Does Sex-Differential Gene Expression Drive Sex-Differential Selection in Humans?

Matthew J. Ming[1], Changde Cheng[2], Mark Kirkpatrick[1], and Arbel Harpak[1,3†]

[1]Department of Integrative Biology, University of Texas at Austin, Austin, TX, USA
[2]School of Medicine, University of Alabama at Birmingham, Birmingham, AL, USA
[3]Department of Population Health, University of Texas at Austin, Austin, TX, USA
[†]Correspondence: arbelharpak@utexas.edu

Abstract

Sex differences in human transcriptomes have been argued to drive sex-differential natural selection (SDS). Here, we show that previous evidence supporting this hypothesis has been largely unfounded. We develop a new method to test for a genome-wide relationship between sex differences in expression and selection on expression-influencing alleles (eQTLs). We apply it across 39 human tissues and find no evidence for a general relationship. We offer possible explanations for the lack of evidence, including that it is due in part to eQTL ascertainment bias towards sites under weak selection. We conclude that the drivers of ongoing SDS in humans remain to be identified.

Introduction

Sex-differences in gene expression have been theorized to be the *result* of long-term sex-differential selection (SDS), where allelic fitness effects differ between males and females[1,2]. When a gene product is only beneficial in one sex, it is expected that expression modifiers will evolve to increase expression for that sex and decrease expression in the other [3]. The reverse causality has also been proposed: sex differences in gene expression might be the *driver* of SDS on expression modifiers[1,4].

SDS acting on viability within the current generation can generate between-sex differences in allele frequency[5,6] at expression modifiers (expression Quantitative Trait Loci, or eQTLs)[3]. Sexually dimorphic traits (diseases being one example with strong repercussions on survival[6,7,8,9]) may be subject to SDS acting through viability. Previous work has proposed and tested theoretical models for relating divergence in eQTL allele frequencies with differential expression between sexes[1,10,11]. Some of the empirical results have, however, been called into question[12,13,14].

A controversial result by Cheng and Kirkpatrick (2016)[1] (hereafter "CK16") is a characteristic pattern relating between-sex F_{ST}[1,5,15] (henceforth "F_{ST}") to sex differences in gene expression (henceforth "Δ"). They observed high F_{ST} values at intermediate values of Δ and near-zero values of F_{ST} when a gene is expressed evenly between the sexes ($\Delta = 0$) or is only expressed in one sex ($\Delta = \pm 1$). They nicknamed this bimodal pattern "Twin Peaks". The Twin Peaks pattern has been used as a signal to detect SDS in several species[1,4,16,17].

Here, we revisit the CK16 model and statistical approach. We find that their interpretation was based on previously unappreciated statistical artifacts. We then refine their model and apply it to new, more extensive data on gene expression and allele frequencies. Across 39 human tissues, we find no evidence for a genome-wide relationship between viability SDS and Δ. We discuss how a bias in eQTL discovery towards eQTLs under weaker selection can explain the lack of signal for a relationship between SDS and Δ, and how the drivers of SDS may still be investigated.

Results

Twin Peaks is a Statistical Artifact

We begin by revisiting the Twin Peaks pattern with a critical eye towards caveats in its application and interpretation. Most importantly, we reconsider the model that CK16 proposed to explain the pattern. It is based on two key assumptions. First, the relationship between a gene's expression levels and its effect on fitness in each sex is linear. Second, sexually antagonistic selection is symmetric, meaning selection coefficients are equal in magnitude and opposite in sign between sexes. This yields a quadratic relationship between F_{ST} and Δ at a biallelic site affecting expression. At small absolute values of Δ,

$$F_{ST} = 4pq\Delta^2 A, \quad (1)$$

where p is the allele frequency in zygotes and $q = 1 - p$. F_{ST} is the between-sex fixation index[18] which is used to quantify allele frequency differences between males and females. In the model, these differences are due solely to the sex differences in post-zygotic fitness effects of alleles. The quantity Δ is the sex difference in gene expression (**Methods**). Finally, A is a compound parameter involving the within-sex effect of gene expression level on fitness.

The expectation in CK16 for F_{ST} at extreme expression differences ($|\Delta| \to 1$), however, is based on intuition rather than a model. The authors suggest that if a gene is not expressed in one sex ($\Delta = \pm 1$), then selection will not act on it. Selection on the other sex should optimize expression levels, so under the symmetrical selection assumption there will be no ongoing directional selection in that sex either. As neither sex experiences selection, there will be no force driving increased F_{ST}. CK16 then interpolate a bimodal shape by joining the quadratic relationship at low $|\Delta|$ and the expectation for $F_{ST} = 0$ at $\Delta = \pm 1$ (**Fig. 1**).

