1   **Supplementary Note for**
2   *Modeling tissue co-regulation estimates tissue-specific contributions to disease*
3
4   **Supplementary Table Captions**
5
6   **Supplementary Table 1. Numerical results for robustness and power of TCSC regression in**
7   **simulations.** Across six different eQTL sample sizes, we evaluate the causal and null bias in
8   estimates of disease heritability explained by the *cis*-genetic component of gene expression in
9   tissue t' ($h^2_{ge(t')}$), the type I error, and the power of TCSC. The standard errors (SE) are
10   computed as the standard deviation of measurements across simulations divided by the square
11   root of the number of simulations, e.g. 1,000. Type I error is measured as the percentage of
12   estimates of $h^2_{ge(t')}$ for non-causal tissues that were significantly positive for non-causal tissues
13   at $p < 0.05$ for nominal significance or at 5% FDR across tissues using a one-sided z-test. Power
14   is measured as the percentage of estimates of $h^2_{ge(t')}$ for causal tissues that were significantly
15   positive at $p < 0.05$ for nominal significance or at 5% FDR across tissues using a one-sided z-test.
16
17   **Supplementary Table 2. Type I error and power of RTC Coloc, LDSC-SEG, RolyPoly, and**
18   **CoCoNet in simulations.** We implemented all methods as previously described and applied it to
19   our TCSC simulation framework, such that the same eQTL effect sizes and co-regulation was
20   used. We performed 1,000 simulations of LDSC-SEG, RolyPoly, and CoCoNet and 100
21   simulations of RTC Coloc, due to the complexity and prohibitively large computation time of
22   RTC Coloc.
23
24   **Supplementary Table 3. Numerical results for robustness and power of cross-trait TCSC in**
25   **simulations.** Across six different eQTL sample sizes, we evaluate the causal and null bias on the
26   estimate of tissue-specific contributions to covariance, the type I error, and the power of cross-
27   trait TCSC. The standard errors (SE) are computed as the standard deviation of measurements
28   across simulations divided by the square root of the number of simulations, e.g. 1,000. Type I
29   error is measured as the percentage of estimates of $\omega_{ge(t')}$ for non-causal tissues that were
30   significantly positive at $p < 0.05$ for nominal significance or at 5% FDR across tissues using a
31   one-sided z-test. Power is measured as the percentage of estimates of $\omega_{ge(t')}$ for causal tissues
32   that were significantly positive at $p < 0.05$ for nominal significance or at 5% FDR across tissues
33   using a one-sided z-test.
34
35   **Supplementary Table 4. List of 78 diseases and complex traits analyzed in primary analyses.**
36   We selected 78 diseases/traits, 33 of which are from UK Biobank, such that all summary
37   statistics have SNP-heritability z-score > 6 and no pair of traits have a squared genetic
38   correlation greater than 0.1 as well as substantial sample overlap. We report the SNP-
39   heritability, standard error, z-score, GWAS sample size, trait nickname used to index traits in
40   plots and tables, and the name of the most closely related trait analyzed in previous studies[1,2].
41
42   **Supplementary Table 5. Numerical results for tissue-specific contributions to disease and**
43   **complex trait heritability.** For each significant tissue-trait pair identified by TCSC as reported in

44 **Fig. 3**, we report the value of $\pi_{t'}$, or the proportion of SNP-heritability explained by tissue-
45 specific predicted gene expression, and the false discovery rate for each finding. We also report
46 the result of an independent analysis for traits for which we could easily obtain independent
47 GWAS summary statistics.

48

49 **Supplementary Table 6. Numerical results for tissue-specific contributions to disease and**
50 **complex trait heritability for all tissues and diseases/traits analyzed.** For every tissue-trait pair
51 analyzed by TCSC, across 39 tissues and 78 diseases/traits, we report the phenotypic variance
52 explained by tissue-specific predicted gene expression, $h^2_{ge(t')}$, the jackknife standard error of
53 this quantity, the nominal P value from a one-sided z-test, the FDR calculated across tissues
54 per-trait, the value of $\pi_{t'}$, and the standard error of $\pi_{t'}$. We estimated the standard error of
55 this quantity using a genomic block jackknife. We note that no value of $h^2_{ge(all\ tissues)}$ exceeds
56 the SNP-heritability for a given trait. The largest value of $h^2_{ge(all\ tissues)}$ is 0.68, which is for red
57 blood cell distribution width.

58

59 **Supplementary Table 7. Median jackknife P values across traits for each tissue.** Here we
60 report the median jackknife P value across traits for each tissue. For pairs of tissues with high
61 genetic correlation, if the median jackknife P value is substantially different across traits, this
62 means TCSC is systematically more likely to identify the tissue with lower median jackknife P
63 value as a causal tissue relative to the other, and this might suggest an issue in quality of gene
64 expression prediction models from the tissue with larger median jackknife P value.

65

66 **Supplementary Table 8. Numerical results for tissue-specific contributions to disease and**
67 **complex trait heritability in secondary analysis of 23 tissues, removing tissues with small**
68 **eQTL sample size.** To increase the power of TCSC to identify causal tissues, we removed tissues
69 with eQTL sample size less than 320. As a result, we analyzed 23 tissues across 78
70 diseases/traits. We report the same quantities reported in **Supplementary Table 6**.

71

72 **Supplementary Table 9. Statistical significance of differences between TCSC estimates in**
73 **primary and secondary analyses.** For every significant tissue-trait pair identified in the primary
74 analysis (analysis of 39 tissues, **Fig. 4**), we assessed if the value of $\pi_{t'}$ was significantly different
75 than the value produced in the secondary analysis (analysis of 23 tissues, **Supplementary Table**
76 **8**). We used a genomic block jackknife to assess the difference and using a two-sided test,
77 identified that no differences were significantly nonzero at 10% FDR.

78

79 **Supplementary Table 10. List of 41 brain diseases/traits analyzed in brain-specific analysis.**
80 We performed a brain-specific TCSC analysis to exploit the diversity of brain tissues provided by
81 GTEx (n = 13 brain tissues). We analyzed 41 brain diseases/traits using a similar trait selection
82 procedure as was used to select the 78 diseases/traits previously analyzed. However, we first
83 selected for diseases/traits that were behavioral or a known cognitive disorder and iteratively
84 removed traits until no pair of traits had a squared genetic correlation greater than 0.25.

85

86   **Supplementary Table 11. List of 13 GTEx brain tissues analyzed in brain-specific analysis.** For
87   the brain-specific TCSC analysis, we built gene expression prediction models using all European
88   samples from each of 13 GTEx brain tissues, without subsampling or meta-analysis. Here we list
89   the name and eQTL sample size of each GTEx brain tissue, e.g. tissue names beginning with
90   "Brain_". For this analysis, we excluded several tissues relevant to the central nervous system,
91   including pituitary (N = 220) and tibial nerve (N > 320).

92

93   **Supplementary Table 12. Numerical results for tissue-specific contributions to disease and**
94   **complex trait heritability in brain-specific analysis.** For every brain tissue and brain trait
95   analyzed in the brain-specific TCSC analysis, we report $\pi_{t'}$, its standard error, and false
96   discovery rate.

97

98   **Supplementary Table 13. Numerical results for comparison of disease-critical tissues**
99   **identified by RTC Coloc, LDSC-SEG and TCSC for 5 representative traits.** For every tissue-trait
100 pair shown in **Fig. 4**, we report the FDR and -$\log_{10}$FDR of the association statistic for each
101 method (enrichment statistic for Ongen 2017 RTC Coloc, tau* S-LDSC statistic for Finucane 2018
102 LDSC-SEG, and $\pi_{t'}$ for TCSC). The seven traits are the ones having at least one significantly
103 associated tissue across the three methods with the largest SNP-heritability z-score. The tissues
104 reported here are the causal tissues for each of the five traits as well as the most genetically
105 correlated tissue (using marginal eQTL effect sizes).

106

107 **Supplementary Table 14. Numerical results for comparison of disease-critical tissues**
108 **identified by RTC Coloc, LDSC-SEG and TCSC for all 21 diseases/traits with causal tissue-trait**
109 **associations identified by TCSC.** We report the FDR and -$\log_{10}$FDR of the association statistic for
110 each method across all traits shown in **Fig. 3** and each tissue with an association statistic with
111 FDR < 5%.

112

113 **Supplementary Table 15. Numerical results for comparison of disease-critical tissues**
114 **identified by RTC Coloc, LDSC-SEG and TCSC for all diseases/traits and tissues included in**
115 **these comparisons.** We report the FDR and -$\log_{10}$FDR for every tissue-trait pair (39 tissues, 78
116 traits) across each of three compared methods. A value of NA indicates that the tissue-trait pair
117 was not analyzed by the corresponding method.

118

119 **Supplementary Table 16. Numerical results for comparison of disease-critical tissues**
120 **identified by RTC Coloc, LDSC-SEG and TCSC for brain-specific analysis.** We report the FDR and
121 -$\log_{10}$FDR for the brain-specific analysis for each of 41 brain traits, 13 brain tissues, and 3
122 methods, restricting to tissues and traits with a TCSC finding at FDR < 10%, plotted in **Extended**
123 **Data Fig. 8**. A value of NA indicates that the tissue-trait pair was not analyzed by the
124 corresponding method.

125

126 **Supplementary Table 17. List of 262 pairs of diseases/traits analyzed by cross-trait TCSC.** We
127 computed the genetic correlation between all pairs of 78 diseases/traits and selected those
128 pairs with genetic correlation two-sided z-test P value < 0.05/3,003 pairs of traits, e.g. using a

129   Bonferroni correction threshold. Here, we report the genetic correlation *z*-score of these pairs
130   and the estimate of the covariance.

131

132   **Supplementary Table 18. Numerical results for tissue-specific contributions to the genetic**
133   **covariance of two diseases/traits (Figure 6A).** For all tissue-trait covariance pairs identified by
134   TCSC at 10% FDR, we report the value of $\zeta_{t'}$, or the proportion of covariance explained by
135   predicted gene expression in tissue t' and the FDR.

136

137   **Supplementary Table 19. Numerical results for tissue-specific contributions to the genetic**
138   **covariance of two diseases/traits for all tissues and disease/trait pairs analyzed.** For all tissue-
139   trait covariance pairs analyzed by TCSC, we report the estimated tissue-specific covariance, its
140   jackknife standard error, nominal *P* value from a two-sided z-test, $\zeta_{t'}$ and corresponding
141   standard error, FDR, and genome-wide covariance for the trait pair.

142

143   **Supplementary Table 20. List of tissue-trait covariance pairs and reported differences in**
144   **tissue-specific contributions to genetic covariance vs. constituent trait heritability.** For every
145   pair of traits implicated by **Figure 6** and for each of 38 tissues, we assess the difference
146   between $\zeta_{t'}$ and $\pi_{t'}$ for each trait. We identified five tissue-trait covariance pairs for which the
147   difference was significantly nonzero while the value of $\pi_{t'}$ was not significantly different than
148   zero at a significance threshold of 5% FDR across tissues per-trait.

149

150   **Supplementary Table 21. Scenarios where TCSC has more power in the cross-trait analysis**
151   **than in the single-trait analysis.** We used primary simulations and performed new simulations
152   in which tissue-specific contributions to covariance where greater than tissue-specific
153   contributions to heritability in order to report the percentage of simulations in which the causal
154   tissue was detected in the cross-trait analysis but not detected in both of the single-trait
155   analyses.

156
157
158
159
160
161
162
163
164

Supplementary Note

Simulation Framework

We employed a widely used TWAS simulation framework (Mancuso Lab TWAS Simulator, see **Code Availability**) to assess the power, bias, and calibration of TCSC in the presence of co-regulation across genes and tissues. We simulated a genome in which there are 1,000 protein-coding genes from chromosome 1, of which 100 (10%) are causal[3]. For each tissue, 500 genes were chosen to be *cis*-heritable; in the causal tissue and the three most highly genetically correlated tagging tissues, all 100 causal genes were *cis*-heritable. Each primary simulation consists of 10 tissues, of which at least one is causal, defined as having nonzero gene- disease effect sizes. We create a covariance structure among tissues mimicking empirical GTEx data. We use a previously published method to estimate the causal cross-tissue correlation of eQTL effect sizes which is 0.75[4]. Briefly, this method extends cross-trait LD score regression and leverages *cis*-eQTL summary statistics across all expressed genes in a tissue to compute cross-tissue genetic correlations. We observe that not all GTEx tissues are equally correlated to one another. We estimate three different cross-tissue eQTL correlation quantities: (1) average correlation across all pairs of tissues = 0.75, (2) average correlation across similar tissues = 0.80, e.g. brain (13 in GTEx) or adipose (2 in GTEx) tissues, and (3) average correlation across dissimilar tissues, e.g. pairs of brain and adipose tissues = 0.74. To represent these biological modules, we let simulated tissues 1-3 have higher correlation of true eQTL effects to one another than to other tissues; likewise for tissues 4-6 and 7-10. We set covariance parameters, described below, such that the similar tissues had an average eQTL correlation of 0.795 across genes, dissimilar tissues have an average eQTL correlation of 0.722, and the average eQTL correlation across any pair of tissues is 0.753. We use real genotypes from European individuals in the 1000 Genomes Project to define the pairwise SNP LD structure which is used to simulate genotypes, gene expression traits, and complex traits/diseases. We simulate each gene having 5 true *cis*-eQTLs, based on the upper bound of empirical data from GTEx[5] and others[6], as well as the value used in other TWAS simulation methods[7]. Between pairs of co-regulated tissues, the same gene shares 3 *cis*-eQTLs. Between pairs of co-regulated genes in the same tissue, 3 *cis*-eQTLs are shared. The minimum allowed *cis*-heritability of a gene is 0.01 in our simulations. *Cis*-heritability is approximated as the sum of squared true *cis*-eQTL effect sizes, as done previously[8]. The *cis*-heritability of each gene was sampled from an exponential distribution, and neighboring co-regulated genes were assigned the same heritability to maximize gene-gene co-regulation. In each tissue, the average *cis*-heritability (across genes) was set to 0.08 (sd = 0.05, ranging from 0.01 to 0.40) in order to achieve an average estimated *cis*-heritability (*across significantly cis-heritable genes, estimated by GCTA[9]*) varying from 0.11 to 0.31 (across gene expression sample sizes), which matches empirical values from GTEx[5]. Effect sizes for the 3 shared eQTLs across tissues are sampled from a multivariate normal distribution with mean 0 and a variance-covariance matrix. We define the variance and covariance terms of this matrix such that (1) the proportion of genes detected as significantly *cis*-heritable by GCTA at a given sample size and (2) the average *cis* heritability of detected genes at a given sample size match empirical observations from GTEx data at sample sizes N = 100, 200, 300 and 500. As a result, the diagonal of the variance-covariance matrix, e.g.

209   the variance term, is set to 0.075, and the off-diagonal elements are set to the product of the
210   variance term and the desired correlation for each tissue pair, described above.
211   For each of 1,000 independent simulations per analysis, we simulate a GWAS (N =
212   10,000) by creating a complex trait which is the summation of the genetic components of
213   causal gene expression (in the causal tissue). We use simulated genotypes based on the LD
214   structure of 1000 Genomes. Gene-disease effect sizes are drawn from a normal distribution
215   with mean 0 and variance 1. In cross-trait TCSC analysis, effect sizes across genes between the
216   two traits are correlated with default R_g = 0.5. To simulate a GWAS trait, we first compute the
217   genetic component of each gene, which is the product of GWAS cohort genotypes and eQTL
218   effects, such that we have 100 gene-specific traits. We then add noise to each gene-specific
219   trait such that the total variance of the phenotype explained by the five eQTLs from the causal
220   tissue is equal to a specified value; the value of $h^2_{ge(t)}$ in primary simulations is 10%. Then, we
221   multiply each gene-specific trait by the causal gene-disease effect size, consistent with the
222   additive generative model of gene-level effects on trait (see above). Finally, we take the sum
223   across all gene-specific traits to make one complex trait, where the total variance of the trait
224   explained by gene effects from the causal tissue is $h^2_{ge(t)}$, e.g. 10%.
225   We simulate an eQTL cohort of various gene expression sample sizes (N = 100, 200, 300,
226   500, 1000, 1500) using simulated genotypes based on the LD structure of 1000 Genomes. We
227   simulate total gene expression in the eQTL cohort by adding a desired amount of noise to the
228   genetic component of gene expression, e.g. the product of individual genotypes and true eQTL
229   effect sizes, with variance equal to one minus the gene expression heritability, which is the sum
230   of squared eQTL effects. Next, we fit gene expression prediction models by regressing the total
231   gene expression on eQTL cohort genotypes of *cis* variants using lasso regularization, a standard
232   approach used in TWAS. We define significantly *cis*-heritable genes as genes with GCTA
233   heritability *P* value < 0.01[10] and heritability estimate > 0, and adjusted-R^2 > 0 in cross-validation
234   prediction.
235   Then we estimate co-regulation scores at each different eQTL sample size by predicting
236   gene expression into a cohort of 500 individuals, to approximate the size of the European
237   sample of 1000 Genomes (*N* = 489). Using significantly *cis*-heritable genes from each tissue at a
238   given sample size, we estimate gene and tissue co-regulation scores $l(g, t; t')$ as described
239   above, including bias correction. In simulations, *cis* genes are defined as genes within the same
240   1 Mb block.
241   Then we apply TWAS to individual-level simulated GWAS data and gene expression
242   prediction models. We predict gene expression into each of the 10,000 GWAS cohort individuals
243   across all significantly *cis*-heritable genes for each tissue. We regress each complex trait on
244   predicted gene expression to obtain TWAS z-scores. Finally, we run TCSC by regressing TWAS χ^2
245   statistics, or products of TWAS z-scores, on bias-corrected gene and tissue co-regulation scores.
246
247   We simulated four tissue-trait association methods: RTC Coloc[1], LDSC-SEG[2], RolyPoly[11],
248   or CoCoNet[12]. First, we simulated the RTC Coloc method[1] by leveraging our existing TCSC
249   simulation framework such that both methods could be compared via application to same
250   simulated data (**Code Availability,** ref.[13]). We used the same simulated GWAS cohort of 10,000
251   individuals as in our TCSC simulations and then followed the steps of the RTC Coloc method as

252   published. Briefly, we perform a genome-wide association study using our simulated complex
253   trait and the genotypes of our simulated GWAS cohort and select null variants with similar LD
254   properties. Then, we simulate an eQTL cohort consisting of total gene expression and
255   genotypes, using the same underlying true eQTL effect sizes as for TCSC simulations. Then, we
256   perform colocalization analysis of GWAS variants with eQTLs, across 10 tissues at 6 different
257   eQTL sample sizes, to obtain the regulatory trait concordance (RTC) score. This is repeated for
258   the set of null variants. Next, we perform colocalization analysis of eQTL variants between pairs
259   of tissues to obtain tissue-sharing RTC scores, and similarly repeat this for null variants. GWAS-
260   eQTL RTC scores are divided by tissue-sharing RTC scores summed across variants. Tissue-
261   specific enrichment is computed as the ratio of this quotient to the null quotient. The
262   enrichment P value is obtained using a Wilcox test comparing the values of the quotient to the
263   values of the null quotient.
264   Second, we simulated the three methods that utilize GWAS data and total expression
265   across tissues: LDSC-SEG[2] (using S-LDSC v1.0.0), RolyPoly[11] (v0.1.0), and CoCoNet[12] (v1.0). To
266   this end, we retained the full GWAS summary statistics from the RTC Coloc analysis above. We
267   separately simulated total expression across tissues in which the 100 causal genes in addition to
268   200 randomly selected genes were positively differentially expressed in the causal tissue and
269   the two tagging tissues in the same simulated "module" as the causal tissue, e.g. with higher
270   genetic correlation of gene regulatory effects. We also selected 100 random non-causal genes
271   to be negatively differentially expressed in the causal tissue and the other two module tissues.
272   For the remaining 7 tagging tissues, we randomly selected 300 genes to be positively
273   differentially expressed, some of which at random will be causal genes, and let the remaining
274   700 genes be negatively differentially expressed. Then, as previously done[2], we calculated the t-
275   statistics for the specific expression of each gene in each tissue. While we have modules of
276   tissues that are more highly correlated to one another, these within-module tissues were
277   excluded from the calculation of t-statistics, as previously done[2]. Finally, we created SNP-based
278   annotations for each tissue, across 1000 simulations, and across 6 sample sizes, in which SNPs
279   within +/- 100 kb of a specifically expressed gene is assigned a value of 1 and 0 otherwise, as
280   previously done[2]. Then, we calculated LD scores and partitioned the heritability of our
281   simulated complex traits. For the simulations of RolyPoly and CoCoNet, we installed the
282   following R packages: rolypoly (v0.1.0) and CoCoNet (v1.0) and used the simulated data above
283   to run each method. While CoCoNet does not technically use GWAS summary statistics, but
284   rather gene-based "outcome variables", we used the label of causal or non-causal for each gene
285   in each tissue of our simulations as the outcome variable.
286
287
288   ***Single-trait simulation analysis at large sample size***
289
290   We simulated four larger gene expression sample sizes: 10K, 50K, 100K, and infinite sample size
291   (for infinite sample size, we used the true eQTL effect sizes in place of estimated effect sizes
292   from the gene expression prediction model). Due to the computational intractability of running
293   GCTA at these sample sizes across thousands of simulations, we used cross-validation adjusted-
294   $R^2 > 0$ in lieu of GCTA to define significantly *cis*-heritable genes in analyses at very large sample
295   sizes. We determined that the type I error of TCSC plateaus at 1,500 individuals (**Fig. 2C**); we

