Population-specific causal disease effect sizes in functionally important regions impacted by selection

Huwenbo Shi[1,2,*], Steven Gazal[1,2], Masahiro Kanai[2,3,4,5,6], Evan M. Koch[7,8], Armin P. Schoech[1,2,9], Samuel S. Kim[1,2,10], Yang Luo[2,5,7,11,12], Tiffany Amariuta[2,5,11,12,13], Yukinori Okada[6,14], Soumya Raychaudhuri[2,5,7,11,12,15], Shamil R. Sunyaev[7,8], and Alkes L. Price[1,2,9,†]

[1]Department of Epidemiology, Harvard T.H. Chan School of Public Health, Boston, MA, USA
[2]Program in Medical and Population Genetics, Broad Institute of MIT and Harvard, Cambridge, MA, USA
[3]Analytic and Translational Genetics Unit, Massachusetts General Hospital, Boston, MA, USA
[4]Stanley Center for Psychiatric Research, Broad Institute of Harvard and MIT, Cambridge, MA, USA
[5]Department of Biomedical Informatics, Harvard Medical School, Boston, MA, USA
[6]Department of Statistical Genetics, Osaka University Graduate School of Medicine, Suita, Japan
[7]Division of Genetics, Brigham and Women's Hospital, Harvard Medical School, Boston, MA, USA
[8]Department of Medicine, Harvard Medical School, Boston, MA, USA
[9]Department of Biostatistics, Harvard T.H. Chan School of Public Health, Boston, MA, USA
[10]Department of Electrical Engineering and Computer Science, Massachusetts Institute of Technology, Cambridge, MA, USA
[11]Division of Rheumatology, Immunology, and Allergy, Brigham and Women's Hospital, Harvard Medical School, Boston, MA, USA
[12]Center for Data Sciences, Brigham and Women's Hospital, Harvard Medical School, Boston, MA, USA
[13]Graduate School of Arts andSciences, Harvard University, Cambridge, MA, USA
[14]Laboratory of Statistical Immunology, Immunology Frontier Research Center (WPI-IFReC), Osaka University, Suita, Japan
[15]Arthritis Research UK Centre for Genetics and Genomics, Centre for Musculoskeletal Research, Manchester Academic Health Science Centre, The University of Manchester, Manchester, UK

Correspondence: *hshi@hsph.harvard.edu (HS), †aprice@hsph.harvard.edu (ALP)

Abstract

Many diseases and complex traits exhibit population-specific causal effect sizes with trans-ethnic genetic correlations significantly less than 1, limiting trans-ethnic polygenic risk prediction. We developed a new method, S-LDXR, for stratifying squared trans-ethnic genetic correlation across genomic annotations, and applied S-LDXR to genome-wide association summary statistics for 30 diseases and complex traits in East Asians (EAS) and Europeans (EUR) (average N_{EAS}=93K, N_{EUR}=274K) with an average trans-ethnic genetic correlation of 0.83 (s.e. 0.01). We determined that squared trans-ethnic genetic correlation was 0.81× (s.e. 0.01) smaller than the genome-wide average at SNPs in the top quintile of background selection statistic, implying more population-specific causal effect sizes. Accordingly, causal effect sizes were more population-specific in functionally important regions, including coding, conserved, and regulatory regions. In analyses of regions surrounding specifically expressed genes, causal effect sizes were most population-specific for skin and immune genes and least population-specific for brain genes. Our results could potentially be explained by stronger gene-environment interaction at loci impacted by selection, particularly positive selection.

Introduction

Trans-ethnic genetic correlations are significantly less than 1 for many diseases and complex traits,[1–6] implying that population-specific causal disease effect sizes contribute to the incomplete portability of genome-wide association study (GWAS) findings and polygenic risk scores to non-European populations.[6–12] However, current methods for estimating genome-wide trans-ethnic genetic correlations assume the same trans-ethnic genetic correlation for all categories of SNPs,[2,5,13] providing little insight into why causal disease effect sizes are population-specific. Understanding the biological processes contributing to population-specific causal disease effect sizes can help inform polygenic risk prediction in non-European populations and alleviate health disparities.[6,14,15]

Here, we introduce a new method, S-LDXR, for stratifying squared trans-ethnic genetic correlation across functional categories of SNPs using GWAS summary statistics and population-matched linkage disequilibrium (LD) reference panels (e.g. the 1000 Genomes Project[16]); we stratify the *squared* trans-ethnic genetic correlation across functional categories to robustly handle noisy heritability estimates. We confirm that S-LDXR yields robust estimates in extensive simulations. We apply S-LDXR to 30 diseases and complex traits with GWAS summary statistics available in both East Asian (EAS) and European (EUR) populations, leveraging recent large studies in East Asian populations from the CONVERGE consortium and Biobank Japan;[17–19] we analyze a broad set of genomic annotations from the baseline-LD model,[20–22] as well as tissue-specific annotations based on specifically expressed gene sets.[23]

Results

Overview of methods

Our method (S-LDXR) for estimating stratified trans-ethnic genetic correlation is conceptually related to stratified LD score regression[20,21] (S-LDSC), a method for partitioning heritability from GWAS summary statistics. The S-LDSC method determines that a category of SNPs is enriched for heritability if SNPs with high LD to that category have higher expected χ^2 statistic than SNPs with low LD to that category. Analogously, the S-LDXR method determines that a category of SNPs is enriched for trans-ethnic genetic covariance if SNPs with high LD to that category have higher expected product of Z-scores than SNPs with low LD to that category. Unlike S-LDSC, S-LDXR models per-allele effect sizes (accounting for differences in minor allele frequency (MAF) between populations), and employs a shrinkage estimator to reduce noise.

In detail, the product of Z-scores of SNP j in two populations, $Z_{1j}Z_{2j}$, has the expectation

$$\mathrm{E}[Z_{1j}Z_{2j}] = \sqrt{N_1N_2}\sum_{C}\ell_{\times}(j,C)\theta_C \ , \tag{1}$$

where N_p is the sample size for population p; $\ell_{\times}(j,C) = \sum_k r_{1jk}r_{2jk}\sigma_{1j}\sigma_{2j}a_C(k)$ is the trans-ethnic LD score of SNP j with respect to annotation C, whose value for SNP k, $a_C(k)$, can be either binary or continuous; r_{pjk} is the LD (Pearson correlation) between SNP j and k in population p; σ_{pj} is the standard deviation of SNP j genotypes in population p; and θ_C represents the per-SNP contribution to trans-ethnic genetic covariance of the *per-allele* causal disease effect size of annotation C. Here, r_{pjk} and σ_{pj} can be estimated from population-matched reference panels (e.g. 1000 Genomes Project[16]). We estimate θ_C for each annotation C using weighted least square regression. Subsequently, we estimate the trans-ethnic genetic covariance of each binary annotation C ($\rho_g(C)$) as $\sum_{j\in C}\sum_{C'}a_{C'}(j)\theta_{C'}$, using coefficients ($\theta_{C'}$) for both binary and continuous-valued annotations C'; the heritabilities in each population ($h^2_{g1}(C)$ and $h^2_{g2}(C)$) are estimated analogously. We then estimate the stratified *squared* trans-ethnic genetic correlation, defined as

$$r^2_g(C) = \frac{\rho^2_g(C)}{h^2_{g1}(C)h^2_{g2}(C)} \ . \tag{2}$$

In this work, we only estimate $r^2_g(C)$ for SNPs with MAF greater than 5% in both populations. We estimate $r^2_g(C)$ instead of $r_g(C)$ to avoid bias (or undefined values) from computing square roots of noisy (possibly negative) heritability estimates, and use a bootstrap method[24] to correct for bias in estimating a ratio. We further employ a shrinkage estimator, with shrinkage parameter α (between 0 and 1, where larger values imply more shrinkage; the default value is 0.5), to reduce noise. We do not constrain estimates of $r_g^2(C)$ to their plausible range (between 0 and 1), which would introduce bias. We define the enrichment/depletion of squared trans-ethnic genetic correlation as $\lambda^2(C) = \frac{r_g^2(C)}{r_g^2}$, where r_g^2 is the genome-wide squared trans-ethnic genetic correlation; $\lambda^2(C)$ can be meta-analyzed across traits with different r_g^2. We compute standard errors via block-jackknife, as in previous work.[20] We estimate $\lambda^2(C)$ for binary annotations only, such as functional annotations[20] or quintiles of continuous-valued annotations.[21] Further details of the S-LDXR method are provided in the Methods section; we have publicly released open-source software implementing the method (see URLs). We note that all genetic correlations are defined using *causal* effect sizes, as opposed to joint-fit effect sizes.[2,5]

We apply S-LDXR to 62 annotations, defined in both EAS and EUR populations (Table S1, Figure S1, S2). 61 of these annotations (54 binary annotations and 7 continuous-valued annotations) are from the baseline-LD model (v1.1; see URLs), which includes a broad set of coding, conserved, regulatory and LD-related annotations; we modified the definition of two MAF-adjusted continuous-valued annotations (level of LD (LLD) and predicted allele age) to make them compatible with both populations. We also added one new continuous-valued annotation, SNP-specific F_{ST} between EAS and EUR populations. We did not include MAF bins from the baseline-LD model, due to the complexity of defining MAF bins in both populations. We refer to our final set of annotations as the baseline-LD-X model (Methods). We have publicly released all baseline-LD-X model annotations and LD scores for EAS and EUR populations (see URLs). We also apply S-LDXR to specifically expressed gene annotations for 53 tissues[23] (Table S2).

Simulations

We evaluated the accuracy of S-LDXR in simulations using genotypes that we simulated using HAPGEN2[25] from phased haplotypes of 481 EAS and 489 EUR individuals from the 1000 Genomes Project[16] (35,378 simulated EAS-like and 36,836 simulated EUR-like samples, after removing genetically related samples; ~2.5 million SNPs on chromosomes $1-3$) (Methods); we did not have access to individual-level EAS data at sufficient sample size to perform simulations with real genotypes. For each population, we randomly selected a subset of 500 simulated samples to serve as the reference panel for estimating LD scores. We performed both null simulations (heritable trait with functional enrichment but no enrichment/depletion of squared trans-ethnic genetic correlation; $\lambda^2(C) = 1$) and causal simulations ($\lambda^2(C) \neq 1$). In our main simulations, we randomly selected 10% of the SNPs as causal SNPs in both populations, set genome-wide heritability to 0.5 in each population, and adjusted genome-wide genetic covariance to attain a genome-wide r_g of 0.60 (unless otherwise indicated). In the null simulations, we used heritability enrichments from analyses of real traits in EAS samples to specify per-SNP causal effect size variances and covariances. In the causal simulations, we directly specified per-SNP causal effect size variances and covariances to attain $\lambda^2(C) \neq 1$ values from analyses of real traits, as these were difficult to attain using the heritability and trans-ethnic genetic covariance enrichments from analyses of real traits.

First, we assessed the accuracy of S-LDXR in estimating genome-wide trans-ethnic genetic correlation (r_g). Across a wide range of simulated r_g values (0.20 to 0.96), S-LDXR yielded approximately unbiased estimates and well-calibrated jackknife standard errors (Table S3, Figure S3).

