

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. Khrantsova, 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. Dermizakis, 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 an-

notation, 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 Lys²⁷ 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. ■

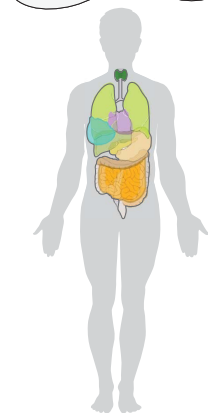
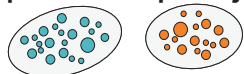
The list of author affiliations and a full list of the GTEx authors and their affiliations are available in the full article online.

†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

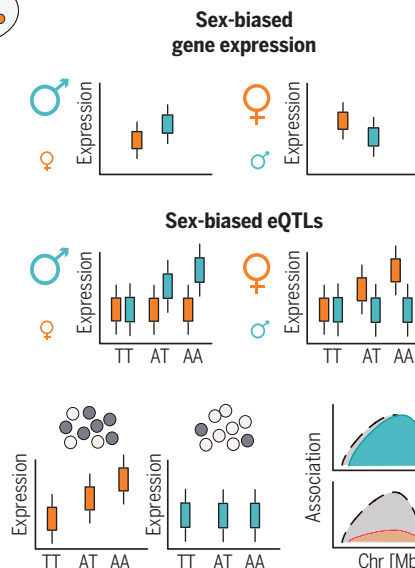
READ THE FULL ARTICLE AT
<https://doi.org/10.1126/science.aba3066>

Sex biases in biological processes and pathways

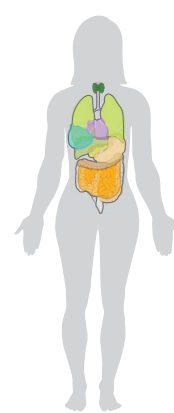
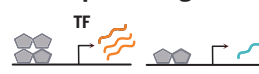


Male

Sex-biased eQTL mediation by cellular abundances



Sex biases in transcriptional regulation



Female

Sex biases in gene-trait association

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.

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 Wucher⁴, Ariel D. H. Gewirtz⁹, Daniel J. Cotter¹⁰, Princy Parsana¹¹, Silva Kasela^{7,8}, Brunilda Balliu¹², Ana Viñuela¹³, Stephane E. Castel^{7,8}, Pejman Mohammadi¹⁴, François Aguet¹⁵, Yuxin Zou¹⁶, Ekaterina A. Khrantsova^{1,17}, Andrew D. Skol^{1,2,18,19}, Diego Garrido-Martín⁴, Ferran Reverter²⁰, Andrew Brown²¹, Patrick Evans²², Eric R. Gamazon^{22,23}, Anthony Payne²⁴, Rodrigo Bonazzola¹, Alvaro N. Barbeira¹, Andrew R. Hamel^{15,25}, Angel Martinez-Perez²⁶, José Manuel Soria²⁶, GTEx Consortium[§], Brandon L. Pierce³, Matthew Stephens^{16,27}, Eleazar Eskin²⁸, Emmanouil T. Dermizakis¹³, Ayellet V. Segre^{15,25}, Hae Kyung Im¹, Barbara E. Engelhardt^{29,30}, Kristin G. Ardlie¹⁵, 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

¹Section of Genetic Medicine, Department of Medicine, University of Chicago, Chicago, IL, USA. ²Institute for Genomics and Systems Biology, University of Chicago, Chicago, IL, USA. ³Department of Public Health Sciences, University of Chicago, Chicago, IL, USA. ⁴Centre for Genomic Regulation, Barcelona Institute for Science and Technology, Barcelona, Catalonia, Spain. ⁵Department of Statistics and Operations Research, Universitat Politècnica de Catalunya, Barcelona, Catalonia, Spain. ⁶Statistical Genetics, Max Planck Institute of Psychiatry, Munich, Germany. ⁷New York Genome Center, New York, NY, USA. ⁸Department of Systems Biology, Columbia University, New York, NY, USA. ⁹Lewis-Sigler Institute for Integrative Genomics, Princeton University, Princeton, NJ, USA. ¹⁰Department of Genetics, Stanford University, Stanford, CA, USA. ¹¹Department of Computer Science, Johns Hopkins University, Baltimore, MD, USA. ¹²Department of Computational Medicine, University of California, Los Angeles, CA, USA. ¹³Department of Genetic Medicine and Development, University of Geneva Medical School, Geneva, Switzerland. ¹⁴Department of Integrative Structural and Computational Biology, The Scripps Research Institute, Scripps Research Translational Institute, La Jolla, CA, USA. ¹⁵Broad Institute of MIT and Harvard, Cambridge, MA, USA. ¹⁶Department of Statistics, University of Chicago, Chicago, IL, USA. ¹⁷Computational Sciences, Janssen Pharmaceuticals, Spring House, PA, USA. ¹⁸Center for Translational Data Science, University of Chicago, Chicago, IL, USA. ¹⁹Department of Pathology and Laboratory Medicine, Ann and Robert H. Lurie Children's Hospital of Chicago, Chicago, IL, USA. ²⁰Department of Genetics, Microbiology and Statistics, Faculty of Biology, University of Barcelona, Barcelona, Spain. ²¹University of Dundee, Dundee, Scotland, UK. ²²Division of Genetic Medicine, Vanderbilt University Medical Center, Nashville, TN, USA. ²³Clare Hall, University of Cambridge, Cambridge, UK. ²⁴Wellcome Centre for Human Genetics, Nuffield Department of Medicine, University of Oxford, Oxford, UK. ²⁵Massachusetts Eye and Ear, Harvard Medical School, Boston, MA, USA. ²⁶Genomics of Complex Diseases Group, Research Institute Hospital de la Sant Creu i Sant Pau, IIB Sant Pau, Barcelona, Spain. ²⁷Department of Human Genetics, University of Chicago, Chicago, IL, USA. ²⁸Departments of Computational Medicine, Computer Science, and Human Genetics, University of California, Los Angeles, CA, USA. ²⁹Department of Computer Science, Center for Statistics and Machine Learning, Princeton University, Princeton, NJ, USA. ³⁰Genomics plc, Oxford, UK. ³¹Department of Pathology, Stanford University, Stanford, CA, USA. ³²Department of Biomedical Engineering, Johns Hopkins University, Baltimore, MD, USA. ³³Universitat Pompeu Fabra, Barcelona, Catalonia, Spain. ³⁴Center for Genetic Medicine, Department of Pharmacology, Northwestern University, Chicago, IL, USA.

