The landscape of pervasive horizontal pleiotropy in human genetic variation is driven by extreme polygenicity of human traits and diseases

Daniel M. Jordan, Ph.D.[1,2,3*], Marie Verbanck, Ph.D.[1,2,3*], Ron Do, Ph.D.[1,2,3]

[1]The Charles Bronfman Institute for Personalized Medicine, Icahn School of Medicine at Mount Sinai, 1468 Madison Avenue, New York, NY, USA

[2]The Icahn Institute for Genomics and Multiscale Biology, Icahn School of Medicine at Mount Sinai, 1425 Madison Ave, New York, NY, USA

[3]Department of Genetics and Genomic Sciences, Icahn School of Medicine at Mount Sinai, 1425 Madison Avenue, New York, NY, USA

*Dr. Jordan and Dr. Verbanck contributed equally to this work.

Addresses for correspondence:
Ron Do, Ph.D.
The Charles Bronfman Institute for Personalized Medicine
Department of Genetics and Genomic Sciences
Icahn School of Medicine at Mount Sinai
One Gustave L. Levy Pl., Box 1003
New York, NY 10029
Email: ron.do@mssm.edu
Tel: 212-241-6206

Abstract

Horizontal pleiotropy, where one variant has independent effects on multiple traits, is important for our understanding of the genetic architecture of human phenotypes. We applied a method to quantify horizontal pleiotropy using genome-wide association summary statistics to 372 heritable phenotypes measured in 361,194 UK Biobank individuals. We observed horizontal pleiotropy is: 1) pervasive throughout the human genome and across a wide range of phenotypes; 2) especially prominent among highly polygenic phenotypes; 3) detected in 24,968 variants in 7,831 loci, a majority of which have never reported in any genome-wide association study; and 4) enriched in active regulatory regions. Our results highlight the central role horizontal pleiotropy plays in the genetic architecture of human phenotypes, and the importance of modeling horizontal pleiotropy in genomic medicine.

Introduction

The term “pleiotropy” refers to a single genetic variant having multiple distinct phenotypic effects. In general terms, the existence and extent of pleiotropy has far-reaching implications on our understanding of how genotypes map to phenotypes [1], of the genetic architectures of traits [2,3], of the biology underlying common diseases [4] and of the dynamics of natural selection [5]. However, beyond this general idea of the importance of pleiotropy, it quickly becomes difficult to discuss in specifics, because of the difficulty in defining what counts as a direct causal effect and what counts as a separate phenotypic effect.

One particularly important dividing line in these conflicting definitions is the distinction between vertical pleiotropy and horizontal pleiotropy [6,7]. When a genetic variant has a phenotypic effect that then has its own downstream effects in turn, that variant exhibits “vertical” pleiotropy. For example, a variant that increases low density lipoprotein (LDL) cholesterol might also have an additional corresponding effect on coronary artery disease risk due to the causal relationship between these two traits, thus exhibiting vertical pleiotropy. Vertical pleiotropy has been conceptualized and measured by explicit genetic methods like Mendelian randomization.

In contrast, a genetic cause that directly influences multiple traits, without one trait being mediated by another, would exhibit “horizontal” pleiotropy. Horizontal pleiotropy contains some conceptual difficulties, and consequently can be difficult to measure. In principle, we might imagine selecting a variant and counting how many phenotypes are associated with it. Indeed, several versions of this analysis have been performed for different lists of traits [8,2,3,9]. However, the results of these analyses are highly dependent on the exact list of traits used, and traits of interest to researchers previously tend to involve only a small number of phenotypes and/or be heavily biased towards a small set of disease-relevant biological systems and processes. Due to these limitations, it is unknown to what extent horizontal pleiotropy affects genetic variation in the human genome at the genome-wide level.

The proliferation of data sources like large-scale biobanks and metabolomics data that include a wide array of phenotypes in one dataset, combined with the growing public availability of genome-wide association studies (GWASs) summary statistic data, especially for extremely large meta-analyses, has allowed the development of methods that use these summary statistics to gain insight into human biology, and particularly into the genetic architecture of complex traits and diseases [10].

Here, we present a method to measure horizontal pleiotropy using publicly available GWAS summary statistics. We focus on measuring horizontal pleiotropy of SNVs on observable traits, meaning a scenario where a single SNV affects multiple independent phenotypes that do not have a detectable causal relationship. Using this framework, we are able to score each SNV in the human genome for horizontal pleiotropy, giving us broad insight into the genetic architecture of pleiotropy. Because our framework explicitly removes correlations between the input phenotypes, and because these phenotypes represent a diverse array of traits and diseases, these insights are largely robust to the specific list of traits studied, and pertain to human biology overall rather than relationships between specific traits.

Results

A quantitative score for pleiotropy

We have developed a method to measure horizontal pleiotropy using summary statistics data from GWASs on multiple traits. Our method relies on applying a statistical whitening procedure to a set of input variant-trait associations, which removes correlations between traits and normalizes effect sizes across all traits. Using the decorrelated association Z-scores, we measure two related but distinct components of pleiotropy: the total magnitude of effect on independent traits ("magnitude" score, denoted P_m) and the total number of independent traits affected by a variant ("number of traits" score, denoted P_n). Both scores are then scaled by the number of traits and multiplied by 100, so that the final score represents the value as it would be measured in a dataset of 100 traits. This two-component quantitative pleiotropy score allows us to measure both the magnitude (pleiotropy magnitude score P_m) and quantity (pleiotropy number of traits score P_n) of horizontal pleiotropy for all SNVs in the human genome. Since we expect these scores to be heavily influenced by linkage disequilibrium (LD), we regress P_m and P_n against LD scores to produce an LD-corrected score (P_m^{LD} and P_m^{LD}) (**Figure 1**; **Methods**).

Calculating significance of pleiotropy

We compute *P-values* for the two components of our pleiotropy score using two different procedures, corresponding to two different null expectations.

1. Theoretical *P*-values (Raw pleiotropy score [P_m and P_n] or LD-corrected pleiotropy score [P_m^{LD} and P_n^{LD}]), calculated analogously to *P*-values for genetic association studies including GWAS, based on a null scenario where variants do not exhibit pleiotropic effects on observed traits.
2. Empirical *P*-values (Polygenicity/LD-corrected pleiotropy score [P_m^P and P_n^P]), calculated by permutation of the observed distributions of whitened traits. These *P*-values are based on a null scenario where variants may have significant effects on one or more traits, but the effects of each variant on each trait are independent and the number of variants with effects on multiple traits is no more than would be expected by chance.

This empirical correction for polygenicity is required because polygenicity is a major factor that can produce pleiotropy. For example, it has been estimated that approximately 100,000 independent loci are causal for height in humans [11]. If the total number of independent loci in the human genome is approximately 1 million, this corresponds to about 10% of the human genome having an effect on height. If we imagine multiple phenotypes with this same highly polygenic genetic architecture, we should expect substantial overlap between causal loci for multiple different traits, even in the absence of any true causal relationship between the traits, resulting in horizontal pleiotropy.

