Escape from X inactivation drives sex differences and female trait variation

Carrie Zhu[1,2], Liaoyi Xu[1,2], Arbel Harpak[1,2,+]

[1]Departmentof Integrative Biology, University of Texas at Austin, Austin, TX, USA

[2]Department of Population Health, University of Texas at Austin, Austin, TX, USA

[+]To whom correspondence should be addressed: arbelharpak@utexas.edu

Abstract

X chromosome inactivation (XCI) partially balances gene dosage between sexes, yet expression from the inactive X (Xi) is variable across genes. In this study, we investigate whether gene-level Xi expression predicts transcriptional and phenotypic consequences of X-linked variation. We find that Xi expression levels are a strong linear predictor of female-male expression differences, suggesting that other compensatory or regulatory mechanisms play a more minor role in sex differences in X-linked gene expression. Among females, we identify three traits—lymphocyte percentage, HbA1c, and schizophrenia—for which higher Xi expression correlates with the strength of evidence for dominance effects. We hypothesize that an underappreciated mechanism could generate dominance effects of X-linked variants on a trait—specifically when the variant influences skew in X inactivation. This work establishes Xi expression as essential for understanding sex differences and the female-specific genetic basis of disease.

Introduction

X chromosome inactivation (XCI) in females is traditionally viewed as an equalizer, bringing X-linked gene dosage to be on par with dosage in males by silencing one X chromosome[1–3]. However, many genes "escape" this silencing and are expressed from the inactive X (Xi)[4,5]. Roughly 15–30% of X-linked genes show some degree of escape, with some escaping consistently across tissues and individuals, and others showing context-dependent escape[6,7]. The extent of escape can vary continuously, from low-level to equal expression between Xi and the active X (Xa). Understanding the consequences of expression from Xi is critical for interpreting both sex differences and individual variation among females[8–10].

An immediate expectation is that levels of Xi expression would impact dosage balance between the sexes. Genes in the pseudoautosomal regions (PAR) of the X chromosome, which recombine with the Y chromosome, are not subject to XCI and thus maintain expression from both sex chromosomes[11,12]. Yet, even PAR genes show unequal expression between the sexes, with higher expression in males reported in several tissues. This demonstrates that expression from Xi does not typically reach the full levels of the active X chromosome nor that of the Y chromosome[6,7,13]. In contrast, many non-PAR (NPX) genes have no Y homologs and exhibit higher expression in females due to unopposed Xi expression[6].

Beyond sex differences, Xi expression may also influence phenotypic variation within females[14,15]. Heterozygote loss of function mutations provide compelling case studies. For example, such mutations in DDX3X increase the risks of developmental delay and disability[16], and in KDM6A the risk of Kabuki syndrome[17–19]. These observations imply that Xi gene dosage can be impactful for development and disease. Further, Agarwal et al. recently showed that in a group of 17 NPX genes (those with Y homologs), the fitness cost of heterozygote loss of function mutations is higher than in any other group of genes genome-wide[20] .

Here, we test whether gene-level Xi expression predicts transcriptional and phenotypic consequences of X-linked variation. We find that, across genes, Xi expression levels quantitatively predict both (i) the magnitude of female-male expression differences and (ii) the strength of genetic associations for disease risk and physiological traits in females. We also suggest a novel hypothesis for non-additive effects of X-linked variants on a trait, specifically when the variant also influences a skew in X inactivation. Testing for dominance effects in X-linked variants, we find weak evidence for these effects being mediated by genetically encoded skews at individual loci. However, integrating signals across Xi expressed genes, we identify three

traits—lymphocyte percentage, HbA1c, and schizophrenia—for which the strength of Xi expression predicts the significance of dominance effects.