*Corresponding author. Email: meritxello@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.

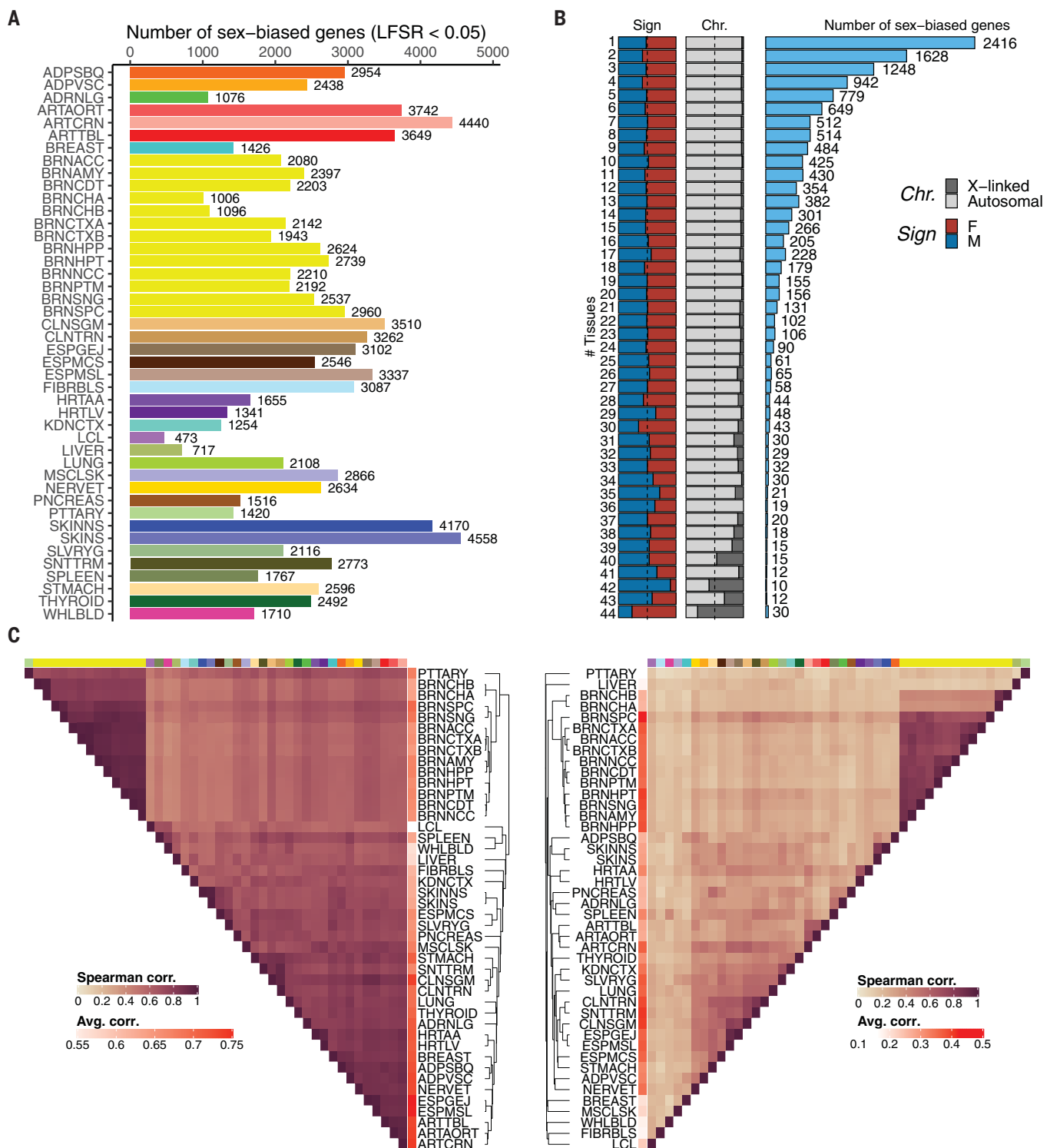


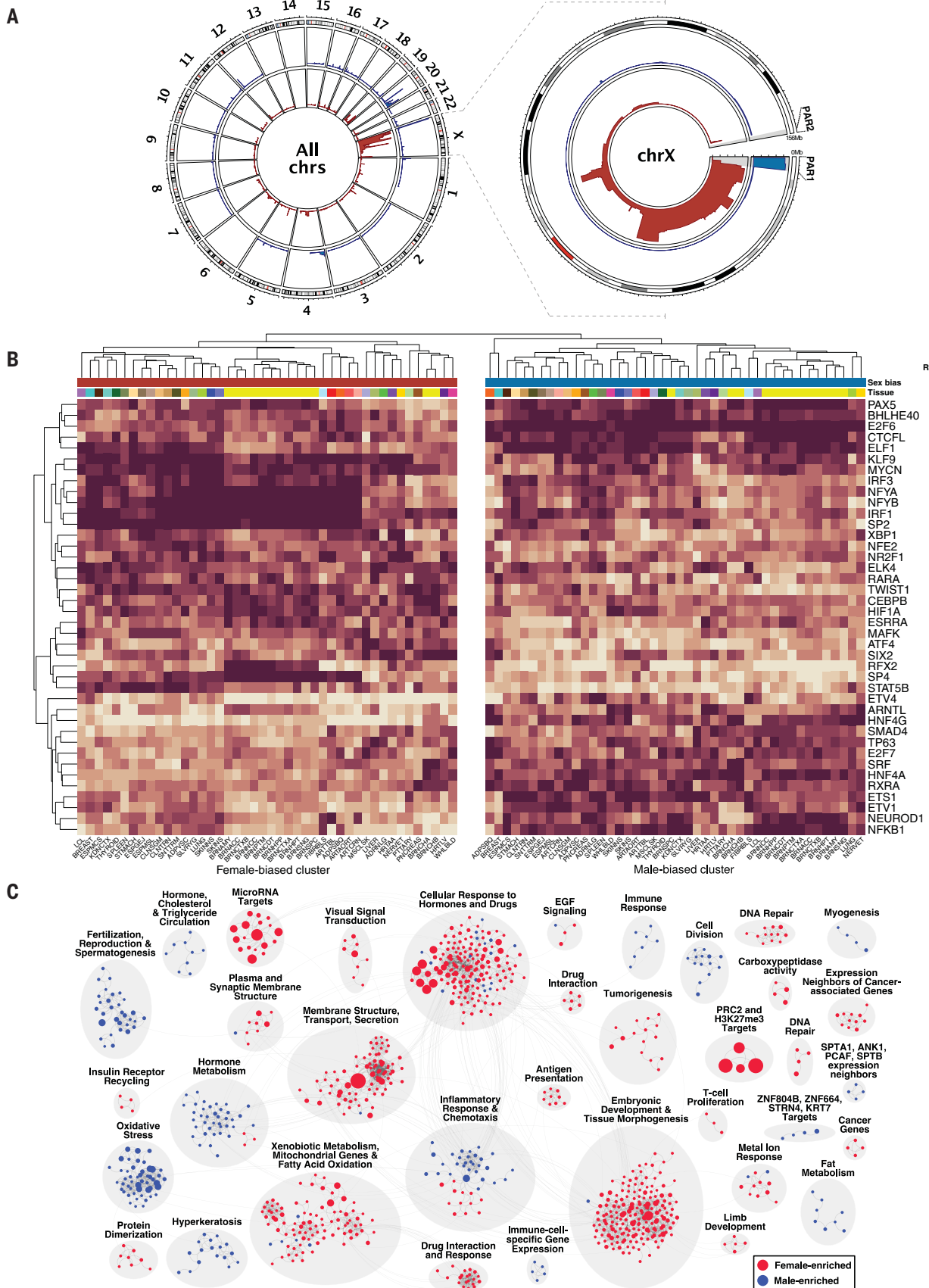
Fig. 2. Sex-differential gene expression. (A) Number of sex-differentially expressed genes (sex-biased genes) per tissue. Tissue colors are as in Fig. 1. (B) Sex-biased gene discovery (histogram, number of sex-biased genes) and characteristics of sex-biased genes (stacked bar plots) as a function of 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 \leq 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, pseudo-autosomal region PARI and the remainder of the X-chromosome short arm *p* were enriched



Downloaded from <http://science.sciencemag.org/> on September 10, 2020

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 *SPI*, *E2F6*, *NRF1*, *KLIF9*, 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 *ETSI* (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 Lys²⁷ (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 \times sex (G \times Sex) interaction on expression (9). Notably, this approach captures G \times 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 \leq 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 ($P_{G \times \text{Sex}} \leq 0.01$). Only 36 sb-eGenes (14%) exhibited sex-biased expression in the discovery tissue [multivariate adaptive shrinkage (MASH) LFSR \leq 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-*ADRALA* 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 \times 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 \times 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 \times 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 \times epithelial cell abundance estimates in the model ($\beta_{G \times \text{Sex}} = 0.187$, 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