Power to detect pleiotropy in simulations

We conducted a simulation study to evaluate the performance of our two-component pleiotropy score. We simulated 800,000 variants with a realistic LD structure controlling 100 traits, varying

the per-trait liability scale heritability of all traits h^2 and the proportion of pleiotropic and non-pleiotropic causal variants. Under the null model, all trait-variant associations were independent, and no horizontal pleiotropy was added. Under the added-pleiotropy models, we randomly chose a fraction of causal variants and forced them to have simultaneous associations with multiple traits. The simulation study showed that both components of the pleiotropy score were well-powered to detect horizontal pleiotropy (**Figure 2**). Under the null hypothesis of no added horizontal pleiotropy, the false positive rate was well controlled for both scores when there was low heritability or few causal variants. However, when there are many causal variants and high per-variant heritability, the LD-corrected pleiotropy score (P_m^{LD} and P_n^{LD}) detects a large excess of pleiotropic variants, due to serendipitous overlap between causal variants without explicitly induced pleiotropy. The LD/polygenicity-corrected empirical *P*-value (P_m^P and P_n^P) does not detect this serendipitous pleiotropy at the same high rate.

In the presence of added horizontal pleiotropy, our approach was powered to detect pleiotropy with per-variant heritability h^2 as small as 0.002 if there are no non-pleiotropic causal variants. In the presence of both pleiotropic and non-pleiotropic causal variants, detecting pleiotropy was more difficult, but our approach still had appreciable power to detect pleiotropic variants, which increased with increasing per-variant heritability and decreased with increasing numbers of non-pleiotropic causal variants. Adding the correction for polygenic architecture (P_m^P and P_n^P) reduced this power only slightly. The power of our method was not substantially reduced by increasing the number of traits affected by pleiotropic variants (**Figure S1**) or by adding a realistic correlation structure between the traits (**Figure S2**).

Genome-Wide Pleiotropy Study (GWPS) reveals pervasive pleiotropy

To apply our method to real human association data, we used GWAS association statistics for 372 heritable medical traits measured in 337,119 individuals from the UK Biobank [12,13] . We successfully computed our two-component pleiotropy score for 767,057 variants genome-wide and conducted a genome-wide pleiotropy study (GWPS), by analogy to a standard GWAS (**Figure 1**; **Methods**). **Figure S3** shows the resulting quantile-quantile plots (Q-Q plots). We observed significant inflation for both the LD-corrected magnitude score P_m^{LD} and number of traits score P_n^{LD} (Mann-Whitney U test $P < 10^{-300}$ for both). Furthermore, we observed across both scores that horizontal pleiotropy was widely distributed across the genome, rather than being localized to a few specific loci (**Figure S4**). Testing an alternative strategy for computing the phenotype-correlation matrix using all SNVs produced comparable results (Pearson r = 0.995 and 0.964 for P_m^{LD} and P_n^{LD} respectively) to our strategy of using a pruned set of SNVs to account for LD (r^2 < 0.1) (**Figure S5**).

Pleiotropy is driven by polygenicity

We applied the permutation-based empirical *P*-value calculation (Polygenicity/LD corrected pleiotropy score: P_m^P and P_n^P) to correct for the known polygenic architecture of traits and test whether any loci are pleiotropic to a greater extent than would be expected due to polygenicity. **Figures S6 and S7** show the resulting Q-Q plots and Manhattan plots. In contrast to the results

from the LD-corrected pleiotropy score (P_m^{LD} and P_n^{LD}), we find no evidence for pleiotropy in excess of what would be expected from the known polygenic architecture of traits: there are no loci with genome-wide significant levels of pleiotropy after correcting for polygenic architecture, and the genome-wide distribution of pleiotropy score shows less pleiotropy than expected (Mann-Whitney U test $P < 10^{-300}$ for both P_m^P and P_n^P).

As an additional test of whether the pleiotropy we observe is driven by polygenicity, we calculated the polygenicity of the same 372 heritable traits from the UK Biobank. We measured polygenicity using a version of the genomic inflation factor corrected using LD score λ_{GC}^c[14]. We then stratified these traits by λ_{GC}^c after controlling for heritability (**Methods**), and calculated the two-component LD-corrected pleiotropy score [P_m^{LD} and P_n^{LD}]) and *P-values* for each component independently for every variant in the genome using each of these bins of traits. We observed that both scores are highly dependent on polygenicity, with the lowest-polygenicity bins in each heritability class showing very little inflation. (**Figure 3**; **Table S1**). Taken together, these results suggest that extreme polygenicity drives pleiotropy, and that this has an extremely large effect on the genetic architecture of human phenotypes.

Genome-wide distribution of pleiotropy score gives insight into genetic architecture

In addition to observing genome-wide inflation of the pleiotropy score, we can also gain insight from the distribution of the pleiotropy score on a more granular level.

Figure 4a shows the distribution of pleiotropy score for independent SNVs (LD pruned to a threshold of $r^2 < 0.1$) compared to the expectation under the null hypothesis of no pleiotropic effect. We observe a large excess in the number of traits score P_n^{LD}, and a smaller but still highly significant excess in total magnitude of pleiotropic effect P_m^{LD}. This excess comes in part from a long tail of highly pleiotropic loci that pass the threshold of genome-wide significance (dashed line in **Figure 4a**), but is primarily driven by weak pleiotropy among loci that do not reach genome-wide significance.

Pleiotropy score is correlated with molecular and biological function

To further investigate the properties of pleiotropic variants, we examined the effects of various functional and biochemical annotations on our LD-corrected pleiotropy score (P_m^{LD} and P_n^{LD}) **(Table 1; Methods)**. Using annotations from Ensembl Variant Effect Predictor [15], we observed that both components of the pleiotropy score are higher on average in transcribed regions (coding and UTR) than in intergenic noncoding regions. This result was confirmed and expanded by annotations from Roadmap Epigenomics [16], which showed that regions whose chromatin configurations were associated with actively transcribed regions, promoters, enhancers, and transcription factor binding sites had significantly higher levels of both components of the pleiotropy score, while heterochromatin and quiescent chromatin states had significantly lower levels. Investigating individual histone marks, we found that both the repressive histone mark H3K27me3 and the activating histone mark H3K27ac were associated

with elevated levels of pleiotropy, although the activating mark H3K27ac had a larger effect. This may indicate that being under active regulation at all produces higher levels of pleiotropy, whether that regulation is repressive or activating.

We also used data from the Genotype-Tissue Expression [17] project to measure the connection between transcriptional effects and our pleiotropy score **(Table 1)**. Consistent with the previous observation that functional regions had higher pleiotropy scores, we found that variants that were identified as *cis*-eQTLs for any gene in any tissue had higher pleiotropy scores on average. Within eQTLs, we also observed significant correlations between our pleiotropy score and the numbers of genes (P_m^{LD}: r = 0.036, $P < 2.2 \times 10^{-16}$; P_n^{LD}: r = 0.035, $P < 2.2 \times 10^{-16}$) and tissues (P_m^{LD}: r = 0.062, $P < 2.2 \times 10^{-16}$; P_n^{LD}: r = 0.059, $P < 2.2 \times 10^{-16}$) where the variant was annotated as an eQTL, showing that our pleiotropy score is related to transcriptional measures of pleiotropy.

Finally, we found that variants that are eQTLs for genes whose orthologs are associated with multiple measurable phenotypes in mice or yeast have higher pleiotropy scores, demonstrating that our pleiotropy score is also related to pleiotropy in model organisms.

All these results are consistent when using the Polygenicity/LD-corrected pleiotropy score (P_m^P and P_n^P), indicating that the association of pleiotropy with molecular and biological function is not exclusively driven by highly polygenic architecture (**Table S2**).

