An eQTL Landscape of Kidney Tissue in Human Nephrotic Syndrome.

Christopher E. Gillies
Rosemary Putler
Rajasree Menon
Edgar Otto
Kalyn Yasutake

See next page for additional authors

Follow this and additional works at: https://scholarlyexchange.childrensmercy.org/papers

Part of the Congenital, Hereditary, and Neonatal Diseases and Abnormalities Commons, Health Information Technology Commons, Medical Genetics Commons, Nephrology Commons, Pediatrics Commons, and the Urogenital System Commons

Recommended Citation

This Article is brought to you for free and open access by SHARE @ Children's Mercy. It has been accepted for inclusion in Manuscripts, Articles, Book Chapters and Other Papers by an authorized administrator of SHARE @ Children's Mercy. For more information, please contact library@cmh.edu.
An eQTL Landscape of Kidney Tissue in Human Nephrotic Syndrome

Christopher E. Gillies,1,8 Rosemary Putler,1,8 Rajasree Menon,2,8 Edgar Otto,3 Kalyn Yasutake,1 Viji Nair,3 Paul Hoover,4,5 David Lieb,5 Shuqiang Li,5 Sean Eddy,3 Damian Fermin,1 Michelle T. McNulty,1 Nephrotic Syndrome Study Network (NEPTUNE), Nir Hacohen,4,5 Krzysztof Kryluk,6 Matthias Kretzler,2,3,9 Xiaoquan Wen,7,9 and Matthew G. Sampson1,9,10,*

Expression quantitative trait loci (eQTL) studies illuminate the genetics of gene expression and, in disease research, can be particularly illuminating when using the tissues directly impacted by the condition. In nephrology, there is a paucity of eQTL studies of human kidney. Here, we used whole-genome sequencing (WGS) and microdissected glomerular (GLOM) and tubulointerstitial (TI) transcripts from 187 individuals with nephrotic syndrome (NS) to describe the eQTL landscape in these functionally distinct kidney structures. Using MatrixEQTL, we performed cis-eQTL analysis on GLOM (n = 136) and TI (n = 166). We used the Bayesian “Deterministic Approximation of Posteriors” (DAP) to fine-map these signals, eQTLBMA to discover GLOM- or TI-specific eQTLs, and single-cell RNA-seq data of control kidney tissue to identify the cell type specificity of significant eQTLs. We integrated eQTL data with an IgA Nephropathy (IgAN) GWAS to perform a transcriptome-wide association study (TWAS). We discovered 894 GLOM eQTLs and 1,767 TI eQTLs at FDR < 0.05. 14% and 19% of GLOM and TI eQTLs, respectively, had >1 independent signal associated with its expression. 12% and 26% of eQTLs were GLOM specific and TI specific, respectively. GLOM eQTLs were most significantly enriched in podocyte transcripts and TI eQTLs in proximal tubules. The IgAN TWAS identified significant GLOM and TI genes, primarily at the HLA region. In this study, we discovered GLOM and TI eQTLs, identified those that were tissue specific, deconvoluted them into cell-specific signals, and used them to characterize known GWAS alleles. These data are available for browsing and download via our eQTL browser, “nephQTL.”

Introduction

Nephrotic syndrome (NS) is a rare disease of glomerular filtration barrier failure,1,2 causing massive urinary excretion of protein, that can progress to chronic kidney disease (CKD) and end-stage renal disease (ESRD).3,4 NS is a heterogeneous disease, so we use the histologic descriptions of glomeruli on kidney biopsy to diagnose individuals with “minimal change disease (MCD)” and “focal segmental glomerulosclerosis (FSGS).” Additionally, we use an individual’s response to these treatments to give them a post hoc diagnosis of steroid-sensitive NS (SSNS) or steroid-resistant NS (SRNS).

Understanding how human genetic variation contributes to the development and progression of NS has been a fruitful strategy in gaining a more precise understanding of the molecular underpinnings of NS.5 More than 50 genes have been discovered that harbor rare variants sufficient to cause SRNS (“Mendelian” NS).6 Through genome-wide association studies (GWASs) and exome-chip studies,7–10 common genetic variants have been discovered that contribute to the pathogenesis of FSGS, pediatric SSNS, and membranous nephropathy (MN). Rare variant association studies in FSGS have implicated additional forms of NS-associated genetic variation to gain biological and clinical insights.

Expression quantitative trait loci (eQTL) studies, which use mRNA expression as a proximal (and continuous) cellular endophenotype, have increased power for the discovery of statistically significant genetic effects as compared to GWASs and provides inherent biological meaning in the associations between a regulatory variant and its associated gene.11–14 The GTEx project has generated eQTL data which is publicly available and has been used extensively to help interpret GWAS signals emerging from other complex traits.15 Meaningful eQTL discoveries using the affected tissues in other human diseases suggest their potential for NS genomic discovery as well. This is appealing given that we often obtain kidney tissue via biopsy from affected individuals.

With regards to kidney eQTLs, the final release of GTEx will have only 73 kidney cortex samples. There is also an

1Department of Pediatrics-Nephrology, University of Michigan School of Medicine, Ann Arbor, MI 48109, USA; 2Department of Computational Medicine and Bioinformatics, University of Michigan School of Medicine, Ann Arbor, MI 48109, USA; 3Department of Medicine-Nephrology, University of Michigan School of Medicine, Ann Arbor, MI 48109, USA; 4Department of Medicine, Massachusetts General Hospital Cancer Center, Boston, MA 02114, USA; 5Broad Institute of the Massachusetts Institute of Technology (MIT) and Harvard, Cambridge, MA 02142, USA; 6Department of Medicine, Division of Nephrology, College of Physicians & Surgeons, Columbia University, New York, NY, USA; 7Department of Biostatistics, University of Michigan School of Public Health, Ann Arbor, MI 48109, USA
8These authors contributed equally to this work
9These authors contributed equally to this work
10Twitter: @kidneyomicsamps
*Correspondence: mgsamps@med.umich.edu
https://doi.org/10.1016/j.ajhg.2018.07.004.
© 2018 American Society of Human Genetics.
absence of any other major public kidney eQTL datasets. This represents a significant barrier for genomic discovery in nephrology.

