DZ pairs (DZ 1 and DZ 2) tends to be larger than that between MZ pairs (MZ 1 and MZ 2), suggesting that genetic factors influence expression variability for most of these genes.

To determine the influence of single-cell expression variability on population-level expression variability, we computed the CVs of expression for a selection of genes whose expression levels have been measured in single LCL cells (Livak et al. 2013). No correlation between the single-cell CVs and the between-individual CVs measured was detected for MZ 1 (Spearman’s correlation test, \( P = 0.21, n = 59 \); Figure 1D). This suggests a limited contribution of single-cell expression variability (or transcriptional noise at the single-cell level) to the variability between individuals (or transcriptional noise at the population level).

Finally, we hypothesized that variable gene expression may be due to different mRNA decay rates for different genes. To test this, we used the mRNA decay rate data from the study of Pai et al. (2012). The correlation between mean mRNA decay rate and CV of expression among genes is not specific as shown by the opposite signs of two correlation coefficients: Spearman’s \( \rho = -0.027 \) (\( P = 0.00498 \)) and Pearson’s \( r = 0.044 \) (\( P = 4e-6, n = 11,083 \); Figure 1E). Thus, gene expression variability showed no signs of correlation with the mRNA decay rates of genes.

**Genetic variants underlying expression variability**

To systematically assess the genetic influence on expression variability, we identified genome-wide expression variability QTLs (evQTLs) using the method we previously established (Hulse and Cai 2013). We focused on cis-acting evQTLs by limiting our search to those SNPs that flanked probes within 1.0 Mb on either side.

After filtering for quality control (Materials and Methods), a total of 35,078 probes were available for analysis. On average, each probe corresponded to 1,212 SNPs in the 2-Mb cis region (i.e., 6 SNPs per 10 kb). For each SNP-probe pair, we conducted a three-step test to determine the evQTL relationship as described previously (Hulse and Cai 2013).
Briefly, we first tested for the homogeneity of variances in gene expression among different genotype groups using Fligner-Killeen (F-K) test (Fligner and Killeen 1976). Only those SNPs with a \( P < 0.01 \) [following (Fraser and Schadt 2010)] were carried on to the next step of analysis. We then applied the dglm method (Ronnegard and Valdar 2011) to each SNP-probe pair, ultimately computing \( P_{\text{dispersion}} \) for a total of 1,251,611 SNP-probe pairs. To account for multiple tests performed between these probe and SNP pairs, we used the threshold of \( P_{\text{dispersion}} < 1 \times 10^{-8} \), which is roughly equivalent to Bonferroni-adjusted \( P < 0.01 \), to assess the genome-wide significance. Finally, we conducted permutation tests for each significant SNP-probe pair to control for the influence of outlier data points on the dglm results (Materials and Methods). The detection of evQTLs was performed independently for each of the two groups of twin sets. The assignment of individual twins to each group was random and did not influence the overall results. Each evQTL detected with one group of twin sets and was then validated with the other group of twin sets to confirm its authenticity. For all three tissues, concordance was prevalent (Figure 2A) and the cases of discordance were mostly due to outliers present in one group of twins but not in the other group. The direction of effect (association with increased or decreased gene expression variability) was the same between the two groups of the twin sets for all evaluated SNPs.

A total of 99, 75 and 59 genes were identified and confirmed to have at least one validated cis-evQTL SNP (abbreviated as “evSNP”) in LCLs, fat and skin, respectively (Supplementary Table S1). Eight genes (corresponded to 4.2% of all unique evQTL genes) were shared in all three tissues (Figure 2B). One of these shared evQTL genes, \( \text{SEMA4G} \), is given as an example to illustrate the consistent influence of genotypes on the variance of gene expression across the three tissues (Figure 2C). All evQTLs shared across tissues showed the same directional effect, defined as either increasing or decreasing the variance of gene expression. That is to say, the directionality of some evQTL effects is not tissue- or cell-type specific. However, understanding regulatory variability in many different tissues might yield insights into the basic biological processes that influence tissue differentiation.
Given that many evQTL genes have more than one *cis*-evSNP, we examined the structure of haplotypes of these multiple *cis*-evSNPs. We found that *cis*-evSNPs of the same gene are likely to be located within same linkage disequilibrium (LD) block and that typically these blocks contained only few prominent haplotypes (see Supplementary Figure S2 for an example involving gene ALG11). This suggests that multiple evSNPs are likely to be linked with the same causal variant. We furthermore retrieved the ancestral allele information for SNPs from the 1,000 Genomes Project (THE-1000-GENOMES-PROJECT-CONSORTIUM et al. 2012). The prediction of ancestral alleles was based on the phylogenetic trees constructed with sequences of human, chimpanzee, orangutan, and rhesus macaque. We found that ancestral alleles of evSNPs are more likely to be associated with smaller expression variability than derived alleles (Fisher’s exact test: \( P = 0.0036, 0.022 \) and 0.036 for LCLs, skin and fat, respectively).

