Heritability and genomics of gene expression in peripheral blood

Fred A Wright1–3,13, Patrick F Sullivan4,13, Andrew I Brooks5, Fei Zou6, Wei Sun6, Kai Xia6, Vered Madar6, Rick Jansen7, Wonil Chung6, Yi-Hui Zhou1,2, Abdel Abdellaoui8, Sandra Batista9, Casey Butler9, Guanhua Chen6, Ting-Huei Chen6, David D’Ambrosio10, Paul Gallins4, Min Jin Ha6, Jouke Jan Hottenga8, Shunping Huang9, Mathijs Kattenberg8, Jaspreet Kochar10, Christel M Middeldorp8, Ani Qu10, Andrey Shabalin11, Jay Tischfield5, Raleigh, North Carolina, USA. 3Department of Biological Sciences, North Carolina State University, Raleigh, North Carolina, USA. 4Department of Genetics, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina, USA. 5Department of Psychiatry, VU Medical Center, Amsterdam, The Netherlands. 6Department of Biostatistics, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina, USA. 7Department of Genetics, Rutgers University, New Brunswick, New Jersey, USA. 8Department of Biostatistics, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina, USA. 9Department of Mathematics, VU University, Amsterdam, The Netherlands. 10Department of Biostatistics, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina, USA. 11Environmental and Occupational Health Sciences Institute, Rutgers University, New Brunswick, New Jersey, USA. 12Department of Computer Science, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina, USA. 13These authors contributed equally to this work. Correspondence should be addressed to F.A.W. (fred_wright@ncsu.edu) or P.F.S. (pfsullivan@email.unc.edu).

Received 28 October 2012; accepted 14 March 2014; published online 13 April 2014; doi:10.1038/ng.2951

We assessed gene expression profiles in 2,752 twins, using a classic twin design to quantify expression heritability and quantitative trait loci (eQTLs) in peripheral blood. The most highly heritable genes (~777) were grouped into distinct expression clusters, enriched in gene-poor regions, associated with specific gene function or ontology classes, and strongly associated with disease designation. The design enabled a comparison of twin-based heritability to estimates based on dizygotic identity-by-descent sharing and distant genetic relatedness. Consideration of sampling variation suggests that previous heritability estimates have been upwardly biased. Genotyping of 2,494 twins enabled powerful identification of eQTLs, which we further examined in a replication set of 1,895 unrelated subjects. A large number of non-redundant local eQTLs (6,756) met replication criteria, whereas a relatively small number of distant eQTLs (165) met quality control and replication standards. Our results provide a new resource toward understanding the genetic control of transcription.

Determining the biological relevance of findings from genome-wide association studies (GWAS) has emerged as a major challenge for complex trait analysis, as over 90% of significant associations are noncoding. Several lines of evidence suggest that genetic variation implicated in GWAS alters transcription1–5. eQTLs5–7 overlap markedly with GWAS-identified SNPs, both collectively6–8 and for specific traits (for example, height, adiposity, cardiovascular risk factors, chemotherapy-induced cytotoxicity, autism, schizophrenia and Crohn’s disease)9–16. An estimated 55% of eQTL SNPs lie in DNase I hypersensitivity sites, and 77% of significant GWAS SNPs are in or correlated with these sites2,17,18. Although understanding of eQTLs has progressed rapidly, important questions remain. Most eQTL catalogs are incomplete, and few studies have had sample sizes of > 1,000 (refs. 15,19,20), although > 3,000 may be necessary for more complete eQTL identification21. Many eQTLs do not replicate, even using the same HapMap lymphoblastoid cell lines (LCLs) under standardized procedures19. Replication of distant (trans) eQTLs has been particularly elusive22. Potential sources of variation include tissue type8,10,21, ancestry7, winner’s curse and batch effects5,7,24–26, and cell heterogeneity27,28.

To achieve large sample sizes in humans, tissues must be accessible. An attractive choice is peripheral venous blood, although most but not all20–23 human blood-derived eQTL studies have used LCLs. However, gene expression differs between LCLs and peripheral blood20, and LCLs can be influenced by factors such as Epstein-Barr virus (EBV) copy number and growth rates31. The Multiple Tissue Human Expression Resource (MuTHER) LCL study of expression in female twins found a large impact of the common ‘environment’ shared by twins: 32% of transcripts showed common environmental effects of >30%, compared to 2% in adipose and 8% in skin8. The authors attributed this dramatic effect to correlated sample handling rather than to environmental exposures shared by twins, suggesting possible biases with LCLs.

Despite these challenges, quantifying human transcriptomic heritability is important. Although genes with genome-wide significant eQTLs are by definition ‘heritable’, additional polygenic variation may be widespread and fail to reach statistical significance by standard
RESULTS

Twin-based heritability in the peripheral blood transcriptome

We first report the heritability of steady-state transcription in peripheral blood for 43,628 probe sets from 18,392 genes from 2,752 individual twins in the Netherlands Twin Registry (NTR; Table 1). The U219 platform includes alternate 3′ sequences of well-annotated genes, and we refer to each of the probe sets as a ‘transcript’ (1–18 transcripts per gene, mean of 2.4). We performed careful annotation for the platform, which compares favorably to RNA sequencing (RNA-seq) (Supplementary Note)33. Subjects were from 1,444 twin pairs (both members of 1,308 pairs, 95.1% of subjects and 1 member from 136 pairs). The 1,308 complete pairs consisted of 690 monozygotic pairs (52.8%; 209 male and 481 female monozygotic pairs) and 618 dizygotic pairs (47.2%; 110 male, 256 female and 252 opposite-sex dizygotic pairs). Expression quality control included zygosity and sex confirmation, randomization for sex and zygosity balance, sample identity checks and removal of low-quality samples. Primary analyses were based on robust multi-array average (RMA) expression estimates, filtered to exclude probes containing SNPs or mapping non-uniquely, with each transcript transformed to an exact normal distribution for robust analysis.

The Supplementary Note lists ~140 covariates used, including blood cell counts and the genotypes of blood count–associated SNPs. We computed the proportion of variance explained ($r^2$) attributable to covariates and the effect of covariate control on heritability ($h^2$) and variance explained by common ($c^2$) and unique ($c^2$) environment, where these values were measured using a covariance (ACE) model that includes additive genetic, common or shared environment, and non-shared environment terms (Supplementary Fig. 1). Variance components were not constrained to be positive, so the model would be unbiased for $h^2$ estimation, and to indicate whether genetic non-additive effects (dominance) might be present (by estimating $c^2$ as negative). Covariate correction notably increased evidence for highly heritable transcripts, whereas no transcript was significant for $c^2$ values (Supplementary Fig. 1b–c), in contrast to the MuTHER study8.

