RESEARCH ARTICLE SUMMARY

HUMAN GENOMICS

The impact of sex on gene expression across human tissues

Meritxell Oliva*†, Manuel Muñoz-Aguirre†, Sarah Kim-Hellmuth†, Valentin Wucher, Ariel D. H. Gewirtz, Daniel J. Cotter, Princy Parsana, Silva Kasela, Brunilda Balliu, Ana Viñuela, Stephane E. Castel, Pejman Mohammadi, François Aguet, Yuxin Zou, Ekaterina A. Khramtsova, Andrew D. Skol, Diego Garrido-Martín, Ferran Reverter, Andrew Brown, Patrick Evans, Eric R. Gamazon, Anthony Payne, Rodrigo Bonazzola, Alvaro N. Barbeira, Andrew R. Hamel, Angel Martinez-Perez, José Manuel Soria, GTEx Consortium, Brandon L. Pierce, Matthew Stephens, Eleazar Eskin, Emmanouil T. Dermitzakis, Ayellet V. Segrè, Hae Kyung Im, Barbara E. Engelhardt, Kristin G. Ardlie, Stephen B. Montgomery, Alexis J. Battle, Tuuli Lappalainen, Roderic Guigó, Barbara E. Stranger*

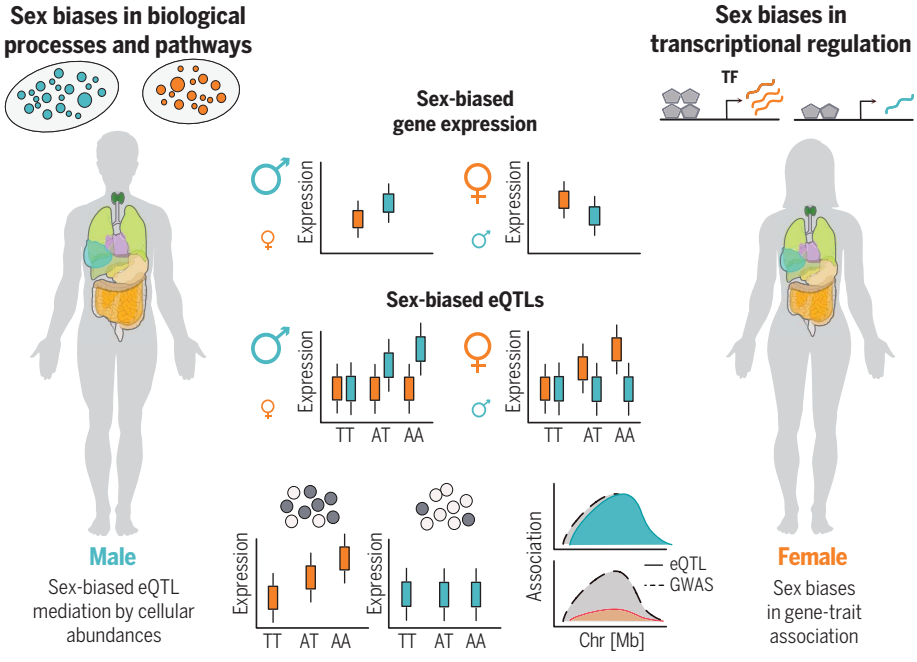
INTRODUCTION: Many complex human phenotypes, including diseases, exhibit sex-differentiated characteristics. These sex differences have been variously attributed to hormones, sex chromosomes, genotype × sex effects, differences in behavior, and differences in environmental exposures; however, their mechanisms and underlying biology remain largely unknown. The Genotype-Tissue Expression (GTEx) project provides an opportunity to investigate the prevalence and genetic mechanisms of sex differences in the human transcriptome by surveying many tissues that have not previously been characterized in this manner.

RATIONALE: To characterize sex differences in the human transcriptome and its regulation, and to discover how sex and genetics interact to influence complex traits and disease, we generated a catalog of sex differences in gene expression and its genetic regulation across 44 human tissue sources surveyed by the GTEx project (v8 data release), analyzing 16,245 RNA-sequencing samples and genotypes of 838 adult individuals. We report sex differences in gene expression levels, tissue cell type composition, and cis expression quantitative trait loci (cis-eQTLs). To assess their impact, we integrated these results with gene function, transcription factor binding annotation, and genome-wide association study (GWAS) summary statistics of 87 GWASs.

RESULTS: Sex effects on gene expression are ubiquitous (13,294 sex-biased genes across all tissues). However, these effects are small and largely tissue-specific. Genes with sex-differentiated expression are not primarily driven by tissue-specific gene expression and are involved in a diverse set of biological functions, such as drug and hormone response, embryonic development and tissue morphogenesis, fertilization, sexual reproduction and spermatogenesis, fat metabolism, cancer, and immune response. Whereas X-linked genes with higher expression in females suggest candidates for escape from X-chromosome inactivation, sex-biased expression of autosomal genes suggests hormone-related transcription factor regulation and a role for additional transcription factors, as well as sex-differentiated distribution of epigenetic marks, particularly histone H3 Lys^{27} trimethylation (H3K27me3).

Sex differences in the genetic regulation of gene expression are much less common (369 sex-biased eQTLs across all tissues) and are highly tissue-specific. We identified 58 gene-trait associations driven by genetic regulation of gene expression in a single sex. These include loci where sex-differentiated cell type abundances mediate genotype-phenotype associations, as well as loci where sex may play a more direct role in the underlying molecular mechanism of the association. For example, we identified a female-specific eQTL in liver for the hexokinase HKDC1 that influences glucose metabolism in pregnant females, which is subsequently reflected in the birth weight of the offspring.

CONCLUSION: By integrating sex-aware analyses of GTEx data with gene function and transcription factor binding annotations, we describe tissue-specific and tissue-shared drivers and mechanisms contributing to sex differences in the human transcriptome and eQTLs. We discovered multiple sex-differentiated genetic effects on gene expression that colocalize with complex trait genetic associations, thereby facilitating the mechanistic interpretation of GWAS signals. Because the causative tissue is unknown for many phenotypes, analysis of the diverse GTEx tissue collection can serve as a powerful resource for investigations into the basis of sex-biased traits. This work provides an extensive characterization of sex differences in the human transcriptome and its genetic regulation. ■



Sex affects gene expression and its genetic regulation across tissues. Sex effects on gene expression were measured in 44 GTEx human tissue sources and integrated with genotypes of 838 subjects. Sex-biased expression is present in numerous biological pathways and is associated to sex-differentiated transcriptional regulation. Sex-biased expression quantitative trait loci in cis (sex-biased eQTLs) are partially mediated by cellular abundances and reveal gene-trait associations. TT, AT, and AA are genotypes for a single-nucleotide polymorphism; TF, transcription factor.

The list of author affiliations and a full list of the GTEx authors and their affiliations are available in the full article online.
†These authors contributed equally to this work.
***Corresponding author. Email: meritxellop@uchicago.edu (M.O.); barbara.stranger@northwestern.edu (B.E.S.)**
Cite this article as M. Oliva *et al.*, *Science* 369, eaba3066 (2020). DOI: 10.1126/science.aba3066

READ THE FULL ARTICLE AT
https://doi.org/10.1126/science.aba3066

RESEARCH ARTICLE

HUMAN GENOMICS

The impact of sex on gene expression across human tissues

Meritxell Oliva[1,2,3]*†, Manuel Muñoz-Aguirre[4,5]†, Sarah Kim-Hellmuth[6,7,8]†‡, Valentin Wucher[4], Ariel D. H. Gewirtz[9], Daniel J. Cotter[10], Princy Parsana[11], Silva Kasela[7,8], Brunilda Balliu[12], Ana Viñuela[13], Stephane E. Castel[7,8], Pejman Mohammadi[14], François Aguet[15], Yuxin Zou[16], Ekaterina A. Khramtsova[1,17], Andrew D. Skol[1,2,18,19], Diego Garrido-Martín[4], Ferran Reverter[20], Andrew Brown[21], Patrick Evans[22], Eric R. Gamazon[22,23], Anthony Payne[24], Rodrigo Bonazzola[1], Alvaro N. Barbeira[1], Andrew R. Hamel[15,25], Angel Martinez-Perez[26], José Manuel Soria[26], GTEx Consortium§, Brandon L. Pierce[3], Matthew Stephens[16,27], Eleazar Eskin[28], Emmanouil T. Dermitzakis[13], Ayellet V. Segrè[15,25], Hae Kyung Im[1], Barbara E. Engelhardt[29,30], Kristin G. Ardlie[15], Stephen B. Montgomery[10,31], Alexis J. Battle[11,32], Tuuli Lappalainen[7,8], Roderic Guigó[4,33], Barbara E. Stranger[1,2,18,34]*

Many complex human phenotypes exhibit sex-differentiated characteristics. However, the molecular mechanisms underlying these differences remain largely unknown. We generated a catalog of sex differences in gene expression and in the genetic regulation of gene expression across 44 human tissue sources surveyed by the Genotype-Tissue Expression project (GTEx, v8 release). We demonstrate that sex influences gene expression levels and cellular composition of tissue samples across the human body. A total of 37% of all genes exhibit sex-biased expression in at least one tissue. We identify cis expression quantitative trait loci (eQTLs) with sex-differentiated effects and characterize their cellular origin. By integrating sex-biased eQTLs with genome-wide association study data, we identify 58 gene-trait associations that are driven by genetic regulation of gene expression in a single sex. These findings provide an extensive characterization of sex differences in the human transcriptome and its genetic regulation.

Many complex human phenotypes, such as anthropometric traits (e.g., waist-to-hip ratio), exhibit sex-differentiated distributions; disease features such as prevalence, progression, age of onset, and response to treatment often differ by sex (*1*–*5*). These sex differences have been variously attributed to hormones, sex chromosomes, genotype × sex effects, differences in behavior, and differences in environmental exposures (*6*), but the mechanisms and underlying biology of the sex differences remain largely unknown. The Genotype-Tissue Expression (GTEx) project (*7*) provides an opportunity to investigate the prevalence and genetic mechanisms of sex differences in transcriptomes and to identify how sex and genetics interact to influence complex traits and disease. The analyses presented here characterize sex differences in a relatively large population sample, including many tissues that generally lack characterization. Because the causative tissue is unknown for many diseases and disorders, analysis of this diverse tissue set can serve as a powerful resource for investigations into the basis of sex-differentiated phenotypes.

We present an extensive characterization of sex differences in the human transcriptome across 44 tissue sources of the GTEx project [v8 data release (*8*)] from 838 individuals (557 males, 281 females), constituting a large collection of multi-tissue bulk gene expression and genotype data (Fig. 1) (*9*). We quantify and characterize sex differences in gene expression levels (sex-biased gene expression) and cis sex-biased expression quantitative trait loci (sb-eQTLs). By incorporating the results of these sex-aware analyses of GTEx data with gene features and transcription factor binding annotation, we describe tissue-specific and tissue-nonspecific drivers and mechanisms contributing to sex differences in the human transcriptome and eQTLs. By integrating data from genome-wide association studies (GWASs), we report multiple sex-differentiated genetic effects on the transcriptome that colocalize with complex trait associations, highlighting the power of characterizing sex bias in GTEx samples for the mechanistic interpretation of GWAS signals.

Sex effects on gene expression are ubiquitous but small

Using GTEx v8 data (table S1), we quantified sex-biased gene expression in each of the 44 tissue sources for all genes expressed in at least one tissue. We considered a total of 35,431 X-linked and autosomal genes, including protein coding, long intergenic noncoding RNA (lincRNA), and other less-characterized gene types such as transcribed pseudogenes (*9*). For each tissue, we first fit a linear model that accounts for known sample and donor characteristics, as well as surrogate variables that capture hidden technical or biological factors of expression variability, including tissue cell type composition (fig. S1, A to C). Consequently, we are able to identify sex-biased gene expression that does not derive from sex differences in cell type abundances. We next modeled sex bias effects across tissues. We discovered a total of 13,294 differentially expressed genes [sex-biased genes; local false sign rate (LFSR) ≤ 0.05], with 473 to 4558 genes discovered per tissue, representing 1.3% to 12.9% of all tested genes, respectively (Fig. 2A, fig. S1, D to F, and table S2). Previous studies have reported widespread sex-biased gene expression (*10*–*12*) and described breast as the most sex-differentiated tissue (*10*, *11*, *13*). However, we did not observe this in the present study after controlling for sex differences in tissue cell type composition

[1]Section of Genetic Medicine, Department of Medicine, University of Chicago, Chicago, IL, USA. [2]Institute for Genomics and Systems Biology, University of Chicago, Chicago, IL, USA. [3]Department of Public Health Sciences, University of Chicago, Chicago, IL, USA. [4]Centre for Genomic Regulation, Barcelona Institute for Science and Technology, Barcelona, Catalonia, Spain. [5]Department of Statistics and Operations Research, Universitat Politècnica de Catalunya, Barcelona, Catalonia, Spain. [6]Statistical Genetics, Max Planck Institute of Psychiatry, Munich, Germany. [7]New York Genome Center, New York, NY, USA. [8]Department of Systems Biology, Columbia University, New York, NY, USA. [9]Lewis-Sigler Institute for Integrative Genomics, Princeton University, Princeton, NJ, USA. [10]Department of Genetics, Stanford University, Stanford, CA, USA. [11]Department of Computer Science, Johns Hopkins University, Baltimore, MD, USA. [12]Department of Computational Medicine, University of California, Los Angeles, CA, USA. [13]Department of Genetic Medicine and Development, University of Geneva Medical School, Geneva, Switzerland. [14]Department of Integrative Structural and Computational Biology, The Scripps Research Institute, Scripps Research Translational Institute, La Jolla, CA, USA. [15]Broad Institute of MIT and Harvard, Cambridge, MA, USA. [16]Department of Statistics, University of Chicago, Chicago, IL, USA. [17]Computational Sciences, Janssen Pharmaceuticals, Spring House, PA, USA. [18]Center for Translational Data Science, University of Chicago, Chicago, IL, USA. [19]Department of Pathology and Laboratory Medicine, Ann and Robert H. Lurie Children's Hospital of Chicago, Chicago, IL, USA. [20]Department of Genetics, Microbiology and Statistics, Faculty of Biology, University of Barcelona, Barcelona, Spain. [21]University of Dundee, Dundee, Scotland, UK. [22]Division of Genetic Medicine, Vanderbilt University Medical Center, Nashville, TN, USA. [23]Clare Hall, University of Cambridge, Cambridge, UK. [24]Wellcome Centre for Human Genetics, Nuffield Department of Medicine, University of Oxford, Oxford, UK. [25]Massachusetts Eye and Ear, Harvard Medical School, Boston, MA, USA. [26]Genomics of Complex Diseases Group, Research Institute Hospital de la Sant Creu i Sant Pau, IIB Sant Pau, Barcelona, Spain. [27]Department of Human Genetics, University of Chicago, Chicago, IL, USA. [28]Departments of Computational Medicine, Computer Science, and Human Genetics, University of California, Los Angeles, CA, USA. [29]Department of Computer Science, Center for Statistics and Machine Learning, Princeton University, Princeton, NJ, USA. [30]Genomics plc, Oxford, UK. [31]Department of Pathology, Stanford University, Stanford, CA, USA. [32]Department of Biomedical Engineering, Johns Hopkins University, Baltimore, MD, USA. [33]Universitat Pompeu Fabra, Barcelona, Catalonia, Spain. [34]Center for Genetic Medicine, Department of Pharmacology, Northwestern University, Chicago, IL, USA.
*Corresponding author. Email: meritxellop@uchicago.edu (M.O.); barbara.stranger@northwestern.edu (B.E.S.)
†These authors contributed equally to this work. ‡Present address: Department of Pediatrics, Dr. von Hauner Children's Hospital, University Hospital LMU Munich, Munich, Germany.
§A full list of the GTEx authors and their affiliations is available at the end of this article.