**Dissect genetic and nongenetic effects of evQTLs**

Twin data facilitated the dissection of the contributions of genetic and nongenetic factors. Variability measured between pairs of DZ twins is expected to be larger than that between pairs of MZ twins, as the phenotypic difference between DZ pairs may result from both genetic and environmental (nongenetic) effects while differences between genetically-identical MZ pairs are attributable to the environment, assuming that the environments influencing MZ and DZ twin individuals are essentially identical. Figure 3 depicts the difference in expression level of evQTL gene AXIN2 in three genotypes (GG, AG, and AA) defined by rs740026. Figure 3A and B illustrate genotypes at rs740026 by linking the two data points for each twin pair by a straight line: Figure 3A shows genotype similarities between MZ twins, while in Figure 3B, similarities between DZ twin pairs are shown. Note that linkers between DZ twin pairs with different genotypes at the SNP site (i.e., DZ 1 ≠ DZ 2) are not plotted. The expression difference between a pair of twins can be visually quantified by the slope of the straight line: a steeper line reflects a more dissimilar expression level between the twins. In the case of AXIN2, it is apparent that expression differences between DZ pairs tend to be larger than between MZ pairs. This is especially true for the AA genotype group, which shows a larger variance in expression between individuals.
For each evSNP and its associated genes in LCLs, we computed the RMD in gene expression between all pairs of MZ or DZ twins, as long as the genotypes of two individuals of the pair of twins were both identical to each other and homozygous at the SNP site. By definition, each evSNP allele is associated with either larger (L) or smaller (S) variance in gene expression. Thus, the RMD values (for evSNPs and associated genes) were separated according to whether homozygous genotypes defined by evSNPs were associated with larger (L) or smaller (S) variances in gene expression. The cumulative distribution functions (CDFs) of these RMD values were plotted (Figure 3C). The curves were based on the RMD values calculated between all possible twin pairs for all evSNPs and genes, and classified into four groups: MZ-S, MZ-L, DZ-S, and DZ-L. The MZ-S and DZ-S groups included pairs whose genotypes showed a small (S) amount of variance, while the MZ-L and DZ-L groups included pairs whose genotypes were associated with large (L) variances. In the end, the four groups—MZ-S, MZ-L, DZ-S, and DZ-L—contained 3,629, 2,548, 3,825, and 2,520 RMD values, respectively. Detailed statistics for the distributions of RMD values in each of these four groups are provided (Supplementary Table S2). We found that CDF curves for the large-variance groups (MZ-L and DZ-L) were shifted toward the right compared to those for small-variance groups (MZ-S and DZ-S) [Kolmogorov-Smirnov (K-S) test, all \( P < 10^{-5} \)]. This indicated that the distribution of RMD between twin pairs (either MZ or DZ) in the large-variance groups was significantly different from that of the small-variance groups, with a larger RMD median for the large-variance groups. This difference (in RMD distribution between L and S groups) remained even when we randomly assigned the identities of MZ and DZ pairs (see insert of Figure 3C). Together, these results suggested that the increased discrepancy in gene expression between twin pairs (shown as a larger median RMD) contributed to the elevated variability in expression, which is true for both MZ and DZ twins. Because MZ twins are genetically identical, the increased RMD between MZ pairs was likely due to an increased sensitivity of gene expression to environmental factors.

More importantly, we found a significant discrepancy in distribution of RMD between MZ and DZ: DZ groups tended to have larger RMD values than MZ groups. This trend applied to both L and S groups, but was more obvious in the L group (all K-S tests, \( P < \)
0.01) (Figure 3C). These results suggested that the different genetic backgrounds resulted in a larger difference in gene expression between DZ twin pairs, which is more pronounced than that observed between MZ twin pairs.