Figure 1a shows a $P$-value Manhattan plot for twin-based $h^2$ estimation for 18,392 genes (selecting for each gene the transcript with the largest $h^2$ value), based on twin zygosity comparisons. The $h^2$ value had mean $±$ s.d. of 0.101 ± 0.142 (0.138 ± 0.153 for expressed genes), with maximum estimated $h^2 = 0.905$. We conservatively highlight 777 genes with significant heritability ($q$$<$$0.05; 4.2% of the genes on the microarray), applying $k$-means clustering and analysis of genomic location. The 777 genes yielded 9 expression clusters (Fig. 1b and Supplementary Table 1). Mean within-cluster expression correlation $r$ ranged from 0.46 to 0.006. Cluster identity was supported by significantly ($P < 0.05$) higher connectivity in protein–protein interaction databases34 and Gene Ontology (GO) pathways35 (Supplementary Table 1). Numerous clustered genes showed expression patterns similar to those observed in other tissues, including brain32, suggesting broader tissue relevance. Regional clustering indicated enrichment for immune function (Supplementary Table 2; for example, IgG Fc fragment receptors encoded at chr: 1: 161–162 Mb and the major histocompatibility complex (MHC) region at chr: 6: 31–33 Mb), whereas other regions showed fewer heritable genes (for example, the neuronal protocadherin gene cluster at chr: 5: 140–141 Mb and epidermal keratin gene clusters on chr: 17: 39–40 Mb and chr: 21: 31–32 Mb). Heritability was strongly associated with mean expression ($r = 0.356; P < 1 × 10^{-200}$, Fig. 1c), with a striking increase above an array-specific detection threshold, showing detectable expression for 21,971 transcripts (50.3%).

We next compared $h^2$ values for all genes to multiple external ‘predictors’ (refs. 1,36–44) using an enrichment statistic rigorously evaluated under permutations of twin zygosity (Table 2). Heritability was strongly associated with expression mean and variance. Regional GC content was negatively associated with $h^2$ after correction for mean expression. This negative association was surprising, as GC content $±$ 5 kb from the transcription start site (TSS) was positively correlated with gene density ($r = 0.40$), and each correlated modestly with mean expression ($r = 0.11$ and 0.10, respectively). Accordingly, after correction for mean expression, the negative association with gene density was even stronger (Fig. 2a and Table 2). Genes with recent evolutionary acceleration in primates and humans42 showed significant ($P = 7.11 × 10^{-5}$ and $3.73 × 10^{-5}$) positive association with $h^2$ after correction for mean expression (Fig. 2b and Table 2). HomoloGene conservation was highly significant ($P = 2.00 × 10^{-17}$), although it was attenuated after correction. Associations between $h^2$ and numerous Kyoto Encyclopedia of Genes and Genomes (KEGG) and GO pathways were also highly significant (Supplementary Table 3). Interestingly, all pathway associations with $h^2$ were positive, except for two related to sensory perception and smell (GO:0050907 and GO:0050911, respectively).

To investigate disease relevance, we used the National Human Genome Research Institute (NHGRI) GWAS catalog (17 July 2013)1, identifying the nearest gene (GWAS genes) for each of 3,628 significantly disease-associated SNPs ($P ≤ 5 × 10^{-8}$), for a total of 3,243 GWAS genes. There was a highly significantly positive association between heritability and GWAS genes (Fig. 2b and Table 2). Enrichment remained elevated for genes that were nearby but not necessarily the closest to the GWAS-identified SNP, and genes with numerous nearby GWAS SNPs were especially heritable (Supplementary Fig. 2). Enrichment was attenuated by removing chromosome 6 genes (including the MHC region) and for immune-related diseases43.

Table 1 Demography of 2,752 subjects from 1,444 twin pairs for twin-based heritability analyses

| Variable                      | Median or proportion | Quartiles         |
|-------------------------------|----------------------|-------------------|
| Age (years)                   | 32                   | 28–39             |
| Body mass index (kg/m²)       | 23.3                 | 21.3–25.8         |
| White blood cell count (10³/ℓ) | 6.3                  | 5.3–7.4          |
| Hematocrit (fraction)         | 0.42                 | 0.40–0.45         |
| Female sex                    | 0.658                |                   |
| Blood draw between 7:00 and 11:00 a.m. | 0.940              |                   |
| Fasting at time of blood draw | 0.947                |                   |
| Current smoker                | 0.216                |                   |
| Alcohol user (12 drinks/year) | 0.771                |                   |
| Quartiles, interquartile range. |                     |                   |
Figure 1 Transcriptome-wide estimates of heritability based on 2,752 twins. (a) Manhattan plot of heritability \( P \) values for the transcript with the highest \( h^2 \) estimate for each of 38,392 genes. The inset (PADIRZ) shows that the evidence for heritability is based on higher correlation between monozygotic pairs (MZ) than between dizygotic pairs (DZ). The dashed line marks the threshold for genes with \( q < 0.05 \). (b) Clustering of 777 genes with \( q < 0.05 \) for \( h^2 \) estimates. The most heritable genes belong to the cluster with the lowest intergene correlation, but many significant genes belong to clusters with high intergene correlation. (c) Among 43,628 transcripts, the significant proportion (in terms of FDR \( q \) value) is dependent on mean transcript expression, increasing rapidly for transcripts above an approximate detection threshold (RMA expression \( \geq 3.584 \), determined as the 90th percentile of chromosome Y RMA ‘expression’ in females).

(Supplementary Table 4). GWAS phenotypes included ones relevant to blood and immunity along with the central nervous system, the bowel, cancers and morphological traits. Given that GWAS genes were designated only on the basis of proximity to NHGRI-listed SNPs, these results may reflect an even stronger true tendency of disease-causing genes to be highly heritable (Supplementary Fig. 2). These results are complementary to observations that disease-associated SNPs show eQTL enrichment. Additionally, the Online Mendelian Inheritance in Man (OMIM) database shows similar heritability enrichment, even though NHGRI GWAS and OMIM only partly overlap (of genes in either list, 10% are in both). The OMIM genes with significant heritability (\( q < 0.05 \)) are also quite diverse, further supporting the potential relevance of peripheral blood to other tissues and developmental processes (Supplementary Table 5). Moreover, evolutionary associations are consistent with the observation that heritability is necessary for responsiveness to selection.