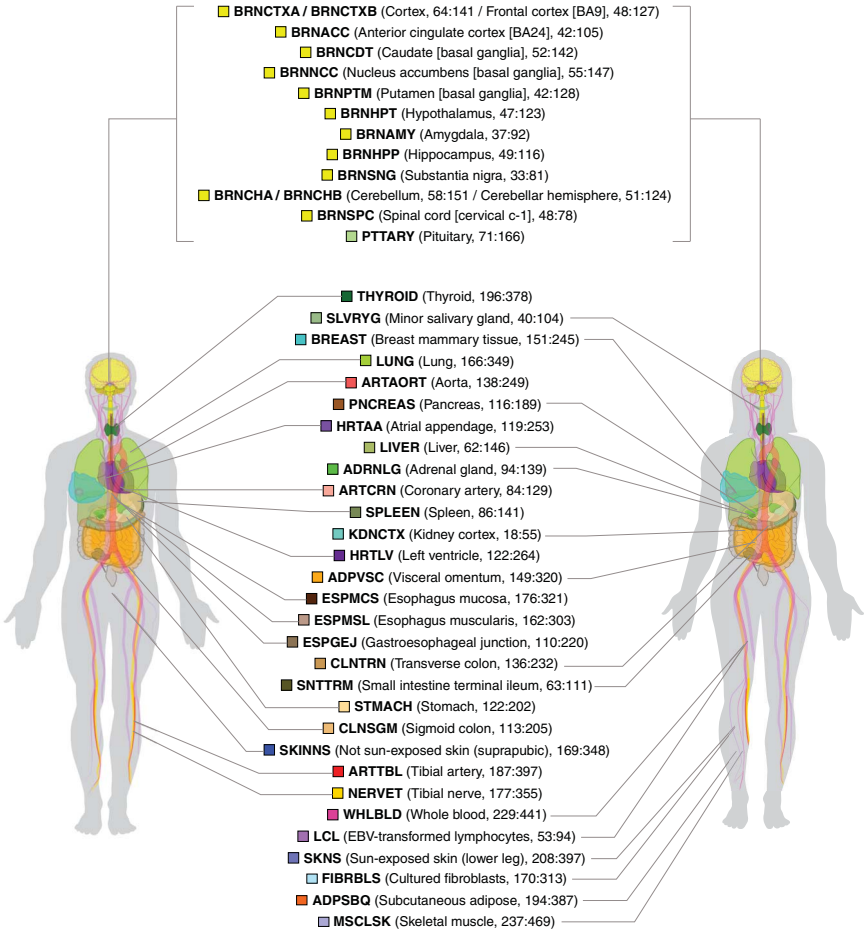


Fig. 1. Sample, data types, and discovery sets in the study of sex differences in GTEx v8. Tissue types (including 11 distinct brain regions and two cell lines) are illustrated, with sample numbers from GTEx v8 genotyped donors (females:males, in parentheses) and color coding indicated for each. This study included $N = 44$ tissue sources present in both sexes with ≥70 samples. Tissue sources comprised two cell lines, 40 tissues, and two additional replicates for brain cerebellum and cortex tissues. Tissue name abbreviations are shown in bold. See (*9*) for specific numbers of donors used in each analysis.

(fig. S1A). We next assessed replication of sex-biased genes in independent gene expression datasets for four tissues (brain cerebellum, brain cortex, heart left ventricle, and lymphocytes; table S2). We observed moderate to strong replication (average $\pi_1 = 0.62$, average effect size Spearman's $\rho = 0.78$). In total, 37.5% (13,294/35,431) of the human transcriptome was differentially expressed in at least one tissue. Of these, 531 genes (4%) were X-linked and 12,763 genes (96%) were autosomal, representing 47% and 37% of all tested X-linked and autosomal genes, respectively. Although abundant, sex effects were mostly small (fig. S2A), particularly for autosomal genes (*9*) (fig. S2B). X-linked genes with higher expression in females (female-biased genes) exhibited larger sex effects [median fold change (FC) = 1.13] than either X-linked genes with higher expression in males (male-biased genes; median FC = 1.08) or autosomal sex-biased genes (median FC_M and $FC_F = 1.04$; fig. S2B), potentially as a result of escape from X-chromosome inactivation (XCI) (*14*). The number of sex-biased genes and the effect sizes were not dominated by either sex (fig. S2C).

Sex-biased gene expression is largely tissue-specific

Sex-biased genes exhibited a skewed pattern of tissue sharing; they were likely to be differentially expressed in only a small subset of tissues (Fig. 2B), as previously reported (*10*–*13*). Of 13,294 total sex-biased genes, 2416 (18.2%) were differentially expressed in only a single tissue (Fig. 2B), suggesting tissue-dependent regulation. Only 30 genes (0.23%), 22 of which are known constitutive XCI escapees (table S3), exhibited consistent sex bias across all 44 tissue sources (Fig. 2B). This tissue specificity did not simply reflect patterns of gene expression across tissues; sex-biased genes tended to be ubiquitously expressed across tissues, whereas sex-biased expression was limited to one or a few tissues (*9*) (Fig. 2C and fig. S2D). The majority (8241/10,878 genes, 76%) of genes with sex bias in two or more tissues exhibited consistent effect direction across tissues, especially for X-linked genes (fig. S2E). Notably, whole blood and cell lines, the most widely studied biospecimen types, were not representative of sex-biased expression across tissues; sex-biased genes in whole blood constituted only 12.9% (1710/13,294) of all sex-biased genes. Although hierarchical clustering of tissues based on gene expression and on sex-biased expression is highly concordant (cophenetic correlation coefficient = 0.75) (*9*) (Fig. 2C and fig. S3, A to C), the intersection between the cluster-defining gene sets (table S4) is less than expected by chance ($P < 2.2 \times 10^{-16}$, hypergeometric test). For example, both gene expression and sex-biased expression supported a cluster of brain subregions that is clearly differentiated from other tissues (Fig. 2C and fig. S3, B and D). However, the cluster based on sex-biased expression was driven by 194 genes, whereas the transcriptome-based brain cluster was driven by 982 genes, from which only six were common with those defining the sex-based brain cluster. Among drivers of the sex-based liver cluster, we identified CYP450 genes—*CYP1A2*, *CYP3A7*, *CYP3A4*—as previously reported (*15*), but we also found genes less well characterized for sex bias, such as *PZP*, *H19*, and *VWCE*, which were previously shown to be sex-differentially expressed as a result of liver-specific sex differences in DNA methylation (*16*). These results suggest that the tissue specificity of sex-biased expression is not driven primarily by tissue-specific gene expression.

X-linked female-biased genes accurately predict sex and suggest tissue-specific candidates for escape from X-chromosome inactivation

We accurately predicted sex from gene expression, as previously explored (*17*), using X-linked genes (*9*) (fig. S4, A to D) with gradient boosted trees. Although the most predictive X-linked genes (fig. S4E) are those known to escape XCI, we identified 40 X-linked female-biased genes predictive of sex (within the top tertile with respect to their Shapley values) not previously described as XCI escapees (table S3). These results suggest further evaluation of these genes as potential XCI escapees; we did not directly test escape from XCI, and female-biased expression of X-linked genes may originate from other mechanisms. Sex prediction from autosomal genes was less accurate (mean accuracy = 84%), less specific (mean specificity = 56%, sensitivity = 96%; fig. S4D), and required more genes (fig. S4F) than prediction based on X-linked genes. However, in two tissues—breast and muscle—autosomal genes predicted sex with specificity ≥ 90% and sensitivity ≥ 98% (fig. S4G).

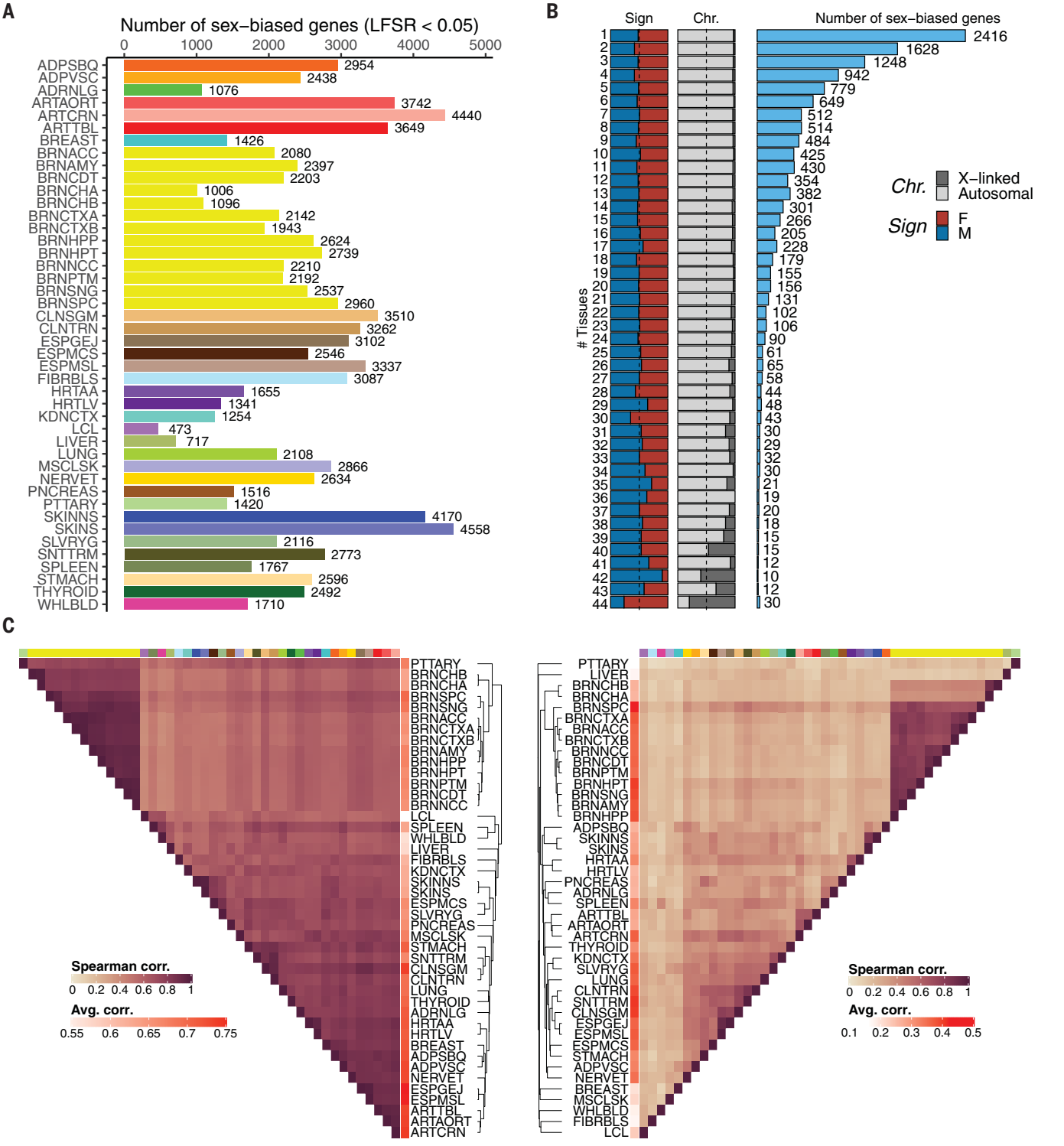


Fig. 2. Sex-differential gene expression. (**A**) Number of sex-differentially expressed genes (sex-biased genes) per tissue. Tissue colors are as in Fig. 1. (**B**) Sex-biased gene discovery (histogram, number of sex-biased genes) and characteristics of sex-biased genes (stacked bar plots) as a function of tissue sharing. Proportions of X-linked and autosomal sex-biased genes (Chr.) and of female- and male-biased genes (Sign) are indicated. (**C**) Hierarchical clustering of tissues based on gene expression (left) and the effect size of sex-biased genes (right). See (*9*) for further details.

Sex-biased genes exhibit nonrandom and tissue-specific genomic distribution

Except for the enrichment of female-biased genes on the X chromosome, little is known about the genome-wide distribution of sex-biased genes. We applied a positional gene enrichment analysis method (*18*) separately for male- and female-biased genes (LFSR ≤ 0.05) from each tissue (*9*) (fig. S5A). We discovered clustering of a total of 1559 sex-biased genes in 134 autosomal and five X-linked regions ($P \leq 0.001$, hypergeometric test) (Fig. 3A and table S5). On the X chromosome, pseudoautosomal region PAR1 and the remainder of the X-chromosome short arm *p* were enriched