Second, we assessed the accuracy of S-LDXR in estimating $\lambda^2(C)$ in quintiles of the 8 continuous-valued annotations of the baseline-LD-X model. We performed both null simulations ($\lambda^2(C) = 1$) and causal simulations ($\lambda^2(C) \neq 1$). Results are reported in Figure 1a and Tables S4 – S9 . At default parameter settings, S-LDXR yielded approximately unbiased estimates of $\lambda^2(C)$ for most annotations. As a secondary analysis, we tried varying the S-LDXR shrinkage parameter, α, which has a default value of 0.5. We determined that reducing the shrinkage parameter led to less accurate estimates of $\lambda^2(C)$ for annotations depleted for heritability, whereas increasing the shrinkage parameter biased results towards $\lambda^2(C) = 1$ in causal simulations (Figure S4, Tables S5, S8). Results were similar at other values of the proportion of causal SNPs (1% and 100%; Tables S4, S6, S7, S9). We also confirmed that S-LDXR produced well-calibrated jackknife standard errors (Tables S4-S9).

Finally, we assessed the accuracy of S-LDXR in estimating $\lambda^2(C)$ for the 28 main binary annotations of the baseline-LD-X model (inherited from the baseline model of ref.[20]). We discarded $\lambda^2(C)$ estimates with the highest standard errors (top 5%), as estimates with large standard errors (which are particularly common for annotations of small size) are uninformative for evaluating unbiasedness of the estimator (in analyses of real traits, trait-specific estimates with large standard errors are retained, but contribute very little to meta-analysis results). Results are reported in Figure 1b and Tables S5, S8. At default parameter settings, S-LDXR yielded approximately unbiased estimates of $\lambda^2(C)$ for functional annotations of large size in both null and causal simulations; however, estimates were slightly downward biased in null simulations for functional annotations of small size (e.g. 5' UTR; 0.5% of SNPs). This is likely because the bootstrap method for correcting bias in ratio estimation (Methods) has limited capability when heritability estimates in the denominator of Equation (2) are noisy,[24] as is the case for small annotations. Increasing the shrinkage parameter above its default value of 0.5 and extending the functional annotations by 500bp on each side[20] ameliorated the downward bias (and reduced standard errors) for annotations of small size in null simulations (Figure S5, S6);. However, increasing the shrinkage parameter also biased results towards the null ($\lambda^2(C) = 1$) in causal simulations (Tables S7, S8, S9), and $\lambda^2(C)$ estimates for the extended annotations are less biologically meaningful than for the corresponding main annotations. To ensure robust estimates, we focus on the 20 main binary annotations of large size ($>$ 1% of SNPs) in analyses of real traits (see below). Results were similar at other values of the proportion of causal SNPs (1% and 100%; Tables S4, S6, S7, S9). We also confirmed that S-LDXR produced well-calibrated jackknife standard errors (Tables S4-S9).

In summary, S-LDXR produced approximately unbiased estimates of enrichment/depletion of squared trans-ethnic genetic correlation in both null and causal simulations of both quintiles of continuous-valued annotations and binary annotations of large size ($>$ 1% of SNPs).

Analysis of baseline-LD-X model annotations across 30 diseases and complex traits

We applied S-LDXR to 30 diseases and complex traits with summary statistics in East Asians (average N = 93K) and Europeans (average N = 274K) available from Biobank Japan, UK Biobank, and other sources (Table S10 and Methods). First, we estimated the trans-ethnic genetic correlation (r_g) (as well as population-specific heritabilies) for each trait. Results are reported in Figure S7 and Table S10. The average r_g across 30 traits was 0.83 (s.e. 0.01) (average $r_g^2 = 0.69$ (s.e. 0.02)). 28 traits had $r_g < 1$, and 11 traits had r_g significantly less than 1 after correcting for 30 traits tested ($P < 0.05/30$); the lowest r_g was 0.34 (s.e. 0.07) for Major Depressive Disorder (MDD), although this may be confounded by different diagnostic criteria in the two populations.[26] These estimates were consistent with estimates obtained using Popcorn[2] (Figure S8) and those reported in previous studies.[2,5,6]

Second, we estimated the enrichment/depletion of squared trans-ethnic genetic correlation ($\lambda^2(C)$) in quintiles of the 8 continuous-valued annotations of the baseline-LD-X model, meta-analyzing results across traits; these annotations are moderately correlated (Figure 2a and Table S1). We used the default shrinkage parameter ($\alpha = 0.5$) in all analyses. Results are reported in Figure 2b and Table S11. We consistently observed a depletion of $r_g^2(C)$ ($\lambda^2(C) < 1$, implying more population-specific causal effect sizes) in functionally important regions. For example, we estimated $\lambda^2(C) = 0.81$ (s.e. 0.01) for SNPs in the top quintile of background selection statistic (defined as 1 − McVicker B statistic / 1000;[27] see ref.[21]); $\lambda^2(C)$ estimates were less than 1 for 27/30 traits (including 7 traits with two-tailed $p < 0.05/30$).

The background selection statistic quantifies the genetic distance of a site to its nearest exon; regions with high background selection statistic have higher per-SNP heritability, consistent with the action of selection, and are enriched for functionally important regions.[21] We observed the same pattern for CpG content and SNP-specific F_{st} (which are positively correlated with background selection statistic; Figure 2a) and the opposite pattern for nucleotide diversity (which is negatively correlated with background selection statistic). We also estimated $\lambda^2(C) = 0.85$ (s.e. 0.03) for SNPs in the top quintile of average LLD (which is positively correlated with background selection statistic), although these SNPs have *lower* per-SNP heritability due to a competing positive correlation with predicted allele age.[21] Likewise, we estimated $\lambda^2(C) = 0.83$ (s.e. 0.02) for SNPs in the *bottom* quintile of recombination rate (which is negatively correlated with background selection statistic), although these SNPs have average per-SNP heritability due to a competing negative correlation with average LLD.[21] However, $\lambda^2(C) < 1$ estimates for the bottom quintile of GERP (NS) (which is positively correlated with both background selection statistic and recombination rate) and the middle quintile of predicted allele age are more difficult to interpret. For all annotations analyzed, heritability enrichments did not differ significantly between EAS and EUR, consistent with previous studies.[19,28] Results were similar at a more stringent shrinkage parameter value ($\alpha = 1.0$; Figure S9), and for a meta-analysis across a subset of 20 approximately independent traits (Methods; Figure S10).

Finally, we estimated $\lambda^2(C)$ for the 28 main binary annotations of the baseline-LD-X model (Table S1), meta-analyzing results across traits. Results are reported in Figure 3a and Table S12. Our primary focus is on the 20 annotations of large size ($>$ 1% of SNPs), for which our simulations yielded robust estimates; results for remaining annotations are reported in Table S12. We consistently observed a depletion of $\lambda^2(C)$ (implying more population-specific causal effect sizes) within these annotations: 17 annotations had $\lambda^2(C) < 1$, and 8 annotations had $\lambda^2(C)$ significantly less than 1 after correcting for 20 annotations tested ($P < 0.05/20$); these annotations included Coding ($\lambda^2(C) = 0.90$ (s.e. 0.03)), Conserved ($\lambda^2(C) = 0.92$ (s.e. 0.02)), Promoter ($\lambda^2(C) = 0.88$ (s.e. 0.03)) and Super Enhancer ($\lambda^2(C) = 0.91$ (s.e. 0.01)), each of which was significantly enriched for per-SNP heritability, consistent with ref.[20]. For all annotations analyzed, heritability enrichments did not differ significantly between EAS and EUR (Figure 3a), consistent with previous studies.[19,28] Results were similar at a more stringent shrinkage parameter value ($\alpha = 1.0$; Figure S9), and for a meta-analysis across a subset of 20 approximately independent traits (Methods; Figure S11).

Since the functional annotations are moderately correlated with the 8 continuous-valued annotations (Table S1c, Figure S1), we investigated whether the depletions of squared trans-ethnic genetic correlation ($\lambda^2(C) < 1$) within the 20 binary annotations could be explained by the 8 continuous-valued annotations. For each binary annotation, we estimated its expected $\lambda^2(C)$ based on values of the 8 continuous-valued annotations for SNPs in the binary annotation (Methods), meta-analyzed this quantity across traits, and compared observed vs. expected $\lambda^2(C)$ (Figure 3b and Table S13). We observed strong concordance, with a slope of 0.63 (correlation of 0.56) across the 20 binary annotations. This implies that the depletions of $r_g^2(C)$ ($\lambda^2(C) < 1$) within binary annotations are largely explained by corresponding values of continuous-valued annotations.

In summary, our results show that causal disease effect sizes are more population-specific in functionally important regions impacted by selection. Further interpretation of these findings, including the role of positive and/or negative selection, is provided in the Discussion section.

Analysis of specifically expressed gene annotations

We analyzed 53 specifically expressed gene (SEG) annotations, defined in ref.[23] as ±100kb regions surrounding the top 10% of genes specifically expressed in each of 53 GTEx[29] tissues (Table S2), by applying S-LDXR with the baseline-LD-X model to the 30 diseases and complex traits (Table S10). We note that although SEG annotations were previously used to prioritize disease-relevant tissues based on disease-specific heritability enrichments,[19,23] enrichment/depletion of squared trans-ethnic genetic correlation ($\lambda^2(C)$) is standardized with respect to heritability, hence not expected to produce disease-specific signals. Thus, for each tissue, we meta-analyzed $\lambda^2(C)$ estimates across the 30 diseases and complex traits.

Results are reported in Figure 4a and Table S14. $\lambda^2(C)$ estimates were less than 1 for all 53 tissues and significantly less than 1 ($p < 0.05/53$) for 39 tissues, with statistically significant heterogeneity across tissues ($p < 10^{-20}$; Methods). The strongest depletions of squared trans-ethnic genetic correlation were observed in skin tissues (e.g. $\lambda^2(C) = 0.81$ (s.e. 0.02) for Skin Sun Exposed (Lower Leg)), Prostate and Ovary (e.g. $\lambda^2(C) = 0.82$ (s.e. 0.02) for Prostate) and immune-related tissues (e.g. $\lambda^2(C) = 0.83$ (s.e. 0.02) for Spleen), and the weakest depletions were observed in Testis ($\lambda^2(C) = 0.97$ (s.e. 0.02)) and brain tissues (e.g. $\lambda^2(C) = 0.96$ (s.e. 0.02) for Brain Nucleus Accumbens (Basal Ganglia)). Results were similar at less stringent and more stringent shrinkage parameter values ($\alpha = 0.0$ and $\alpha = 1.0$; Figures S12, S13 and Table S14). A comparison of 14 blood-related traits and 16 other traits yielded highly consistent $\lambda^2(C)$ estimates ($R = 0.82$; Figure S14, Table S15), confirming that these findings were not disease-specific.