However, the assumption of symmetric selection used at low values of Δ may be inappropriate for large Δ values. In particular, when $\Delta = \pm 1$, the lack of expression in one sex plausibly suggests different selection between males and females. It is therefore not intuitive that F_{ST} should simply go to zero at sites regulating expression in these genes. Additionally, although the quantity $2pq$ appears in **Eq. 1**, CK16 did not include that term in fitting the model, effectively assuming it to only contribute random noise to the relationship between F_{ST} and Δ.

Because of these caveats to the model, we decided to revisit the support for Twin Peaks. We first ask whether the Twin Peaks pattern is due to SDS by applying the statistical tests of CK16 to data generated under a null hypothesis of no relationship between Δ and SDS. We replicated the pattern shown in CK16 by using the same data and methods. Namely, we performed a 4^{th} degree polynomial regression of F_{ST} on Δ, using allele count data from 1000

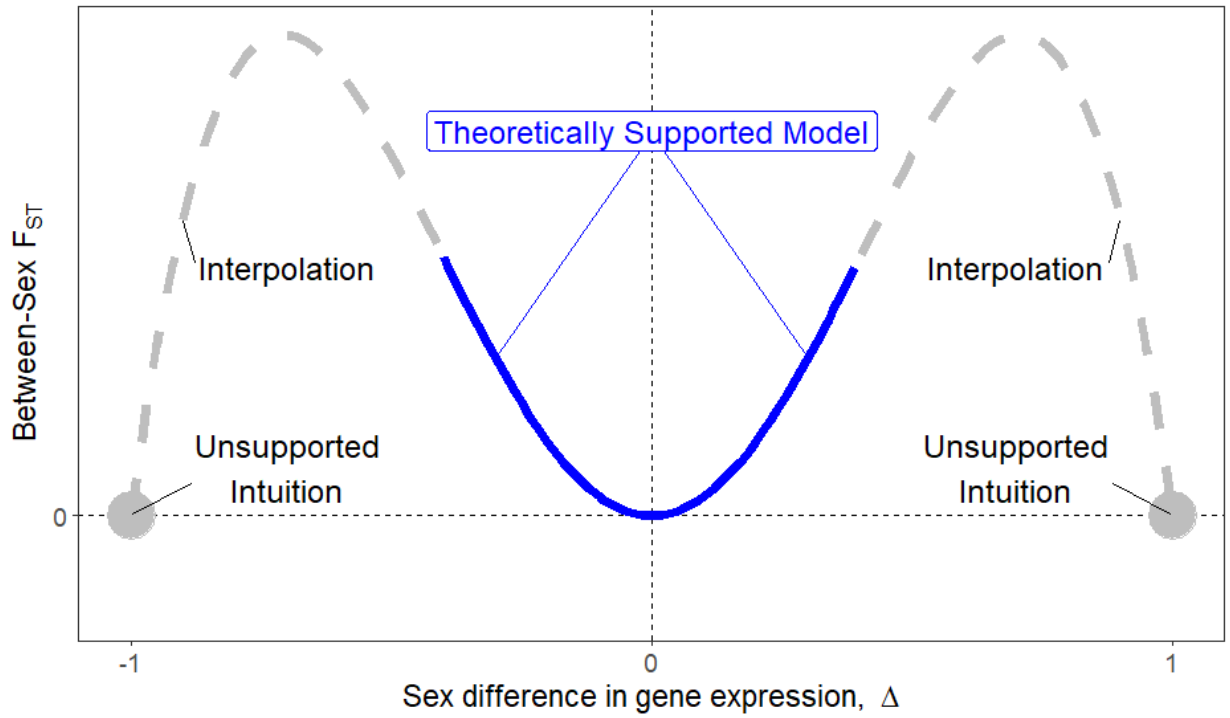

Figure 1: Revisiting the Twin Peaks theoretical expectation. Cheng and Kirkpatrick (2016) derive an expectation for a quadratic relationship at low values of Δ (solid blue line). However, their expectation for "Twin Peaks" is based on a reasoning for $F_{ST} = 0$ at genes with sex-specific expression ($\Delta = \pm 1$) that is unsupported. Additionally, values between the theoretically supported region and the extremes are based on qualitative interpolation (dashed line). The portions of the model which are not supported are shown in grey.

Genomes[19] and expression data from the gonads (ovaries and testes) in GTEx v3[20]. The curve and associations generated using these datasets we refer to as the "real data" (blue line in **Fig. 2a**).

We then generated an empirical null by permuting ovary and testes tissue labels in the expression data but retaining sex labels associated with F_{ST} values (**Methods**), then recomputed Δ values, and again fit a 4th-degree polynomial. We find 21% of permutations qualify as Twin Peaks according to the criteria used by CK16, and that the polynomial fit to the real data is not visually distinct from those fits to the null (**Fig. 2a**). Higher order polynomial regressions can also yield spurious fits because distant points have an oversized impact[21,22]. We therefore binned genes by Δ values and examined the relationship with mean F_{ST} in each bin. Again, the real data shows no unusual relationship between F_{ST} and Δ compared to null data (**Fig. 2b**). From these results, we conclude the Twin Peaks pattern is not statistically significant.