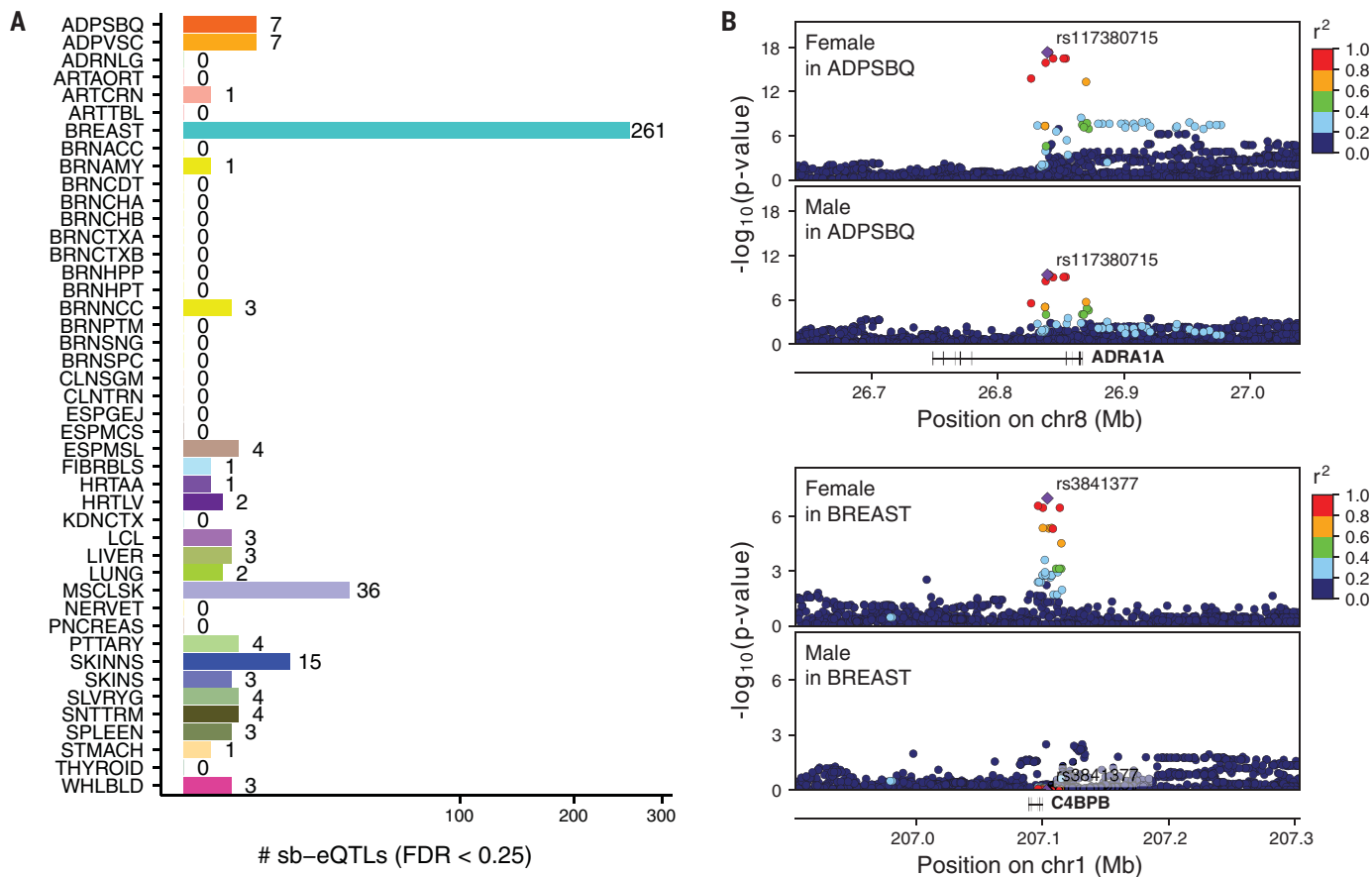


Fig. 4. Sex-biased eQTLs (sb-eQTLs). (A) Number of sb-eQTLs discovered per tissue. Square-root transformation was applied to the x axis. See Fig. 1 for tissue abbreviations. (B) Association P values of the female-stratified (top) and male-stratified (bottom) cis-eQTLs in the *ADRA1A* locus in adipose subcutaneous tissue (upper panels; $\beta_F = -0.78$, $P_F = 4.64 \times 10^{-18}$, $\beta_M = -0.47$, $P_M = 3.98 \times 10^{-10}$, $P_{G \times \text{Sex}} = 1.05 \times 10^{-5}$) and *C4BPB* locus in breast mammary tissue (lower panels; $\beta_F = 0.40$, $P_F = 2.68 \times 10^{-7}$, $\beta_M = -0.02$, $P_M = 0.89$, $P_{G \times \text{Sex}} = 7.22 \times 10^{-5}$). 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 $G \times \text{Sex}$ (x axis) and $G \times \text{Epithelial cells}$ (y axis) derived from a linear regression model with both interaction terms. Gray lines represent confidence intervals of the effect sizes of $G \times \text{Sex}$ (horizontal lines) and $G \times \text{Epithelial cells}$ (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 appar-

ent 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 (PP_4) > 0.5; Fig. 5, A to C]. Of these, 58 were colocalized ($PP_4 > 0.5$) 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 *HKDCI* 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 ($PP_{4F} > PP_{4M}$) 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 *HKDCI* 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) ($rs35696875$ -*HKDCI* $P_F = 2.73 \times 10^{-8}$, $P_M = 1.60 \times 10^{-4}$, z -test $P = 0.004$; fig. S10B). *HKDCI* 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 *HKDCI* have been associated with gestational diabetes mellitus risk (56) and glycemic traits during pregnancy (54). Here, we confirmed that the *HKDCI* female eQTL signal in the liver colocalizes with maternal glucose levels in plasma during pregnancy ($PP_4 = 0.92$; fig. S10C). Recently, regulatory variants spanning multiple enhancers were found to have a coordinated allelic effect on *HKDCI* expression in hepatocyte-derived cells (57). Estimates of hepatocyte abundance in GTEx liver samples did not differ by sex ($P = 0.30$), and the $rs35696875$ -*HKDCI* sb-eQTL showed no evidence of being a hepatocyte ieQTL ($P_{G \times \text{Hepatocytes}} = 0.11$) (29). Thus, unlike many sb-eQTLs in breast, the *HKDCI* sb-eQTL in liver did not seem to be driven by sex-differentiated cell type abundances. The *HKDCI* sb-eQTL alternative allele is associated with lower *HKDCI* expression, higher maternal glucose levels, and increased birth weight. These results suggest that the *HKDCI* 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 *HKDCI* 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