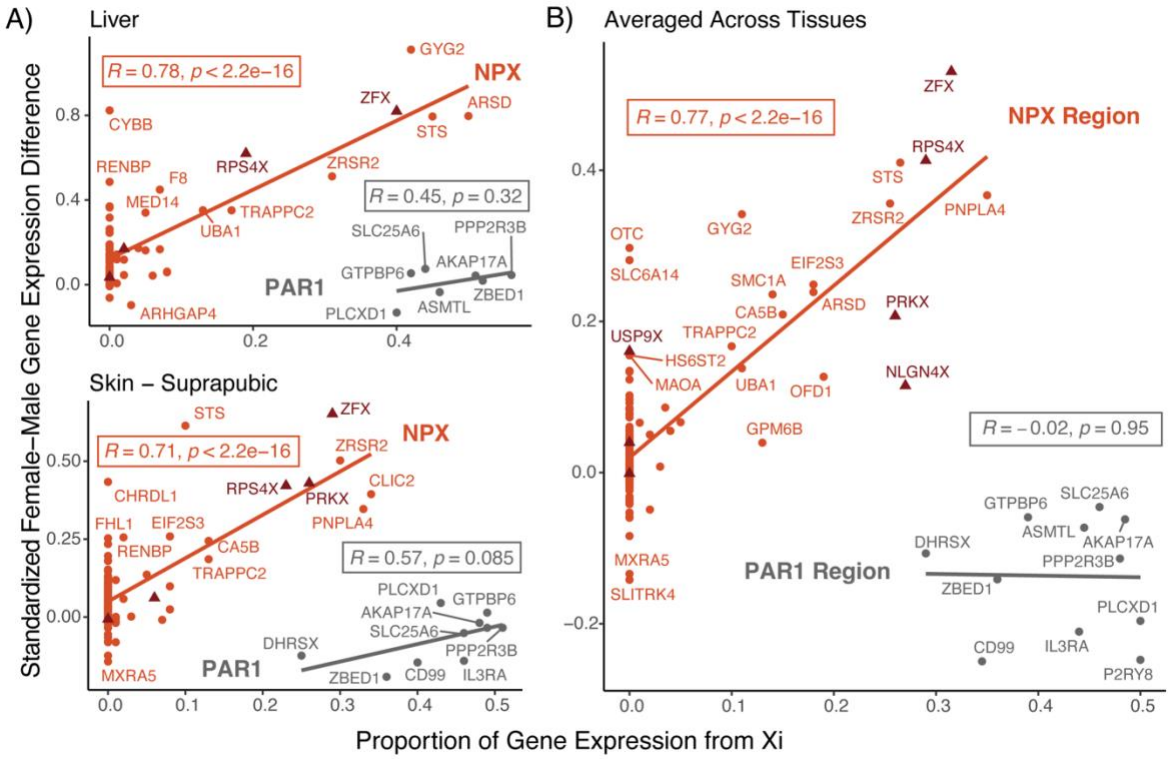
Results

We used RNA-seq data from the Genotype-Tissue Expression (GTEx) v8 project[21] for 643 X-linked genes across 14 tissue types in 635 XY and 312 XX adult individuals. While recognizing that sex chromosome karyotype does not always align with biological sex, we henceforth label XX individuals as female and XY individuals as male. We quantify Xi expression using allele-specific expression data obtained from a single GTEx female donor whose X inactivation is non-random across cells and tissues ("skewed XCI"), i.e. the inactivated chromosome is the one inherited from the same parent in nearly all cells[6]. Importantly, our reliance on data from this single GTEx donor implies that, throughout this work, we do not consider variation in XCI and escape among females. Instead, our focus is on variation in Xi expression across genes and these genes' contribution to sex differences and within-female variation.

Xi expression predicts sex differences in gene expression. Across genes in the non-pseudoautosomal (NPX) region, expression from Xi is strongly correlated with sex differences in median expression level at both the individual tissue level (e.g., liver and skin in **Fig. 1A**) and when examining the per-gene medians across tissues (Pearson $r = 0.77$; $p < 2.2e\text{-}16$; **Fig. 1B**). In contrast, in the pseudoautosomal region (PAR), 11 out of 11 genes had higher expression in males (by 19% on average), and sex differences were not correlated with Xi expression (Pearson p=0.95; **Fig. 1B**). These results are in line with prior work by Tukiainen et al.[6] and Tomofuji et al.[22] who also report greater female expression in escape genes in the NPX region, and greater male expression in the PAR1 region. This is, to our knowledge, the first illustration of the linear, continuous and tight relationship between Xi expression and subsequent sex differences, which suggests that other compensatory or regulatory mechanisms play a more minor role in sex differences in X-linked gene expression.

Many Xi-expressed genes have a highly consistent expression levels across tissues—18 out of 43 (42%) per a previous report[6]. In such genes, e.g. in ZRSR2 and UBA1, we also see little variance in sex differences in gene expression (**Fig. S1**). Genes such as ARSD and FHL1, which have more variable Xi expression, tend to also exhibit greater variance in sex differences in expression. In general, variance across tissues in sex differences is proportional to variance across tissues in Xi expression (**Fig. S1**). One outlier to this trend is GYG2, which is involved in

88 glycogenesis[23]. GYG2 shows disproportionately larger variance in sex differences across tissues
89 and is in fact expressed from Xi only in the liver, pancreas, and stomach (between 27% and



90

91 **Figure 1: Xi expression in a single female predicts sex differences in gene expression. (A)** The x-
92 axis shows the ratio of expression from Xi and total expression in the tissue (liver on top, suprapubic skin
93 on the bottom) estimated in a single GTEx female donor. The y-axis displays the difference between median
94 female and male transcripts per million (TPM), standardized by the median male TPM for each gene. For
95 example, y=0 confers to equal medians and 0.5 confers to the median expression of the gene among
96 females being 50% higher than that of males. Gene expression data for the respective tissues was obtained
97 from GTEx v8. We estimate the Pearson correlation statistics independently for the non-PAR X (orange)
98 and PAR1 (gray) regions. Genes with homologs on the Y chromosome are indicated as dark red triangles.
99 **(B)** The x-axis shows the median proportion of expression from Xi across tissues. The y-axis shows the
100 median, across tissues, of standardized female-male differences.

101

102 111% higher median expression in females, across these tissues; **Figs. 1B, S2**). In other tissues
103 with no expression from Xi, sex differences still varied extensively. For example, in coronary
104 artery tissue, median expression was 83% higher in females while in the esophagus' muscularis
105 layer it was only 20% higher. Such tissue-specific sex differences may reflect additional regulatory
106 factors, such as epigenetic modifications or tissue-specific chromatin states, that are not
107 necessarily unique to Xi expression[24].

Beyond the impact of Xi expression on sex differences, it may also contribute uniquely to female variation. We therefore also inquired into consequences for phenotypic variation among females. Once again, we do so by examining, across genes, the relationship between Xi expression (inferred from a single individual) and measures of gene-level contribution to variation. We performed X-wide association studies (XWAS) for 13 diseases and continuous physiological traits. We chose these traits based on either having relatively high SNP heritabilities or documented phenotypic sex differences for greater statistical power[25–28]. We tested for standard, additive effects. We discuss the relationship between additive effects and Xi expression later in the section "**Xi expression predicts contribution to trait variation among females**".