Genome-wide pleiotropy study identifies novel biological loci

By analogy to standard GWAS, our GWPS methodology can identify individual variants that have a genome-wide significant level of horizontal pleiotropy. Using the LD-corrected magnitude score P_m^{LD}, we identified 74,335 variants in 8,093 independent loci with a genome-wide significant level of horizontal pleiotropy, while using the LD-corrected number of traits score P_n^{LD} identified 18,393 variants in 2,859 independent loci with a genome-wide significant level of horizontal pleiotropy (**Methods**, **Table S3**). Strikingly, a majority of these loci significant for P_n^{LD} (1,519 of 2,859) has no entry in the NHGRI-EBI GWAS catalog, meaning that they have never been reported as an associated locus in any published GWAS. These loci represent an under-recognized class of genetic variation that has multiple weak to intermediate effects on many independent phenotypes, but no specific strong effect on any one particular trait. Functional enrichment analysis on genes near genome-wide significant loci includes a wide range of biological functions, including cell adhesion, post-translational modification of proteins, cytoskeleton, transcription factors, and intracellular signaling cascades (**Table S4**). The role of these novel loci and these biological processes in human genetics and biology may be a fruitful area for future study, with the potential for biological discovery.

Pleiotropic loci replicate in independent GWAS datasets

As replication datasets, we used two additional sources of GWAS summary statistics to calculate our LD-corrected pleiotropy score (P_m^{LD} and P_n^{LD}): previously published GWASs and meta-analyses for 73 human complex traits and diseases, which we collected and curated manually from the literature (**Methods, Table S5**) [18]; and a previously published study of 430

blood metabolites measured in 7,824 European adults [19]. For all variants covered by the UK Biobank, we were able to compute our pleiotropy score independently using these two datasets (**Figure 5**). In the traits and diseases dataset, we observed that 57% of P_m^{LD} loci and 38% of P_n^{LD} loci replicated, while in the blood metabolites dataset, we observed that 17% of P_m^{LD} loci and 12% of P_n^{LD} loci replicated, compared to 5% of P_m^{LD} loci and 6% of P_n^{LD} loci expected by chance according to a permutation-based null model. This high level of replication using independent sets of GWAS summary statistics suggests that our pleiotropy score is capturing an underlying biological property, rather than an artifact of the UK Biobank study.

Pleiotropy is correlated with specific complex traits and diseases

To characterize the phenotypic associations of these loci, we used our replication dataset of published GWAS summary statistics for 73 human quantitative traits and diseases, plus nine additional traits we excluded from our replication dataset for a total of 82 (**Methods**). **Figure 4c** shows the correlation between our LD-corrected pleiotropy score (P_m^{LD} and P_n^{LD}) and the association statistics for these 82 traits and diseases. The most strongly correlated traits were anthropometric traits like body mass index, waist and hip circumference, and height; certain blood lipid levels, including total cholesterol and triglycerides; and schizophrenia. These are all known to be highly polygenic and heterogeneous traits. The least correlated traits include several measurements of insulin sensitivity and glucose response, certain features of brain morphology, and the inflammatory biomarker lipoprotein(a). This is likely due to low sample size of the corresponding GWASs. However, these correlations do not appear to be driven exclusively by sample size: in cases where multiple GWASs for the same trait have been performed on subsamples of the population (for example, males only, female only, and combined), the sample size only marginally affects the correlation (**Table S6**).

Discussion

We have presented a framework for scoring horizontal pleiotropy across human genetic variation. In contrast to previous analyses, our framework explicitly distinguishes between horizontal pleiotropy and vertical pleiotropy or biological causation. After applying both components of our pleiotropy score to 372 heritable medical traits from the UK Biobank, we made the following observations: 1) horizontal pleiotropy is pervasive and widely distributed across the genome; 2)) horizontal pleiotropy is driven by extreme polygenicity of traits; 3) horizontal pleiotropy is significantly enriched in actively transcribed regions and active regulatory regions, and is correlated with the number of genes and tissues for which the variant is an eQTL; 4) there are thousands of loci that exhibit extreme levels of horizontal pleiotropy, a majority of which have no previously reported associations; and 5) pleiotropic loci are enriched in specific complex traits including body mass index, height, and schizophrenia.

This analysis is enabled by the technique of whitening trait associations to remove correlations between traits. This lets us count pleiotropic effects in a more objective and systematic way, as opposed to manually selecting putatively independent traits to count, or manually grouping traits into independent blocks. However, it does come with two major limitations compared to these approaches. First, it is somewhat more difficult to tell which specific traits are driving a signal of pleiotropy at a particular locus. Our whitened traits are linear combinations of real observed traits, and do not necessarily correspond to any specific biological traits of interest. However, it is relatively easy to inspect the input GWAS summary statistics for a particular variant of interest to see which traits it is associated with. Furthermore, since pleiotropic loci are by definition associated with a large cross-section of traits, this kind of inspection is not likely to be very informative about specific traits. Second, the whitening procedure has the counterintuitive property that a variant that has a narrow effect on a single trait without also affecting correlated traits can appear to be highly pleiotropic. For example, if a variant had a strong risk-increasing effect on coronary artery disease (CAD), but no effect on any of the known upstream risk factors of CAD (such as blood lipid levels or adiposity) or any of the known downstream consequences of CAD (such as inflammatory biomarkers or increased mortality), such a variant would appear as highly pleiotropic in our analysis. Our analysis would interpret the variant as increasing the risk of CAD while suppressing these upstream and downstream factors. We believe this treatment is appropriate, however counterintuitive. Regardless, these kinds of isolated effects are fairly rare: in our dataset of 372 heritable traits from UK Biobank, only 6% of variants (42,684 of 767,057) reach genome-wide significance for only a single trait. Indeed, it is unlikely by definition that a variant is associated with only one trait from a set of correlated traits, since we compute our correlations from observed association statistics.

Our findings are in keeping with several recent studies that have found abundant pleiotropy in the genome [20,21,8,2,9]. Our pleiotropy score goes a step further than many of these studies by explicitly removing vertical pleiotropy between traits, which are indicative of fundamental biological relationships between traits [8,18,22]. Furthermore, the current study has evaluated horizontal pleiotropy in human genetic variation genome-wide, whereas previous studies have focused on only a small subset of disease-associated variants identified from GWAS. Our

results therefore suggest that there is substantial complexity and heterogeneity not only in causal relationships between human traits, but also in the genetic architecture of individual traits.

Our findings have several important implications for the field of human genetics. First, our observation of ubiquitous horizontal pleiotropy is problematic for Mendelian Randomization (MR) methods, which assumes horizontal pleiotropy to be absent. Recent developments in the field of MR include methods that account for horizontal pleiotropy explicitly [18,22,23]; our results reinforce the importance of these methods. The presence of widespread horizontal pleiotropy suggests that single-instrument methods that independently account for every variant, each of which presumably has pleiotropic effects on many different distinct traits, should be considered in addition to multi-instrument methods for MR, which collapse many variants into a single polygenic score for analysis, and therefore treat all variants equivalently.

Second, our results appear to support the “network pleiotropy” hypothesis of Boyle, Li, and Pritchard [11], which proposes widespread pleiotropy driven by small perturbations of densely connected functional networks, where any perturbation in a relevant cell type will have at least a small effect on all phenotypes affected by that cell type. A subsequent paper detailed a more specific mechanism, where causal effects are driven by many biological components that are only indirectly related to the phenotype itself [24]. Many of the functional enrichments we observe, including transcription factors, cytoskeleton, and intracellular signaling cascades, represent components that can plausibly influence a wide variety of cell types and processes, providing evidence for this model over one where a specific biological component is largely responsible for pleiotropy.

