Sexual dimorphism in gene expression and regulatory networks across human tissues

Cho-Yi Chen[1,2,†], Camila Lopes-Ramos[1,2,†], Marieke L. Kuijjer[1,2], Joseph N. Paulson[1,2], Abhijeet R. Sonawane[3], Maud Fagny[1,2], John Platig[1,2], Kimberly Glass[1,2,3], John Quackenbush[1,2,3,4,*], Dawn L. DeMeo[3,5,*]

[1]Department of Biostatistics and Computational Biology, Dana-Farber Cancer Institute, Boston, MA 02215, USA
[2]Department of Biostatistics, Harvard School of Public Health, Boston, MA 02215, USA
[3]Channing Division of Network Medicine, Brigham and Women's Hospital, and Harvard Medical School, Boston, MA 02115, USA
[4]Department of Cancer Biology, Dana-Farber Cancer Institute, Boston, MA 02215, USA
[5]Division of Pulmonary and Critical Care Medicine, Brigham and Women's Hospital, Boston, MA 02115, USA

[*]To whom correspondence should be addressed: JQ, johnq@jimmy.harvard.edu; DLD, dawn.demeo@channing.harvard.edu.
[†]These authors contributed equally to this work

Summary

Sexual dimorphism manifests in many diseases and may drive sex-specific therapeutic responses. To understand the molecular basis of sexual dimorphism, we conducted a comprehensive assessment of gene expression and regulatory network modeling in 31 tissues using 8716 human transcriptomes from GTEx. We observed sexually dimorphic patterns of gene expression involving as many as 60% of autosomal genes, depending on the tissue. Interestingly, sex hormone receptors do not exhibit sexually dimorphic expression in most tissues; however, differential network targeting by hormone receptors and other transcription factors (TFs) captures their downstream sexually dimorphic gene expression. Furthermore, differential network wiring was found extensively in several tissues, particularly in brain, in which not all regions exhibit strong differential expression. This systems-based analysis provides a new perspective on the drivers of sexual dimorphism, one in which a repertoire of TFs plays important roles in sex-specific rewiring of gene regulatory networks.

Keywords: Sexual dimorphism, GTEx, transcriptome, gene regulatory network, single-sample network.

Highlights

1.   Sexual dimorphism manifests in both gene expression and gene regulatory networks

2.   Substantial sexual dimorphism in regulatory networks was found in several tissues

3.   Many differentially regulated genes are not differentially expressed

4.   Sex hormone receptors do not exhibit sexually dimorphic expression in most tissues

Introduction

Sexual dimorphism refers to the phenotypic difference between males and females of the same species (Poissant et al., 2010; Williams and Carroll, 2009). These differences present themselves in the characteristics of morphology, physiology, psychology and behavior. Physiological differences primarily exist in sex organs and reproductive systems, but also manifest in other systems, including the musculoskeletal, respiratory, and nervous systems (Blecher and Erickson, 2007; Morris et al., 2004).

Sexual dimorphism is also prevalent in human diseases. A wide range of diseases present differently in females and males, including atherosclerosis, diabetes, osteoporosis, asthma, neuropsychological disorders, and autoimmune diseases (Kaminsky et al., 2006; Morrow, 2015; Ober et al., 2008). For example, systemic lupus erythematosus is an autoimmune disease predominantly occurring in females, in a ratio of 9:1 female-to-male (Lisnevskaia et al., 2014). Differences in incidence, prevalence, severity, and response to treatment between the sexes can complicate our understanding and hinder our ability to cure and prevent diseases. Phenotypic differences between sexes may have a genetic basis (Ober et al., 2008; Poissant et al., 2010; Williams and Carroll, 2009). However, the extent to which genetics drives these effects and the mechanisms responsible are not yet fully understood (Morrow, 2015), partly because the sex-related differences in gene expression in the autosomes are usually subtle and difficult to detect. A systems-based analysis that integrates multi-omics data and collected from a large cohort of research subjects will advance our understanding of the molecular basis of sexual dimorphism.

We conducted a comprehensive expression and network analysis of 8,716 human transcriptomes collected from 28 solid tissues, whole blood, and two cell lines using the Genotype-Tissue Expression (GTEx) data set version 6.0. We compared autosomal gene expression between males and females for each tissue and found that sexual dimorphism in gene expression was seemingly ubiquitous across a wide range of human tissues. However, both the variability in expression and the number of genes involved were tissue-dependent. Gene expression of sex hormone receptors was not differentially expressed between males and females in most tissues, implying that downstream sexually dimorphic regulation is not solely explained by differences in mRNA levels of these receptors. We used a network modeling approach (PANDA+LIONESS) to infer sample-specific gene regulatory networks (Glass et al., 2013; Kuijjer et al., 2015) and compared the networks between males and females. We found significant differences in the structure of gene regulatory networks between males and females in several tissues. This suggests that differences in gene regulatory networks and their associated processes could be a characteristic of sexual dimorphism, and that changes in gene regulatory networks may explain not only differences in tissue expression, but may also help us to understand the mechanism of sexual dimorphism in diseases and other complex traits.

Results
Sexual dimorphism in autosomal gene expression

An overview of this study is shown in Figure 1A. RNA-Seq data from GTEx version 6.0 release were downloaded from dbGaP, preprocessed and normalized as outlined in the supplemental methods and described in more detail in (Paulson et al.). We examined only tissues with samples collected from both men and women, which included 8,716 total samples covering 31 tissues (including 28 solid organ tissues, whole blood, and two derived cell lines) from 549 research subjects (Table S1 and Table S2). Figure 1B shows the demographic information for the 549 subjects included in our analysis.

We used voom (Law et al., 2014) to identify autosomal genes that were differentially expressed (DE) between males and females at a false discovery rate (FDR) less than 0.1 for each tissue (referred to as the sexually dimorphic DE genes) (Materials and Methods; File S1). We also defined a transcriptomic signal-to-noise ratio (tSNR) as a measure of the overall distance between male and female transcriptomes. For the tSNR, the signal was defined as the Euclidean distance of average gene expression profiles between groups (in this case, females and males), and the noise was defined as the overall variation among individuals (Materials and Methods). The tSNR measures the overall divergence of transcriptomes between males and females while the proportion of DE genes focuses on gene-specific sexually dimorphic expression.

An overview of sexually dimorphic gene expression is shown by tissue in Figure 1C. We found the tSNR and the proportion of DE genes are highly correlated (Pearson's r = 0.98) across tissues. We observed sexually dimorphic gene expression across autosomes in most tissues, with 22/31 tissues having tSNRs significantly higher than expected by chance ($P < 0.05$, permutation test) (Table S3). Breast, skin, thyroid, brain, and adipose tissues were the most sexually dimorphic (more than 10% of autosomal genes were DE), whereas the gastrointestinal tract is the least (only two genes were DE).

Tissues that are anatomically close or compositionally similar also demonstrated differing levels of sexually dimorphic expression. For example, subcutaneous adipose tissue exhibited greater sexually dimorphic gene expression than visceral adipose tissue (17% vs 4% of autosomal genes were DE, respectively). In the brain, the cerebral cortex and related structures (Other Brain) were more sexually dimorphic than cerebellum and basal ganglia (14%, 6%, and 2% of autosomal genes were DE, respectively).

We performed a pre-ranked Gene Set Enrichment Analysis (GSEA) based on the weighted t-statistics derived from the differential expression analysis to identify sex-biased enrichment of biological

processes and associations of diseases and phenotypes in the sexually dimorphic gene signature in each tissue (Materials and Methods). We found a great diversity of functional gene sets enriched across tissues (File S2-S3). For example, energy metabolism-related genes were enriched in the adipose tissues, while the neurotransmitter secretion/transport and immune response-related gene sets were enriched in brain (Figure S1).

Additionally, disease and phenotype associations were highly enriched in a tissue-specific manner. For example, in brain, gene sets associated with neurodegenerative disorders and immune-related diseases were enriched in the brain sexually dimorphic gene signature; in heart, gene sets associated with heart block, syncope, ventricular arrhythmia, atrial fibrillation, and palpitations were enriched; in lung, we found enrichment for genes associated with chronic obstructive pulmonary disease and bronchial abnormalities; and in liver, we found ascites and hepatitis-associated gene sets. Additional findings are detailed in Files S2-S3. Overall, our results suggest that the genes with sexually dimorphic expression patterns may be involved in biological processes linked to human diseases and phenotypes, which may help explain sexual dimorphism in development and disease.