For comparison, we randomly selected the same number of genes and cis-SNPs and conducted the same analysis of RMD distribution. There was no difference between CDFs of RMD in these non-evQTL genes in regard to either MZ or DZ twins, larger or smaller variance groups, as well as before or after shuffling of the twin identities. CDFs of all groups were more similar to each other (K-S test, all $P > 0.025$, except between MZ-S and DZ-L, $P = 2.9e-4$, Supplementary Table S2, Figure 3D). That is to say, the influence of genetic and/or environmental effects on variable expression was not detected at the genomic level for all genes, but was limited to evQTL regions.

Finally, we repeated the CDF analyses using the RMD values computed from the Box-Cox normalized log$_2$-transformed expression data, as well as using the absolute difference (instead of RMD) in gene expression. In both cases, our findings were highly similar to those obtained above (Supplementary Figure S3), which supports the robustness of the results presented above.

**Possible evQTLs replicated by using RNA-seq data and SNPs of the 1,000 Genomes Project**

We obtained genotype data for fully-sequenced samples of European ancestry (CEU) from the phase 1 release of the 1,000 Genomes Project (The-1000-Genomes-Project-Consortium et al. 2012), along with short reads from RNA-seq experiments in LCLs for these same individuals ($n = 43$) (Montgomery et al. 2010). After mapping the short reads, we estimated the expression level in FPKM [fragments per kilobase of exon per million fragments (Trapnell et al. 2013)] for all genes. For the same evQTL gene-SNP pairs detected in LCLs (i.e., those included in Supplementary Table S1), we plotted the relationships between genotype and FPKM value for each. Many evQTL relationships could be recognized by visual inspection (examples are presented in Supplementary Figure S4), though none were statistically significant due to the small sample size. It is noteworthy that the algorithm for computing FPKM models the dispersion in a transcript’s fragment count with a negative binomial distribution.
(Trapnell et al. 2013), which may introduce a relationship between the mean and variance of the count. The relationship should be taken into account in FPKM-based evQTL analysis.

**Partially-linked SNPs contribute to variable gene expression**

Recent theoretical work showed that the within-genotype variance of a quantitative trait varies when a non-additive genetic interaction or epistasis is present (Pare et al. 2010; Struchalin et al. 2010). Alternatively, the variance of a quantitative trait may be from the result of the interaction between genetic variants additively associated with the mean of the quantitative trait. To test this, we employed a two-step procedure to identify SNPs partially associated with (or interacting) with evSNPs through an incomplete haplotype structure (Materials and Methods). In an ideal scenario (Figure 4A), the genotype heterozygosity of the partially-linked SNPs is large among individuals whose evSNP genotype is associated with larger expression variability (L-group), while the genotype heterozygosity is small or equals zero among individual whose evSNP genotype is associated with smaller expression variability (S-group). If the interacting SNP is associated with the mean level of gene expression, then the association between the evSNP genotype and greater expression variability is likely due to the partial linkage between the evSNP and the interacting SNP.

Given these considerations, we performed a genome-wide search to identify a set of candidate interacting SNPs for each evQTL SNP, and then used simple linear regression analysis to evaluate whether the potential interacting SNPs are significantly associated with gene expression among L-group individuals (Materials and Methods). For the 99 evQTL in LCLs, we identified 56 with at least one interacting SNP (Supplementary Table S3). Among these interacting SNPs, 54 are located within the cis-region of the evSNPs with which they interact. Figure 4B presents one such relationship between evSNP rs742090 and interacting SNP rs3799378, both at BTN3A2. Individuals with a CC genotype for evSNP rs742090 were further sorted by rs3799378 genotypes. Clearly, the expression level of BTN3A2 in individuals with the rs742090-CC genotype is significantly influenced by rs3799378 genotypes. The increased variability in gene expression showed in individuals with rs742090-CC genotype is caused by the
heterogeneity of rs3799378 genotypes. These results suggest that local haplotype structure between SNPs contributed to the creation of evQTLs.