We emphasize that these results do not imply causality, and, in particular, disease associations should be interpreted with caution. Enrichment of disease-associated heritability may reflect other underlying sources of commonality but still point to transcription as an important intermediary in disease risk.

Local genetic contributions and bias in \( h^2 \) estimation

After genotyping quality control and imputation, 8.3 million SNPs were available for eQTL mapping in 2,494 individual twins (90.4% of the expression data set). We evaluated multiple predictors of heritability, including association \( r^2 \) values based on the most significant local SNP within 1 Mb, \( r^2 \) values for the top distant SNP, local SNP heritability estimation based on genetic relatedness among unrelated subjects using Genome-wide Complex Trait Analysis (GCTA) and variance-component results from complete local identity-by-descent inference among the dizygotic pairs (local IBD). We computed ratios of each component to the overall \( h^2 \) estimate (Supplementary Fig. 3). Mean and median values for \( r^2_{\text{localSNP}}/h^2 \) (0.04 and 0.09, respectively) were similar to those reported in the MuTHER study\(^8\), whereas the \( h^2_{\text{localIBD}}/h^2 \) ratio was higher (median = 0.11, mean = 0.30), consistent with higher explained variation when the total local contribution was considered. However, in published studies, estimates have been complicated by bias and variability in \( h^2 \) estimation. MuTHER reported mean \( h^2 \) values in expressed genes of 0.16 (skin), 0.21 (LCLs) and 0.26 (adipose), with >20% of expressed genes displaying \( h^2 > 0.3 \) (ref. 8). Our study, although much larger, produced lower values of 0.14 and 12.3%. Each of our \( h^2 \) estimates should be unbiased, as we allowed for negative estimates (even if \( h^2 \geq 0 \)), whereas variance-component methods\(^8\) can produce bias by forcing estimates to be non-negative, and sampling variability further complicates the view.

To more definitively assess the true extent of transcriptomic heritability for our study, we modeled true \( h^2 \) values as following a gamma distribution, with sampling variation determined by the ACE model. The result (Fig. 3a) was a shrunken distribution with a similar mean \( h^2 \) value but markedly less variation. The model estimated that the true proportion of expressed genes with heritability of \( >0.3 \) was actually only 7.9%. With high heritability thresholds, differing results across studies can appear dramatic—whereas the MuTHER report estimated >700 expressed genes in both skin and LCLs with heritability of >0.5, we estimate the true number in our study as ~100. The studies differed in tissue type and platform (the MuTHER study used the Illumina HT-12 BeadChip platform), NTR mean age was ~20 years younger and the NTR samples included both sexes. Results when age was removed as a covariate (Supplementary Note) suggested that it was not an important heritability determinant in NTR. However, the important effect of sampling variation has not been fully explored. First, we assessed the gamma fit by artificially adding sampling error to the true distribution, showing that it fit our estimated \( h^2 \) distribution (Fig. 3a). A similar approach quantified the impact of sample size (Supplementary Note), again using the gamma model.
Table 2 Predictors of high heritability expression levels

| Predictor                                | Mean $h^2$ change | Enrichment $z$ | $P$     | Expression-corrected enrichment $z$ | $P$     |
|------------------------------------------|------------------|----------------|--------|-------------------------------------|--------|
| Mean expression                           | 11.25            | 2.43 × 10$^{-29}$ | –      | –                                   | –      |
| Variance in expression                   | 14.14            | 2.23 × 10$^{-45}$ | 14.89  | 4.02 × 10$^{-50}$                   | –      |
| GC content, +5 kb of TSS                | –1.42            | 0.155           | –5.33  | 9.60 × 10$^{-8}$                    | –      |
| GC content, –5 kb of TSS                | –0.72            | 0.471           | –5.00  | 5.73 × 10$^{-7}$                    | –      |
| DHS near TSS                             | 9.45             | 3.55 × 10$^{-21}$ | 4.01   | 6.00 × 10$^{-5}$                    | –      |
| DHS near TSS, blood                      | 8.87             | 7.02 × 10$^{-16}$ | 1.30   | 0.195                               | –      |
| Gene density                             | –6.98            | 2.98 × 10$^{-12}$ | –10.85 | 2.09 × 10$^{-27}$                   | –      |
| Local recombination rate                 | 8.07             | 7.02 × 10$^{-16}$ | 11.30  | 1.27 × 10$^{-29}$                   | –      |
| Size of LD block                         | –0.05            | 0.959           | –0.49  | 0.622                               | –      |
| Gene conservation score                  | 8.49             | 2.00 × 10$^{-17}$ | 1.14   | 0.255                               | –      |
| Genes under selection (185)              | 0.013            | 1.60 × 10$^{-10}$ | 1.82   | 0.068                               | –      |
| Genes under positive selection (549)     | 0.007            | 1.32 × 10$^{-16}$ | 1.78   | 0.074                               | –      |
| Genes under balancing selection (47)     | 0.042            | 2.65 × 10$^{-8}$  | 2.83   | 0.0046                              | –      |
| Genes under adaptive selection (174)     | 0.019            | 2.26 × 10$^{-24}$ | 1.13   | 0.260                               | –      |
| Human accelerated genes (161)            | 0.024            | 3.05 × 10$^{-23}$ | 4.12   | 3.73 × 10$^{-5}$                    | –      |
| Primate accelerated genes (137)          | 0.024            | 2.86 × 10$^{-22}$ | 3.97   | 7.11 × 10$^{-5}$                    | –      |
| NHGRI GWAS catalog (2,343)               | 0.018            | 7.42 × 10$^{-13}$ | 7.52   | 5.53 × 10$^{-14}$                   | –      |
| NHGRI, chr. 6 genes removed (2,142)      | 0.016            | 6.06 × 10$^{-10}$ | 6.42   | 1.36 × 10$^{-10}$                   | –      |
| NHGRI, immune diseases (720)             | 0.032            | 7.22 × 10$^{-13}$ | 5.77   | 7.99 × 10$^{-9}$                    | –      |
| NHGRI, non-immune diseases (1,623)       | 0.011            | 3.71 × 10$^{-10}$ | 4.88   | 1.03 × 10$^{-6}$                    | –      |
| OMIM disease entries (3,089)             | 0.018            | 8.87 × 10$^{-10}$ | 7.54   | 4.81 × 10$^{-14}$                   | –      |
| NHGRI + OMIM (4,809)                     | 0.019            | 10.81 × 10$^{-27}$ | 9.84   | 7.81 × 10$^{-23}$                   | –      |