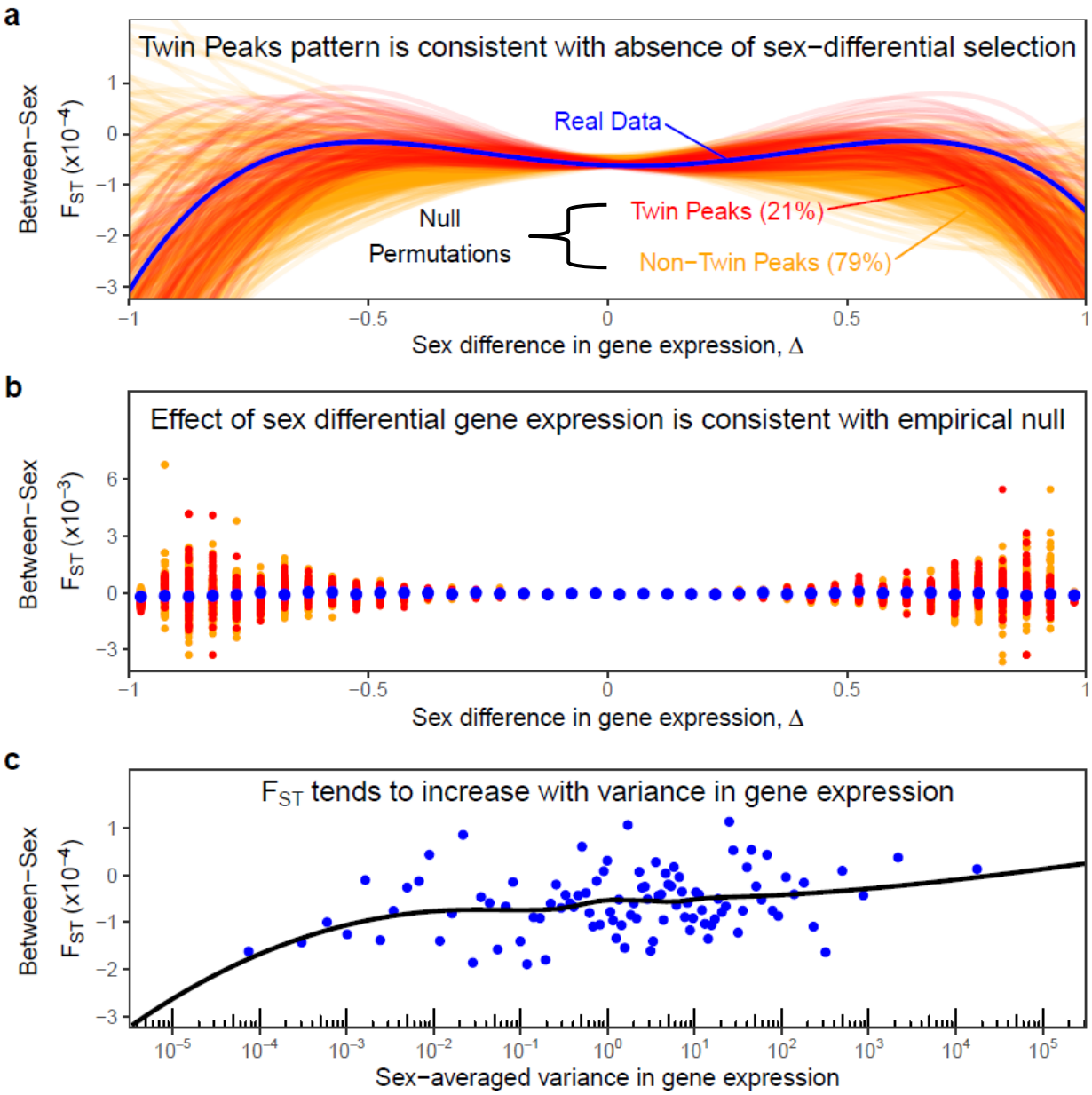

Figure 2. The Twin Peaks pattern is the result of statistical artifacts. a) A 4th-degree polynomial regression of between-sex F_{ST} to sex differences in gene expression for the real data is shown in blue. Red and yellow curves shown are fits based on samples from an empirical null generated by permuting the sex labels in expression data. In red (yellow) are null iterations classified as consistent (inconsistent) with the Twin Peaks expectation according to criteria employed by Cheng and Kirkpatrick (2016). **b)** The data was split into 40 evenly spaced bins of sex difference in gene expression to visualize mean trends. The x-axis values show the midpoint of each bin, and the y-axis values show mean male-female F_{ST} of biallelic sites within each bin. **c)** A positive relationship between F_{ST} and variance in gene expression (regardless of sex) may contribute to the genome-wide relationship between sex differences in gene expression and selection. Genes are divided into 100 evenly sized bins of sex-averaged sample variance in gene

expression. The y-axis values show the mean male-female F_{ST} for all biallelic sites for all genes in each the bin. The black curve shows a LOESS fit to the non-binned data.