A new hypothesis for dominance effects of X-linked variants. In our XWAS, we also tested whether the heterozygote effect deviates from the average of homozygote effects (a "dominance-deviation effect"). Hypothetically, dominance may be observed on the X chromosome for the same reasons as they may exist in autosomes. However, dominance effects are thought to be negligible among common variants in autosomes[29,30]. Should XCI itself, or perhaps escape from XCI, alter our expectations for X-linked variants? We distinguish three biological scenarios for a variant in an escape gene that yield different expectations for the nature of X-linked variants' effects.

First, we consider a biallelic SNP with an allelic effect β when the effect allele is expressed from Xa. If the locus completely escapes X inactivation, the expected effect of a heterozygote is equal to the average of the two homozygotes ($\alpha = 1$ in **Table 1-i**) just like in autosomes. When inactivation is complete and random, one of the two alleles' effects would be zeroed at random and so the genotypic effect is halved ($\alpha = 0$ in **Table 1-i**). For some intermediate degree of Xi expression (ranging from 0 to 1), the expected effect in homozygotes is $(1 + \alpha)\beta$, and in heterozygotes it is $\frac{(1+\alpha)}{2}\beta$ (**Table 1-i**). In conclusion, the presence of a variant in an Xi-expressed locus, in and within itself, does not lead us to expect dominance-deviation effects.

Second, we consider an additional effect of the variant—modulating Xi expression levels—when it is present on the inactive chromosome. Still, in this scenario, the variant does not influence X inactivation choice. While the effect size now depends on the allele-specific level of Xi expression, the effect remains additive (**Table 1-ii**).

Case	Expected Effect of Genotype	Dominance Deviation Effect?
(i) Inactivation and expression from Xi (at level α) are independent of allele.	AA: 0 AT: $\frac{(1+\alpha)}{2}\beta$ TT: $(1+\alpha)\beta$	No
(ii) Inactivation is independent, but expression from Xi depends on allele (at level α_A for allele A on Xi and level α_T for allele T on Xi)	AA: 0 AT: $\frac{(1+\alpha_T)}{2}\beta$ TT: $(1+\alpha_T)\beta$	No
(iii) Inactivation is allele-dependent: In heterozygotes, an X chromosome with the T allele has probability δ of being the active X. (Expression from Xi is fixed at level α.)	AA: 0 AT: $\delta\beta+(1-\delta)\alpha\beta$ TT: $(1+\alpha)\beta$	Yes

Table 1. When an X-linked allele affects both a trait and the likelihood of inactivation, we expect it to have a dominance deviation effect on the trait.

A third case we consider is when the variant influences which parental X is inactivated. Heterozygotes will then exhibit skewing of XCI with unequal probabilities of each of the alleles being on Xa and, as a result, having the more pronounced effect. Therefore, under this scenario, we expect dominance effects in our XWAS (**Table 1-iii**). In other words, the allelic effect on the trait in this case is mediated by a pleiotropic effect on XCI.

Dominance effects may also arise for mechanisms that are not unique to the X or Xi. We emphasize, however, that the mechanism in case (iii) hinges on genetically encoded biases in XCI. While there is ample literature on skewed XCI in blood and other tissues[31–34], the majority of skewing is thought arise from stochastic processes, clonal expansion, or aging[35–37], rather than from heritable genetic variation. Nonetheless, some evidence supports the possibility of heritable components to skewed XCI, especially in extreme or familial cases[33,38–41]. These findings may also motivate the testing for dominance effects as a way to identify potential loci that influence XCI.

Across all 13 traits, we found only 17 SNPs with genome-wide significant (p<5e-8) dominance effects. All of the SNPs were in the NPX, but Xi expression data was not available for nearby

161 genes. Notably, the nearest genes also had no Y homologs. These effects were suggestive of
162 complete dominance (**Fig. S3**). To illustrate, **Fig. 2** shows the mean phenotype across genotypes
163 for three SNPs with significant dominance effects on HbA1c (**Fig. 2A-C**) and one with a significant
164 effect on reported migraine cases (**Fig. 2D**). We caveat these observations in noting that sample
165 sizes for the minor allele homozygotes were small, and the allelic effect of the SNPs among males
166 (per a simple comparison of mean phenotypes conferring to hemizygote genotypes; **Fig. S3**) was
167 often discordant. One intriguing case is a SNP close to ZFP92 (**Fig. 2A**), which encodes a zinc
168 finger protein thought to regulate DNA transcription and transposable element silencing[42].
169 Therefore, a linked variant may influence X-inactivation, consistent with Case (iii) in **Table 1**.
170
171 **Xi expression predicts contribution to trait variation among females.** We next asked
172 whether higher Xi expression suggests a greater role in explaining trait variation among females.
173 We did so by testing for a correlation between Xi expression level (using the median across
174 tissues) and the strength of a gene's association with the trait. To assign a strength of association,
175 we annotated each gene with the minimal p-value association among SNPs for which the gene is
176 the closest gene[43]. We estimated this correlation across 13 traits and two types of effects, additive
177 and dominance deviation. For height, the significance of additive effects and expression from Xi
178 were highly correlated (Pearson $r = 0.45$, $p = 0.029$; **Fig S4**).
179
180 Another intriguing correlation is observed between Xi expression levels and significance of
181 dominance-deviation effects on percentage of lymphocytes in leukocytes (Pearson $r = 0.52$, $p =$
182 0.012; **Fig. S4**). If XCI skewness is heritable and more polygenic than previously appreciated, this
183 correlation would fall in line with Case (iii) in **Table 1**. Overall, for both types of effects, we find
184 that genes with higher tissue-averaged Xi expression tend to have stronger GWAS associations
185 near them. For additive effects, there was a positive Pearson correlation for 11/13 traits (1
186 significant at $p < 0.05$). For dominance-deviation effects, this was true for 12/13 traits (1 significant
187 at $p < 0.05$; **Fig. S4**).
188
189 We investigated this relationship further, in a tissue-specific manner, namely by examining
190 correlations of XWAS significance with tissue-specific Xi expression. In some tissues, most traits
191 we examined showed significant correlation between Xi expression and strength of XWAS
192 association ($p < 0.05$). This includes stomach (7 traits for additive GWAS; 10 traits for dominance
193 deviation), thyroid (4 additive, 7 dominance deviation), as well as in lymphoblastoid cell lines
194 (LCLs) (4 additive, 8 dominance deviation) (**Fig. 3C**). These tissues were also highly correlated