These $\lambda^2(C)$ results were consistent with the higher background selection statistic[27] in Skin Sun Exposed (Lower Leg) ($R = 0.17$), Prostate ($R = 0.16$) and Spleen ($R = 0.14$) as compared to Testis ($R = 0.02$) and Brain Nucleus Accumbens (Basal Ganglia) ($R = 0.08$) (Figure S15, Table S2), and similarly for CpG content (Figure S16, Table S2). Although these results could in principle be confounded by gene size,[30] the low correlation between gene size and background selection statistic ($R = 0.06$) or CpG content ($R = -0.20$) (in ±100kb regions) implies limited confounding. We note the well-documented action of recent positive selection on genes impacting skin pigmentation[31–35] and the immune system;[31–34,36] we are not currently aware of any evidence of positive selection impacting Prostate and Ovary. We further note the well-documented action of negative selection on fecundity- and brain-related traits,[37–39] but it is possible that recent positive selection may more closely track differences in causal disease effect sizes across human populations, which have split relatively recently[40] (see Discussion).

More generally, since SEG annotations are moderately correlated with the 8 continuous-valued annotations (Figure S17, Table S2), we investigated whether these $\lambda^2(C)$ results could be explained by the 8 continuous-valued annotations (analogous to Figure 3b). Results are reported in Figure 4b and Table S16. We observed strong concordance, with a slope of 1.01 (correlation of 0.75) across the 53 SEG annotations. This implies that the depletions of $\lambda^2(C)$ within SEG annotations are explained by corresponding values of continuous-valued annotations.

In summary, our results show that causal disease effect sizes are more population-specific in regions surrounding specifically expressed genes. This effect was strongest in tissues impacted by positive selection (as opposed to negative selection), suggesting a possible connection between positive selection and population-specific causal effect sizes (see Discussion).

Discussion

We developed a new method (S-LDXR) for stratifying squared trans-ethnic genetic correlation across functional categories of SNPs that yields approximately unbiased estimates in extensive simulations. By applying S-LDXR to East Asian and European summary statistics across 30 diseases and complex traits, we determined that SNPs with high background selection statistic[27] have substantially lower squared trans-ethnic genetic correlation (vs. the genome-wide average), implying that causal effect sizes are more population-specific. Accordingly, squared trans-ethnic genetic correlations were substantially lower for SNPs in many functional categories. In analyses of specifically expressed gene annotations, we observed substantial depletion of squared trans-ethnic genetic correlation for SNPs near skin and immune-related genes, which are strongly impacted by recent positive selection, but not for SNPs near brain genes.

Reductions in trans-ethnic genetic correlation have several possible underlying explanations, including gene-environment (G×E) interaction, gene-gene (G×G) interaction, and dominance variation (but not differences in heritability across populations, which would not affect trans-ethnic genetic correlation and were not observed in our study). Given the increasing evidence of the role of G×E interaction in complex trait architectures,[41] and evidence that G×G interaction and dominance variation explain limited heritability,[42–44] we hypothesize that depletions of squared trans-ethnic genetic correlation in the top quintile of background selection statistic and in functionally important regions may be primarily attributable to stronger G×E interaction in these regions. Interestingly, a recent study on plasticity in Arabidopsis observed a similar phenomenon: lines with more extreme phenotypes exhibited stronger G×E interaction.[45] Distinguishing between stronger G×E interaction in regions impacted by selection and stronger G×E interaction in functionally important regions as possible explanations for our findings is a challenge, because functionally important regions are more strongly impacted by selection. To this end, we constructed an annotation that is similar to the background selection statistic but does not make use of recombination rate, instead relying solely on a SNP's physical distance to the nearest exon (Methods). Applying S-LDXR to the 30 diseases and complex traits using a joint model incorporating baseline-LD-X model annotations and the nearest exon annotation, the background selection statistic remained highly conditionally informative for trans-ethnic genetic correlation, whereas the nearest exon annotation was not conditionally informative (Table S17). This result implicates stronger G×E interaction in regions with reduced effective population size that are impacted by selection, and not just proximity to functional regions, in explaining depletions of squared trans-ethnic genetic correlation; however, we emphasize that selection acts on allele frequencies rather than causal effect sizes, and could help explain our findings only in conjunction with other explanations such as G×E interaction. Our results on specifically expressed genes implicate stronger G×E interaction near skin and immune genes and weaker G×E interaction near brain genes, potentially implicating positive selection (as opposed to negative selection). This conclusion is further supported by the lack of variation in squared trans-ethnic genetic correlation across genes in different deciles of probability of loss-of-function intolerance[46] (Methods, Figure S18, S19, Table S18). We conclude that depletions of squared trans-ethnic genetic correlation could potentially be explained by stronger G×E interaction at loci impacted by positive selection. We caution that other explanations are also possible; in particular, evolutionary modeling using an extension of the Eyre-Walker model[47] to two populations suggests that our results for the background selection statistic could also be consistent with negative selection (Supplementary Note, Figure S20, S21, Table S19). Additional information, such as genomic annotations that better distinguish different types of selection or data from additional diverse populations, may help elucidate the relationship between selection and population-specific causal effect sizes.

Our study has several implications. First, polygenic risk scores in non-European populations that make use of European training data[6,9] may be improved by reweighting SNPs based on the expected enrichment/depletion of squared trans-ethnic genetic correlation, helping to alleviate health disparities;[6,14,15] specifically, although the impact of population-specific LD patterns on trans-ethnic polygenic risk scores is well-documented,[6,9] population-specific causal effect sizes also merit thorough investigation. Second, modeling population-specific genetic architectures may improve trans-ethnic fine-mapping, moving beyond the standard assumption that all causal variants are shared across populations.[28,48] Third, modeling population-specific genetic architectures may also increase power in trans-ethnic meta-analysis,[49] e.g. by adapting MTAG[50] to two populations (instead of two traits). Fourth, it may be of interest to stratify G×E interaction effects[41] across genomic annotations. Fifth, the S-LDXR method could potentially be extended to stratify squared *cross-trait* genetic correlations[51] across genomic annotations.[52]

We note several limitations of this study. First, S-LDXR is designed for populations of homogeneous continental ancestry (e.g. East Asians and Europeans) and is not currently suitable for analysis of admixed populations[53] (analogous to LDSC and its published extensions[20,51,54]). However, a recently proposed extension of LDSC to admixed populations[55] could be incorporated into S-LDXR, enabling its application to the growing set of large studies in admixed populations.[10] Second, since S-LDXR applies shrinkage to reduce standard error in estimating stratified squared trans-ethnic genetic correlation and its enrichment, estimates are slightly conservative – true depletions of squared trans-ethnic genetic correlation in functionally important regions may be stronger than the estimated depletions. Third, the specifically expressed gene (SEG) annotations analyzed in this study are defined primarily based on gene expression measurements of Europeans.[23] However, genetic architectures of gene expression differ across diverse populations.[12,56,57] Thus, SEG annotations derived from gene expression data from diverse populations may provide additional insights into population-specific causal effect sizes. Fourth, we restricted our analyses to SNPs that were relatively common (MAF>5%) in both populations, due to the lack of a large LD reference panel for East Asians. Extending our analyses to lower-frequency SNPs may provide further insights into the role of negative selection in shaping population-specific genetic architectures, given the particular importance of negative selection for low-frequency SNPs.[58] Fifth, we did not consider population-specific variants in our analyses, due to the difficulty in defining trans-ethnic genetic correlation for population-specific variants;[2,5] a recent study[59] has reported that population-specific variants substantially limit trans-ethnic genetic risk prediction accuracy. Sixth, estimates of genome-wide trans-ethnic genetic correlation may be confounded by different trait definitions or diagnostic criteria in the two populations, particularly for major depressive disorder. However, this would not impact estimates of enrichment/depletion of squared trans-ethnic genetic correlation ($\lambda^2(C)$), which is defined relative to genome-wide values. Seventh, we have not pinpointed the exact underlying phenomena (e.g. environmental heterogeneity coupled with gene-environment interaction) that lead to population-specific causal disease effect sizes at functionally important regions. Despite these limitations, our study provides an improved understanding of the underlying biology that contribute to population-specific causal effect sizes, and highlights the need for increasing diversity in genetic studies.

URLs

- S-LDXR software: https://github.com/huwenboshi/s-ldxr/
- Python code for simulating GWAS summary statistics: https://github.com/huwenboshi/s-ldxr-sim/
- baseline-LD-X model annotations and LD scores: https://data.broadinstitute.org/alkesgroup/LDSCORE/baseline-LD-X/
- Distance to nearest exon annotation and LD scores: https://data.broadinstitute.org/alkesgroup/LDSCORE/baseline-LD-X/
- baseline-LD model annotations: https://data.broadinstitute.org/alkesgroup/LDSCORE/readme_baseline_versions
- 1000 Genomes Project: https://www.internationalgenome.org/
- PLINK2: https://www.cog-genomics.org/plink/2.0/
- HAPGEN2: https://mathgen.stats.ox.ac.uk/genetics_software/hapgen/hapgen2.html
- UCSC Genome Browser: https://genome.ucsc.edu/
- Exome Aggregation Consortium (ExAC): https://exac.broadinstitute.org/

Methods

Definition of stratified squared trans-ethnic genetic correlation

We model a complex phenotype in two populations using linear models, $\boldsymbol{Y}_1 = \boldsymbol{X}_1\boldsymbol{\beta}_1 + \boldsymbol{\epsilon}_1$ and $\boldsymbol{Y}_2 = \boldsymbol{X}_2\boldsymbol{\beta}_2 + \boldsymbol{\epsilon}_2$, where $\boldsymbol{Y}_1$ and $\boldsymbol{Y}_2$ are vectors of phenotype measurements of population 1 and population 2 with sample size N_1 and N_2 , respectively; $\boldsymbol{X}_1$ and $\boldsymbol{X}_2$ are mean-centered *but not normalized* genotype matrices at M SNPs in the two populations; $\boldsymbol{\beta}_1$ and $\boldsymbol{\beta}_2$ are *per-allele causal* effect sizes of the M SNPs; and $\boldsymbol{\epsilon}_1$ and $\boldsymbol{\epsilon}_2$ are environmental effects in the two populations. We assume that in each population, genotypes, causal effect sizes, and environmental effects are independent from each other. We assume that the per-allele effect size of SNP j in the two populations has variance and covariance,

$$\begin{gathered}\mathrm{Var}[\beta_{1j}] = \sum_C a_C(j)\tau_{1C},\ \mathrm{Var}[\beta_{2j}] = \sum_C a_C(j)\tau_{2C}, \\ \mathrm{Cov}[\beta_{1j}, \beta_{2j}] = \sum_C a_C(j)\theta_C,\end{gathered} \tag{3}$$

where $a_C(j)$ is the value of SNP j for annotation C, which can be binary or continuous-valued; τ_{1C} and τ_{2C} are the net contribution of annotation C to the variance of β_{1j} and β_{2j}, respectively; and θ_C is the net contribution of annotation C to the covariance of β_{1j} and β_{2j}.