296 also confirmed that the alternative definition of *cis*-heritable genes did not impact results at
297 intermediate sample sizes (**Fig. 2C** vs. **Fig. 2B**).
298
299 ***Cross-trait simulation analyses***
300
301 We first evaluated the bias in TCSC estimates of the genetic covariance explained by the *cis*-
302 genetic component of gene expression in tissue t' ($\omega_{ge(t')}$), for both causal and non-causal
303 tissues (**Extended Data Fig. 2A**, **Supplementary Table 3**). For causal tissues, TCSC produced
304 unbiased estimates of $\omega_{ge(t')}$ (conservative estimates when setting $G_{t'}$ to the number of
305 significantly *cis*-heritable genes, rather than the number of true *cis*-heritable genes), analogous
306 to single-trait simulations. For non-causal tissues, TCSC again produced estimates of $\omega_{ge(t')}$ that
307 were significantly positive when averaged across all simulations, but not large enough to
308 substantially impact type I error. We next evaluated the type I error of cross-trait TCSC for non-
309 causal tissues. TCSC was well-calibrated with type I error ranging from 5.4%-6.7% at $p < 0.05$
310 (**Extended Data Fig. 2B**). Finally, we evaluated the power of cross-trait TCSC for causal tissues.
311 We determined that cross-trait TCSC was modestly powered at realistic eQTL sample sizes, with
312 power ranging from 8%-27% across eQTL sample sizes at $p < 0.05$ (**Extended Data Fig. 2C**) (and
313 1-6% power at $p < 0.004$ corresponding to 5% per-trait FDR across tissues in these simulations;
314 **Supplementary Table 3**); as noted above, the power of TCSC varies greatly with the choice of
315 parameter settings (see below). In ROC curve analysis, TCSC attained an AUC of 0.67 (**Extended
316 Data Fig. 1**).
317
318 ***Secondary simulation analyses***
319
320 We performed twelve secondary analyses. First, we varied the eQTL sample size across tissues.
321 Specifically, we set the eQTL sample size of the causal tissue to 300 individuals and the eQTL
322 sample sizes of the non-causal tissues to range between 100 and 1,500 individuals. We
323 observed inflated type I error for non-causal tissues (particularly those with larger eQTL sample
324 sizes), implying that large variations in eQTL sample sizes may compromise type I error
325 (**Supplementary Fig. 1**).
326   Second, we evaluated the robustness of TCSC when varying the number of expressed
327 genes in the causal tissue under four scenarios: (i) only the 500 *cis*-heritable genes are
328 expressed in the causal tissue, (ii) only 375 *cis*-heritable genes (including all 100 causal genes)
329 are expressed in the causal tissue, (iii) only 225 *cis*-heritable genes (including all 100 causal
330 genes) are expressed in the causal tissue, and (iv) only the 100 causal genes are expressed in
331 the causal tissue. We determined that type I error remained approximately well-calibrated in all
332 scenarios, and that power was dramatically improved and bias for non-causal tissues decreased
333 as the number of tagging genes in the causal tissue decreased (**Supplementary Fig. 2-3**); for
334 causal tissues, estimates of $h^2_{ge(t')}$ were upward biased when setting $G_{t'}$ to the number of true
335 *cis*-heritable genes and unbiased when setting $G_{t'}$ to the number of significantly *cis*-heritable
336 genes across tissues.
337   Third, we varied the true values of $h^2_{ge(t')}$ (or $\omega_{ge(t')}$) for causal tissues. We determined
338 that patterns of bias, type I error, and power were generally robust across different parameter

339   values, although the smallest values resulted in lower power and greater bias for non-causal
340   tissues (**Extended Data Figs. 3-4**). Specifically, in **Extended Data Fig. 3**, we varied the value of
341    $h^2_{ge(t\prime)}$ for causal tissues across different eQTL sample sizes. In panel A, we observe that at type
342   I error is more consistent across different values of $h^2_{ge(t\prime)}$ at smaller eQTL sample sizes. At
343   larger eQTL sample sizes, smaller values of $h^2_{ge(t\prime)}$ have the lowest error rates, with the smallest
344   value not significantly different than 5%. In panel B, we observe that small values of $h^2_{ge(t\prime)}$
345   result in low power, medium values of $h^2_{ge(t\prime)}$ result in the greatest observed power, and
346   (potentially unrealistically) large values of $h^2_{ge(t\prime)}$ result in mediocre power. In panel C,
347   estimates of $h^2_{ge(t\prime)}$ are unbiased across sample sizes and different true values of $h^2_{ge(t\prime)}$. In
348   panel D, there is greater null bias for small eQTL sample sizes and larger values of $h^2_{ge(t\prime)}$ in the
349   causal tissue. These patterns are similar when varying the value of $\omega_{ge(t\prime)}$ in **Extended Data Fig.**
350   **4**.
351   Fourth, we varied the number of causal tissues, considering 1, 2, or 3 causal tissues. We
352   observed that the power of TCSC decreased with multiple causal tissues but did not differ
353   greatly between 2 and 3 causal tissues (**Supplementary Figs. 4-5**); for causal tissues, estimates
354   of $h^2_{ge(t\prime)}$ were upward biased when setting $G_{t\prime}$ to the number of true *cis*-heritable genes.
355   Specifically, in **Supplementary Fig. 4**, we varied the number of causal tissues in the TCSC model.
356   In panel A, the type I error tends to decrease with an increasing number of causal tissues. In
357   panel B, the power to detect two or three causal tissues was significantly less than the power to
358   detect a single causal tissue. In panel C, estimates of $h^2_{ge(t\prime)}$ for causal tissues have anti-
359   conservative bias for two or more tissues. In panel D, null bias decreases when there are
360   multiple causal tissues. We observe similar patterns in **Supplementary Fig. 5** when varying the
361   number of causal tissues in cross-trait TCSC simulations.
362   Fifth, we varied the number of non-causal tissues from 0 to 9. For causal tissues, TCSC
363   estimates were upward biased with fewer tagging tissues but unbiased with more tagging
364   tissues (**Extended Data Figs. 5-6**). TCSC type I error and power were generally higher with fewer
365   tagging tissues; this finding does not compromise our real trait analysis, which involve a large
366   number of tissues. Specifically, in **Extended Data Fig. 5**, we varied the number of non-causal
367   tissues in the TCSC regression. In panel A, type I error decreased when increasing the number of
368   non-causal tissues. In panel B, power was greatest for one or two tagging tissues, but decreased
369   with every additional tagging tissue. In panel C, the estimate of $h^2_{ge(t\prime)}$ for causal tissues had
370   anti-conservative bias when there were fewer than 9 tagging tissues and unbiased when there
371   were 9 tagging tissues. In panel D, null bias of $h^2_{ge(t\prime)}$ for non-causal tissues is not significantly
372   different than zero where there is only one tagging tissue; the most extreme case of anti-
373   conservative null bias occurs at middle numbers of tagging tissues. We observed similar
374   patterns for cross-trait TCSC in **Extended Data Fig. 6**.
375   Sixth, we modified TCSC to not correct for bias in tissue co-regulation scores arising
376   from differences between *cis*-genetic and *cis*-predicted expression. We determined that
377   removal of bias correction resulted in conservative bias in estimates for causal tissues,
378   increased type I error, and similar power (**Supplementary Figs. 6-7**).
379   Seventh, we modified TCSC to apply bias correction to the calculation of all correlations
380   of *cis*-predicted expression contributing to co-regulation scores rather than only those involving

381   the same gene and tissue, which resulted in a decrease in power, anti-conservative bias in
382   estimates for causal tissues, and similar type I error rate (**Supplementary Figs. 8-9**).
383   Eighth, we modified TCSC to use bias-corrected co-regulation scores in the calculation of
384   regression weights, which resulted in similar performance to the default setting
385   (**Supplementary Figs. 10-11**). We note that regression weights pertain to maximizing signal to
386   noise and not avoiding bias in estimates of $h^2_{ge(t)}$; we continue to not perform bias correction
387   when calculating regression weights, consistent with GCSC[8].
388   Ninth, we violated the model assumption that gene-disease effects are independent and
389   identically distributed (i.i.d.) across tissues by including a second causal tissue whose gene-
390   disease effects correlate with varying degree to the gene-disease effects of the original causal
391   tissue (**Supplementary Figs. 12-13**). We determined that while this increases noise to TCSC
392   estimates, the estimates are generally unbiased and TCSC is able to powerfully identify the
393   causal tissue, similar to the addition of a causal tissue where there are no shared gene-disease
394   effects (see **Supplementary Figs. 4-5**).
395   Tenth, we violated the i.i.d. model assumption by duplicating the causal tissue. We
396   determined that TCSC performs well, (e.g. frequently identifies both tissues as causal and
397   estimates $h^2_{ge(t\prime)}$ for both tissues without bias) despite the violation of model assumption
398   (**Supplementary Fig. 14-15**), similar to the previous analysis.
399   Eleventh, we evaluated the robustness of TCSC in the presence of disease heritability
400   that is not mediated via gene expression. We observed that all areas of TCSC performance are
401   affected, with slightly increased type I error rates, decreased power in the case of larger non-
402   mediated heritability, and upward bias in estimates of $h^2_{ge(t\prime)}$ for causal tissues (**Supplementary
403   Fig. 16-17**). However, in simulations incorporating the four larger gene expression sample sizes
404   (see above), we determined that the type I error of TCSC at a given level of non-mediated
405   disease heritability does not increase with larger gene expression sample size (**Supplementary
406   Fig. 18-19**).
407   Finally, we evaluated the robustness of TCSC to variation in the window size used to
408   identify co-regulated genes in the calculation of co-regulation scores and determined that TCSC
409   performance was robust and type I error decreased with larger window sizes (**Supplementary
410   Fig. 20-21**).
411
412   ***Analyzing GTEx tissues***
413
414   We downloaded GTEx v8 gene expression data for 49 tissues. We excluded tissues with
415   fewer than 100 samples, e.g. kidney cortex (n = 69). We retained only European samples for
416   each tissue, as labeled by GTEx via PCA of genotypes. We constructed gene expression models
417   for two scenarios: (1) subsampling to 320 individuals including meta-analyzed tissues (**Table 1**)
418   or (2) using all European samples per tissue. We recommend meta-analyzing gene expression
419   prediction models across tissues in the case of tissues with low eQTL sample size (e.g. < 320
420   samples) and high pairwise genetic correlation (e.g. > 0.93). We determined in simulations that
421   TCSC is sensitive to eQTL sample size differences, such that a tagging tissue with larger sample
422   size than a causal tissue can produce false positive results; the subsampling approach was
423   designed to mitigate this issue. For the subsampling procedure, we first set aside tissues with

424   more than 320 samples; we chose 320 based on the average GTEx tissue sample size (N = 271)
425   and robustness of TCSC in simulations at N = 300. Then, we grouped tissues with genetic
426   correlation, e.g. marginal effect size correlation as reported by GTEx, with R_g > 0.93, an arbitrary
427   threshold that produced biologically plausible groups of related tissues, separating groups of
428   brain tissues based on cranial compartment. We meta-analyzed gene expression prediction
429   models for these grouped tissues in order to achieve a total sample size of 320 individuals
430   where each tissue contributed an approximately equal number of samples, using an inverse-
431   variance weighted meta-analysis across genes that were significantly *cis*-heritable in two or
432   more constituent tissues. The prediction weights of genes that were significantly *cis*-heritable in
433   a single constituent tissue were left unmodified.
434
435   ***Extended primary analysis of tissue-specific contributions to diseases and complex traits.***
436
437   WHRadjBMI (waist-hip-ratio conditional on body mass index) and subcutaneous adipose tissue
438   ($\pi_{t'}$ = 0.10, s.e. = 0.037, $P = 2.4 \times 10^{-3}$). A previous study comparing subcutaneous adipose
439   tissue to visceral adipose tissue found that the level of adiponectin, a hormone released by
440   adipose tissue to regulate insulin, is specifically associated with subcutaneous adipose tissue
441   and not visceral adipose tissue; and, adiponectin levels are significantly negatively correlated
442   with waist-hip-ratio[14]. Furthermore, LDSC-SEG found WHRadjBMI to be associated not only
443   with subcutaneous adipose, but also with visceral adipose tissue. While RTC Coloc finds many
444   WHR-associated tissues, it was able to distinguish subcutaneous adipose (FDR = 2.9×10^{-4})
445   from visceral adipose (FDR = 1).
446
447   HDL (high density lipoprotein) with subcutaneous adipose tissue ($\pi_{t'}$ = 0.159, s.e. = 0.054, P =
448   1.5×10^{-3}) and whole blood ($\pi_{t'}$ = 0.098, s.e. = 0.034, $P = 1.8 \times 10^{-3}$). Previous work has
449   implicated subcutaneous adipose tissue in mediating HDL levels, as this tissue stores cholesterol
450   and expresses genes involved in cholesterol transport and HDL lipidation[15]. The relationship
451   with whole blood is likely due to the role that red blood cells play in cholesterol transport, while
452   being a large proportion of cells in whole blood samples[16]. Notably, TCSC did not identify liver
453   as a causal tissue for HDL, and this might be due to the smaller eQTL sample size of liver which
454   limits the power to detect this association.
455
456   BMI (body mass index) and brain cerebellum ($\pi_{t'}$ = 0.042, s.e. = 0.015, $P = 2.6 \times 10^{-3}$). While
457   several studies have found that the central nervous system is enriched for genetic variation
458   associated with BMI and obesity[17,2,18], the precise causal brain tissue is uncertain. Neither LDSC-
459   SEG nor RTC Coloc can distinguish between highly co-regulated brain tissues, such as the
460   cerebellum and cortex. Previous studies have indicated that the brain cerebellum takes part in
461   regulating feeding control (for example via connection to the hypothalamus) and therefore can
462   have substantial impacts on obesity related traits and diseases[19]. Moreover, differential activity
463   has been observed in the brain cerebellum in individuals experiencing hunger, thirst, or
464   satiation[19]. Furthermore, a different study associated the brain cerebellum with endocrine
465   homeostasis, suggesting that the cerebellum plays several important biological roles, rather
466   than strictly motor control[20]. A more recent multi-omics approach identified that cerebellar

467   nuclei in mice are activated when they are eating and even suggests a potential therapeutic
468   target for the management of excessive eating behavioral traits[21].
469
470   Fecundity and brain cerebellum ($\pi_{t'}$ = 0.075, s.e. = 0.024, P = 9.1×10^{-4}). This is consistent
471   with the known relationship between fertility and energy metabolism, involving hormone
472   secretion, which is largely regulated by the brain. However, previous studies have specifically
473   linked fertility-related hormonal dysregulation to the hypothalamus and brainstem[22,23].
474
475   Total protein and fibroblasts ($\pi_{t'}$ = 0.079, s.e. = 0.025, P = 7.0×10^{-4}) and whole blood ($\pi_{t'}$ =
476   0.081, s.e. = 0.027, P = 1.5×10^{-3}). Fibroblasts are cells that play diverse roles across the
477   tissues of the body, markedly producing protein complexes that constitute the extracellular
478   matrices that define the structure of fibroblasts[24]. Serum protein is a quantity measured from
479   whole blood, explaining the second relationship.
480
481   Cerebral cortex surface area and fibroblasts ($\pi_{t'}$ = 0.10, s.e. = 0.034, P = 1.8×10^{-3}). Tissue
482   surface areas are likely related to developmental processes governing body proportions. As
483   stated in the main text, TCSC identified fibroblasts (and skeletal muscle) as causal tissues for
484   height, the most commonly studied anthropometric phenotype, which suggests that
485   fibroblasts, as a connective tissue, likely regulates the growth of different organs and tissues.
486
487   Lipid traits and liver: AST ($\pi_{t'}$ = 0.077, s.e. = 0.025, P = 9.2×10^{-4}), RBC width ($\pi_{t'}$ = 0.077, s.e.
488   = 0.027, P = 1.9×10^{-3}), total cholesterol ($\pi_{t'}$ = 0.14, s.e. = 0.044, P = 5.3×10^{-4}), Bilirubin
489   ($\pi_{t'}$ = 0.11, s.e. = 0.036, P = 1.0×10^{-3}). These causal tissue-trait pairs are reasonable as the
490   liver is the production center of cholesterol and phospholipids.
491
492   Blood cell traits and whole blood: eosinophil count ($\pi_{t'}$ = 0.17, s.e. = 0.052, P = 6.5×10^{-4}),
493   lymphocyte count ($\pi_{t'}$ = 0.22, s.e. = 0.053, P = 2.1×10^{-5}), monocyte count ($\pi_{t'}$ = 0.25, s.e. =
494   0.078, P = 7.5×10^{-4}). These causal tissue-trait pairs are reasonable as these different blood
495   cell populations are present in whole blood.
496
497   MDD (Major depressive disorder) and whole blood ($\pi_{t'}$ = 0.068, s.e. = 0.022, P = 1.3×10^{-3}).
498   This is consistent with reports of elevated immune system cytokines in MDD cases[25].
499
500   ***Secondary analysis of tissue-specific contributions to diseases and complex traits (N = 320***
501   ***tissues only)***
502
503   Tissues with smaller eQTL sample size may be underpowered in TCSC analysis. This prompted
504   us to remove tissues with eQTL sample size less than 320 individuals. The number of causal
505   tissue-trait pairs with significantly positive contributions to disease/trait heritability (at 5% FDR)
506   increased from 21 to 23, likely due to a decrease in multiple hypothesis testing burden from
507   removing underpowered tissues. The 23 significant tissue-trait pairs reflect a gain of 8 newly
508   significant tissue-trait pairs (and a loss of 6 formerly significant tissue-trait pairs, of which 5
509   were lost because the tissue was removed), but estimates of $\pi_{t'}$ for each significant tissue-trait
510   pair were not statistically different from our primary analysis (**Supplementary Table 9**).