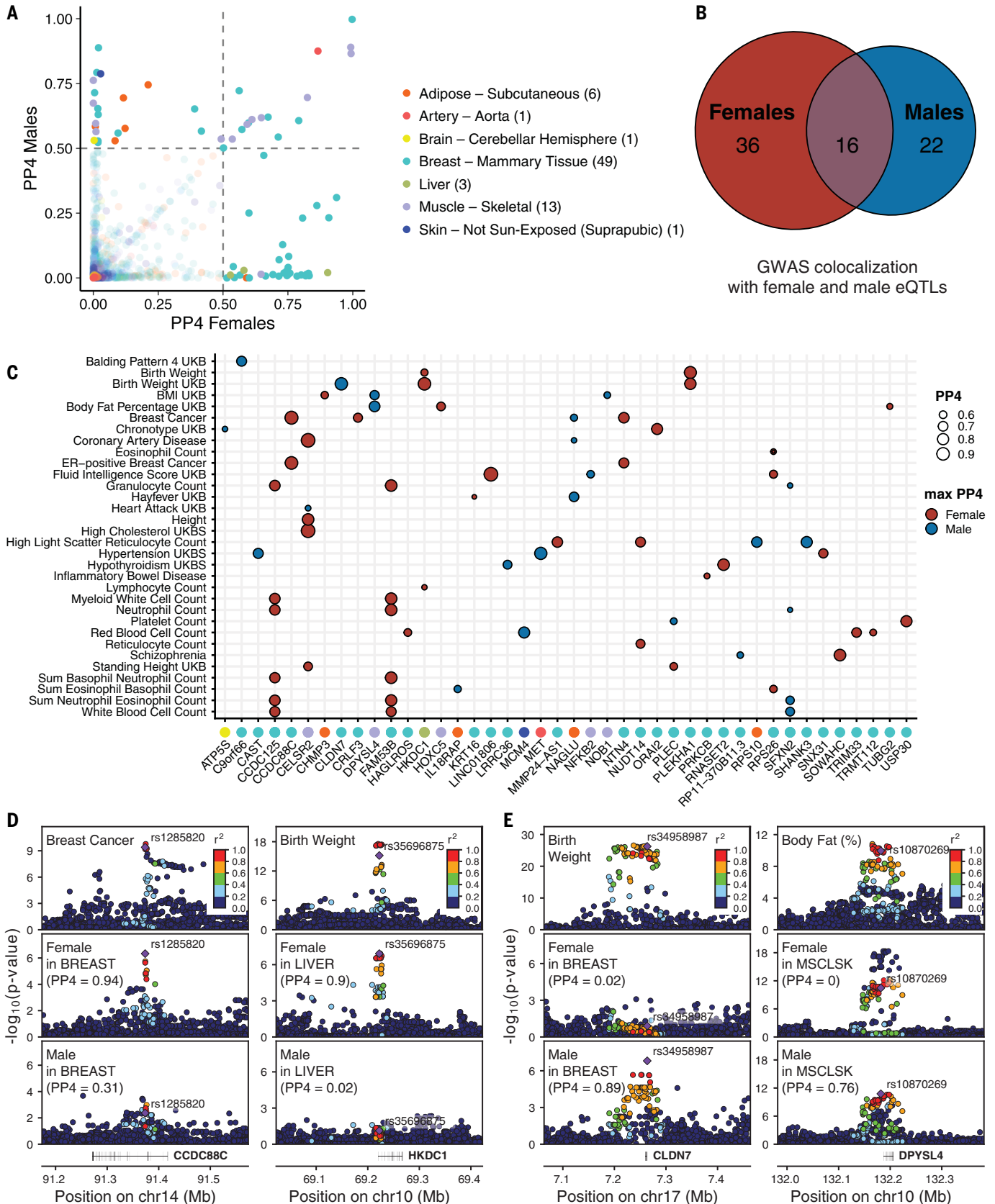


Fig. 5. Colocalization of sb-eQTLs with GWAS traits. (A) Posterior probability (PP4) of 74 colocalized gene-trait pairs where a GWAS shows evidence of colocalization with the female-stratified and/or male-stratified cis-eQTL signal (PP4 > 0.5). Numbers of colocalizing loci per tissue are shown in parentheses. (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 \times 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 $G \times$ 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 pvcult (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 $G \times$ Sex by $G \times$ 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).

31. 3718–3720 (2015). doi: [10.1093/bioinformatics/btv428](https://doi.org/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.data.bio.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](https://doi.org/10.1101/gr.1239303); PMID: 14597658
72. H. Ongen, A. Buil, A. A. Brown, E. T. Dermizakis, O. Delaneau, Fast and efficient QTL mapper for binding sites of molecular phenotypes. *Bioinformatics* **32**, 1479–1485 (2016). doi: [10.1093/bioinformatics/btv722](https://doi.org/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. Struwing 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 HHSN261200800001E (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.S.), 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.), UO1HG007593 (S.B.M.), R01MH101822 (C.D.B.), UO1HG007598 (M.O., B.E.S.), UO1MH104393 (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.), UMH1G008901 (T.L.), R01GM124486 (T.L.), R01HG100667 (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.), B1O2015-70777-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, R01HG101480 (A.B.), 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-I) (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-eQTLs; 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 Freonome; G.G. receives research funds from IBM and Pharmalytics, and is an inventor on patent applications related to MuTect, ABSOLUTE, MutSig, MSMTect, 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 MyOrme; 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/histology>). 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](https://doi.org/10.5281/zenodo.3939042)).

SUPPLEMENTARY MATERIALS

[science.science.org/content/369/6509/eaba3066/suppl/DC1](https://www.science.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 Aguet¹, Shankara Anand¹, Kristin G. Ardlie², Stacey Gabriel¹, Gad A. Getz^{1,2,3}, Aaron Graubert¹, Kane Hadley¹, Robert E. Handsaker^{4,5,6}, Katherine H. Huang¹, Seva Kashin^{4,5,6}, Xiao Li¹, Daniel G. MacArthur^{5,7}, Samuel R. Meier¹, Jared L. Nedzel¹, Duyen T. Nguyen¹, Ayellet V. Segre^{1,8}, Ellen Todes¹