We define stratified trans-ethnic genetic correlation of a binary annotation C (e.g. functional annotations[20] or quintiles of continuous-valued annotations[21]) as,

$$r_g(C) = \frac{\rho_g(C)}{\sqrt{h^2_{g1}(C)}\sqrt{h^2_{g2}(C)}}, \tag{4}$$

where $\rho_g(C) = \sum_{j\in C}\mathrm{Cov}[\beta_{1j}, \beta_{2j}] = \sum_{j\in C}\sum_{C'} a_{C'}(j)\theta_{C'}$ is the trans-ethnic genetic covariance of annotation C; and $h^2_{gp}(C) = \sum_{j\in C}\mathrm{Var}[\beta_{pj}] = \sum_{j\in C}\sum_{C'} a_{C'}(j)\tau_{pC'}$ is the heritability (sum of per-SNP variance of causal effect sizes) of annotation C in population p. Here, C' includes both binary and continuous-valued annotations. Since estimates of $h^2_{gp}(C)$ can be noisy (possibly negative), we estimate *squared* stratified trans-ethnic genetic correlation,

$$r^2_g(C) = \frac{\rho^2_g(C)}{h^2_{g1}(C)h^2_{g2}(C)}, \tag{5}$$

to avoid bias or undefined values in the square root. In this work, we only estimate $r^2_g(C)$ for SNPs with minor allele frequency (MAF) greater than 5% in both populations. To assess whether causal effect sizes are more or less correlated for SNPs in annotation C compared with the genome-wide average, r_g^2, we define the enrichment/depletion of stratified squared trans-ethnic genetic correlation as

$$\lambda^2(C) = \frac{r_g^2(C)}{r_g^2}. \tag{6}$$

We meta-analyze $\lambda^2(C)$ instead of $r_g^2(C)$ across diseases and complex traits. We note that the average value of $\lambda^2(C)$ across quintiles of continuous-valued annotations is not necessarily equal to 1, as squared trans-ethnic genetic correlation is a non-linear quantity.

S-LDXR method

S-LDXR is conceptually related to stratified LD score regression[20,21] (S-LDSC), a method for stratifying heritability from GWAS summary statistics, to two populations. The S-LDSC method determines that a category of SNPs is enriched for heritability if SNPs with high LD to that category have higher expected χ^2 statistic than SNPs with low LD to that category. Analogously, the S-LDXR method determines that a category of SNPs is enriched for trans-ethnic genetic covariance if SNPs with high LD to that category have higher expected product of Z-scores than SNPs with low LD to that category.

S-LDXR relies on the regression equation

$$\mathrm{E}[Z_{1j}Z_{2j}] = \sqrt{N_1N_2}\sum_C \ell_\times(j,C)\theta_C \tag{7}$$

to estimate θ_C, where Z_{pj} is the Z-score of SNP j in population p; $\ell_\times(j,C) = \sum_k r_{1jk}r_{2jk}\sigma_{1j}\sigma_{2j}a_C(k)$ is the trans-ethnic LD score of SNP j with respect to annotation C, whose value for SNP k, $a_C(k)$, can be either binary or continuous; r_{pjk} is the LD between SNP j and k in population p; and σ_{pj} is the standard deviation of SNP j in population p. We obtain unbiased estimates of $\ell_\times(j,C)$ using genotype data of 481 East Asian and 489 European samples in the 1000 Genomes Project.[16] To account for heteroscedasticity and increase statistical efficiency, we use weighted least square regression to estimate θ_C. We include only well-imputed (imputation INFO>0.9) and common (MAF>5% in both populations) SNPs that are present in HapMap 3[60] in the regression, as in our previous work.[20,51,54] We use regression equations analogous to those described in ref.[20] to estimate τ_{1C} and τ_{2C}.

Let $\hat{\tau}_{1C}$, $\hat{\tau}_{1C}$, and $\hat{\theta}_C$ be the estimates of τ_{1C}, τ_{1C}, and θ_C, respectively. For each binary annotation C, we estimate the stratified heritability of annotation C in each population,

$h^2_{g1}(C)$ and $h^2_{g2}(C)$, and trans-ethnic genetic covariance, $\rho_g(C)$, as

$$\hat{h}^2_{g2}(C) = \sum_{j\in C}\sum_{C'} a_{jC'}\hat{\tau}_{2C'},\ \hat{h}^2_{g1}(C) = \sum_{j\in C}\sum_{C'} a_{jC'}\hat{\tau}_{1C'},\ \hat{\rho}_g(C) = \sum_{j\in C}\sum_{C'} a_{jC'}\hat{\theta}_{C'}, \tag{8}$$

respectively, using coefficients ($\tau_{1C'}$, $\tau_{2C'}$, and $\theta_{C'}$) of both binary and continuous-valued annotations. We then estimate $r^2_g(C)$ as

$$\hat{r}^2_g(C) = \frac{\hat{\rho}^2_g(C) - \hat{\text{S.E.}}^2[\hat{\rho}_g(C)]}{\hat{h}^2_{g1}(C)\hat{h}^2_{g2}(C) - \hat{\text{Cov}}[\hat{h}^2_{g1}(C), \hat{h}^2_{g2}(C)]} - \hat{\text{bias}}(C), \tag{9}$$

where $\hat{\text{bias}}(C)$ is obtained using bootstrap to correct for bias in estimating the ratio.[24] We do not constrain the estimate of $r^2_g(C)$ to its plausible range of $[-1, 1]$ to be unbiased. Subsequently, we obtain enrichment of stratified squared trans-ethnic genetic correlation as

$$\hat{\lambda}^2(C) = \frac{\hat{r}^2_g(C)}{\hat{r}^2_g}, \tag{10}$$

where $\hat{r}^2_g$ is the estimate of genome-wide squared trans-ethnic genetic correlation r^2_g. We use block jackknife over 200 non-overlapping and equally sized blocks to obtain standard error of all estimates. The standard error of $\lambda^2(C)$ typically depends on sample size of the GWAS and overall heritability of annotation C in the two populations (i.e. $h^2_{g1}(C)$ and $h^2_{g2}(C)$).

To assess the informativeness of each annotation in explaining disease heritability and trans-ethnic genetic covariance, we define standardized annotation effect size on heritability and trans-ethnic genetic covariance for each annotation C analogous to ref.[21],

$$\begin{gathered}\tau^*_{1C} = \frac{Mh^2_{g1}}{h^2_{g1}(C)} \times \sigma_C \times \tau_{1C},\ \tau^*_{2C} = \frac{Mh^2_{g2}}{h^2_{g2}(C)} \times \sigma_C \times \tau_{2C},\\ \theta^*_C = \frac{M\rho_g}{\rho_g(C)} \times \sigma_C \times \theta_C,\end{gathered} \tag{11}$$

where τ^*_{1C}, τ^*_{2C}, and θ^*_C represent proportionate change in per-SNP heritability in population 1 and 2 and trans-ethnic genetic covariance, respectively, per standard deviation increase in annotation C; τ_{1C}, τ_{2C}, and θ_C are the corresponding unstandardized effect sizes, defined in Equation (3); and σ_C is the standard deviation of annotation C.

We provide a more detailed description of the method, including derivations of the regression equation and unbiased estimators of the LD scores, in the **Supplementary Note**.

S-LDXR shrinkage estimator

Estimates of $r^2_g(C)$ can be imprecise with large standard errors if the denominator, $h^2_{g1}(C)h^2_{g2}(C)$, is close to zero and noisily estimated. This is especially the case for annotations of small size ($<1\%$ SNPs). We introduce a shrinkage estimator to reduce the standard error in estimating $r^2_g(C)$.

Briefly, we shrink the estimated per-SNP heritability and trans-ethnic genetic covariance of annotation C towards the genome-wide averages, which are usually estimated with smaller standard errors, prior to estimating $r^2_g(C)$. In detail, let M_C be the number of SNPs in annotation C, we shrink $\frac{\hat{h}^2_{1g}(C)}{M_C}$, $\frac{\hat{h}^2_{2g}(C)}{M_C}$, and $\frac{\hat{\rho}_g(C)}{M_C}$ towards $\frac{\hat{h}^2_{1g}}{M}$, $\frac{\hat{h}^2_{2g}}{M}$, and $\frac{\hat{\rho}_g}{M}$, respectively, where $\hat{h}^2_{g1}$, $\hat{h}^2_{g2}$, $\hat{\rho}_g$ are the genome-wide estimates, and M the total number of SNPs. We obtain the shrinkage as follows. Let $\gamma_1 = 1/\left(1+\alpha\frac{\mathrm{Var}\left[\hat{h}^2_{g1}(C)\right]}{\mathrm{Var}\left[\hat{h}^2_{g1}\right]}\frac{M}{M_C}\right)$, $\gamma_2 = 1/\left(1+\alpha\frac{\mathrm{Var}\left[\hat{h}^2_{g2}(C)\right]}{\mathrm{Var}\left[\hat{h}^2_{g2}\right]}\frac{M}{M_C}\right)$, and $\gamma_3 = 1/\left(1+\alpha\frac{\mathrm{Var}[\hat{\rho}_g(C)]}{\mathrm{Var}[\hat{\rho}_g]}\frac{M}{M_C}\right)$ be the shrinkage obtained separately for $\hat{h}^2_{g1}(C)$, $\hat{h}^2_{g2}(C)$ and $\hat{\rho}_g(C)$, respectively, where $\alpha \in [0,1]$ is the shrinkage parameter adjusting magnitude of shrinkage. We then choose the most stringent shrinkage, $\gamma = \min\{\gamma_1, \gamma_2, \gamma_3\}$, as the final shared shrinkage for both heritability and trans-ethnic genetic covariance.

We shrink heritability and trans-ethnic genetic covariance of annotation C using γ as, $\bar{h}^2_{g1}(C) = M_C\left(\gamma\frac{\hat{h}^2_{g1}(C)}{M_C} + (1-\gamma)\frac{\hat{h}^2_{g1}}{M}\right)$, $\bar{h}^2_{g2}(C) = M_C\left(\gamma\frac{\hat{h}^2_{g2}(C)}{M_C} + (1-\gamma)\frac{\hat{h}^2_{g2}}{M}\right)$, and $\bar{\rho}_g(C) = M_C\left(\gamma\frac{\hat{\rho}_g(C)}{M_C} + (1-\gamma)\frac{\hat{\rho}_g}{M}\right)$, where $\bar{h}^2_{g1}(C)$, $\bar{h}^2_{g2}(C)$, and $\bar{\rho}_g(C)$ are the shrunk counterparts of $\hat{h}^2_{g1}(C)$, $\hat{h}^2_{g2}(C)$, and $\hat{\rho}_g(C)$, respectively. We shrink $\hat{r}^2_g(C)$ by substituting $\hat{h}^2_{g1}(C)$, $\hat{h}^2_{g2}(C)$, and $\hat{\rho}_g(C)$ with $\bar{h}^2_{g1}(C)$, $\bar{h}^2_{g2}(C)$, $\bar{\rho}_g(C)$, respectively, in Equation (9), to obtain its shrunk counterpart, $\bar{r}^2_g(C)$. Finally, we shrink $\hat{\lambda}^2(C)$, by plugging in $\bar{r}^2_g(C)$ in Equation (10) to obtain its shrunk counterpart, $\bar{\lambda}^2(C)$. We recommend $\alpha = 0.5$ as the default shrinkage parameter value, as this value provides robust estimates of $\lambda^2(C)$ in simulations.

Baseline-LD-X model

We include a total of 54 binary functional annotations in the baseline-LD-X model. These include 53 annotations introduced in ref.,[20] which consists of 28 main annotations including conserved annotations (e.g. Coding, Conserved) and epigenomic annotations (e.g. H3K27ac, DHS, Enhancer) derived from ENCODE[61] and Roadmap,[62] 24 500-base-pair-extended main annotations, and 1 annotation containing all SNPs. We note that although chromatin accessibility can be population-specific, the fraction of such regions is small.[63] Following ref,[21] we created an additional annotation for all genomic positions with number of rejected substitutions[64] greater than 4. Further information for all functional annotations included in the baseline-LD-X model is provided in Table S1a.