**Linking evQTLs with complex disease phenotypes**

Several studies have utilized eQTL data to interpret the discoveries of association studies of complex traits (EMILSSON et al. 2008; NICA et al. 2010; NICOLAE et al. 2010). Along this same vein, we identified evQTLs associated with complex traits from the catalog of published genome-wide association (GWA) studies (http://www.genome.gov/gwastudies/). From the results of these GWA studies, we identified 61 reported genes are evQTL genes (Supplementary Table S4). In four cases, the exact same SNP was found to be both an evSNP and a marker SNP associated with risk or susceptibility to the complex trait (Table 1). Intriguingly, the “T” allele of rs8070463, associated with smaller expression variability of TBKP1, is a reported risk allele in multiple sclerosis (PATSOPoulos et al. 2011) while the “C” allele for this same SNP, associated with larger expression variability, is linked with risk for ankylosing spondylitis (EVANS et al. 2011).

**Discussion**

There is empirical evidence across several species that the variance among phenotypes is genotype-dependent (ANSEL et al. 2008; WOLC et al. 2009; HILL and MULDER 2010; JIMENEZ-GOMEZ et al. 2011). Understanding genetic control of phenotypic variability has become increasingly important in evolutionary biology, human medicine, the agricultural industry and other branches of biological science (GIBSON 2009; YANG et al. 2012). Despite the importance, few research programs focus on genetic variants associated with trait variance, while studies of trait averages abound. Recently, a powerful statistical framework based on the dglm model has been developed for studying phenotypic variability of complex traits (RONNEGARD and VALDAR 2011). Given that gene expression is a complex trait with highly variable and heritable patterns (STRANGER et al. 2005; WILLIAMS et al. 2007; MONTGOMERY and DERMITZAKIS 2011), we have previously adapted the dglm method to investigate genetic variants controlling expression variability (HULSE and CAI 2013).
In this study, we further investigated the relative contributions of genetic and nongenetic (environmental) factors to expression variability and the roles of these factors in the formation of evQTLs. We started by exploring basic gene expression statistics measured in the TwinsUK cohort. For all genes, expression level dispersions were highly similar in and between both MZ and DZ twins. No correlations with expression variability were detected when compared between individuals, between single cells, or relative to the average mRNA decay rate, highlighting the marked discrepancies in variability measured at population and molecular levels. Further results showed that the discordance in expression between each pair of DZ twins was more pronounced than that between MZ twins, implying that the increased amount of genetic variation between DZ twins influences expression variability. Next, we systematically identified cis-acting evQTLs in three tissues of the TwinsUK cohort. Twin data greatly facilitated the validation of detected evQTLs and revealed overall robust signals that would otherwise not be appreciable in studies of non-twin design. Focusing on the detected evQTLs, we showed that the discordance in expression between DZ pairs was larger than that between MZ pairs, and further showed that the discordance in expression between MZ pairs whose genotypes were associated with large expression variability was significantly larger than that between MZ pairs whose genotypes were associated with small expression variability. It is intriguing to find that the phenotypic discordance remained even in the absence of genetic variation between MZ twins. This might be explained by incomplete penetrance of mutations, which is frequent in isogenic model organisms in homogeneous environments (HORVITZ and SULSTON 1980; GARTNER 1990). This might also be epigenetic: for example, DNA methylation, which can be influenced by environmental factors such as diet and lifestyle, is known to affect gene expression (BADANO and KATSANIS 2002; BARANZINI et al. 2010). Lastly, much to our surprise, we found that more than half of evQTLs could be explained by a conceptually simple scenario in which the evSNP was occasionally associated with a nearby SNP that influenced gene expression both additively and independently. We suspect there should be many different ways of non-epistatic interaction between two or more genetic variants, such as the mode of partial linkage we have described here, giving rise to genotype-dependent expression variances. That is to say, the majority of phenotypic
variability across individuals might be explained without invoking epistasis (Hill et al. 2008; Powell et al. 2013).

In light of our new findings, several related considerations are discussed below.

**Methodological considerations for studying phenotypic variability**

The procedure we used for identifying evQTLs [Materials and Methods, and (Hulse and Cai 2013)] consisted of three steps. First, the F-K test was applied to test for the heterogeneity of variances of gene expression between different genotypes and identify corresponding SNPs. Next, the dglm method was applied to the selected SNPs. The significant results of dglm test were then subjected to permutation tests to reduce the influence of outliers in the data. This procedure is less likely to be susceptible to issues related to multiple testing and outliers in input data, though a formal assessment of its statistical power remains to be done.