A core set of sexually dimorphic genes shared by multiple tissues

We used Fisher's method to combine the results from the sex-specific differential expression analyses to find a common set of genes that share sexually dimorphic expression patterns across multiple tissues (Materials and Methods). We identified a total of 1,568 genes as sexually dimorphic in multiple tissues (FDR < 0.1) (Figure 2A; File S4). Many of the top significant genes, such as *FRG1B, DDX43, SPESP1, NOX5,* and *OOEP*, have been linked to sex-specific functions: *FRG1B* is a putative target gene of human androgen receptor; *SPESP1* encodes sperm equatorial segment protein involved in sperm-egg binding and fusion (Wolkowicz et al., 2008); NOX5 regulates redox-dependent processes in lymphocytes and spermatozoa (El Jamali et al., 2008); OOEP is a component of a subcortical maternal complex which was

found essential during embryonic development in a mouse model (Tashiro et al., 2010). *OOEP*, *DDX43*,

SPESP1, and *NOX5* are all highly expressed in human testis.

The expression patterns of the 1,568 multi-tissue sexually dimorphic genes showed similarities

between anatomically close or compositionally comparable tissue sites (Figure 2A). For example,

subcutaneous adipose tissue was clustered with visceral adipose tissue; different brain regions were also

clustered together. Closer inspection shows that the two adipose sites shared similar sexually dimorphic

gene signature (Figure 2B); this was also observed in the brain regions (Figure 2C). These results suggest

that the sub-regions of the same primary tissue often share similar sexually dimorphic gene expression

patterns for a core set of genes.

We were also interested in whether there were characteristics of sexually dimorphic genes

specific to each of the sexes. We categorized the 1,568 multi-tissue sexually dimorphic genes into

female- (N = 919) or male-biased genes (N = 649) based on their mean t-statistics across all the tissues

(Figure 2D). We found male-biased genes are under higher negative selective pressure (lower dN/dS

ratio) than female-biased genes (P = 0.004, two-sample t-test) (Figure 2E). Both male- and female-biased

genes were overrepresented as human genome-wide association (GWAS) catalog disease genes (female,

P = 1.67 × 10^{-16}; male, P = 7.96 × 10^{-17}) (Figure 2F). Many immunoglobulin genes were female-biased

(Figure 2D & 2G), consistent with previous reports (Bouman et al., 2005; Gonzalez-Quintela et al., 2008;

Verthelyi, 2001). This is consistent with the enrichment of sexually dimorphic DE genes involved in

immune response processes and associated with autoimmune diseases, such as rheumatoid arthritis,

lupus erythematosus, ulcerative colitis, multiple sclerosis, and atopic dermatitis, that we noted

previously.

As reported by others, we found long intergenic noncoding RNAs (lincRNAs) to be sexually

dimorphic (Chen et al., 2012; Mele et al., 2015; Melia et al., 2015). Among the 1,568 multi-tissue

sexually dimorphic genes, 50 (5%) of the female-biased genes and 46 (7%) of male-biased genes encode long intergenic noncoding RNAs (lincRNAs). Several of them were the most strongly tissue-wide sexually dimorphic genes. For example, *RP4-610C12.3* and *RP4-610C12.4* (*LINC01597*) were highly male-biased across tissues; *RP11-575H3.1* was highly female-biased (Figure 2A).

Differential network targeting mediated by estrogen receptors

The gene-expression differences between men and women that we observed likely represent distinct, sex-specific gene regulatory processes (Williams and Carroll, 2009). One of the candidate drivers of these differences could be estrogen. The estrogen receptor (ER) is a ligand-inducible TF that can be activated by estrogen and then regulates the gene expression of a large number of target genes through binding to specific palindromic DNA sequences called estrogen response elements (EREs) (Jin et al., 2004; Welboren et al., 2007). By orchestrating the transcription of target genes, ERs mediate the physiological effects of estrogen in a diverse range of mammalian tissues and thus play an important role in growth, development, reproduction, immunology, mental health, and human disease (Au et al., 2016; Hara et al., 2015).

Given the important role of estrogen and ER in sexual development and reproductive function, we might expect the gene expression of ER to be sexually dimorphic. However, both estrogen receptor genes, estrogen receptor alpha (*ESR1*) and beta (*ESR2*), were not differentially expressed between males and females in most of the tissues we analyzed, including highly sexually dimorphic tissues such as breast, brain, and muscle (Figure S2). This suggests that estrogen may exert its sexually dimorphic effects by altering gene regulatory networks without differential expression of its receptors.

We used PANDA+LIONESS (Glass et al., 2013; Kuijjer et al., 2015) to infer gene regulatory network models for each sample in each tissue and compared those network models between men and women (Figure 1A; Materials and Methods). Briefly, we applied PANDA to create a consensus network

model that comprises 652 TFs and 27,175 target genes by integrating multiple sources of data including gene expression profiles (GTEx), TF binding motifs (CIS-BP) (Weirauch et al., 2014), and protein-protein interactions (StringDB) (Szklarczyk et al., 2015). We then used LIONESS to infer individual networks for each sample in the population via linear interpolation. PANDA+LIONESS produced 8,716 inferred gene regulatory network models, one for each RNA-Seq transcriptome from the 549 GTEx research subjects. These network models represent the strength of the estimated regulatory interaction between TFs and their target genes as weighted bipartite directed graphs, in which nodes represent TFs or their target genes and edges represent inferred regulatory relationships between nodes.

Analogous to how we identified sexually dimorphic DE genes, we used limma to compare the weight of each edge between males and females in each tissue and identified "differentially targeting" edges, which represent a set of TF-target relationships that exhibit differential regulatory potential between the sexes (File S5; Materials and Methods). This allowed us to identify regulatory network "wiring patterns" that differed between males and females in each tissue and to identify key regulators that may contribute to the shaping of the sexually dimorphic gene expression landscape in each tissue.

We started by investigating estrogen receptors in the breast tissue. Although the estrogen receptors *ESR1* and *ESR2* were not differentially expressed between males and females, we found that the regulatory targeting by these receptors was different between the sexes. We highlight two of the most sexually dimorphic DE genes that are targets of ESR1 in our network models, *IGLV4-69* and *MYOT*, as examples. *IGLV4-69* is an immunoglobulin gene that had increased expression in females yet increased targeting by ESR1 in males (Figure 3A). In contrast, *MYOT*, a gene that encodes a cytoskeletal protein (myotilin) that stabilizes thin filaments during muscle contraction, showed over-expression in males but greater ESR1-targeting in females (Figure 3B). In both cases, edge weights connecting ESR1 to its target genes linearly reflected the expression of these target genes (Figure 3C-D, upper). However,

ESR1 itself was not differentially expressed between males and females, and its expression poorly

correlated with its target gene expression (Figure 3C-D, bottom). This implies that the edge weights in

the network models capture the regulatory activities of ESR1 while the gene expression of *ESR1* alone

does not.

By considering the putative set of ESR1 target genes in the breast tissue, we found that the

sexually dimorphic differential targeting by ESR1 explains much of the variance (adjusted R^2 = 0.88) in

the differential expression of its targets (Figure 3E), suggesting that the regulatory influence of ESR1 may

drive the sexually dimorphic gene expression of its target genes. We note that ESR2 shows a similar

regulatory pattern in the regulatory networks in the breast tissue (adjusted R^2 = 0.81) (Figure S3).

Figure 3F shows the 20 genes most differentially targeted by ESR1 in the breast network models.

Several of these genes are implicated in breast function and disease. For example, *MST1R,* encodes a

macrophage stimulating receptor associated breast cancer progression (Privette Vinnedge et al., 2015).

ATP6V0A4 mediates invasion of breast cancer cell lines in the transwell Matrigel assay (Hinton et al.,

2009). *EDN3* has been suggested as a tumor suppressor gene in the human mammary gland (Wiesmann

et al., 2009) and its expression was shown to be regulated by estrogen and progesterone in the rhesus

macaque (Keator et al., 2011). *LTF* (Lactotransferrin) is an estrogen-inducible gene (Das et al., 1998;

Ghosh et al., 1999; Moggs et al., 2004; Stuckey et al., 2006); its protein product exists in high

concentration in human milk (Hakansson et al., 1995).