Importantly, this permutation method differs from the method used in CK16, which permuted Δ values across genes. Both methods break the associations between F_{ST} and Δ as desired. Our sex -label permutation method, however, preserves F_{ST} associations with the gene's overall expression, maintaining gene features such as expression variance which the CK16 method does not. This explains why their reported *p*-value for Twin Peaks curves (0.016) is lower than our (0.21).

We hypothesized that one reason for the spurious Twin Peaks pattern is confounding. In particular, variance in expression (regardless of sex) is positively correlated with both Δ and F_{ST}. Previous work has shown that genes subject to weaker stabilizing selection show higher variance in expression[23]. Higher variance means larger differences between randomly selected subgroups, and therefore it should translate to larger values of Δ even in the absence of SDS. In turn, stronger (sex-agnostic) selection can lead to stronger drift at linked sites[24,25,26,27]. Taken together, the confounding with variance in gene expression could generate a relationship between SDS and |Δ| in the absence of SDS. Indeed, expression variance and F_{ST} are positively correlated (Pearson $p = 0.011$; **Fig. 2c**), consistent with this hypothesis. In sum, we do not find support for Cheng and Kirkpatrick's conclusion that there is a genome-wide relationship between F_{ST} and Δ.

No evidence for genome-wide SDS on eQTLs

Although we do not find support for a relationship between sex differences in gene expression and selection using CK16's methodology for generating Twin Peaks, one may still exist. We test this hypothesis across many tissues using improved statistical modeling, data, and methods. Despite the caveats to portions of the model discussed above, we believe that CK16's theoretical expectation of a quadratic relationship between F_{ST} and small values of Δ is valid. We therefore built on that model by introducing the compound parameter $\delta^2 = 4qp\Delta^2$ and rewriting **Eq. 1** as

$$F_{ST} = A\delta^2. \quad (2)$$

(**Methods**). To estimate A, for each gene-tissue pair, we used the single *cis*-eQTL from GTEx v8[28] with the strongest association with its expression. Our estimator of A is then the inverse-variance-weighted linear regression of F_{ST} on δ^2 (**Eq. 2**). The advantages of this formulation over CK16's are that it allows direct estimation of A, the strength of SDS on sex differences in

expression. This model also accounts for variation in allele frequencies across sites. Further, by using a single eQTL to calculate F_{ST} for the whole gene, we circumvent biases which can arise when using a simple mean to estimate gene-wide F_{ST}[29,30].

To accompany the updated regression methodology, we also updated the datasets for both allele frequencies and gene expression. For allele frequencies, we used the Non-Finnish European subset in gnomAD v3 (averaging over 60,000 allele samples per site)[31]. We used expression data from 49 tissues from the GTEx v8 dataset[28] (averaging over 200 male samples and 150 female samples). Both datasets greatly expand our sample size compared to CK16, and GTEx v8 provides sample-specific sex labels for calculating Δ across multiple tissues instead of just the gonads—ovaries and testes—as in the original Twin Peaks paper.

Using the updated statistical framework method, we find no evidence that A differs significantly from zero in any tissue (**Fig. 3; Methods**). One explanation for the absence of a pattern may be found in recent work by Mostafavi et al. (2023). The authors contrasted how selection impacts the discovery of genome-wide association study (GWAS) hits with the discovery of eQTLs. Variants with large effects on phenotypes are expected to segregate at low frequencies, reducing discovery power. In GWAS, this is counterbalanced by increased power due to their large effect sizes. Consequently, in GWAS, low frequency variants can still be detected if their effect is large enough. In contrast, eQTL discovery is based only on the effect of genotype on gene expression which does not necessarily translate to fitness-relevant trait variation. Detection of strongly selected sites is therefore less likely[32]. This ascertainment bias can weaken the relationship between Δ and F_{ST} at eQTLs, as sites with high F_{ST} are less likely to be eQTLs.

In "**Twin Peaks is a Statistical Artifact**", we suggested that large sex differences in expression may be entirely unrelated to SDS—for instance, merely tagging genes with high expression variance. There are other potential explanations for the lack of a genome-wide relationship. While some sex differences in expression may be driven by or drive SDS, these are the exception rather than the rule. The majority of sex differences in expression may be a regulatory side effect of SDS on different genes. Lastly, sex-differential gene expression may be due to past SDS, but not drivers of current SDS[33]. Regardless of the reason, if a causal relationship between Δ and FST is rare in the genome, it would be difficult to detect using models that assume a pervasive, persistent relationship between the two.