511 Notably, among the newly significant tissue-trait pairs, whole blood was associated with
512 hypothyroidism ($\pi_{t'}$ = 0.100, s.e. = 0.032, P = 8.9×10^{-4}); we note that thyroid had a
513 quantitatively large but only nominally significant association ($\pi_{t'}$ = 0.452, s.e. = 0.225, P = 0.02,
514 FDR = 26%). Esophagus muscularis (rather than lung tissue) was associated with the lung trait
515 FEV1/FVC[26] ($\pi_{t'}$ = 0.167, s.e. = 0.056, P = 1.4×10^{-3}). This result may be explained by the fact
516 that smooth muscle in the lung is known to affect FEV1/FVC and influence pulmonary disease
517 pathopysiology[27], and this unobserved causal tissue is likely highly co-regulated with the
518 smooth muscle of the esophagus, which is indeed the site from which the GTEx study sampled
519 the esophagus muscularis tissue[5]. Other newly significant findings are discussed below and
520 numerical results for all tissues and diseases/traits are reported in **Supplementary Table 8**.

522 BMI and tibial nerve: This is broadly consistent with the role of the central nervous system in
523 BMI[28,17,29,30,2,31], although the precise causal relationship that might exist between tibial nerve
524 and BMI is not straightforward.

526 Additional causal tissues for platelet count identified in this secondary analysis include brain
527 cortex, esophagus muscularis, and fibroblasts. Regarding the brain, platelets are often found in
528 blood vessels and are key participants in thrombosis, or the clotting of blood vessels[32,33].
529 Moreover, platelets have been linked to inflammation of death of neurons in the cortex and
530 hippocampus[34]. Regarding the esophagus muscularis, high platelet counts are associated with
531 greater severity of esophageal cancer, likely due to the angiogenic properties of platelets, e.g.
532 creating new blood vessels[35]. Regarding fibroblasts, these cells are known to be recruited to
533 sites of blood clots, caused by platelets, to remedy the clot[36]. Therefore, increased presence of
534 fibroblasts likely reduces platelet activity in individuals with greater susceptibility to vascular
535 clotting. While it is possible that platelet count may have a diverse tissue-specific genetic basis,
536 this result could also be caused by an absent causal tissue or cell type that is co-regulated with
537 these three newly detected tissues.

539 Sleep duration and breast tissue: melatonin is a hormone whose levels are considered
540 protective for breast cancer risk[37]. Melatonin is also a common supplement taken to promote
541 sleep. However, melatonin is produced in the brain, and therefore the causal relationship from
542 breast tissue to sleep duration is unclear.

544 Height with fibroblasts and muscle skeletal tissue. Skin tissue has been shown to widely express
545 growth factors, including embryonic growth factor which plays a key role in fetal
546 development[38]. Fibroblasts are the predominant cell type of skin tissue. Skeletal muscle is one
547 of the most likely causal tissues for anthropometric, or skeletal growth, traits such as height,
548 consistent with previous genetic studies identifying enrichments of height-associated genetic
549 variation near genes regulated in skeletal muscle[17,2], which includes colocalization with eQTLs
550 regulating key growth factors such as IGFBP-3[39].

552 RBC count with fibroblasts and whole blood: Red blood cells and fibroblasts work together
553 during tissue remodeling processes of extracellular matrices[40,41]. However, these studies
554 suggest that red blood cells stimulate fibroblasts to secrete important tissue remodeling

555 molecules, such as interleukin-8 and metalloproteinases. As a blood cell population, the causal
556 relationship between whole blood and red blood cell count is expected.

557

558 Eosinophil count with fibroblasts and muscle skeletal tissue: Similar to the role of red blood
559 cells in tissue remodeling described above, eosinophils also interact with fibroblasts in tissue
560 remodeling and fibrosis, although typically in response to inflammation and allergy[42].
561 Eosinophils have previously been implicated in myopathy, or muscular disease[43], likely due to
562 their recruitment in response to allergy, infection, or cancer.

563

564 Testosterone and muscle skeletal tissue: atrophy of skeletal muscle is associated with lower
565 levels of testosterone, a hormone produced by the testes and understood to be regulated by
566 brain tissues[44,45]. These studies suggest that there is a causal relationship of testosterone on
567 muscle skeletal tissue, rather than the reverse relationship suggested by TCSC.

568

569 ***Secondary analysis of brain-specific contributions to diseases and complex traits***
570 While a subset of our diseases/traits are psychiatric and behavioral phenotypes, we sought to
571 increase the power of our TCSC analysis by restricting tissues to those that are in the brain. We
572 identified 41 independent brain-related traits, reflecting a less stringent squared genetic
573 correlation threshold of 0.25. We relaxed our threshold so that we would have a substantial
574 number of brain traits to analyze, as many would were excluded under the original threshold of
575 0.1. The 13 GTEx brain tissues were analyzed without merging tissues into meta-tissues, and
576 irrespective of eQTL sample size (range: N = 101-189 individuals); we expected power to be
577 limited due to the eQTL small sample sizes and substantial co-regulation among individual brain
578 tissues. TCSC identified 8 brain tissue-brain trait pairs at 5% FDR (**Extended Data Fig. 8,**
579 **Supplementary Table 12**). For ADHD, TCSC identified brain hippocampus as a causal tissue
580 ($\pi_{t'}$ = 0.127, s.e. = 0.045, P = 2.5 $\times 10^{-3}$), consistent with the correlation between
581 hippocampal volume and ADHD diagnosis in children[46]. A recent ADHD GWAS identified a locus
582 implicating the *FOXP2* gene[47], which has been reported to regulate dopamine secretion in
583 mice[48]; hippocampal activation results in the firing of dopamine neurons[49]. For BMI, TCSC
584 identified brain amygdala ($\pi_{t'}$ = 0.054, s.e. = 0.023, P = 8.3 $\times 10^{-3}$) and brain cerebellum ($\pi_{t'}$ =
585 0.039, s.e. = 0.016, P = 7.0 $\times 10^{-3}$) as causal tissues, consistent with previous work linking the
586 amygdala to obesity and dietary self-control[50], although no previous study has implicated the
587 amygdala in genetic regulation of BMI. As for brain cerebellum, previous research has
588 implicated the cerebellar function in dietary behavior, rather than strictly regulation motor
589 control function[20,19,21]. We note that the brain-specific analysis is expected to have greater
590 power to identify tissue-trait pairs than the analysis of **Fig. 3** due to the smaller number of total
591 tissues in the model (as simulations show higher power for TCSC when there are fewer tagging
592 tissues; **Extended Data Fig. 5**). Numerical results for all brain tissues and brain traits analyzed
593 are reported in **Supplementary Table 12**.

594

595 Caudate volume and accumbens: In individuals with major depressive disorder, the basal
596 ganglia, of which the nucleus accumbens is a component, has an attenuated response to
597 positive stimuli compared to healthy controls; and, it has been observed that this associates
598 with reduced caudate volume[51].

599

600 Anisotropy mode with accumbens and cerebellum: Mode of anisotropy reflects the
601 organization of white matter fibers in the brain and is used to suggest abnormalities in brain
602 connections[52]. Therefore, any brain tissue connected to white matter could be causal for
603 morphological anisotropy mode; indeed the nucleus accumbens and cerebellum have
604 connections to white matter[53,54].

605

606 Schizophrenia with brain frontal cortex, brain cerebellum, and brain caudate. The association
607 with the frontal cortex is consistent with previous studies reporting differences in gray and
608 white matter volumes in schizophrenia cases vs. controls within the prefrontal cortex[55,56].
609 Previous large-scale genetic studies identified enrichments of schizophrenia-associated variants
610 in gene sets regulating excitatory and inhibitory neurons[2,57,58], but did not distinguish the origin
611 of this enrichment among the cortex, hippocampus, and amygdala. The association with the
612 cerebellum might be due to its large proportion of neurons, and is also consistent with previous
613 reports of decreased blood flow within the cerebellum in schizophrenia patients[59]. The
614 association with caudate is consistent with early studies reporting schizophrenia-like
615 characteristics in patients with damaged caudate projections[60,61]. While TCSC often identifies
616 one causal tissue for a given trait, the identification of three causal tissues for schizophrenia
617 may reflect a diverse tissue-specific genetic basis for the disease, the absence of the true causal
618 tissue or cell type and its co-regulation with analyzed tissues, or the common presence of the
619 true causal cell type among each of the three tissues.

620

621 Bipolar disorder with caudate and cerebellum: This is consistent with previous work linking
622 reduced cerebellar volume to anxiety-related disorders[62,63] and is similarly consistent with
623 previous work associating reduced caudate volumes with bipolar disorder[64].

624

625 Reaction time and cerebellum: This is consistent with previous studies in patients and monkeys
626 with reduced reaction time and cerebellar lesions[65].

627

628 Cerebral cortex width with frontal cortex and spinal cord: Intuitively, the frontal cortex has a
629 causal effect on the tissue of the same name. While the connection between spinal cord and
630 cerebral cortex is not as straightforward, the spinal cord and hypothalamus are connected via
631 hypothalamic projections[66] and hypothalamic projections to the cerebral cortex are responsible
632 for propagating autonomic signaling[67].

633

634 Starting age of smoking habit and frontal cortex: This is consistent with previous work reporting
635 that development of the frontal cortex during adolescence is associated with behaviors and
636 lifestyle choices, such as smoking[68].

637

638 Brainstem volume and spinal cord: This is consistent with the brainstem being the connection
639 point of the brain to the spinal cord[69].

640

641 ***Brain-specific comparison across RTC Coloc, LDSC-SEG and TCSC***

642

643   In the brain-specific analysis, patterns of LDSC-SEG and RTC Coloc were striking. First, LDSC-SEG
644   did not identify heritability enrichments in any brain tissues other than cerebellum and cortex,
645   suggesting that these two tissues are the only disease relevant parts of the brain, although this
646   is highly unlikely. For example, for four traits LDSC-SEG produced very similar enrichments for
647   the frontal cortex and the cortex. TCSC attributed these associations to the brain cerebellum,
648   and in the specific case of schizophrenia, also implicated the frontal cortex. Second, six of the
649   ten brain traits, for which TCSC identified a causal tissue at 10% FDR, had no associated tissue
650   according to LDSC-SEG; these traits coincided with traits not analyzed by the RTC Coloc study.
651   We note that the RTC Coloc study did not analyze all GTEx tissues; brain amygdala, spinal cord,
652   and substantia nigra were omitted from their study. Lastly, RTC Coloc found 8 of 8 tested
653   tissues shown in **Extended Data Fig. 10** to be associated with schizophrenia and four of 8 tested
654   tissues to be associated with BMI, a superset of the tissues implicated by TCSC.

656   ***Extended analysis of tissue-specific contributions to genetic covariance***

658   We note that the direction of effect of tissue-specific contributions to the genetic covariance
659   between two traits may be in the opposite direction of the global covariance between two
660   traits, analogous to how local contributions to genome-wide genetic correlation may be in the
661   opposite direction of the genome-wide genetic correlation[70-73].

663   Before discussing all significant findings, we discuss two particularly compelling examples. First,
664   brain substantia nigra had a significantly negative contribution to the genetic covariance of age
665   at first birth and height ($\zeta_{t'}$ = -0.11, s.e. = 0.032, $P = 4.5 \times 10^{-4}$). Previous work in *C. elegans*
666   reported that fecundity is positively regulated by dopamine[74,75], which is produced in the
667   substantia nigra[76]. Therefore, it is plausible that reproductive outcomes related to fecundity,
668   such as age at first birth, are also regulated by dopamine via the substantia nigra. Dopamine
669   also plays a role in regulating the levels of key growth hormones such as IGF-1 and IGF-BP3[77]
670   and has been previously shown to be associated with height[78]. Second, pituitary had a
671   significantly negative contribution to the genetic covariance of vitamin D and WHR | BMI ($\zeta_{t'}$ =
672   -0.19, s.e. = 0.057, $P = 4.5 \times 10^{-4}$). Irregularities in pituitary development are associated with
673   decreased vitamin D levels and decreased IGF-1 levels, the latter of which is integral for bone
674   development and is directly proportional to body proportion phenotypes such as WHR | BMI[79-
675   81].

677   Negative contribution of brain cortex to the genetic covariance of neuroticism and years of
678   education ($\zeta_{t'}$ = -0.10, s.e. = 0.029, $P = 2.1 \times 10^{-4}$). When certain personality traits underlie
679   neuroticism, such as conscientiousness, neuroticism has been shown to be positively correlated
680   with educational success[82]. The specific implication of the brain cortex, as opposed to other
681   brain tissues, has not be reported previously in the literature.

683   Positive contribution of the brain spinal cord to the genetic covariance of type 2 diabetes (T2D)
684   and vitamin D ($\zeta_{t'}$ = 0.17, s.e. = 0.052, $P = 5.5 \times 10^{-4}$). Vitamin D is a known neurosteroid,
685   which affects various brain functions including calcium signaling and cellular differentiation[83,

686   and reduced vitamin D is a prominent risk factor for infectious diabetes as well as diabetes
687   mellitus (a subset of which is T2D)[84], explaining the negative covariance identified by TCSC.
688
689   Negative contribution of breast tissue to the genetic covariance of white blood cell count and
690   BMI ($\zeta_{t'}$ = -0.16, s.e. = 0.041, $P = 4.1 \times 10^{-5}$). This observation is consistent with many previous
691   studies reporting an association of elevated white blood cell counts with breast cancer, as these
692   cells are a biomarker of inflammation and are predictive of other cancers and cardiovascular
693   disease[85-87]. One of these studies investigated this relationship in the context of BMI and found
694   that in premenopausal women, individuals with lower BMI and breast cancer had elevated
695   white blood cell counts[87]. This direction of effect is consistent with TCSC's detection of tissue-
696   specific negative covariance between white blood cell count and BMI, despite a genome-wide
697   positive genetic correlation of these two traits.
698
699   Negative contribution of lung to the genetic covariance of age at first birth and intelligence ($\zeta_{t'}$
700   = -0.096, s.e. = 0.026, $P = 1.2 \times 10^{-4}$). First, previous work has found that older age at first birth
701   is associated with reduced risk of lung cancer involving regulation by steroid hormones; and
702   while some studies consider age at first birth to be a causal protective factor, this relationship
703   might be better explained by reverse causality[88-91]. Indeed, TCSC is not impacted by reverse
704   causality as phenotype cannot influence gene expression-modifying genetic variation. Second,
705   positive health outcomes, including lung function, are genetically associated with cognitive
706   traits in GWAS, although the causal mechanisms are poorly understood[92-94]. However, the
707   direction of effect estimated by TCSC is inconsistent with these findings, possibly suggesting
708   distinct causal mechanisms of lung tissue on these traits.
709
710   Negative contribution of lung to the genetic covariance of intelligence and years of education
711   ($\zeta_{t'}$ = -0.045, s.e. = 0.013, $P = 2.2 \times 10^{-4}$). As stated above, we would expect lung genes with a
712   positive effect on intelligence to have a consistent direction of effect on years of education.
713   While this is not what TCSC concludes, this may suggest distinct causal mechanisms of lung
714   tissue on these traits.
715
716   Negative contribution of pituitary to the genetic covariance of vitamin D and WHRadjBMI ($\zeta_{t'}$ =
717   -0.19, s.e. = 0.057, $P = 4.5 \times 10^{-4}$). Previous work has established a relationship between
718   vitamin D and bone structure development, which is directly related to WHRadjBMI[80]. Other
719   work has suggested that this may be due to the positive correlation between vitamin D levels
720   and growth hormone levels, such as IGF-1[79]. Moreover, irregularities in pituitary development
721   (specifically pituitary stalk interruption syndrome) are associated with reduced IGF-1, in which
722   individuals also have reduced serum levels of vitamin D[81].
723
724   Contributions to the genetic covariance of eosinophil count and platelet count by lung($\zeta_{t'}$ = -
725   0.20, s.e. = 0.068, $P = 1.5 \times 10^{-3}$), ovary ($\zeta_{t'}$ = -0.15, s.e. = 0.052, $P = 2.2 \times 10^{-3}$), skin ($\zeta_{t'}$ =
726   0.28, s.e. = 0.087, $P = 7.1 \times 10^{-4}$), and whole blood ($\zeta_{t'}$ = 0.30, s.e. = 0.105, $P = 2.2 \times 10^{-3}$).
727   Eosinophils and platelets are highly co-regulated, with eosinophils secreting platelet-activating
728   enzymes[95]. Therefore, it is expected that across multiple tissues, genes have pleiotropic effects

729   on eosinophil count and platelet counts. It is also possible that such genes have a direct effect
730   on eosinophil count and a secondary effect mediated be eosinophils on platelet count.
731
732   Negative contribution of vagina to the genetic covariance of testosterone and vitamin D ($\zeta_{t'}$ = -
733   0.25, s.e. = 0.079, $P = 6.5 \times 10^{-4}$). Previous work has shown that vaginal tissue growth and
734   differentiation were improved as a result of increased vitamin D levels[96]. Similarly, testosterone
735   is a hormone that plays a key role in healthy vaginal function[97]. Since TCSC detected negative
736   covariance for these two traits, it is likely that they are regulated by distinct sets of genes.
737
738   Negative contribution of whole blood to the genetic covariance of age at first birth and
739   rheumatoid arthritis ($\zeta_{t'}$ = -0.17, s.e. = 0.053, $P = 6.6 \times 10^{-4}$). The association between
740   rheumatoid arthritis (RA) and whole blood, which is comprised of many immune cell types, is
741   logical. However, previous work has not reported an association between whole blood, or the
742   immune system, and age at first birth. Moreover, it is not immediately clear why genes that
743   increase risk for RA would also increase age at first birth. We hypothesize that the underlying
744   mechanism pertains to age-related changes in an individual's immune system which might
745   affect reproductive behavior later in life, as risk for RA and other autoimmune diseases
746   increases.
747
748   Negative contribution of spleen to the genetic covariance of major depressive disorder (MDD)
749   and BMI ($\zeta_{t'}$ = -0.29, s.e. = 0.039, $P = 7.3 \times 10^{-4}$). As discussed above, whole blood was
750   detected as causal tissue for MDD likely due to the role of the cytokines in regulation of MDD;
751   the spleen plays a key role in the immune system. It is widely known that obesity, or high BMI,
752   is associated with irregularities in immune cell counts[98].
753
754   Negative contribution of coronary artery to the genetic covariance of years of education and
755   menopause age ($\zeta_{t'}$ = -0.13, s.e. = 0.115, $P = 7.3 \times 10^{-4}$). This is consistent with previous work
756   indicating that reduced coronary artery disease risk is associated with more years of education
757   via a Mendelian randomization study[99]. Late menopause is considered a protective factor
758   coronary artery disease[100]. This biological consistency would suggest a positive covariance,
759   therefore we might conclude that distinct sets of genes regulate years of education and
760   menopause age in coronary artery.
761
762   Positive contribution of aorta artery to the genetic covariance of mode of anisotropy and
763   menopause age ($\zeta_{t'}$ = 0.27, s.e. = 0.084, $P = 7.8 \times 10^{-4}$). Previous work has associated
764   calcification in the aorta with bone loss, specifically in postmenopausal women[101]. Mode of
765   anisotropy, from brain MRI which is a measure of structural cellular organization, is not a well-
766   studied complex trait and has a lack of literature evidence to support any role for aorta, and
767   similarly for any other tissue.
768
769   Contributions to the genetic covariance of anorexia and insomnia by tibial nerve ($\zeta_{t'}$ = 0.20, s.e.
770   = 0.063, $P = 7.8 \times 10^{-4}$), testis ($\zeta_{t'}$ = 0.18, s.e. = 0.062, $P = 1.7 \times 10^{-3}$), and whole blood ($\zeta_{t'}$ = -
771   0.14, s.e. = 0.053, $P = 3.9 \times 10^{-3}$). The explanation for tibial nerve can be found in the main
772   text. The association with whole blood might be explained by previous work demonstrating the

773   role of the immune system in anorexia[102] and insomnia[103]. The immune regulation impacting
774   insomnia is specifically discussed in the context of the central nervous system, further
775   supporting the association with the tibial nerve, a central nervous system tissue. Testis size has
776   also been associated with sleep irregularities[104]. Separately, the reduced production of the
777   androgen hormone in the testis, or hypogonadism, is a comorbidity of male anorexia[105].
778
779   Negative contribution of prostate to the genetic covariance of medication use and years of
780   education ($\zeta_{t'}$ = -0.075, s.e. = 0.024, P = 7.8×10^{-4}). This is consistent with previous studies
781   establishing a negative association between drug use and prostate health outcomes[106]. The
782   positive covariance detected with years of education is not supported by any literature
783   evidence and could be a false positive.
784
785   Positive contribution of muscle skeletal to the genetic covariance of brain accumbens volume
786   and caudate volume ($\zeta_{t'}$ = 0.20, s.e. = 0.063, P = 8.0×10^{-4}). Previous work indicates that
787   musculoskeletal tissue likely influences biological processes the brain via regulation of energy
788   metabolism[107].
789
790   Negative contribution of muscle skeletal to the genetic covariance of total protein and
791   WHRadjBMI ($\zeta_{t'}$ = -0.27, s.e. = 0.084, P = 8.0×10^{-4}). This is consistent with known regulation
792   in musculoskeletal tissue influencing waist-hip-ratio[108]. Musculoskeletal tissue is also related to
793   protein levels, as restricted protein intake leads to dysregulation and morphology of skeletal
794   muscle[109].
795
796   Negative contribution of skin to the genetic covariance of height and FVC ($\zeta_{t'}$ = -0.47, s.e. =
797   0.155, P = 1.2×10^{-3}). Height and forced vital capacity (FVC) are genetically correlated as both
798   are affected by body proportions and growth-regulating processes. Skin tissue has been shown
799   to widely express growth factors, including embryonic growth factor which plays a key role in
800   fetal development[38].
801
802   Negative contribution of lung to the genetic covariance of age at first birth and menopause age
803   ($\zeta_{t'}$ = -0.20, s.e. = 0.067, P = 1.3×10^{-3}). This is consistent with a negative association between
804   earlier menopause age and healthy pulmonary function[110,111]. As described above, older age at
805   first birth is associated with improved lung cancer outcomes[112].
806
807   Negative contribution of adipose subcutaneous to the genetic covariance of bipolar disorder
808   and major depressive disorder ($\zeta_{t'}$ = -0.18, s.e. = 0.051, P = 2.7×10^{-4}). This is consistent with
809   an expanding body of literature supporting a bidirectional link between obesity and
810   depression[113].
811
812   Negative contribution of the meta-tissue brain limbic to the genetic covariance of risk tolerance
813   and schizophrenia ($\zeta_{t'}$ = -0.15, s.e. = 0.049, P = 8.5×10^{-4}). Risk taking (or impulsivity) has
814   previously been linked to both schizophrenia and bipolar disorder[114].
815
816   ***Extended analysis of differences in tissue-specific contributions to heritability vs. covariance***

817

818 We note that $\zeta_{t'}$ and $\pi_{t'}$ are both signed proportions and are therefore on the same scale, thus
819 the scenario in which these two quantities are equal is a natural and parsimonious null.