Statistical methods, including Levene’s test (Pare et al. 2010; Struchalin et al. 2010), SVLM (squared residual value linear modeling) (Struchalin et al. 2012), and the methods by Yang et al. (2011), have been applied in studying phenotypic variability [see review (Geiler-Samerotte et al. 2013)]. As a full parametric approach, the dglm method (Ronnegard and Valdar 2011) has several advantages. For example, it accounts for the uncertainty of fitted parameters for both the mean and the variance aspects of the model, and also allows fitting of covariates (Ronnegard and Valdar 2012); it is also highly flexible, allowing for any response distribution from the exponential family (Smyth 2002) (such as binomial, Poisson, or gamma) to be modeled [see Section 2 in the Supporting Theory of (Ronnegard and Valdar 2011) for a sample of modelling the gamma-distributed traits].

Given the flexibility of the dglm method, we acknowledge that the results of our evQTL analysis are likely to be dependent on how the dglm analysis was set up. For this study, we adapted the Gaussian error distribution and link function because no significant departure from normality was found in the expression data. The effect of different methods of normalization on statistical interpretation of gene expression remains subject to careful scrutiny (Bolker et al. 2009; Qin et al. 2012; Geiler-Samerotte et al. 2013).
For example, normalizations may perturb the covariance structure of input data or change the scale of the resulting data. Thus, the impacts of different methods of data transformation and normalization should be carefully considered in future studies involving evQTL analysis. Finally, we acknowledge that the dglm analysis described in this paper may be influenced by the scale effect (e.g. mean-variance relationship). It is not uncommon for trait variance to change with trait mean, often causing trait skewness. If this occurs, any SNP associated with a large increase in mean expression would also be associated with an increase in variability (Ronnegard and Valdar 2011). Analyses studying a specific phenotype and/or with a more narrowly-targeted focus than that of the broad-based study described in this paper should employ a more conservative approach in which QTL associated strictly with variance (i.e. those affecting only variability and not the mean) are identified, using the procedure proposed by Ronnegard and Valdar (2011).

Additive vs. epistatic effect of genotypes on phenotypic variation in a population

Quantitative geneticists partition the genetic effect on phenotypic variation between individuals into additive, dominance and epistatic components. The additive component describes the variance associated with the independent contributions of alleles, while dominance describes the variance contributed by interactions between alleles at the same locus, and epistasis refers to the contribution of interactions between alleles at different loci. For most complex traits, quantitative genetic theory (Hill et al. 2008; Crow 2010) suggests that epistasis is unlikely to contribute substantially to the between-individual variation. That is to say, most of the variation in a population will be due to the additive effects of specific allelic combinations. Yet this assertion is not without controversy. The results of empirical linkage mapping and association studies suggest that epistasis seems to explain considerable variation in complex trait characteristics within natural populations (Carlborg and Haley 2004; Zuk et al. 2012).

Our results showed that >50% of evQTLs can be explained by a partial linkage (i.e., partial association between haplotypes) of the evQTL SNP and another SNP nearby. Our interacting SNP analysis only considered a simplistic scenario of the association. There
are many other possible ways of partial associations in which SNPs interact. For example, consider the genotyped SNP “A/a” and the causative expression QTL “Q/q”, with only three haplotypes segregating in the population: AQ, aQ and aq (as would occur if the novel “q” allele arose on the “a” haplotype). Then the “a” SNP allele would be associated with a changed trait mean and a higher trait variance as the expression QTL segregates within that genotype. If we could take all possible partial associations into account, we would anticipate that even more evQTLs could be explained by the effect of partial linkage, rather than epistasis. We therefore conclude that much variance in a quantitative trait may be explained by partial linkage between local genetic variants, each additively associated with the trait. Our view is supported by the results of recent studies. Powell et al. (2013) conducted a gene expression study using blood samples from 862 individuals from nuclear families containing MZ or DZ twin pairs, using both pedigree and genotype information. They found that the genetic architecture of gene expression is predominantly additive, with a minority of transcripts displaying non-additive effects. Hill et al. (2008) evaluated the evidence from empirical studies of genetic variance components and found that additive variance typically accounts for over half and often close to 100% of the total genetic variance.