While our results largely support this network pleiotropy hypothesis, we have also demonstrated an alternate view of horizontal pleiotropy in the context of highly polygenic causation. In our simulations, introducing extreme polygenicity at the levels suggested by these papers inherently results in high levels of horizontal pleiotropy detectable by our score, independent of any assumptions about the mechanism of pleiotropy or of polygenicity. Indeed, our null hypothesis of no horizontal pleiotropy, that 5% of the genome is independently causal to each trait with $P < 0.05$, is trivially rejected when a single trait is influenced by an unexpectedly large fraction of the genome. This means that, on some level, widespread horizontal pleiotropy in human genetic variation is simply a logical consequence of widespread polygenicity of human traits, regardless of the specific mechanism of either. In simple terms, the more loci are associated with each trait, the more chances there are for associations with multiple traits to overlap. Supporting this result, we find that controlling for the polygenic architecture of the input traits significantly attenuates our signal of pleiotropy, as does restricting to oligogenic traits. It may be the case that horizontal pleiotropy is only truly widespread among the most complex and polygenic subset of human traits.

In this study, we have presented a quantitative score for horizontal pleiotropy in human genome variation. Using this score, we have identified a genome-wide trend of highly inflated levels of horizontal pleiotropy, an underappreciated relationship between horizontal pleiotropy with polygenicity and functional biology, and a large number of specific novel loci with high levels of

horizontal pleiotropy. We expect further investigations using this score to yield deep insights into the genetic architecture of human traits and to uncover important novel biology.

Acknowledgments

We thank the various genome-wide association consortia as well as Dr. Benjamin Neale's group for generously sharing the genome-wide association summary statistics. R.D is supported by R35GM124836 from the National Institute Of General Medical Sciences of the National Institutes of Health, R01HL139865 from the National Heart, Lung, Blood Institute of the National Institutes of Health and previously an American Heart Association Cardiovascular Genome-Phenome Discovery grant (15CVGPSD27130014). D.M.J. is supported by T32HL00782 from the National Heart, Lung, and Blood Institute of the National Institutes of Health. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

Author Contributions

D.M.J. and M.V. contributed to study conception, data analysis, interpretation of the results and drafting of the manuscript. R.D. contributed to study conception, interpretation of the results and critical revision of the manuscript.

Declaration of Interests

R.D has received research support from AstraZeneca and Goldfinch Bio.

References

1. Zhan, J., Arking, D. E. & Bader, J. S. Discovering patterns of pleiotropy in genome-wide association studies. *bioRxiv* 273540 (2018). doi:10.1101/273540
2. Chesmore, K., Bartlett, J. & Williams, S. M. The ubiquity of pleiotropy in human disease. *Hum Genet* 1–6 (2017). doi:10.1007/s00439-017-1854-z
3. Socrates, A. *et al.* Polygenic risk scores applied to a single cohort reveal pleiotropy among hundreds of human phenotypes. *bioRxiv* 203257 (2017). doi:10.1101/203257
4. Solovieff, N., Cotsapas, C., Lee, P. H., Purcell, S. M. & Smoller, J. W. Pleiotropy in complex traits: challenges and strategies. *Nat. Rev. Genet.* **14**, 483–495 (2013).
5. Keightley, P. D. & Hill, W. G. Variation maintained in quantitative traits with mutation–selection balance: pleiotropic side-effects on fitness traits. *Proc. R. Soc. Lond. B* **242**, 95–100 (1990).
6. Paaby, A. B. & Rockman, M. V. The many faces of pleiotropy. *Trends in Genetics* **29**, 66–73 (2013).
7. Tyler, A. L., Asselbergs, F. W., Williams, S. M. & Moore, J. H. Shadows of complexity: what biological networks reveal about epistasis and pleiotropy. *BioEssays* **31**, 220–227 (2009).
8. Pickrell, J. K. *et al.* Detection and interpretation of shared genetic influences on 42 human traits. *Nat. Genet.* **48**, 709–717 (2016).
9. Kanai, M. *et al.* Genetic analysis of quantitative traits in the Japanese population links cell types to complex human diseases. *Nat. Genet.* **50**, 390–400 (2018).
10. Pasaniuc, B. & Price, A. L. Dissecting the genetics of complex traits using summary association statistics. *Nature Reviews Genetics* **18**, 117–127 (2017).
11. Boyle, E. A., Li, Y. I. & Pritchard, J. K. An Expanded View of Complex Traits: From Polygenic to Omnigenic. *Cell* **169**, 1177–1186 (2017).
12. Neale, Benjamin. UK Biobank. (2018).

13. Zhou, W. *et al.* Efficiently controlling for case-control imbalance and sample relatedness in large-scale genetic association studies. *Nature Genetics* **50**, 1335–1341 (2018).
14. Bulik-Sullivan, B. K. *et al.* LD Score regression distinguishes confounding from polygenicity in genome-wide association studies. *Nature Genetics* **47**, 291 (2015).
15. McLaren, W. *et al.* The Ensembl Variant Effect Predictor. *Genome Biology* **17**, 122 (2016).
16. Bernstein, B. E. *et al.* The NIH Roadmap Epigenomics Mapping Consortium. *Nature Biotechnology* **28**, 1045 (2010).
17. GTEx Consortium *et al.* Genetic effects on gene expression across human tissues. *Nature* **550**, 204–213 (2017).
18. Verbanck, M., Chen, C.-Y., Neale, B. & Do, R. Detection of widespread horizontal pleiotropy in causal relationships inferred from Mendelian randomization between complex traits and diseases. *Nature Genetics* 1 (2018). doi:10.1038/s41588-018-0099-7
19. Shin, S.-Y. *et al.* An atlas of genetic influences on human blood metabolites. *Nat Genet* **46**, 543–550 (2014).
20. Wang, Z., Liao, B.-Y. & Zhang, J. Genomic patterns of pleiotropy and the evolution of complexity. *PNAS* **107**, 18034–18039 (2010).
21. Sivakumaran, S. *et al.* Abundant pleiotropy in human complex diseases and traits. *Am. J. Hum. Genet.* **89**, 607–618 (2011).
22. Bowden, J. *et al.* A framework for the investigation of pleiotropy in two-sample summary data Mendelian randomization. *Statist. Med.* n/a-n/a (2017). doi:10.1002/sim.7221
23. Bowden, J., Smith, G. D. & Burgess, S. Mendelian randomization with invalid instruments: effect estimation and bias detection through Egger regression. *Int. J. Epidemiol.* **44**, 512–525 (2015).
24. Liu, X., Li, Y. I. & Pritchard, J. K. Trans effects on gene expression can drive omnigenic inheritance. *bioRxiv* 425108 (2018). doi:10.1101/425108

Methods

We developed a statistical method to measure horizontal pleiotropy using a two-component pleiotropy score. For a given variant, we measured 1) the total magnitude of pleiotropic effect the variant has and 2) the number of independent traits affected by the variant.

Z-scores decorrelation strategy

Observable traits and diseases can be highly correlated, which can lead to inflation of our pleiotropy score if the correlation is not properly accounted for. Therefore, we developed an efficient strategy to remove this correlation and obtain independent traits. Let Z^{raw} denote the matrix of raw Z-scores, with variants in columns and traits in rows, and Σ denote the corresponding correlation matrix between the Z-scores. Under the null hypothesis of no horizontal pleiotropy, Z-scores for each trait are assumed to follow a Gaussian distribution $N(0,1)$, and the columns of Z^{raw} collectively follow a multivariate Gaussian distribution $N(0,\Sigma)$. Our goal is to eliminate the extra-diagonal terms of the correlation matrix Σ. To achieve this, we use a Mahalanobis whitening transformation on the matrix Z^{raw} to obtain a whitened Z-score matrix Z. The procedure to obtain Z can be formally expressed as:

$$Z = \Sigma^{\frac{-1}{2}} Z^{raw}$$

Under the null hypothesis of no horizontal pleiotropy, we expect Z to follow a multivariate Gaussian distribution $N(0, Id_l)$, where Id_l is the identity matrix of size l, l being the number of traits.