820

821 Negative contribution of skin (sun exposed) to the genetic covariance of height and FVC: Skin
822 does not explain a nonzero proportion of heritability for height or FVC; however, skin does
823 explain a significant amount of positive covariance, although the genome-wide covariance of
824 this trait pair is negative.

825

826 Negative contribution of breast to the genetic covariance of WBC count and BMI: Breast is not a
827 causal tissue for either trait, although it does explain a significant amount of negative
828 covariance between the two traits, although the genome-wide covariance of this trait pair is
829 positive.

830

831 Negative contribution of brain cortex to the genetic covariance of years of education and
832 neuroticism: Brain cortex is not a causal tissue for either trait, consistent with what we found in
833 the brain-specific analysis. However, brain cortex explains a significant amount of positive
834 covariance between the two traits, although the genome-wide covariance is negative.

835

836 Negative contribution of pituitary to the genetic covariance of vitamin D and WHRadjBMI:
837 While pituitary does not explain a nonzero proportion of the heritability of either trait, it does
838 explain a significant amount of positive covariance between the two traits, although the
839 genome-wide covariance is negative.

840

841 ***Other tissue-trait association methods***

842

843 MaxCPP models contributions to heritability enrichment of fine-mapped eQTL variants across
844 tissues or meta-tissues[115]; although this approach proved powerful when analyzing eQTL effects
845 that were meta-analyzed across all tissues, it has limited power to identify disease-critical
846 tissues: fine-mapped eQTL annotations for blood (resp. brain) were significant conditional on
847 annotations constructed using all tissues only when meta-analyzing results across a large set of
848 blood (resp. brain) traits (Fig. 4 of ref.[115]). eQTLenrich compares eQTL enrichments of disease-
849 associated variants across tissues[116]; this approach produced compelling findings for eQTL that
850 were aggregated across tissues, but tissue-specific analyses often implicated many tissues (Fig.
851 1d of ref.[116]). MESC estimates the proportion of heritability causally mediated by gene
852 expression in assayed tissues[117]; this study made a valuable contribution in its strict definition
853 and estimation of mediated effects (see below), but did not jointly model distinct tissues and
854 had limited power to distinguish disease-critical tissues (Fig. 3 of ref.[117]). CAFEH leverages multi-
855 trait fine-mapping methods to simultaneously evaluate all tissues for colocalization with
856 disease[118]; however, this locus-based approach does not produce genome-wide estimates and
857 it remains the case that many (causal or tagging) tissues may colocalize with disease under this
858 framework. Likewise, methods for identifying tissues associated to disease/trait covariance do
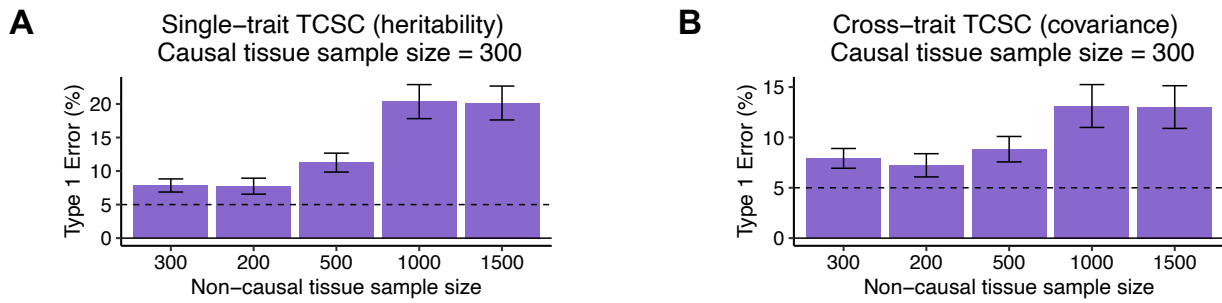859 not distinguish causal tissues from tagging tissues[119,120].

860

Other limitations

1. TCSC has low power at small eQTL sample sizes; in addition, TCSC estimates are impacted by the number of significantly *cis*-heritable genes in a focal tissue, which can lead to conservative bias at small eQTL sample sizes. We anticipate that these limitations will become less severe as eQTL sample sizes increase.

2. TCSC is susceptible to large variations in eQTL sample size, which may compromise type I error; therefore, there is a tradeoff between maximizing the number of tissues analyzed and limiting the variation in eQTL sample size.

3. TCSC assumes that causal gene expression-disease effects are independent across tissues; this assumption may become invalid for tissues and cell types assayed at high resolution. However, we verified via simulations that TCSC performs well when this model assumption is violated (**Supplementary Figs. 12-15**).

4. TCSC does not formally model measurement error in tissue co-regulation scores, but instead applies a heuristic bias correction. We determined that the bias correction generally performs well in simulations.

5. Eighth, TCSC does not produce locus-specific estimates or identify causal tissues at specific loci. However, genome-wide results from TCSC may be used as a prior for locus-based methods (analogous to GWAS fine-mapping with functional priors[121]).

6. We did not apply TCSC to single-cell RNA-seq (scRNA-seq) data, which represents a promising new direction as scRNA-seq sample sizes increase[122-124,6]; we caution that scRNA-seq data may require new eQTL modeling approaches[122].

7. Finally, we focused our cross-trait analyses on relatively independent traits from the single-trait analysis, to enable comparisons with single-trait results (**Fig. 5B, 5C**); cross-trait analysis of more strongly genetically correlated traits is a future direction of high interest.

904 **Supplementary Figures**

A Single–trait TCSC (heritability)
Causal tissue sample size = 300

B Cross–trait TCSC (covariance)
Causal tissue sample size = 300

905

906 **Supplementary Figure 1. Type I error of TCSC regression in simulations with large variations in**
907 **eQTL sample size of non-causal tissues.** We performed n = 1,000 independent simulated
908 genetic architectures in which each simulation had one causal tissue (gene expression sample
909 size = 300 individuals) and nine non-causal tissues with the following sample sizes: 300, 200,
910 300, 500, 1,000, 1500, 200, 300, 500. (A) We report the false positive rate for non-causal tissues
911 as $h^2_{ge(t\prime)}$ > 0 at $p < 0.05$ which is not well-controlled, demonstrating the need for comparable
912 gene expression sample sizes across tissues in TCSC. (B) We report the false positive rate for
913 non-causal tissues as $\omega_{ge(t\prime)}$ > 0 at $p < 0.05$ which is not well-controlled, demonstrating the
914 need for comparable gene expression sample sizes across tissues in TCSC. For panels A and B,
915 we used a one-sided z-test and the genomic block jackknife standard error to obtain p-values
916 and data are presented as mean values +/- 1.96 x SEM.
917

918
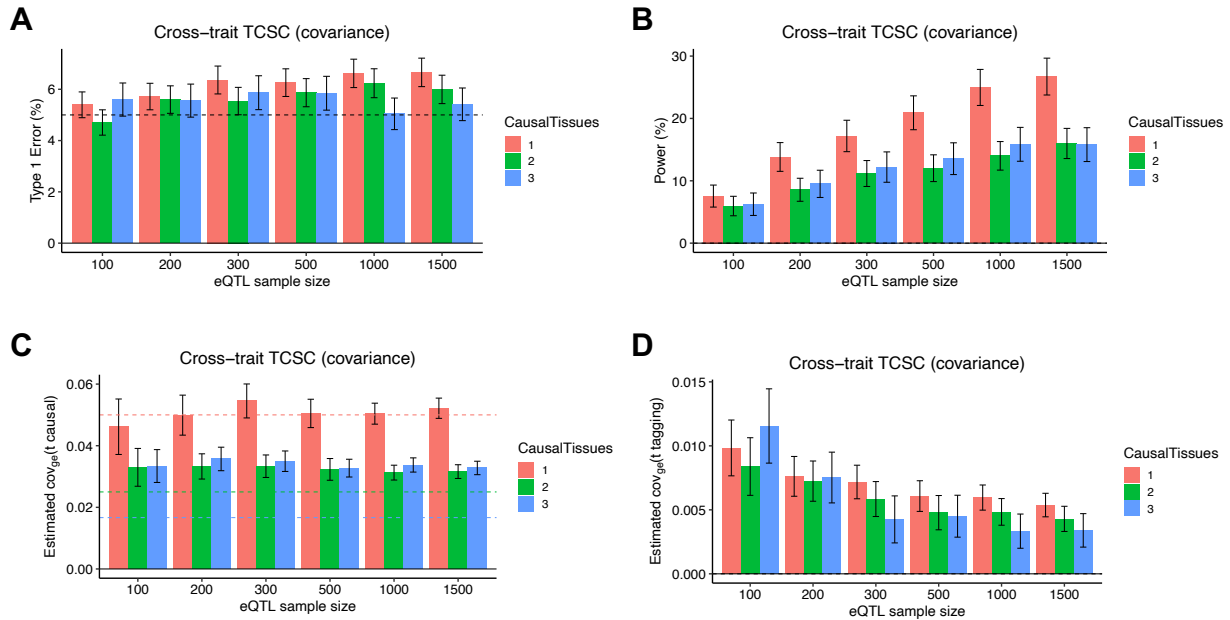

919
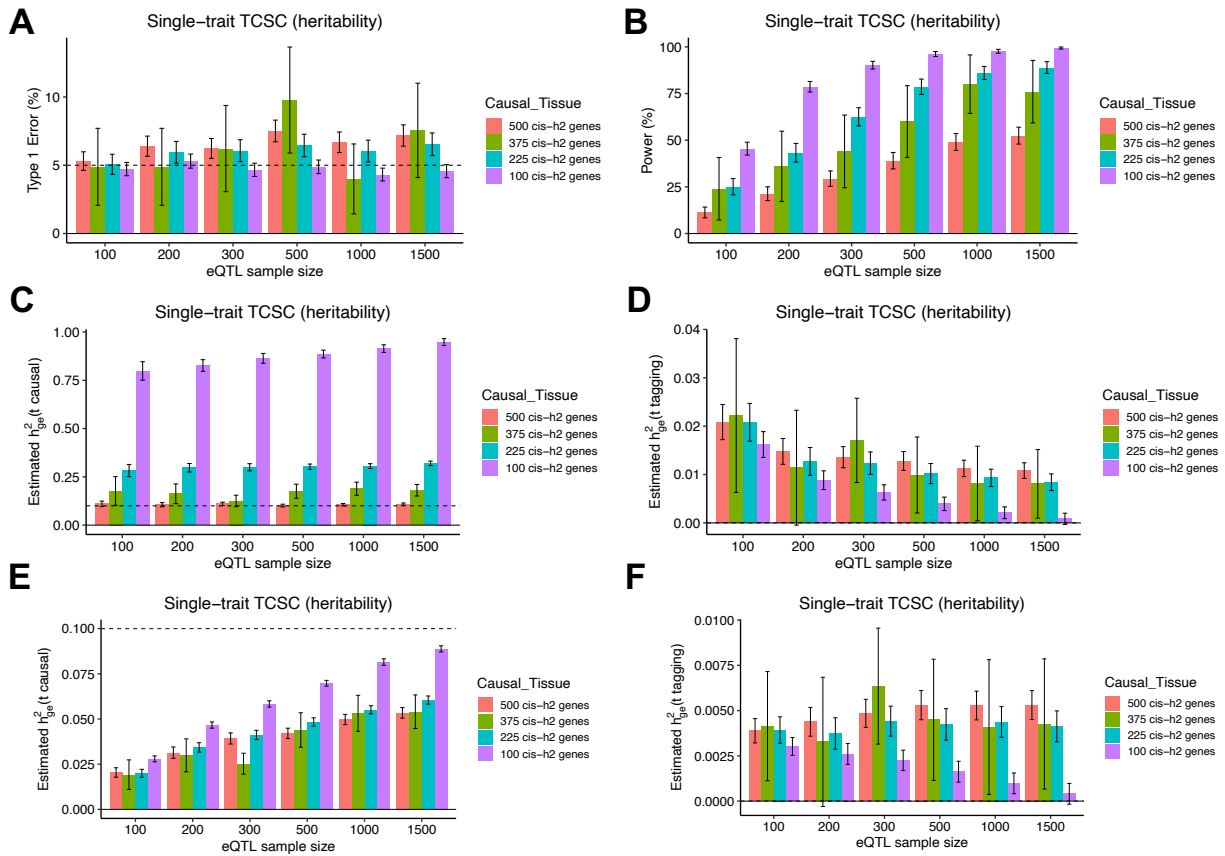
Supplementary Figure 2. Robustness and power of TCSC regression in simulations when the causal tissue has fewer cis-heritable genes than tagging tissues. (A) Type I error when changing the number of expressed genes in the causal tissue. False positive event is defined as $h^2_{ge(t\prime)} > 0$ for non-causal tissues at $p < 0.05$. (B) Power to detect the causal tissue per scenario. A true positive event is defined as $h^2_{ge(t\prime)} > 0$ for causal tissues at $p < 0.05$. (C) Bias on causal estimates of $h^2_{ge(t\prime)}$ for different scenarios. (D) Bias on non-causal estimates of $h^2_{ge(t\prime)}$ for different scenarios. (E) Bias on causal estimates of $h^2_{ge(t\prime)}$ for different scenarios. (F) Bias on non-causal estimates of $h^2_{ge(t\prime)}$ for different scenarios. For panels C and D, $G_{t\prime}$ is set to the total number of unique *cis*-heritable genes across all tissues. For panels E and F, $G_{t\prime}$ is set to the number of significantly *cis*-heritable genes detected in each tissue. For panels C and E, dashed lines indicate true value of $h^2_{ge(t\prime)}$. In all panels, we performed n = 1,000 independent simulated genetic architectures across different eQTL sample sizes (n = 100, 200, 300, 500, 1000, 1500); we used a one-sided z-test and the genomic block jackknife standard error to obtain p-values and data are presented as mean values +/- 1.96 x SEM.

934

Supplementary Figure 3. Robustness and power of TCSC regression in simulations when the causal tissue has fewer cis-heritable genes than tagging tissues. (A) Type I error when changing the number of expressed genes in the causal tissue. False positive event is defined as $\omega_{ge(t\prime)} > 0$ for non-causal tissues at $p < 0.05$. (B) Power to detect the causal tissue per scenario. A true positive event is defined as $\omega_{ge(t\prime)} > 0$ for causal tissues at $p < 0.05$. (C) Bias on causal estimates of $\omega_{ge(t\prime)}$ for different scenarios. (D) Bias on non-causal estimates of $\omega_{ge(t\prime)}$ for different scenarios. (E) Bias on causal estimates of $\omega_{ge(t\prime)}$ for different scenarios. (F) Bias on non-causal estimates of $\omega_{ge(t\prime)}$ for different scenarios. For panels C and D, $G_{t\prime}$ is set to the total number of unique *cis*-heritable genes across all tissues. For panels E and F, $G_{t\prime}$ is set to the number of significantly *cis*-heritable genes detected in each tissue. For panels C and E, dashed lines indicate true value of $\omega_{ge(t\prime)}$. In all panels, we performed n = 1,000 independent simulated genetic architectures across different eQTL sample sizes (n = 100, 200, 300, 500, 1000, 1500); we used a one-sided z-test and the genomic block jackknife standard error to obtain p-values and data are presented as mean values +/- 1.96 x SEM.

Supplementary Figure 4. Robustness and power of TCSC regression in simulations with different numbers of causal tissues. (A) Type I error for each different causal tissue architecture. A single causal tissue (pink) represents the primary simulation analysis. Other architectures include two causal tissues (green) and three causal tissues (blue). False positive event defined as $h^2_{ge(t\prime)} > 0$ for non-causal tissues at $p < 0.05$. (B) Power to detect the causal tissue, defined by $h^2_{ge(t\prime)} > 0$ for causal tissues at $p < 0.05$. (C) Bias on estimates of $h^2_{ge(t\prime)}$ in causal tissues for different numbers of causal tissues in the model. The dashed line indicates that the true value of $h^2_{ge(t\prime)} = 0.1$. (D) Bias on estimates of $h^2_{ge(t\prime)}$ in non-causal tissues for different numbers of causal tissues in the model. In all panels, we performed n = 1,000 independent simulated genetic architectures across different eQTL sample sizes (n = 100, 200, 300, 500, 1000, 1500); we used a one-sided z-test and the genomic block jackknife standard error to obtain p-values and data are presented as mean values +/- 1.96 x SEM. The value of $G_{t\prime}$ is set to the total number of unique *cis*-heritable genes across all tissues.

A Cross−trait TCSC (covariance)

B Cross−trait TCSC (covariance)

C Cross−trait TCSC (covariance)

D Cross−trait TCSC (covariance)

971

972 **Supplementary Figure 5. Robustness and power of cross-trait TCSC regression in simulations**

973 **with different numbers of causal tissues.** (A) Type I error for each different causal tissue

974 architecture. A single causal tissue (pink) represents the primary simulation analysis. Other

975 architectures include two causal tissues (green) and three causal tissues (blue). False positive

976 event defined as $\omega_{ge(t\prime)} > 0$ for non-causal tissues at $p < 0.05$. (B) Power to detect the causal

977 tissue, defined by $\omega_{ge(t\prime)} > 0$ for causal tissues at $p < 0.05$. (C) Bias on estimates of $\omega_{ge(t\prime)}$ in

978 causal tissues for different numbers of causal tissues in the model. The dashed line indicates

979 that the true value of $\omega_{ge(t\prime)} = 0.05$. (D) Bias on estimates of $\omega_{ge(t\prime)}$ in non-causal tissues for

980 different numbers of causal tissues in the model. In all panels, we performed n = 1,000

981 independent simulated genetic architectures across different eQTL sample sizes (n = 100, 200,

982 300, 500, 1000, 1500); we used a one-sided z-test and the genomic block jackknife standard

983 error to obtain p-values and data are presented as mean values +/- 1.96 x SEM. The value of G_t,

984 is set to the total number of unique *cis*-heritable genes across all tissues.