We also include a total of 8 continuous-valued annotations in the baseline-LD-X model. First, we include 5 continuous-valued annotations introduced in ref.[21] (see URLs), without modification: background selection statistic,[27] CpG content (within a ±50 kb window), GERP (number of substitutation) score,[64] nucleotide diversity (within a ±10 kb window), and Oxford map recombination rate (within a ±10 kb window).[65] Second, we include 2 minor allele frequency (MAF) adjusted annotations introduced in ref.,[21] with modification: level of LD (LLD) and predicted allele age. We created analogous annotations applicable to both East Asian and European populations. To create an analogous LLD annotation, we estimated LD scores for each population using LDSC,[54] took the average across populations, and then quantile-normalized the average LD scores using 10 average MAF bins. We call this annotation "average level of LD". To create analogous predicted allele age annotation, we quantile-normalized allele age estimated by ARGweaver[66] across 54 multi-ethnic genomes using 10 average MAF bins. Finally, we include 1 continuous-valued annotation based on F_{ST} estimated by PLINK2,[67] which implements the Weir & Cockerham estimator of F_{ST}.[68] Further information for all continuous-valued annotations included in the baseline-LD-X model is provided in Table S1b.

Code and data availability

Python code implementing S-LDXR is available at https://github.com/huwenboshi/s-ldxr. Python code for simulating GWAS summary statistics under the baseline-LD-X model is available at https://github.com/huwenboshi/s-ldxr-sim. baseline-LD-X model annotations and LD scores are available at https://data.broadinstitute.org/alkesgroup/LDSCORE/baseline-LD-X/.

Simulations

We used simulated East Asian (EAS) and European (EUR) genotype data to assess the performance our method, as we did not have access to real EAS genotype data at sufficient sample size to perform simulations with real genotypes. We simulated genotype data for 100,000 East-Asian-like and 100,000 European-like individuals using HAPGEN2[25] (see URLs), starting from phased haplotypes of 481 East Asians and 489 Europeans individuals available in the 1000 Genomes Project[16] (see URLs), restricting to ∼2.5 million SNPs on chromosome $1-3$ with minor allele count greater than 5 in either population. Since excessive relatedness arose from HAPGEN2 simulations,[2] we used PLINK2[67] (see URLs) to remove simulated individuals with genetic relatedness greater than 0.05. From the filtered set of individuals, we randomly selected 500 individuals in each simulated population to serve as reference panels, and used the remaining 35,378 East-Asian-like and 36,836 European-like individuals to simulate GWAS summary statistics.

We performed both null simulations, where enrichment of squared trans-ethnic genetic correlation, $\lambda^2(C)$, is 1 across all functional annotations, and causal simulations, where $\lambda^2(C)$ varies across annotations, under various degrees of polygenicity (1%, 10%, and 100% causal SNPs). In the null simulations, we set τ_{1C}, τ_{2C}, θ_C to be the meta-analyzed τ_C in real-data analyses of EAS GWASs, and followed Equation (3) to obtain variance, $\text{Var}[\beta_{1j}]$ and $\text{Var}[\beta_{2j}]$, and covariance, $\text{Cov}[\beta_{1j}, \beta_{2j}]$, of per-SNP causal effect sizes β_{1j}, β_{2j}, setting all negative per-SNP variance and covariance to 0. In the causal simulations, we directly specified per-SNP causal effect size variances and covariances using self-devised τ_{1C}, τ_{2C}, and θ_C coefficients, to attain $\lambda^2(C) \neq 1$, as these were difficult to attain using the coefficients from analyses of real traits.

We randomly selected a subset of SNPs to be causal for both populations, and set $\text{Var}[\beta_{1j}]$, $\text{Var}[\beta_{2j}]$, and $\text{Cov}[\beta_{1j}, \beta_{2j}]$ to be 0 for all remaining non-causal SNPs. We scaled the trans-ethnic genetic covariance to attain a desired genome-wide r_g. Next, we drew causal effect sizes of each causal SNP j in the two populations from the bi-variate Gaussian distribution,

$$\begin{bmatrix} \beta_{1j} \\ \beta_{2j} \end{bmatrix} \sim N\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \text{Var}[\beta_{1j}] & \text{Cov}[\beta_{1j}, \beta_{2j}] \\ \text{Cov}[\beta_{1j}, \beta_{2j}] & \text{Var}[\beta_{2j}] \end{bmatrix} \right), \tag{12}$$

and scaled the drawn effect sizes to match the desired total heritability and trans-ethnic genetic covariance. We simulated genetic component of the phenotype in population p as $\boldsymbol{X}_p\boldsymbol{\beta}_p$, where $\boldsymbol{X}_p$ is column-centered genotype matrix, and drew environmental effects, $\boldsymbol{\epsilon}_p$, from the Gaussian distribution, $N\left(0, 1 - \text{Var}[\boldsymbol{X}_p\boldsymbol{\beta}_p]\right)$, such that the total phenotypic variance in each population is 1. Finally, we simulated GWAS summary association statistics for population p, $\boldsymbol{Z}_p$, as $Z_{pj} = \frac{\boldsymbol{X}_{pj}^\intercal \boldsymbol{Y}_p}{\sqrt{N_p}\sigma_{pj}}$, where σ_{pj} is the standard deviation of SNP j in population p. We have publicly released Python code for simulating GWAS summary statistics for 2 populations (see URLs).

Summary statistics for 30 diseases and complex traits

We analyzed GWAS summary statistics of 30 diseases and complex traits, primarily from UK Biobank,[69] Biobank Japan,[19] and CONVERGE.[17] These include: atrial fibrillation (AF),[70,71] age at menarche(AMN),[72,73] age at menopause (AMP),[72,73] basophil count(BASO),[19,74] body mass index (BMI),[19,75] blood sugar(BS),[19,75] diastolic blood pressure (DBP),[19,75] eosinophil count(EO),[19,75] estimated glomerular filtration rate (EGFR),[19,76] hemoglobin A1c(HBA1C),[19,75] height (HEIGHT),[75,77] high density lipoprotein (HDL),[19,75] hemoglobin (HGB),[19,74] hematocrit (HTC),[19,74] low density lipoprotein (LDL),[19,75] lymphocyte count(LYMPH),[19,75] mean corpuscular hemoglobin (MCH),[19,75] mean corpuscular hemoglobin concentration (MCHC),[19,74] mean corpuscular volume (MCV),[19,74] major depressive disorder (MDD),[17,78] monocyte count (MONO),[19,75] neutrophil count(NEUT),[19,74] platelet count (PLT),[19,75] rheumatoid arthritis(RA),[79] red blood cell count (RBC),[19,75] systolic blood pressure (SBP),[19,75] type 2 diabetes (T2D),[80,81] total cholesterol (TC),[19,75] triglyceride (TG),[19,75] and white blood cell count (WBC).[19,75] Further information for the GWAS summary statistics analyzed is provided in Table S10. In our main analyses, we performed random-effect meta-analysis to aggregate results across all 30 diseases and complex traits. We also defined a set of 20 approximately independent diseases and complex traits with cross-trait r_g^2 (estimated using cross-trait LDSC[51]) less than 0.25 in both populations: AF, AMN, AMP, BASO, BMI, EGFR, EO, HBA1C, HEIGHT, HTC, LYMPH, MCHC, MCV, MDD, NEUT, PLT, RA, SBP, TC, TG.

Expected enrichment of stratified squared trans-ethnic genetic correlation from 8 continuous-valued annotations

To obtain expected enrichment of squared trans-ethnic genetic correlation of a binary annotation C, $\lambda^2(C)$, from 8 continuous-valued annotations, we first fit the S-LDXR model using these 8 annotations together with the base annotation for all SNPs, yielding coefficients, $\tau_{1C'}$, $\tau_{2C'}$, and $\theta_{C'}$, for a total of 9 annotations. We then use Equation (3) to obtain per-SNP variance and covariance of causal effect sizes, β_{1j} and β_{1j}, substituting τ_{1C}, τ_{2C}, θ_C with $\tau_{1C'}$, $\tau_{2C'}$, and $\theta_{C'}$, respectively. We apply shrinkage with default parameter setting ($\alpha = 0.5$), and use Equation (9) and (10) to obtain expected stratified squared trans-ethnic genetic correlation, $r_g^2(C)$, and subsequently $\lambda^2(C)$.

Analysis of specifically expressed gene annotations

We obtained 53 specifically expressed gene (SEG) annotations, defined in ref.[23] as ±100k-base-pair regions surrounding genes specifically expressed in each of 53 GTEx[29] tissues. A list of the SEG annotations is provided in Table S2. Correlations between SEG annotations and the 8 continuous-valued annotations are reported in Figure S17 and Table S2. Most SEG annotations are moderately correlated with the background selection statistic and CpG content annotations.

To test whether there is heterogeneity in enrichment of squared trans-ethnic genetic correlation, $\lambda^2(C)$, across the 53 SEG annotations, we first computed the average $\lambda^2(C)$ across the 53 annotations, $\bar{\lambda}^2(C)$, using fixed-effect meta-analysis. We then computed the test statistic $\sum_{i=1}^{53} \frac{\left(\hat{\lambda}^2(C_i)-\bar{\lambda}^2(C_i)\right)^2}{\mathrm{Var}[\hat{\lambda}^2(C_i)]}$, where C_i is the i-th SEG annotation, and $\hat{\lambda}^2(C_i)$ the estimated $\lambda^2(C)$. We computed a p-value for this test statistic based on a χ^2 distribution with 53 degrees of freedom.

Analysis of distance to nearest exon annotation

We created a continuous-valued annotation, named "distance to nearest exon annotation", based on a SNP's physical distance (number of base pairs) to its nearest exon, using 233,254 exons defined on the UCSC genome browser[82] (see URLs). This annotation is moderately correlated with the background selection statistic annotation[21] ($R = -0.21$), defined as (1 - McVicker B statistic / 1000), where the McVicker B statistic quantifies a site's genetic distance to its nearest exon.[27] We have publicly released this annotation (see URLs).

To assess the informativeness of functionally important regions versus regions impacted by selection in explaining the depletions of squared trans-ethnic genetic correlation, we applied S-LDXR on the distance to nearest exon annotation together with the baseline-LD-X model annotations. We used both enrichment of squared trans-ethnic genetic correlation ($\lambda^2(C)$) and standardized annotation effect size (τ^*_{1C}, τ^*_{2C}, and θ^*_C) to assess informativeness.

Analysis of probability of loss-of-function intolerance decile gene annotations

We created 10 annotations based on genes in deciles of probability of being loss-of-function intolerant (pLI) (see URLs), defined as the probability of assigning a gene into haplosufficient regions, where protein-truncating variants are depleted.[46] Genes with high pLI (e.g. > 0.9) have higly constrained functionality, and therefore mutations in these genes are subject to negative selection. We included SNPs within a 100kb-base-pair window around each gene, following ref.[23] A correlation heat map between pLI decile gene annotations and the 8 continuous-valued annotations is provided in Figure S18. All pLI decile gene annotations are moderately correlated with the background selection statistic and CpG content annotations.