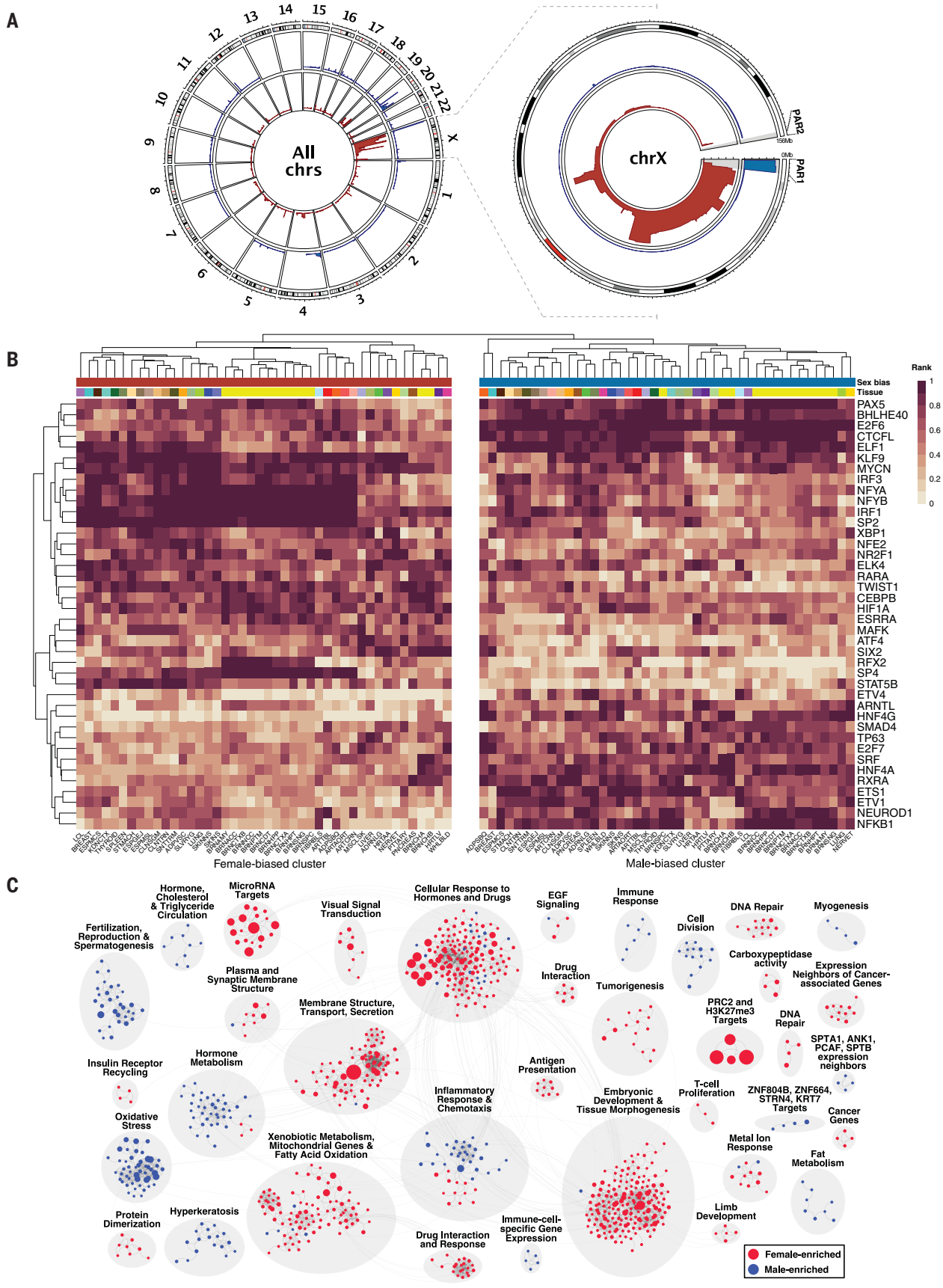

A
All chrs
chrX
PAR2
PAR1
B
Sex bias
Tissue
Rank
PAX5
BHLHE40
E2F6
CTCFL
ELF1
KLF9
MYCN
IRF3
NFYA
NFYB
IRF1
SP2
XBP1
NFE2
NR2F1
ELK4
RARA
TWIST1
CEBPB
HIF1A
ESRRA
MAFK
ATF4
SIX2
RFX2
SP4
STAT5B
ETV4
ARNTL
HNF4G
SMAD4
TP63
E2F7
SRF
HNF4A
RXRA
ETS1
ETV1
NEUROD1
NFKB1
Female-biased cluster
Male-biased cluster
C
Hormone, Cholesterol & Triglyceride Circulation
MicroRNA Targets
Visual Signal Transduction
Cellular Response to Hormones and Drugs
EGF Signaling
Immune Response
DNA Repair
Myogenesis
Fertilization, Reproduction & Spermatogenesis
Cell Division
Carboxypeptidase activity
Expression Neighbors of Cancer-associated Genes
Plasma and Synaptic Membrane Structure
Drug Interaction
Tumorigenesis
Membrane Structure, Transport, Secretion
PRC2 and H3K27me3 Targets
DNA Repair
SPTA1, ANK1, PCAF, SPTB expression neighbors
Insulin Receptor Recycling
Hormone Metabolism
Antigen Presentation
Inflammatory Response & Chemotaxis
Embryonic Development & Tissue Morphogenesis
T-cell Proliferation
ZNF804B, ZNF664, STRN4, KRT7 Targets
Cancer Genes
Oxidative Stress
Xenobiotic Metabolism, Mitochondrial Genes & Fatty Acid Oxidation
Metal Ion Response
Fat Metabolism
Protein Dimerization
Hyperkeratosis
Drug Interaction and Response
Immune-cell-specific Gene Expression
Limb Development
Female-enriched
Male-enriched

Fig. 3. Regulatory mechanisms and biological functions of sex-biased genes. (**A**) Genomic position enrichment of sex-biased genes, as indicated by male-biased (blue) and female-biased (red) genes across all chromosomes (left) and chromosome X (right). The height of each rim represents the tissue sharing of the significant genomic enrichment signal and ranges from 1 to 44 (number of tissue sources). See (*9*) for further details. (**B**) Transcription factor binding site (TFBS) enrichment in promoter regions of sex-biased genes. Of 92 enriched TFBS profiles, the top 40 with the largest difference across all tissues in the enrichment profile derived from male-biased and female-biased genes are displayed. Values represent the TFBS enrichment ranking transformed to [0, 1] per tissue and per sex; a value of 1 corresponds to the highest enrichment. See (*9*) for further details. (**C**) Clusters (gray circles) of gene sets enriched for genes highly expressed (blue and red balloons) in females (red) or males (blue) across tissues. Balloon size corresponds to the *P* value for the across-tissue meta-analysis of GSEA. Faint lines connecting balloons correspond to shared leading-edge genes between gene sets. See (*9*) for further details.

for male-biased and female-biased genes, respectively (Fig. 3A, right), as previously reported (*14*). Female-biased gene enrichment was stronger (Spearman's $\rho = 0.51$, $P = 1.63 \times 10^{-15}$) in the younger strata of arm *p* (fig. S5B), likely driven by escape from XCI (*14*, *19*). Although enriched X-chromosome regions spanned ~126 Mb, only 25% of subregions were enriched in at least two-thirds of the tissues. Among autosomal sex-biased genes, we observed a cluster of male-biased genes on chromosome 20 that was identified in 70% (30/44) of tissues (fig. S5C), but the majority of the 134 autosomal enriched regions were tissue-specific, identified on average in ~7% (3/44) of tissues (fig. S5D and table S5). These results are compatible with tissue-variable escape from XCI (*14*, *20*) and with tissue-specific topologically associating domains, possibly mediated by hormones (*21*). Further investigation is warranted to corroborate these and other hypotheses, as observed patterns may originate from a variety of mechanisms.

Promoters of sex-biased genes are enriched for hormone-related and other transcription factor binding sites

We hypothesized that transcription factor (TF) activity might drive observed patterns of differential expression, because sex-biased gene regulation by TFs has recently been reported (*13*) and TFs contribute to evolutionary changes in sex bias (*12*). We tested for enrichment of TF binding sites (TFBSs) of 231 TFs previously identified through chromatin immunoprecipitation sequencing (*22*) in promoter regions (i.e., 2 kb upstream of the transcription start site) of male- and female-biased genes (*9*) (fig. S5E). We discovered enrichment for TFBSs of a total of 92 TFs (fig. S5F), two of which were X-linked (*AR*, *ELK1*). TFBSs for 54 TFs were enriched among female-biased genes and 60 TFs among male-biased genes, with 22 TFs enriched among both sets of genes (table S6). The 92 TFs include (i) known hormone-related TFs estrogen (*ESR1*), androgen (*AR*), and glucocorticoid (*NR3C1*) receptors, (ii) 10 TFs that colocalize with steroid receptors, and (iii) TFs with a nonreported or less-characterized hormone association, including *SP1*, *E2F6*, *NRF1*, *KLF9*, and *SP2*, the top five TFs with consistent TFBS enrichment across tissues (*9*).

The strongest difference between male- and female-biased enrichment profiles was observed for TFBSs of *SP2*, *SP4*, *NFYB*, *TWIST1*, and *STAT5B* (female-biased) and of *HNF4G*, *NFKB1*, *E2F6*, *HNF4A*, and *ETS1* (male-biased), respectively, which were detected across most tissues (Fig. 3B and table S6). In contrast, we observed tissue specificity for enrichment of TFBSs of several TFs, such as *RFX2* and *ETV4* for brain and breast tissues, respectively (Fig. 3B and fig. S5F). Although *STAT5B* and *HNF4A* play known roles in sex differences in body growth rates and liver gene expression (*15*), less is known about their roles and sex biases across all tissues. The effect of sex on most of the remaining TFs is uncharacterized. Together, these results suggest that hormone-related TFs regulate sex-biased expression as expected, but they also indicate that additional TFs play a role in sex-biased gene expression, in some cases in a tissue-specific manner (table S6). Notably, TFBS enrichment is not driven by sex-biased expression of the TFs themselves (*9*), consistent with the observation that sex-biased TF targeting of genes is independent of sex-biased gene expression (*13*). However, this scenario cannot be discarded if such differences occur at an earlier developmental time point and translate into a more constitutive sex-biased TF binding profile (*23*). Alternatively, other mechanisms involving TFs could be causal drivers [e.g., posttranslational modifications as reported in mice (*24*)].

Sex-biased genes are involved in a highly diverse set of biological functions and suggest sex-specific deposition of epigenetic marks

To gain insight into cellular functions affected by sex-biased genes, we performed gene set enrichment analysis (GSEA) in each tissue, considering the direction of the sex effect (*9*) (fig. S6A and tables S7 and S8). To identify gene sets that are enriched across multiple tissues, we performed a meta-analysis using Fisher's combined probability test and identified 2134 enriched gene sets [false discovery rate (FDR) ≤ 0.05; table S9]. We applied a community detection approach to identify common features across enriched gene sets and defined 36 clusters (table S9). Among the top-scoring clusters (*9*), we identified enrichment of genes in pathways involved in drug and hormone response, epigenetic marks, embryonic development and tissue morphogenesis, fertilization, sexual reproduction and spermatogenesis, fat metabolism, cancer, immune response, and other functions (Fig. 3C and table S9). The top-scoring cluster corresponds to targets of polycomb repressive complex 2 (PRC2) and trimethylation of histone H3 at Lys^{27} (H3K27me3), which is predominantly driven by female-biased genes—a pattern also reported for other epigenetic modifications (*13*). This complex induces gene silencing and is involved in XCI (*25*). Sex-specific deposition of H3K27me3 marks has been previously reported, resulting in sex-biased gene expression in mammalian placenta (*26*) and adult liver (*27*). These differences have been hypothesized to be regulated by sex differences in the secretion of placental glycosyltransferase OGT and pituitary growth hormone. The observed association of H3K27me3 with sex-biased expression in the tissues of this study (table S9) has not been previously reported. We also identified clusters related to drug metabolism that include CYP450 genes. Sex-biased expression of CYP450 has been reported in liver (*15*) and linked to sex-differentiated growth hormone profiles; we observed sex-biased expression in additional tissues (fig. S6B). Sex-biased expression was also identified for clusters related to gonad tissue functions (e.g., meiotic synapsis), which comprise genes expressed largely in testis (fig. S6B). It is possible that some of the cross-tissue sex-biased expression patterns observed in adult tissues are derived from gamete formation and embryogenesis (*28*). Together, these results indicate that sex-biased genes are involved in a wide range of biological functions and pathways, many of which have not been previously associated with sex differences.

Sex and disease influence tissue cellular composition

The GTEx tissue samples are mixtures of heterogeneous cell types, with variation among individuals and tissues (*29*). In whole blood, cell type composition differs between sexes (*30*, *31*), but little is known about sex differences in composition of other tissues. Using a *t* test, we examined each GTEx tissue for sex differences in cellular composition on the basis of estimated abundances of seven cell types (*9*, *29*). We discovered significant (FDR ≤ 0.05) differences for four cell types—keratinocytes, neutrophils, adipocytes, and epithelial cells—in three tissues (fig. S7A and table S10). We hypothesize that additional cell types uncharacterized in this study may influence the cell type composition of GTEx tissues, particularly

of immune cells, because marked sex differences in immune cell abundances have been reported (*30*, *32*). To investigate cellular abundances in disease, we used histological annotations from pathology review of GTEx tissue samples (*9*). We discovered six pathological phenotypes with altered cell type composition (fig. S7, B to E, and table S10). Together, these results suggest that sex is correlated with tissue cellular composition, and that disease may alter cellular abundances in a sex-differentiated manner or in sex-specific pathologies.

Sex differences in the genetic regulation of gene expression are highly tissue-specific and less common than sex effects on gene expression

Sex-differentiated human phenotypes and disease characteristics may derive in part from sex-differentiated genetic effects (*6*, *33*–*36*), some of which may have an impact on gene expression. For each of 491,694 conditionally independent cis-eQTLs identified in the sex-combined cis-eQTL analysis of the GTEx v8 project (*8*), we performed sex-biased cis-eQTL (sb-eQTL) analysis in each of 44 tissues present in both sexes (Fig. 1). We used a linear regression model including genotype, sex, and covariates, and tested for significance of a genotype × sex (G×Sex) interaction on expression (*9*). Notably, this approach captures G×Sex interactions that derive both from sex and from sex-correlated factors, including cell type abundances or environmental factors. Although the contribution of cell type heterogeneity to sb-eQTLs is currently unknown, we observed sex differences in tissue cell type composition (fig. S7A), which may affect sb-eQTL discovery. Hence, we characterized the impact of cell type–specific eQTLs on sb-eQTLs (see below). We discovered a total of 369 sb-eQTLs, corresponding to 366 genes (sb-eGenes) (FDR ≤ 0.25; table S11). The majority of sb-eQTLs were identified in breast tissue (261 sb-eQTLs), but also in muscle (36 sb-eQTLs), skin (18 sb-eQTLs), and adipose tissues (14 sb-eQTLs) (Fig. 4A and fig. S8, A and B). Overall, sb-eQTLs showed strong evidence for tissue specificity (*9*); only one sb-eQTL was significant in two tissues (table S11), and only 21% displayed patterns suggestive of tissue-sharing even at a lenient significance threshold ($P_{G\times Sex} \leq 0.01$). Only 36 sb-eGenes (14%) exhibited sex-biased expression in the discovery tissue [multivariate adaptive shrinkage (MASH) LFSR ≤ 0.05; table S12], similar to recent observations (*37*). This is compatible with small sb-eQTL effects not translating into significant sex-biased gene expression, or with different functional mechanisms contributing to each sex bias type.

To provide additional support for the sb-eQTLs, we used two approaches to assess differential allele-specific expression (ASE) between sexes: allelic fold change (ASE aFC) (*38*) and environment ASE through generalized linear modeling (EAGLE) (*9*, *39*). Allele-specific expression can result from cis-regulatory genetic effects in heterozygous individuals. Differential ASE therefore indicates condition-specific cis effects (*39*), including sex specificity. We observed that both approaches, despite limited power when restricted to heterozygous individuals and differences in methodology, indicate that a portion of the detected sb-eQTLs correspond to sex differences in ASE (fig. S8C): sb-eQTLs were enriched for sex-biased ASE aFC (all tissues, $\pi_1 = 0.36$; breast, $\pi_1 = 0.41$; fig. S8, D and E) and for EAGLE associations ($\pi_1 = 0.13$, empirical test, $P \leq 0.001$). Of the 243 and 163 sb-eQTLs tested by ASE aFC and EAGLE methods, respectively, 65 (26.7%) were supported by ASE aFC (Wilcoxon $P \leq 0.05$) (fig. S7, F and G), 29 (17.8%) were supported by significant EAGLE associations, and 16 sb-eQTLs (10.4% of the 154 sb-eQTLs tested by both methods) were supported by both methods (table S11).