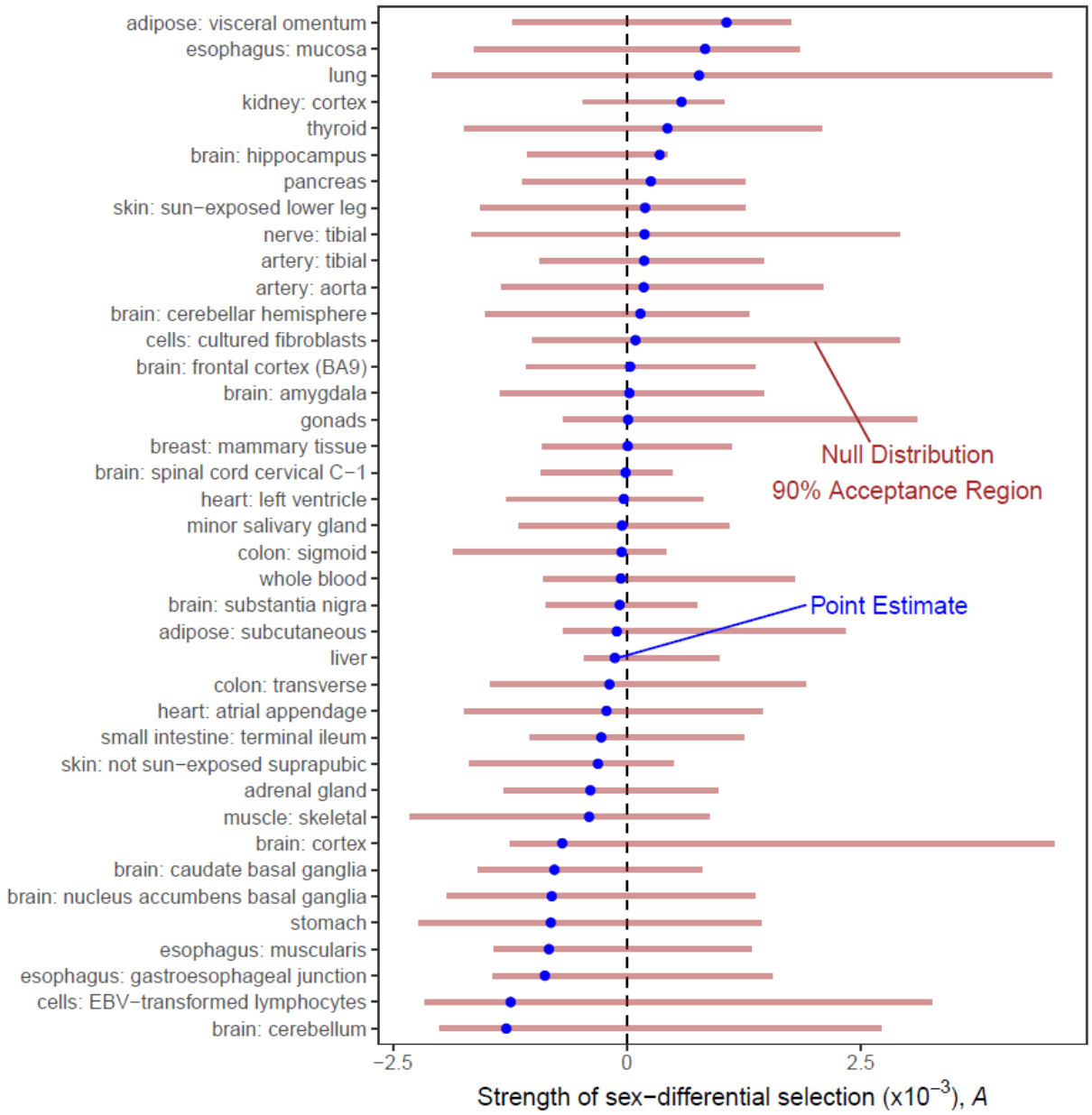

Figure 3. Lack of evidence for sex-differential selection on sex-differential gene expression using newer data. Shown are estimates of A (**Eq. 2**). Acceptance regions are based on the 5th to 95th quantiles of an empirical null distribution generated by permuting sex labels in the gene expression data.

Conclusion Previous work suggested a genome-wide relationship between sex differences in expression and SDS. However, this work was based on statistical artifacts and confounded effects, such as stabilizing selection and variance in gene expression. Even when using newer data and improved statistical methods, we found no evidence for a genome-wide relationship

between sex-differential expression and contemporary SDS. In contrast, studies that measured SDS irrespective of expression or trait variation have reported pervasive, genome-wide signals of SDS in the human genome[5,10,34]. While causal relationships, past and present, between SDS and sex-differential gene expression in humans remain plausible, they are yet to be fully elucidated.