A repertoire of TFs modulate sexually dimorphic gene expression

To investigate whether the targeting of genes by ESRs is equally important in tissues other than

breast, and to investigate if other TFs also drive sexually dimorphic gene expression, we extended our

analysis to include all 652 TFs and analyzed them in each tissue. We found that ESR1 influences gene

expression differently across tissues. In subcutaneous adipose and thyroid, for example, the differential

targeting of ESR1 largely explains the sexually dimorphic gene expression of its target genes (adjusted R^2 = 0.78 and 0.59, respectively) (Figure 4A-B), whereas in colon the relationship was relatively poor (adjusted R^2 = 0.18) (Figure 4C).

By ranking all 652 TFs in our network models based on how well their differential targeting explains the variance in the sexually dimorphic differential expression patterns of their target (adjusted R^2), we found that the highest-ranking TFs in each tissue are not always hormone receptors (Figure 4D; File S6). However, the top TFs often have documented roles in hormone-responsive pathways and sex-related biological processes and diseases (Table 1). These results suggest that a wide variety of TFs mediate the sexually dimorphic regulatory programs active in different tissues.

Brain exhibits extensive differential network wiring

Sex differences in gene regulatory networks involve both TFs and target genes. To identify differentially targeted genes between males and females, we calculated the in-degree for each gene as the sum of its network edge weights and compared in-degrees between males and females. As shown in Figure 5A, we found a large proportion of genes were both DE and differentially targeted (DT) in several tissues including breast, thyroid, muscle, and liver.

Importantly, we also found that a surprisingly large number of genes in brain, especially in basal ganglia, are differentially targeted but not differentially expressed (6,479 DT genes at FDR < 0.1). These DT-only genes in basal ganglia are associated with brain functions, with enrichment in synaptic and behavior-related biological processes (Figure 5B). For instance, *CACNA1A*, which encodes a voltage-dependent calcium channel, has been associated with multiple diseases such as episodic ataxia (Tomlinson et al., 2016), epileptic encephalopathy (Damaj et al., 2015), and autism (Damaj et al., 2015; Li et al., 2015). As an example, we found that there are 185 TFs differentially targeting 128 genes associated with synaptic transmission process in basal ganglia between males and females (Figure 5C-D).

These results reveal an additional layer of complexity of sexual dimorphism, in which differential network wiring also characterizes sex-specific gene regulatory networks.

Discussion

Sexual dimorphism manifests in human development, physiology, and in the incidence and progression of diseases. Sexual dimorphism influences how the genetic program drives the phenotypes we observe and the ways in which those phenotypes differ between the sexes. A natural hypothesis is that the phenotypic differences between men and women in different tissues should be reflected in sex-specific gene expression in these tissues. However, a global picture of sexual dimorphism in gene expression is difficult to depict, partly because the sex-related differences in the autosomal gene expression are usually subtle and varied across tissues.

One strategy to tackle this problem is to include more samples in the analysis to increase statistical power. Our analysis was based on the GTEx version 6.0 cohort, which includes 8,716 samples from 549 individuals; we excluded the sex chromosome genes so that they would not skew estimates of sexually dimorphic expression. This large sample size enabled us to detect subtle differences in autosomal gene expression at a higher resolution than in previous studies. We discovered many tissue-specific sexually dimorphic differentially expressed genes on the autosomes, many of which have important known biological and disease-related functions.

Mele and colleagues (Mele et al., 2015) have also reported sex-biased gene expression across human tissues based on a relatively small GTEx pilot data set which included only 1,641 postmortem tissue samples from 175 individuals. They included both autosomal and sex chromosome genes and found that most sexually dimorphic DE genes were on sex chromosomes. The only exception was breast, in which 715 differentially expressed autosomal genes were found. Then they used WGCNA to identify modules of correlated gene expression and showed that those co-expression modules were enriched

with GO terms that included spermatid, ectoderm, and epidermis development. However, WGCNA is a correlation-based method that captures general patterns of co-expression but does not distinguish between transcription factors and their targets, and so may fail to extract sex-specific regulatory processes.

The emergence of sex-specific traits likely involves modifications in sexually dimorphic gene expression patterns, which may be driven by sex-dependent gene regulation (Blekhman et al., 2010; Ober et al., 2008; Williams and Carroll, 2009). Ligand-dependent transcriptional regulation by steroid sex hormones and their receptors was suggested to play a role in sexual development and reproductive function by mediating downstream target gene expression with other TFs (Bouman et al., 2005; Ikeda et al., 2015; Verthelyi, 2001). To our surprise, we did not observe significant differences in gene expression of steroid hormone receptors between men and women in most of the analyzed tissues.

A hypothesis is that sexual dimorphism is a manifestation of different regulatory programs in male and female tissues. We used PANDA and LIONESS, a systems-based method that uses a prior based on TF binding motifs and integrates multi-omics data to model regulatory processes in both populations and individuals. We discovered that male and female networks have significant differences in their regulatory structure, even when the targets of specific TFs are not differentially expressed. This is a subtle but important point as tissue-specific patterns of gene expression may be maintained differentially in males and females. This means that sexual dimorphism may result from perturbations that affect the regulatory processes in one sex rather than the other, altering the expression of the downstream genes.

As an example, we presented an analysis of sexual dimorphism in the human brain. We found that the structure of gene regulatory networks differed between males and females. Differentially targeted genes in brain were enriched in biological functions and pathways relevant to brain diseases,

many of which are known to exhibit differences in men and women. Indeed, sex differences have been found in brain structure, cognitive functions, behaviors, and in several brain diseases and mental disorders, including Parkinson's and Alzheimer's disease, drug abuse, anxiety, and depression (Frick and Gresack, 2003; Gillies and McArthur, 2010; McCarthy, 2008; Pol et al., 2006; Rizk-Jackson et al., 2006; Tsai et al., 2009).

Our systems-based approach to gene regulatory network modeling provides a new perspective on sexual dimorphism, one in which steroid hormones and non-steroid transcription factors affect the regulatory program by altering the structure of sex-specific gene regulatory networks. These sex-specific processes often regulate genes associated with development and disease, and may help to explain sexual dimorphism we observe. These findings across 31 tissues reiterate a phenomenon that we had observed in the study of male and female differences in chronic obstructive pulmonary disease (Glass et al., 2014).

Sexual dimorphism has been recognized as one of the most significantly understudied aspects of human disease research and the 2015 guidelines from the National Institutes of Health (NOT-OD-15-102) have identified consideration of sex as biological variable should be factored into research designs, analyses, and reporting in vertebrate animal and human studies (Health, 2015). The GTEx data provide an unprecedented opportunity to explore how sex influences gene expression. We see great diversity in the degree of sexually dimorphic gene expression, with some tissues appearing to be very different while others are almost indistinguishable. But by using PANDA and LIONESS to model regulatory processes, we find higher order differences in regulatory processes that distinguish men and women. Our analysis of 31 tissues provides an important baseline for any disease-based studies and underscores the importance of looking beyond expression to consider sexually dimorphic patterns of transcriptional regulation when studying human development and disease.

Experimental procedures

More details are available in Supplemental Experimental Procedures.

GTEx data set

The Genotype-Tissue Expression (GTEx) version 6.0 RNA-Seq data set (phs000424.v6.p1, 2015-10-05 released) was downloaded from dbGaP (approved protocol #9112). Using YARN in Bioconductor [bioconductor.org/packages/yarn], we performed quality control, gene filtering, and normalization preprocessing (Paulson et al.). We identified and removed GTEX-11ILO due to potential sex misannotation. We grouped related body regions using gene expression similarity. For example, skin samples from the lower leg (sun exposed) and from the suprapubic region (sun unexposed) were grouped as "skin." We removed sex-specific tissues (prostate, testis, uterus, vagina, and ovary) and those with fewer than 30 samples of one sex (kidney cortex and minor salivary gland). The final data set contains 8716 samples from 31 tissues (which included 28 solid organ tissues, whole blood, and two derived cell lines) from 549 research subjects (188 females and 361 males) (Table S1 and Table S2). We removed sex-chromosome and mitochondrial genes (retaining 29,242 genes).