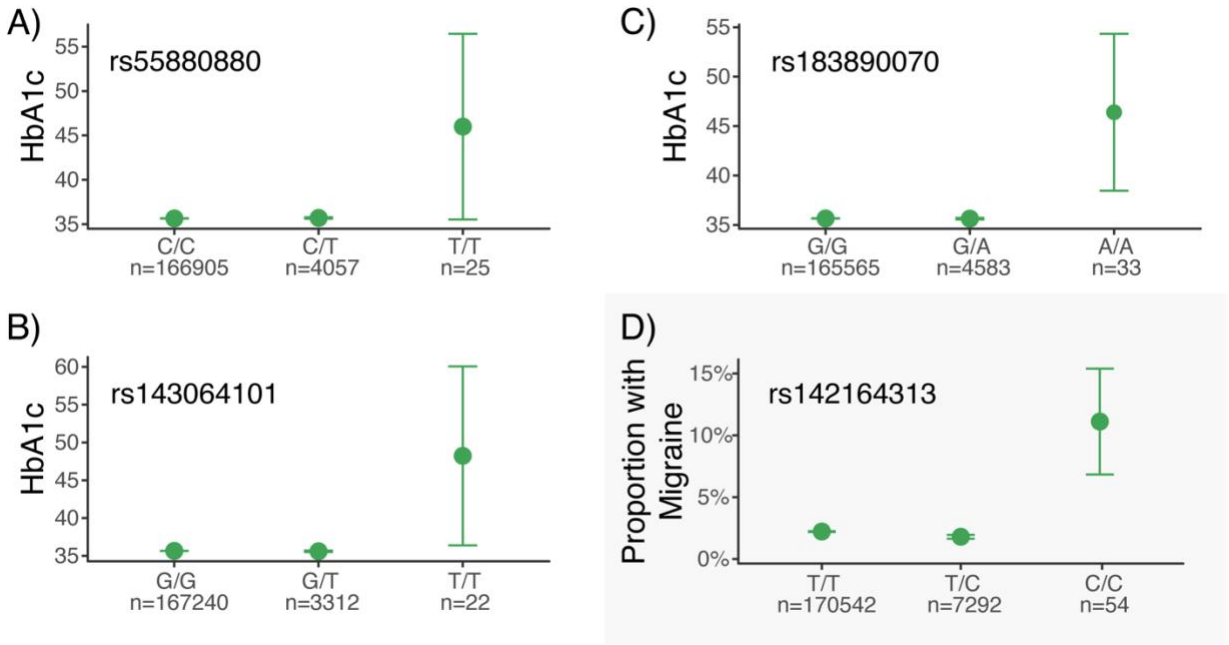


Figure 2: Examples of NPX SNPs with significant dominance effects. Shown are four SNPs that exhibited genome-wide significant dominance deviation effects in an XWAS for HbA1c (**A-C**) and reported case of migraine (**D**). The y-values confer to the mean phenotype value (±SE), or proportion, in (D), for females in the White British subset of the UKB with the respective genotype. Sample sizes (n) are printed below each genotype.

with each other in terms of Xi expression (see dendrogram on the left hand side of **Fig. 3C**). However, we were unable to discern whether this points to pleiotropic contributions from Xi expressed genes. Other tissues such as esophagus mucosa, kidney cortex, suprapubic skin, and lung, had no significant correlations. This may result from low variation in Xi expression among genes or a lesser impact of Xi-expressed genes on variation in these traits. Since Xi expression is generally similar across tissues (about 43% of genes with nonzero Xi expression were characterized as having similar expression levels across tissues by Tukiainen et al.[6], the latter is more likely.

For some traits, Xi expression and significance of allelic effects were significantly correlated across multiple tissues for only one type of effect. Interestingly, lymphocyte percentage in leukocytes, HbA1c, and schizophrenia were only significant for the dominance deviation model across 5, 4, and 4 tissues respectively **(Fig. 3A, 3B)**. We therefore hypothesize that X inactivation skew may more commonly mediate the effects of Xi expression in these tissues. Recognizing the multiple hypothesis burden of testing 13 traits, 14 tissues and 2 effects types, we note that only 4