Acknowledgements

We are grateful to L. O'Connor, H. Finucane, D. Kassler, S. Mallick, N. Patterson, B. Neale, R. Walters, K. Siewert, A. Martin, B. Brown, F. Hormozdiari, M. Hujoel, and B. Pasaniuc for helpful discussions. This research was conducted using the UK Biobank Resource under Application 16549 and was funded by NIH grants R01 HG006399, U01 HG009379, R01 MH107649 and R01 MH101244.

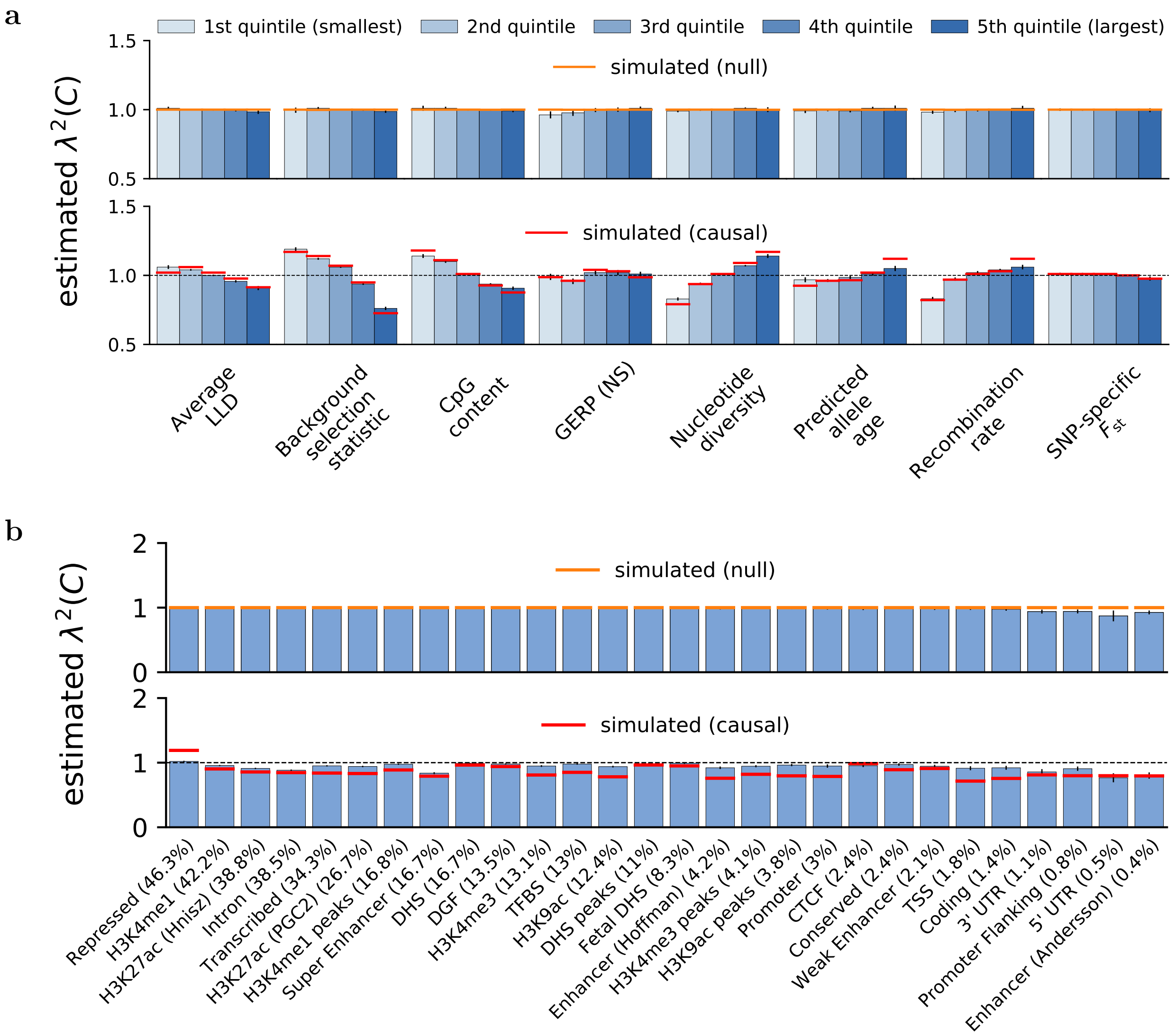

Figure 1: **Accuracy of S-LDXR in null and causal simulations.** We report estimates of the enrichment/depletion of squared trans-ethnic genetic correlation ($\lambda^2(C)$) in both null and causal simulations, for (a) quintiles of 8 continuous-valued annotations and (b) 28 main binary annotations (sorted by proportion of SNPs, displayed in parentheses). Results are averaged across 1,000 simulations. Error bars denote ±1.96× standard error. Numerical results are reported in Table S5 and S8.

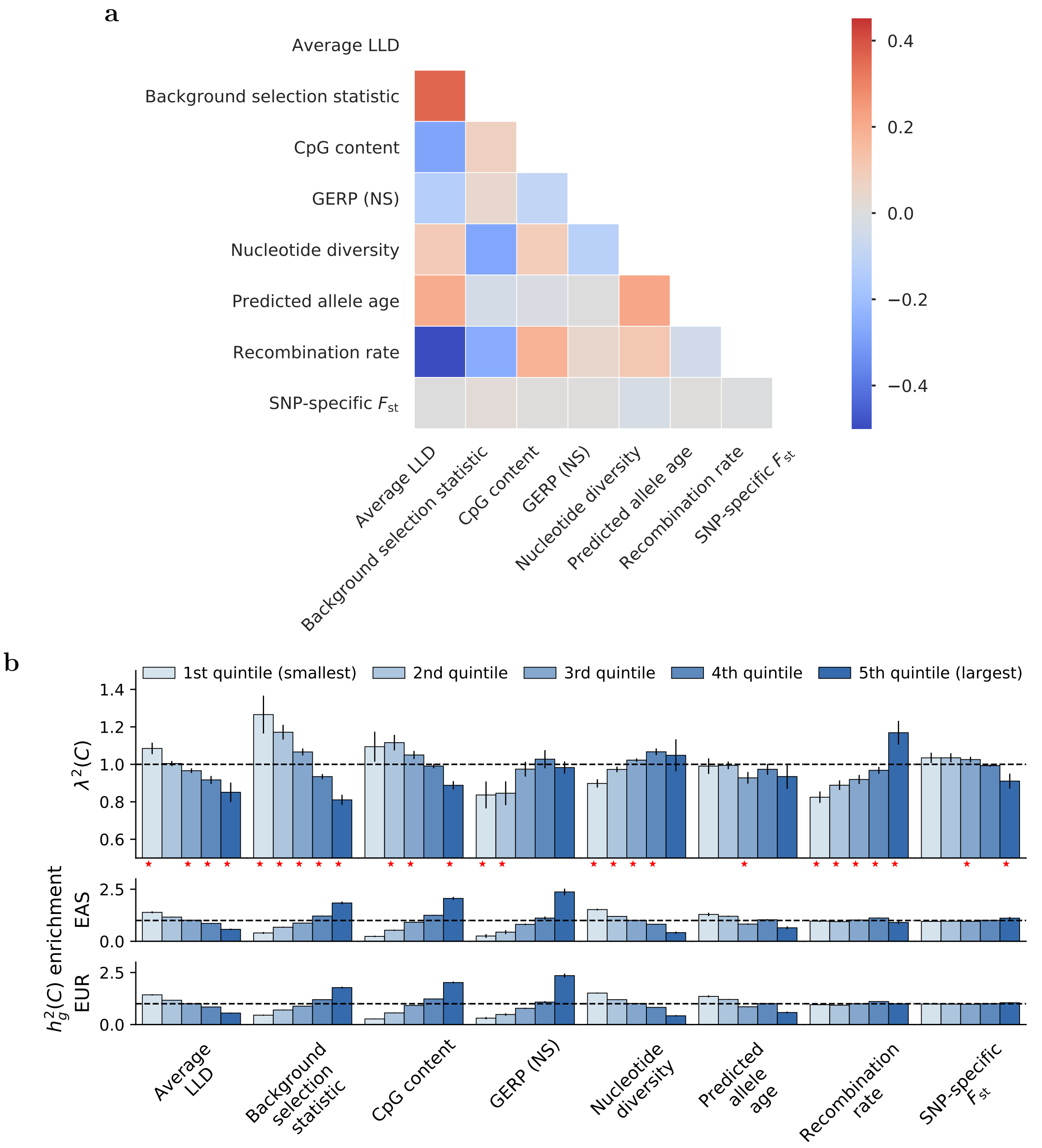

Figure 2: **S-LDXR results for quintiles of 8 continuous-valued annotations across 30 diseases and complex traits.** (a) We report correlations between each continuous-valued annotation; diagonal entries are not shown. Numerical results are reported in Table S1. (b) We report estimates of the enrichment/depletion of squared trans-ethnic genetic correlation ($\lambda^2(C)$), as well as population-specific estimates of heritability enrichment, for quintiles of each continuous-valued annotation. Results are meta-analyzed across 30 diseases and complex traits. Error bars denote ±1.96× standard error. Red stars (⋆) denote two-tailed p<0.05/40. Numerical results are reported in Table S11.