We were limited in our ability to replicate sb-eQTLs because the majority of sb-eQTLs were discovered in breast tissue, and matching well-powered datasets do not exist. We performed internal validation, splitting GTEx breast samples into discovery and validation cohorts, and observed moderate replication (mean $\pi_1 = 0.28$) (*9*) (fig. S8H). We next assessed sb-eQTL replication (considering sb-eQTLs from breast, whole blood, and all tissues) in independent larger (~900 subjects) whole-blood eQTL datasets, including DGN (*40*) and GAIT2 (*41*) cohorts (*9*) (table S13). We observed weak replication ($\pi_1 = 0$ to 0.12, depending on sb-eQTL set and replication cohort). Poor replication of sb-eQTLs has been reported (*40*, *42*, *43*) and has been, in part, attributed to low power (*44*) but also to methodological and study design differences.

For each sb-eGene, we also performed sex-stratified cis-eQTL analysis for each tissue, downsampling males to match the female sample size (*9*). We observed strong correlation (Spearman's rank correlation $\rho = 0.78$, $P \leq 2.2 \times 10^{-16}$) between male and female cis-eQTL effect sizes. For 58% of sb-eQTLs, sex-stratified cis-eQTL analysis revealed associations in both sexes with concordant allelic effect but different effect sizes. For example, rs117380715-*ADRA1A* in adipose subcutaneous tissue showed a stronger effect in females than in males ($\beta_F = -0.78$, $P_F = 4.64 \times 10^{-18}$, $\beta_M = -0.47$, $P_M = 3.98 \times 10^{-10}$) (Fig. 4B and fig. S8I). For the remainder of the sb-eQTLs, a cis-eQTL was detected exclusively in either females (70, 19%) or males (84, 23%). For example, we identified a female-specific cis-eQTL for rs8942-*C4BPB* in breast ($\beta_F = 0.40$, $P_F = 2.68 \times 10^{-7}$, $\beta_M = -0.02$, $P_M = 0.89$) (Fig. 4B and fig. S8I). *C4BPB* encodes the beta unit of the C4b-binding protein and controls activation of the complement cascade (*45*). We also identified a male-specific cis-eQTL for rs2273535-*AURKA* in skeletal muscle ($\beta_M = 0.47$, $\beta_F = 0.01$), described in (*8*). *AURKA*, encoding Aurora kinase A, is a member of the serine/threonine kinase family involved in mitotic chromosomal segregation and muscle differentiation (*46*) and is a known risk factor for several cancers (*47*). These results demonstrate that sex-biased genetic effects on gene expression exist for a small proportion of previously identified cis-eQTLs, and that some sb-eQTLs affect genes implicated in human phenotypes.

Sex differences in genetic regulation of gene expression are partially mediated by cell type–specific eQTLs

Given that the G×Sex interaction term of our eQTL model captures interactions that derive from sex as well as interactions with sex-correlated factors, we next characterized the fraction of sex-biased eQTLs that are driven by cell type–specific eQTLs (fig. S9A). We focused on breast, the tissue with the most sb-eQTLs and the largest sex differences in cellular composition (figs. S7A and S8B). We tested 261 breast sb-eQTLs for enrichment of cell type interacting cis-eQTLs (ieQTLs) (*9*, *29*). These ieQTLs correspond to cis-eQTLs where the effect varies depending on estimated cell type abundances (*29*). Breast sb-eQTLs were strongly enriched ($\pi_1 = 0.66$ and 0.89) for ieQTL signal corresponding to adipocytes and epithelial cells (fig. S9B). After including an interaction term for genotype × epithelial cell abundance estimates in the sb-eQTL model, 58% of breast sb-eQTLs (152/261) remained significant, whereas for 42% of sb-eQTLs (109/261), the genotype × sex effect was strongly attenuated (fig. S9C and table S14). For example, the strongest breast sb-eQTL, rs2289149-*LINC00920* ($P = 4.83 \times 10^{-11}$), was not significant after incorporating the genotype × epithelial cell abundance estimates in the model ($\beta_{G\times Sex} = 0.187$, 95% confidence interval = [−0.004, 0.378]; fig. S9C and table S14).

To formally test the impact of cell type composition on sb-eQTL detection, we performed a mediation analysis, using genotype interactions with estimated epithelial cell abundance as a potential mediator (*9*) (fig. S9D). We discovered that 60 sb-eQTLs (23%) were mediated by cell type abundances (average causal mediation effects $P \leq 0.001$) (Fig. 4C and table S14). Mediation by other cell types cannot be excluded, particularly by immune cells: We observed that breast sb-eGenes are enriched for immunoglobulin variable chain genes (Fisher's exact test, odds ratio = 12, $P = 9.2 \times 10^{-8}$). In all cases, the eQTL effect size is larger in females (table S11). Because immunoglobulin genes are mainly expressed in B cells and are among the most sex-discriminative genes in breast (fig. S7D), we hypothesize that immunoglobulin sb-eQTLs may be driven

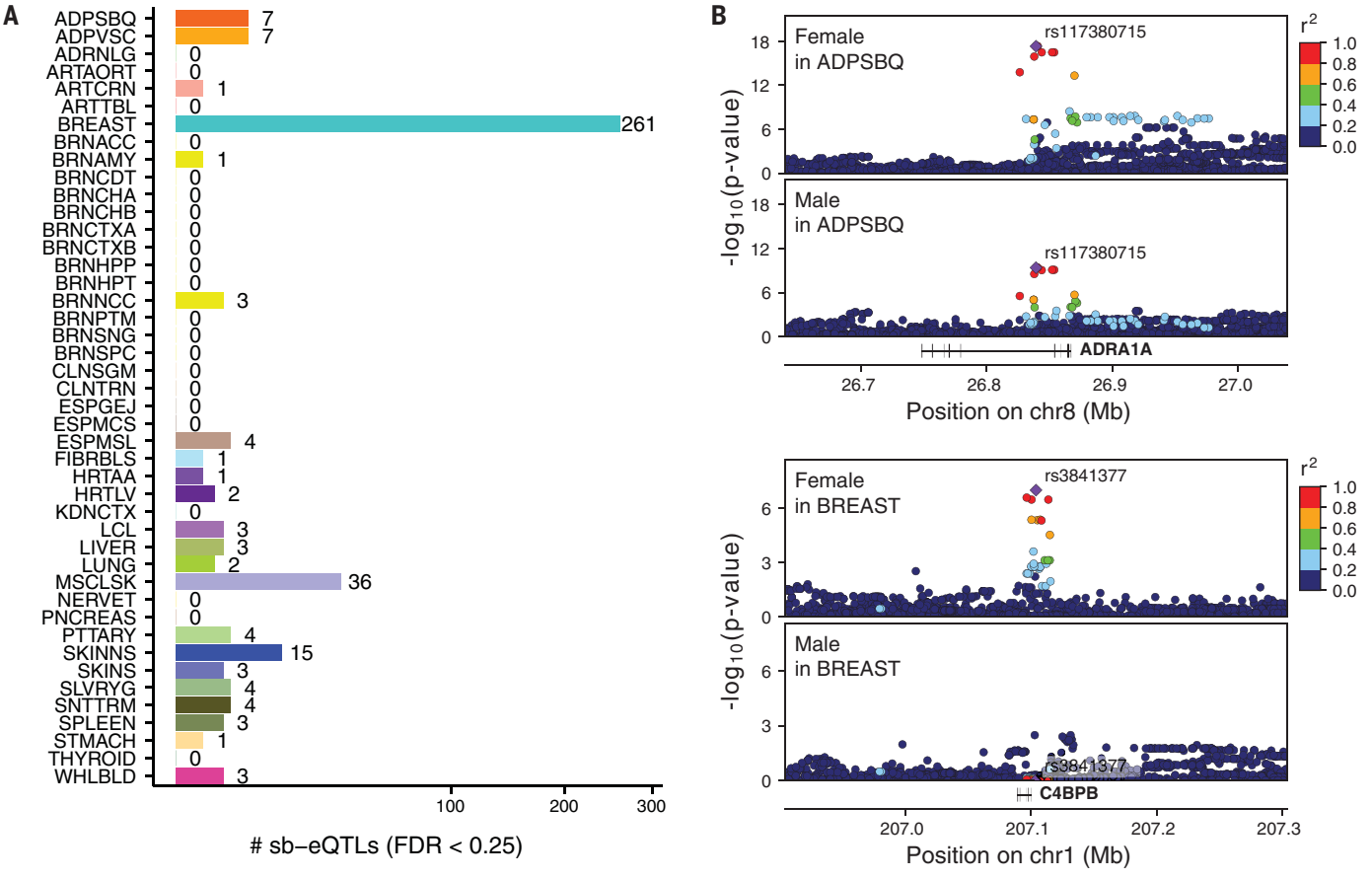


Fig. 4. Sex-biased eQTLs (sb-eQTLs). (**A**) Number of sb-eQTLs discovered per tissue. Square-root transformation was applied to the *x* axis. See Fig. 1 for tissue abbreviations. (**B**) Association *P* values of the female-stratified (top) and male-stratified (bottom) cis-eQTLs in the *ADRA1A* locus in adipose subcutaneous tissue (upper panels; $\beta_F = -0.78$, $P_F = 4.64 \times 10^{-18}$, $\beta_M = -0.47$, $P_M = 3.98 \times 10^{-10}$, $P_{G\times Sex} = 1.05 \times 10^{-5}$) and *C4BPB* locus in breast mammary tissue (lower panels; $\beta_F = 0.40$, $P_F = 2.68 \times 10^{-7}$, $\beta_M = -0.02$, $P_M = 0.89$, $P_{G\times Sex} = 7.22 \times 10^{-5}$). Linkage disequilibrium between loci is quantified by squared Pearson coefficient of correlation (r^2). Diamond-shaped point represents the top significant eQTL variant across sex-stratified *P* values. (**C**) sb-eQTL mediation analysis of 261 breast sb-eQTLs. Point coordinates represent the effect size of G×Sex (*x* axis) and G×Epithelial cells (*y* axis) derived from a linear regression model with both interaction terms. Gray lines represent confidence intervals of the effect sizes of G×Sex (horizontal lines) and G×Epithelial cells (vertical lines). Point size represents sb-eQTL significance; color corresponds to mediation significance. See (9) for further details.

by greater abundances of this cell type in female breasts. Collectively, these results indicate that a large proportion of sb-eQTLs in breast are driven by cell type–specific genetic effects on gene expression that become apparent when cell types differ between sexes, although our analysis cannot distinguish whether the tested cell types or others correlated with them (fig. S9E) are the true mediators of the signal.

Sex-aware eQTL-GWAS colocalization provides insights into the genetic basis of complex traits

To assess whether sb-eQTLs are useful as a means of dissecting the molecular basis of

complex trait associations, we performed colocalization (*48*) between sex-stratified cis-eQTLs and 87 GWASs, representing 74 distinct complex traits, for 1089 sb-eGenes at a more relaxed FDR (≤0.50) (*9*). We identified 74 colocalized gene-trait pairs [posterior probability of sharing the same causal variant (PP4) > 0.5; Fig. 5, A to C]. Of these, 58 were colocalized (PP4 > 0.5) in one sex but not in the other—36 for females and 22 for males—corresponding to 36 unique genetic loci and 27 distinct traits (Fig. 5, A to C, and table S15). For 24/36 (67%) female-stratified and 10/22 (45%) male-stratified cis-eQTL–trait pairs, evidence for colocalization was also found using the male and female combined GTEx v8 cis-eQTLs (fig. S10A). For these 34 loci that colocalized in the sex-combined approach, we found evidence that the colocalization signal is driven by regulatory effects in a single sex. The remaining 12/36 (33%) female and 12/22 (55%) male gene-trait colocalizations were not discovered with the sex-combined approach.

The strongest colocalizations between a trait and a female-stratified cis-eQTL were identified for *CCDC88C* and breast cancer, and for *HKDC1* and birth weight (Fig. 5, C and D). Conversely, the strongest colocalizations between a trait and a male-stratified cis-eQTL were identified for *DPYSL4* and percentage of body fat, and for *CLDN7* and birth weight (Fig. 5, C and E). *CCDC88C* is a negative regulator of the Wnt signaling pathway, a key mechanism in cancer progression (*49*), and the *CCDC88C* female cis-eQTL signal in breast colocalizes with risk of breast cancer (Fig. 5D, left), a trait with highly sex-differentiated incidence and presentation (*50*). For breast cancer, we identified two additional female-driven ($PP4_F$ > $PP4_M$) colocalized sb-eGenes, *NTN4* and *CRLF3* (table S15), previously reported as breast cancer–relevant genes (*51*, *52*).

We also discovered a preferential colocalization of blood and immune traits with female-stratified relative to male-stratified cis-eQTLs (odds ratio = 2.22; $P = 0.047$, Fisher's exact test). This includes inflammatory bowel diseases, which show a higher prevalence in females with increasing age (*53*), and immune cell abundances in blood, which also exhibit sex differences (*30*, *31*). Together, these results suggest that sex-biased genetic regulation of gene expression may contribute to the etiology of diseases with marked sex differences.

Moreover, we identified colocalization signal for eQTLs and GWAS of sex-specific traits as well as signal possibly derived from sex-specific conditions, such as pregnancy in females and balding patterns in males. The *C9orf66* male-stratified cis-eQTL signal in breast colocalized with balding patterns in males, and the *HKDC1* female-stratified cis-eQTL signal in liver colocalized with birth weight, which is strongly influenced by maternal factors (Fig. 5D, right) (*54*). The sb-eQTL for this locus in liver was replicated in an independent dataset (*55*) (rs35696875-*HKDC1* $P_F = 2.73 \times 10^{-8}$, $P_M = 1.60 \times 10^{-4}$, z-test $P = 0.004$; fig. S10B). *HKDC1* encodes a member of the hexokinase protein family and is involved in glucose metabolism. Multiple variants in perfect or high linkage disequilibrium with rs35696875 that cause reduced expression of *HKDC1* have been associated with gestational diabetes mellitus risk (*56*) and glycemic traits during pregnancy (*54*). Here, we confirmed that the *HKDC1* female eQTL signal in the liver colocalizes with maternal glucose levels in plasma during pregnancy (PP4 = 0.92; fig. S10C). Recently, regulatory variants spanning multiple enhancers were found to have a coordinated allelic effect on *HKDC1* expression in hepatocyte-derived cells (*57*). Estimates of hepatocyte abundance in GTEx liver samples did not differ by sex ($P = 0.30$), and the rs35696875-*HKDC1* sb-eQTL showed no evidence of being a hepatocyte ieQTL ($P_{G\times Hepatocytes} = 0.11$) (*29*). Thus, unlike many sb-eQTLs in breast, the *HKDC1* sb-eQTL in liver did not seem to be driven by sex-differentiated cell type abundances. The *HKDC1* sb-eQTL alternative allele is associated with lower *HKDC1* expression, higher maternal glucose levels, and increased birth weight. These results suggest that the *HKDC1* female cis-eQTL influences glucose metabolism in the pregnant female, which is reflected in the birth weight of the offspring. Further investigation is needed, however, to prove causality.