The most comprehensive kidney eQTL study thus far discovered them using unaffected portions of 96 nephrectomy samples from The Cancer Genome Atlas.16 The investigators integrated these eQTLs with risk loci for CKD to establish links between these risk alleles and molecular mechanisms.16 A limitation of this study was that bulk renal cortex was used for association, which is known to be 80% proximal tubule cells. The preponderance of this cell type may obscure eQTL signals emerging from the structurally and cellularly heterogeneous kidney. This study also exclusively used healthy tissue, which prevents an opportunity to potentially discover disease-context-specific eQTL effects.

Microdissecting bulk renal cortex tissue into its two main functional structures, the glomerulus (GLOM) and tubulointerstitium (TI), increases specificity for kidney transcriptomics studies. For instance, targeted GLOM and TI eQTL studies have led to discoveries of the transcriptomic impact of diabetic kidney disease GWAS alleles in individuals with diabetic nephropathy.17-19 In an NS GLOM eQTL study of apolipoprotein L1 (APOL1), ubiquitin D (UBD) was significantly upregulated in individuals with a high-risk (HR) APOL1 genotype.20 These findings were subsequently followed up in an admixture mapping study that identified enriched African ancestry at the UBD locus in people with a HR genotype and FSGS versus no FSGS and protective effects of UBD expression on the viability of cells overexpressing APOL1 risk variants,21 providing additional support for UBD’s involvement in APOL1-attributed NS.

The success of previous tissue-specific eQTL studies in affected populations motivated us to use this strategy for NS genomic discovery. Here, we report the results of a genome-wide GLOM and TI cis-eQTL study from 187 biopsied NS participants enrolled in the Nephrotic Syndrome Study Network (NEPTUNE).22 In addition to this report, we have created “nephQTL,” a publicly available eQTL browser to share these data with the wider community (Web Resources).

Material and Methods

Data Sources and Participant Inclusion

The Nephrotic Syndrome Study Network (NEPTUNE) is a prospective, longitudinal cohort recruiting participants with substantial proteinuria at the time of first kidney biopsy for clinical suspicion of minimal change disease (MCD), focal segmental glomerulosclerosis (FSGS), or membranous nephropathy (MN).22 Phenotypic data, urine, and blood samples are collected at baseline and over time. DNA is collected for a variety of genotyping studies, and a research renal biopsy core is collected for transcriptomic analysis.20,22-24 All procedures were done in accordance with the ethical standards of the IRBs overseeing the NEPTUNE study, and proper informed consent was obtained from each participant.

Whole-Genome Sequencing

We used Illumina Hi-Seq to perform low-depth WGS on 322 NEPTUNE participants, allowing us to take advantage of shared haplotypes and make accurate genotype calls.25 We used GotCloud26 as our standard pipeline for alignment and variant calling (including insertion-deletions [indels]). A subset of participants also underwent Illumina Exome Chip SNP genotyping. Using Exome chip genotypes as our gold standard, we calculated their concordance with non-monomorphic sites on WGS at 77,769 shared sites. Among heterozygous sites present in both the exome chip and WGS with minor allele frequency (MAF) > 0.05 and MAF between 0.03 and 0.05, there was 98.7% and 97.8% concordance, respectively. Site-level genotype data from WGS and the Exome Chip, summarized across this cohort, are available at our NephVS browser (Web Resources).

Gene Expression

Gene expression data from microdissected GLOM and TI were generated using Affymetrix 2.1 ST chips27 and quantified with a Custom CDF file from BrainArray for EntrezG, v.19.28 Expression was normalized across genes using robust multi-array average (RMA).29 PEER factors were computed as previously described.20

Participant Inclusion

There were 187 NEPTUNE participants eligible for this study because they had WGS data, transcriptomic data for at least GLOM or TI, and clinical and demographic information. There were 136 and 166 available for GLOM and TI eQTL discovery, respectively. Of these, 115 participants were part of both the GLOM and TI eQTL analyses (Figure S1).

eQTL Analysis

We used MatrixEQTL24 for the initial step in the eQTL study. We focused solely on cis-eQTLs because we lacked substantial power for trans studies. Eligible variants were those with a MAF > 0.03 in our cohort and within 500 kb of qualifying genes’ transcription start or end sites (and within the gene itself). We adjusted for age, sex, principal components of genetic ancestry, and PEER factors. Principal components of genetic ancestry were calculated using EPACTS (Web Resources) on LD-pruned WGS data across all 187 individuals from whom we had expression information (Figure S2). We adjusted for the first four PCs based on visual inspection of the eigenvalue plot (Figure S3). PEER factors were created utilizing the PEER framework as previously described,30,32 adjusting for age, sex, and microarray batch. We adjusted for 31 PEER factors in GLOM and 25 in TI, the number which, at an empirical p value threshold (p < 1 × 10⁻⁴), maximized the number of eQTLs discovered on chromosome 21 for that tissue. We examined the correlation of our MatrixEQTL results with and without the inclusion of a metric of NS severity, choosing baseline estimated glomerular filtration rate (eGFR). To control the FDR at gene level from the MatrixEQTL output, we used TORUS, a computationally efficient method that uses an EM (expectation maximization) algorithm to compute the prior enrichment parameters for variants.33 The enrichment parameters computed from this procedure define a prior distribution per gene (Figure S4). The estimated priors from the GLOM and TI data can be summarized across all genes in relation to distance from the transcription start site (TSS).

Fine Mapping with DAP

We utilized the estimated priors from TORUS, accounting for variant distances to the TSS of the target genes, to perform variant
fine-mapping of eQTLs for each gene, using the Deterministic Approximation of Posteriors (DAP) algorithm.\textsuperscript{34} Rather than solely calling eGenes (genes with significant eQTLs), DAP identifies independent eQTLs contributing significantly to changes in expression and assigns a posterior inclusion probability (PIP) to them. DAP incorporates functional genomic annotations and accounts for patterns of linkage disequilibrium (LD) among single-nucleotide polymorphisms (SNPs). Compared to calling eGenes, implicating independent eQTLs using DAP requires higher statistical stringency.

Independent association signals were either single SNPs or indels (or groups of them in LD \([r^2 > 0.25]\)). The significance of each signal was characterized by local FDR. Those with local FDR < 0.05 were considered significant association signals in which we could predict the driving variants. Due to high LD and similar effect, the exact variant that drives the association signal was not always easily determined. DAP provides within-signal posterior inclusion probabilities (PIP) for each variant to help identify the most likely driver. The member SNPs for each signal group naturally form a 95% Bayesian credible set.