In reality, the true correlation matrix Σ is unknown, and we must use an estimated correlation matrix $\hat{\Sigma}$ obtained by measuring the genome-wide correlation between actual Z-scores. We tested two approaches to obtain $\hat{\Sigma}$, either using all genotyped variants genome-wide or using a subset of variants pruned to $r^2 < 0.1$ in the 1000 Genomes European population to account for the effects of linkage disequilibrium (LD). Both approaches produced similar results (See **Figure S3**). In all subsequent analysis, we used covariance matrices estimated from pruned variants.

Computation of the pleiotropy score

We computed two different scores to capture both the magnitude and number of traits of pleiotropy. First, we quantify the total pleiotropic magnitude of effect a variant using the magnitude pleiotropy score P_m:

$$P_m = \frac{100}{l} \sqrt{\sum_{1}^{l} {z_i}^2}$$

where z_i is the whitened Z-score for trait i for a given variant. Second, we quantify the number of independent traits affected by a variant using the number of pleiotropic traits score P_n:

$$P_n = \frac{100}{l}\sum_{1}^{l} H(z_i - 2)$$

where z_i is the whitened Z-score for trait i for the tested variant and $H()$ is the Heaviside step function which equals 1 if $|z_i| > 2$ and 0 otherwise. 2 represents a standard value of the Z-score which represents the normal threshold for nominal significance ($P < 0.05$).

LD-corrected pleiotropy score

Similarly to LD score regression, each component of the pleiotropy score was regressed on the LD scores for all variants. Then, we regressed out the effect of LD on each component of the pleiotropy score independently to obtain an LD-corrected pleiotropy score. The LD-corrected pleiotropy score components P_m^{LD} and P_n^{LD} are given by:

$$P_m^{LD} = P_m - \beta_m l$$

$$P_n^{LD} = P_n - \beta_n l$$

where l is the LD score of the variant site, and β_m and β_n are the regression coefficients for LD score on P_m and P_n, respectively.

Computation of theoretical *P-values* for the pleiotropy score

Based on the observation that Z follows a multivariate standard Gaussian distribution $N(0, Id_l)$ under the null hypothesis of no pleiotropy, *P-values* can easily be computed for P_m and P_n. Under the null hypothesis, the square of P_m (or P_m^{LD}) follows a chi-square distribution $\chi^2(l)$ where l is the total number of traits. Likewise, P_n (or P_n^{LD}) follows a binomial distribution $B(l, p)$ where l is the total number of traits and *p* the probability to get a Z-score greater than 2 under the standard Gaussian distribution ($P \approx 0.045$).

Computation of empirical (polygenicity/LD-corrected) *P-values* for the pleiotropy score

To correct for the known polygenic architecture of traits in addition to LD, we additionally computed empirical permutation-based *P*-values for both P_m^{LD} and P_n^{LD}. We performed 25 random permutations of the input Z-scores for each observable trait, producing millions of permuted variants. We calculated P_m and P_n for each of these permuted variants, and then rank ordered the resulting scores. The empirical *P*-value corresponding to a value of P_m^{LD} or P_n^{LD} is given by the fraction of permuted variants with higher scores than the given value. We converted these *P*-values into polygenicity/LD-corrected P_m^P and P_n^P scores by converting each *P*-value into the score it would correspond to under the expected (theoretical) distributions described above.

Simulation framework

We simulated a realistic matrix of Z-scores Z with 100 traits and 800,000 genotyped variants. For non-causal variants, Z-scores for each trait were drawn from an independent Gaussian distribution $N(0,1)$. A subset of variants was designated as causal, either pleiotropically or non-pleiotropically. For these causal variants, Z-scores were drawn from a different Gaussian distribution $N(0, h^2)$,where h^2 is a parameter representing the per-variant heritability of each trait. Non-pleiotropic variants were selected independently for each trait, while pleiotropic variants were selected globally and each forced to be causal for a specified number of traits ν.

Simulations were run for all combinations of the following parameters: 1) correlation structure: absent or present; 2) proportion of pleiotropic causal variants: 0.1% (800/800,000 variants) or 1% (8,000/800,000 variants); 3) proportion of non-pleiotropic causal variants: 0 (0/800,000 variants), 0.1% (800/800,000 variants), or 1% (8,000/800,000 variants); 4) number of traits involved in horizontal pleiotropy ν: 10 or 20; 5) per-variant heritability of traits h^2: 0.0002, 0.002, 0.02, or 0.2. Each scenario was replicated 10,000 times.

Collection of genome-wide association (GWA) summary statistics datasets

First, we retrieved GWA publicly available summary statistics from 545 continuous traits in 361,194 samples from the UK Biobank (see http://www.nealelab.is/uk-biobank/), and 1,403 binary traits from the same dataset calculated using SAIGE (see https://github.com/weizhouUMICH/SAIGE#uk-biobank-gwas-results). We used LD score regression to calculate heritability for each trait, using the liability scale for binary traits, and restricted the sample to only traits with a significant *P*-value for nonzero heritability after Bonferroni correction. For every pair of traits with correlation coefficient between Z-scores $r^2 > 0.8$, we additionally removed the member of the pair with lower heritability. This left a total of 372 traits.

Second, we retrieved publicly available genome-wide association (GWA) summary statistics data for 82 complex traits and diseases [1–36] (**Table S9**). For each dataset, we retrieved the appropriate variant annotation (build, rsid, chromosome, position, reference and alternate alleles) and summary statistics (effect size, standard errors, *P-values* and sample size of the study). All variant coordinates (chr, pos) were lifted over to hg19 using the UCSC Genome Browser LiftOver Tool and aligned to the reference and alternate alleles retrieved from the Ensembl variation database. After performing the same pruning of highly correlated phenotypes, we were left with 73 traits and diseases.

Third, we retrieved GWA summary statistics datasets from a GWAS of 453 blood metabolites in 7,824 individuals [37]. After performing the same pruning of highly correlated phenotypes, we were left with 430 metabolites.

Genome-wide pleiotropy study (GWPS)

Using the two components of the pleiotropy score, we can run a genome-wide pleiotropy study (GWPS) which consists of computing two *P-values* for each component of the score (P_m^{LD} and P_n^{LD}) and for all variants genome-wide. We computed the pleiotropy score separately for each of

the three datasets described above (372 UK Biobank phenotypes, 73 traits and diseases, and 430 blood metabolites). Additionally, we computed the pleiotropy score on a subset of 372 traits with genome-wide significant heritability based on the heritability estimates downloaded from https://sites.google.com/broadinstitute.org/ukbbgwasresults (univariate heritability significant after Bonferroni correction). The 372 UK Biobank heritable traits were used for discovery, and the 73 traits and diseases and 430 blood metabolites datasets were used for replication. There was a total of 768,756 variants genotyped across all three datasets.