Additionally, the *DPYSL4* male-stratified cis-eQTL signal in skeletal muscle colocalized with genetic signal associated with percentage of body fat (Fig. 5E, right). *DPYSL4* is linked to the pathophysiology of obesity and cancer: p53-inducible *DPYSL4* associates with mitochondrial supercomplexes and regulates energy metabolism in adipocytes and cancer cells. Low *DPYSL4* expression is associated with poor survival of breast cancer patients (*58*). Of note, although the colocalizing signal was detected with the male-stratified cis-eQTL signal, the low probability of colocalization appears to be due to the presence of an additional cis-eQTL in females that is absent in males. These results suggest that characterizing sex differences in the genetic associations of complex traits and molecular phenotypes can prove useful to dissect allelic heterogeneity.

Five colocalized sb-eGenes (*CLDN7*, *CCDC125*, *FAM53B*, *PLEC*, and *SOWAHC*), corresponding to cell type interaction cis-eQTL (cell type ieQTL) signals, also colocalized with reported GWAS signals (birth weight, blood cell counts, height, platelet counts, and schizophrenia, respectively) (*29*). For instance, the male-biased cis-eQTL rs34958987-*CLDN7* in breast (Fig. 5E, left, and fig. S10D) was identified as an epithelial cell ieQTL in breast (*29*). Both the sb-eQTL and cell type ieQTL signals colocalized with the birth weight GWAS signal (fig. S10E). This suggests that the origin of these sex differences in gene-trait associations may be in sex-differentiated cell type abundances.

Finally, to assess whether sex-biased eQTL signals are reflected in sex-biased GWAS effects, we obtained sex-stratified GWAS data for 36 of the 58 colocalized gene-trait pairs (*9*) (table S15). We identified two of 36 loci with sex differences in GWAS effect size (FDR ≤ 0.05, Bonferroni correction). These two signals correspond to *RNASET2* and *CELSR2* genes, which are more strongly associated to hyperthyroidism in females and to heart attack in males, respectively. However, with the current GWAS sample sizes, we observed that, in general, sex-biased effects at the eQTL level do not readily translate into sex-biased effects at the GWAS level, in line with recent power calculations where millions of GWAS samples were estimated to be needed to address this question (*37*).

Overall, our colocalization results identified loci where sex-differentiated cell type abundances mediate genotype-phenotype associations, and also loci where sex may play a more direct role in the underlying molecular mechanism of the association, as in the *HKDC1* locus. For future studies, accounting for context or environment (sex in the present study) in colocalization approaches is a promising approach to the discovery of gene-trait associations and their underlying origins.

Discussion

We identified widespread sex-biased gene expression in all tissues, with 37% of genes exhibiting sex bias in at least one tissue, but with overall small (median FC = 1.04) sex effects. These results derive from overlapping male and female distributions of interindividual expression variation, indicative of differential expression as opposed to completely dimorphic expression. These genes represent diverse molecular and biological functions, and they include genes relevant to disease and clinical phenotypes. As expected, the strongest sex bias was observed for X-chromosome genes, whereas the vast majority of sex-biased genes were autosomal, which suggests the influence of sex on genome-wide regulatory programs. As reported in (*59*) but not well characterized to date, we discovered that a portion of these genes were nonrandomly distributed across the genome, suggesting sex differences in regional regulation. Integration of these results with sex-aware analysis of epigenetic and chromosome conformation capture Hi-C data may provide mechanistic insights into these patterns.

Although we identified a set of X-linked genes with sex-biased expression across many tissues, the overall sharing of sex-biased expression among tissues was strongly skewed

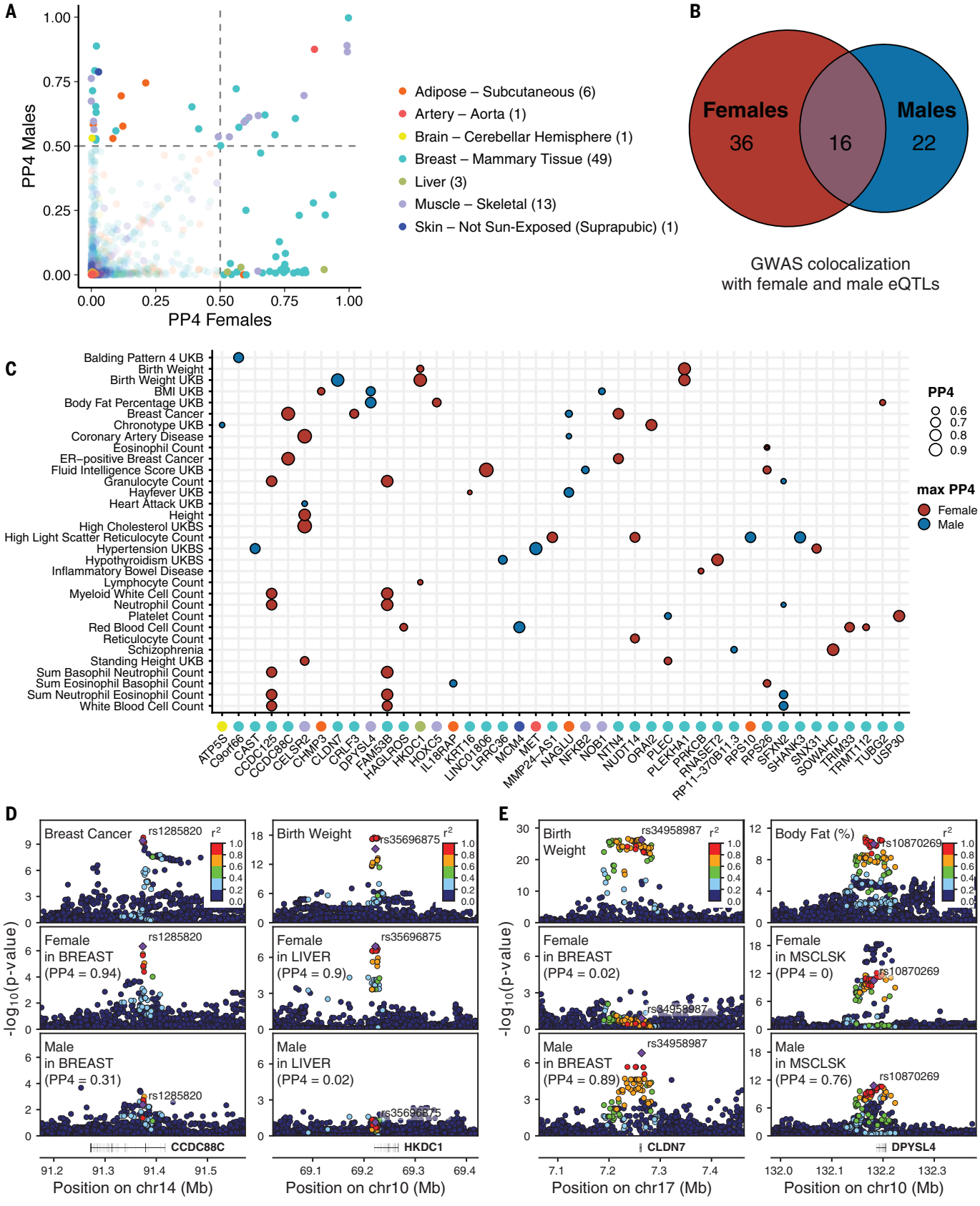

A
PP4 Males
PP4 Females
1.00
0.75
0.50
0.25
0.00
Adipose – Subcutaneous (6)
Artery – Aorta (1)
Brain – Cerebellar Hemisphere (1)
Breast – Mammary Tissue (49)
Liver (3)
Muscle – Skeletal (13)
Skin – Not Sun-Exposed (Suprapubic) (1)
B
Females
36
16
Males
22
GWAS colocalization
with female and male eQTLs
C
Balding Pattern 4 UKB
Birth Weight
Birth Weight UKB
BMI UKB
Body Fat Percentage UKB
Breast Cancer
Chronotype UKB
Coronary Artery Disease
Eosinophil Count
ER−positive Breast Cancer
Fluid Intelligence Score UKB
Granulocyte Count
Hayfever UKB
Heart Attack UKB
Height
High Cholesterol UKBS
High Light Scatter Reticulocyte Count
Hypertension UKBS
Hypothyroidism UKBS
Inflammatory Bowel Disease
Lymphocyte Count
Myeloid White Cell Count
Neutrophil Count
Platelet Count
Red Blood Cell Count
Reticulocyte Count
Schizophrenia
Standing Height UKB
Sum Basophil Neutrophil Count
Sum Eosinophil Basophil Count
Sum Neutrophil Eosinophil Count
White Blood Cell Count
ATP5S C9orf66 CAST CCDC125 CCDC88C CELSR2 CHMP3 CLDN7 CRLF3 DPYSL4 FAM53B HAGLROS HKDC1 HOXC5 IL18RAP KRT16 LINC01806 LRRC36 MCM4 MET MMP24-AS1 NAGLU NFKB2 NOB1 NTN4 NUDT14 ORAI2 PLEC PLEKHA1 PRKCB RNASET2 RP11-370B11.3 RPS10 RPS26 SFXN2 SHANK3 SNX31 SOWAHC TRIM33 TRMT112 TUBG2 USP30
PP4
0.6
0.7
0.8
0.9
max PP4
Female
Male
D
Breast Cancer
rs1285820
Female in BREAST (PP4 = 0.94)
Male in BREAST (PP4 = 0.31)
Birth Weight
rs35696875
Female in LIVER (PP4 = 0.9)
Male in LIVER (PP4 = 0.02)
-log10(p-value)
CCDC88C
HKDC1
Position on chr14 (Mb)
Position on chr10 (Mb)
E
Birth Weight
rs34958987
Female in BREAST (PP4 = 0.02)
Male in BREAST (PP4 = 0.89)
Body Fat (%)
rs10870269
Female in MSCLSK (PP4 = 0)
Male in MSCLSK (PP4 = 0.76)
CLDN7
DPYSL4
Position on chr17 (Mb)
Position on chr10 (Mb)

Fig. 5. Colocalization of sb-eQTLs with GWAS traits. (**A**) Posterior probability (PP4) of 74 colocalized gene-trait pairs where a GWAS shows evidence of colocalization with the female-stratified and/or male-stratified cis-eQTL signal (PP4 > 0.5). Numbers of colocalizing loci per tissue are shown in parentheses. (**B**) Numbers of colocalizing loci for female and male cis-eQTLs. (**C**) GWAS-eQTL colocalizing genes (PP4 > 0.5) color-labeled by eQTL tissue of origin according to labels in (A) (*x* axis) are categorized by the sex where the colocalization signal is maximized with the corresponding GWAS trait (*y* axis). Comparing the colocalization PP4 values for male and female cis-eQTL signals, the estimates can be maximum in females (red) or males (blue). (**D**) Genotype-phenotype association *P* values of the *CCDC88C* (left) and *HKDC1* (right) loci. For the *CCDC88C* locus, panels illustrate GWAS signal for breast cancer (top) and *CCDC88C* cis-eQTL signal for females (middle) and males (bottom) in breast mammary tissue. For the *HKDC1* locus, panels illustrate GWAS signal for birth weight (top) and *HKDC1* cis-eQTL signal for females (middle) and males (bottom) in liver. (**E**) Genotype-phenotype association *P* values of the *CLDN7* (left) and *DPYSL4* (right) loci. For the *CLDN7* locus, panels illustrate GWAS signal for birth weight (top) and *CLDN7* cis-eQTL signal for females (middle) and males (bottom) in breast mammary tissue. For the *DPYSL4* locus, panels illustrate GWAS signal for body fat (top) and *DPYSL4* cis-eQTL signal for females (middle) and males (bottom) in muscle skeletal tissue. In (D) and (E), linkage disequilibrium between loci is quantified by squared Pearson coefficient of correlation (r^2). Diamond-shaped point represents the top significant cis-eQTL variant across sex-stratified *P* values.

toward tissue specificity, with 18.2% of sex-biased genes discovered in only a single tissue. The high tissue specificity of sex-biased gene expression and the enrichment of TFBSs in sex-biased gene promoters implicate specific TFs in mediating sex-biased expression. Functional experiments to assess sex-differentiated TF binding are needed to evaluate the role of TF function in observed patterns.

In contrast to the large impact of sex on gene expression levels, the overall extent of sex effects on genetic regulation in cis is much less (369 sb-eQTLs). This observation is consistent with an overall weaker role of sex in genetic regulation but is also affected by differences in power of the two analyses (*60*). For sb-eQTLs, the combination of small genotype × sex interaction effect sizes, high interindividual expression heterogeneity, and the sex imbalance in the GTEx collection affects the power of the interaction test. This implies that much larger cohorts are needed to fully characterize this phenomenon, particularly to assess sex effects for all cis variants and genes. The relatively modest number of G×Sex interactions for a factor as impactful as sex suggests that other, more subtle genotype-interacting environmental factors are likely to be challenging to identify [as noted in (*39*)]. The sb-eQTL analysis is also affected by cell type heterogeneity within tissues. We demonstrated that a portion of sb-eQTLs are mediated by cell type composition, which suggests that a portion of the sb-eQTL signal may derive from the combination of cell type–specific eQTLs and sex differences in the tissue's cell type composition. The remaining loci for which we had no evidence of cell type mediation may represent true sex differences in genetic regulation of these genes, but might also derive from unknown factors confounded with sex, including cell types that were not part of our analysis. Thus, the full impact of cell type differences across tissues remains to be determined.

The identification of sb-eQTLs that are unequivocally not derived from sex differences in cell type abundances cannot be assessed with analysis of sb-eQTLs in bulk tissue. We anticipate that single-cell sb-eQTL analysis will help to disentangle sex effects on the genetics of gene expression that derive from sex differences in tissue composition versus those that derive from sex chromosome status. However, this approach also has limitations due to the removal of cells from the in situ tissue environment—including, for example, the presence of other cell types and diverse hormonal environments.

In efforts to understand the molecular basis of sex differences in disease and other phenotypes, it is important to note that the connection between the molecular changes observed here and complex phenotypes is likely to be complicated by many compensatory and buffering effects (*61*). Despite extensive sex differences at the transcriptome level, the majority of biology at all phenotype levels is shared between males and females. Furthermore, the sex differences observed here are based on a snapshot of mostly older individuals. Sex differences that occur during different developmental stages, in specific environments, or in specific disease states are not well represented in our analysis. For example, sex biases are observed in many cancers (*1*). Our results provide a resource of sex effects in "nondiseased" tissues to compare with those of disease cohorts. We note that sex is highly correlated with many features of behavior and external environments [e.g., smoking (*62*)], and disentangling sex differences driven by inherent biology versus gendered environments is an important further challenge.