Differential expression analysis

Differential expression (DE) analysis was performed using voom (Law et al., 2014) to transform RNA-Seq read counts to log counts per million (log-cpm) with associated precision weights, followed by linear modeling and empirical Bayes procedure using limma. In each tissue, we adopted the following linear regression model to detect sexually dimorphic gene expression:

$$Y \sim \beta_0 + \beta_1 Batch + \beta_2 Race + \beta_3 Age + \beta_4 BMI + \beta_5 Sex + \varepsilon$$

where Y is the gene expression; *Batch* denotes the type of nucleic acid isolation; TrueSeq RNA library preparation kit v1 was used in all GTEx RNA-Seq experiments; *Race* denotes the race of the subject; *Age* denotes the age of the subject, *BMI* denotes the body mass index of the subject, and *Sex* denotes the

reported sex of the subject. The linear model is fitted to the expression data for each gene using a least-squares fitting method. The p-values for the estimated coefficient of *Sex* were adjusted for multiple testing using the Benjamini-Hochberg method. A false discovery rate of 0.1 was used as the significance threshold to report sexually dimorphic DE genes in each tissue. To identify differential expression across multiple tissues, we used Fisher's method (Fisher, 1970) to combine the adjusted p-values from the results of differential expression analysis in each of the GTEx tissues.

Transcriptomic signal-to-noise ratio

Transcriptomic signal-to-noise ratio (tSNR) between expression profiles of females (*X*) and males (*Y*) was calculated as:

$$tSNR(X, Y) = \frac{\|\bar{X} - \bar{Y}\|_2}{\sqrt{\frac{\sigma_X^2}{F} + \frac{\sigma_Y^2}{M}}}$$

$$\sigma_X^2 = \frac{\sum_{i=1}^{F} \|x_i - \bar{X}\|_2^2}{F - 1}$$

$$\sigma_Y^2 = \frac{\sum_{i=1}^{M} \|y_i - \bar{Y}\|_2^2}{M - 1}$$

where *F* and *M* represent the sample sizes of females and males, respectively. To derive an empirical p-value for each observed tSNR in each tissue, we used a permutation test that shuffled the sex labels to samples while maintaining the total numbers of females and males unchanged for 1,000 iterations. If there was zero occurrence from the 1,000 times permutation more extreme than the observed tSNR, the empirical p-value would be set to 0.5/1,000.

Functional enrichment analysis

Gene Set Enrichment Analysis (GSEA) was performed based on the pre-ranked list of *t*-statistics derived from the differential expression analysis (Subramanian et al., 2005). Figure S1 was visualized by using Enrichment map (Isserlin et al., 2014) in Cytoscape (Smoot et al., 2011). If not stated explicitly otherwise,

15

default parameters and a default background set were used when performing the analyses. Gene

annotations were obtained from the Gene Ontology Consortium (Gene Ontology, 2015), FunDO (Du et

al., 2009), and Human Phenotype Ontology (HPO). Annotation files were compiled into the Gene Matrix

Transposed (GMT) file format and were loaded to GSEA program (v2-2.2.2).

Network inference using PANDA+LIONESS

Single-sample networks were reconstructed using PANDA+LIONESS (Glass et al., 2013). We performed a

motif scan to generate a gene regulatory prior. We downloaded position weight matrices (PWM) for

Homo sapiens motifs from the Catalog of Inferred Sequence Binding Preferences (CIS-BP) (Weirauch et

al., 2014). We mapped the PWMs to promoter regions of Ensembl genes (GRCh37.p13) using FIMO

(Grant et al., 2011). Motif mappings were parsed to only retain those below p-value cutoff of 10^{-5} and

those that were within a promoter range of [-750, +250] around the transcription start site (TSS). This

resulted in a regulatory prior of 652 TFs targeting 27,249 target ENSG identifiers. We generated a

protein-protein interaction (PPI) prior based on the *Homo sapiens* PPI data set from StringDB version 10.

We then parsed our gene expression data to match the genes in the motif prior (27,175 genes) and ran

PANDA followed by LIONESS.

Author contributions

CYC, CLR, MK, JNP, AS, MF, JP, KG, JQ, and DLD conceived the study; CYC, CLR, MK, and JNP analyzed the

data; CYC, CLR, KG, JQ, and DLD interpreted the results. All authors contributed to the reviewing and

editing of the manuscript. All authors read and approved the final manuscript.

Acknowledgments

Table

Table 1. TFs that may drive downstream sexually dimorphic gene expression across tissues

TF	Gene name	Documented roles	References
BRCA1	Breast Cancer 1	Responsive to estrogen; regulating acetylation and ubiquitination of ESR1; implicated in breast cancer	(Chrzan and Bradford, 2007; Crowe and Lee, 2006; Ma et al., 2010; Wang et al., 2005)
CEBPD/ CEBPE	CCAAT/enhancer binding protein, delta/epsilon	Responsive to estradiol	(Buterin et al., 2006; Tee et al., 2004)
FOXP1/F OXP4	Forkhead Box P1/P4	Co-expressed with estrogen receptors; estrogen-inducible; implicated in breast and ovarian cancers; important in heart development	(Fox et al., 2004; Ijichi et al., 2013; Ikeda et al., 2015; Shigekawa et al., 2011)
GMEB2	Glucocorticoid Modulatory Element Binding Protein 2	Responsive to ethinyl estradiol	(Fong et al., 2007)
KLF4	Kruppel-Like Factor 4	Yamanaka factor (iPS factor) required for skin development; suppressing transcriptional activity of ESR1	(Akaogi et al., 2009; Rowland and Peeper, 2006; Takahashi et al., 2007)
NR2F6	Nuclear Receptor Subfamily 2 Group F Member 6	Involved in modulation of hormonal responses and development of brain circadian clock; suppressing transcriptional activity of ESR1 induced by estrogens	(Warnecke et al., 2005; Zhu et al., 2000)
SP1/SP3	Sp1/Sp3 Transcription Factor	Involved in liver development; engaged in estrogen signaling pathway via PPI with ER-ligand complexes	(Kazi et al., 2005; Moggs and Orphanides, 2001; Mukherjee et al., 2005)
TET1	Tet Methylcytosine Dioxygenase 1	Involved in DNA methylation processes; implicated in lateral myocardial infarction and coronary aneurysm	(Li et al., 2013; Rovai et al., 2016; Takai et al., 2014)
TOPORS	TOP1 Binding Arginine/Serine-Rich Protein	Binding to Parkinson's disease-associated protein (DJ-1) and protecting against neuronal apoptosis by regulating p53 transcriptional activity	(Shinbo et al., 2005; Xu et al., 2005)

References

Akaogi, K., Nakajima, Y., Ito, I., Kawasaki, S., Oie, S.H., Murayama, A., Kimura, K., and Yanagisawa, J. (2009). KLF4 suppresses estrogen-dependent breast cancer growth by inhibiting the transcriptional activity of ERalpha. Oncogene *28*, 2894-2902.

Au, A., Feher, A., McPhee, L., Jessa, A., Oh, S., and Einstein, G. (2016). Estrogens, inflammation and cognition. Frontiers in neuroendocrinology *40*, 87-100.

Blecher, S.R., and Erickson, R.P. (2007). Genetics of sexual development: a new paradigm. Am J Med Genet A *143A*, 3054-3068.

Blekhman, R., Marioni, J.C., Zumbo, P., Stephens, M., and Gilad, Y. (2010). Sex-specific and lineage-specific alternative splicing in primates. Genome Research *20*, 180-189.

Bouman, A., Heineman, M.J., and Faas, M.M. (2005). Sex hormones and the immune response in humans. Human reproduction update *11*, 411-423.

Buterin, T., Koch, C., and Naegeli, H. (2006). Convergent transcriptional profiles induced by endogenous estrogen and distinct xenoestrogens in breast cancer cells. Carcinogenesis *27*, 1567-1578.

Chen, H., Palmer, J.S., Thiagarajan, R.D., Dinger, M.E., Lesieur, E., Chiu, H., Schulz, A., Spiller, C., Grimmond, S.M., Little, M.H., *et al.* (2012). Identification of novel markers of mouse fetal ovary development. PLoS One *7*, e41683.

Chrzan, B.G., and Bradford, P.G. (2007). Phytoestrogens activate estrogen receptor beta1 and estrogenic responses in human breast and bone cancer cell lines. Mol Nutr Food Res *51*, 171-177.

Crowe, D.L., and Lee, M.K. (2006). New role for nuclear hormone receptors and coactivators in regulation of BRCA1-mediated DNA repair in breast cancer cell lines. Breast Cancer Res *8*, R1.

Damaj, L., Lupien-Meilleur, A., Lortie, A., Riou, E., Ospina, L.H., Gagnon, L., Vanasse, C., and Rossignol, E. (2015). CACNA1A haploinsufficiency causes cognitive impairment, autism and epileptic encephalopathy with mild cerebellar symptoms. Eur J Hum Genet *23*, 1505-1512.