**Tissue-Specific Mapping with eQTLBMA**

To assess the extent of tissue specificity, we used eQTLBMA (eQTL by Bayesian Model Averaging) to estimate the proportion of eQTLs shared across GLOM and TI,\textsuperscript{35} using summary statistics generated by MatrixeQTL. Partial overlapping of samples is explicitly modeled in eQTLBMA. In comparing between GLOM and TI, genes were filtered prior to running eQTLBMA to make sure that they were expressed in both tissues. This filtering process ensured that the effect sizes for each SNP/transcript pair were reliably measured in both compartments. This framework also has the benefit of maximizing power over tissue-by-tissue analysis by jointly analyzing the tissues, allowing for differences in eQTLs by tissue, and improving power for detection of shared eQTLs.

**Gene Set Enrichment Analysis**

To better understand the biologic context of the gene sets emerging from our DAP and eQTLBMA analyses, we used the Genomatix Software Suite (Genomatix) and Ingenuity Pathway Analysis (IPA; QIAGEN)\textsuperscript{36} for functional annotation. In the IPA analysis, a complex network was constructed among the pathways that shared three or more genes to highlight “clusters” and pathways that crosstalk between the clusters.

**Generation of Single-Cell RNA-Seq Data**

We combined and analyzed three single-cell RNA-seq (scRNA-seq) datasets generated from healthy portions of tumor-nephrectomy samples specifically harvested for single-cell analysis using 10x Genomics methodology. Single-cell dissociations were done as reported.\textsuperscript{37} Individual cells were labeled with barcodes, and transcripts within each cell were tagged with distinct UMIs (unique molecular identifiers) in order to determine absolute transcript abundance. The library quality was assessed with “high sensitivity cDNA arrays” on an Agilent Bioanalyzer (Thermo Fischer) platform. Sequencing was done on Illumina HiSeq2500 with 2 × 75 paired-end kits using the following read lengths: 26 bp Read1 and 110 bp Read2.

The sequencing data were first analyzed using the Cell Ranger software (10x Genomics Inc.) in order to extract the gene expression data matrix. Each element of the matrix is the number of UMIs associated with a gene and a barcoded cell. Next, we filtered out cells with fewer than 500 genes per cell and with more than 25% mitochondrial read content. Further downstream analysis steps used the Seurat R package including normalization, identification of highly variable genes across the single cells, scaling based on the number of UMI and batch effect, dimensionality reduction (PCA and t-SNE), standard unsupervised clustering, and the discovery of differentially expressed cell type-specific markers. Differential gene expression to identify cell type-specific genes was performed using the non-parametric Wilcoxon rank sum test.

**Cell-Type Deconvolution using Kidney Single-Cell RNA-Seq**

We wanted to discover whether our significant GLOM and TI eQTLs were enriched in specific kidney cell types. We began by analyzing our adult kidney scRNA-seq data to identify genes whose expression were enriched in a particular cell type. We then computed the enrichment of GLOM and TI eGenes in these cell type-enriched gene sets using a Fisher’s exact test.

**Enrichment of eQTLs for Roadmap Fetal Kidney Annotations**

We downloaded the full dataset of epigenetic annotations of fetal kidney from the NIH Roadmap project (Web Resources) to examine whether our eQTL variants were enriched in these annotations. As input, we used all variants within all clusters identified by DAP, and we stored the variants’ PIP and their corresponding Roadmap fetal kidney annotation.

\[ \log(\text{PIP}) \sim \text{Annotation}, \]

In the above equation, “\( \text{PIP} \)” corresponds to the posterior inclusion probability of variant \( i \), and \( \text{Annotation} \), is the state of the ChromHMM model that intersects with variant \( i \). \( \text{Annotation} \), was coded as indicator variables for each state of the 15-state ChromHMM model, with state number 15 (“quiescent/low”) as the baseline state. The result of this is a logistic regression model with 1 intercept term and 14 indicator variables, where each variant was assigned to one annotation. Significance was determined using a Wald-test of each annotation level while controlling for all other levels.

**Comparison with GTEx eGenes and eQTLs**

We performed three analyses to compare DAP-derived eQTLs and eGenes found in GLOM and TI with those from GTEx v.6. First, we looked up the GTEx v.6 tissue with the smallest p value for each significant SNP-eGene pair found in GLOM and TI. We then compared the difference in the –log10 p values of GLOM and TI eQTLs with those from GTEx v.6. First, we performed three analyses to compare DAP-derived eQTLs and eGenes found in GLOM and TI with those from GTEx v.6. First, we looked up the GTEx v.6 tissue with the smallest p value for each significant SNP-eGene pair found in GLOM and TI. We then compared the difference in the –log10 p values of GLOM and TI SNP-eGene pairs with the minimum GTEx tissue. These analyses were limited to autosomes.

Second, to gain insight about eGenes unique to GLOM and TI as compared to other tissues, we used the following approach. If a significant GLOM or TI eGene was also present in the GTEx dataset and classified as an eGene in 1 of the 44 GTEx v6 tissues (per their previously described procedure\textsuperscript{38}), we labeled the corresponding tissue an “eGene-tissue.” We then summarized the distribution of the number of eGene-tissues across GTEx for all significant eGenes discovered in GLOM and TI. We also plotted GTEx liver (which has a similar number of eGenes identified as in TI) as a comparator.

Finally, to identify the magnitude of overlap between our eQTLs and those from whole blood in GTEx, we first selected GLOM and TI eQTLs with a p value < 1 × 10\textsuperscript{-5}. We chose the most significant
eSNP in this eQTL to take forward (a "SNP-eGene" pair). We then identified overlapping SNP-eGene pairs in whole blood from GTEx v6. Whole blood SNP-eGene pairs with \( p \) value \(< 1 \times 10^{-3} \) were empirically defined as nominally significant and classified as overlapping with GLOM or TI eQTLs.

**Comparison to Renal Cortex eGenes and eQTLs**

Ko et al. have previously reported kidney cortex eQTLs \(^{16}\) and published 1,886 eGenes and 124,612 SNP-eGene pairs that passed a statistical threshold of \( q \)-value \(< 0.05 \). We compared these with our GLOM and TI data in two separate analyses, recognizing that biological heterogeneity, technical and methodological differences, and availability of only a subset of cortex eQTL data must be taken into account when interpreting the data.