Beyond gene expression, sex-biased genetic regulation may also contribute to higher-order phenotypes such as complex traits and diseases; colocalization analysis of sex-stratified cis-eQTLs and sex-combined GWAS summary statistics yielded variant-gene-trait associations that were not detected in combined-sex cis-eQTL colocalization analysis. In general, context-aware colocalization analyses may help to elucidate the origin of gene-trait associations, as hypothesized here for *HKDC1*'s impact on birth weight through alteration of glucose metabolism in a pregnant female's liver. We show that sex-biased gene-trait associations are likely attributable to either allelic heterogeneity in the combined-sex cohort or genetic effects on gene expression that are (predominantly) driven by a single sex; colocalized sb-eGenes cannot be considered as proxies of loci harboring sex differences in the genetic architecture of the linked trait. Because sex-aware colocalizations can provide insights into the sex-differentiated genetic architecture of disease, we expect future work in this area combining sex-stratified cis-eQTLs with summary statistics from sex-stratified GWASs to enable us to fully comprehend the impact of sex on human health and disease. The extension of analytical approaches to facilitate widespread genetic analysis of sex chromosomes is an important step toward these new research directions.

Methods summary

Sex-differential expression was performed with voom-limma (*63*) and MASH (*64*) (fig. S1A). Sex-differential effect sizes and gene expression levels were investigated for tissue specificity with the Tau index (*65*), clustered with pvclust (*66*), and compared with dendextend (*67*) (fig. S3A). Sex predictivity of sex-biased genes per tissue was quantified through gradient-boosted tree classifier models (*68*) (fig. S4A). Positional gene enrichment analysis of sex-biased genes was performed with PGE (*18*) (fig. S5A). Transcription factor binding site enrichment in promoter regions of sex-biased genes was performed with Unibind (*22*) and runLOLA (*69*) (fig. S5E). Gene set enrichment analysis was performed with fgsea (*70*) (fig. S6A) and results characterized with Cytoscape (*71*). Sex differences in cell type abundances and their effect on histopathological phenotypes were explored using linear regression. sb-eQTL mapping was implemented using an adaptation of FastQTL (*72*) (fig. S8A); sb-eQTLs were validated using haplotype-level allelic expression data generated with phASER and allele-specific expression modeling using EAGLE. Characterization of sex-specific cis-eQTL effects was performed with linear regression. Mediation of G×Sex by G×Epithelial interactions was tested with the mediation R package. Colocalization of GWAS and eQTLs was performed with coloc (*48*). Further details for each analysis are provided in (*9*).

REFERENCES AND NOTES

1. D. Zheng *et al*., Sexual dimorphism in the incidence of human cancers. *BMC Cancer* **19**, 684 (2019). doi: 10.1186/s12885-019-5902-z; pmid: 31299933
2. G. D. Anderson, Sex and racial differences in pharmacological response: Where is the evidence? Pharmacogenetics, pharmacokinetics, and pharmacodynamics. *J. Womens Health* **14**, 19–29 (2005). doi: 10.1089/jwh.2005.14.19; pmid: 15692274
3. V. Kuan *et al*., A chronological map of 308 physical and mental health conditions from 4 million individuals in the English National Health Service. *Lancet Digit. Health* **1**, e63–e77 (2019). doi: 10.1016/S2589-7500(19)30012-3; pmid: 31650125
4. S. T. Ngo, F. J. Steyn, P. A. McCombe, Gender differences in autoimmune disease. *Front. Neuroendocrinol.* **35**, 347–369 (2014). doi: 10.1016/j.yfrne.2014.04.004; pmid: 24793874
5. D. Westergaard, P. Moseley, F. K. H. Sørup, P. Baldi, S. Brunak, Population-wide analysis of differences in disease progression patterns in men and women. *Nat. Commun.* **10**, 666 (2019). doi: 10.1038/s41467-019-08475-9; pmid: 30737381
6. E. A. Khramtsova, L. K. Davis, B. E. Stranger, The role of sex in the genomics of human complex traits. *Nat. Rev. Genet.* **20**, 173–190 (2019). doi: 10.1038/s41576-018-0083-1; pmid: 30581192
7. J. Lonsdale *et al*., The Genotype-Tissue Expression (GTEx) project. *Nat. Genet.* **45**, 580–585 (2013). doi: 10.1038/ng.2653; pmid: 23715323
8. The GTEx Consortium, The GTEx Consortium atlas of genetic regulatory effects across human tissues. *Science* **369**, 1318 (2020). doi: 10.1126/science.aaz1776; pmid: 25954002
9. See supplementary materials.
10. M. Melé *et al*., The human transcriptome across tissues and individuals. *Science* **348**, 660–665 (2015). doi: 10.1126/science.aaa0355; pmid: 25954002
11. M. Gershoni, S. Pietrokovski, The landscape of sex-differential transcriptome and its consequent selection in human adults. *BMC Biol.* **15**, 7 (2017). doi: 10.1186/s12915-017-0352-z; pmid: 28173793
12. S. Naqvi *et al*., Conservation, acquisition, and functional impact of sex-biased gene expression in mammals. *Science* **365**, eaaw7317 (2019). doi: 10.1126/science.aaw7317; pmid: 31320509
13. C. M. Lopes-Ramos *et al*., Sex Differences in Gene Expression and Regulatory Networks across 29 Human Tissues. *Cell Rep.* **31**, 107795 (2020). doi: 10.1016/j.celrep.2020.107795; pmid: 32579922
14. L. Carrel, H. F. Willard, X-inactivation profile reveals extensive variability in X-linked gene expression in females. *Nature* **434**, 400–404 (2005). doi: 10.1038/nature03479; pmid: 15772666
15. J. L. Rinn, M. Snyder, Sexual dimorphism in mammalian gene expression. *Trends Genet.* **21**, 298–305 (2005). doi: 10.1016/j.tig.2005.03.005; pmid: 15851067
16. S. García-Calzón, A. Perfilyev, V. D. de Mello, J. Pihlajamäki, C. Ling, Sex Differences in the Methylome and Transcriptome of the Human Liver and Circulating HDL-Cholesterol Levels. *J. Clin. Endocrinol. Metab.* **103**, 4395–4408 (2018). doi: 10.1210/jc.2018-00423; pmid: 29846646
17. S. E. Ellis, L. Collado-Torres, A. Jaffe, J. T. Leek, Improving the value of public RNA-seq expression data by phenotype prediction. *Nucleic Acids Res.* **46**, e54 (2018). doi: 10.1093/nar/gky102; pmid: 29514223
18. K. De Preter, R. Barriot, F. Speleman, J. Vandesompele, Y. Moreau, Positional gene enrichment analysis of gene sets for high-resolution identification of overrepresented chromosomal regions. *Nucleic Acids Res.* **36**, e43 (2008). doi: 10.1093/nar/gkn114; pmid: 18346969
19. A. Kelkar, V. Thakur, R. Ramaswamy, D. Deobagkar, Characterisation of inactivation domains and evolutionary strata in human X chromosome through Markov segmentation. *PLOS ONE* **4**, e7885 (2009). doi: 10.1371/journal.pone.0007885; pmid: 19946363
20. T. Tukiainen *et al*., Landscape of X chromosome inactivation across human tissues. *Nature* **550**, 244–248 (2017). doi: 10.1038/nature24265; pmid: 29022598
21. F. Le Dily *et al*., Distinct structural transitions of chromatin topological domains correlate with coordinated hormone-induced gene regulation. *Genes Dev.* **28**, 2151–2162 (2014). doi: 10.1101/gad.241422.114; pmid: 25274727
22. M. Gheorghe *et al*., A map of direct TF-DNA interactions in the human genome. *Nucleic Acids Res.* **47**, e21 (2019). doi: 10.1093/nar/gky1210; pmid: 30517703
23. F. Spitz, E. E. M. Furlong, Transcription factors: From enhancer binding to developmental control. *Nat. Rev. Genet.* **13**, 613–626 (2012). doi: 10.1038/nrg3207; pmid: 22868264
24. N. Leuenberger, S. Pradervand, W. Wahli, Sumoylated PPARα mediates sex-specific gene repression and protects the liver from estrogen-induced toxicity in mice. *J. Clin. Invest.* **119**, 3138–3148 (2009). doi: 10.1172/JCI39019; pmid: 19729835
25. N. Brockdorff, Polycomb complexes in X chromosome inactivation. *Philos. Trans. R. Soc. London Ser. B* **372**, 20170021 (2017). doi: 10.1038/nrg3207; pmid: 22868264
26. B. M. Nugent, C. M. O'Donnell, C. N. Epperson, T. L. Bale, Placental H3K27me3 establishes female resilience to prenatal insults. *Nat. Commun.* **9**, 2555 (2018). doi: 10.1038/s41467-018-04992-1; pmid: 29967448
27. D. Lau-Corona, W. K. Bae, L. Hennighausen, D. J. Waxman, Sex-biased genetic programs in liver metabolism and liver fibrosis are controlled by EZH1 and EZH2. bioRxiv [Preprint]. 14 March 2019. pmid: 577056
28. D. F. Deegan, N. Engel, Sexual Dimorphism in the Age of Genomics: How, When, Where. *Front. Cell Dev. Biol.* **7**, 186 (2019). doi: 10.3389/fcell.2019.00186; pmid: 31552249
29. S. Kim-Hellmuth *et al*., Cell type specific genetic regulation of gene expression across human tissues. *Science* **369**, eaaz8528 (2020).
30. Y. Chen *et al*., Difference in Leukocyte Composition between Women before and after Menopausal Age, and Distinct Sexual Dimorphism. *PLOS ONE* **11**, e0162953 (2016). doi: 10.1371/journal.pone.0162953; pmid: 27657912
31. E. Bongen *et al*., Sex Differences in the Blood Transcriptome Identify Robust Changes in Immune Cell Proportions with Aging and Influenza Infection. *Cell Rep.* **29**, 1961–1973.e4 (2019). doi: 10.1016/j.celrep.2019.10.019; pmid: 31722210
32. W. J. Astle *et al*., The Allelic Landscape of Human Blood Cell Trait Variation and Links to Common Complex Disease. *Cell* **167**, 1415–1429.e19 (2016). doi: 10.1016/j.cell.2016.10.042; pmid: 27863252
33. K. Rawlik, O. Canela-Xandri, A. Tenesa, Evidence for sex-specific genetic architectures across a spectrum of human complex traits. *Genome Biol.* **17**, 166 (2016). doi: 10.1186/s13059-016-1025-x; pmid: 27473438
34. D. Shungin *et al*., New genetic loci link adipose and insulin biology to body fat distribution. *Nature* **518**, 187–196 (2015). doi: 10.1038/nature14132; pmid: 25673412
35. S. L. Pulit *et al*., Meta-analysis of genome-wide association studies for body fat distribution in 694 649 individuals of European ancestry. *Hum. Mol. Genet.* **28**, 166–174 (2019). doi: 10.1093/hmg/ddy327; pmid: 30239722
36. J. Martin *et al*, Examining sex-differentiated genetic effects across neuropsychiatric and behavioral traits. bioRxiv [Preprint]. 5 May 2020. pmid: 076042
37. E. Porcu *et al*, The role of gene expression on human sexual dimorphism: Too early to call. bioRxiv [Preprint]. 17 April 2020. pmid: 042986
38. S. E. Castel *et al*, A vast resource of allelic expression data spanning human tissues. bioRxiv [Preprint]. 3 October 2019. pmid: 792911
39. D. A. Knowles *et al*., Allele-specific expression reveals interactions between genetic variation and environment. *Nat. Methods* **14**, 699–702 (2017). doi: 10.1038/nmeth.4298; pmid: 28530654
40. K. R. Kukurba *et al*., Impact of the X Chromosome and sex on regulatory variation. *Genome Res.* gr.197897.115 (2016). doi: 10.3389/fgene.2019.00313; pmid: 31024623
41. The GAIT2 Project; http://ugcd.github.io/pages/projects/gait2/.
42. J. J. Shen, Y.-F. Wang, W. Yang, Sex-Interacting mRNA- and miRNA-eQTLs and Their Implications in Gene Expression Regulation and Disease. *Front. Genet.* **10**, 313 (2019). doi: 10.3389/fgene.2019.00313; pmid: 31024623
43. C. Yao *et al*., Sex- and age-interacting eQTLs in human complex diseases. *Hum. Mol. Genet.* **23**, 1947–1956 (2014). doi: 10.1093/hmg/ddt582; pmid: 24242183
44. A. C. Leon, M. Heo, Sample Sizes Required to Detect Interactions between Two Binary Fixed-Effects in a Mixed-Effects Linear Regression Model. *Comput. Stat. Data Anal.* **53**, 603–608 (2009). doi: 10.1016/j.csda.2008.06.010; pmid: 20084090
45. K. Iida, V. Nussenzweig, Complement receptor is an inhibitor of the complement cascade. *J. Exp. Med.* **153**, 1138–1150 (1981). doi: 10.1084/jem.153.5.1138; pmid: 6910481
46. K. Dhanasekaran *et al*., Unraveling the role of aurora A beyond centrosomes and spindle assembly: Implications in muscle differentiation. *FASEB J.* **33**, 219–230 (2019). doi: 10.1096/fj.201800997; pmid: 29995440
47. A. Tang *et al*., Aurora kinases: Novel therapy targets in cancers. *Oncotarget* **8**, 23937–23954 (2017). doi: 10.18632/oncotarget.14893; pmid: 28147341
48. C. Giambartolomei *et al*., Bayesian test for colocalisation between pairs of genetic association studies using summary statistics. *PLOS Genet.* **10**, e1004383 (2014). doi: 10.1371/journal.pgen.1004383; pmid: 24830394
49. T. Zhan, N. Rindtorff, M. Boutros, Wnt signaling in cancer. *Oncogene* **36**, 1461–1473 (2017). doi: 10.1038/onc.2016.304; pmid: 27617575
50. D. Ly, D. Forman, J. Ferlay, L. A. Brinton, M. B. Cook, An international comparison of male and female breast cancer incidence rates. *Int. J. Cancer* **132**, 1918–1926 (2013). doi: 10.1002/ijc.27841; pmid: 22987302
51. X. Xu, Q. Yan, Y. Wang, X. Dong, NTN4 is associated with breast cancer metastasis via regulation of EMT-related biomarkers. *Oncol. Rep.* **37**, 449–457 (2017). doi: 10.3892/or.2016.5239; pmid: 27840993
52. M. M. Marjaneh *et al*, High-throughput allelic expression imbalance analyses identify candidate breast cancer risk genes. bioRxiv [Preprint]. 16 January 2019. pmid: 521013
53. J. D. Betteridge, S. P. Armbruster, C. Maydonovitch, G. R. Veerappan, Inflammatory bowel disease prevalence by age, gender, race, and geographic location in the U.S. military health care population. *Inflamm. Bowel Dis.* **19**, 1421–1427 (2013). doi: 10.1097/MIB.0b013e318281334d; pmid: 23518811
54. M. G. Hayes *et al*., Identification of HKDC1 and BACE2 as genes influencing glycemic traits during pregnancy through genome-wide association studies. *Diabetes* **62**, 3282–3291 (2013). doi: 10.2337/db12-1692; pmid: 23903356
55. T. Strunz *et al*., A mega-analysis of expression quantitative trait loci (eQTL) provides insight into the regulatory architecture of gene expression variation in liver. *Sci. Rep.* **8**, 5865 (2018). doi: 10.1007/s11892-015-0709-z; pmid: 26803651
56. W. L. Lowe Jr., D. M. Scholtens, V. Sandler, M. G. Hayes, Genetics of Gestational Diabetes Mellitus and Maternal Metabolism. *Curr. Diab. Rep.* **16**, 15 (2016). doi: 10.1007/s11892-015-0709-z; pmid: 26803651
57. C. Guo *et al*., Coordinated regulatory variation associated with gestational hyperglycaemia regulates expression of the novel hexokinase HKDC1. *Nat. Commun.* **6**, 6069 (2015). doi: 10.1038/ncomms7069; pmid: 25648650
58. H. Nagano *et al*., p53-inducible DPYSL4 associates with mitochondrial supercomplexes and regulates energy metabolism in adipocytes and cancer cells. *Proc. Natl. Acad. Sci. U.S.A.* **115**, 8370–8375 (2018). doi: 10.1073/pnas.1804243115; pmid: 30061407
59. B. J. Matthews, D. J. Waxman, CTCF and Cohesin link sex-biased distal regulatory elements to sex-biased gene expression in mouse liver. bioRxiv [Preprint]. 14 March 2019. pmid: 577577
60. G. Shieh, Effect size, statistical power, and sample size for assessing interactions between categorical and continuous variables. *Br. J. Math. Stat. Psychol.* **72**, 136–154 (2019). doi: 10.1111/bmsp.12147; pmid: 30468259
61. Y. Liu, A. Beyer, R. Aebersold, On the Dependency of Cellular Protein Levels on mRNA Abundance. *Cell* **165**, 535–550 (2016). doi: 10.1016/j.cell.2016.03.014; pmid: 27104977
62. A. M. Allen, T. S. Scheuermann, N. Nollen, D. Hatsukami, J. S. Ahluwalia, Gender Differences in Smoking Behavior and Dependence Motives Among Daily and Nondaily Smokers. *Nicotine Tob. Res.* **18**, 1408–1413 (2016). doi: 10.1093/ntr/ntv138; pmid: 26136526
63. C. W. Law, Y. Chen, W. Shi, G. K. Smyth, voom: Precision weights unlock linear model analysis tools for RNA-seq read counts. *Genome Biol.* **15**, R29 (2014). doi: 10.1186/gb-2014-15-2-r29; pmid: 24485249
64. S. M. Urbut, G. Wang, P. Carbonetto, M. Stephens, Flexible statistical methods for estimating and testing effects in genomic studies with multiple conditions. *Nat. Genet.* **51**, 187–195 (2019). doi: 10.1038/s41588-018-0268-8; pmid: 30478440
65. N. Kryuchkova-Mostacci, M. Robinson-Rechavi, A benchmark of gene expression tissue-specificity metrics. *Brief. Bioinform.* **18**, 205–214 (2017). pmid: 26891983
66. R. Suzuki, H. Shimodaira, pvclust: Hierarchical Clustering with P-Values via Multiscale Bootstrap Resampling (2015).
67. T. Galili, dendextend: An R package for visualizing, adjusting and comparing trees of hierarchical clustering. *Bioinformatics*