TSS, transcription start site; DHS, DNase I–hypersensitive site; NHGRI, National Human Genome Research Institute; GWAS, genome-wide association study; OMIM, Online Mendelian Inheritance in Man. Values in bold correspond to $P < 0.00022$, for Bonferroni significance at $\alpha = 0.05$ for the 23 tests in each of uncorrected and corrected analyses. *From Encyclopedia of DNA Elements (ENCODE) Duke UCSC tracks. **Defined as the reverse ranked variance of the base-pair positions of the gene and two flanking genes. ***The end transcription base-pair position minus the start transcription base-pair position. ****ENCODE sex-averaged standardized recombination maps in 10-kb bins. *****Linkage disequilibrium block boundaries as described in the Supplementary Note. ******NCBI HomoloGene (Build 66) score, defined as the ratio of the number of appearances in other organisms to the total of 21. **Ref. 36, genes with the property shown in parentheses. ***Ref. 36–38. Ref. 39. Ref. 35. Ref. 41. Ref. 1, for SNPs with $P < 5 × 10^{-8}$. ****Following classification in ref. 42. Ref. 43.

Figure 2 Gene density and other predictors of heritability, using 2,616 paired twins and 18,392 genes. (a) Mean $h^2$ estimates (corrected for gene expression levels) versus density of protein-coding genes per autosome, showing that heritability is considerably higher for gene-poor chromosomes. Plot symbol area is proportional to the number of genes present on the array per chromosome. (b) Histograms of the permuted enrichment $z$ statistics for two predictors listed and defined in Table 2. Observed values (blue circles) are extreme compared to the permutations. The value for this ratio of $0.0042$ will require more complex modeling to handle correlation structures in the measurements and underlying true structure. However, the results from our large sample support the view that local genetic variation explains only a minority of transcriptomic heritability and much of the unexplained variation is among genes with modest $h^2$ estimates. A regression approach (Supplementary Fig. 5) showed that ~35% of the variation in estimated $h^2$ values could be explained by the predictors.

eQTL analyses of peripheral blood

We next analyzed genotypes as predictors of transcription (a GWAS for each transcript) for 2,494 twins, using an REMI (restricted maximum-likelihood) model accounting for twin status and covariates. eQTLs within 1 Mb upstream of the TSS and 1 Mb downstream of the transcription end site of a gene were classified as ‘local’, and all others were classified as ‘distant’, with separate false discovery rate (FDR) control. Genes with at least one local eQTL ($q < 0.01$) had significantly higher expression levels and heritability ($P < 1 × 10^{-200}$ for both).

The effect of sample size on local eQTL identification is shown in Figure 4a, which includes nearly all published blood-derived eQTL studies7,8,15,20,31,48–52 (comparisons to the large meta-analysis in ref. 29 are described separately), the full NTR data ($n = 2,494$) and random subsamples of our data. We reanalyzed the data sets using a common quality control pipeline on inverse quantile-normalized data19 (except where unavailable8,15). For comparison, we selected a set of unrelated twins (1,263 individuals) and performed local

higher total heritability8. A definitive statement of average per-gene ratios ($h_{localIBD}^2/h^2$)
eQTL mapping on random subsets of varying sample size, using fewer covariates (no blood counts or SNPs) and ~600,000 genotyped SNPs. We also evaluated our robustness approaches (normal quantile transformation and normal quantile transformation with SNP minor allele frequency (MAF) of >0.005 or >0.01 in each subsample). For local eQTLs, there was little difference among the transformations.

There is considerable interstudy variability in the number of significant eQTLs (Fig. 4a), even with consistent quality control and analysis19,31. With increasing sample size, it seems that most expressed genes (>10,000) show evidence of local eQTL influence in peripheral blood. For NTR, the number of genes with significant eQTLs (q < 0.01) was 11,834, and, after employing final quality control steps, there were 9,640 significant genes. Replication was examined in 1,895 unrelated samples from the Netherlands Study of Depression and Anxiety (NESDA), which had a similar sex distribution (68% female) and age range (from 18 through 65 years). Reproducibility of eQTLs between NTR and the 1,895 unrelated NESDA samples is shown in Supplementary Figure 6, and enrichment and deficits of regulatory features for local eQTL SNPs are shown in Supplementary Figure 7. Of the 9,640 genes with local eQTLs in NTR (at least 1 SNP with q < 0.01), 9,148 (94.9%) replicated (q < 0.1) in NESDA (using a less stringent replication q-value threshold to allow for winner’s curse attenuation). This approach was not intended to control the per-gene FDR but to focus on genes with the greatest evidence of replication. Of the genes with the strongest evidence of local eQTLs in NTR (q < 0.001), 6,756 of 6,941 (97.3%) replicated in NESDA. There was strong overlap (P = 1 × 10−18) of genes with local eQTLs in the full NTR sample, with the same gene having a local eQTL in a meta-analysis of HapMap LCL studies19. For genes with local eQTLs (q < 0.1) in the LCL meta-analysis, 56.1% (2,417/4,306) also had significant local eQTLs in NTR. Genes that replicated had smaller meta-analysis q values (P = 1 × 10−18), along with higher expression (P = 2 × 10−119) and higher heritability (P = 8 × 10−131) in NTR. The lack of overlap among smaller HapMap samples is likely an example of winner’s curse: considering two larger studies15,20, among the genes annotated in all three studies, replication in NTR was 66.8% (2,799/4,189 genes) and 77.2% (3,404/4,412 genes), respectively (Fig. 4b). Similarly, for local gene-SNP pairs with q < 0.05 from the peripheral blood eQTL

Figure 4 Comparison and replication of eQTL results. (a) Number of unique genes with evidence of local association (q < 0.01; within 1 Mb of gene), depicted for published leukocyte eQTL studies (LCLs19, monocytes15 and peripheral blood leukocytes (PBLs)20), as well as subsampling of NTR data (PBLs) using only genotyped markers and moderate quality control (n = 2,494; 43,628 transcripts examined). Sample sizes are corrected for the number of covariates used. The “NTR” with final QC value applies q < 0.001. qnorm refers to the rank inverse quantile-normalized expression data. Ancestries are CEU (Northern and Western European), YRI (Yoruba in Ibadan, Nigeria), CHB (Han Chinese in Beijing, China), JPT (Japanese in Tokyo, Japan), LWK (Luhya in Webuye, Kenya) and GIH (Gujarati Indians in Houston, Texas). (b) Overlap of local eQTL findings with two other large blood studies, at q < 0.01. (c) Number of unique genes with evidence (q < 0.01) for distant association (>1 Mb from gene). The implausible non-monotone pattern for NTR on original expression values shows the importance of robust association methods. Using final quality control on NTR data and q < 0.001 drops the number of distant eQTLs from over 800 to ~300. The results suggest that many distant associations remain to be discovered, but careful quality control is essential. (d) Overlap of distant eQTL findings (q < 0.001) with previous studies (within 1 Mb of gene).
meta-analysis of Westra et al.\(^\text{29}\) (\(n = 5,311\)), the estimated true discovery rates in NTR and NESDA were 59.6\% and 59.7\%, respectively (Supplementary Fig. 8 and Supplementary Note).