**Detecting evQTL as a shortcut for detecting epistasis?**

Detection of the variance of a quantitative trait in genetic association studies is thought to increase knowledge about the interaction between genetic variants. More specifically, detecting variability QTL (e.g., evQTL) is considered to be a shortcut for detecting genetic interactions (Ronnegard and Valdar 2011; Ronnegard and Valdar 2012). So far, many methods for detecting genetic interactions are based on testing for different variances of phenotype between genotypes, with the underlying assumption that the variance of a quantitative trait is likely to differ under the influence of epistasis (Pare et al. 2010; Ronnegard and Valdar 2011). However, our new discovery that evQTLs are formed due to the partial haplotype association between SNPs refutes this assumption. As stated above, more than half (and probably many more) of the evQTLs identified could be explained by partial linkage between SNPs with additive effects. Both additive and epistatic effects can result in increased phenotypic variation (as schematically illustrated in **Supplementary Figure S5**). Merely detecting the variance of a
quantitative trait cannot in itself distinguish between the additive and epistatic effects; thus, no specific conclusions can be made. The relationship between partially associated SNPs, each additively associated with phenotypic variation, needs to be integrated more carefully in the study of phenotypic variability. Thus, the variance of a quantitative trait should not serve as a hallmark of genetic interaction or epistasis.

**Phenotypic variability and implications in complex traits and diseases**

High-throughput sequencing and genotyping technologies have spurred an increasing number of studies detecting genotype-phenotype relationships and mapping in complex, polygenic traits and human diseases (HINDORFF et al. 2009). The remarkable success of genome-wide association studies (GWAS) is accompanied by the issue of the “missing heritability” (MANOLIO et al. 2009), namely the fact that the trait-associated SNPs identified through GWAS often account for only a small proportion of the observed correlations in phenotype between relatives. The reason behind this issue has been thought to be that additional genetic factors remain to be found, and that the presence of epistasis is a particular cause for concern (CARLBOG and HALEY 2004; MOORE and WILLIAMS 2009; UEKI and CORDELL 2012). In reality, if the effect of one locus is altered or masked by effects at another locus, power to detect the first locus is likely to be reduced and elucidation of the joint effects at the two loci will be hindered by their interaction. Consequently, a large amount of research has been devoted to the detection and investigation of epistatic interactions; a number of methods for detecting the interaction between SNPs have been proposed (PARE et al. 2010; STRUCHALIN et al. 2010; SHANG et al. 2011; DAVE et al. 2012; STRUCHALIN et al. 2012), yet there has been much confusion in the literature over definitions and interpretations of epistasis (CORDELL 2002).

This study, along with other studies (HILL et al. 2008; POWELL et al. 2013), have clearly pointed out that a detailed investigation of local haplotype structure between SNPs at the same locus is necessary to reveal their combined influences on phenotypes of complex traits. For example, we have identified a list of evSNPs that are also associated with human complex trait (see Table 1). Further investigations on partial linkage between SNPs that may influence these traits should be performed. The same should
also be done for FTO whose genotype is associated with phenotypic variability of body mass index (Yang et al. 2012).

Finally, we point out that an interaction detected via statistical models is different from the biological interaction (Phillips 1998; Correll 2002; Wang et al. 2010). The lack of direct correspondence between statistical and biological interactions makes it difficult to make strong inferences concerning biological mechanisms based on interaction terms from a statistical model (Ueki and Correll 2012). Therefore, detection of statistical interaction merely provides a good starting point for a more focused investigation of the joint involvement of the relevant factors, which can perhaps be better addressed through other types of experimental data. Our findings suggest that there is a lot that can be done at the statistical level to prioritize those loci that are most likely to produce significant experimental results.

Conclusions
In conclusion, we used evQTLs as a statistical model system for studying phenotypic variability and dissected the genetic and nongenetic effects by using the twin data. Our findings concerning evQTLs offer new insights into relative contribution of genetic and environmental factors in the formation of evQTLs. Dissecting the genetic components underlying phenotypic variability into additive and epistatic effects allowed the dominant role of additive effect to be revealed.