31, 3718–3720 (2015). doi: 10.1093/bioinformatics/btv428; pmid: 26209431
68. T. Chen, C. Guestrin, in *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (2016), pp. 785–794.
69. N. C. Sheffield, C. Bock, LOLA: Enrichment analysis for genomic region sets and regulatory elements in R and Bioconductor. *Bioinformatics* (2016); http://code.databio.org/LOLA.
70. A. A. Sergushichev, An algorithm for fast preranked gene set enrichment analysis using cumulative statistic calculation. bioRxiv [Preprint]. 20 June 2016. pmid: 060012
71. P. Shannon *et al*., Cytoscape: A software environment for integrated models of biomolecular interaction networks. *Genome Res.* **13**, 2498–2504 (2003). doi: 10.1101/gr.1239303; pmid: 14597658
72. H. Ongen, A. Buil, A. A. Brown, E. T. Dermitzakis, O. Delaneau, Fast and efficient QTL mapper for thousands of molecular phenotypes. *Bioinformatics* **32**, 1479–1485 (2016). doi: 10.1093/bioinformatics/btv722; pmid: 26708335

ACKNOWLEDGMENTS

We thank the donors and their families for their generous gifts of biospecimens to the GTEx research project; the Genomics Platform at the Broad Institute for data generation; J. Struewing for his support and leadership of the GTEx project; B. H. F. Weber and T. Strunz for assistance with replicating the sb-eQTL for *HKDC1* in liver; D. Nicolae and L. Chen for advice on mediation analysis; J. Witkos for comments on an earlier version of the manuscript; M. Gloudemans for making the sex-stratified eQTL data available on LocusCompare (http://locuscompare.com/); D. Muehlschlegel for assistance with replicating sex-biased genes; E. Flynn for assistance in the interpretation of sex-biased gene expression patterns, and G. Hayes for providing GWAS summary statistics for maternal glycemic traits. The summary figure and Fig. 1 were partially generated using www.biorender.com, and M. Khan and C. Stolte contributed to the design. This work was completed in part with computational resources provided by the Center for Research Informatics at the University of Chicago and the Centre for Genomic Regulation. **Funding:** Supported by the Common Fund of the Office of the Director, U.S. National Institutes of Health, and by NCI, NHGRI, NHLBI, NIDA, NIMH, NIA, NIAID, and NINDS through NIH contracts HHSN2612008000001E (Leidos Prime contract with NCI: A.M.S., D.E.T., N.V.R., J.A.M., L.S., M.E.B., L.Q., T.K., D.B., K.R., A.U.), 10XS170 (NDRI: W.F.L., J.A.T., G.K., A.M., S.S., R.H., G.Wal., M.J., M.Wa., L.E.B., C.J., J.W., B.R., M.Hu., K.M., L.A.S., H.M.G., M.Mo., L.K.B.), 10XS171 (Roswell Park Cancer Institute: B.A.F., M.T.M., E.K., B.M.G., K.D.R., J.B.), 10X172 (Science Care Inc.), 12ST1039 (IDOX), 10ST1035 (Van Andel Institute: S.D.J., D.C.R., D.R.V.), HHSN268201000029C (Broad Institute: F.A., G.G., K.G.A., A.V.S., Xiao. L., E.T., S.G., A.G., S.A., K.H.H., D.T.N., K.H., S.R.M., J.L.N.), 5U41HG009494 (F.A., G.G., K.G.A.), and through NIH grants R01 DA006227-17 (Univ. of Miami Brain Bank: D.C.M., D.A.D.), Supplement to University of Miami grant DA006227 (D.C.M., D.A.D.), R01 MH090941 (Univ. of Geneva), R01 MH090951 and R01 MH090937 (Univ. of Chicago), R01 MH090936 (Univ. of North Carolina–Chapel Hill), R01MH101814 (M.M-A., V.W., S.B.M., R.G., E.T.D., D.G-M., A.V.), U01HG007593 (S.B.M.), R01MH101822 (C.D.B.), U01HG007598 (M.O., B.E.S.), U01MH104393 (A.P.F.), extension H002371 to 5U41HG002371 (W.J.K.) as well as other funding sources: R01MH106842 (T.L., P.M., E.F., P.J.H.), R01HL142028 (T.L., Si.Ka., P.J.H.), R01GM122924 (T.L., S.E.C.), R01MH107666 (H.K.I.), P30DK020595 (H.K.I.), UM1HG008901 (T.L.), R01GM124486 (T.L.), R01HG010067 (Y.Pa.), R01HG002585 (G.Wan., M.St.), Gordon and Betty Moore Foundation GBMF 4559 (G.Wa., M.St.), 1K99HG009916-01 (S.E.C.), R01HG006855 (Se.Ka., R.E.H.), BIO2015-70777-P, Ministerio de Economía y Competitividad and FEDER funds (M.M-A., V.W., R.G., D.G-M.), la Caixa Foundation ID 100010434 under agreement LCF/BQ/SO15/52260001 (D.G-M.), NIH CTSA grant UL1TR002550-01 (P.M.), Marie-Skłodowska Curie fellowship H2020 Grant 706636 (S.K-H.), R35HG010718 (E.R.G.), FPU15/03635, Ministerio de Educación, Cultura y Deporte (M.M-A.), R01MH109905, 1R01HG010480 (A.Ba.), Searle Scholar Program (A.Ba.), R01HG008150 (S.B.M.), 5T32HG000044-22, NHGRI Institutional Training Grant in Genome Science (N.R.G.), EU IMI program (UE7-DIRECT-115317-1) (E.T.D., A.V.), FNS funded project RNA1 (31003A_149984) (E.T.D., A.V.), DK110919 (F.H.), F32HG009987 (F.H.), Massachusetts Lions Eye Research Fund Grant (A.R.H.), 2R01GM108711 (L.C.), R01MH101820 (B.E.S.), Supplement to R01MH101820 (E.A.K., P.E., B.E.S.), Consolidate Research Group. Generalitat de Catalunya SGR 1736 and CERCA program (A.M.-P., J.M.S.); Rhodes Trust and Natural Sciences and Engineering Research Council of Canada (A.J.P.). All CRG authors acknowledge the support of the Spanish Ministry of Science, Innovation and Universities to the EMBL partnership, the Centro de Excelencia Severo Ochoa and the CERCA Programme / Generalitat de Catalunya. The University of Chicago Center for Research Informatics is funded by the Biological Sciences Division at the University of Chicago with additional funding from NIH award UL1TR000430. **Author contributions:** B.E.S. conceived the study; M.O. and B.E.S. led the writing and editing of the manuscript and supplement; M.O. and B.E.S. coordinated analyses of all contributing authors; M.O., M.M.-A. and V.W. performed differential gene expression analysis; B.B., D.J.C., M.M.-A., M.O., and V.W. characterized effect sizes of sex-biased genes; M.M.-A and M.O., performed sex-biased genes replication in independent datasets, M.M.-A., M.O., and V.W. performed analysis of transcription factor binding sites; M.M.-A., M.O., and V.W. performed tissue clustering based on gene expression levels and sex bias; M.M.-A. built the expression-based sex classifier; Y.Z. performed *MASH* analyses; M.S. and S.K.-H. provided advice on MASH analysis; M.O. generated sb-eQTL pipelines and performed sb-eQTL mapping; F.A., B.B., A.J.B., B.E.E., E.E., P.E., E.R.G., S.K.-H., S.K., E.A.K., S.B.M., P.P., A.D.S., and B.E.S. contributed to sb-eQTL analysis approach; D.J.C., S.K.-H., E.A.K., M.O., and V.W. characterized sb-eGenes; F.A., B.B., and A.V. performed sb-eQTL replication analysis in external datasets; S.K.-H., A.M.-P., and J.-M.S. contributed to sb-eQTL replication analysis; S.K. performed ASE aFC validation of sb-eQTLs; S.E.C., S.K.-H., and P.M. contributed to ASE aFC validation of sb-eQTLs; P.P. performed EAGLE ASE validation of sb-eQTLs; A.J.B. and S.K.-H. provided advice on EAGLE ASE validation; S.K.-H. performed coloc analysis; M.O. performed mediation analysis; S.K.-H. and B.L.P. provided advice on mediation analysis; F.A., D.G., S.K.-H., D.J.C., M.O., M.M.-A., and V.W. generated figures; F.A., K.G.A., and A.V.S. generated and oversaw GTEx v8 data generation, LDACC, and pipelines; A.N.B., R.B., and H.K.I. generated GWAS data; F.A., S.K.-H., and M.O. generated cell type abundances and ieQTL data; S.K.-H., M.M.-A., and M.O. characterized sex differences in cell type abundances; M.M.-A. and V.W. characterized phenotype relationships with cell type abundances; A.B., A.D.H.G., A.R.H., E.A.K., A.J.P., B.E.E., D.G., E.R.G., S.K.-H., A.M.-P., F.R., and A.D.S. performed analysis or provided feedback that significantly shaped this work but was not included in this final version; M.M.-A. and V.W. managed data and code in the GitHub repository; A.J.B., B.E.E., E.T.D., R.G., H.K.I., T.L., S.B.M., B.L.P., M.S., A.V.S., and B.E.S. supervised the work of trainees in their labs; D.J.C., S.K., S.K.-H., M.M.-A., M.O., P.P., V.W., Y.Z., and B.E.S. wrote manuscript text; and B.B., A.J.B., B.E.E., A.D.H.G., R.G., S.K.-H., H.K.I., E.A.K., T.L., M.M.-A., M.O., S.B.M., L.C., S.K., P.P., B.L.P., A.D.S., B.E.S., A.V., and V.W. edited the manuscript. All authors read and approved the final manuscript. **Competing interests:** F.A. is an inventor on a patent application related to TensorQTL; S.E.C. is a co-founder, chief technology officer, and stock owner at Variant Bio; D.G.M. is co-founder with equity in Goldfinch Bio, and has received research support from AbbVie, Astellas, Biogen, BioMarin, Eisai, Merck, Pfizer, and Sanofi-Genzyme; E.A.K. is an employee of Janssen Pharmaceuticals; H.I. has received speaker honoraria from GSK and AbbVie; E.T.D. is chairman and member of the board of Hybridstat LTD; T.L. is a scientific advisory board member of Variant Bio with equity, and Goldfinch Bio. Other GTEx members: E.R.G. is on the Editorial Board of Circulation Research, and does consulting for the City of Hope / Beckman Research Institute; E.T.D. is chairman and member of the board of Hybridstat Ltd.; B.E.E. is on the scientific advisory boards of Celsius Therapeutics and Freenome; G.G. receives research funds from IBM and Pharmacyclics, and is an inventor on patent applications related to MuTect, ABSOLUTE, MutSig, MSMuTect, MSMutSig, POLYSOLVER, and TensorQTL. G.G. is a founder, consultant and holds privately held equity in Scorpion Therapeutics; S.B.M. is on the scientific advisory board of MyOme; P.F. is member of the scientific advisory boards of Fabric Genomics Inc., and Eagle Genomes Ltd. P.G.F. is a partner of Bioinf2Bio. **Data and materials availability:** All GTEx open-access data, including summary statistics of sex-biased genes and sex-biased eQTLs, are available on the GTEx Portal (https://gtexportal.org/home/datasets). Histological images and their annotations are also available on the portal (https://gtexportal.org/home/histologyPage). GTEx v8 sex-stratified eQTL data are available on LocusCompare (http://locuscompare.com/). All GTEx protected data are available via dbGaP (accession phs000424.v8). Access to the raw sequence data are now provided through the AnVIL platform (https://gtexportal.org/home/protectedDataAccess). Code for the sex-biased gene expression analysis is deposited at https://zenodo.org/record/3939042 (doi: 10.5281/zenodo.3939042).

SUPPLEMENTARY MATERIALS

science.sciencemag.org/content/369/6509/eaba3066/suppl/DC1
Materials and Methods
Figs. S1 to S10
Tables S1 to S15
References (*73*–*102*)

View/request a protocol for this paper from *Bio-protocol*.