**Characteristics of distant eQTLs**

Robust distant eQTL results (Fig. 4c, expression transformed to an exact normal) were again consistent with published studies, roughly linear (log-log scale) with sample size\(^\text{15}\). For NTR, we obtained a robust set of 348 distant eQTLs by applying stricter significance criteria (\(q < 0.001\)) followed by additional careful quality control. Extrapolating to larger sample sizes, we anticipate the identification of 1,000 replicating eQTLs, even for sample sizes exceeding 5,000. Overlap of genes with significant distant eQTLs (\(q < 0.001\)) among the large studies is shown in Figure 4d, with much lower overlap for distant eQTLs than for local eQTLs. For significant distant gene-SNP pairs from Westra et al.\(^\text{29}\) (\(n = 5,311\)), the estimated true discovery rates in NTR and NESDA were 23.1\% and 23.0\%, respectively (Supplementary Fig. 8).

Our 601 distant eQTLs with \(q < 0.001\) (Fig. 5) involved 581 genes and 538 non-redundant SNPs (for each gene, only the most significant SNP per chromosome was retained). We applied additional quality control to these highly significant distant eQTLs (Supplementary Note), reducing the number of eQTLs to 348 (57.9\%), of which 165 (47.4\%) replicated in NESDA (\(q < 0.01\)) (Fig. 5, Supplementary Fig. 6 and Supplementary Table 6). Genes in the 348 eQTLs were analyzed using the Database for Annotation, Visualization and Integrated Discovery (DAVID) \(p\) values for KEGG and GO enrichment, which are mediated by unconventional myosins, which bind to membranous compartments and serve in intracellular movements. SOX13 encodes a transcription factor that modulates the WNT-TCF signaling pathway\(^\text{30}\), and several other distant eQTL genes are involved in cellular signaling (for example, TMEM134, RGS12 and SYT13). Additional network estimation was performed for the replicating hotspots (Supplementary Fig. 12), but the relatively few genes influenced by hotspots or distant eQTLs suggest that such networks do not have a predominant role in steady-state transcription in peripheral blood.

---

**Figure 5** Properties of distant eQTLs. (a) In total, 348 eQTLs (gene-SNP pairs) were significant (\(q < 0.001\)) and passed the quality control procedures; of these, 165 replicated (\(q < 0.01\)) in 1,895 NESDA individuals. (b) We examined 304 SNPs in significant eQTLs for overlap with regulatory features, including DNase I/FAIRE (formaldehyde-assisted isolation of regulatory features, including DNase I/FAIRE) in 1,895 NESDA individuals. (c) The \(P\) value represents the estimated proportion of the transcriptome influenced by the 304 SNPs that passed quality control in significant eQTLs. Across all significant bins, the cumulative proportion is only \(-3\%\). (d) A distant eQTL hotspot on chromosome 19 was associated with the expression of 12 distant genes and 1 local gene (MYO1F). The partial correlation graph suggests that MYO1F expression is independent of the expression of the other distant genes, given the expression of the transcription factor SOX13.
Biomedical relevance

This catalog of eQTLs can be used to generate in silico hypotheses for biomedical follow-up using peripheral blood as a proxy tissue. Using the NHGRI GWAS catalog, after stringent filtering ($P < 1 \times 10^{-8}$), there were significant results for 3,415 SNPs, 498 traits and 4,167 SNP-trait pairs from 927 reports. The greatest numbers of SNPs associated with a trait or disease were found for height ($n = 248$), high-density lipoprotein (HDL) cholesterol ($n = 92$), Crohn’s disease ($n = 155$), type 2 diabetes ($n = 98$) and ulcerative colitis ($n = 81$). The extended MHC region (chr. 6: 25–34 Mb, 0.3% of the genome) was the second most gene-dense region of the genome and contained the greatest number of SNPs implicated by GWAS (6.8%). Of the 4,167 SNP-trait pairs implicated by GWAS, 534 (12.8%) were part of a local eQTL (either directly or via a proxy SNP with $r^2 > 0.5$).

To complement the analyses, we evaluated genes cataloged in OMIM (downloaded 17 July 2013) of the 3,118 genes in OMIM, 74.4% were part of a SNP-gene local eQTL pair ($q < 0.05$). These included many genes related to immune and hematological abnormalities, muscular dystrophy (21 genes) and genes implicated in nervous system diseases. Examples include Alzheimer’s disease (APP and PSEN2), deafness (42 genes), amyotrophic lateral sclerosis (15 genes), Charcot-Marie-Tooth disease (25 genes), epilepsies (21 genes) and candidate genes for schizophrenia (DISC1, D4OA and RGS4). Of the 517 genes implicated in mendelian autism spectrum disorders or mental retardation or mental retardation (44,58,59, 69.6% were part of a local eQTL SNP-gene pair. Of the 3,294 genes with a copy number variant implicated in disease status was significantly enriched ($P = 1.87 \times 10^{-13}$), which we attribute to the neglect of the X chromosome.

Finally, we combined heritability predictors and gene-disease designations into several multiple regressions (Supplementary Table 7). Predictors were as shown in Table 2, with the addition of eQTL evidence (best local and distant $r^2$ values), chromosome 6 (human leukocyte antigen (HLA) genes), chromosome 19 (an outlier in gene density analysis), the X chromosome (under-represented in GWAS) and a blood DNase I hypersensitivity–gene conservation interaction (identified in exploratory analyses). eQTL evidence alone (top local and distant SNPs) explained 23.9% of the variation in $h^2$ estimates, and the full model explained 32.9%. $h^2$ estimates remained significantly predictive of OMIM and NHGRI GWAS disease status except for the smaller sets of NHGRI-cataloged genes subdivided by immune designation, even when the best local and distant eQTLs were no longer significant. Gene conservation was highly predictive of OMIM status. Gene density showed strong negative association with disease status, but this effect was attenuated for OMIM. NHGRI disease status was significantly enriched ($P = 1.70 \times 10^{-10}$; Supplementary Table 7) for chromosome 6 loci and showed a deficit on the X chromosome ($P = 1.87 \times 10^{-13}$), which we attribute to the neglect of the X chromosome in GWAS. OMIM showed enrichment of the X chromosome, consistent with the importance of X-linked disorders in medical genetics.