First, using the SNP-eGene pairs from cortex, we compared the magnitude and direction of effect of the same SNP-eGene pairs found in GLOM and TI. We then assessed the \( z \)-scores computed from cortex, GLOM, and TI via visualization and Pearson correlation. Second, we compared our significant eGene lists to those 1,886 significant cortical eGenes to identify eGenes uniquely found significant in GLOM and TI.

**Transcriptome-wide Association Study with IgA Nephropathy GWASs**

Previous studies have shown that risk alleles from GWASs are enriched for eQTLs, \(^{16}\) and others have used eQTL data to pinpoint genes whose expression was regulated by GWAS alleles. \(^{16}\) Here, we integrated our DAP-derived GLOM and TI eQTL data with summary statistics from the largest published GWAS of glomerular disease, in IgA Nephropathy \(^{40}\) (IgAN), and performed a transcriptome-wide association study (TWAS) using an approach adapted from the PrediXcan and MetaXcan methods. \(^{41,42}\) See Appendix A for details.

**Results**

Characteristics of the included participants are presented in Table 1. Approximately 30% of participants had childhood onset of disease and the median duration of disease prior to biopsy was 4 months, suggesting that we were not studying the transcriptome of participants with long-standing proteinuria. Based on genotype data, 53% and 29% of participants were of European (EUR) and African (AFR) continental ancestry, respectively. The remaining \( \approx 20\% \) of participants were closely divided between admixed (AMR) and Asian (ASN). The histologic diagnoses were almost equally divided between MCD, FSGS, MN, and other glomerular conditions (most often IgA nephropathy).
Using Matrix eQTL, we analyzed 76,979,158 cis-pairs across 9,114,417 SNPs and expression from 22,893 genes. Based on the minimum FDR values across all eQTLs, there were 1,055 eGenes in GLOM and 3,217 eGenes in TI, at FDR < 0.05. We found a strong correlation between models that included and did not include a measure of disease severity (in this case, baseline eGFR). We used FDR of the eQTL as our metric of comparison. The overall correlation was 0.99 in both GLOM and TI. For eQTLs in which at least one model resulted in an FDR < 0.05, the correlation was 0.83 in GLOM and 0.89 in TI. Given these results, we used a model without a metric of disease severity.

Multi-SNP and Multi-tissue Analyses

The output from MatrixeQTL plus the local FDR from TORUS (Figure S4) were “fine-mapped” using DAP. Rather than solely identifying that an eQTL exists, DAP identifies eQTLs in which the specific variants or “clusters” of variants predicted to be driving the association can be confidently identified. In addition, DAP can discover eQTLs in which multiple independent SNPs or clusters of SNPs are associated with the gene’s expression.

Using DAP, we discovered 894 GLOM eQTLs and 1,767 TI eQTLs at <5% FDR level (Table S1). The majority of eQTLs had one independent signal responsible for the association. Multiple independent signals associated with expression were found in 112 GLOM eQTL (14%) and in 337 in TI (19%) (Figure S5).

To identify eQTLs that were specific to, or shared between, the GLOM and TI, we used eQTLBMA, with the MatrixeQTL data as input. We estimated 12.2% and 26.3% of eQTLs were GLOM specific and TI specific, respectively (Figure 1; Table S2). Note that this tissue specificity estimate of eQTLs is not a simple tally of individual eQTL signals. It is obtained by pooling all genes across the two tissues simultaneously, which takes the power difference in each tissue into account.

We performed gene set enrichment analysis of the DAP-derived significant GLOM and TI eQTLs using Genomatix and IPA software. Analyzing the IPA networks showed a number of immunity pathways in GLOM (Figure S6A). The TI pathways were also enriched for immunity; however, metabolic and oxidative pathways were equally as enriched (Figure S6B).

A striking aspect of the pathway enrichment analyses was the shift from inflammatory/immune-related pathways and processes in the DAP GLOM gene set to those more specific to podocyte biology with the eQTLBMA GLOM gene set. For example, eight of the top ten GO (gene ontology) biologic processes for the DAP GLOM gene set were related to antigen presentation or interferon...
gamma signaling. By contrast, the top ten GO biologic processes for GLOM-specific eQTLs via eQTLBMA included actin cytoskeleton rearrangement, calcium signaling, and cell and biologic adhesion (Table 2). A StringDB network created from the eQTLBMA GLOM-specific genes showed a network significantly enriched for interactions, with the two most interconnected genes being transforming growth factor beta 1 (TGFBI) and integrin alpha V (ITGAV) (Figure S7). TGFBI has been implicated in renal disease pathways, including pathology of NS glomerulonephritis.

Among the most significant GLOM-specific eQTLs as computed by eQTLBMA were those of notable relevance to NS, particularly with phospholipase C gamma 2 (PLCG2) and vacuolar protein sorting 33b (VPS33B). PLCG2 has been implicated in pediatric SSNS via genome-wide rare-variant association study. VPS33B is a vacuolar protein and functions in vesicle-mediated protein sorting, predominantly in the late endosome/lysosome.

Importantly, mutations in VPS33B cause arthrogryposis, renal dysfunction, and cholestasis (ARC) syndrome, in which the renal phenotype includes NS.

Single-cell transcriptome analyses of 4,734 cells identified 14 clusters of specific cell types as defined by differentially expressed cell type-specific genes (Figures 2A and S8; Table S3). The number of cells per cluster ranged from 49 in the podocyte to 1,712 within the proximal tubule. 772 and 973 genes were significantly differentially expressed in the podocyte and proximal tubule cell clusters, respectively.

As shown in Figure 2B, GLOM and TI eGenes were most significantly enriched in podocyte (OR: 2.5, \(p = 8 \times 10^{-11}\)) and proximal tubule cells (OR: 3.4, \(p = 1 \times 10^{-43}\)), respectively. We note that the GLOM eQTLs are second most enriched in proximal cells. We believe that this is a function of the microdissection process, in which we are unable to remove all proximal tubular cells from the glomeruli.

Assignations of GLOM and TI eQTLs to specific cell type clusters is provided in Table S4.

We used Genomatix to annotate the 79 podocyte-specific GLOM eQTLs. The top three biologic processes (all \(p < 7 \times 10^{-07}\)) were “vesicle-mediated transport,” “endocytosis,” and “regulation of locomotion,” the top molecular function was “extracellular matrix binding” (\(p < 7 \times 10^{-06}\)), and the top signal transduction pathway was “integrin signaling” (\(p < 5 \times 10^{-05}\)).