985

986
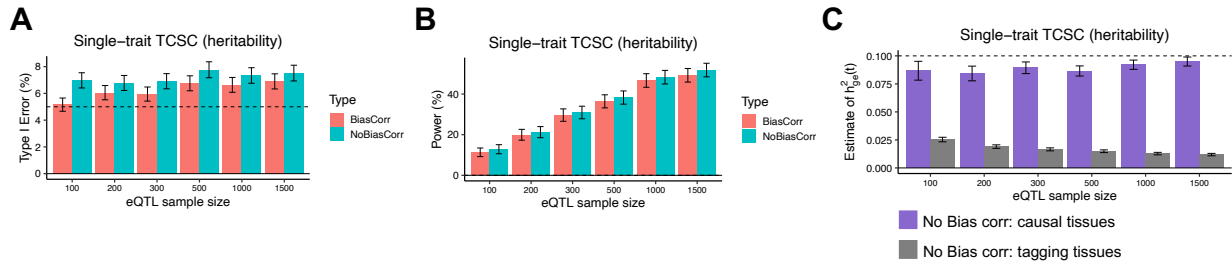
987

988 **Supplementary Figure 6. Robustness and power of TCSC regression with or without correction**
989 **for bias in tissue co-regulation scores in simulations.** (A) Type I error for each of two scenarios:
990 (1) "BiasCorr": tissue co-regulation scores estimated using bias correction as in primary
991 simulations (pink) vs (2) "NoBiasCorr": tissue co-regulation scores estimated without bias
992 correction (green). False positive event defined as $h^2_{ge(t')} > 0$ for non-causal tissues at $p < 0.05$.
993 (B) Power to detect the causal tissue, in which $h^2_{ge(t')} > 0$ for causal tissues at $p < 0.05$. (C) Bias
994 on estimates of causal and non-causal $h^2_{ge(t')}$ whose true values are 0.1 (purple bars) and 0
995 (gray bars), respectively, in the scenario of using no bias correction on tissue co-regulation
996 scores. In all panels, we performed n = 1,000 independent simulated genetic architectures
997 across different eQTL sample sizes (n = 100, 200, 300, 500, 1000, 1500); we used a one-sided z-
998 test and the genomic block jackknife standard error to obtain p-values and data are presented
999 as mean values +/- 1.96 x SEM. The value of $G_{t'}$ is set to the total number of unique *cis*-
1000 heritable genes across all tissues.

1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017

A Cross−trait TCSC (covariance)
B Cross−trait TCSC (covariance)
C Cross−trait TCSC (covariance)
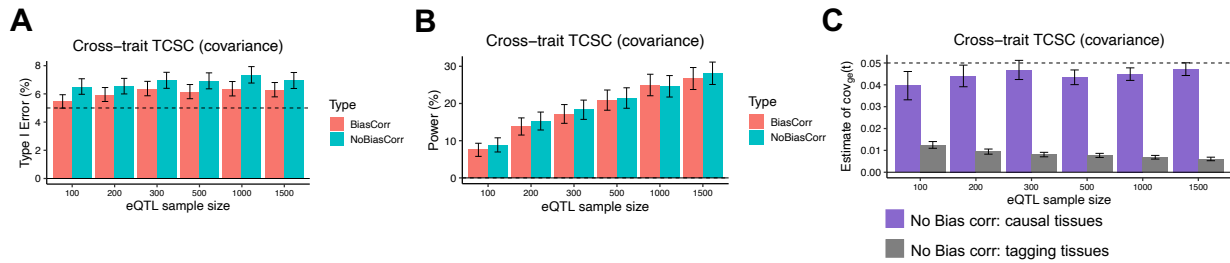
1018
1019 **Supplementary Figure 7. Robustness and power of cross-trait TCSC regression with or without**
1020 **correction for bias in tissue co-regulation scores in simulations.** (A) Type I error for each of two
1021 scenarios: (1) "BiasCorr": tissue co-regulation scores estimated using bias correction as in
1022 primary simulations (pink) vs (2) "NoBiasCorr": tissue co-regulation scores estimated without
1023 bias correction (green). False positive event defined as $\omega_{ge(t')} > 0$ for non-causal tissues at $p <$
1024 0.05. (B) Power to detect the causal tissue, in which $\omega_{ge(t')} > 0$ for causal tissues at $p < 0.05$. (C)
1025 Bias on estimates of causal and non-causal $\omega_{ge(t')}$ whose true values are 0.05 (purple bars) and
1026 0 (gray bars), respectively, in the scenario of using no bias correction on tissue co-regulation
1027 scores. In all panels, we performed n = 1,000 independent simulated genetic architectures
1028 across different eQTL sample sizes (n = 100, 200, 300, 500, 1000, 1500); we used a one-sided z-
1029 test and the genomic block jackknife standard error to obtain p-values and data are presented
1030 as mean values +/- 1.96 x SEM. The value of $G_{t'}$ is set to the total number of unique *cis*-
1031 heritable genes across all tissues.
1032
1033
1034
1035
1036
1037
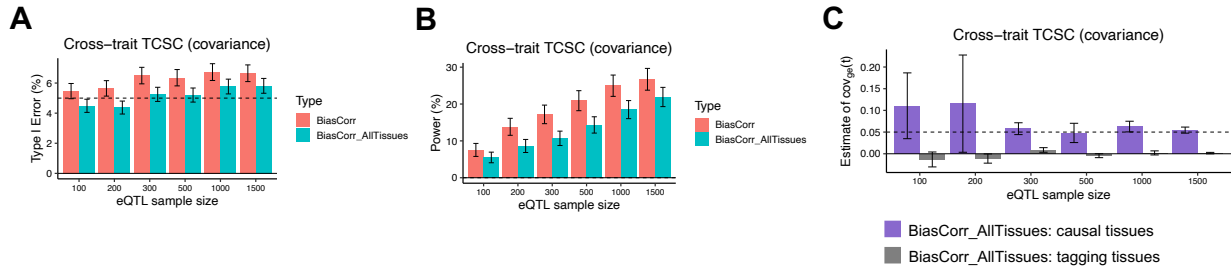1038
1039
1040
1041
1042
1043
1044
1045
1046
1047
1048
1049
1050

A Single–trait TCSC (heritability)

B Single–trait TCSC (heritability)

C Single–trait TCSC (heritability)
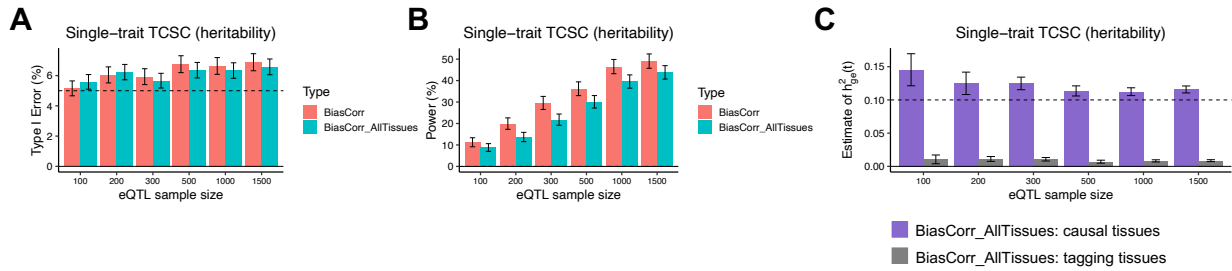
1051

Supplementary Figure 8. Robustness and power of single-trait TCSC regression without (default) or with bias correction applied to all pairs of tissues in simulations. (A) Type I error for each of two scenarios: (1) "BiasCorr": tissue co-regulation scores estimated using bias correction as in primary simulations, e.g. when t = t' (pink) vs (2) "BiasCorr_AllTissues": tissue co-regulation scores estimated using bias correction applied to all correlations of predicted gene expression (green). False positive event defined as $h^2_{ge(t\prime)} > 0$ for non-causal tissues at $p < 0.05$. (B) Power to detect the causal tissue, in which $h^2_{ge(t\prime)} > 0$ for causal tissues at $p < 0.05$. (C) Bias on estimates of causal and non-causal $h^2_{ge(t\prime)}$ whose true values are 0.1 (purple bars) and 0 (gray bars), respectively, in the scenario of using bias correction applied to all correlations of predicted gene expression in co-regulation scores. In all panels, we performed n = 1,000 independent simulated genetic architectures across different eQTL sample sizes (n = 100, 200, 300, 500, 1000, 1500); we used a one-sided z-test and the genomic block jackknife standard error to obtain p-values and data are presented as mean values +/- 1.96 x SEM. The value of $G_{t\prime}$ is set to the total number of unique *cis*-heritable genes across all tissues.

1066

1067

1068

1069

1070

1071

1072

1073

A Cross−trait TCSC (covariance)

B Cross−trait TCSC (covariance)

C Cross−trait TCSC (covariance)

BiasCorr_AllTissues: causal tissues

BiasCorr_AllTissues: tagging tissues

1074

Supplementary Figure 9. Robustness and power of cross-trait TCSC regression without (default) or with bias correction applied to all pairs of tissues in simulations. (A) Type I error for each of two scenarios: (1) "BiasCorr": tissue co-regulation scores estimated using bias correction as in primary simulations, e.g. when t = t' (pink) vs (2) "BiasCorr_AllTissues": tissue co-regulation scores estimated using bias correction applied to all correlations of predicted gene expression (green). False positive event defined as $\omega_{ge(t\prime)} > 0$ for non-causal tissues at $p < 0.05$. (B) Power to detect the causal tissue, in which $\omega_{ge(t\prime)} > 0$ for causal tissues at $p < 0.05$. (C) Bias on estimates of causal and non-causal $\omega_{ge(t\prime)}$ whose true values are 0.05 (purple bars) and 0 (gray bars), respectively, in the scenario of using bias correction applied to all correlations of predicted gene expression in co-regulation scores. In all panels, we performed n = 1,000 independent simulated genetic architectures across different eQTL sample sizes (n = 100, 200, 300, 500, 1000, 1500); we used a one-sided z-test and the genomic block jackknife standard error to obtain p-values and data are presented as mean values +/- 1.96 x SEM. The value of $G_{t\prime}$ is set to the total number of unique *cis*-heritable genes across all tissues.

1089
1090
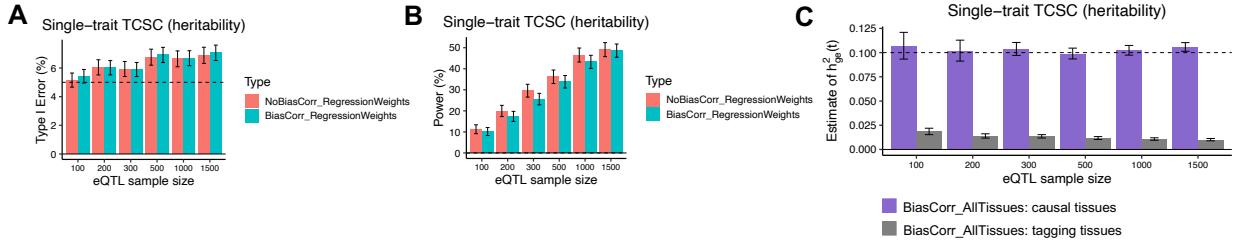1091
1092
1093
1094
1095
1096
1097
1098
1099
1100
1101
1102
1103
1104

Supplementary Figure 10. Robustness and power of TCSC regression without (default) or with bias correction for tissue co-regulation scores used to calculate regression weights in simulations. (A) Type I error for each of two scenarios: (1) "NoBiasCorr_RegressionWeights": regression weights calculated using uncorrected tissue co-regulation scores as in primary simulations (pink) vs (2) "BiasCorr_RegressionWeights": regression weights calculated using bias-corrected tissue co-regulation scores (green). False positive event defined as $h^2_{ge(t')} > 0$ for non-causal tissues at $p < 0.05$. (B) Power to detect the causal tissue, in which $h^2_{ge(t')} > 0$ for causal tissues at $p < 0.05$. (C) Bias on estimates of causal and non-causal $h^2_{ge(t')}$ whose true values are 0.1 (purple bars) and 0 (gray bars), respectively, in the scenario of calculating regression weights using bias-corrected tissue co-regulation scores. In all panels, we performed n = 1,000 independent simulated genetic architectures across different eQTL sample sizes (n = 100, 200, 300, 500, 1000, 1500); we used a one-sided z-test and the genomic block jackknife standard error to obtain p-values and data are presented as mean values +/- 1.96 x SEM. The value of $G_{t'}$ is set to the total number of unique *cis*-heritable genes across all tissues.

A Cross–trait TCSC (covariance)

B Cross–trait TCSC (covariance)

C Cross–trait TCSC (covariance)
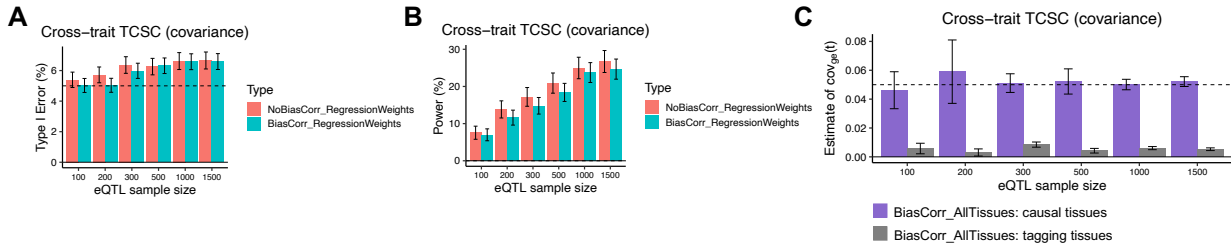
1136

Supplementary Figure 11. Robustness and power of cross-trait TCSC regression with or without correction for bias in tissue co-regulation scores in simulations. (A) Type I error for each of two scenarios: (1) "NoBiasCorr_RegressionWeights": regression weights calculated using uncorrected tissue co-regulation scores as in primary simulations (pink) vs (2) "BiasCorr_RegressionWeights": regression weights calculated using bias-corrected tissue co-regulation scores (green). False positive event defined as $\omega_{ge(t\prime)} > 0$ for non-causal tissues at $p < 0.05$. (B) Power to detect the causal tissue, in which $\omega_{ge(t\prime)} > 0$ for causal tissues at $p < 0.05$. (C) Bias on estimates of causal and non-causal $\omega_{ge(t\prime)}$ whose true values are 0.05 (purple bars) and 0 (gray bars), respectively, in the scenario of calculating regression weights using bias-corrected tissue co-regulation scores. In all panels, we performed n = 1,000 independent simulated genetic architectures across different eQTL sample sizes (n = 100, 200, 300, 500, 1000, 1500); we used a one-sided z-test and the genomic block jackknife standard error to obtain p-values and data are presented as mean values +/- 1.96 x SEM. The value of $G_{t\prime}$ is set to the total number of unique *cis*-heritable genes across all tissues.

1151
1152
1153
1154
1155
1156
1157
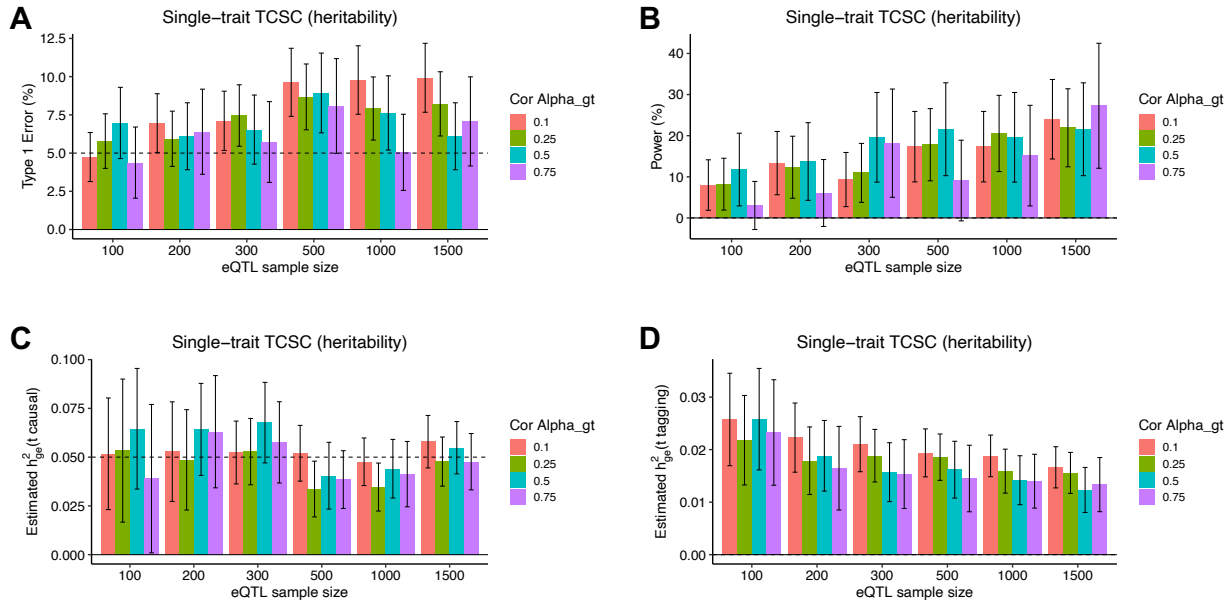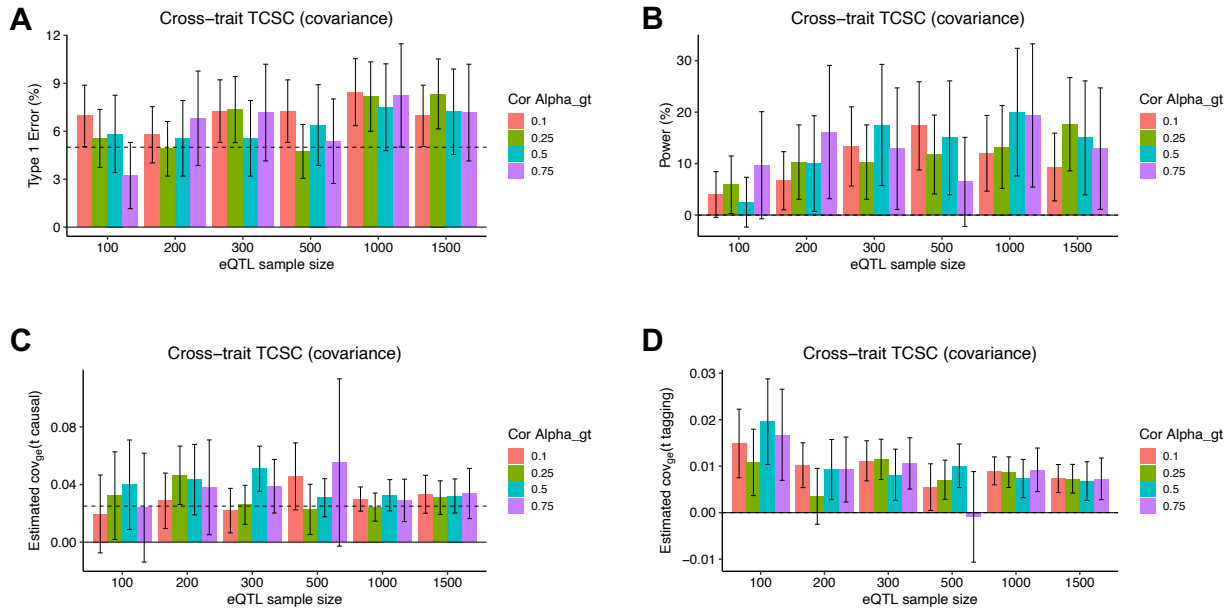1158
1159
1160