GTEx Consortium Laboratory and Data Analysis Coordinating Center (LDACC): François Aguet[1], Shankara Anand[1], Kristin G. Ardlie[1], Stacey Gabriel[1], Gad A. Getz[1,2,3], Aaron Graubert[1], Kane Hadley[1], Robert E. Handsaker[4,5,6], Katherine H. Huang[1], Seva Kashin[4,5,6], Xiao Li[1], Daniel G. MacArthur[5,7], Samuel R. Meier[1], Jared L. Nedzel[1], Duyen T. Nguyen[1], Ayellet V. Segrè[1,8], Ellen Todres[1]
Analysis Working Group Funded by GTEx Project Grants: François Aguet[1], Shankara Anand[1], Kristin G. Ardlie[1], Brunilda Balliu[9], Alvaro N. Barbeira[10], Alexis Battle[11,12], Rodrigo Bonazzola[10], Andrew Brown[13,14], Christopher D. Brown[15], Stephane E. Castel[16,17], Donald F. Conrad[18,19], Daniel J. Cotter[20], Nancy Cox[21], Sayantan Das[22], Olivia M. de Goede[20], Emmanouil T. Dermitzakis[13,23,24], Jonah Einson[16,25], Barbara E. Engelhardt[26,27], Eleazar Eskin[28], Tiffany Y. Eulalio[29], Nicole M. Ferraro[29], Elise D. Flynn[16,17], Laure Fresard[30], Eric R. Gamazon[21,31,32,33], Diego Garrido-Martín[34], Nicole R. Gay[20], Gad A. Getz[1,2,3], Michael J. Gloudemans[29], Aaron Graubert[1], Roderic Guigó[34,35], Kane Hadley[1], Andrew R. Hame[8,1], Robert E. Handsaker[4,5,6], Yuan He[11], Paul J. Hoffman[16], Farhad Hormozdiari[1,36], Lei Hou[1,37], Katherine H. Huang[1], Hae Kyung Im[10], Brian Jo[26,27], Silva Kasela[16,17], Seva Kashin[4,5,6], Manolis Kellis[1,37], Sarah Kim-Hellmuth[16,17,38], Alan Kwong[22], Tuuli Lappalainen[16,17], Xiao Li[1], Xin Li[30], Yanyu Liang[10], Daniel G. MacArthur[5,7], Serghei Mangul[28,39], Samuel R. Meier[1], Pejman Mohammadi[16,17,40,41], Stephen B. Montgomery[20,30], Manuel Muñoz-Aguirre[34,42], Daniel C. Nachun[30], Jared L. Nedzel[1], Duyen T. Nguyen[1], Andrew B. Nobel[43], Meritxell Oliva[10,44], YoSon Park[15,45], Yongjin Park[1,37], Princy Parsana[12], Abhiram S. Rao[46], Ferran Reverter[47], John M. Rouhana[1,8], Chiara Sabatti[48], Ashis Saha[12], Ayellet V. Segrè[1,8], Andrew D. Skol[10,49], Matthew Stephens[50], Barbara E. Stranger[10,51], Benjamin J. Strober[11], Nicole A. Teran[30], Ellen Todres[1], Ana Viñuela[13,23,24,52], Gao Wang[50], Xiaoquan Wen[22], Fred Wright[53], Valentin Wucher[34], Yuxin Zou[54]
Analysis Working Group Not Funded by GTEx Project Grants: Pedro G. Ferreira[55,56,57,58], Gen Li[59], Marta Melé[60], Esti Yeger-Lotem[61,62]
Leidos Biomedical Project Management: Mary E. Barcus[63], Debra Bradbury[63], Tanya Krubit[63], Jeffrey A. McLean[63], Liqun Qi[63], Karna Robinson[63], Nancy V. Roche[63], Anna M. Smith[63], Leslie Sobin[63], David E. Tabor[63], Anita Undale[63]
Biospecimen Collection Source Sites: Jason Bridge[64], Lori E. Brigham[65], Barbara A. Foster[66], Bryan M. Gillard[66], Richard Hasz[67], Marcus Hunter[68], Christopher Johns[69], Mark Johnson[70], Ellen Karasik[66], Gene Kopen[71], William F. Leinweber[71], Alisa McDonald[71], Michael T. Moser[66], Kevin Myer[68], Kimberley D. Ramsey[66], Brian Roe[68], Saboor Shad[71], Jeffrey A. Thomas[71,70], Gary Walters[70], Michael Washington[70], Joseph Wheeler[69]
Biospecimen Core Resource: Scott D. Jewell[72], Daniel C. Rohrer[72], Dana R. Valley[72]
Brain Bank Repository: David A. Davis[73], Deborah C. Mash[73]
Pathology: Mary E. Barcus[63], Philip A. Branton[74], Leslie Sobin[63]
ELSI Study: Laura K. Barker[75], Heather M. Gardiner[75], Maghboeba Mosavel[76], Laura A. Siminoff[75]
Genome Browser Data Integration and Visualization: Paul Flicek[77], Maximilian Haeussler[78], Thomas Juettemann[77], W. James Kent[78], Christopher M. Lee[78], Conner C. Powell[78], Kate R. Rosenbloom[78], Magali Ruffier[77], Dan Sheppard[77], Kieron Taylor[77], Stephen J. Trevanion[77], Daniel R. Zerbino[77]
eGTEx Group: Nathan S. Abell[20], Joshua Akey[79], Lin Chen[44], Kathryn Demanelis[44], Jennifer A. Doherty[80], Andrew P. Feinberg[81], Kasper D. Hansen[82], Peter F. Hickey[83], Lei Hou[1,37], Farzana Jasmine[44], Lihua Jiang[20], Rajinder Kaul[84,85], Manolis Kellis[1,37], Muhammad G. Kibriya[44], Jin Billy Li[20], Qin Li[20], Shin Lin[86], Sandra E. Linder[20], Stephen B. Montgomery[20,30], Meritxell Oliva[10,44], Yongjin Park[1,37], Brandon L. Pierce[44], Lindsay F. Rizzardi[87], Andrew D. Skol[10,49], Kevin S. Smith[30], Michael Snyder[20], John Stamatoyannopoulos[84,88], Barbara E. Stranger[10,51], Hua Tang[20], Meng Wang[20]

NIH Program Management: Philip A. Branton[74], Latarsha J. Carithers[74,89], Ping Guan[74], Susan E. Koester[90], A. Roger Little[91], Helen M. Moore[74], Concepcion R. Nierras[92], Abhi K. Rao[74], Jimmie B. Vaught[74], Simona Volpi[93]

Affiliations
[1]Broad Institute of MIT and Harvard, Cambridge, MA, USA. [2]Cancer Center and Department of Pathology, Massachusetts General Hospital, Boston, MA, USA. [3]Harvard Medical School, Boston, MA, USA. [4]Department of Genetics, Harvard Medical School, Boston, MA, USA. [5]Program in Medical and Population Genetics, Broad Institute of MIT and Harvard, Cambridge, MA, USA. [6]Stanley Center for Psychiatric Research, Broad Institute of MIT and Harvard, Cambridge, MA, USA. [7]Analytic and Translational Genetics Unit, Massachusetts General Hospital, Boston, MA, USA. [8]Ocular Genomics Institute, Massachusetts Eye and Ear, Harvard Medical School, Boston, MA, USA. [9]Department of Biomathematics, University of California, Los Angeles, CA, USA. [10]Section of Genetic Medicine, Department of Medicine, The University of Chicago, Chicago, IL, USA. [11]Department of Biomedical Engineering, Johns Hopkins University, Baltimore, MD, USA. [12]Department of Computer Science, Johns Hopkins University, Baltimore, MD, USA. [13]Department of Genetic Medicine and Development, University of Geneva Medical School, Geneva, Switzerland. [14]Population Health and Genomics, University of Dundee, Dundee, Scotland, UK. [15]Department of Genetics, University of Pennsylvania, Perelman School of Medicine, Philadelphia, PA, USA. [16]New York Genome Center, New York, NY, USA. [17]Department of Systems Biology, Columbia University, New York, NY, USA. [18]Department of Genetics, Washington University School of Medicine, St. Louis, MO, USA. [19]Division of Genetics, Oregon National Primate Research Center, Oregon Health & Science University, Portland, OR, USA. [20]Department of Genetics, Stanford University, Stanford, CA, USA. [21]Division of Genetic Medicine, Department of Medicine, Vanderbilt University Medical Center, Nashville, TN, USA. [22]Department of Biostatistics, University of Michigan, Ann Arbor, MI, USA. [23]Institute for Genetics and Genomics in Geneva (iGE3), University of Geneva, Geneva, Switzerland. [24]Swiss Institute of Bioinformatics, Geneva, Switzerland. [25]Department of Biomedical Informatics, Columbia University, New York, NY, USA. [26]Department of Computer Science, Princeton University, Princeton, NJ, USA. [27]Center for Statistics and Machine Learning, Princeton University, Princeton, NJ, USA. [28]Department of Computer Science, University of California, Los Angeles, CA, USA. [29]Program in Biomedical Informatics, Stanford University School of Medicine, Stanford, CA, USA. [30]Department of Pathology, Stanford University, Stanford, CA, USA. [31]Data Science Institute, Vanderbilt University, Nashville, TN, USA. [32]Clare Hall, University of Cambridge, Cambridge, UK. [33]MRC Epidemiology Unit, University of Cambridge, Cambridge, UK. [34]Centre for Genomic Regulation (CRG), The Barcelona Institute for Science and Technology, Barcelona, Catalonia, Spain. [35]Universitat Pompeu Fabra (UPF), Barcelona, Catalonia, Spain. [36]Department of Epidemiology, Harvard T.H. Chan School of Public Health, Boston, MA, USA. [37]Computer Science and Artificial Intelligence Laboratory, Massachusetts Institute of Technology, Cambridge, MA, USA. [38]Statistical Genetics, Max Planck Institute of Psychiatry, Munich, Germany [39]Department of Clinical Pharmacy, School of Pharmacy, University of Southern California, Los Angeles, CA, USA. [40]Scripps Research Translational Institute, La Jolla, CA, USA. [41]Department of Integrative Structural and Computational Biology, The Scripps Research Institute, La Jolla, CA, USA. [42]Department of Statistics and Operations Research, Universitat Politècnica de Catalunya (UPC), Barcelona, Catalonia, Spain. [43]Department of Statistics and Operations Research and Department of Biostatistics, University of North Carolina, Chapel Hill, NC, USA. [44]Department of Public Health Sciences, The University of Chicago, Chicago, IL, USA. [45]Department of Systems Pharmacology and Translational Therapeutics, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA, USA. [46]Department of Bioengineering, Stanford University, Stanford, CA, USA. [47]Department of Genetics, Microbiology and Statistics, University of Barcelona, Barcelona, Spain. [48]Departments of Biomedical Data Science and Statistics, Stanford University, Stanford, CA, USA. [49]Department of Pathology and Laboratory Medicine, Ann & Robert H. Lurie Children's Hospital of Chicago, Chicago, IL, USA. [50]Department of Human Genetics, University of Chicago, Chicago, IL, USA. [51]Center for Genetic Medicine, Department of Pharmacology, Northwestern University, Feinberg School of Medicine, Chicago, IL, USA. [52]Department of Twin Research and Genetic Epidemiology, King's College London, London, UK. [53]Bioinformatics Research Center and Departments of Statistics and Biological Sciences, North Carolina State University, Raleigh, NC, USA. [54]Department of Statistics, University of Chicago, Chicago, IL, USA. [55]Department of Computer Sciences, Faculty of Sciences, University of Porto, Porto, Portugal. [56]Instituto de Investigação e Inovação em Saúde, University of Porto, Porto, Portugal. [57]Institute of Molecular Pathology and Immunology, University of Porto, Porto, Portugal. [58]Laboratory of Artificial Intelligence and Decision Support, Institute for Systems and Computer Engineering, Technology and Science, Porto, Portugal. [59]Mailman School of Public Health, Columbia University, New York, NY, USA. [60]Life Sciences Department, Barcelona Supercomputing Center, Barcelona, Spain. [61]Department of Clinical Biochemistry and Pharmacology, Ben-Gurion University of the Negev, Beer-Sheva, Israel. [62]National Institute for Biotechnology in the Negev, Beer-Sheva, Israel. [63]Leidos Biomedical, Rockville, MD, USA. [64]Upstate New York Transplant Services, Buffalo, NY, USA. [65]Washington Regional Transplant Community, Annandale, VA, USA. [66]Therapeutics, Roswell Park Comprehensive Cancer Center, Buffalo, NY, USA. [67]Gift of Life Donor Program, Philadelphia, PA, USA. [68]LifeGift, Houston, TX, USA. [69]Center for Organ Recovery and Education, Pittsburgh, PA, USA. [70]LifeNet Health, Virginia Beach, VA. USA. [71]National Disease Research Interchange, Philadelphia, PA, USA. [72]Van Andel Research Institute, Grand Rapids, MI, USA. [73]Department of Neurology, University of Miami Miller School of Medicine, Miami, FL, USA. [74]Biorepositories and Biospecimen Research Branch, Division of Cancer Treatment and Diagnosis, National Cancer Institute, National Institutes of Health, Bethesda, MD, USA. [75]College of Public Health, Temple University, Philadelphia, PA, USA. [76]Virginia Commonwealth University, Richmond, VA, USA. [77]European Molecular Biology Laboratory, European Bioinformatics Institute, Hinxton, UK. [78]Genomics Institute, University of California, Santa Cruz, CA, USA. [79]Carl Icahn Laboratory, Princeton University, Princeton, NJ, USA. [80]Department of Population Health Sciences, The University of Utah, Salt Lake City, UT, USA. [81]Departments of Medicine, Biomedical Engineering, and Mental Health, Johns Hopkins University, Baltimore, MD, USA. [82]Department of Biostatistics, Bloomberg School of Public Health, Johns Hopkins University, Baltimore, MD, USA. [83]Department of Medical Biology, The Walter and Eliza Hall Institute of Medical Research, Parkville, Victoria, Australia. [84]Altius Institute for Biomedical Sciences, Seattle, WA, USA. [85]Division of Genetics, University of Washington, Seattle, WA, USA. [86]Department of Cardiology, University of Washington, Seattle, WA, USA. [87]HudsonAlpha Institute for Biotechnology, Huntsville, AL, USA. [88]Genome Sciences, University of Washington, Seattle, WA, USA. [89]National Institute of Dental and Craniofacial Research, National Institutes of Health, Bethesda, MD, USA. [90]Division of Neuroscience and Basic Behavioral Science, National Institute of Mental Health, National Institutes of Health, Bethesda, MD, USA. [91]National Institute on Drug Abuse, Bethesda, MD, USA. [92]Office of Strategic Coordination, Division of Program Coordination, Planning and Strategic Initiatives, Office of the Director, National Institutes of Health, Rockville, MD, USA. [93]Division of Genomic Medicine, National Human Genome Research Institute, Bethesda, MD, USA.

25 November 2019; accepted 3 August 2020
10.1126/science.aba3066