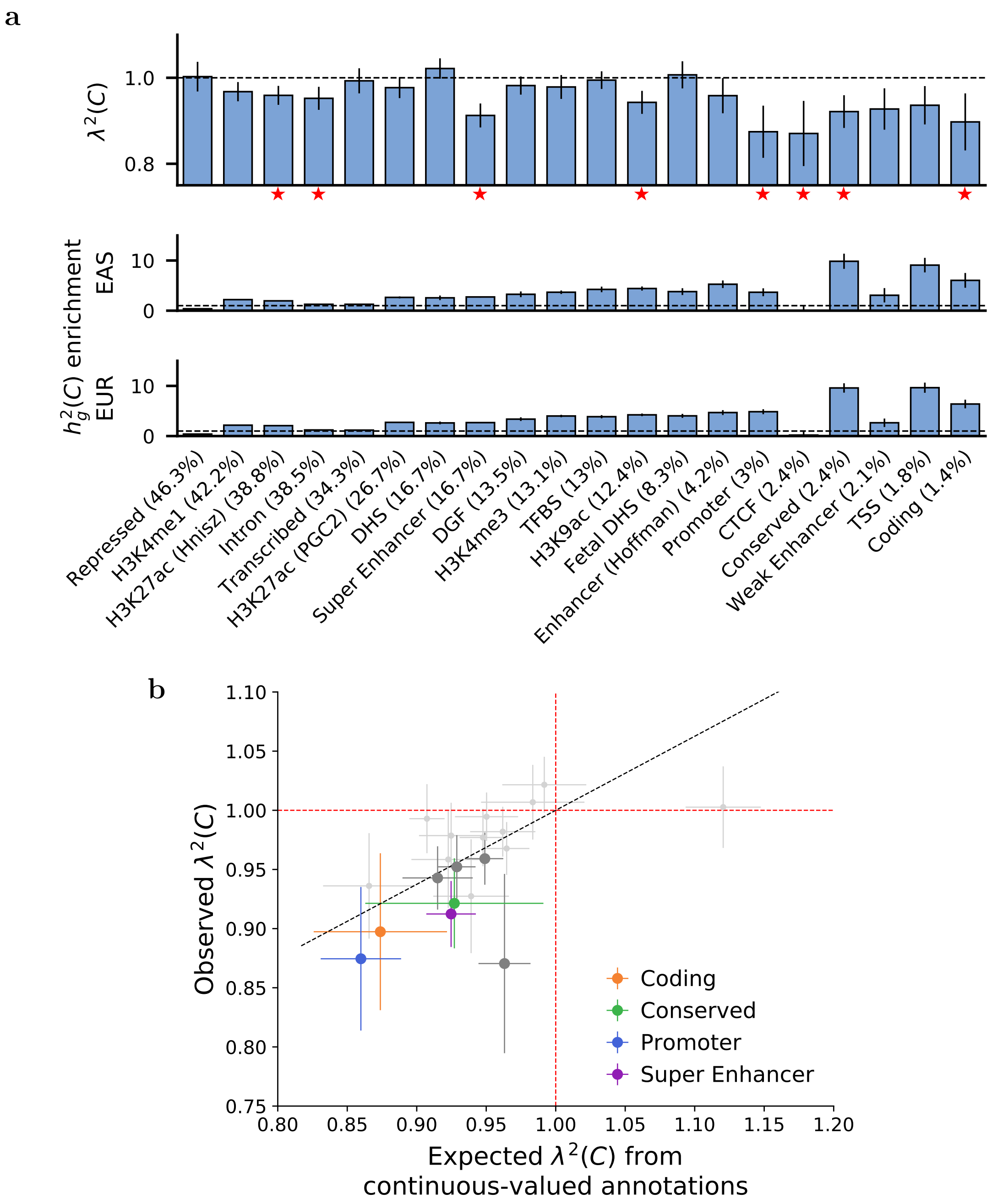

Figure 3: **S-LDXR results for 20 binary functional annotations across 30 diseases and complex traits.** (a) We report estimates of the enrichment/depletion of squared trans-ethnic genetic correlation ($\lambda^2(C)$), as well as population-specific estimates of heritability enrichment, for each binary annotation (sorted by proportion of SNPs, displayed in parentheses). Results are meta-analyzed across 30 diseases and complex traits. Error bars denote ±1.96× standard error. Red stars ($\star$) denote two-tailed p<0.05/20. Numerical results are reported in Table S12. (b) We report observed $\lambda^2(C)$ vs. expected $\lambda^2(C)$ based on 8 continuous-valued annotations, for each binary annotation. Results are meta-analyzed across 30 diseases and complex traits. Error bars denote ±1.96× standard error. Annotations for which $\lambda^2(C)$ is significantly different from 1 ($p<0.05/20$) are denoted in color (see legend) or dark gray. The dashed black line (slope=0.63) denotes a regression of observed $\lambda(C) - 1$ vs. expected $\lambda(C) - 1$ with intercept constrained to 0. Numerical results are reported in Table S13.

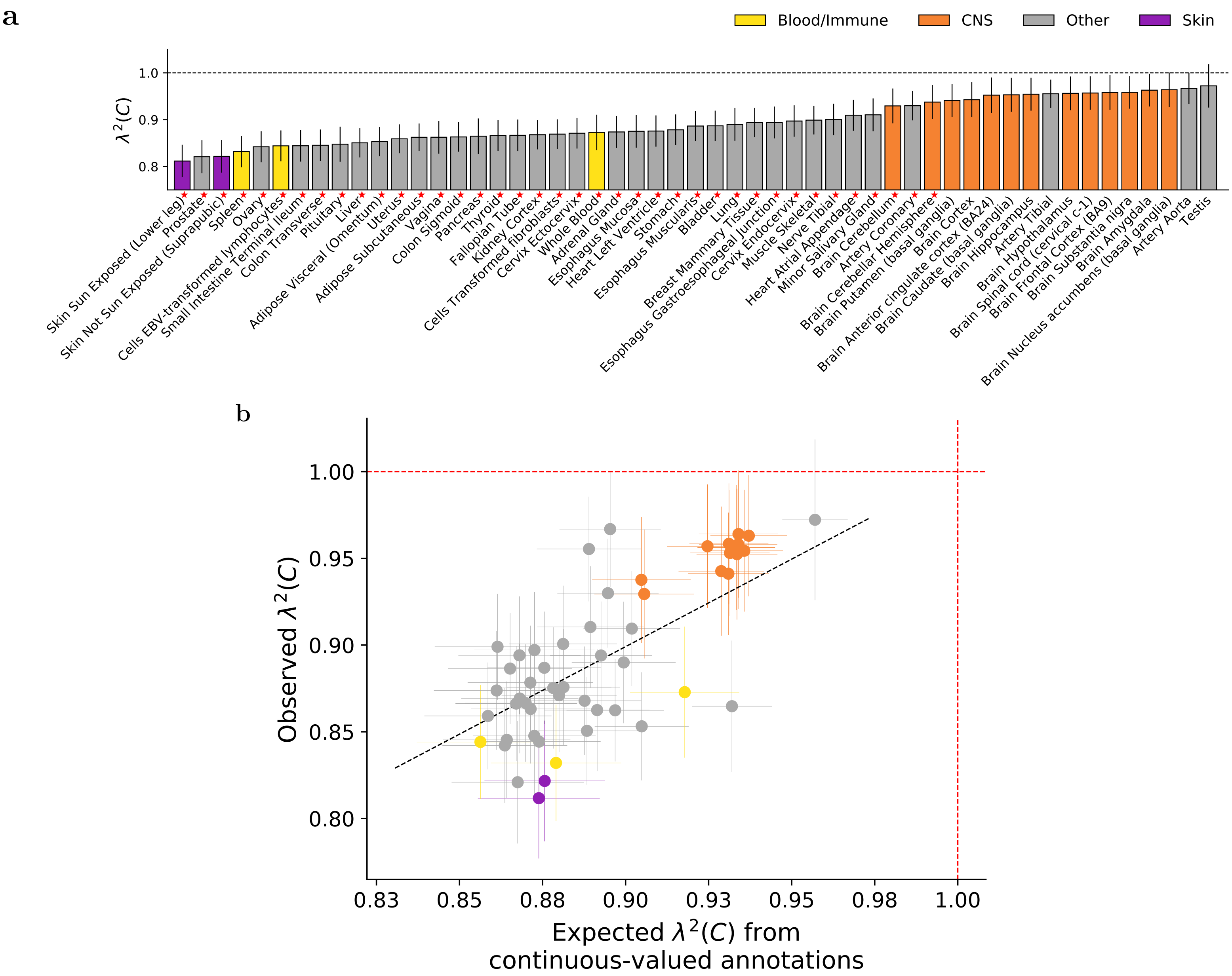

Figure 4: **S-LDXR results for 53 specifically expressed gene (SEG) annotations across 30 diseases and complex traits.** (a) We report estimates of the enrichment/depletion of squared trans-ethnic genetic correlation ($\lambda^2(C)$) for each SEG annotation (sorted by $\lambda^2(C)$). Results are meta-analyzed across 30 diseases and complex traits. Error bars denote ±1.96× standard error. Red stars (★) denote two-tailed p<0.05/53. Numerical results are reported in Table S14. (b) We report observed $\lambda^2(C)$ vs. expected $\lambda^2(C)$ based on 8 continuous-valued annotations, for each SEG annotation. Results are meta-analyzed across 30 diseases and complex traits. Error bars denote ±1.96× standard error. Annotations are color-coded as in (a). The dashed black line (slope=1.01) denotes a regression of observed $\lambda(C) - 1$ vs. expected $\lambda(C) - 1$ with intercept constrained to 0. Numerical results and population-specific heritability enrichment estimates are reported in Table S16.

References

[1] Teresa R de Candia et al. "Additive genetic variation in schizophrenia risk is shared by populations of African and European descent". In: *The American Journal of Human Genetics* 93.3 (2013), pp. 463–470.

[2] Brielin C Brown et al. "Transethnic genetic-correlation estimates from summary statistics". In: *The American Journal of Human Genetics* 99.1 (2016), pp. 76–88.

[3] Nicholas Mancuso et al. "The contribution of rare variation to prostate cancer heritability". In: *Nature genetics* 48.1 (2016), p. 30.

[4] Masashi Ikeda et al. "Genome-Wide Association Study Detected Novel Susceptibility Genes for Schizophrenia and Shared Trans-Populations/Diseases Genetic Effect". In: *Schizophrenia bulletin* 45.4 (2018), pp. 824–834.

[5] Kevin J Galinsky et al. "Estimating cross-population genetic correlations of causal effect sizes". In: *Genetic epidemiology* 43.2 (2019), pp. 180–188.

[6] Alicia R Martin et al. "Clinical use of current polygenic risk scores may exacerbate health disparities". In: *Nature genetics* 51.4 (2019), p. 584.

[7] Christopher S Carlson et al. "Generalization and dilution of association results from European GWAS in populations of non-European ancestry: the PAGE study". In: *PLoS biology* 11.9 (2013), e1001661.

[8] Alicia R Martin et al. "Human demographic history impacts genetic risk prediction across diverse populations". In: *The American Journal of Human Genetics* 100.4 (2017), pp. 635–649.

[9] Carla Márquez-Luna et al. "Multiethnic polygenic risk scores improve risk prediction in diverse populations". In: *Genetic epidemiology* 41.8 (2017), pp. 811–823.

[10] Genevieve L Wojcik et al. "Genetic analyses of diverse populations improves discovery for complex traits". In: *Nature* (2019).

[11] L Duncan et al. "Analysis of polygenic risk score usage and performance in diverse human populations". In: *Nature Communications* 10.1 (2019), p. 3328.

[12] Kevin L Keys et al. "On the cross-population portability of gene expression prediction models". In: *bioRxiv* (2019), p. 552042.

[13] Sang Hong Lee et al. "Estimation of pleiotropy between complex diseases using single-nucleotide polymorphism-derived genomic relationships and restricted maximum likelihood". In: *Bioinformatics* 28.19 (2012), pp. 2540–2542.

[14] Giorgio Sirugo, Scott M Williams, and Sarah A Tishkoff. "The missing diversity in human genetic studies". In: *Cell* 177.1 (2019), pp. 26–31.

[15] Deepti Gurdasani et al. “Genomics of disease risk in globally diverse populations”. In: *Nature Reviews Genetics* (2019).

[16] 1000 Genomes Project Consortium et al. “A global reference for human genetic variation”. In: *Nature* 526.7571 (2015), p. 68.

[17] Na Cai et al. “Sparse whole-genome sequencing identifies two loci for major depressive disorder”. In: *Nature* 523.7562 (2015), p. 588.

[18] Akiko Nagai et al. “Overview of the BioBank Japan Project: study design and profile”. In: *Journal of epidemiology* 27.Supplement_III (2017), S2–S8.

[19] Masahiro Kanai et al. “Genetic analysis of quantitative traits in the Japanese population links cell types to complex human diseases”. In: *Nature genetics* 50.3 (2018), p. 390.