A Single−trait TCSC (heritability)

B Single−trait TCSC (heritability)

C Single−trait TCSC (heritability)

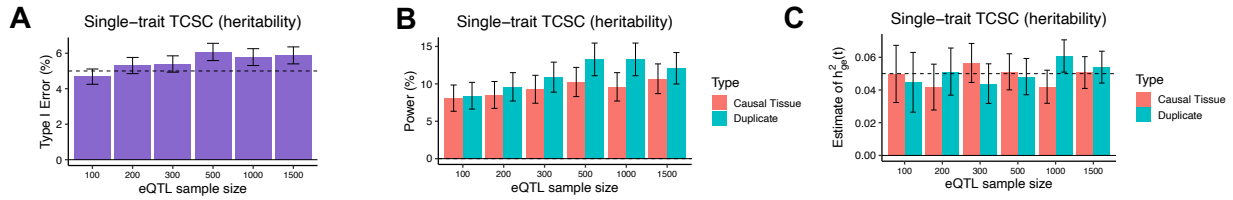**D** Single−trait TCSC (heritability)

1161
1162
Supplementary Figure 12. Robustness of cross-trait TCSC regression in simulations with two causal tissues with varying levels of correlated gene expression-trait effects. We demonstrate how TCSC would behave when there are two causal tissues with correlation gene-trait effects (α_{gt}), each tissue of which contributes 5% heritability to the trait. This is a model violation where TCSC assumes that gene expression-trait effects are i.i.d. (A) Type I error while varying the correlation between the α_{gt} of each causal tissue. False positive event defined as $h^2_{ge(t')} > 0$ for non-causal tissues at $p < 0.05$. (B) Power to detect the causal tissues in which $h^2_{ge(t')} > 0$ for causal tissues at $p < 0.05$. (C) Bias on estimates of $h^2_{ge(t')}$ for the causal tissue, while varying the correlation between the α_{gt} of each causal tissue. The dashed line indicates that the true value of $h^2_{ge(t')}$ for either causal tissue. (D) Bias on estimates of $h^2_{ge(t')}$ for non-causal tissues, while varying the correlation between the α_{gt} of each causal tissue. In all panels, we performed n = 1,000 independent simulated genetic architectures across different eQTL sample sizes (n = 100, 200, 300, 500, 1000, 1500); we used a one-sided z-test and the genomic block jackknife standard error to obtain p-values and data are presented as mean values +/- 1.96 x SEM. The value of $G_{t'}$ is set to the total number of unique *cis*-heritable genes across all tissues.
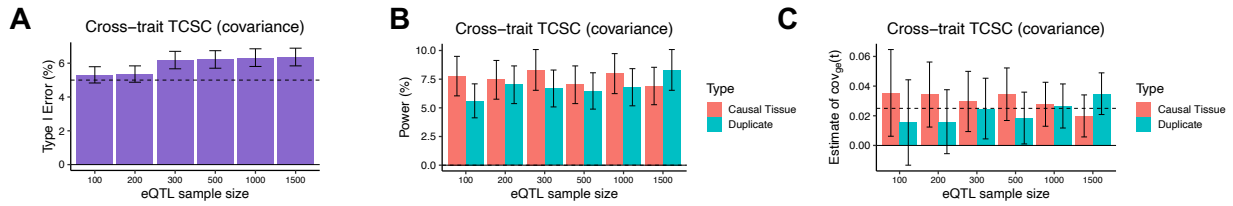1178

Supplementary Figure 13. Robustness of cross-trait TCSC regression in simulations with two causal tissues with varying levels of correlated gene expression-trait effects. We demonstrate how TCSC would behave when there are two causal tissues with correlation gene-trait effects (α_{gt}), each tissue of which contributes 5% heritability to the trait. This is a model violation where TCSC assumes that gene expression-trait effects are i.i.d. (A) Type I error while varying the correlation between the α_{gt} of each causal tissue. False positive event defined as $\omega_{ge(t')} > 0$ for non-causal tissues at $p < 0.05$. (B) Power to detect the causal tissues in which $\omega_{ge(t')} > 0$ for causal tissues at $p < 0.05$. (C) Bias on estimates of $\omega_{ge(t')}$ for the causal tissue, while varying the correlation between the α_{gt} of each causal tissue. The dashed line indicates that the true value of $\omega_{ge(t')}$ for either causal tissue. (D) Bias on estimates of $\omega_{ge(t')}$ for non-causal tissues, while varying the correlation between the α_{gt} of each causal tissue. In all panels, we performed n = 1,000 independent simulated genetic architectures across different eQTL sample sizes (n = 100, 200, 300, 500, 1000, 1500); we used a one-sided z-test and the genomic block jackknife standard error to obtain p-values and data are presented as mean values +/- 1.96 x SEM. The value of $G_{t'}$ is set to the total number of unique *cis*-heritable genes across all tissues.

A Single–trait TCSC (heritability)

B Single–trait TCSC (heritability)

**C** Single–trait TCSC (heritability)

1207

1208 **Supplementary Figure 14. Robustness of single-trait TCSC regression in simulations with two**
1209 **causal tissues with identical gene expression-trait effects.** We demonstrate how TCSC would
1210 behave when there are two identical tissues contributing the same genetic component of gene
1211 expression to the trait. This is a model violation where TCSC assumes that gene expression-trait
1212 effects are i.i.d. To the original and duplicated causal tissue, we added a small amount of noise
1213 (with mean 0, variance 0.0025) to the tissue co-regulation scores of each duplicated tissue to
1214 avoid collinearity in the multiple linear regression. (A) Type I error; false positive event defined
1215 as $h^2_{ge(t\prime)} > 0$ for non-causal tissues at $p < 0.05$. (B) Power to detect each of two causal tissues,
1216 e.g. $h^2_{ge(t\prime)} > 0$ for the causal tissue at $p < 0.05$. (C) TCSC estimates similar values of $h^2_{ge(t\prime)}$ to
1217 each causal tissue, approximately one-half the value of the trait variance explained by the
1218 original causal tissue (0.1). Dashed line at 0.05, the expected tissue-specific contribution to
1219 heritability for both original and duplicated tissue. In all panels, we performed n = 1,000
1220 independent simulated genetic architectures across different eQTL sample sizes (n = 100, 200,
1221 300, 500, 1000, 1500); we used a one-sided z-test and the genomic block jackknife standard
1222 error to obtain p-values and data are presented as mean values +/- 1.96 x SEM. The value of $G_{t\prime}$
1223 is set to the total number of unique *cis*-heritable genes across all tissues.

1224

1225

1226

1227

1228

1229

1230

1231

1232

1233

1234

1235

1236

1237

1238

1239

1240

1241

A Cross–trait TCSC (covariance)

B Cross–trait TCSC (covariance)

C Cross–trait TCSC (covariance)

Supplementary Figure 15. Robustness of cross-trait TCSC regression in simulations with two causal tissues with identical gene expression-trait effects. We demonstrate how TCSC would behave when there are two identical tissues contributing the same genetic component of gene expression to the trait. This is a model violation where TCSC assumes that gene expression-trait effects are i.i.d. To the original and duplicated causal tissue, we added a small amount of noise (with mean 0, variance 0.0025) to the tissue co-regulation scores of each duplicated tissue to avoid collinearity in the multiple linear regression. (A) Type I error; false positive event defined as $\omega_{ge(t')} > 0$ for non-causal tissues at $p < 0.05$. (B) Power to detect each of two causal tissues, e.g. $\omega_{ge(t')} > 0$ for the causal tissue at $p < 0.05$. (C) TCSC estimates similar values of $\omega_{ge(t')}$ to each causal tissue, approximately one-half the value of the trait covariance explained by the original causal tissue (0.05). Dashed line at 0.025, the expected tissue-specific contribution to covariance for both original and duplicated tissue. In all panels, we performed n = 1,000 independent simulated genetic architectures across different eQTL sample sizes (n = 100, 200, 300, 500, 1000, 1500); we used a one-sided z-test and the genomic block jackknife standard error to obtain p-values and data are presented as mean values +/- 1.96 x SEM. The value of $G_{t'}$ is set to the total number of unique *cis*-heritable genes across all tissues.
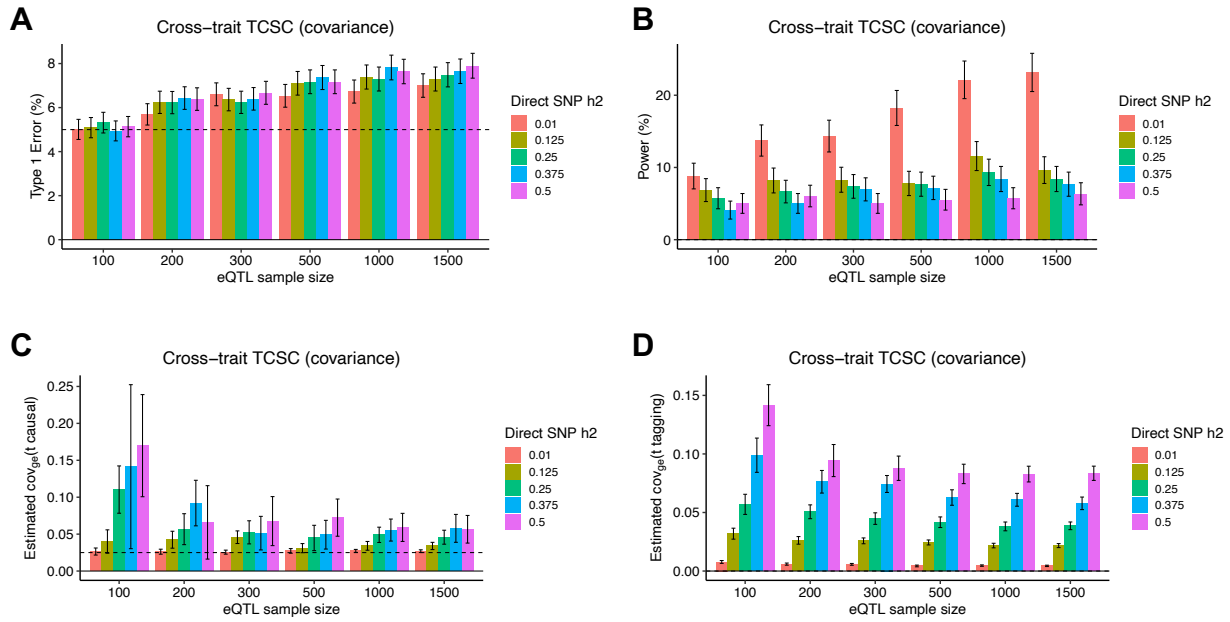
1272
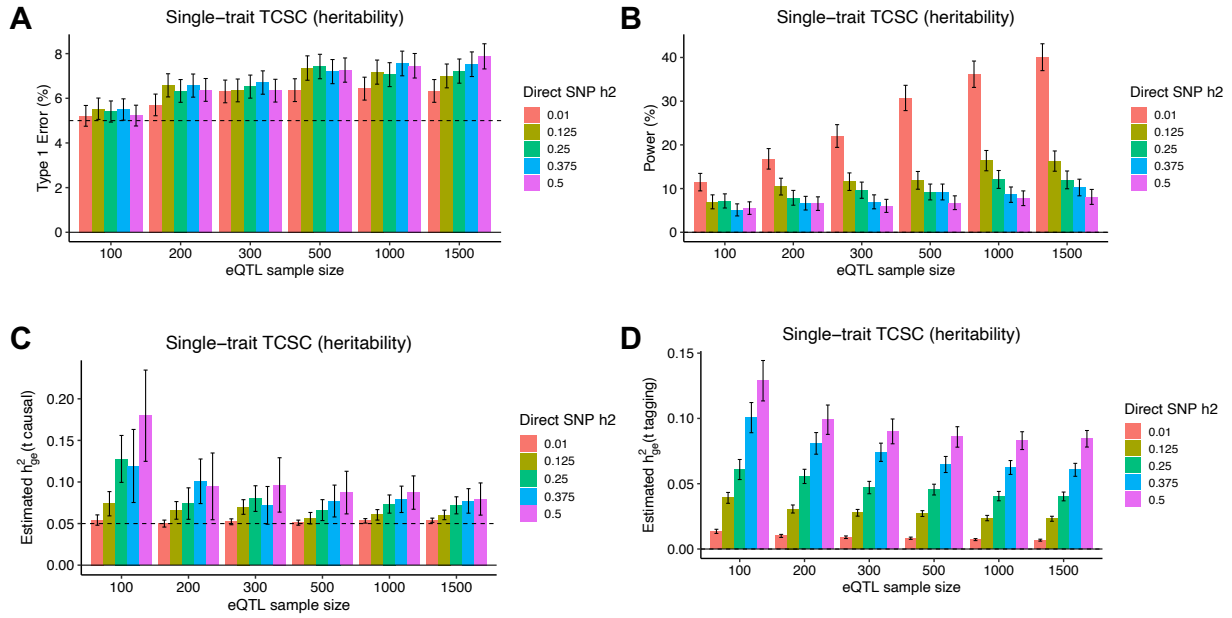
Supplementary Figure 16. **Robustness and power of TCSC regression in simulations with different amounts of direct SNP-trait heritability (h^2_{SNP}) not mediated by gene expression.** (A) Type I error per value of h^2_{SNP}. False positive event is defined as $h^2_{ge(t\prime)} > 0$ for non-causal tissues at $p < 0.05$. (B) Power to detect the causal tissue per value of h^2_{SNP}. A true positive event is defined as $h^2_{ge(t\prime)} > 0$ for causal tissues at $p < 0.05$. (C) Bias on causal estimates of $h^2_{ge(t\prime)}$ for different values of h^2_{SNP}. Dashed line indicates true value of $h^2_{ge(t\prime)}$. (D) Bias on non-causal estimates of $h^2_{ge(t\prime)}$ for different values of h^2_{SNP}. In all panels, we performed n = 1,000 independent simulated genetic architectures across different eQTL sample sizes (n = 100, 200, 300, 500, 1000, 1500); we used a one-sided z-test and the genomic block jackknife standard error to obtain p-values and data are presented as mean values +/- 1.96 x SEM. The value of $G_{t\prime}$ is set to the total number of unique *cis*-heritable genes across all tissues.

A Cross−trait TCSC (covariance)

B Cross−trait TCSC (covariance)

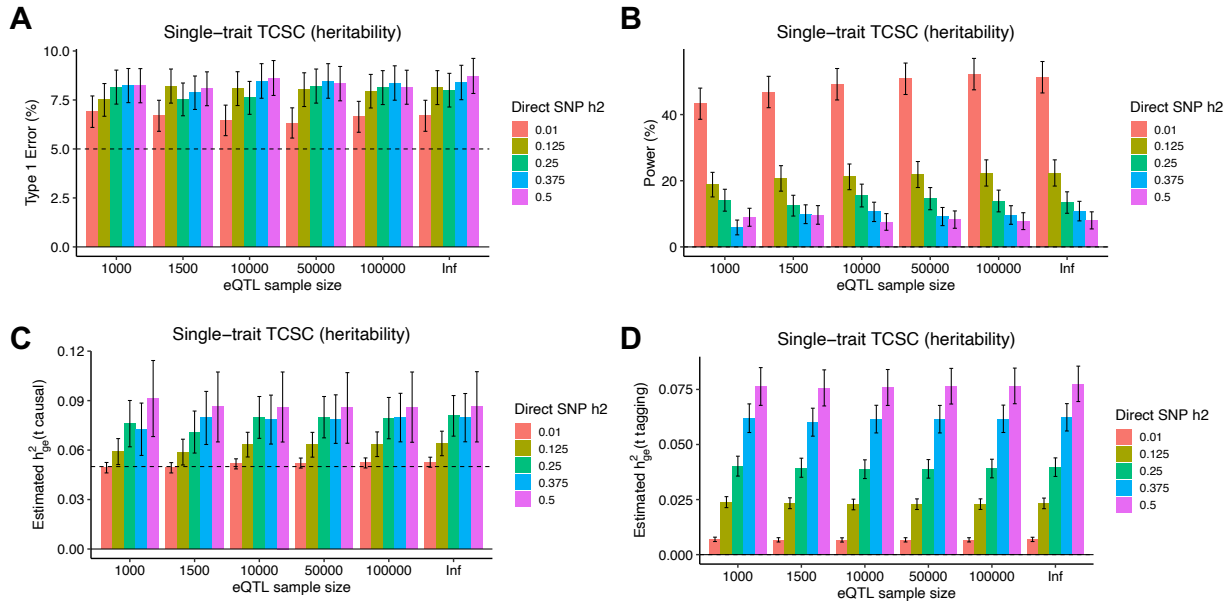**C** Cross−trait TCSC (covariance)

D Cross−trait TCSC (covariance)

1286

Supplementary Figure 17. Robustness and power of cross-trait TCSC regression in simulations with different amounts of direct SNP-trait heritability (h^2_{SNP}) not mediated by gene expression. (A) Type I error per value of h^2_{SNP}. False positive event is defined as $\omega_{ge(t\prime)} > 0$ for non-causal tissues at $p < 0.05$. (B) Power to detect the causal tissue per value of h^2_{SNP}. A true positive event is defined as $\omega_{ge(t\prime)} > 0$ for causal tissues at $p < 0.05$. (C) Bias on causal estimates of $\omega_{ge(t\prime)}$ for different values of h^2_{SNP}. Dashed line indicates true value of $\omega_{ge(t\prime)}$. (D) Bias on non-causal estimates of $\omega_{ge(t\prime)}$ for different values of h^2_{SNP}. In all panels, we performed n = 1,000 independent simulated genetic architectures across different eQTL sample sizes (n = 100, 200, 300, 500, 1000, 1500); we used a one-sided z-test and the genomic block jackknife standard error to obtain p-values and data are presented as mean values +/- 1.96 x SEM. The value of $G_{t\prime}$ is set to the total number of unique *cis*-heritable genes across all tissues.