Study of polygenicity on horizontal pleiotropy

To study the effect of polygenicity on horizontal pleiotropy, we first estimated the liability-scale heritability of all 372 traits in our UK Biobank dataset using LD score regression, and stratified all traits into four equally-sized classes of heritability, in order to control for the effect of high heritability separate from the effect of high polygenicity. Next, we estimated the polygenicity of the 372 traits using a corrected version of the genomic inflation factor λ_{CG}^{c} [38]. The intercept of LD score regression minus one is an estimator of the mean contribution of confounding bias to the inflation in the test statistics. Therefore, we computed a corrected version of the genomic inflation factor by subtracting the quantity (intercept of LD score regression - 1) from λ_{GC}. The 372 phenotypes were then ranked according to λ_{CG}^{c} within each heritability class, and grouped into 5 equal-sized bins of about 20 phenotypes each. We then recomputed the LD-corrected pleiotropy score components ($P_m^{LD}$and P_n^{LD}) for the subset of phenotypes in each bin. The inflation of the pleiotropy score was measured per bin to represent pleiotropy score inflation as a function of polygenicity.

Characterization of the pleiotropic variants

We performed various enrichment analyses for the pleiotropy score to characterize the pleiotropic variants using a variety of annotations that could be a direct consequence of horizontal pleiotropy. Each analysis uses the principle of assigning each variant an annotation category and selecting one category as the reference category. Then, for each category, we selected a set of variants from the corresponding reference category with minor allele frequencies matched to those in the query category, and performed a Student's t-test to test whether the average LD-corrected pleiotropy score ($P_m^{LD}$and P_n^{LD}) of the variants in each given category is different from the average LD-corrected pleiotropy score of the selected reference variants.

First, we used Ensembl Variant Effect Predictor [39] to classify each variant as noncoding, UTR, nonsynonymous, or coding synonymous, treating noncoding as the reference class. These were complemented by annotations from Roadmap Epigenomics [40]. We used the 25-state chromatin state model published by Roadmap Epigenomics to assign each variant 25 scores from 0 to 127, where each score represents the number of epigenomes for which that site is assigned to the corresponding category. We did the same for two specific chromatin marks: the activating mark H3K27ac and the repressive mark H3K27me3. For these annotations, we used a combination of all other categories as a reference set. In other words, the reference set for each category is all variants that are not in that category.

In addition to these molecular annotations, we used expression-related annotations from the Genotype-Tissue Expression project [41]. For each variant, we retrieved the number of genes for which the variant is referenced as a *cis* eQTL (expression quantitative trait loci) in any tissue (eGenes), and the number of tissues where the variant is annotated as a *cis* eQTL for any gene (eTissues). We divided variants into bins by number of eGenes (below 10, between 10 and 15, between 15 and 20, and over 20) and eTissues (below 30, between 30 and 35, between 35 and 40, and above 40). The reference set used for these analyses were variants that are not annotated as eQTLs in any gene or tissue.

Finally, we used model organism phenotypes measured by the International Mouse Phenotyping Consortium (IMPC) [42] and the Saccharomyces Cerevisiae Morphological Database (SCMD) [43]. To map ortholog genes from IMPC and SCMD to human variants, we used orthology annotations of gene orthologs, and eQTLs from GTEx. Thus, variants annotated as associated with a mouse or yeast phenotype are those that are annotated as *cis* eQTLs in any tissue for a gene whose ortholog in mouse or yeast is associated with that phenotype. The reference set for this analysis was variants annotated as *cis* eQTLs for genes that are not associated with mouse or yeast phenotypes.

Genome-wide significant pleiotropy loci

To detect loci with a genome-wide significant pleiotropy, we used the LD-corrected two-component pleiotropy score (P_m^{LD} and P_n^{LD}) computed on the dataset of 372 heritable traits from UK Biobank described above. We used LD clumping as implemented in PLINK to cluster linked variants, with an r^2 threshold of 0.1, a distance threshold of 100 kb, and *P-value* thresholds of 5 x 10^{-8} for genome-wide significance and 0.05 for nominal significance. The resulting loci were assigned to genes using 1) localization of variants within a gene, as annotated by Ensembl Variant Effect predictor, and 2) annotation as a *cis* eQTL in any tissue, as annotated by GTEx. We submitted the resulting list of genome-wide significant genes to DAVID for enrichment analysis [44,45] (https://david.ncifcrf.gov).

Permutation-based null model for replication analysis

In general, we should expect only 5% of loci to replicate by chance in each replication dataset; however, it is possible that this number might increase because of polygenicity in the underlying GWAS statistics and the resulting inflation in our pleiotropy score, which may cause substantially more than 5% of the genome to be assigned $P < 0.05$. To correct for this, we performed random permutations of the whitened Z-scores independently for each trait and used these permuted Z-scores to compute our LD-corrected pleiotropy score components (P_m^{LD} and P_n^{LD}). This generates a null expectation that preserves the polygenicity and inflation within each dataset. For both datasets, our null model expected that 5% of loci for P_m^{LD} loci and 6% of loci for P_n^{LD} should replicate. The fraction that replicated in the actual data was substantially higher (**Figure 5**).

References

1. Prokopenko, I. *et al.* A central role for GRB10 in regulation of islet function in man. *PLoS Genet.* **10**, e1004235 (2014).
2. Bradfield, J. P. *et al.* A genome-wide association meta-analysis identifies new childhood obesity loci. *Nat. Genet.* **44**, 526–531 (2012).
3. Major Depressive Disorder Working Group of the Psychiatric GWAS Consortium *et al.* A mega-analysis of genome-wide association studies for major depressive disorder. *Mol. Psychiatry* **18**, 497–511 (2013).
4. van der Valk, R. J. P. *et al.* A novel common variant in DCST2 is associated with length in early life and height in adulthood. *Hum. Mol. Genet.* **24**, 1155–1168 (2015).
5. Liu, J. Z. *et al.* Association analyses identify 38 susceptibility loci for inflammatory bowel disease and highlight shared genetic risk across populations. *Nat. Genet.* **47**, 979–986 (2015).
6. Schizophrenia Working Group of the Psychiatric Genomics Consortium. Biological insights from 108 schizophrenia-associated genetic loci. *Nature* **511**, 421–427 (2014).
7. Hibar, D. P. *et al.* Common genetic variants influence human subcortical brain structures. *Nature* **520**, 224–229 (2015).
8. Taal, H. R. *et al.* Common variants at 12q15 and 12q24 are associated with infant head circumference. *Nat. Genet.* **44**, 532–538 (2012).
9. Wood, A. R. *et al.* Defining the role of common variation in the genomic and biological architecture of adult human height. *Nat. Genet.* **46**, 1173–1186 (2014).
10. Global Lipids Genetics Consortium *et al.* Discovery and refinement of loci associated with lipid levels. *Nat. Genet.* **45**, 1274–1283 (2013).
11. Pattaro, C. *et al.* Genetic associations at 53 loci highlight cell types and biological pathways relevant for kidney function. *Nat Commun* **7**, 10023 (2016).