DISCUSSION

We have established clear patterns underlying the heritability of steady-state gene transcription in peripheral blood and have demonstrated strong connections to disease annotation. The use of peripheral blood enables further investigation of immune-related diseases but may also be useful for other tissues. Our results supply mechanistic hypotheses that can be evaluated in subsequent experiments. In comparisons across four mouse tissues, we found that genes expressed in multiple tissues tended to have cis regulatory elements (J.J. Crowley, V. Zhabotinsky, W. Sun, S. Huang, I.K. Pakatci et al. unpublished data).

Examination of $h^2$ estimates relative to gene density builds upon a literature demonstrating that essential genes expressed in many tissues can occur in dense clusters of high expression, including instances of transcriptional colocalization. Essential genes (those that when mutated cause lethality at or before birth) identified in mouse mutagenesis screens show high linkage conservation, and intergenic regions in humans have higher SNP densities than in introns, along with higher rates of neutral polymorphisms. Our observations seem concordant with these reports, whether selection directly inhibits its heritability in gene-dense regions or the inhibition is due to the relative paucity of genotype variation in such regions.

The ability of $h^2$ estimates to predict OMIM and NHGRI designations may suggest new approaches to augment association mapping, as current approaches generally focus on the sequence context of associated SNPs rather than the genes themselves. The ability to detect heritability only in expressed genes somewhat complicates interpretation, given the higher average expression in high-density clusters and the lack of information for genes not expressed in this tissue. Critically, full elucidation of these relationships may be possible only with careful cross-tissue eQTL analysis of a large number of individuals.

ACKNOWLEDGMENTS

The work described in this paper was funded by the US National Institute of Mental Health (RC2 MH089951, principal investigator P.F.S.) as part of the American Recovery and Reinvestment Act of 2009. Transport, extraction and preparation of the NTR samples were carried out under a supplement to the NIMH Center for Collaborative Genomics Research on Mental Disorders (U24 MH068457, principal investigator J.T.). We thank T. Lehner (National Institute of Mental Health) for his support. Additional analytic support was provided by grants R01 MH09396, R01 GM074175 and P42 ES003948 and by a Gillings Innovations Award. The Netherlands Study of Depression and Anxiety (NEDSA) and the Netherlands Twin Register (NTR) were funded by the Netherlands Organization for Scientific Research (MagW/ZonMW; grants 904-61-090, 985-10-002, 904-61-193, 480-04-004, 400-05-717 and 912-100-20; Spinozopremie 56-464-11429; and Geestkracht program grant 10-000-1002), the Center for Medical Systems Biology (CMSB2, NWO Genomics), Biobanking and Biomolecular Resources Research Infrastructure (BBMRI-NL), the VU University EMGO+ Institute for Health and Care Research and the Neuroscience Campus Amsterdam, NBIC/BioAssist/RR (2008.024), the European Science Foundation (EU/QLRT-2001-01254), the European Community’s Seventh Framework Programme (FP7/2007–2013), ENGAGE (EU7-F4-2007-201431) and the European Research Council (ERC; 230374).

AUTHOR CONTRIBUTIONS

Study design and writing: F.A.W., P.F.S., A.I.B., E.Z., W.S., B.W.J.H.P. and D.I.B.

Analysis: F.A.W., P.F.S., E.Z., W.S., K.K., V.M., R.J., W.C., Y.-H.Z., A.A., G.C., T.-H.C., P.G., M.J.H., J.J.H., S.H., M.K., J.K., C.M.M., A.Q., A.S., J.-Y.T., Q.W., Wei Wang, Weibo Wang, G.W., J.H.S., E.J.d.G. and Z.Y. Genomic assays: A.I.B., D.D., J.T., A.Q. and Q.W. Phenotype collection: G.W.G. and J.M.V. Project management: T.L. Database design and management: S.B. and C.B.
ONLINE METHODS

Subjects and biological sampling. Subjects were ascertained and sampled using harmonized protocols from two longitudinal cohort studies, the Netherlands Twin Registry (NTR) and the Netherlands Study of Depression and Anxiety (NESDA). NTR is an observational, 25-year longitudinal study of twins and their families. The study protocol was approved by the Central Ethics Committee on Research Involving Human Subjects of the VU University Medical Center. NESDA is a cohort study to investigate the long-term course and consequences of depressive and anxiety disorders and includes persons both with and without emotional disorders.

Third, we verified sample identity and accompanying data. Using RMA normalization (Affymetrix Power Tools, v1.12.0), the study protocol was approved by the Ethical Review Board of the VU University Medical Center and, subsequently, by the local review board of each participating center.

Peripheral venous blood samples were drawn in the morning (NTR, 7:00—11:00 a.m.; NESDA, 8:30—9:30 a.m.) after an overnight fast. For fertile women in NTR, samples were obtained on days 3–5 of their menstrual cycle or in the pill-free week if on oral contraception. Heparinized whole blood was transferred into PAXgene Blood RNA tubes (Qiagen) within 20 min (60 min for NESDA), incubated and stored at −20 °C or −30 °C (NTR). High-molecular-weight genomic DNA was isolated using PureGene DNA isolation kits (Qiagen).

Gene expression assays. Gene expression assays for NTR and NESDA were conducted at the Rutgers University Cell and DNA Repository. Total RNA was extracted at Rutgers (for NESDA, at the VU Medical Center) using the PAXgene Blood RNA MDx kit protocol in 96-well format with the BioRobot Universal System (Qiagen). RNA quality and quantity were assessed by Caliper AMS90 with HT DNA 5K/HT RNA LabChips. Samples were randomized to plates, with checks to ensure sex and zygosity balance. Co-twins were randomized without respect to relationship to avoid bias in family correlation estimates.

For cDNA synthesis, 50 ng of RNA was reverse transcribed and amplified in a plate format on a Biomek FX liquid-handling robot (Beckman Coulter) using Ovation Pico WTA reagents (NuGEN). Products purified from single-primer isothermal amplification (SPIA) were fragmented and labeled with biotin (Encore Biotin Module, NuGEN), and size distributions were verified (Caliper AMS90, HT DNA 5K/RNA LabChips). Samples were hybridized to Affymetrix U219 array plates (Supplementary Note) to enable expression profiling in 96-sample sets. Array hybridization, washing, staining and scanning were carried out in an Affymetrix GeneTitan System according to the manufacturer’s protocol.