Acknowledgments
We thank the TwinUK study for producing the datasets, Ken Livak for sharing the single-cell expression data, and the Whole Systems Genomics Initiative (WSGI) of Texas A&M University for providing computational resources. The TwinUK study was funded by the Wellcome Trust; European Community’s Seventh Framework Programme (FP7/2007-2013). The study also received support from the National Institute for Health Research (NIHR) Clinical Research Facility at Guy’s & St Thomas’ NHS Foundation Trust and NIHR Biomedical Research Centre based at Guy’s and St Thomas’ NHS Foundation Trust and King’s College London. SNP Genotyping was performed by The Wellcome Trust Sanger Institute and National Eye Institute via NIH/CIDR.
References


Nicolae, D. L., E. Gamazon, W. Zhang, S. Duan, M. E. Dolan et al., 2010 Trait-associated SNPs are more likely to be eQTLs: annotation to enhance discovery from GWAS. PLoS Genet 6: e1000888.
Pare, G., N. R. Cook, P. M. Ridker and D. I. Chasman, 2010 On the use of variance per genotype as a tool to identify quantitative trait interaction effects: a report from the Women's Genome Health Study. PLoS Genet 6: e1000981.


**Figures**

**Figure 1.** Distributions of expression variability in LCLs. (A) Distribution of CVs of gene expression (probe \( n = 35,078 \)) measured in MZ and DZ twins. MZ 1 is the set of first pairs of all MZ twins and MZ 2 is the set of second pairs of all MZ twins. Similarly, DZ 1 is the set of first pairs of all DZ twins and DZ 2 is the set of second pairs of all DZ twins. (B) Scatter plot of CVs of gene expression (probe \( n = 35,078 \)) in MZ 1 against those in MZ 2 (blue) or DZ 1 (red) cohorts. (C) Scatter plot of median RMD between pairs of MZ twins against median RMD between pairs of DZ twins. Each blue dot indicates a single expression probe (or a gene) and the position of the blue dot indicates the median value of RMD of expression between all MZ pairs (MZ 1 – MZ 2) on the x-axis and that between all DZ pairs (DZ 1 – DZ 2) on the y-axis. The red line is based on quadratic regression to show a more pronounced difference between MZ and DZ with greater RMD. (D) Scatter plot of CVs of gene expression \( (n = 59) \) in single cells against CVs of gene expression in MZ 1. (E) Scatter plot of mean mRNA decay rate against CVs of gene expression in MZ 1 cohort. The red line is based on the linear regression.
Figure 2. Numbers of evQTLs in LCL, skin, and fat. (A) Venn diagrams of evQTL genes detected in two groups of twin sets. Each group of the twin sets is composed of one set of unrelated twin individuals. Overlapping areas of the Venn diagrams contain numbers of validated evQTL genes identified with both sets of twins. (B) Numbers correspond to evQTL genes within a subset of tissues. (C) One example of evQTLs shared by all three tissues: evQTL at SEMA4G.
Figure 3. Dissection of genetic and nongenetic effects of evQTLs using twins data. (A) The evQTL between AXIN2 and rs740026. The expression data points from pairs of MZ twins are linked. (B) Same as A except that DZ twins are linked. (C) CDFs of RMD between twins classified into four groups, namely MZ-S, DZ-S, MZ-L, and DZ-L (see main text for definitions). The double-arrow highlights the highly significant discrepancy in RMD distribution between MZ-L and DZ-L (K-S test, \( P < 0.01 \)). Insert shows the same CDFs of RMD re-computed after randomly shuffling identities of corresponding MZ and DZ pairs. (D) Same as C except that data are randomly sampled from non-evQTL genes.
**Figure 4.** Schematic and example of an interacting SNP that helps the creation of an evQTLs. (A) L indicates the group of individuals with evSNP genotype (C/C) associated with large variance in gene expression, while S indicates that with evSNP genotype (A/A) associated with small variance. The interacting SNP show large genotype heterogeneity in L group and small or no genotype heterogeneity in S group. (B) Real example of evSNP rs742090 and interacting SNP rs3799378 at BTN3A2. Individuals with rs742090-CC genotype is further broken down by rs3799378 into three sub-genotype groups, which are associated with different means of gene expression levels (shadowed panel).
Tables

Table 1. SNPs associated with gene expression variability and human complex trait. $^L$ and $^S$ indicate that individuals carrying homozygotic genotype of the risk allele have large and small variance in gene expression, respectively.