Acknowledgements

We thank Jared Cole for comments on the manuscript and other members of the Harpak Lab for helpful conversations. The work was funded by NIH grants GM11685307 to M.K., and R35GM151108 and a Pew Scholarship to A.H.

Materials and Methods

Permuting Twin Peaks sex-labels

In the section "**Twin Peaks is a Statistical Artifact**", we demonstrate that the Twin Peaks curve presented in CK16 is not statistically significant compared to a permuted null. To do so, we used the methods and datasets described in CK16 and **Eq. 1**. We computed allele count using the 1000 Genomes dataset[19] and calculating F_{ST}. We filtered out any sites with only a single alternative allele in either males or females (i.e., singletons). We used the Transcripts Per Million (TPM) normalized GTEx v3 (referred to as "pilot" on the download page) dataset[20] for expression levels for calculating Δ. Because GTEx v3 does not have individual sample labels, this analysis only compared expression in the gonads (ovaries and testes) where expression level summaries represent a single biological sex. We used the Ensembl GRCh38 v77[35] annotation file for gene annotations. Following CK16, we limited our analysis to protein-coding genes.

To estimate F_{ST}, we used Hudson's estimator[36] based on the R package used in CK16[37] as presented by Bhatia et al. (2013)[29],

$$\hat{F}_{ST} = \frac{(p_m - p_f)^2 - \frac{p_m(1-p_m)}{n_m - 1} - \frac{p_f(1-p_f)}{n_f - 1}}{p_m(1-p_f) + p_f(1-p_m)}. \quad (3)$$

Here, p_m and p_f are the allele frequencies in males and females respectively, and n_m and n_f are the number of males and females respectively. As a measure of sex-differential expression we used Δ as defined in CK16:

$$\Delta = \frac{x_m - x_f}{x_m + x_f}. \quad (4)$$

Here, x_m and x_f are sex-averaged TPM-normalized expression levels in males and females respectively. Because $\hat{F}_{ST}$ in **Eq. 3** is a site-specific estimator while Δ is gene-wide, we generated a gene-wide estimate of F_{ST} by taking a simple average of all sites within a gene body plus 1000bp upstream and downstream. To generate the Twin Peaks curve, we used a 4th-degree polynomial regression between $\hat{F}_{ST}$ and Δ. We note that here and in CK16, this regression therefore ignores variation in heterozygosity across sites (**Eq. 1**). In our improved model (**Results: "No evidence for SDS on eQTLs across all genes in multiple tissues"** and **Methods: "Applying a new model and method for SDS-expression regression"**) we correct this omission.

To compare the original Twin Peaks curve to curves generated under the null of no SDS, we generated an empirical null distribution of 4th-degree regression curves using sex-label permutations of gene expression data. We permuted the ovary and testes labels for each GTEx sample, then recalculated Δ for all genes. We then re-performed the 4th-degree polynomial regression on the new Δ values (the F_{ST} values remain unchanged). This was repeated 500 times. By permuting sex labels, we break the association between sex-differential expression and sex-differential expression, while preserving other gene-level features. To quantitatively evaluate the significance of Twin Peaks in the null distribution, we used the three criteria laid out by Cheng and Kirkpatrick (2016) for classifying a curve as Twin Peaks. Namely, a 4th-degree polynomial must 1) be significant ($p < 0.05$) for the 4th-degree term by ANOVA, 2) have a negative coefficient for the Δ^4 term, and 3) have three real roots. Any 4th-degree regression passing all three criteria is classified as Twin Peaks.

Applying a new model and method for SDS-expression regression

In the section "**No evidence for SDS on eQTLs across all genes in multiple tissues**", we described a revised method for testing a genome-wide relationship between SDS and Δ. For this analysis, we revised the inference model and data. We used the Non-Finnish European subset in the gnomAD V3 dataset[31] to calculate between-sex F_{ST} and heterozygosity at each eQTL. This set of samples has an order of magnitude more individuals (average of 31,470 samples per site) compared to 1000 Genomes data (average of 2,300 samples per site when combining all ancestry

groups) originally used in CK16. Additionally, we used the normalized TPM from the current GTEx v8 for expression level data, which includes many more samples (average of 236 male and 168 female samples per tissue) than the pilot (8 male samples and 13 female samples in the gonads)[28]. This version also provides sample-specific sex labels, enabling us to use additional tissues beyond gonads.