Quality control was conducted on NTR and NESDA data in parallel. Expression data were required to pass standard Affymetrix Expression Console quality metrics before further undergoing quality control. The array superset consisted of 6,526 U219 arrays (3,516 NTR samples, 2,783 NESDA samples, divided into baseline samples and a smaller portion after 2-year follow-up, and 227 controls) on 69 plates, including 417 samples that were identified as having reduced quality (D < −5.0) and were rehybridized. Expression values were obtained using RNA normalization (Affymetrix Power Tools, v1.12.0).

Probe sequences were mapped to the human genome (hg19) using Bowtie2 and probes with sequences not mapping, mapping to multiple locations or intersecting a polymorphic SNP (HapMap 3 and 1000 Genomes Project data) were removed.

The large sample size enabled additional quality control metric involving intersample comparisons. First, samples showing sex inconsistency were removed (on the basis of X-chromosome and Y-chromosome probe sets). Second, we examined the pairwise correlation matrix of expression profiles. Using $r_{ij}$ as the correlation between arrays i and j, we computed

\[ \bar{r}_i = \frac{1}{n} \sum_{j=1}^{n} r_{ij} \]

the average correlation of array i with all others of the total n arrays. Lower $\bar{r}_i$ values correspond to lower quality and were expressed in terms of median absolute deviations $D_2 = (|\bar{r}_i - \bar{\bar{r}}|)/\text{median}(|\bar{r}_i - \bar{\bar{r}}|)$ to provide a sense of distance from the grand correlation mean $\bar{\bar{r}}$.

Third, we verified sample identity on the basis of U219 gene expression data and Affymetrix 6.0 genotypes, having previously discovered genotype-expression mismatch rates of up to 5% in published eQTL studies.

Briefly, 500 of the most significant SNP-transcript local eQTL pairs were used to estimate a posterior probability for a match between gene expression and genotype profile (similar to in ref. 78). This approach identified sex-mismatched samples and additional samples of poor quality.

Fourth, initial analysis using unrelated participants showed the potential for spurious eQTL identification owing to expression outliers. Thus, conservatively, we transformed expression values using inverse quantile normal transformation, which results in values that precisely fit a normal distribution. These values were used for all primary analyses. Fifth, we evaluated the effects of covariates on gene expression and found significant associations for plate, hybridization well position, age at blood sampling, sex, time interval between extraction and hybridization steps, total white and red blood cell counts, hematocrit and the top five expression principal components (similar to that of surrogate variables).

Imputation was performed to estimate a small proportion of missing covariates (2.1%). All heritability and eQTL analyses corrected for these covariates (93 degrees of freedom), and eQTL analyses additionally corrected for the first 3 genotype principal components.

Sixth, we observed that D values and the posterior probability of mismatch were highly correlated, and we reasoned that D values might be useful in removing additional low-quality samples. To determine the optimal threshold for D, we successively removed individual samples according to D value, and recomputed the intraclass correlation coefficient (ICC)-based estimate of heritability (2($\hat{h}_a^2$ − $\hat{h}_d^2$)) and accompanying P values for all transcripts using covariate-residualized expression data. Benjamini-Hochberg FDR q values for transcripts were computed using p.adjust in R (v2.14).

Removal of 19 samples with the lowest D values resulted in the largest number of significant transcripts (q < 0.10; Supplementary Note). The optimal choice of samples to remove was largely robust to the q-value threshold in the range q = 0.05–0.20 and to the use of non-normalized expression data.

After expression quality control, the U219 gene expression set consisted of 2,752 NTR subjects. An additional 1,895 NESDA subjects (representing the NESDA baseline set) were used for replication in this report. Expression quality control for NESDA followed the same steps as for NTR (except that zygosity did not apply). Expression distributions for monzygotic and dizygotic twins were compared for differing mean expression ($t$ test) and differing variances ($F$ test for normally distributed data), performed separately within twin sets 1 and 2. No transcript showed significantly different mean expression between monzygotic and dizygotic twins, but four transcripts showed significantly different (FDR $q < 0.05$) variances. However, of these four transcripts, none showed $q < 0.05$ for $h^2$ estimates.

Genome-wide SNP assays. Genomic DNA was tested using 96 TaqMan SNP Genotyping assays (RUID panel) with Fluidigm 96 GT Dynamic Array chips, a BioMark Genetic Analysis instrument and SNP Genotyping Analysis Software (v3.0.2). After quality, sex and identity of genomic DNA samples were verified, all samples were randomized to plates. Genotyping was conducted using Affymetrix Genome-Wide Human SNP Array 6.0 (Supplementary Note) according to the manufacturer’s protocol. Resulting data were required to pass standard Affymetrix quality control metrics (contrast QC > 0.4) before further analysis.

SNP quality control is detailed in the Supplementary Note. Briefly, quality control included the removal of SNPs for non-unique probes mapping to NCBI Build 37/UCSC hg19, low MAF (<0.005, determined empirically), substantial deviation from HapMap 3 CEU (Utah residents of Northern and Western European ancestry) founder allele frequencies, deviation from Hardy-Weinberg equilibrium ($P < 1 \times 10^{-5}$) or high missingness ($>0.05$). Subjects were eliminated from analysis for high missingness ($>0.05$), outlying genome-wide homoygosity or ancestry, discrepant genetic and phenotypic sex, or twin relatedness inconsistent with monzygosity or dizygosity. Resulting genotypes were of high quality, with relatively low SNP and subject missingness (97.5 percentiles of 0.035 and 0.020, respectively). Among 714 monzygotic twin pairs, the intrapair agreement for 686,895 autosomal SNPs was 0.9985. Previous genome-wide genotyping using a Perlegen 4-chip platform was available for 2,219 subjects and 110,588 SNPs and showed 0.9996 agreement with Affymetrix 6.0 genotyping.
Phased genotype calls on 379 European samples from the 1000 Genomes Project were used as the reference set for imputation. NTR samples were split into two unrelated sets. SNPs with call rate of <95% or Hardy-Weinberg equilibrium \( P < 1 \times 10^{-7} \) were excluded. Imputation was performed using MACH. For each NTR set, MAF bins of (0.005, 0.1), (0.01, 0.03), (0.03, 0.05) and (0.05, 0.5) were defined, and within each bin an \( r^2 \) threshold was defined such that average \( r^2 \) = 0.8. The \( r^2 \) thresholds were 0.55, 0.4, 0.3 and 0.3, respectively. The final SNP numbers were 8.4 million for each of the twin sets, with the intersection of 8.3 million SNPs used here.