Das, S.K., Tan, J., Johnson, D.C., and Dey, S.K. (1998). Differential spatiotemporal regulation of lactoferrin and progesterone receptor genes in the mouse uterus by primary estrogen, catechol estrogen, and xenoestrogen. Endocrinology *139*, 2905-2915.

Du, P., Feng, G., Flatow, J., Song, J., Holko, M., Kibbe, W.A., and Lin, S.M. (2009). From disease ontology to disease-ontology lite: statistical methods to adapt a general-purpose ontology for the test of gene-ontology associations. Bioinformatics *25*, i63-68.

El Jamali, A., Valente, A.J., Lechleiter, J.D., Gamez, M.J., Pearson, D.W., Nauseef, W.M., and Clark, R.A. (2008). Novel redox-dependent regulation of NOX5 by the tyrosine kinase c-Abl. Free Radic Biol Med *44*, 868-881.

Fisher, R.A. (1970). Statistical methods for research workers, 14 edn (Edinburgh, Scotland: Oliver and Boyd).

Fong, C.J., Burgoon, L.D., Williams, K.J., Forgacs, A.L., and Zacharewski, T.R. (2007). Comparative temporal and dose-dependent morphological and transcriptional uterine effects elicited by tamoxifen and ethynylestradiol in immature, ovariectomized mice. BMC Genomics *8*, 151.

Fox, S.B., Brown, P., Han, C., Ashe, S., Leek, R.D., Harris, A.L., and Banham, A.H. (2004). Expression of the forkhead transcription factor FOXP1 is associated with estrogen receptor alpha and improved survival in primary human breast carcinomas. Clin Cancer Res *10*, 3521-3527.

Frick, K.M., and Gresack, J.E. (2003). Sex differences in the behavioral response to spatial and object novelty in adult C57BL/6 mice. Behavioral neuroscience *117*, 1283-1291.

Gene Ontology, C. (2015). Gene Ontology Consortium: going forward. Nucleic Acids Res *43*, D1049-1056.

Ghosh, D., Taylor, J.A., Green, J.A., and Lubahn, D.B. (1999). Methoxychlor stimulates estrogen-responsive messenger ribonucleic acids in mouse uterus through a non-estrogen receptor (non-ER) alpha and non-ER beta mechanism. Endocrinology *140*, 3526-3533.

Gillies, G.E., and McArthur, S. (2010). Estrogen actions in the brain and the basis for differential action in men and women: a case for sex-specific medicines. Pharmacol Rev *62*, 155-198.

Glass, K., Huttenhower, C., Quackenbush, J., and Yuan, G.C. (2013). Passing messages between biological networks to refine predicted interactions. PLoS One *8*, e64832.

Glass, K., Quackenbush, J., Silverman, E.K., Celli, B., Rennard, S.I., Yuan, G.C., and DeMeo, D.L. (2014). Sexually-dimorphic targeting of functionally-related genes in COPD. BMC Syst Biol *8*, 118.

Gonzalez-Quintela, A., Alende, R., Gude, F., Campos, J., Rey, J., Meijide, L.M., Fernandez-Merino, C., and Vidal, C. (2008). Serum levels of immunoglobulins (IgG, IgA, IgM) in a general adult population and their relationship with alcohol consumption, smoking and common metabolic abnormalities. Clinical and experimental immunology *151*, 42-50.

Grant, C.E., Bailey, T.L., and Noble, W.S. (2011). FIMO: scanning for occurrences of a given motif. Bioinformatics *27*, 1017-1018.

Hakansson, A., Zhivotovsky, B., Orrenius, S., Sabharwal, H., and Svanborg, C. (1995). Apoptosis induced by a human milk protein. Proc Natl Acad Sci U S A *92*, 8064-8068.

Hara, Y., Waters, E.M., McEwen, B.S., and Morrison, J.H. (2015). Estrogen Effects on Cognitive and Synaptic Health Over the Lifecourse. Physiological reviews *95*, 785-807.

National Institute of Health (2015). Consideration of Sex as a Biological Variable in NIH-funded Research (Notice NOT-OD-15–102).

Hinton, A., Sennoune, S.R., Bond, S., Fang, M., Reuveni, M., Sahagian, G.G., Jay, D., Martinez-Zaguilan, R., and Forgac, M. (2009). Function of a subunit isoforms of the V-ATPase in pH homeostasis and in vitro invasion of MDA-MB231 human breast cancer cells. J Biol Chem *284*, 16400-16408.

Ijichi, N., Ikeda, K., Horie-Inoue, K., and Inoue, S. (2013). FOXP1 and estrogen signaling in breast cancer. Vitam Horm *93*, 203-212.

Ikeda, K., Horie-Inoue, K., and Inoue, S. (2015). Identification of estrogen-responsive genes based on the DNA binding properties of estrogen receptors using high-throughput sequencing technology. Acta pharmacologica Sinica *36*, 24-31.

Isserlin, R., Merico, D., Voisin, V., and Bader, G.D. (2014). Enrichment Map - a Cytoscape app to visualize and explore OMICs pathway enrichment results. F1000Research *3*, 141.

Jin, V.X., Leu, Y.W., Liyanarachchi, S., Sun, H., Fan, M., Nephew, K.P., Huang, T.H., and Davuluri, R.V. (2004). Identifying estrogen receptor alpha target genes using integrated computational genomics and chromatin immunoprecipitation microarray. Nucleic Acids Res *32*, 6627-6635.

Kaminsky, Z., Wang, S.C., and Petronis, A. (2006). Complex disease, gender and epigenetics. Annals of medicine *38*, 530-544.

Kazi, A.A., Jones, J.M., and Koos, R.D. (2005). Chromatin immunoprecipitation analysis of gene expression in the rat uterus in vivo: estrogen-induced recruitment of both estrogen receptor alpha and hypoxia-inducible factor 1 to the vascular endothelial growth factor promoter. Mol Endocrinol *19*, 2006-2019.

Keator, C.S., Mah, K., Ohm, L., and Slayden, O.D. (2011). Estrogen and progesterone regulate expression of the endothelins in the rhesus macaque endometrium. Human reproduction *26*, 1715-1728.

Kuijjer, M.L., Tung, M., Yuan, G., Quackenbush, J., and Glass, K. (2015). Estimating sample-specific regulatory networks. arXiv preprint arXiv:150506440.

Law, C.W., Chen, Y., Shi, W., and Smyth, G.K. (2014). voom: Precision weights unlock linear model analysis tools for RNA-seq read counts. Genome Biol *15*, R29.

Li, D., Guo, J., Han, Y., Zhang, Z., Yu, X., Cai, H., Zhang, L., Cui, Q., and Yu, T. (2013). Dual left anterior descending artery arising from left circumflex artery with an intramyocardial coronary aneurysm. The Canadian journal of cardiology *29*, 1742 e1721-1743.

Li, J., You, Y., Yue, W., Jia, M., Yu, H., Lu, T., Wu, Z., Ruan, Y., Wang, L., and Zhang, D. (2015). Genetic Evidence for Possible Involvement of the Calcium Channel Gene CACNA1A in Autism Pathogenesis in Chinese Han Population. PLoS One *10*, e0142887.

Lisnevskaia, L., Murphy, G., and Isenberg, D. (2014). Systemic lupus erythematosus. Lancet *384*, 1878-1888.

Ma, Y., Fan, S., Hu, C., Meng, Q., Fuqua, S.A., Pestell, R.G., Tomita, Y.A., and Rosen, E.M. (2010). BRCA1 regulates acetylation and ubiquitination of estrogen receptor-alpha. Mol Endocrinol *24*, 76-90.

McCarthy, M.M. (2008). Estradiol and the developing brain. Physiological reviews *88*, 91-124.

Mele, M., Ferreira, P.G., Reverter, F., DeLuca, D.S., Monlong, J., Sammeth, M., Young, T.R., Goldmann, J.M., Pervouchine, D.D., Sullivan, T.J., *et al.* (2015). Human genomics. The human transcriptome across tissues and individuals. Science *348*, 660-665.

Melia, T., Hao, P., Yilmaz, F., and Waxman, D.J. (2015). Hepatic Long Intergenic Noncoding RNAs: High Promoter Conservation and Dynamic, Sex-Dependent Transcriptional Regulation by Growth Hormone. Mol Cell Biol *36*, 50-69.