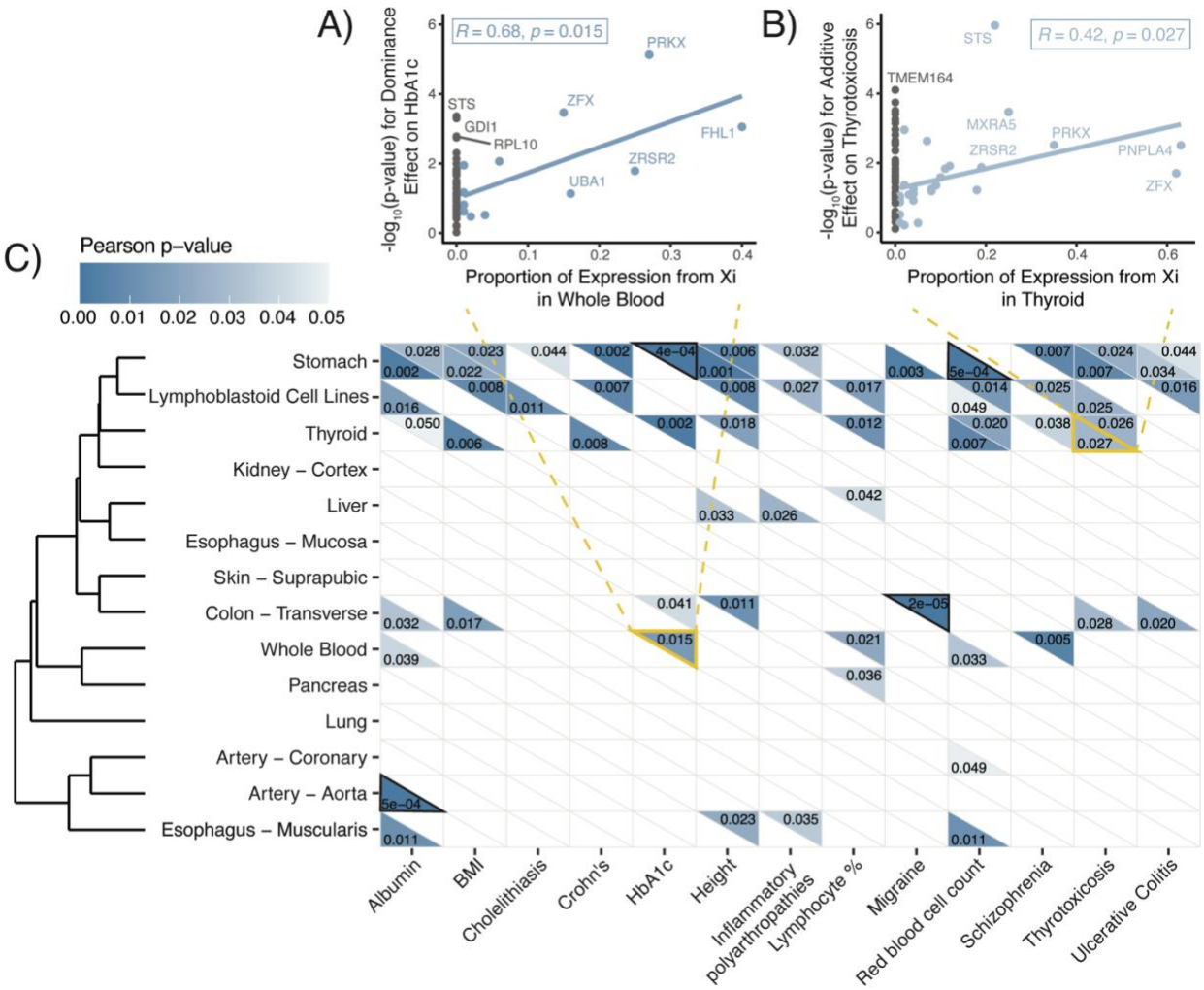


Figure 3: Correlation between Xi expression and gene-trait association across tissues. We estimate, for each tissue, trait and effect type (additive or dominance deviation) the Pearson correlation between the proportion of Xi expression out of total (Xi + Xa) expression and the maximum -$\log_{10}$(p-value) among variants overlapping the gene. Only genes with non-zero expression (blue points) were considered. **(A)** HbA1c in whole blood and dominance deviation effects. **(B)** Thyrotoxicosis in thyroid and dominance deviation effects. **(C)** Correlations across tissue, trait and effect type combinations. Only significant correlations are displayed, with blue hues conferring to the significance levels. The bottom-left triangles of each block show Pearson p-values for Xi expression and additive effect -$\log_{10}$(p-value), and the top-right triangles show the same for dominance-deviation effects. We cluster tissues by pairwise correlations of Xi expression, as shown by the dendrogram on the left. The length of each branch is inversely proportional to the correlation between tissues in Xi expression, For example, Xi expression is more correlated between stomach and lymphoblastoid cell lines (LCL) than it is between thyroid and kidney cortex.

232 tissue-trait-effect combinations are significant at an FDR significance level of 0.05 (bolded
233 triangles in **Fig. 3C**).
234

235 Discussion

236 Our study demonstrates that expression from Xi is a major, quantifiable contributor to
237 transcriptional sex differences and trait variation among females. First, we find that Xi expression,
238 as measured in a single individual with skewed inactivation, robustly predicts sex differences in
239 expression across non-pseudoautosomal (NPX) genes and tissues. While previous work has
240 documented sex differences in expression of escape genes[6], our findings go further by showing
241 that this relationship extends across the full range of Xi expression, even among genes not
242 formally categorized as 'escapees', supporting the view that escape from X inactivation acts along
243 a continuum, rather than as a binary on/off phenomenon (see relevant discussion in [7,44]). The
244 strong correlation (Pearson $r = 0.77$) between Xi expression and higher female expression in NPX
245 regions supports a model by which sex differences arise largely from unequal dosage, rather than
246 from differential regulation of the active X.
247
248 We also propose a simple model to delineate an X-chromosome specific expectation for
249 dominance-deviation effects. Namely, dominance-deviation effects will arise when allele-specific
250 skewing affects which haplotype is inactivated. By distinguishing between effects on expression
251 given that an allele is Xi-linked and effects on XCI, we provide a clear rationale for interpreting
252 dominance effects in the context of XCI. We identify several traits—lymphocyte percentage,
253 HbA1c, and schizophrenia—where dominance-deviation signals may point to genetically encoded
254 XCI bias. These traits are particularly compelling because they align with known biology of XCI
255 skew or tissue-specific gene regulation. Lymphocyte percentage reflects hematopoietic
256 processes, where XCI skew is well-documented due to clonal expansion and somatic
257 selection[33,37]. Some studies also suggest a heritable component to skew[33,38], which may explain
258 why dominance effects on the percentage of lymphocytes in leukocytes is strongly correlated with
259 Xi expression. Such skew may expose or silence functional alleles, helping explain the greater
260 incidence of autoimmune disorders in females. HbA1c, a marker of glucose control, could reflect
261 Xi expression effects in metabolic tissues like the liver or pancreas—tissues where genes like
262 GYG2 and PNPLA4 exhibit high Xi expression. Schizophrenia is a neurodevelopmental disorder
263 with known sex differences in presentation, as well as previous observations of an XCI skew in
264 affected females[45]. Together, these examples suggest that dominance effects on these traits may
265 reveal genetically encoded variation in XCI.