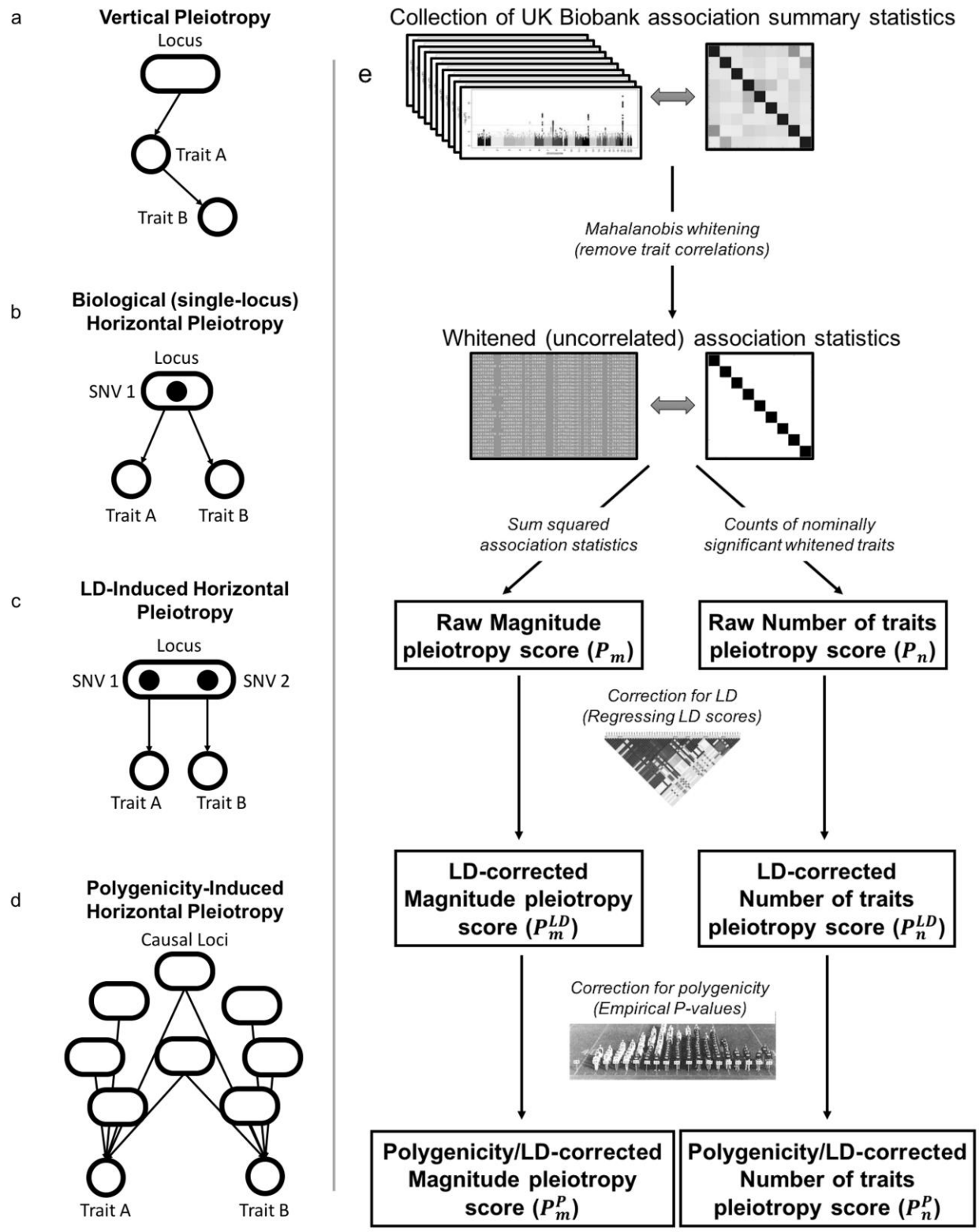
12. Locke, A. E. *et al.* Genetic studies of body mass index yield new insights for obesity biology. *Nature* **518**, 197–206 (2015).
13. International Consortium for Blood Pressure Genome-Wide Association Studies *et al.* Genetic variants in novel pathways influence blood pressure and cardiovascular disease risk. *Nature* **478**, 103–109 (2011).
14. Köttgen, A. *et al.* Genome-wide association analyses identify 18 new loci associated with serum urate concentrations. *Nat. Genet.* **45**, 145–154 (2013).
15. Felix, J. F. *et al.* Genome-wide association analysis identifies three new susceptibility loci for childhood body mass index. *Hum. Mol. Genet.* **25**, 389–403 (2016).
16. Cousminer, D. L. *et al.* Genome-wide association and longitudinal analyses reveal genetic loci linking pubertal height growth, pubertal timing and childhood adiposity. *Hum. Mol. Genet.* **22**, 2735–2747 (2013).
17. Horikoshi, M. *et al.* Genome-wide associations for birth weight and correlations with adult disease. *Nature* **538**, 248–252 (2016).
18. Okbay, A. *et al.* Genome-wide association study identifies 74 loci associated with educational attainment. *Nature* **533**, 539–542 (2016).
19. Okada, Y. *et al.* Genetics of rheumatoid arthritis contributes to biology and drug discovery. *Nature* **506**, 376–381 (2014).
20. Cousminer, D. L. *et al.* Genome-wide association study of sexual maturation in males and females highlights a role for body mass and menarche loci in male puberty. *Hum. Mol. Genet.* **23**, 4452–4464 (2014).
21. Tobacco and Genetics Consortium. Genome-wide meta-analyses identify multiple loci associated with smoking behavior. *Nat. Genet.* **42**, 441–447 (2010).
22. Cross-Disorder Group of the Psychiatric Genomics Consortium. Identification of risk loci with shared effects on five major psychiatric disorders: a genome-wide analysis. *Lancet* **381**, 1371–1379 (2013).

23. Scott, R. A. *et al.* Large-scale association analyses identify new loci influencing glycemic traits and provide insight into the underlying biological pathways. *Nat. Genet.* **44**, 991–1005 (2012).
24. Manning, A. K. *et al.* Meta-analysis of Gene-Environment interaction: joint estimation of SNP and SNP×Environment regression coefficients. *Genet Epidemiol* **35**, 11–18 (2011).
25. the CARDIoGRAMplusC4D Consortium. A comprehensive 1000 Genomes-based genome-wide association meta-analysis of coronary artery disease. *Nat Genet* **47**, 1121–1130 (2015).
26. Morris, A. P. *et al.* Large-scale association analysis provides insights into the genetic architecture and pathophysiology of type 2 diabetes. *Nat. Genet.* **44**, 981–990 (2012).
27. Psychiatric GWAS Consortium Bipolar Disorder Working Group. Large-scale genome-wide association analysis of bipolar disorder identifies a new susceptibility locus near ODZ4. *Nat. Genet.* **43**, 977–983 (2011).
28. Stolk, L. *et al.* Meta-analyses identify 13 loci associated with age at menopause and highlight DNA repair and immune pathways. *Nat. Genet.* **44**, 260–268 (2012).
29. Lambert, J. C. *et al.* Meta-analysis of 74,046 individuals identifies 11 new susceptibility loci for Alzheimer's disease. *Nat. Genet.* **45**, 1452–1458 (2013).
30. Dehghan, A. *et al.* Meta-analysis of genome-wide association studies in >80 000 subjects identifies multiple loci for C-reactive protein levels. *Circulation* **123**, 731–738 (2011).
31. Neale, B. M. *et al.* Meta-analysis of genome-wide association studies of attention-deficit/hyperactivity disorder. *J Am Acad Child Adolesc Psychiatry* **49**, 884–897 (2010).
32. Dupuis, J. *et al.* New genetic loci implicated in fasting glucose homeostasis and their impact on type 2 diabetes risk. *Nat. Genet.* **42**, 105–116 (2010).
33. Shungin, D. *et al.* New genetic loci link adipose and insulin biology to body fat distribution. *Nature* **518**, 187–196 (2015).