The equation we base our regression on is the model relating F_{ST} with δ^2 shown in **Eq. 2**. Now, rather than performing a 4th-degree polynomial regression as in the analysis for "**Twin Peaks is a Statistical Artifact**", we perform a weighted linear regression of F_{ST} to $\delta^2 = 4pq\Delta^2$, isolating A as the coefficient of regression. Note that by including $4pq$ in the independent variable, we allow information about heterozygosity to affect the SDS-expression relationship. We weighted each point of the regression by $1/Var(expression)$, where $Var(expression)$ is the averaged variance in expression of each sex. This should decrease the leverage of points with large $|\Delta|$. To get a gene-wide estimate of F_{ST}, we used eQTLs from GTEx as mapped in the v8 study[31]. For each gene, we chose the single *cis*-eQTL (within 1Mbp of the gene's midpoint) with lowest *p*-value for association with the gene and use that site for calculating F_{ST}, p, and q. By using an eQTL selected this way, we tried to isolate the effect of selection on gene expression to the site presumed to be contributing most to expression changes in that gene.

To determine the significance of our A estimates, we generated 90% null acceptance regions by permuting sex labels. We permuted sex labels as in "**Twin Peaks sex-label permutation**" such that for each tissue, sex labels are permuted in GTEx expression samples and Δ is recalculated across all genes. Then, for each iteration A is recalculated by linear regression using the new Δ values (F_{ST} and $4pq$ remain unchanged). The 90% null acceptance region was obtained by the 5th and 95th percentiles of 1,000 A values calculated on permuted expression data.

Literature Cited

1. Cheng, C., & Kirkpatrick, M. (2016). Sex-specific selection and sex-biased gene expression in humans and flies. *PLoS Genetics*, *12*(9).

2. Naqvi, S., Godfrey, A. K., Hughes, J. F., Goodheart, M. L., Mitchell, R. N., & Page, D. C. (2019). Conservation, acquisition, and functional impact of sex-biased gene expression in mammals. *Science*, *365*(6450).

3. Rice, W. R. (1984). Sex chromosomes and the evolution of sexual dimorphism. *Evolution*, 735-742.

4. Wright, A. E., Fumagalli, M., Cooney, C. R., Bloch, N. I., Vieira, F. G., Buechel, S. D., ... & Mank, J. E. (2018). Male-biased gene expression resolves sexual conflict through the evolution of sex-specific genetic architecture. *Evolution Letters*, *2*(2), 52-61.

5. Zhu, C., Ming, M. J., Cole, J. M., Edge, M. D., Kirkpatrick, M., & Harpak, A. (2023). Amplification is the primary mode of gene-by-sex interaction in complex human traits. *Cell Genomics*, *3*(5).

6. Oliva, M., Muñoz-Aguirre, M., Kim-Hellmuth, S., Wucher, V., Gewirtz, A. D., Cotter, D. J., Parsana, P., Kasela, S., Balliu, B., … & Stranger, B. E. (2020). The impact of sex on gene expression across human tissues. *Science*, *369*(6509).

7. Khramtsova, E. A., Wilson, M. A., Martin, J., Winham, S. J., He, K. Y., Davis, L. K., & Stranger, B. E. (2023). Quality control and analytic best practices for testing genetic models of sex differences in large populations. *Cell*, *186*(10), 2044-2061.

8. Ober, C., Loisel, D. A., & Gilad, Y. (2008). Sex-specific genetic architecture of human disease. *Nature Reviews Genetics*, *9*(12), 911-922.

9. Connallon, T., & Hall, M. D. (2018). Environmental changes and sexually antagonistic selection. *eLS*, 1-7.

10. Ruzicka, F., Holman, L., & Connallon, T. (2022). Polygenic signals of sex differences in selection in humans from the UK Biobank. *PLoS Biology*, *20*(9), e3001768.

11. Ellegren, H., & Parsch, J. (2007). The evolution of sex-biased genes and sex-biased gene expression. *Nature Reviews Genetics*, *8*(9), 689-698.

12. Mank, J. E., Shu, J. J., & Wright, A. E. (2020). Signature of sexual conflict is actually conflict resolved. *Molecular Ecology*, *29*(2): 215–217.

13. Cheng, C., & Kirkpatrick, M. (2020). The signal of sex-specific selection in humans is not an artifact: Reply to Mank et al. *Molecular Ecology*, *29*(8), 1406.

14. Kasimatis, K. R., Abraham, A., Ralph, P. L., Kern, A. D., Capra, J. A., & Phillips, P. C. (2021). Evaluating human autosomal loci for sexually antagonistic viability selection in two large biobanks. *Genetics*, *217*(1).

15. Kirkpatrick, M., & Guerrero, R. F. (2014). Signatures of sex-antagonistic selection on recombining sex chromosomes. *Genetics*, *197*(2), 531-541.