266

267 Finally, the correlation between Xi expression and GWAS association strength tends to cluster by
268 tissue type, suggesting that Xi expression acts through shared regulatory environments. Tissues
269 such as stomach, thyroid, and LCLs show the strongest and most consistent correlations across
270 traits. Given that a large proportion of Xi-expressed genes are similarly expressed across tissues,
271 these results could imply that Xi-driven trait effects may often arise from genes involved in
272 fundamental cellular processes such as metabolism or structure.

273

274 In sum, our study shows that escape from X inactivation—treated as a quantitative, gene-level
275 trait—has widespread implications for sex differences and trait variation in humans. Our findings
276 underscore the importance of Xi expression in shaping sex differences and influencing genetic
277 architecture of physiological and disease traits.

278

279 Methods

280 **GTEx expression data**. To estimate gene expression levels, we used bulk tissue RNA
281 sequencing data from the Genotype-Tissue Expression (GTEx) v8 project[46] for 727 genes across
282 14 tissue types in 635 males and 312 females. Gene expression levels are measured in
283 transcripts per million (TPM). The list of genes on the X chromosome was obtained from San
284 Roman et al.[7] which included 712 genes on the non-PAR (NPX) region and 15 genes on the
285 pseudo-autosomal 1 (PAR1) region of X chromosome. We removed sex-specific tissues and
286 genes with less than 10 individuals with TPM greater than 1. We also removed tissues, including
287 brain tissues, without corresponding Xi expression data described in "**Use of allele-specific**
288 **expression data to measure degree of escape from XCI**". Any gene-tissue pairs with 0 TPM
289 were removed.

290

291 **Use of allele-specific expression data to measure degree of escape from XCI.** We measured
292 the amount of expression from the inactive X chromosome using data obtained from Tukiainen et
293 al.[6] They use allele-specific expression data from a single female GTEx-UPIC donor who exhibits
294 a near complete degree of X inactivation skew, in which the same parental chromosome is
295 inactivated across nearly all cells. For each gene, one ASE site with the highest coverage in the
296 largest number of tissues is chosen. The allele with the lower combined expression across tissues
297 is labeled the inactive X (Xi) allele. XIST is the sole exception, with the Xi allele having higher
298 expression levels across tissues. We then quantify the degree of escape using the ratio of Xi

299 expression to total (inactive + active) X expression. In total, the combination of GTEx expression
300 data and allele-specific expression data is available for 190 NPX and 11 PAR1 genes.
301
302 **UK Biobank.** We used the UK Biobank, a large biomedical database of phenotypic and genotypic
303 information of around half a million UK participants[47,48]. In this study, we focused on female
304 participants, with female labeling based on the presence of an XX karyotype. Quality checks of
305 sample characteristics included removing individuals with sex chromosome aneuploidy, or a self-
306 reported sex other than female. We excluded all related individuals 3rd degree or closer, as listed
307 in data field 22020. We removed individuals who had withdrawn from the UK Biobank at any time
308 before we performed the analysis of UKB data (6/19/2024). Furthermore, we limited participants
309 to those who self-identified as White and as British and who were also clustered in the genetic
310 principal component space. After filtering, we had a total sample of 181012 female individuals.
311
312 **UKB Genotype Data.** We used X-chromosome, bi-allelic SNPs with INFO scores greater than
313 0.8 (Resource 1967 in UKB). Following a similar quality check method as used by Neale Lab[49],
314 we then filtered SNPs with call rate > 0.95, Hardy-Weinberg equilibrium test p-value larger than
315 1e-10 and minor allele frequency larger than 1e-3. Filtering was completed using *plink 2.0*
316 *alpha*[50,51].
317
318 **Phenotype Data**. In this study, we analyzed 13 phenotypes including 6 continuous and 7 binary
319 disease phenotypes. The continuous traits were chosen for their relatively high SNP heritability
320 estimates as provided by Neale Lab[49]. Disease phenotypes were classified using ICD10 coding
321 (data field 41270 in UK Biobank) and chosen based on known sex-specific differences in disease
322 incidence, course or prognosis.
323
324 **XWAS.** We conducted female-specific association analysis for the X chromosome using *plink 2.0*
325 *alpha*[50,51]. We included birth year and the first ten principal components (data field 22009 in UK
326 Biobank) as covariates. We applied two different models, additive and dominance-deviation, using
327 the --genotypic modifier.
328
329 **Gene annotation**. We used *Ensembl Variant Effect Predictor*[43] to annotate highly associated
330 SNPs ($p < 0.05$) obtained from XWAS. We found the closest gene within 5000 bp up- and down-
331 stream each SNP. For **Fig. 3**, we created a list of SNPs for each gene with available Xi expression