1287
1288
1289
1290
1291
1292
1293
1294
1295
1296
1297
1298
1299

A Single–trait TCSC (heritability)

B Single–trait TCSC (heritability)

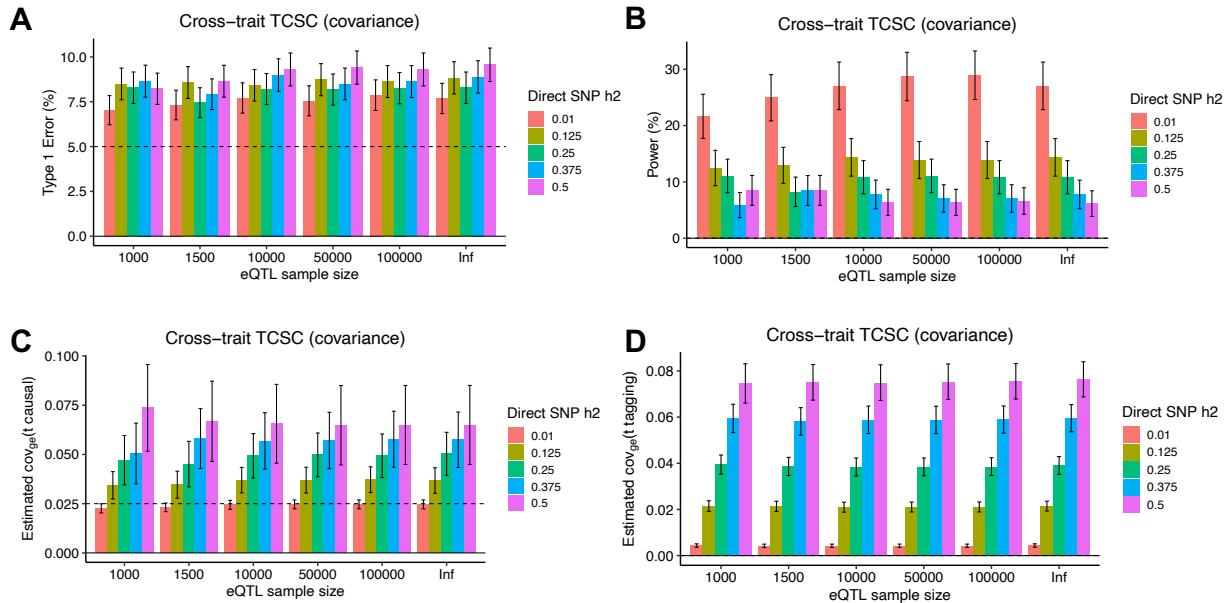**C** Single–trait TCSC (heritability)

D Single–trait TCSC (heritability)

1300

Supplementary Figure 18. **Robustness and power of TCSC regression in simulations with different amounts of direct SNP-trait heritability (h^2_{SNP}) not mediated by gene expression at large eQTL sample size.** (A) Type I error per value of h^2_{SNP}. False positive event is defined as $h^2_{ge(t')} > 0$ for non-causal tissues at $p < 0.05$. (B) Power to detect the causal tissue per value of h^2_{SNP}. A true positive event is defined as $h^2_{ge(t')} > 0$ for causal tissues at $p < 0.05$. (C) Bias on causal estimates of $h^2_{ge(t')}$ for different values of h^2_{SNP}. Dashed line indicates true value of $h^2_{ge(t')}$. (D) Bias on non-causal estimates of $h^2_{ge(t')}$ for different values of h^2_{SNP}. In all panels, we performed n = 1,000 independent simulated genetic architectures across different eQTL sample sizes (n = 1000, 1500, 10K, 50K, 100K, Infinite (true eQTL effects)); we used a one-sided z-test and the genomic block jackknife standard error to obtain p-values and data are presented as mean values +/- 1.96 x SEM. The value of $G_{t'}$ is set to the total number of unique *cis*-heritable genes across all tissues. For these simulations, significantly *cis*-heritable genes were determined using cross-validation adjusted-$R^2 > 0$ rather than GCTA $p < 0.01$, due to the computationally intensive nature of running GCTA on hundreds of thousands of samples across thousands of simulations.

1316

1317

Supplementary Figure 19. Robustness and power of cross-trait TCSC regression in simulations with different amounts of direct SNP-trait heritability (h^2_{SNP}) not mediated by gene expression at large eQTL sample size. (A) Type I error per value of h^2_{SNP}. False positive event is defined as $\omega_{ge(t\prime)} > 0$ for non-causal tissues at $p < 0.05$. (B) Power to detect the causal tissue per value of h^2_{SNP}. A true positive event is defined as $\om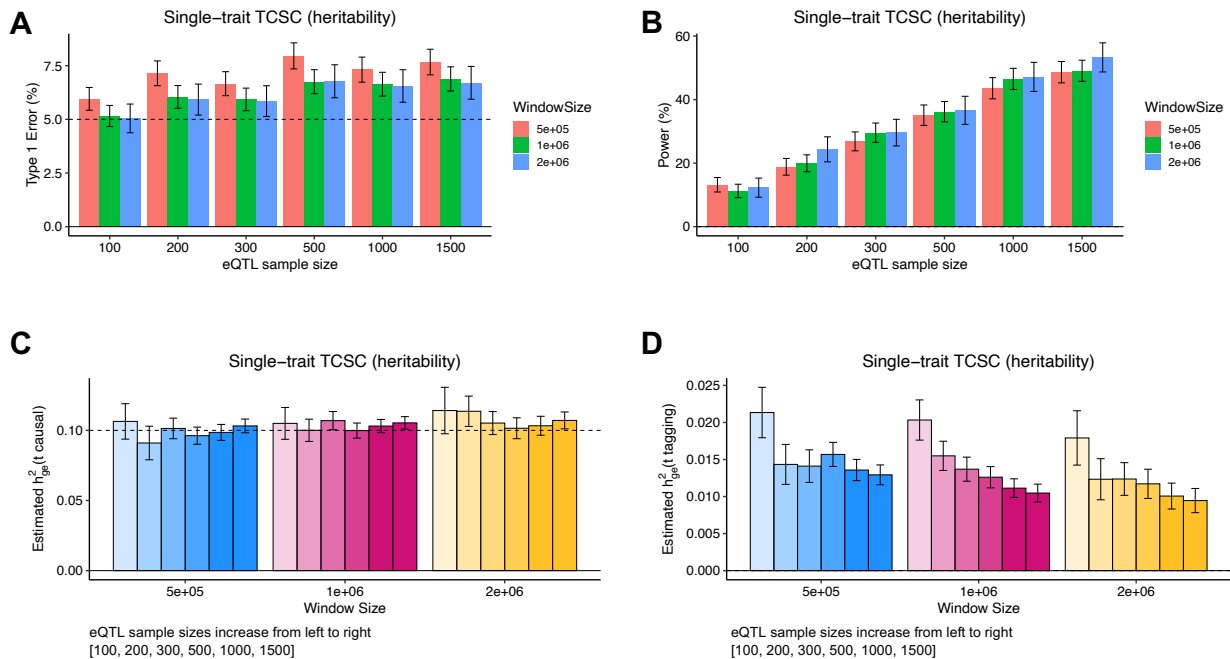ega_{ge(t\prime)} > 0$ for causal tissues at $p < 0.05$. (C) Bias on causal estimates of $\omega_{ge(t\prime)}$ for different values of h^2_{SNP}. Dashed line indicates true value of $\omega_{ge(t\prime)}$. (D) Bias on non-causal estimates of $\omega_{ge(t\prime)}$ for different values of h^2_{SNP}. In all panels, we performed n = 1,000 independent simulated genetic architectures across different eQTL sample sizes (n = 1000, 1500, 10K, 50K, 100K, Infinite (true eQTL effects)); we used a one-sided z-test and the genomic block jackknife standard error to obtain p-values and data are presented as mean values +/- 1.96 x SEM. The value of $G_{t\prime}$ is set to the total number of unique *cis*-heritable genes across all tissues. For these simulations, significantly *cis*-heritable genes were determined using cross-validation adjusted-$R^2 > 0$ rather than GCTA $p < 0.01$, due to the computationally intensive nature of running GCTA on hundreds of thousands of samples across thousands of simulations.

1333
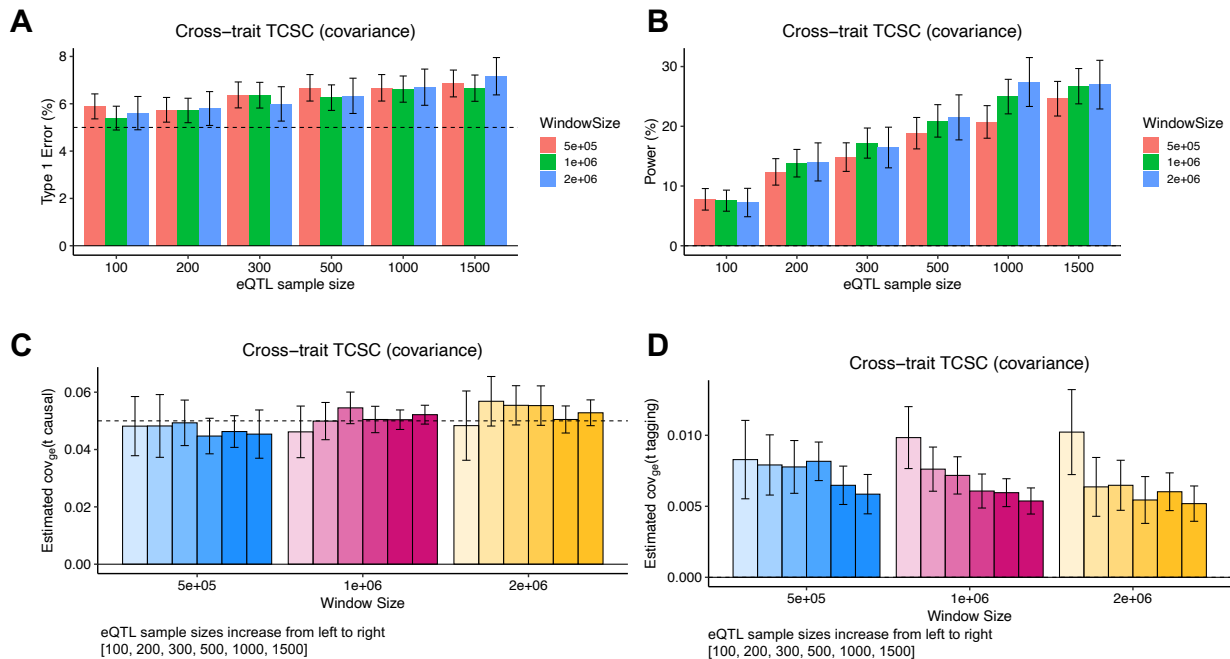
1334

1335

1336

Supplementary Figure 20. Robustness and power of TCSC regression in simulations with different values of window size used to calculate gene-gene co-regulation. (A) Type I error per window size. False positive event is defined as $h^2_{ge(t')} > 0$ for non-causal tissues at $p < 0.05$. (B) Power to detect the causal tissue per window size. A true positive event is defined as $h^2_{ge(t')} > 0$ for causal tissues at $p < 0.05$. (C) Bias on causal estimates of $h^2_{ge(t')}$ for different window sizes. Dashed lines indicate true values of $h^2_{ge(t')}$. (D) Bias on non-causal estimates of $h^2_{ge(t')}$ for different window sizes. In all panels, we performed n = 1,000 independent simulated genetic architectures across different eQTL sample sizes (n = 100, 200, 300, 500, 1000, 1500); we used a one-sided z-test and the genomic block jackknife standard error to obtain p-values and data are presented as mean values +/- 1.96 x SEM. The value of $G_{t'}$ is set to the total number of unique *cis*-heritable genes across all tissues.

1348
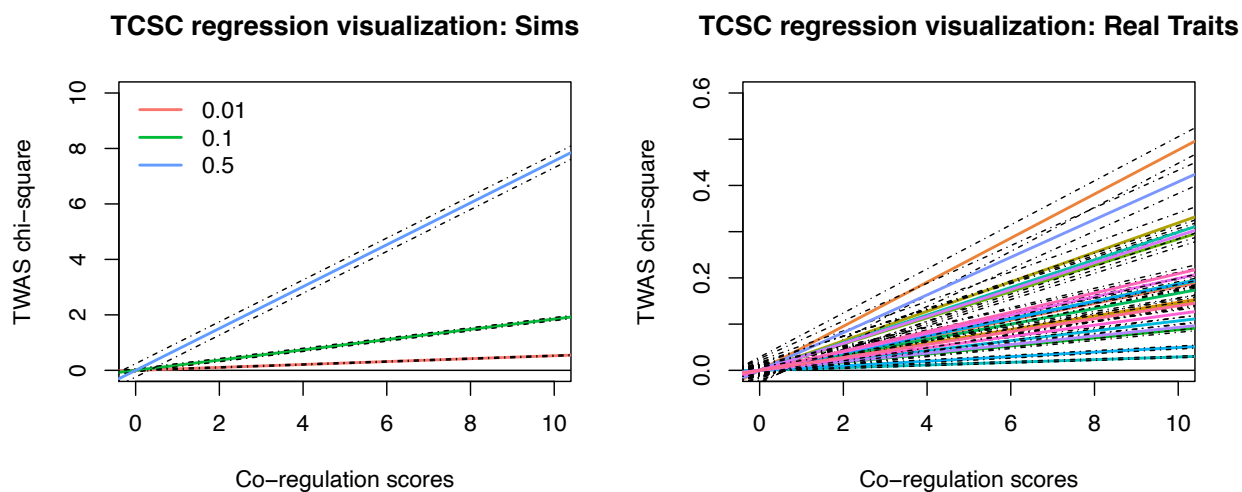
1349

Supplementary Figure 21. Robustness and power of cross-trait TCSC regression in simulations with different values of window size used to calculate gene-gene co-regulation. (A) Type I error per window size. False positive event is defined as $\omega_{ge(t\prime)} > 0$ for non-causal tissues at $p < 0.05$. (B) Power to detect the causal tissue per window size. A true positive event is defined as $\omega_{ge(t\prime)} > 0$ for causal tissues at $p < 0.05$. (C) Bias on causal estimates of $\omega_{ge(t\prime)}$ for different window sizes. Dashed lines indicate true values of $\omega_{ge(t\prime)}$. (D) Bias on non-causal estimates of $\omega_{ge(t\prime)}$ for different window sizes. In all panels, we performed n = 1,000 independent simulated genetic architectures across different eQTL sample sizes (n = 100, 200, 300, 500, 1000, 1500); we used a one-sided z-test and the genomic block jackknife standard error to obtain p-values and data are presented as mean values +/- 1.96 x SEM. The value of $G_{t\prime}$ is set to the total number of unique *cis*-heritable genes across all tissues.

1361
1362
1363

TCSC regression visualization: Sims

TCSC regression visualization: Real Traits

1364

1365 **Supplementary Figure 22. Visualization of TCSC regression.** (A) In simulations, we visualize the

1366 single-trait TCSC estimand ($h^2_{ge(t\prime)}$) as the line of best fit (slope) for one representative

1367 simulation from each of three true values of $h^2_{ge(t\prime)}$ using an intercept of 0. (B) In analysis of real

1368 traits, we visualize the single-trait TCSC estimand ($h^2_{ge(t\prime)}$) as the line of best fit (slope) for each

1369 of 21 significant tissue-trait pairs using an intercept of 0. In panels A and B, solid lines represent

1370 the TCSC estimate; dashed lines represent the estimate +/- 1.96 x the jackknife standard error.

1371 All slopes shown are significantly greater than zero at 5% FDR.

1372

1373

Supplementary references

1.   Ongen, H. *et al.* Estimating the causal tissues for complex traits and diseases. *Nat Genet* **49**, 1676-1683 (2017).
2.   Finucane, H.K. *et al.* Heritability enrichment of specifically expressed genes identifies disease-relevant tissues and cell types. *Nat Genet* **50**, 621-629 (2018).
3.   Gazal, S. *et al.* Combining SNP-to-gene linking strategies to identify disease genes and assess disease omnigenicity. *Nat Genet* **54**, 827-836 (2022).
4.   Liu, X. *et al.* Functional Architectures of Local and Distal Regulation of Gene Expression in Multiple Human Tissues. *Am J Hum Genet* **100**, 605-616 (2017).
5.   Consortium, G.T. The GTEx Consortium atlas of genetic regulatory effects across human tissues. *Science* **369**, 1318-1330 (2020).
6.   Yazar, S. *et al.* Single-cell eQTL mapping identifies cell type-specific genetic control of autoimmune disease. *Science* **376**, eabf3041 (2022).
7.   Li, Z. *et al.* METRO: Multi-ancestry transcriptome-wide association studies for powerful gene-trait association detection. *Am J Hum Genet* **109**, 783-801 (2022).
8.   Siewert-Rocks, K.M., Kim, S.S., Yao, D.W., Shi, H. & Price, A.L. Leveraging gene co-regulation to identify gene sets enriched for disease heritability. *Am J Hum Genet* **109**, 393-404 (2022).
9.   Yang, J., Lee, S.H., Goddard, M.E. & Visscher, P.M. GCTA: a tool for genome-wide complex trait analysis. *Am J Hum Genet* **88**, 76-82 (2011).
10.   Gusev, A. *et al.* Integrative approaches for large-scale transcriptome-wide association studies. *Nat Genet* **48**, 245-52 (2016).
11.   Calderon, D. *et al.* Inferring Relevant Cell Types for Complex Traits by Using Single-Cell Gene Expression. *Am J Hum Genet* **101**, 686-699 (2017).
12.   Shang, L., Smith, J.A. & Zhou, X. Leveraging gene co-expression patterns to infer trait-relevant tissues in genome-wide association studies. *PLoS Genet* **16**, e1008734 (2020).
13.   Amariuta, T. TCSC GitHub Repository. (2023).
14.   Frederiksen, L. *et al.* Subcutaneous rather than visceral adipose tissue is associated with adiponectin levels and insulin resistance in young men. *J Clin Endocrinol Metab* **94**, 4010-5 (2009).
15.   McGillicuddy, F.C., Reilly, M.P. & Rader, D.J. Adipose modulation of high-density lipoprotein cholesterol: implications for obesity, high-density lipoprotein metabolism, and cardiovascular disease. *Circulation* **124**, 1602-5 (2011).
16.   Hung, K.T., Berisha, S.Z., Ritchey, B.M., Santore, J. & Smith, J.D. Red blood cells play a role in reverse cholesterol transport. *Arterioscler Thromb Vasc Biol* **32**, 1460-5 (2012).
17.   Finucane, H.K. *et al.* Partitioning heritability by functional annotation using genome-wide association summary statistics. *Nat Genet* **47**, 1228-35 (2015).
18.   Timshel, P.N., Thompson, J.J. & Pers, T.H. Genetic mapping of etiologic brain cell types for obesity. *Elife* **9**(2020).
19.   Zhu, J.N. & Wang, J.J. The cerebellum in feeding control: possible function and mechanism. *Cell Mol Neurobiol* **28**, 469-78 (2008).
20.   Onat, F. & Cavdar, S. Cerebellar connections: hypothalamus. *Cerebellum* **2**, 263-9 (2003).

1417   21.   Low, A.Y.T. *et al.* Reverse-translational identification of a cerebellar satiation network.
1418   *Nature* **600**, 269-273 (2021).
1419   22.   Bruning, J.C. *et al.* Role of brain insulin receptor in control of body weight and
1420   reproduction. *Science* **289**, 2122-5 (2000).
1421   23.   Fontana, R. & Della Torre, S. The Deep Correlation between Energy Metabolism and
1422   Reproduction: A View on the Effects of Nutrition for Women Fertility. *Nutrients* **8**, 87 (2016).
1423   24.   Kendall, R.T. & Feghali-Bostwick, C.A. Fibroblasts in fibrosis: novel roles and mediators.
1424   *Front Pharmacol* **5**, 123 (2014).
1425   25.   Felger, J.C. & Lotrich, F.E. Inflammatory cytokines in depression: neurobiological
1426   mechanisms and therapeutic implications. *Neuroscience* **246**, 199-229 (2013).
1427   26.   Soler Artigas, M. *et al.* Sixteen new lung function signals identified through 1000
1428   Genomes Project reference panel imputation. *Nat Commun* **6**, 8658 (2015).
1429   27.   Chung, K.F. The role of airway smooth muscle in the pathogenesis of airway wall
1430   remodeling in chronic obstructive pulmonary disease. *Proc Am Thorac Soc* **2**, 347-54; discussion
1431   371-2 (2005).
1432   28.   Farooqi, I.S. Defining the neural basis of appetite and obesity: from genes to behaviour.
1433   *Clin Med (Lond)* **14**, 286-9 (2014).
1434   29.   Locke, A.E. *et al.* Genetic studies of body mass index yield new insights for obesity
1435   biology. *Nature* **518**, 197-206 (2015).
1436   30.   Medic, N. *et al.* Increased body mass index is associated with specific regional
1437   alterations in brain structure. *Int J Obes (Lond)* **40**, 1177-82 (2016).
1438   31.   Yengo, L. *et al.* Meta-analysis of genome-wide association studies for height and body
1439   mass index in approximately 700000 individuals of European ancestry. *Hum Mol Genet* **27**,
1440   3641-3649 (2018).
1441   32.   Kniewallner, K.M., Ehrlich, D., Kiefer, A., Marksteiner, J. & Humpel, C. Platelets in the
1442   Alzheimer's disease brain: do they play a role in cerebral amyloid angiopathy? *Curr Neurovasc*
1443   *Res* **12**, 4-14 (2015).
1444   33.   Kucheryavykh, L.Y. *et al.* Platelets are responsible for the accumulation of beta-amyloid
1445   in blood clots inside and around blood vessels in mouse brain after thrombosis. *Brain Res Bull*
1446   **128**, 98-105 (2017).
1447   34.   Leiter, O. & Walker, T.L. Platelets: The missing link between the blood and brain? *Prog*
1448   *Neurobiol* **183**, 101695 (2019).
1449   35.   Aminian, A. *et al.* Significance of platelet count in esophageal carcinomas. *Saudi J*
1450   *Gastroenterol* **17**, 134-7 (2011).
1451   36.   Burkhardt, M.A. *et al.* Synergistic interactions of blood-borne immune cells, fibroblasts
1452   and extracellular matrix drive repair in an in vitro peri-implant wound healing model. *Sci Rep* **6**,
1453   21071 (2016).
1454   37.   Grant, S.G., Melan, M.A., Latimer, J.J. & Witt-Enderby, P.A. Melatonin and breast cancer:
1455   cellular mechanisms, clinical studies and future perspectives. *Expert Rev Mol Med* **11**, e5 (2009).
1456   38.   Coffey, R.J., Jr. *et al.* Production and auto-induction of transforming growth factor-alpha
1457   in human keratinocytes. *Nature* **328**, 817-20 (1987).
1458   39.   Bartell, E. *et al.* Protein QTL analysis of IGF-I and its binding proteins provides insights
1459   into growth biology. *Hum Mol Genet* **29**, 2625-2636 (2020).