16. Dutoit, L., Mugal, C. F., Bolívar, P., Wang, M., Nadachowska-Brzyska, K., Smeds, L., Yazdi, H. P., Gustafsson, L., & Ellegren, H. (2018). Sex-biased gene expression, sexual antagonism and levels of genetic diversity in the collared flycatcher (Ficedula albicollis) genome. *Molecular Ecology*, *27*(18), 3572-3581.

17. Vaux, F., Rasmuson, L. K., Kautzi, L. A., Rankin, P. S., Blume, M. T. O., Lawrence, K. A., Bohn, S., & O'Malley, K. G. (2019). Sex matters: otolith shape and genomic variation in deacon rockfish (Sebastes diaconus). *Ecology and Evolution, 9,* 13153–13173.

18. Wright, S. (1949). The genetical structure of populations. *Annals of Eugenics*, *15*(1), 323-354.

19. 1000 Genomes Project Consortium. (2015). A global reference for human genetic variation. *Nature*, *526*(7571), 68.

20. Lonsdale, J., Thomas, J., Salvatore, M., Phillips, R., Lo, E., Shad, S., Hasz, R., Walters, G., Garcia, F., Young, N., & Moore, H. F. (2013). The genotype-tissue expression (GTEx) project. *Nature Genetics*, *45*(6), 580-585.

21. Harrell, F. E. (2001). *Regression modeling strategies: with applications to linear models, logistic regression, and survival analysis* (Vol. 608). New York: springer.

22. Magee, L. (1998). Nonlocal behavior in polynomial regressions. *The American Statistician*, *52*(1), 20-22.

23. Glassberg, E. C., Gao, Z., Harpak, A., Lan, X., & Pritchard, J. K. (2019). Evidence for weak selective constraint on human gene expression. *Genetics*, *211*(2), 757-772.

24. Charlesworth, B. (1994). The effect of background selection against deleterious mutations on weakly selected, linked variants. *Genetics Research*, *63*(3), 213-227.

25. Hudson, R. R., & Kaplan, N. L. (1995). Deleterious background selection with recombination. *Genetics*, *141*(4), 1605-1617.

26. Hudson, R. R., & Kaplan, N. L. (1995). The coalescent process and background selection. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, *349*(1327), 19-23.

27. Murphy, D. A., Elyashiv, E., Amster, G., & Sella, G. (2022). Broad-scale variation in human genetic diversity levels is predicted by purifying selection on coding and non-coding elements. *Elife*, *12*.

28. GTEx Consortium. (2020). The GTEx Consortium atlas of genetic regulatory effects across human tissues. *Science*, *369*(6509), 1318-1330.

29. Reynolds, J., Weir, B. S., & Cockerham, C. C. (1983). Estimation of the coancestry coefficient: basis for a short-term genetic distance. *Genetics*, *105*(3), 767-779.

30. Bhatia, G., Patterson, N., Sankararaman, S., & Price, A. L. (2013). Estimating and interpreting FST: the impact of rare variants. *Genome Research*, *23*(9), 1514-1521.

31. Chen, S., Francioli, L. C., Goodrich, J. K., Collins, R. L., Kanai, M., Wang, Q., Alföldi, J. Watts, N. A., Vittal, C., Gautier, L. D., … & Karczewski, K. J. (2024). A genomic mutational constraint map using variation in 76,156 human genomes. *Nature*, *625*(7993), 92-100.

32. Mostafavi, H., Spence, J. P., Naqvi, S., & Pritchard, J. K. (2023). Systematic differences in discovery of genetic effects on gene expression and complex traits. *Nature Genetics*, *55*(11), 1866-1875.

33. Muralidhar, P., & Coop, G. (2024). Polygenic response of sex chromosomes to sexual antagonism. *Evolution*, *78*(3), 539-554.

34. Cole, J., Scott, C.B, Johnson, M.M., Golightly, P.R., Carlson, J., Ming, M.M., Harpak, A., Kirkpatrick, M. (2024). The battle of the sexes in humans is highly polygenic. Preprint.

35. Cunningham, F., Allen, J. E., Allen, J., Alvarez-Jarreta, J., Amode, M. R., Armean, I. M., Austine-Orimoloye, O., Azov, A. G., Barnes, I., Bennett, R., … & Flicek, P. (2022). Ensembl 2022. *Nucleic Acids Research*, *50*(D1), D988-D995.

36. Hudson, R. R., Slatkin, M., & Maddison, W. P. (1992). Estimation of levels of gene flow from DNA sequence data. *Genetics*, *132*(2), 583-589.

37. Pfeifer, B., Wittelsbürger, U., Ramos-Onsins, S. E., & Lercher, M. J. (2014). PopGenome: an efficient Swiss army knife for population genomic analyses in R. *Molecular Biology and Evolution*, *31*(7), 1929-1936.