Heritability. Three methods for estimating heritability are detailed in the Supplementary Note. The primary approach was twin-based heritability via an REML mixed model, with random additive genetic components of variation, along with shared and individual-specific environmental effects plus selected covariates as fixed effects. Random terms were assumed to be mutually independent and normally distributed with mean of 0 and variances \( \sigma_a^2 \) and \( \sigma_e^2 \). This corresponds to a standard ACE model and assumes that dizygotic twins have an average IBD proportion of 0.5 (refs. 81,82). For each transcript, the twin-based heritability and shared environmental effects were estimated, respectively, as \( h^2 = \sigma_a^2 / (\sigma_a^2 + \sigma_c^2 + \sigma_e^2) \) and \( c^2 = \sigma_c^2 / (\sigma_a^2 + \sigma_c^2 + \sigma_e^2) \). The ACE model can be fit using either variance-component maximization, constraining \( h^2 \) and \( c^2 \) to be non-negative or using an unconstrained general covariance structure. After establishing that results from the two approaches were highly concordant, we used the unconstrained approach to best match the intraclass correlation approach\(^{80}\) used for pathway analysis. Under additive assumptions, \( h^2 \) is the heritability estimate, and \( P \) values are reported for the right tail (positive \( h^2 \)) except where noted. \( P \) values for the X chromosome were obtained using separate heritability analysis for males and females (using identical methods as for autosomes) and then combined using Fisher’s method.

For the analyses in Supplementary Table 7, \( h^2 \) values for the X chromosome were obtained by ignoring twin sex, producing an approximate average across the sexes. After calculating results for all 47,628 transcripts, a unique ‘best \( h^2 \)’ set used the most significant transcript for each of the 18,293 genes, with FDR control applied to the best \( h^2 \) set in a manner accounting for all transcripts.

The second heritability estimation approach was dzygotically only heritability following a constrained ACE mixed-model approach for full sibships\(^{83}\). For this approach, an REML mixed model was used to relate observed variation in true IBD proportions among dizygotic pairs to expression phenotypes. \( P \) values were obtained using likelihood ratio tests. The third approach was heritability estimated from the genetic relatedness matrix, as implemented in GCTA\(^{46}\). For this approach, we divided the NTR subjects into unrelated sets (twin set 1, \( n = 1,370 \); twin set 2, \( n = 1,372 \)) and averaged the \( h^2 \) estimates from the 2 twin sets. Results showed almost no correlation with twin-based heritability (data not shown), and we reasoned that genome-wide IBD might have reduced power for those genes influenced largely locally. Thus, we ran GCTA again, using IBD estimation performed in the local region within 1 Mb of each transcript.

Local IBD analysis. Residualized expression data showed a nearly perfect normal distribution, and the bivariate normal model of Wright\(^{84}\) for sibling pair IBD mapping was therefore applied to the dzygotically pairs, offering a potential improvement over the Haseman-Elston approach\(^{85}\). MERLIN\(^{85}\) was run on the thinned set of markers used for stratification analysis, and probabilistic IBD estimates were produced at each marker closest to or within each gene. A full maximum-likelihood approach was applied for an additive model for the effect of each increment of IBD on dzygotically twin correlation as a function of IBD status, thus extracting maximum information. The approach in ref. 84 provides a regression coefficient, which was then converted to local \( h^2 \) equivalents as the proportion of variation in the trait explained by local IBD status.

Heritability enrichment and pathway analysis. A primary question is whether heritability associates with gene sets, pathways or quantitative gene features, which we generically refer to as heritability enrichment. We employed DAVID/ EASE as a descriptive tool to investigate heritable gene clusters\(^{85}\). However, simple methods that ignore transcriptomic correlation produce very high false positive rates\(^{51}\). Furthermore, a large number of genes are heritable, necessitating ‘competitive’ enrichment testing\(^{86}\), contrasting the heritability of each set of genes with that of a complementary set. Accordingly, we devised a rigorous testing approach for each gene set. We used a covariate-residualized version of the expression data, computing the ICC-based estimate for complete twin pairs as \( h^2 = 2(\hat{p}_{MZ} - \hat{p}_{DZ}) \) for all genes using the transcripts with the best \( h^2 \) values. For the observed data, this approach was highly consistent with REML estimates \((r = 0.992; \text{Supplementary Note})\). Twin zyosity status was permuted 1,000 times (additive coding as copies of the minor allele) and with the difference in mean \( h^2 \) for the gene set versus the complementary set. As this difference showed a nearly normal distribution, an enrichment \( z \) statistic was calculated as the observed difference divided by its permutation s.d., and a two-sided \( P \) value was computed assuming normality. A similar approach was used for continuous predictors in which the correlation between \( h^2 \) and the predictor was computed (with \( z \) as the correlation divided by its s.d.). By permuting only zyosity status, the enrichment \( z \)-score approach preserves mean twin pair correlations, as well as gene-gene correlations. To control for the complicating effects of mean expression, some analyses (including all KEGG and GO pathway analyses) were performed in which \( h^2 \) values were corrected for the effect of mean expression in the original and permuted data sets.

eQTL analysis. We refer to eQTLs as local (SNP-transcript associations within 1 Mb of the transcription start and end sites) or distant (the remaining findings). We prefer these terms to cis and trans designations, which connote a greater understanding of underlying mechanisms.

The REML twin-based model can be used for eQTL analysis by including SNP genotype (additive coding as copies of the minor allele) and computing the corresponding Wald statistic, in this manner properly handling covariates and twin correlation structure. This approach is computationally prohibitive for full eQTL analysis, so we used Matrix eQTL\(^{87}\) to rapidly screen for local or distant eQTL relationships. To account for dependence, the full REML model was then applied to all transcript-SNP associations with nominal \( P < 1 \times 10^{-5} \) (a liberal threshold for the \( > 3 \times 10^{16} \) tests performed). Separate FDR \( q \)-value error control was performed for local and distant eQTLs. After FDR correction, it was apparent that all significant results with true RELM q < 0.10 had indeed been captured. Some of the eQTL findings are reported in terms of unique genes, i.e., the most significant transcript-SNP combination for each gene, and in such instances the full testing multiplicity was considered.