Analysis Working Group Funded by GTEx Project Grants: François Aguet¹, Shankara Anand¹, Kristin G. Ardlie¹, Brunilda Balliu⁹, Alvaro N. Barbeira¹⁰, Alexis Battle^{11,12}, Rodrigo Bonazzola¹⁰, Andrew Brown^{13,14}, Christopher D. Brown¹⁵, Stephane E. Castel^{16,17}, Donald F. Conrad^{18,19}, Daniel J. Cotter²⁰, Nancy Cox²¹, Sayantan Das²², Olivia M. de Goede²⁰, Emmanuel T. Dermizakis^{13,23,24}, Jonah Einson^{16,25}, Barbara E. Engelhardt^{26,27}, Eleazar Eskin²⁸, Tiffany Y. Eulalio²⁹, Nicole M. Ferraro²⁹, Elise D. Flynn^{16,17}, Laure Fressard³⁰, Eric R. Gamazon^{21,31,32,33}, Diego Garrido-Martin³⁴, Nicole R. Gay²⁰, Gad A. Getz^{1,2,3}, Michael J. Gloudemans²⁹, Aaron Graubert¹, Roderic Guigó^{34,35}, Kane Hadley¹, Andrew R. Hame^{18,1}, Robert E. Handsaker^{4,5,6}, Yuan He¹¹, Paul J. Hoffman¹⁶, Farhad Hormozdiari^{1,36}, Lei Hou^{1,37}, Katherine H. Huang¹, Hae Kyung Im¹⁰, Brian Jo^{26,27}, Silva Kasela^{16,17}, Seva Kashin^{4,5,6}, Manolis Kellis^{1,37}, Sarah Kim-Hellmuth^{16,17,38}, Alan Kwong²², Tuuli Lappalainen^{16,17}, Xiao Li¹, Xin Li³⁰, Yanyu Liang¹⁰, Daniel G. MacArthur^{5,7}, Serghei Mangul²⁹, Samuel R. Meier¹, Pejman Mohammadi^{16,17,40,41}, Stephen B. Montgomery^{20,30}, Manuel Muñoz-Aguirre^{34,42}, Daniel C. Natchun³⁰, Jared L. Nedzel¹, Duyen T. Nguyen¹, Andrew B. Nobel⁴³, Meritxell Oliva^{10,44}, YoSon Park^{15,45}, Yongjin Park^{1,37}, Princy Parsana¹², Abhiram S. Rao⁴⁶, Ferran Reverter¹⁷, John M. Rouhana^{1,8}, Chiara Sabatti⁴⁸, Ashis Saha¹², Ayellet V. Segre^{1,8}, Andrew D. Skol^{10,49}, Matthew Stephens⁵⁰, Barbara E. Stranger^{10,51}, Benjamin J. Strober¹, Nicole A. Teran⁵⁰, Ellen Todes¹, Ana Viñuela^{13,23,24,52}, Gao Wang⁵⁰, Xiaoquan Wen²², Fred Wright⁵³, Valentin Wucher⁵⁴, Yuxin Zou⁵⁴

Analysis Working Group Not Funded by GTEx Project Grants: Pedro G. Ferreira^{55,56,57,58}, Gen Li⁵⁹, Marta Melé⁶⁰, Esti Yeger-Lotem^{61,62}

Leidos Biomedical Project Management: Mary E. Barcus⁶³, Debra Bradbury⁶³, Tanya Krubitz⁶³, Jeffrey A. McLean⁶³, Lique Q⁶³, Karina Robinson⁶³, Nancy V. Roche⁶³, Anna M. Smith⁶³, Leslie Sobin⁶³, David E. Tabor⁶³, Anita Undale⁶³

Biospecimen Collection Source Sites: Jason Bridge⁶⁴, Lori E. Brigham⁶⁵, Barbara A. Foster⁶⁶, Bryan M. Gillard⁶⁶, Richard Hasz⁶⁷, Marcus Hunter⁶⁸, Christopher Johns⁶⁹, Mark Johnson⁷⁰, Ellen Karasik⁶⁹, Gene Kopen⁷¹, William F. Leinweber⁷¹, Alisa McDonald⁷², Michael T. Moser⁶⁶, Kevin Myer⁶⁸, Kimberley D. Ramsey⁶⁶, Brian Roe⁶⁸, Saboor Shad⁷¹, Jeffrey A. Thomas^{17,70}, Gary Walters⁷⁰, Michael Washington⁷⁰, Joseph Wheeler⁶⁹

Biospecimen Core Resource: Scott D. Jewell⁷², Daniel C. Rohrer⁷², Dana R. Valley⁷²

Brain Bank Repository: David A. Davis⁷³, Deborah C. Mash⁷³

Pathology: Mary E. Barcus⁶³, Philip A. Branton⁷⁴, Leslie Sobin⁶³

ELSI Study: Laura K. Barker⁷⁵, Heather M. Gardiner⁷⁵, Maghboeba Mosavel⁷⁶, Laura A. Siminoff⁷⁵

Genome Browser Data Integration and Visualization: Paul Flicek⁷⁷, Maximilian Haessler⁷⁸, Thomas Juettemann⁷⁷, W. James Kent⁷⁸, Christopher M. Lee⁷⁸, Conner C. Powell⁷⁸, Kate R. Rosenbloom⁷⁸, Magali Ruffier⁷⁷, Dan Sheppard⁷⁷, Kieron Taylor⁷⁷, Stephen J. Trevanion⁷⁷, Daniel R. Zerbino⁷⁷

eGTEx Group: Nathan S. Abell²⁰, Joshua Akey⁷⁹, Lin Chen⁴⁴, Kathryn Demanelis⁴⁴, Jennifer A. Doherty⁸⁰, Andrew P. Feinberg⁸¹, Kasper D. Hansen⁸², Peter F. Hickey⁸³, Lei Hou^{1,37}, Farzana Jasmine⁴⁴, Lihua Jiang²⁰, Rajinder Kaul^{84,85}, Manolis Kellis^{1,37}, Muhammad G. Kibriya⁴⁴, Jin Billy Li²⁰, Qin Li²⁰, Shin Lin⁸⁶, Sandra E. Linder²⁰, Stephen B. Montgomery^{20,30}, Meritxell Oliva^{10,44}, Yongjin Park^{1,37}, Brandon L. Pierce⁴⁴, Lindsay F. Rizzardi⁸⁷, Andrew D. Skol^{10,49}, Kevin S. Smith³⁰, Michael Snyder²⁰, John Stamatoyannopoulos^{84,88}, Barbara E. Stranger^{10,51}, Hua Tang²⁰, Meng Wang²⁰

NIH Program Management: Philip A. Branton⁷⁴, Latarsha J. Carithers^{74,89}, Ping Guan⁷⁴, Susan E. Koester⁹⁰, A. Roger Little⁹¹, Helen M. Moore⁷⁴, Concepcion R. Nierras⁹², Abhi K. Rao⁷⁴, Jimmie B. Vaught⁷⁴, Simona Volpi⁹³