Moggs, J.G., and Orphanides, G. (2001). Estrogen receptors: orchestrators of pleiotropic cellular responses. EMBO reports *2*, 775-781.

Moggs, J.G., Tinwell, H., Spurway, T., Chang, H.S., Pate, I., Lim, F.L., Moore, D.J., Soames, A., Stuckey, R., Currie, R., *et al.* (2004). Phenotypic anchoring of gene expression changes during estrogen-induced uterine growth. Environmental health perspectives *112*, 1589-1606.

Morris, J.A., Jordan, C.L., and Breedlove, S.M. (2004). Sexual differentiation of the vertebrate nervous system. Nature neuroscience *7*, 1034-1039.

Morrow, E.H. (2015). The evolution of sex differences in disease. Biology of sex differences *6*, 5.

Mukherjee, T.K., Reynolds, P.R., and Hoidal, J.R. (2005). Differential effect of estrogen receptor alpha and beta agonists on the receptor for advanced glycation end product expression in human microvascular endothelial cells. Biochim Biophys Acta *1745*, 300-309.

Ober, C., Loisel, D.A., and Gilad, Y. (2008). Sex-specific genetic architecture of human disease. Nat Rev Genet *9*, 911-922.

Paulson, J.N., Chen, C.-Y., Fagny, M., Kuijjer, M., Lopes-Ramos, C., Platig, J., Sonawane, A., Glass, K., and Quackenbush, J. Sparsity-Aware RNA-Seq Processing and Normalization for Heterogeneous Data. (manuscript accompanied submission).

Poissant, J., Wilson, A.J., and Coltman, D.W. (2010). Sex-specific genetic variance and the evolution of sexual dimorphism: a systematic review of cross-sex genetic correlations. Evolution; international journal of organic evolution *64*, 97-107.

Pol, H.E.H., Cohen-Kettenis, P.T., Van Haren, N.E.M., Peper, J.S., Brans, R.G.H., Cahn, W., Schnack, H.G., Gooren, L.J.G., and Kahn, R.S. (2006). Changing your sex changes your brain: influences of testosterone and estrogen on adult human brain structure. European Journal of Endocrinology *155*, S107-S114.

Privette Vinnedge, L.M., Benight, N.M., Wagh, P.K., Pease, N.A., Nashu, M.A., Serrano-Lopez, J., Adams, A.K., Cancelas, J.A., Waltz, S.E., and Wells, S.I. (2015). The DEK oncogene promotes cellular proliferation through paracrine Wnt signaling in Ron receptor-positive breast cancers. Oncogene *34*, 2325-2336.

Rizk-Jackson, A.M., Acevedo, S.F., Inman, D., Howieson, D., Benice, T.S., and Raber, J. (2006). Effects of sex on object recognition and spatial navigation in humans. Behav Brain Res *173*, 181-190.

Rovai, D., Gimelli, A., Coceani, M., Sbrana, F., Masini, G., and Rossi, G. (2016). T wave abnormalities identify patients with previous lateral wall myocardial infarction and circumflex artery disease. Journal of electrocardiology *49*, 216-222.

Rowland, B.D., and Peeper, D.S. (2006). KLF4, p21 and context-dependent opposing forces in cancer. Nat Rev Cancer *6*, 11-23.

Shigekawa, T., Ijichi, N., Ikeda, K., Horie-Inoue, K., Shimizu, C., Saji, S., Aogi, K., Tsuda, H., Osaki, A., Saeki, T., *et al.* (2011). FOXP1, an estrogen-inducible transcription factor, modulates cell proliferation in breast cancer cells and 5-year recurrence-free survival of patients with tamoxifen-treated breast cancer. Horm Cancer *2*, 286-297.

Shinbo, Y., Taira, T., Niki, T., Iguchi-Ariga, S.M., and Ariga, H. (2005). DJ-1 restores p53 transcription activity inhibited by Topors/p53BP3. Int J Oncol *26*, 641-648.

Smoot, M.E., Ono, K., Ruscheinski, J., Wang, P.L., and Ideker, T. (2011). Cytoscape 2.8: new features for data integration and network visualization. Bioinformatics *27*, 431-432.

Stuckey, R., Aldridge, T., Lim, F.L., Moore, D.J., Tinwell, H., Doherty, N., Davies, R., Smith, A.G., Kimber, I., Ashby, J., *et al.* (2006). Induction of iron homeostasis genes during estrogen-induced uterine growth and differentiation. Molecular and cellular endocrinology *253*, 22-29.

Subramanian, A., Tamayo, P., Mootha, V.K., Mukherjee, S., Ebert, B.L., Gillette, M.A., Paulovich, A., Pomeroy, S.L., Golub, T.R., Lander, E.S., *et al.* (2005). Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. Proc Natl Acad Sci U S A *102*, 15545-15550.

Szklarczyk, D., Franceschini, A., Wyder, S., Forslund, K., Heller, D., Huerta-Cepas, J., Simonovic, M., Roth, A., Santos, A., Tsafou, K.P., *et al.* (2015). STRING v10: protein-protein interaction networks, integrated over the tree of life. Nucleic Acids Res *43*, D447-452.

Takahashi, K., Tanabe, K., Ohnuki, M., Narita, M., Ichisaka, T., Tomoda, K., and Yamanaka, S. (2007). Induction of pluripotent stem cells from adult human fibroblasts by defined factors. Cell *131*, 861-872.

Takai, H., Masuda, K., Sato, T., Sakaguchi, Y., Suzuki, T., Suzuki, T., Koyama-Nasu, R., Nasu-Nishimura, Y., Katou, Y., Ogawa, H., *et al.* (2014). 5-Hydroxymethylcytosine plays a critical role in glioblastomagenesis by recruiting the CHTOP-methylosome complex. Cell reports *9*, 48-60.

Tashiro, F., Kanai-Azuma, M., Miyazaki, S., Kato, M., Tanaka, T., Toyoda, S., Yamato, E., Kawakami, H., Miyazaki, T., and Miyazaki, J. (2010). Maternal-effect gene Ces5/Ooep/Moep19/Floped is essential for oocyte cytoplasmic lattice formation and embryonic development at the maternal-zygotic stage transition. Genes to cells : devoted to molecular & cellular mechanisms *15*, 813-828.

Tee, M.K., Rogatsky, I., Tzagarakis-Foster, C., Cvoro, A., An, J., Christy, R.J., Yamamoto, K.R., and Leitman, D.C. (2004). Estradiol and selective estrogen receptor modulators differentially regulate target genes with estrogen receptors alpha and beta. Molecular biology of the cell *15*, 1262-1272.

Tomlinson, S.E., Tan, S.V., Burke, D., Labrum, R.W., Haworth, A., Gibbons, V.S., Sweeney, M.G., Griggs, R.C., Kullmann, D.M., Bostock, H., *et al.* (2016). In vivo impact of presynaptic calcium channel dysfunction on motor axons in episodic ataxia type 2. Brain *139*, 380-391.

Tsai, H.W., Grant, P.A., and Rissman, E.F. (2009). Sex differences in histone modifications in the neonatal mouse brain. Epigenetics *4*, 47-53.

Verthelyi, D. (2001). Sex hormones as immunomodulators in health and disease. Int Immunopharmacol *1*, 983-993.

Wang, C., Fan, S., Li, Z., Fu, M., Rao, M., Ma, Y., Lisanti, M.P., Albanese, C., Katzenellenbogen, B.S., Kushner, P.J., *et al.* (2005). Cyclin D1 antagonizes BRCA1 repression of estrogen receptor alpha activity. Cancer Res *65*, 6557-6567.

Warnecke, M., Oster, H., Revelli, J.P., Alvarez-Bolado, G., and Eichele, G. (2005). Abnormal development of the locus coeruleus in Ear2(Nr2f6)-deficient mice impairs the functionality of the forebrain clock and affects nociception. Genes Dev *19*, 614-625.

Weirauch, M.T., Yang, A., Albu, M., Cote, A.G., Montenegro-Montero, A., Drewe, P., Najafabadi, H.S., Lambert, S.A., Mann, I., Cook, K., *et al.* (2014). Determination and inference of eukaryotic transcription factor sequence specificity. Cell *158*, 1431-1443.

Welboren, W.J., Stunnenberg, H.G., Sweep, F.C., and Span, P.N. (2007). Identifying estrogen receptor target genes. Molecular oncology *1*, 138-143.