### Enrichment of eQTLs for Roadmap Fetal Kidney Annotations

Both the glomerular and the tubulointerstitial eQTLs were significantly enriched in active regulatory regions in fetal kidney. In GLOM, the most significant association was annotated as “Active TSS” (\(PGLOM = 5.7 \times 10^{-23}\)) and the second as “Enhancer” (Figure 3). The directions of effect for active regulatory regions were positively associated with higher variant PIP as compared to quiescent/low regions.

### Comparison with GTex eGenes and eQTLs

We used three different strategies to assess overlap between our DAP-derived eQTLs and those in GTEx. First, to identify GLOM and TI eQTLs whose magnitude of effect was significantly greater than in any GTEx tissue, we selected NEPTUNE eQTLs with \(p < 1 \times 10^{-5}\) (most significant per gene) and then we queried the GTEx v6 multi-tissue eQTL dataset for the same variant and gene pair (Figure S9). 53% and 66% of GLOM and TI eQTLs, respectively, had a \(p < 0.001\) among at least one of any GTEx tissues. 38% and 44% of GLOM and TI eQTLs, respectively, had a more significant \(p\) value in a GTEx tissue. 10% and 14% of GLOM and TI eQTLs, respectively, had a \(p\) value five orders of magnitude more significant than that found in any GTEx tissue. Given that this analysis is in part derived from a set of significant GLOM and TI eQTLs emerging from a discovery study using a 5% FDR threshold, the comparisons reported here may be interpreted with this degree of uncertainty.

Second, we observed that the majority of eGenes discovered in GLOM and TI are replicated in multiple GTEx tissues, but a substantial proportion can be confidently

---

**Table 2. Top GO Biologic Processes for GLOM-Specific eQTLs as Defined by eQTLBMA**

<table>
<thead>
<tr>
<th>Biological Process (GO Term)</th>
<th>Adjusted p Value*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cell adhesion</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Biological adhesion</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Receptor metabolic process</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Positive regulation of cytosolic calcium ion concentration</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Regulation of protein kinase activity</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Calcium ion transmembrane transport</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Inorganic ion homeostasis</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Actin cytoskeleton reorganization</td>
<td>.001</td>
</tr>
<tr>
<td>Cellular calcium ion homeostasis</td>
<td>.001</td>
</tr>
<tr>
<td>Calcium ion homeostasis</td>
<td>.001</td>
</tr>
</tbody>
</table>

*Adjusted \(p\) value estimated from results of 1,000 simulated null hypothesis queries (GeneRanker, Genomatix).
identified only in our data and a few other GTEx tissues (Figure S10; Table S5). Namely, 105 GLOM eGenes (12%) were not significant in any GTEx tissue; this number was 169 (9.9%) for TI eGenes. When comparing this to GTEx liver, a tissue with a similar number of eQTLs as compared to TI, we note that more liver eGenes are common to multiple other tissues.

Finally, we made a close comparison of our SNP-eGene pairs \( p < 1 \times 10^{-5} \) with the same pairs in GTEx whole blood. We discovered that 26% of GLOM eQTLs with \( p < 1 \times 10^{-5} \) also had a \( p < 1 \times 10^{-3} \) in whole blood. In TI, 33% of eQTLs with \( p < 1 \times 10^{-5} \) also had a \( p < 1 \times 10^{-3} \) in whole blood.

Comparison with Renal Cortex eQTLs
Among the 124,612 significant SNP-eGene pairs identified from kidney cortex by Ko et al. and shared in their supplement, \(^{16} 111,739 \) were also found in our data. For this common set of gene-SNP pairs, we compared the z-scores computed from GLOM and TI to that of cortex and found strong correlations. The Pearson correlations for GLOM-cortex and TI-cortex were 0.69 (\( p < 2 \times 10^{-16} \)) and 0.74 (\( p < 2 \times 10^{-16} \)), respectively. The scatterplots show that the vast majority of the cortex eSNPs exhibit consistent directional effect (i.e., positive or negative) in our GLOM and TI data (Figure S11). At the same time, the plots also indicate considerable effect size heterogeneity between different tissues. We also identified 655 GLOM eGenes (73%) and 1,199 TI eGenes (68%) that were not significant eGenes in cortex (Table S1, column G). These differences likely reflect a combination of tissue heterogeneity between GLOM, TI, and cortex, differences in sample sizes, and the lack of knowledge of the magnitude and direction of effect of cortex eGenes that did not pass significance thresholds.

Kidney Tissue-Based TWAS for IgAN
Using a strategy based on the MetaXcan\(^{42} \) approach with minor modifications to accommodate microarray data, we predicted gene expression with a cross-validated \( r^2 > 0.01 \) (correlation > 0.1) in 3,294 genes from GLOM and 3,869 genes from TI. 39% of GLOM models and 42% of TI models had a cross-validated correlation > 0.2 (Table S6). Among prediction models with \( r^2 > 0.01 \), 73% of models in the GLOM and TI had at least ten variants selected in the prediction models. The number of variants selected in the model did not correspond to the number of independent eQTLs because the elastic net had a grouping effect where correlated variables tend to be selected or not as a group.\(^{53} \) This results in linked variants being selected in the prediction model with the effect size distributed across linked variants. We discovered 13 GLOM and 12 TI eQTLs associated with IgAN with FDR < 0.1. In both tissues, decreased expression of HLA-DRB5 had the strongest associated with IgAN (FDR\(_{\text{GLOM}} = 8.3 \times 10^{-6} \), FDR\(_{\text{TI}} = 3.3 \times 10^{-5} \) (Table 3; Figure S12).