[20] Hilary K Finucane et al. “Partitioning heritability by functional annotation using genome-wide association summary statistics”. In: *Nature genetics* 47.11 (2015), p. 1228.

[21] Steven Gazal et al. “Linkage disequilibrium–dependent architecture of human complex traits shows action of negative selection”. In: *Nature genetics* 49.10 (2017), p. 1421.

[22] Steven Gazal et al. “Reconciling S-LDSC and LDAK functional enrichment estimates”. In: *Nature genetics* 51.8 (2019), pp. 1202–1204.

[23] Hilary K Finucane et al. “Heritability enrichment of specifically expressed genes identifies disease-relevant tissues and cell types”. In: *Nature genetics* 50.4 (2018), p. 621.

[24] James Durbin. “A note on the application of Quenouille’s method of bias reduction to the estimation of ratios”. In: *Biometrika* 46.3/4 (1959), pp. 477–480.

[25] Zhan Su, Jonathan Marchini, and Peter Donnelly. “HAPGEN2: simulation of multiple disease SNPs”. In: *Bioinformatics* 27.16 (2011), pp. 2304–2305.

[26] Na Cai, Kenneth Kendler, and Jonathan Flint. “Minimal phenotyping yields GWAS hits of low specificity for major depression”. In: *BioRxiv* (2018), p. 440735.

[27] Graham McVicker et al. “Widespread genomic signatures of natural selection in hominid evolution”. In: *PLoS genetics* 5.5 (2009), e1000471.

[28] Gleb Kichaev and Bogdan Pasaniuc. “Leveraging functional-annotation data in trans-ethnic fine-mapping studies”. In: *The American Journal of Human Genetics* 97.2 (2015), pp. 260–271.

[29] GTEx Consortium et al. “Genetic effects on gene expression across human tissues”. In: *Nature* 550.7675 (2017), p. 204.

[30] Soumya Raychaudhuri et al. “Accurately assessing the risk of schizophrenia conferred by rare copy-number variation affecting genes with brain function”. In: *PLoS genetics* 6.9 (2010), e1001097.

[31] Pardis C Sabeti et al. "Positive natural selection in the human lineage". In: *science* 312.5780 (2006), pp. 1614–1620.

[32] Rasmus Nielsen et al. "Recent and ongoing selection in the human genome". In: *Nature Reviews Genetics* 8.11 (2007), p. 857.

[33] John Novembre and Anna Di Rienzo. "Spatial patterns of variation due to natural selection in humans". In: *Nature Reviews Genetics* 10.11 (2009), p. 745.

[34] Kevin N Laland, John Odling-Smee, and Sean Myles. "How culture shaped the human genome: bringing genetics and the human sciences together". In: *Nature Reviews Genetics* 11.2 (2010), p. 137.

[35] Sandra Wilde et al. "Direct evidence for positive selection of skin, hair, and eye pigmentation in Europeans during the last 5,000 y". In: *Proceedings of the National Academy of Sciences* 111.13 (2014), pp. 4832–4837.

[36] Harald von Boehmer. "Positive selection of lymphocytes". In: *Cell* 76.2 (1994), pp. 219–228.

[37] Jian Zeng et al. "Signatures of negative selection in the genetic architecture of human complex traits". In: *Nature genetics* 50.5 (2018), p. 746.

[38] Luke J O'Connor et al. "Extreme Polygenicity of Complex Traits Is Explained by Negative Selection". In: *The American Journal of Human Genetics* (2019).

[39] Armin P Schoech et al. "Quantification of frequency-dependent genetic architectures in 25 UK Biobank traits reveals action of negative selection". In: *Nature communications* 10.1 (2019), p. 790.

[40] Krishna R Veeramah and Michael F Hammer. "The impact of whole-genome sequencing on the reconstruction of human population history". In: *Nature Reviews Genetics* 15.3 (2014), p. 149.

[41] Matthew R Robinson et al. "Genotype–covariate interaction effects and the heritability of adult body mass index". In: *Nature genetics* 49.8 (2017), p. 1174.

[42] William G Hill, Michael E Goddard, and Peter M Visscher. "Data and theory point to mainly additive genetic variance for complex traits". In: *PLoS genetics* 4.2 (2008), e1000008.

[43] Asko Mäki-Tanila and William G Hill. "Influence of gene interaction on complex trait variation with multilocus models". In: *Genetics* 198.1 (2014), pp. 355–367.

[44] Zhihong Zhu et al. "Dominance genetic variation contributes little to the missing heritability for human complex traits". In: *The American Journal of Human Genetics* 96.3 (2015), pp. 377–385.

[45] Maaike de Jong et al. "Natural variation in Arabidopsis shoot branching plasticity in response to nitrate supply affects fitness". In: *PLoS genetics* 15.9 (2019), e1008366.

[46] Monkol Lek et al. "Analysis of protein-coding genetic variation in 60,706 humans". In: *Nature* 536.7616 (2016), p. 285.

[47] Adam Eyre-Walker. "Genetic architecture of a complex trait and its implications for fitness and genome-wide association studies". In: *Proceedings of the National Academy of Sciences* (2010), p. 200906182.

[48] Reedik Mägi et al. "Trans-ethnic meta-regression of genome-wide association studies accounting for ancestry increases power for discovery and improves fine-mapping resolution". In: *Human molecular genetics* 26.18 (2017), pp. 3639–3650.

[49] Andrew P Morris. "Transethnic meta-analysis of genomewide association studies". In: *Genetic epidemiology* 35.8 (2011), pp. 809–822.

[50] Patrick Turley et al. "Multi-trait analysis of genome-wide association summary statistics using MTAG". In: *Nature genetics* 50.2 (2018), p. 229.

[51] Brendan Bulik-Sullivan et al. "An atlas of genetic correlations across human diseases and traits". In: *Nature genetics* 47.11 (2015), p. 1236.

[52] Qiongshi Lu et al. "A powerful approach to estimating annotation-stratified genetic covariance via GWAS summary statistics". In: *The American Journal of Human Genetics* 101.6 (2017), pp. 939–964.

[53] Michael F Seldin, Bogdan Pasaniuc, and Alkes L Price. "New approaches to disease mapping in admixed populations". In: *Nature Reviews Genetics* 12.8 (2011), p. 523.

[54] Brendan K Bulik-Sullivan et al. "LD Score regression distinguishes confounding from polygenicity in genome-wide association studies". In: *Nature genetics* 47.3 (2015), p. 291.

[55] Yang Luo et al. "Estimating heritability and its enrichment in tissue-specific gene sets in admixed populations". In: *bioRxiv* (2019), p. 503144.

[56] Alicia R Martin et al. "Transcriptome sequencing from diverse human populations reveals differentiated regulatory architecture". In: *PLoS genetics* 10.8 (2014), e1004549.

[57] Lauren S Mogil et al. "Genetic architecture of gene expression traits across diverse populations". In: *PLoS genetics* 14.8 (2018), e1007586.

[58] S Gazal et al. "Functional architecture of low-frequency variants highlights strength of negative selection across coding and non-coding annotations." In: *Nature genetics* 50.11 (2018), p. 1600.

[59] Arun Durvasula and Kirk E Lohmueller. "Negative selection on complex traits limits genetic risk prediction accuracy between populations". In: *bioRxiv* (2019), p. 721936.

[60] International HapMap 3 Consortium et al. "Integrating common and rare genetic variation in diverse human populations". In: *Nature* 467.7311 (2010), p. 52.

[61] ENCODE Project Consortium et al. “An integrated encyclopedia of DNA elements in the human genome”. In: *Nature* 489.7414 (2012), p. 57.

[62] Anshul Kundaje et al. “Integrative analysis of 111 reference human epigenomes”. In: *Nature* 518.7539 (2015), p. 317.

[63] Maya Kasowski et al. “Extensive variation in chromatin states across humans”. In: *Science* 342.6159 (2013), pp. 750–752.

[64] Eugene V Davydov et al. “Identifying a high fraction of the human genome to be under selective constraint using GERP++”. In: *PLoS computational biology* 6.12 (2010), e1001025.

[65] Simon Myers et al. “A fine-scale map of recombination rates and hotspots across the human genome”. In: *Science* 310.5746 (2005), pp. 321–324.

[66] Matthew D Rasmussen et al. “Genome-wide inference of ancestral recombination graphs”. In: *PLoS genetics* 10.5 (2014), e1004342.

[67] Christopher C Chang et al. “Second-generation PLINK: rising to the challenge of larger and richer datasets”. In: *Gigascience* 4.1 (2015), p. 7.

[68] Bruce S Weir and C Clark Cockerham. “Estimating F-statistics for the analysis of population structure”. In: *evolution* 38.6 (1984), pp. 1358–1370.

[69] Clare Bycroft et al. “The UK Biobank resource with deep phenotyping and genomic data”. In: *Nature* 562.7726 (2018), p. 203.

[70] Siew-Kee Low et al. “Identification of six new genetic loci associated with atrial fibrillation in the Japanese population”. In: *Nature genetics* 49.6 (2017), p. 953.

[71] Jonas B Nielsen et al. “Biobank-driven genomic discovery yields new insight into atrial fibrillation biology”. In: *Nature genetics* 50.9 (2018), p. 1234.

[72] Momoko Horikoshi et al. “Elucidating the genetic architecture of reproductive ageing in the Japanese population”. In: *Nature communications* 9.1 (2018), p. 1977.

[73] Felix R Day et al. “Large-scale genomic analyses link reproductive aging to hypothalamic signaling, breast cancer susceptibility and BRCA1-mediated DNA repair”. In: *Nature genetics* 47.11 (2015), p. 1294.

[74] William J Astle et al. “The allelic landscape of human blood cell trait variation and links to common complex disease”. In: *Cell* 167.5 (2016), pp. 1415–1429.

[75] Po-Ru Loh et al. “Mixed-model association for biobank-scale datasets”. In: *Nature genetics* 50.7 (2018), p. 906.

[76] Cristian Pattaro et al. “Genetic associations at 53 loci highlight cell types and biological pathways relevant for kidney function”. In: *Nature communications* 7 (2016), p. 10023.

[77] Masato Akiyama et al. “Characterizing rare and low-frequency height-associated variants in the Japanese population”. In: *Nature Communications* 10.1 (2019), pp. 1–11.

[78] Naomi R Wray et al. "Genome-wide association analyses identify 44 risk variants and refine the genetic architecture of major depression". In: *Nature genetics* 50.5 (2018), p. 668.

[79] Yukinori Okada et al. "Genetics of rheumatoid arthritis contributes to biology and drug discovery". In: *Nature* 506.7488 (2014), p. 376.

[80] Ken Suzuki et al. "Identification of 28 new susceptibility loci for type 2 diabetes in the Japanese population". In: *Nature genetics* 51.3 (2019), p. 379.

[81] Robert A Scott et al. "An expanded genome-wide association study of type 2 diabetes in Europeans". In: *Diabetes* 66.11 (2017), pp. 2888–2902.

[82] Donna Karolchik, Angie S Hinrichs, and W James Kent. "The UCSC genome browser". In: *Current protocols in bioinformatics* 40.1 (2012), pp. 1–4.