Affiliations

¹Broad Institute of MIT and Harvard, Cambridge, MA, USA. ²Cancer Center and Department of Pathology, Massachusetts General Hospital, Boston, MA, USA. ³Harvard Medical School, Boston, MA, USA. ⁴Department of Genetics, Harvard Medical School, Boston, MA, USA. ⁵Program in Medical and Population Genetics, Broad Institute of MIT and Harvard, Cambridge, MA, USA. ⁶Stanley Center for Psychiatric Research, Broad Institute of MIT and Harvard, Cambridge, MA, USA. ⁷Analytic and Translational Genetics Unit, Massachusetts General Hospital, Boston, MA, USA. ⁸Ocular Genomics Institute, Massachusetts Eye and Ear, Harvard Medical School, Boston, MA, USA. ⁹Department of Biomathematics, University of California, Los Angeles, CA, USA. ¹⁰Section of Genetic Medicine, Department of Medicine, The University of Chicago, Chicago, IL, USA. ¹¹Department of Biomedical Engineering, Johns Hopkins University, Baltimore, MD, USA. ¹²Department of Computer Science, Johns Hopkins University, Baltimore, MD, USA. ¹³Department of Genetic Medicine and Development, University of Geneva Medical School, Geneva, Switzerland. ¹⁴Population Health and Genomics, University of Dundee, Dundee, Scotland, UK. ¹⁵Department of Genetics, University of Pennsylvania, Perelman School of Medicine, Philadelphia, PA, USA. ¹⁶New York Genome Center, New York, NY, USA. ¹⁷Department of Systems Biology, Columbia University, New York, NY, USA. ¹⁸Department of Genetics, Washington University School of Medicine, St. Louis, MO, USA. ¹⁹Division of Genetics, Oregon National Primate Research Center, Oregon Health & Science University, Portland, OR, USA. ²⁰Department of Genetics, Stanford University, Stanford, CA, USA. ²¹Division of Genetic Medicine, Department of Medicine, Vanderbilt University Medical Center, Nashville, TN, USA. ²²Department of Biostatistics, University of Michigan, Ann Arbor, MI, USA. ²³Institute for Genetics and Genomics in Geneva (IG3), University of Geneva, Geneva, Switzerland. ²⁴Swiss Institute of Bioinformatics, Geneva, Switzerland. ²⁵Department of Biomedical Informatics, Columbia University, New York, NY, USA. ²⁶Department of Computer Science, Princeton University, Princeton, NJ, USA. ²⁷Center for Statistics and Machine Learning, Princeton University, Princeton, NJ, USA. ²⁸Department of Computer Science, University of California, Los Angeles, CA, USA. ²⁹Program in Biomedical Informatics, Stanford University School of Medicine, Stanford, CA, USA. ³⁰Department of Pathology, Stanford University, Stanford,

CA, USA. ³¹Data Science Institute, Vanderbilt University, Nashville, TN, USA. ³²Clare Hall, University of Cambridge, Cambridge, UK. ³³MRC Epidemiology Unit, University of Cambridge, Cambridge, UK. ³⁴Centre for Genomic Regulation (CRG), The Barcelona Institute for Science and Technology, Barcelona, Catalonia, Spain. ³⁵Universitat Pompeu Fabra (UPF), Barcelona, Catalonia, Spain. ³⁶Department of Epidemiology, Harvard T.H. Chan School of Public Health, Boston, MA, USA. ³⁷Computer Science and Artificial Intelligence Laboratory, Massachusetts Institute of Technology, Cambridge, MA, USA. ³⁸Statistical Genetics, Max Planck Institute of Psychiatry, Munich, Germany. ³⁹Department of Clinical Pharmacy, School of Pharmacy, University of Southern California, Los Angeles, CA, USA. ⁴⁰Scripps Research Translational Institute, La Jolla, CA, USA. ⁴¹Department of Integrative Structural and Computational Biology, The Scripps Research Institute, La Jolla, CA, USA. ⁴²Department of Statistics and Operations Research, Universitat Politècnica de Catalunya (UPC), Barcelona, Catalonia, Spain. ⁴³Department of Statistics and Operations Research and Department of Biostatistics, University of North Carolina, Chapel Hill, NC, USA. ⁴⁴Department of Public Health Sciences, The University of Chicago, Chicago, IL, USA. ⁴⁵Department of Systems Pharmacology and Translational Therapeutics, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA, USA. ⁴⁶Department of Bioengineering, Stanford University, Stanford, CA, USA. ⁴⁷Department of Pathology and Laboratory Medicine, Ann & Robert H. Lurie Children's Hospital of Chicago, Chicago, IL, USA. ⁴⁸Departments of Biomedical Data Science and Statistics, Stanford University, Stanford, CA, USA. ⁴⁹Department of Pathology and Laboratory Medicine, Ann & Robert H. Lurie Children's Hospital of Chicago, Chicago, IL, USA. ⁵⁰Department of Human Genetics, University of Chicago, Chicago, IL, USA. ⁵¹Center for Genetic Medicine, Department of Pharmacology, Northwestern University, Feinberg School of Medicine, Chicago, IL, USA. ⁵²Department of Twin Research and Genetic Epidemiology, King's College London, London, UK. ⁵³Bioinformatics Research Center and Departments of Statistics and Biological Sciences, North Carolina State University, Raleigh, NC, USA. ⁵⁴Department of Statistics, University of Chicago, Chicago, IL, USA. ⁵⁵Department of Computer Sciences, Faculty of Sciences, University of Porto, Porto, Portugal. ⁵⁶Instituto de Investigação e Inovação em Saúde, University of Porto, Porto, Portugal. ⁵⁷Institute of Molecular Pathology and Immunology, University of Porto, Porto, Portugal. ⁵⁸Laboratory of Artificial Intelligence and Decision Support, Institute for Systems and Computer Engineering, Technology and Science, Porto, Portugal. ⁵⁹Mailman School of Public Health, Columbia University, New York, NY, USA. ⁶⁰Life Sciences Department, Barcelona Supercomputing Center,