Wiesmann, F., Veeck, J., Galm, O., Hartmann, A., Esteller, M., Knuchel, R., and Dahl, E. (2009). Frequent loss of endothelin-3 (EDN3) expression due to epigenetic inactivation in human breast cancer. Breast Cancer Res *11*, R34.

Williams, T.M., and Carroll, S.B. (2009). Genetic and molecular insights into the development and evolution of sexual dimorphism. Nat Rev Genet *10*, 797-804.

Wolkowicz, M.J., Digilio, L., Klotz, K., Shetty, J., Flickinger, C.J., and Herr, J.C. (2008). Equatorial segment protein (ESP) is a human alloantigen involved in sperm-egg binding and fusion. Journal of andrology *29*, 272-282.

Xu, J., Zhong, N., Wang, H., Elias, J.E., Kim, C.Y., Woldman, I., Pifl, C., Gygi, S.P., Geula, C., and Yankner, B.A. (2005). The Parkinson's disease-associated DJ-1 protein is a transcriptional co-activator that protects against neuronal apoptosis. Human molecular genetics *14*, 1231-1241.

Zhu, X.G., Park, K.S., Kaneshige, M., Bhat, M.K., Zhu, Q., Mariash, C.N., McPhie, P., and Cheng, S.Y. (2000). The orphan nuclear receptor Ear-2 is a negative coregulator for thyroid hormone nuclear receptor function. Mol Cell Biol *20*, 2604-2618.

Figure legends

Figure 1. Overview of this study. (**A**) The study workflow. The GTEx v6 data set was used and annotated based on GENCODE v19. (**B**) Demographics of the 549 study subjects. The "Others" class includes 2 American Native or Alaska Native subjects and 2 subjects of unknown race. P-values were calculated using *t*-test (for age and BMI) and Chi-square test (for race). BMI is the abbreviation of Body Mass Index (kg/m^2). (**C**) Male-vs-female transcriptomic distance (tSNR) and percentage of differentially expressed (DE) genes in the genome (at FDR < 0.1) were used to quantify sexual dimorphism in gene expression across 31 tissues. Brain other: cerebral cortex and a set of associated structures.

Figure 2. Multi-tissue sexually dimorphic DE genes. (**A**) Hierarchical clustering of the top 1568 genes with tissue-wide sexually dimorphic gene expression (combined q < 0.1 under Fisher's method). The color code of heat map depicts the *t*-statistics of sexually dimorphic differential expression. The color bar on top of the heatmap depicts germ layers of origin. Annotated on the right are the top 10 female- or male-biased DE genes. (**B-C**) Comparison of *t*-statistics of sexually dimorphic differential expression for all autosomal genes between two adipose tissue types (**B**) and between three brain clusters (**C**). *t* > 0: overexpression in males, depicted in shades of blue; *t* < 0: overexpression in females, depicted in shades of reds. Grey shades represent the significance level of data points (combined q-value). (**D**) Composition of gene biotypes of these 1568 multi-tissue sexually dimorphic DE genes. Female: genes overexpressed in females. Male: genes overexpressed in males. (**E-F**) Characteristics of the protein coding genes in the list of 1568 top multi-tissue sexually dimorphic DE genes. (**E**) Distributions of dN/dS ratio against mouse homologs. P-values were calculated using two-sample *t*-test. (**F**) Enrichment of GWAS disease-associated genes. P-values were calculated using Fisher's Exact Test (*: p < 0.001). (**G**) Word clouds of enriched Gene Ontology terms. Word size reflects term frequency.

Figure 3. Network inference suggests regulatory potential of ESR1 in driving the sexually dimorphic gene expression of its target genes. (**A-B**) Comparison of gene expression levels and targeting levels by ESR1 between females and males in breast for *IGLV4-69* (**A**) and *MYOT* (**B**). (**C-D**) Expression of *IGLV4-69* (**C**) and *MYOT* (**D**) are highly correlated with targeting levels by ESR1 (upper) but not with *ESR1* expression (bottom). Each data point represents a single sample. (**E**) Regression between differential expression levels and differential targeting levels by ESR1 on its putative targets. (**C-E**) R^2: adjusted coefficient of determination. (**F**) Network visualization of the top 20 genes significantly differentially targeted by ESR1. Edge color represents the direction and strength of differential targeting. Node color represents the direction and strength of differential expression.

Figure 4. Key transcription factors that shape the sexually dimorphic gene expression landscape. (**A**-**C**) Inferring differential targeting of ESR1 to its putative target genes versus differential expression levels in three representative tissues. (**D**) Adjusted coefficient of determination (R^2) was used to quantify and rank the regulatory potential of 652 TFs in driving the sexually dimorphic gene expression of their target genes. Positions of representative TFs are highlighted approximately in the plot. See Table 1 for TFs' documented roles in hormone-responsive pathways and sex-related biological processes and diseases. Adipose: subcutaneous adipose tissue. Heart: left ventricle. Brain other: cerebral cortex and a set of associated structures. Colon: sigmoid colon.

Figure 5. Brain networks are differentially wired between males and females. (**A**) Number of genes that are differentially expressed (DE), differentially targeted (DT), and both differentially expressed and targeted (DE&DT) across tissues. (**B**) In basal ganglia, brain functions are enriched in DT genes. X-axis: top nine enriched GO terms; y-axis: genes with shared enriched GO terms. Cells in the matrix indicate if a gene is associated with a term. (**C-D**) Visualization of the synaptic transmission subnetwork in a force-directed layout shows differential wiring in the network between males (**C**) and females (**D**) in basal ganglia.

Supplemental figure legends

Figure S1. GSEA Enrichment maps for brain and adipose tissues, related to Figure 1. Maps of female-vs-male functional enrichments using GSEA based on Gene Ontology (GO) vocabulary for (**A**) adipose and (**B**) brain tissues. Nodes are GO terms (gene sets). Edges represent term distance (mutual overlap between two connected gene sets). Node size indicates gene set size. In (**A**), filled and unfilled circles represent enrichment in subcutaneous and visceral adipose tissues, respectively. Edge colors indicate the source of enrichment: green corresponds to subcutaneous adipose tissue and blue corresponds to visceral adipose tissue.

Figure S2. Gene expression of sex hormone receptors is not ubiquitously sexually dimorphic, related to Figure 2. (**A-C**) Distribution of mRNA expression (log-cpm, Y-axis) of (**A**) estrogen receptor alpha (ESR1), (**B**) estrogen receptor beta (ESR2), and (**C**) androgen receptor (AR) in females (red) and males (blue) across 31 sites (X-axis).

Figure S3. Network inference suggests regulatory potential of transcription factor ESR2 in driving the sexually dimorphic gene expression of its putative target genes, related to Figure 3. (**A**) Inferring sex differential targeting of transcription factor ESR2 on its putative target genes versus differential expression levels. (**B-C**) Comparison of gene expression levels and inferred ESR2 targeting edge weights between females and males for EDN3 (**B**) and MYOT (**C**), as highlighted in (**A**). (**D-E**) Expression of EDN3 (**D**) and MYOT (**E**) are highly correlated with ESR2 regulation (upper) but not with ESR2 expression (bottom). (**F**) Network visualization of the top 20 significantly differentially targeted genes of ESR2.