332 data using the *Ensembl REST API*[52] overlap endpoint, which retrieves variants that lie within the
333 gene region. We chose the SNP with the lowest p-value to represent the gene.
334
335 **Summarizing the correlation distance between tissues in Xi expression using a tree.** To
336 create the dendrogram shown on the left in **Fig. 3C**, we first estimated the pairwise Pearson
337 correlation of proportion of Xi expression levels for 14 tissues using the data from Tukiainen et
338 al[6]. We treated the inverse of these correlations as a distance measure between each pair of
339 tissues. Using the resulting pairwise distance matrix, we performed hierarchical clustering using
340 the function *hclust()* in *R version 4.3.1* using default settings. The length of (horizontal) branches
341 shown in **Fig. 3C** confers to the distance between each pair or cluster of tissues.
342

343 ## Acknowledgements
344 We thank Ipsita Agarwal, Andrés Bendesky, Jared Cole, Mark Kirkpatrick and Molly Przeworski
345 for helpful discussions and comments on the manuscript. This work was funded by NIH grant
346 R35GM151108 and a Pew Biomedical Scholarship to A.H. This study was conducted using the
347 UK Biobank resource under application 92741, as approved by the University of Texas at Austin
348 institutional review board (study protocol 00003287). The authors acknowledge the Texas
349 Advanced Computing Center (TACC) at The University of Texas at Austin for providing
350 computational resources that have contributed to the research results reported in this manuscript.

Literature Cited

1. Lyon, M. F. Gene Action in the X-chromosome of the Mouse (Mus musculus L.). *Nature* 190, 372–373 (1961).
2. Disteche, C. M. Dosage Compensation of the Sex Chromosomes. *Annu Rev Genet* 46, 537–560 (2012).
3. Disteche, C. M., Filippova, G. N. & Tsuchiya, K. D. Escape from X inactivation. *Cytogenet Genome Res* 99, 36–43 (2002).
4. Posynick, B. J. & Brown, C. J. Escape From X-Chromosome Inactivation: An Evolutionary Perspective. *Front Cell Dev Biol* 7, (2019).
5. Balaton, B. P., Dixon-McDougall, T., Peeters, S. B. & Brown, C. J. The eXceptional nature of the X chromosome. *Hum Mol Genet* 27, R242–R249 (2018).
6. Tukiainen, T. *et al.* Landscape of X chromosome inactivation across human tissues. *Nature* 550, 244–248 (2017).
7. San Roman, A. K. *et al.* The human inactive X chromosome modulates expression of the active X chromosome. *Cell Genomics* 3, 100259 (2023).
8. Khramtsova, E. A. *et al.* Quality control and analytic best practices for testing genetic models of sex differences in large populations. *Cell* 186, 2044–2061 (2023).
9. Fass, S. B. *et al.* Relationship between sex biases in gene expression and sex biases in autism and Alzheimer's disease. *Biol Sex Differ* 15, 47 (2024).
10. Ungar, R. A. *et al.* Transcriptomic signatures of rare variant impacts across sex and the X chromosome. *Human Genetics and Genomics Advances* 6, 100463 (2025).
11. Graves, J. The origin and evolution of the pseudoautosomal regions of human sex chromosomes. *Hum Mol Genet* 7, 1991–1996 (1998).
12. Blaschke, R. J. & Rappold, G. The pseudoautosomal regions, SHOX and disease. *Curr Opin Genet Dev* 16, 233–239 (2006).
13. Johnston, C. M. *et al.* Large-Scale Population Study of Human Cell Lines Indicates that Dosage Compensation Is Virtually Complete. *PLoS Genet* 4, e9 (2008).
14. Fu, Y., Kenttämies, A., Ruotsalainen, S., Pirinen, M. & Tukiainen, T. Role of X chromosome and dosage-compensation mechanisms in complex trait genetics. *The American Journal of Human Genetics* 112, 1330–1343 (2025).
15. Sidorenko, J. *et al.* The effect of X-linked dosage compensation on complex trait variation. *Nat Commun* 10, 3009 (2019).
16. Tang, L. *et al.* Prospective and detailed behavioral phenotyping in DDX3X syndrome. *Mol Autism* 12, 36 (2021).
17. Van Laarhoven, P. M. *et al.* Kabuki syndrome genes *KMT2D* and *KDM6A* : functional analyses demonstrate critical roles in craniofacial, heart and brain development. *Hum Mol Genet* 24, 4443–4453 (2015).
18. Lederer, D. *et al.* Deletion of KDM6A, a Histone Demethylase Interacting with MLL2, in Three Patients with Kabuki Syndrome. *The American Journal of Human Genetics* 90, 119–124 (2012).
19. Cheon, C. K. *et al.* Identification of KMT2D and KDM6A mutations by exome sequencing in Korean patients with Kabuki syndrome. *J Hum Genet* 59, 321–325 (2014).
20. Agarwal, I., Fuller, Z. L., Myers, S. R. & Przeworski, M. Relating pathogenic loss-of-function mutations in humans to their evolutionary fitness costs. *Elife* 12, (2023).