Barcelona, Spain. ⁶¹Department of Clinical Biochemistry and Pharmacology, Ben-Gurion University of the Negev, Beer-Sheva, Israel. ⁶²National Institute for Biotechnology in the Negev, Beer-Sheva, Israel. ⁶³Leidos Biomedical, Rockville, MD, USA. ⁶⁴Upstate New York Transplant Services, Buffalo, NY, USA. ⁶⁵Washington Regional Transplant Community, Annandale, VA, USA. ⁶⁶Therapeutics, Roswell Park Comprehensive Cancer Center, Buffalo, NY, USA. ⁶⁷Gift of Life Donor Program, Philadelphia, PA, USA. ⁶⁸LifeGift, Houston, TX, USA. ⁶⁹Center for Organ Recovery and Education, Pittsburgh, PA, USA. ⁷⁰LifeNet Health, Virginia Beach, VA, USA. ⁷¹National Disease Research Interchange, Philadelphia, PA, USA. ⁷²Van Andel Research Institute, Grand Rapids, MI, USA. ⁷³Department of Neurology, University of Miami Miller School of Medicine, Miami, FL, USA. ⁷⁴Biorepositories and Biospecimen Research Branch, Division of Cancer Treatment and Diagnosis, National Cancer Institute, National Institutes of Health, Bethesda, MD, USA. ⁷⁵College of Public Health, Temple University, Philadelphia, PA, USA. ⁷⁶Virginia Commonwealth University, Richmond, VA, USA. ⁷⁷European Molecular Biology Laboratory, European Bioinformatics Institute, Hinxton, UK. ⁷⁸Genomics Institute, University of California, Santa Cruz, CA, USA. ⁷⁹Carl Icahn Laboratory, Princeton University, Princeton, NJ, USA. ⁸⁰Department of Population Health Sciences, The University of Utah, Salt Lake City, UT, USA. ⁸¹Departments of Medicine, Biomedical Engineering, and Mental Health, Johns Hopkins University, Baltimore, MD, USA. ⁸²Department of Biostatistics, Bloomberg School of Public Health, Johns Hopkins University, Baltimore, MD, USA. ⁸³Department of Medical Biology, The Walter and Eliza Hall Institute of Medical Research, Parkville, Victoria, Australia. ⁸⁴Altius Institute for Biomedical Sciences, Seattle, WA, USA. ⁸⁵Division of Genetics, University of Washington, Seattle, WA, USA. ⁸⁶Department of Cardiology, University of Washington, Seattle, WA, USA. ⁸⁷HudsonAlpha Institute for Biotechnology, Huntsville, AL, USA. ⁸⁸Genome Sciences, University of Washington, Seattle, WA, USA. ⁸⁹National Institute of Dental and Craniofacial Research, National Institutes of Health, Bethesda, MD, USA. ⁹⁰Division of Neuroscience and Basic Behavioral Science, National Institute of Mental Health, National Institutes of Health, Bethesda, MD, USA. ⁹¹National Institute on Drug Abuse, Bethesda, MD, USA. ⁹²Office of Strategic Coordination, Division of Program Coordination, Planning and Strategic Initiatives, Office of the Director, National Institutes of Health, Rockville, MD, USA. ⁹³Division of Genomic Medicine, National Human Genome Research Institute, Bethesda, MD, USA.

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

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 and Barbara E. Stranger

Science **369** (6509), eaba3066.
DOI: 10.1126/science.aba3066

The role of sex in the human transcriptome

In humans, the inheritance of the XX or XY set of sex chromosomes is responsible for most individuals developing into adults expressing male or female sex-specific traits. However, the degree to which sex-biased gene expression occurs in tissues, especially those that do not contribute to characteristic sexually dimorphic traits, is unknown. Oliva *et al.* examined Genotype-Tissue Expression (GTEx) project data and found that 37% of genes in at least one of the 44 tissues studied exhibit a tissue-specific, sex-biased gene expression. They also identified a sex-specific variation in cellular composition across tissues. Overall, the effects of sex on gene expression were small, but they were genome-wide and mostly mediated through transcription factor binding. With sex-biased gene expression associated with loci identified in genome-wide association studies, this study lays the groundwork for identifying the molecular basis of male- and female-based diseases.

Science, this issue p. eaba3066

ARTICLE TOOLS

<http://science.sciencemag.org/content/369/6509/eaba3066>

SUPPLEMENTARY MATERIALS

<http://science.sciencemag.org/content/suppl/2020/09/09/369.6509.eaba3066.DC1>

RELATED CONTENT

<http://science.sciencemag.org/content/sci/369/6509/1318.full>
<http://science.sciencemag.org/content/sci/369/6509/1316.full>
<http://science.sciencemag.org/content/sci/369/6509/eaaz8528.full>
<http://science.sciencemag.org/content/sci/369/6509/eaaz5900.full>
<http://science.sciencemag.org/content/sci/369/6509/eaaz6876.full>
<http://stm.sciencemag.org/content/scitransmed/6/258/258fs40.full>
<http://stm.sciencemag.org/content/scitransmed/11/473/eaao5253.full>

REFERENCES

This article cites 100 articles, 15 of which you can access for free
<http://science.sciencemag.org/content/369/6509/eaba3066#BIBL>

Use of this article is subject to the [Terms of Service](#)

Science (print ISSN 0036-8075; online ISSN 1095-9203) is published by the American Association for the Advancement of Science, 1200 New York Avenue NW, Washington, DC 20005. The title *Science* is a registered trademark of AAAS.

Copyright © 2020 The Authors, some rights reserved; exclusive licensee American Association for the Advancement of Science. No claim to original U.S. Government Works

PERMISSIONS

<http://www.sciencemag.org/help/reprints-and-permissions>

Use of this article is subject to the [Terms of Service](#)

Science (print ISSN 0036-8075; online ISSN 1095-9203) is published by the American Association for the Advancement of Science, 1200 New York Avenue NW, Washington, DC 20005. The title *Science* is a registered trademark of AAAS.

Copyright © 2020 The Authors, some rights reserved; exclusive licensee American Association for the Advancement of Science. No claim to original U.S. Government Works