1460   40.   Fredriksson, K. *et al.* Red blood cells stimulate human lung fibroblasts to secrete
1461   interleukin-8. *Inflammation* **27**, 71-8 (2003).
1462   41.   Fredriksson, K. *et al.* Red blood cells increase secretion of matrix metalloproteinases
1463   from human lung fibroblasts in vitro. *Am J Physiol Lung Cell Mol Physiol* **290**, L326-33 (2006).
1464   42.   Gomes, I. *et al.* Eosinophil-fibroblast interactions induce fibroblast IL-6 secretion and
1465   extracellular matrix gene expression: implications in fibrogenesis. *J Allergy Clin Immunol* **116**,
1466   796-804 (2005).
1467   43.   Schroder, T., Fuchss, J., Schneider, I., Stoltenburg-Didinger, G. & Hanisch, F. Eosinophils
1468   in hereditary and inflammatory myopathies. *Acta Myol* **32**, 148-53 (2013).
1469   44.   Herbst, K.L. & Bhasin, S. Testosterone action on skeletal muscle. *Curr Opin Clin Nutr*
1470   *Metab Care* **7**, 271-7 (2004).
1471   45.   Atkinson, R.A. *et al.* Effects of testosterone on skeletal muscle architecture in
1472   intermediate-frail and frail elderly men. *J Gerontol A Biol Sci Med Sci* **65**, 1215-9 (2010).
1473   46.   Plessen, K.J. *et al.* Hippocampus and amygdala morphology in attention-
1474   deficit/hyperactivity disorder. *Arch Gen Psychiatry* **63**, 795-807 (2006).
1475   47.   Demontis, D. *et al.* Discovery of the first genome-wide significant risk loci for attention
1476   deficit/hyperactivity disorder. *Nat Genet* **51**, 63-75 (2019).
1477   48.   Enard, W. *et al.* A humanized version of Foxp2 affects cortico-basal ganglia circuits in
1478   mice. *Cell* **137**, 961-71 (2009).
1479   49.   Floresco, S.B., Todd, C.L. & Grace, A.A. Glutamatergic afferents from the hippocampus
1480   to the nucleus accumbens regulate activity of ventral tegmental area dopamine neurons. *J*
1481   *Neurosci* **21**, 4915-22 (2001).
1482   50.   Kim, M.S. *et al.* Prefrontal Cortex and Amygdala Subregion Morphology Are Associated
1483   With Obesity and Dietary Self-control in Children and Adolescents. *Front Hum Neurosci* **14**,
1484   563415 (2020).
1485   51.   Pizzagalli, D.A. *et al.* Reduced caudate and nucleus accumbens response to rewards in
1486   unmedicated individuals with major depressive disorder. *Am J Psychiatry* **166**, 702-10 (2009).
1487   52.   Yoncheva, Y.N. *et al.* Mode of Anisotropy Reveals Global Diffusion Alterations in
1488   Attention-Deficit/Hyperactivity Disorder. *J Am Acad Child Adolesc Psychiatry* **55**, 137-45 (2016).
1489   53.   Travis, K.E., Leitner, Y., Feldman, H.M. & Ben-Shachar, M. Cerebellar white matter
1490   pathways are associated with reading skills in children and adolescents. *Hum Brain Mapp* **36**,
1491   1536-53 (2015).
1492   54.   Olivo, G., Latini, F., Wiemerslage, L., Larsson, E.M. & Schioth, H.B. Disruption of
1493   Accumbens and Thalamic White Matter Connectivity Revealed by Diffusion Tensor
1494   Tractography in Young Men with Genetic Risk for Obesity. *Front Hum Neurosci* **12**, 75 (2018).
1495   55.   Wible, C.G. *et al.* Prefrontal cortex, negative symptoms, and schizophrenia: an MRI
1496   study. *Psychiatry Res* **108**, 65-78 (2001).
1497   56.   Selemon, L.D. & Zecevic, N. Schizophrenia: a tale of two critical periods for prefrontal
1498   cortical development. *Transl Psychiatry* **5**, e623 (2015).
1499   57.   Bryois, J. *et al.* Genetic identification of cell types underlying brain complex traits yields
1500   insights into the etiology of Parkinson's disease. *Nat Genet* **52**, 482-493 (2020).
1501   58.   Trubetskoy, V. *et al.* Mapping genomic loci implicates genes and synaptic biology in
1502   schizophrenia. *Nature* **604**, 502-508 (2022).

1503   59.   Andreasen, N.C. & Pierson, R. The role of the cerebellum in schizophrenia. *Biol*
1504   *Psychiatry* **64**, 81-8 (2008).
1505   60.   Middleton, F.A. & Strick, P.L. Anatomical evidence for cerebellar and basal ganglia
1506   involvement in higher cognitive function. *Science* **266**, 458-61 (1994).
1507   61.   Heckers, S. Neuropathology of schizophrenia: cortex, thalamus, basal ganglia, and
1508   neurotransmitter-specific projection systems. *Schizophr Bull* **23**, 403-21 (1997).
1509   62.   Baldacara, L. *et al.* Is cerebellar volume related to bipolar disorder? *J Affect Disord* **135**,
1510   305-9 (2011).
1511   63.   Schutter, D.J., Koolschijn, P.C., Peper, J.S. & Crone, E.A. The cerebellum link to
1512   neuroticism: a volumetric MRI association study in healthy volunteers. *PLoS One* **7**, e37252
1513   (2012).
1514   64.   Beyer, J.L. *et al.* Caudate volume measurement in older adults with bipolar disorder. *Int*
1515   *J Geriatr Psychiatry* **19**, 109-14 (2004).
1516   65.   Nixon, P.D. & Passingham, R.E. Predicting sensory events. The role of the cerebellum in
1517   motor learning. *Exp Brain Res* **138**, 251-7 (2001).
1518   66.   Cechetto, D.F. & Saper, C.B. Neurochemical organization of the hypothalamic projection
1519   to the spinal cord in the rat. *J Comp Neurol* **272**, 579-604 (1988).
1520   67.   Saper, C.B. Hypothalamic connections with the cerebral cortex. *Prog Brain Res* **126**, 39-
1521   48 (2000).
1522   68.   Goriounova, N.A. & Mansvelder, H.D. Nicotine exposure during adolescence alters the
1523   rules for prefrontal cortical synaptic plasticity during adulthood. *Front Synaptic Neurosci* **4**, 3
1524   (2012).
1525   69.   Basinger, H. & Hogg, J.P. Neuroanatomy, Brainstem. in *StatPearls* (Treasure Island (FL),
1526   2022).
1527   70.   Cotsapas, C. *et al.* Pervasive sharing of genetic effects in autoimmune disease. *PLoS*
1528   *Genet* **7**, e1002254 (2011).
1529   71.   Shi, H., Mancuso, N., Spendlove, S. & Pasaniuc, B. Local Genetic Correlation Gives
1530   Insights into the Shared Genetic Architecture of Complex Traits. *Am J Hum Genet* **101**, 737-751
1531   (2017).
1532   72.   Zhang, Y. *et al.* SUPERGNOVA: local genetic correlation analysis reveals heterogeneous
1533   etiologic sharing of complex traits. *Genome Biol* **22**, 262 (2021).
1534   73.   Werme, J., van der Sluis, S., Posthuma, D. & de Leeuw, C.A. An integrated framework for
1535   local genetic correlation analysis. *Nat Genet* **54**, 274-282 (2022).
1536   74.   Schafer, W.R. & Kenyon, C.J. A calcium-channel homologue required for adaptation to
1537   dopamine and serotonin in Caenorhabditis elegans. *Nature* **375**, 73-8 (1995).
1538   75.   Weinshenker, D., Garriga, G. & Thomas, J.H. Genetic and pharmacological analysis of
1539   neurotransmitters controlling egg laying in C. elegans. *J Neurosci* **15**, 6975-85 (1995).
1540   76.   Triarhou, L.C. Introduction. Dopamine and Parkinson's disease. *Adv Exp Med Biol* **517**, 1-
1541   14 (2002).
1542   77.   Zielonka, M. *et al.* Dopamine-Responsive Growth-Hormone Deficiency and Central
1543   Hypothyroidism in Sepiapterin Reductase Deficiency. *JIMD Rep* **24**, 109-13 (2015).
1544   78.   Comings, D.E. *et al.* The dopamine D2 receptor (DRD2) as a major gene in obesity and
1545   height. *Biochem Med Metab Biol* **50**, 176-85 (1993).

1546   79.   Ameri, P. *et al.* Interactions between vitamin D and IGF-I: from physiology to clinical
1547   practice. *Clin Endocrinol (Oxf)* **79**, 457-63 (2013).
1548   80.   Wacker, M. & Holick, M.F. Vitamin D - effects on skeletal and extraskeletal health and
1549   the need for supplementation. *Nutrients* **5**, 111-48 (2013).
1550   81.   Delecroix, C., Brauner, R. & Souberbielle, J.C. Vitamin D in children with growth
1551   hormone deficiency due to pituitary stalk interruption syndrome. *BMC Pediatr* **18**, 11 (2018).
1552   82.   Nechita, F., Alexandru, D.O., Turcu-Stiolica, R. & Nechita, D. The Influence of Personality
1553   Factors and Stress on Academic Performance. *Curr Health Sci J* **41**, 47-61 (2015).
1554   83.   Groves, N.J., McGrath, J.J. & Burne, T.H. Vitamin D as a neurosteroid affecting the
1555   developing and adult brain. *Annu Rev Nutr* **34**, 117-41 (2014).
1556   84.   Di Somma, C. *et al.* Vitamin D and Neurological Diseases: An Endocrine View. *Int J Mol*
1557   *Sci* **18**(2017).
1558   85.   Margolis, K.L. *et al.* Prospective study of leukocyte count as a predictor of incident
1559   breast, colorectal, endometrial, and lung cancer and mortality in postmenopausal women. *Arch*
1560   *Intern Med* **167**, 1837-44 (2007).
1561   86.   Akinbami, A. *et al.* Full blood count pattern of pre-chemotherapy breast cancer patients
1562   in Lagos, Nigeria. *Caspian J Intern Med* **4**, 574-9 (2013).
1563   87.   Park, B., Lee, H.S., Lee, J.W. & Park, S. Association of white blood cell count with breast
1564   cancer burden varies according to menopausal status, body mass index, and hormone receptor
1565   status: a case-control study. *Sci Rep* **9**, 5762 (2019).
1566   88.   Wu, A.H., Yu, M.C., Thomas, D.C., Pike, M.C. & Henderson, B.E. Personal and family
1567   history of lung disease as risk factors for adenocarcinoma of the lung. *Cancer Res* **48**, 7279-84
1568   (1988).
1569   89.   Kreuzer, M., Gerken, M., Heinrich, J., Kreienbrock, L. & Wichmann, H.E. Hormonal
1570   factors and risk of lung cancer among women? *Int J Epidemiol* **32**, 263-71 (2003).
1571   90.   Kabat, G.C., Miller, A.B. & Rohan, T.E. Reproductive and hormonal factors and risk of
1572   lung cancer in women: a prospective cohort study. *Int J Cancer* **120**, 2214-20 (2007).
1573   91.   Baik, C.S., Strauss, G.M., Speizer, F.E. & Feskanich, D. Reproductive factors, hormone
1574   use, and risk for lung cancer in postmenopausal women, the Nurses' Health Study. *Cancer*
1575   *Epidemiol Biomarkers Prev* **19**, 2525-33 (2010).
1576   92.   Wraw, C., Deary, I.J., Gale, C.R. & Der, G. Intelligence in youth and health at age 50.
1577   *Intelligence* **53**, 23-32 (2015).
1578   93.   Backhouse, E.V., McHutchison, C.A., Cvoro, V., Shenkin, S.D. & Wardlaw, J.M. Early life
1579   risk factors for cerebrovascular disease: A systematic review and meta-analysis. *Neurology* **88**,
1580   976-984 (2017).
1581   94.   Deary, I.J., Harris, S.E. & Hill, W.D. What genome-wide association studies reveal about
1582   the association between intelligence and physical health, illness, and mortality. *Curr Opin*
1583   *Psychol* **27**, 6-12 (2019).
1584   95.   Shah, S.A., Page, C.P. & Pitchford, S.C. Platelet-Eosinophil Interactions As a Potential
1585   Therapeutic Target in Allergic Inflammation and Asthma. *Front Med (Lausanne)* **4**, 129 (2017).
1586   96.   Riazi, H., Ghazanfarpour, M., Taebi, M. & Abdolahian, S. Effect of Vitamin D on the
1587   Vaginal Health of Menopausal Women: A Systematic Review. *J Menopausal Med* **25**, 109-116
1588   (2019).

1589   97.   Maseroli, E. & Vignozzi, L. Testosterone and Vaginal Function. *Sex Med Rev* **8**, 379-392
1590   (2020).
1591   98.   de Heredia, F.P., Gomez-Martinez, S. & Marcos, A. Obesity, inflammation and the
1592   immune system. *Proc Nutr Soc* **71**, 332-8 (2012).
1593   99.   Tillmann, T. *et al.* Education and coronary heart disease: mendelian randomisation
1594   study. *BMJ* **358**, j3542 (2017).
1595   100.   El Khoudary, S.R. *et al.* Menopause Transition and Cardiovascular Disease Risk:
1596   Implications for Timing of Early Prevention: A Scientific Statement From the American Heart
1597   Association. *Circulation* **142**, e506-e532 (2020).
1598   101.   Hak, A.E., Pols, H.A., van Hemert, A.M., Hofman, A. & Witteman, J.C. Progression of
1599   aortic calcification is associated with metacarpal bone loss during menopause: a population-
1600   based longitudinal study. *Arterioscler Thromb Vasc Biol* **20**, 1926-31 (2000).
1601   102.   Gibson, D. & Mehler, P.S. Anorexia Nervosa and the Immune System-A Narrative
1602   Review. *J Clin Med* **8**(2019).
1603   103.   Sabharwal, P. *et al.* A Case of Recurrent Insomnia: Extending the Spectrum of
1604   Autoimmune Encephalitis. *J Clin Sleep Med* **12**, 763-5 (2016).
1605   104.   Zhang, W., Piotrowska, K., Chavoshan, B., Wallace, J. & Liu, P.Y. Sleep Duration Is
1606   Associated With Testis Size in Healthy Young Men. *J Clin Sleep Med* **14**, 1757-1764 (2018).
1607   105.   Wong, H.K., Hoermann, R. & Grossmann, M. Reversible male hypogonadotropic
1608   hypogonadism due to energy deficit. *Clin Endocrinol (Oxf)* **91**, 3-9 (2019).
1609   106.   Dahlman, D., Magnusson, H., Li, X., Sundquist, J. & Sundquist, K. Drug use disorder and
1610   risk of incident and fatal breast cancer: a nationwide epidemiological study. *Breast Cancer Res
1611   Treat* **186**, 199-207 (2021).
1612   107.   Delezie, J. & Handschin, C. Endocrine Crosstalk Between Skeletal Muscle and the Brain.
1613   *Front Neurol* **9**, 698 (2018).
1614   108.   Elsayed, E.F. *et al.* Waist-to-hip ratio and body mass index as risk factors for
1615   cardiovascular events in CKD. *Am J Kidney Dis* **52**, 49-57 (2008).
1616   109.   Thalacker-Mercer, A.E., Fleet, J.C., Craig, B.A., Carnell, N.S. & Campbell, W.W.
1617   Inadequate protein intake affects skeletal muscle transcript profiles in older humans. *Am J Clin
1618   Nutr* **85**, 1344-52 (2007).
1619   110.   Triebner, K. *et al.* Menopause as a predictor of new-onset asthma: A longitudinal
1620   Northern European population study. *J Allergy Clin Immunol* **137**, 50-57 e6 (2016).
1621   111.   van der Plaat, D.A. *et al.* Age at menopause and lung function: a Mendelian
1622   randomisation study. *Eur Respir J* **54**(2019).
1623   112.   Peng, H. *et al.* Age at first birth and lung cancer: a two-sample Mendelian randomization
1624   study. *Transl Lung Cancer Res* **10**, 1720-1733 (2021).
1625   113.   Lee, J.I. *et al.* Association between visceral adipose tissue and major depressive disorder
1626   across the lifespan: A scoping review. *Bipolar Disord* **24**, 375-391 (2022).
1627   114.   Reddy, L.F. *et al.* Impulsivity and risk taking in bipolar disorder and schizophrenia.
1628   *Neuropsychopharmacology* **39**, 456-63 (2014).
1629   115.   Hormozdiari, F. *et al.* Leveraging molecular quantitative trait loci to understand the
1630   genetic architecture of diseases and complex traits. *Nat Genet* **50**, 1041-1047 (2018).
1631   116.   Gamazon, E.R. *et al.* Using an atlas of gene regulation across 44 human tissues to inform
1632   complex disease- and trait-associated variation. *Nat Genet* **50**, 956-967 (2018).

1633 117.   Yao, D.W., O'Connor, L.J., Price, A.L. & Gusev, A. Quantifying genetic effects on disease
1634 mediated by assayed gene expression levels. *Nat Genet* **52**, 626-633 (2020).
1635 118.   Arvanitis, M., Tayeb, K., Strober, B.J. & Battle, A. Redefining tissue specificity of genetic
1636 regulation of gene expression in the presence of allelic heterogeneity. *Am J Hum Genet* **109**,
1637 223-239 (2022).
1638 119.   Leyden, G.M. *et al.* Harnessing tissue-specific genetic variation to dissect putative causal
1639 pathways between body mass index and cardiometabolic phenotypes. *Am J Hum Genet* **109**,
1640 240-252 (2022).
1641 120.   Thom, C.S., Wilken, M.B., Chou, S.T. & Voight, B.F. Body mass index and adipose
1642 distribution have opposing genetic impacts on human blood traits. *Elife* **11**(2022).
1643 121.   Weissbrod, O. *et al.* Functionally informed fine-mapping and polygenic localization of
1644 complex trait heritability. *Nat Genet* **52**, 1355-1363 (2020).
1645 122.   Nathan, A. *et al.* Single-cell eQTL models reveal dynamic T cell state dependence of
1646 disease loci. *Nature* **606**, 120-128 (2022).
1647 123.   Perez, R.K. *et al.* Single-cell RNA-seq reveals cell type-specific molecular and genetic
1648 associations to lupus. *Science* **376**, eabf1970 (2022).
1649 124.   Soskic, B. *et al.* Immune disease risk variants regulate gene expression dynamics during
1650 CD4(+) T cell activation. *Nat Genet* **54**, 817-826 (2022).
1651