A

GTEx

549 individuals
(361 males and 188 females)

Prior Knowledge

TF

AAAGGTCA   TGTTTACTTAS

Transcription factor binding motifs
& protein-protein interaction data

Multi-tissue Transcriptomes

8716 samples from 31 sites

Single-sample Networks

Estimating sample-specific regulatory
networks using PANDA+LIONESS

Differential Expression

Profiling sexually dimorphic
differential expression across tissues

Differential Targeting

Identifying sexually dimorphic
differential targeting across tissues

B

Demographics	Male (n=361)	Female (n=188)	Male (%)	Female (%)	P-value
Age (mean±SD) (yrs)	52.5 (±12.4)	52.6 (±12.8)			0.88
BMI (mean±SD)	27.3 (±4.0)	26.9 (±4.2)			0.26
Race					0.67
Asian	4	2	1%	1%	
African American	53	25	15%	13%	
White	300	161	83%	86%	
Others	4	0	1%	0%	

C

tSNR
0.5 1.0 1.5 2.0 2.5 3.0 3.5 4.0 4.5 5.0

Breast
Skin
Thyroid
Adipose subcutaneous
Brain other
Skeletal muscle
Heart left ventricle
Brain cerebellum
Adipose visceral
Esophagus muscularis
Tibial nerve
Artery tibial
Artery aorta
Fibroblast cell line
Brain basal ganglia
Esophagus mucosa
Whole blood
Heart atrial appendage
Pituitary
Gastroesophageal junction
Pancreas
Liver
Lung
Adrenal gland
Artery coronary
Stomach
Spleen
Colon transverse
Colon sigmoid
Lymphoblastoid cell line
Intestine terminal ileum

— ● — tSNR
█ %DE genes

0% 10% 20% 30% 40% 50% 60% 70%
%DE genes (FDR < 0.1)

A — 1568 multi-tissue DE genes (combined q < 0.1) (genes ordered by mean t statistics)

Tissue categories: Ectoderm, Mesoderm, Endoderm, Cell

t scale: -6, -3, 0, 3, 6 (F to M)

Upper gene list (red):
EIF1AXP1
EIF2S3L
RP11-575H3.1
PRSS30P
TSPEAR-AS1
C21orf90
NLRP2
AP001065.15
RP11-204M4.2
AC068279.3

Lower gene list (blue):
AC006946.16
OOEP
RP11-333A23.4
RP11-725P16.2
NOX5
SPESP1
DDX43
FRG1B
LINC01597
RP4-610C12.3

Tissue labels: Breast, Adipose subcutaneous, Adipose visceral, Thyroid, Lung, Stomach, Intestine terminal ileum, Pancreas, Adrenal gland, Spleen, Brain cerebellum, Brain other, Brain basal ganglia, Tibial nerve, Skeletal muscle, Heart atrial appendage, Heart left ventricle, Whole blood, Liver, Esophagus mucosa, Esophagus muscularis, Gastroesophageal junction, Colon sigmoid, Colon transverse, Artery coronary, Artery aorta, Artery tibial, Skin, Pituitary, Lymphoblastoid cell line, Fibroblast cell line

B — Adipose visceral vs Adipose subcutaneous. ♂ / ♀. -log(q): 1, 5, 9.
Labeled genes: SPESP1, AMER2, CADPS, SGCZ, EIF2S3L, EIF1AXP1, SALL1, FOXL2, C3orf72, HOXA11, HOXD11

C — 3D plot: Basal ganglia, Cerebellum, Other brain. ♂ / ♀. -log(q): 1, 3, 5. t scale.
Labeled genes: RP4-610C12.3, LINC01597, FRG1BP, NOX5, RP11-725P16.2, SPESP1, EIF1AXP1, RP11-575H3.1, PRSS30P

D

Gene Biotype	Female	Male
Protein coding genes	634 (69%)	504 (78%)
lincRNA genes	50 (5%)	46 (7%)
Misc RNA genes	3 (1%)	1 (0%)
Ig genes	131 (14%)	0 (0%)
Other genes	101 (11%)	98 (15%)
Total	919	649

E — dN/dS, P=0.004, F vs M

F — GWAS genes (odds ratio), P<0.001, F vs M

G — Word clouds: Female, Male

A Putative target genes of ESR1 — Adipose — $R^2 = 0.78$
Differential targeting (t) vs. Differential expression (t)

B Putative target genes of ESR1 — Thyroid — $R^2 = 0.59$
Differential targeting (t) vs. Differential expression (t)

C Putative target genes of ESR1 — Colon — $R^2 = 0.18$
Differential targeting (t) vs. Differential expression (t)

D Explained variance (R^2), High to Low

Breast: CEBPE, CEBPD, TET1, GMEB2, RFX2, ESR1, ESR2, BRCA1, PPARG, ESRRA

Adipose: SPIC, SPI1, IRF5, SPIB, CEBPG, ESR1, ESR2

Heart: TET1, TOPORS, PITX2, CEBPD, FOXP4, FOXP1, BRCA1, ESR1, ESR2

Thyroid: TET1, CEBPD, CEBPG, PITX2, MEF2D, ESR1, ESR2

Liver: SP3, BHLHA15, SP1, HNF4A, ZNF263, ESR1, ESR2

Other brain: NR2F6, TOPORS, EVX2, GBX1, NR6A1, ESR1, ESR2

Basal ganglia: NR2F6, NR2F2, TOPORS, RARG, TLX2, ESR1, ESR2

Skin: ELF2, KLF4, ZNF263, FOSB, KLF16, ESR2, ESR1

Cerebellum: BACH2, ZNF219, TCF4, TOPORS, TFAP2D, ESR2, ESR1

Colon: RFX3, ATF5, RARA, RFX2, BCL11A, ESR1, ESR2

A

Tissues (top to bottom): Breast, Brain other, Brain basal ganglia, Thyroid, Skin, Adipose subcutaneous, Skeletal muscle, Heart left ventricle, Brain cerebellum, Fibroblast cell line, Esophagus muscularis, Adipose visceral, Tibial nerve, Artery tibial, Artery aorta, Esophagus mucosa, Whole blood, Heart atrial appendage, Pituitary, Gastroesophageal junction, Liver, Pancreas, Lung, Adrenal gland, Artery coronary, Stomach, Spleen, Colon transverse, Colon sigmoid, Lymphoblastoid cell line, Intestine terminal ileum

x-axis: #Genes — 0, 5k, 10k, 15k, 20k, 25k

Inset Venn diagrams: DE, DE&DT, DT

B

Genes (top to bottom): GRIK1, CACNA1A, GLRB, CACNB4, NRXN1, RASGRF1, GRM5, KCNJ10, BRAF, HTR6, CRH, DRD2, SNCA, GRIK2, GRIN2A, PTEN, NETO1, SHANK3, JPH3, EPHB2

Columns (left to right): Synaptic transmission, Regulation of synaptic transmission, Regulation of neuronal synaptic plasticity, Regulation of synaptic plasticity, Behavior, Single-organism behavior, Neuron-neuron synaptic transmission, Heterophilic cell-cell adhesion, Regulation of membrane potential

C

D

A

MITOCHONDRION ORGANIZATION

MITOCHONDRIAL ENVELOPE

GLYCOSYLATION

VACUOLE & LYSOSOME

FATTY ACID METABOLIC PROCESS

LIPOPROTEIN BIOSYNTHETIC PROCESS

PEROXISOME

COENZYME BINDING

AMINE METABOLIC PROCESS

AMINO ACID ACTIVATION

MORPHOGENESIS

TRANSLATION

NCRNA METABOLIC PROCESS

VIRAL REPRODUCTION

NUCLEOLUS

CELL-CELL JUNCTION

CELL CYCLE CHECKPOINT

TRANSFERASE

RNA POLYMERASE COMPLEX

TELOMERE ORGANIZATION

NUCLEOSIDE-TRIPHOSPHATASE REGULATOR ACTIVITY

ELECTRON CARRIER ACTIVITY

LYASE ACTIVITY

REGULATION OF WNT RECEPTOR SIGNALING PATHWAY

VOLTAGE-GATED CATION CHANNEL ACTIVITY

Node: GO term

Edge: Term distance

● Male enrichment

● Female enrichment

B

Immune response

Synaptic transmission

Potassium ion transport

Telomere maintenance

Angiogenesis

Dopamine receptor signaling pathway

Endocrine system development

Regulatory region DNA binding

Cytokine binding

RNA splicing

RNA transport

Cholesterol homeostasis

Respiratory electron transport chain

Response to unfolded protein

Helicase activity

A Putative target genes of ESR2

$R^2 = 0.81$

Differential targeting (t) (y-axis, -15 to 15)
Differential expression (t) (x-axis, -15 to 15)

EDN3 (red point)
MYOT (blue point)

B EDN3 / ESR2-EDN3

Expression (F, M)
Edge Weight (F, M)

C MYOT / ESR2-MYOT

Expression (F, M)
Edge Weight (F, M)

D Edge weight of ESR2-EDN3

$R^2 = 0.68$

Female (red), Male (blue)

Expression of ESR2

$R^2 = 0.04$

Expression of EDN3

E Edge weight of ESR2-MYOT

$R^2 = 0.36$

Female (red), Male (blue)

Expression of ESR2

$R^2 = 0.02$

Expression of MYOT

F ESR2 network

LOXHD1, EMC8, HECW1, CCDC69, TRPV6, RPH3AL, COL7A1, TINF2, LCN2, FES, TMPRSS3, ARL2, FHAD1, SLC2A14, FER1L4, ADA, SYT8, SIX2, EDN3, ATP6V0A4

F — t — M