Characterizing Known Kidney GWAS Alleles
We hypothesized that GLOM and TI eQTLs could be used to enhance interpretations of known GWAS alleles beyond those for NS or other glomerular diseases. For example, in one of the first GWASs of CKD, the lead SNP was rs12917707, 3.6 kb upstream of uromodulin (UMOD),\(^{54} \) a finding that has been replicated numerous times and has led to an entire field of study regarding UMOD in CKD. rs12917707 is not an eQTL across 44 tissues in GTEx. By contrast, it is a significant TI eQTL in our data, with the risk allele associated with increased expression of UMOD (Figure S13). Our TI results are concordant with reports that the risk allele is associated with increased UMOD transcript expression in tumor nephrectomy tissue\(^{55} \) and increased urinary protein UMOD expression.\(^{56} \)

Discussion
Biologic insights from eQTL studies become particularly powerful when using the tissues or cells directly involved in the disease, such as adipose and muscle eQTLs for type 2 diabetes (T2D),\(^{57} \) leukocytes in diverse immune-mediated diseases,\(^{58,59} \) adipocytes in obesity,\(^{60} \) or brain in schizophrenia.\(^{61} \) Here, we used a cis-eQTL study design to discover genetic variants associated with steady-state mRNA levels of genes expressed in the glomeruli and
tubulointerstitium of nephrotic syndrome individuals. Through the use of eQTLBMA and the integration of scRNA-seq data, we gained greater specificity with regards to compartment- and cell-specificity of the eQTLs we discovered. Through the integration of our results with fetal kidney tissue from Roadmap and eQTL results from GTEx, we gained additional insights into the generalizability of our results to kidney tissues without NS and human tissues other than kidney.

At a genome-wide level of significance, we discovered 894 GLOM eQTLs and 1,767 TI eQTLs in 136 and 166 eligible participants, respectively. By using multi-SNP analysis, we could fine-map these eQTLs, identifying specific SNPs with the highest probability of driving the signal, and also discovering a subset of GLOM and TI eQTL whose association is most likely to be driven by multiple, independent loci. Moving forward, narrowing the sets of implicated SNPs to identify the one that is causal, and understanding the genome biology underlying its impact on regulation will be important as we seek to discover targets and mechanisms upon which we can intervene therapeutically. Given the tissue- and cell-specific nature of genomic regulatory elements, the publication and sharing of epigenetic data derived from kidney and GLOM and TI cells will be vital.

Despite biologic and technical heterogeneity between this study and that which discovered eQTLs from the renal cortex of individuals without NS, we found a strong correlation in the direction of effect and their magnitude, albeit with substantial differences in effect sizes. While the majority of GLOM and TI eGenes were not reported as significant in the cortex, we posit that these differences are driven in part by cortex eGenes that displayed the same behavior but were not significant and thus unreported. As our field begins to generate kidney-derived eQTL datasets, it should become increasingly possible to use established joint and meta-analytic methods that explicitly consider heterogeneity to compare eQTLs between studies. A rigorous approach such as this could help to identify eQTLs that are compartment specific (or cell specific) as well as those that differ in disease versus healthy tissue. Making complete eQTL data available from future studies could facilitate this type of joint or meta-analysis.

When functionally annotating the DAP-significant GLOM and TI eQTLs, we observed an enrichment of immunity functions and pathways in both tissues, a not unexpected result given the known role of these pathways in NS. The cellular origins of these eQTLs are unclear, as podocytes can express these genes65 and there are both resident immune cells64 and circulating immune cells in the kidney. Future studies that compare eQTLs in circulating immune cells to those in the kidneys from the same individual may inform us of the contribution of each in individuals with NS. If we can ultimately link these eQTLs to clinical phenotypes, this may provide an opportunity to target the mechanism or the gene’s expression level for therapeutics.

The use of eQTLBMA, integration of scRNA-seq data, and comparison to fetal kidney tissue annotations and GTEx multi-tissue eQTL provided additional insights into the inherent discoveries in this study as well as their broader applicability. The GLOM- and TI-specific eQTLs identified by eQTLBMA were enriched for processes related to the normal physiology of these compartments rather than the immune-related functions shared across tissues. Using the scRNA-seq data, we found that GLOM eQTLs were most enriched in podocytes, TI eQTLs were most enriched in the proximal tubule, and the majority of eQTLs could be assigned to a cell type. In analyzing our eQTLs in the context of fetal kidney annotations, we discovered that GLOM and TI eQTLs from NS kidney were significantly enriched within active regulatory regions annotated from fetal kidney, most strongly at active transcription start sites and enhancers. Finally, in comparing our eQTLs with those found in GTEx, we found, as expected, substantial overlap of the eQTLs between datasets. However, there were also noticeable differences, with a set of 10% and 14% of eQTLs...
seemingly specific to GLOM or TI, respectively. The degree
to which this heterogeneity is attributed to differences in
technologies or sample size is unclear. Nevertheless, this
provides an eQTL set that can be further examined for sup-
port of biologic differences driving this heterogeneity.
Together, these findings should provide confidence that
the data and discoveries emerging from this study may
be useful beyond NS discovery.

Through a TWAS of IgAN and a single SNP lookup of the
lead SNP at the CKD-associated UMOD locus, we dem-
stated the utility of these eQTL data to gain further in-
sights from existing GWASs of kidney disease. We focused
on IgAN because, among glomerular disorders, it has the
largest GWAS readily available. The ancestral composition
of NEPTUNE participants (63% European or East Asian)
differed from that of IgAN GWAS (exclusively Europeans
and East Asians). Previous work has shown that most
eQTL effect sizes are consistent across populations, and
this is the assumption we made in this TWAS. Substantial
heterogeneity across ancestries in this particular analysis
would diminish power. However, this would not increase
type 1 error.

In this TWAS, we identified a limited number of genes
in which change in genetically predicted expression was
associated with IgAN status. The most significant TWAS
findings for IgAN that were detected independently in both
tissue compartments involved decreased mRNA expression
of HLA-DRB5 (p = 8.29 × 10⁻⁶ in GLOM and
p = 3.29 × 10⁻⁹ in TI compartments) with concur-
rent increased expression of HLA-DPB1 (p = 2.24 × 10⁻⁵ in
GLOM and p = 1.19 × 10⁻³ in TI compartments)
and HLA-DPA1 (p = 4.82 × 10⁻⁴ in GLOM and
p = 3.59 × 10⁻³ in TI compartments). These genes reside
under the strongest GWAS association peak for IgAN on
chromosome 6p21, but presently very little is known about
the specific role of these genes in the disease
pathogenesis. HLA-DPA1 and -DPB1 are expressed in
antigen-presenting cells and encode alpha and beta
chains of HLA-DP molecule, a membrane-anchored
heterodimer responsible for presenting peptides derived
from extracellular proteins. HLA-DRB5 represents a pa-
ralog of HLA-DRB1, one of the most polymorphic class
II molecules. The expression of DRB5 is linked with allelic
variants of DRB1; therefore, it is possible that this associ-
ation marks the presence of specific IgAN risk alleles of
the HLA-DRB1 gene. Because these genes have not yet
been experimentally tested in IgAN, further functional
studies will be required to explore potential causality of
these associations. Similarly, not much is presently
known about the role of the other top TWAS genes
(Table 3) in IgAN.