21. Lonsdale, J. *et al.* The Genotype-Tissue Expression (GTEx) project. *Nat Genet* 45, 580–585 (2013).
22. Tomofuji, Y. *et al.* Quantification of escape from X chromosome inactivation with single-cell omics data reveals heterogeneity across cell types and tissues. *Cell Genomics* 4, 100625 (2024).
23. Mu, J. & Roach, P. J. Characterization of Human Glycogenin-2, a Self-glucosylating Initiator of Liver Glycogen Metabolism. *Journal of Biological Chemistry* 273, 34850–34856 (1998).
24. Qadir, M. M. F. *et al.* Single cell regulatory architecture of human pancreatic islets suggests sex differences in β cell function and the pathogenesis of type 2 diabetes. Preprint at https://doi.org/10.1101/2024.04.11.589096 (2024).
25. Nyholt, D. *et al.* Evidence for an X-linked genetic component in familial typical migraine. *Hum Mol Genet* 7, 459–463 (1998).
26. Bauer, M., Glenn, T., Pilhatsch, M., Pfennig, A. & Whybrow, P. C. Gender differences in thyroid system function: relevance to bipolar disorder and its treatment. *Bipolar Disord* 16, 58–71 (2014).
27. Miquel, C. H., Faz-Lopez, B. & Guéry, J. C. Influence of X chromosome in sex-biased autoimmune diseases. *J Autoimmun* 137, 102992 (2023).
28. van Vollenhoven, R. F. Sex differences in rheumatoid arthritis: more than meets the eye... *BMC Med* 7, 12 (2009).
29. Visscher, P. M. *et al.* Genome Partitioning of Genetic Variation for Height from 11,214 Sibling Pairs. *The American Journal of Human Genetics* 81, 1104–1110 (2007).
30. Palmer, D. S. *et al.* Analysis of genetic dominance in the UK Biobank. *Science (1979)* 379, 1341–1348 (2023).
31. Minks, J., Robinson, W. P. & Brown, C. J. A skewed view of X chromosome inactivation. *Journal of Clinical Investigation* 118, 20–23 (2008).
32. Zito, A. *et al.* Escape from X-inactivation in twins exhibits intra- and inter-individual variability across tissues and is heritable. *PLoS Genet* 19, e1010556 (2023).
33. Bolduc, V. *et al.* No evidence that skewing of X chromosome inactivation patterns is transmitted to offspring in humans. *Journal of Clinical Investigation* 118, 333–341 (2008).
34. Sharp, A., Robinson, D. & Jacobs, P. Age- and tissue-specific variation of X chromosome inactivation ratios in normal women. *Hum Genet* 107, 343–349 (2000).
35. Newall, A. E. T. Primary non-random X inactivation associated with disruption of Xist promoter regulation. *Hum Mol Genet* 10, 581–589 (2001).
36. Busque, L. *et al.* Nonrandom X-inactivation patterns in normal females: lyonization ratios vary with age. *Blood* 88, 59–65 (1996).
37. Werner, J. M., Ballouz, S., Hover, J. & Gillis, J. Variability of cross-tissue X-chromosome inactivation characterizes timing of human embryonic lineage specification events. *Dev Cell* 57, 1995-2008.e5 (2022).
38. Amos-Landgraf, J. M. *et al.* X Chromosome–Inactivation Patterns of 1,005 Phenotypically Unaffected Females. *The American Journal of Human Genetics* 79, 493–499 (2006).
39. Thorvaldsen, J. L., Krapp, C., Willard, H. F. & Bartolomei, M. S. Nonrandom X Chromosome Inactivation Is Influenced by Multiple Regions on the Murine X Chromosome. *Genetics* 192, 1095–1107 (2012).

40. Chadwick, L. H. & Willard, H. F. Genetic and parent-of-origin influences on X chromosome choice in Xce heterozygous mice. *Mammalian Genome* 16, 691–699 (2005).
41. Chadwick, L. H., Pertz, L. M., Broman, K. W., Bartolomei, M. S. & Willard, H. F. Genetic Control of X Chromosome Inactivation in Mice: Definition of the *Xce* Candidate Interval. *Genetics* 173, 2103–2110 (2006).
42. Osipovich, A. B. *et al.* ZFP92, a KRAB domain zinc finger protein enriched in pancreatic islets, binds to B1/Alu SINE transposable elements and regulates retroelements and genes. *PLoS Genet* 19, e1010729 (2023).
43. McLaren, W. *et al.* The Ensembl Variant Effect Predictor. *Genome Biol* 17, 122 (2016).
44. San Roman, A. K. *et al.* The human Y and inactive X chromosomes similarly modulate autosomal gene expression. *Cell Genomics* 4, 100462 (2024).
45. Zhang, X. *et al.* A new sex-specific underlying mechanism for female schizophrenia: accelerated skewed X chromosome inactivation. *Biol Sex Differ* 11, 39 (2020).
46. Aguet, F. *et al.* The GTEx Consortium atlas of genetic regulatory effects across human tissues. *Science (1979)* 369, 1318–1330 (2020).
47. Sudlow, C. *et al.* UK Biobank: An Open Access Resource for Identifying the Causes of a Wide Range of Complex Diseases of Middle and Old Age. *PLoS Med* 12, e1001779 (2015).
48. Bycroft, C. *et al.* The UK Biobank resource with deep phenotyping and genomic data. *Nature* 562, 203–209 (2018).
49. Abbot, L. *et al.* GWAS Results. http://www.nealelab.is/uk-biobank/ (2017).
50. Purcell, S. & Chang, C. PLINK 2.00 alpha. Preprint at www.cog-genomics.org/plink/2.0/ (2020).
51. Chang, C. C. *et al.* Second-generation PLINK: rising to the challenge of larger and richer datasets. *Gigascience* 4, 7 (2015).
52. Yates, A. *et al.* The Ensembl REST API: Ensembl Data for Any Language. *Bioinformatics* 31, 143–145 (2015).