<table>
<thead>
<tr>
<th>Gene (evSNP)</th>
<th>Tissue</th>
<th>GWAS complex trait</th>
<th>Risk allele</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>PAX8 (rs11123170)</td>
<td>LCL, Fat, Skin</td>
<td>Renal function related traits (BUN)</td>
<td>rs11123170-$^L$</td>
<td>(OKADA et al. 2012)</td>
</tr>
<tr>
<td>WDR41 (rs163030)</td>
<td>LCL, Fat, Skin</td>
<td>Caudate nucleus volume</td>
<td>rs163030-$^L$</td>
<td>(STEIN et al. 2011)</td>
</tr>
<tr>
<td>HCG22 (rs2517532)</td>
<td>LCL</td>
<td>Hypothyroidism</td>
<td>rs2517532-$^S$</td>
<td>(ERIKSSON et al. 2012)</td>
</tr>
<tr>
<td>TBKBP1 (rs8070463)</td>
<td>LCL</td>
<td>Multiple sclerosis Ankylosing spondylitis</td>
<td>rs8070463-$^T^S$</td>
<td>(PATSOPoulos et al. 2011)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>rs8070463-$^C^L$</td>
<td>(EVANS et al. 2011)</td>
</tr>
</tbody>
</table>
Supplementary Figures Legends

Supplementary Figure S1. Normality of expression data measured in LCLs. (A) Distributions of skewness and kurtosis. Red dashed lines indicate -0.8 and +0.8 skewness; red solid line indicates kurtosis = 3. (B) Profile distributions of expression data for selected probes (i.e., probes involved in evQTLs). (Left) Quantile-normalized expression data; (Right) Box-Cox normalized expression data.

Supplementary Figure S2. Linkage disequilibrium patterns of the genomic region surrounding evQTL at ALG11. The entire region of the analysis included 38 SNPs over a ~400 kb span. The cis-evSNPs are indicated with red boxes. The haplotypes in the LD block accommodating evSNPs are displayed in insert, with corresponding haplotype frequencies. Of note, alleles of evSNPs resulting in larger variance of gene expression are allocated in one haplotype highlighted with blue box.

Supplementary Figure S3. Comparison between results of the CDF analysis for the expression difference between twin pairs in evQTL genes. (A) Results obtained using RMD of Box-Cox normalized log₂-transformed data between twin pairs. (B) Results obtained using absolute difference of log₂-transformed data between twin pairs.

Supplementary Figure S4. Comparison between evQTLs depicted using TwinsUK samples (left) and those of the same gene-SNP pairs depicted using RNA-seq data (MONTGOMERY et al. 2010) and the 1,000 Genomes Project genotype data (THE-1000-GENOMES-PROJECT-CONSORTIUM et al. 2012) (right). $P$-value of Levene’s test for the equality of variances for expression levels calculated for two homozygous genotype groups is shown in the subtitle of the right panel in each figure.

Supplementary Figure S5. Schematic shows that both additive (left) and epistatic (right) effects create similar evQTL signals. “A” and “a” are two alleles of evSNP, while “B” and “b” are alleles of interacting SNP.
**Supplementary Tables Legends**

**Supplementary Table S1.** List of evQTLs detected in LCL, skin, and fat tissues.

**Supplementary Table S2.** $P$-values of Kolmogorov-Smirnov (K-S) test of pairwise comparison between groups in the RMD distribution and basic descriptive statistics for the RMD values (A). Upper diagonal $P$-values are from randomly selected non-evQTL genes. Results obtained from the Box-Cox normalized and not normalized data are given. The per-probe statistics of RMD values for larger variance groups (MZ-L and DZ-L). (B) Probes with no SNP in either groups are not shown.

**Supplementary Table S3.** LCL evQTLs and interacting SNPs. $P$-value indicates the linear regression significance between interacting SNP genotype and gene expression of evQTL gene (see Methods). Two genes close to interacting SNPs at trans-genomic location are given.

**Supplementary Table S4.** The evQTL genes associated with complex traits and reported in the GWAS catalog (accessed in May 2013).