We appreciate that the causal cells for IgAN may also
reside outside the kidney or that multiple hits across cell
types are needed for disease pathogenesis. Future studies
that perform TWASs for IgAN in other cell types and tissues
will help us to clarify this issue, perhaps identifying causal
genes that act in a tissue-specific manner, as has been done
for other conditions. Independent of discovering
causal, tissue-specific genes, another potentially fruitful
strategy could be to study the set of intrarenal genes (and
their resultant networks) jointly implicated via TWASs.

We also recognize that the detected associations based
on cases with NS can be quite different from the associa-
tions derived from normal kidney tissue. We view this as
a potential strength of our eQTL mapping approach, since
some genetic effects may be detectable only in the context
of disease. At the same time, we recognize that using dis-
ease context-specific eQTLs can create an ascertainment
bias in TWASs. As larger GWAS datasets for other primary
glomerular disorders, such as FSGS, MN, or MCD, become
available, the TWAS approach should become even more
relevant.

<table>
<thead>
<tr>
<th>Table 3. Top GLOM and TI Genes Associated with IgA Nephropathy Case Status via Transcriptome-wide Association Study</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Gene</strong></td>
</tr>
<tr>
<td><strong>Glomerulus</strong></td>
</tr>
<tr>
<td>HLA-DRB5</td>
</tr>
<tr>
<td>RNFS</td>
</tr>
<tr>
<td>HLA-DPB1</td>
</tr>
<tr>
<td>HLA-DQA1</td>
</tr>
<tr>
<td>HLA-DPA1</td>
</tr>
<tr>
<td>HLA-DQA2</td>
</tr>
<tr>
<td>HLA-DQB1</td>
</tr>
<tr>
<td>HLA-DQB2</td>
</tr>
<tr>
<td>BTN3A2</td>
</tr>
<tr>
<td>ZKSCAN8</td>
</tr>
<tr>
<td>HLA-DRB1</td>
</tr>
<tr>
<td>CRYBA2</td>
</tr>
<tr>
<td>C2orf74</td>
</tr>
<tr>
<td><strong>Tubulointerstitium</strong></td>
</tr>
<tr>
<td>HLA-DRB5</td>
</tr>
<tr>
<td>HLA-DQB2</td>
</tr>
<tr>
<td>HLA-DQB1</td>
</tr>
<tr>
<td>HLA-DRB1</td>
</tr>
<tr>
<td>HLA-DPB1</td>
</tr>
<tr>
<td>ATF6B</td>
</tr>
<tr>
<td>HLA-DPA1</td>
</tr>
<tr>
<td>VWA7</td>
</tr>
<tr>
<td>SERPINA4</td>
</tr>
<tr>
<td>LAMB4</td>
</tr>
<tr>
<td>HIST1H4L</td>
</tr>
<tr>
<td>C2orf74</td>
</tr>
</tbody>
</table>

*FDR < 0.1
*Associated with IgAN cases
In interpreting this study overall and trying to place it in the context of similarly designed eQTL studies, there are a couple of unique characteristics to explicitly point out. First, the inclusion criteria for participants enrolled in NEPTUNE was a need for a kidney biopsy for suspicion of primary NS. There was no limitation to self-reported race, specific histologic diagnoses, eGFR, or response to immunosuppression. As such, this study should be interpreted in this light; namely that we are identifying eQTLs that are observed in the kidneys of individuals of diverse races with proteinuric glomerular filtration barrier failure.

Second, the recruitment of participants of diverse ancestries created a risk of population stratification confounding the results of this study. By following community standards, we worked to mitigate this risk through the inclusion of four PCs of genetic ancestry in the matrix eQTL model. In addition, the inclusion of PEER factors in the model should also help to account for gene expression variation attributable to population differences. Nonetheless, the transethnic design of this study should be taken into account when interpreting results. Our understanding of the potential impact of population stratification will be aided by future kidney eQTL studies from diverse ancestral populations.

With regards to heterogeneity in the causes of proteinuric kidney disease in this cohort, we found computational and experimental strategies that can be applied to GLOM and TI eQTL data to derive insights that are broadly relevant to underlying biology of these structures and specific to particular cell types. Future GLOM and TI eQTL studies in normal kidneys and those with a specific histologic or molecular diagnosis (and of similar sample sizes) would complement this study with insights about similarities and differences in the genetics of gene expression as a function of disease state.

In our opinion, sharing these eQTL and single-cell data in an easily accessible manner is just as important as any of the specific discoveries that we report here. To this end, we have added a stand-alone eQTL browser “nephQTL,” to our existing NephVS software. nephQTL has a searchable browser of the summary-level MatrixEQTl and DAP output for GLOM and TI, with both summary statistics and visualizations of the eQTLs. The full MatrixEQTl output is also available for download and secondary use. For the single-cell RNA-seq data, we have also uploaded to our file sharing folder a single data matrix files with the read counts from all three tumor nephrectomy databases (see Web Resources for link). Our hope is that unrestricted access to this unique data will be useful to the wider community, catalyzing and accelerating discoveries that will ultimately lead to improved health for individuals with NS and beyond.

Appendix A: Transcriptome-wide Association Study

Using the same genotype and expression data from the eQTL analysis, we first adjusted each gene’s expression by age, sex, 4 genetic PCs, and 31 PEER factors in TI and 25 in GLOM. Using the residuals for each gene, we used the R package “glmnet” to fit a regression equation penalized using an elastic net with $\alpha = 0.5$, which is a mixture of an L1 and L2 penalty. For each gene we allowed non-ambiguous biallelic variants within 500 kb of the start and end positions of each gene that were also present in the IgAN GWAS and eQTL dataset (MAF > 0.03 in the eQTL dataset). To select an appropriate hyperparameter, we used 30-fold cross-validation and selected the parameter that maximized the prediction $R^2$ of the validation set. We selected genes with a cross-validated prediction $R^2 > 0.01$.

Using a reference panel comprising East Asian and European samples from the 1000 Genomes data ($n = 1,007$ samples), for each gene $g$, we computed the variance of SNP $l$ ($\hat{\sigma}_g^2$) and variance of predicted expression $\hat{o}_g^2$. The variance $\hat{o}_g^2$ was defined in the MetaXcan paper as:

$$\hat{o}_g^2 = w_g^T \hat{A} w_g,$$

where $w_g$ is a vector of weights for gene $g$ and $\hat{A}$ is a covariance matrix of the SNPs included for gene $g$ computed across the samples selected above.

Finally, we computed the $Z_g$ statistic for gene $g$ as:

$$Z_g = \sum_{l \in \text{Model}_g} w_l \hat{\sigma}_g \frac{\hat{\beta}_l}{\text{se}(\hat{\beta}_l)},$$

where $\text{Model}_g$ is the set of SNPs for gene $g$, $w_l$ is the weight learned for gene $g$ and SNP $l$ from the elastic net, $\hat{\beta}_l$ is the effect size from the GWAS result for SNP $l$, and $\text{se}(\hat{\beta}_l)$ is the standard error for the effect size. We performed gene-level association using only the genes for which we could predict expression with a cross-validated $r^2 > 0.01$.

Supplemental Data

Supplemental Data include 13 figures, consortium information, and 6 tables and can be found with this article online at https://doi.org/10.1016/j.ajhg.2018.07.004.

Consortia

Members of the Nephrotic Syndrome Study Network (NEPTUNE) are as follows: John Sedor, Katherine Dell, Marleen Schachere, Kevin Lemley, Lauren Whitted, Tarak Srivastava, Connie Haney, Christine Sethna, Kalliopi Grammatikopoulos, Gerald Appel, Michael Toledos, Laurence Greenbaum, Chia-shia Wang, Brian Lee, Sharon Adler, Cynthia Nast, Janine LaPage, Ambarish Athavale, Alice Neu, Sara Boynton, Fernando Fervenza, Marie Hogan, John C. Lieske, Vladimir Cherniiskiy, Frederick Kaskell, Neelja Kuchar, Patricia Flynn, Jeffrey Kopp, Eveleen Castro-Rubio, Jodi Blake, Howard Trachtman, Olga Zhidanova, Frank Modersitzki, Suzanne Vento, Richard Lafayette, Kishma Mehta, Crystal Gadebeku, Duncan Johnstone, Daniel Catrann, Michelle Hladunewich, Heather Reich, Paul Ling, Martin Romano, Alessia Fornoni, Laura Barisoni, Carlos Bidot, Matthias Kretzer, Debbie Gipson, Amanda Williams, Renee Pitter, Patrick Nachman, Keisha Gibson, Sandra Grubbs, Anne Froment, Lawrence Holzman, Kevin Meyers, Krishna Kallem, Fumei Cerecino, Kamal Sambandam, Elizabeth Grubbs, Anne Froment, Lawrence Holzman, Kevin Meyers, Krishna Kallem, Fumei Cerecino, Kamal Sambandam, Elizabeth Grubbs, Anne Froment, Lawrence Holzman, Kevin Meyers, Krishna Kallem, Fumei Cerecino, Kamal Sambandam, Elizabeth Grubbs, Anne Froment, Lawrence Holzman, Kevin Meyers, Krishna Kallem, Fumei Cerecino, Kamal Sambandam, Elizabeth Grubbs, Anne Froment, Lawrence Holzman, Kevin Meyers, Krishna Kallem, Fumei Cerecino, Kamal Sambandam, Elizabeth...
Brown, Natalie Johnson, Ashley Jefferson, Sangeeta Hingorani, Kathleen Tuttle, Laura Curtis, S Dismuke, Ann Cooper, Barry Freedman, Jen Jar Lin, Stefanie Gray, Matthias Kretzler, Larra Barisoni, Crystal Gadegebeku, Brenda Gillespie, Debbie Gipson, Lawrence Holzman, Laura Mariani, Matthew G. Sampson, Peter Song, Johnathan Troost, Jarcy Zee, Emily Herreshoff, Colleen Kincaid, Chrysta Liczewska, Tina Mainieri, Amanda Williams, Kevin Abbott, Cindy Roy, Tiina Urv, and John Brooks

Acknowledgments

M.G.S. is supported by the Charles Woodson Clinical Research Fund, the Ravitz Foundation, and National Institutes of Health RO1-DK108805. K.K. is supported by National Institutes of Health RO1-DK105124. The Nephrotic Syndrome Study Network Consortium (NEPTUNE), US4-DK-083912, is a part of the National Center for Advancing Translational Sciences (NCATS) Rare Disease Clinical Research Network (RDCRN), supported through a collaboration between the Office of Rare Diseases Research (ORDR), NCATS, and the National Institute of Diabetes, Digestive, and Kidney Diseases. RDCRN is an initiative of ORD, NCATS. Additional funding and/or programmatic support for this project has also been provided by the University of Michigan, NephCure Kidney International, and the Halpin Foundation.

Declaration of Interest

The authors declare no competing interests.

Received: March 14, 2018
Accepted: June 29, 2018
Published: July 26, 2018

Web Resources

EPACTS, https://genome.sph.umich.edu/wiki/EPACTS
File sharing folder to access all MatrixEQTL output and single cell data matrix, https://umich.app.box.com/s/6jrvgmdh0ppwqv griuvijiau0xcb4ohm
GTEx Portal, https://www.gtexportal.org/home/
NephQTL browser, http://nephQTL.org
Nephrotic syndrome variant server, http://nephvs.org

References


15. Battle, A., Brown, C.D., Engelhardt, B.E., Montgomery, S.B.; GTEx Consortium; Laboratory, Data Analysis &Coordinating Center (LDACC)—Analysis Working Group; Statistical Methods groups—Analysis Working Group; Enhancing GTEx (eGTEx) groups; NIH Common Fund; NIH/NCI; NIH/NHGRI; NIH/NIMH; NIH/NIDA; Biospecimen Collection Source Site—NDRI; Biospecimen Collection Source Site—RPCI; Biospecimen Core Resource—VARIS; Brain Bank Repository—University of Miami Brain Endowment Bank; Leidos Biomedical—Project Management; ELSI Study; Genome Browser Data Integration &Visualization—EBI; Genome Browser Data Integration &Visualization—UCSC Genomics Institute, University of California Santa Cruz; Lead analysts; Laboratory, Data Analysis &Coordinating Center (LDACC); NIH program management; Biospecimen collection; Pathology; and eQTL manuscript working group (2017). Genetic effects on gene expression across human tissues. Nature 550, 204–213.