34. Dastani, Z. *et al.* Novel loci for adiponectin levels and their influence on type 2 diabetes and metabolic traits: a multi-ethnic meta-analysis of 45,891 individuals. *PLoS Genet.* **8**, e1002607 (2012).
35. Perry, J. R. B. *et al.* Parent-of-origin-specific allelic associations among 106 genomic loci for age at menarche. *Nature* **514**, 92–97 (2014).
36. van der Harst, P. *et al.* Seventy-five genetic loci influencing the human red blood cell. *Nature* **492**, 369–375 (2012).
37. Shin, S.-Y. *et al.* An atlas of genetic influences on human blood metabolites. *Nat Genet* **46**, 543–550 (2014).
38. Bulik-Sullivan, B. K. *et al.* LD Score regression distinguishes confounding from polygenicity in genome-wide association studies. *Nature Genetics* **47**, 291 (2015).
39. McLaren, W. *et al.* The Ensembl Variant Effect Predictor. *Genome Biology* **17**, 122 (2016).
40. Bernstein, B. E. *et al.* The NIH Roadmap Epigenomics Mapping Consortium. *Nature Biotechnology* **28**, 1045 (2010).
41. GTEx Consortium *et al.* Genetic effects on gene expression across human tissues. *Nature* **550**, 204–213 (2017).
42. Dickinson, M. E. *et al.* High-throughput discovery of novel developmental phenotypes. *Nature* **537**, 508–514 (2016).
43. Saito, T. L. *et al.* SCMD: Saccharomyces cerevisiae Morphological Database. *Nucleic Acids Res.* **32**, D319-322 (2004).
44. Huang, D. W., Sherman, B. T. & Lempicki, R. A. Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. *Nucleic Acids Res.* **37**, 1–13 (2009).
45. Huang, D. W., Sherman, B. T. & Lempicki, R. A. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat Protoc* **4**, 44–57 (2009).

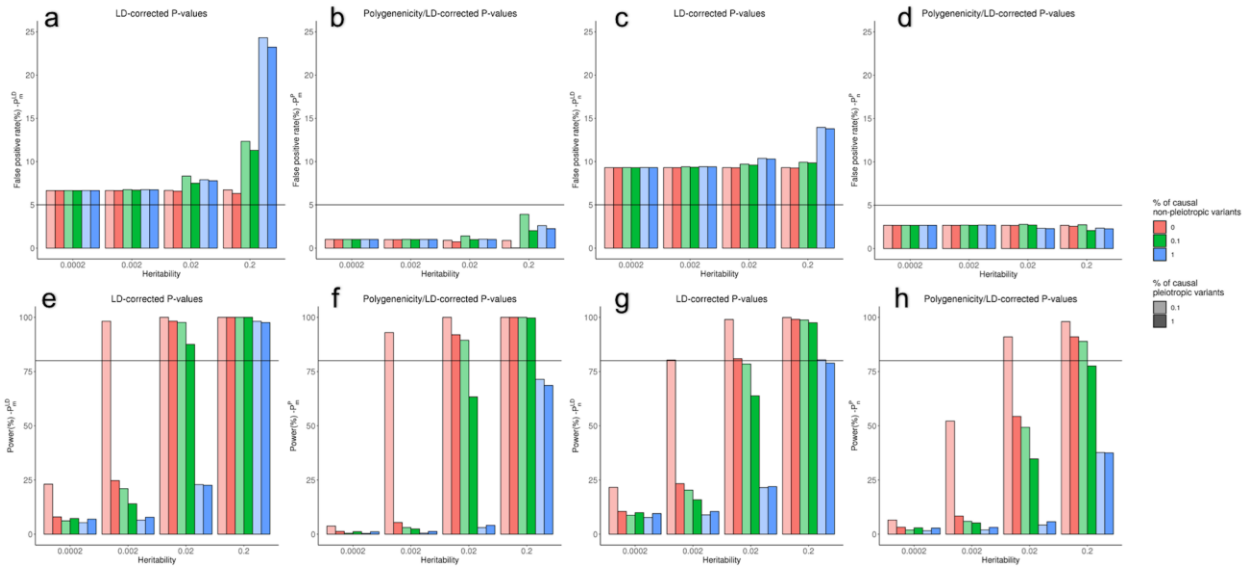
Figures

Figure 1: Schematic of different types of pleiotropy (a-d) and two-component pleiotropy score method (e).



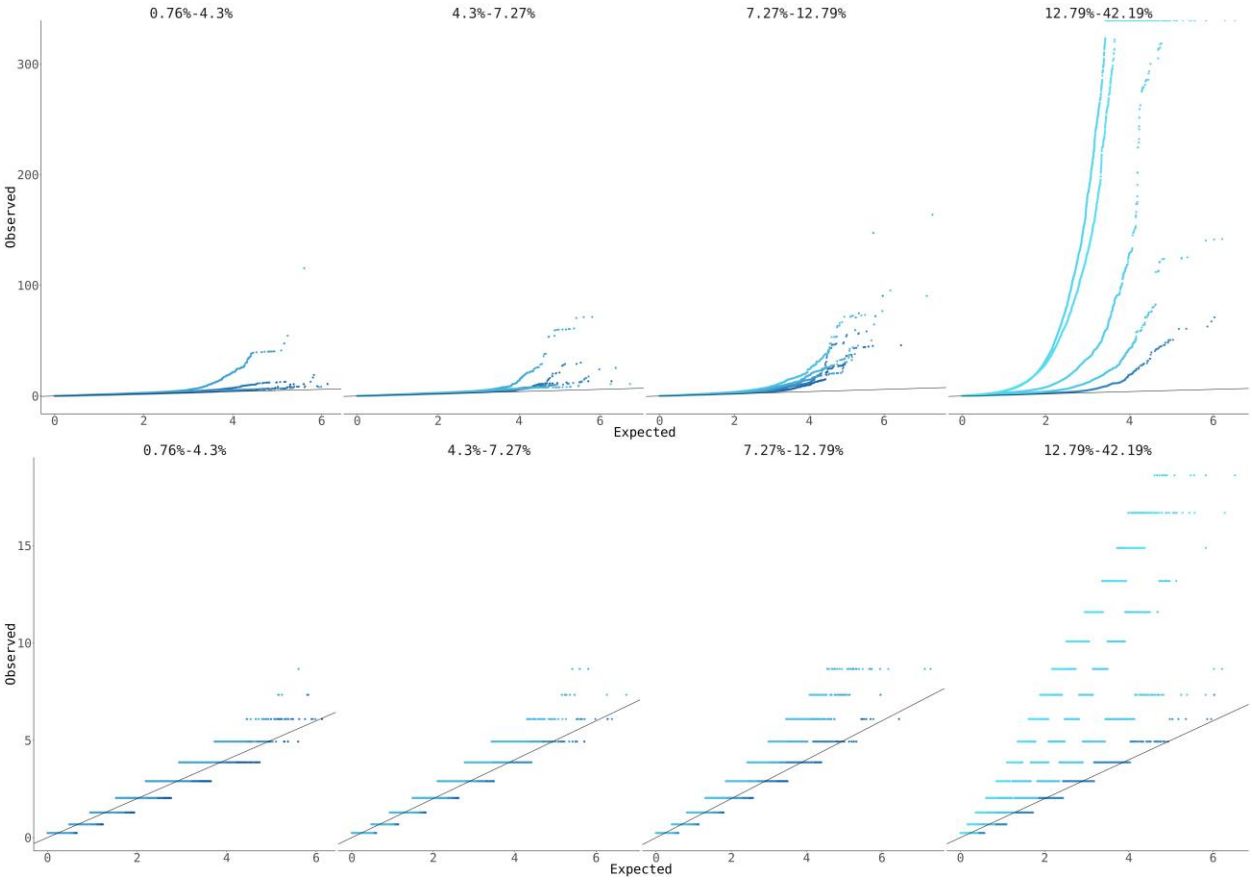
Previous studies distinguish between vertical pleiotropy (a), where effects on one trait are mediated through effects on another trait, and horizontal pleiotropy (b), where effects on multiple traits are independent. In this study, we also discuss linkage disequilibrium (LD)-induced pleiotropy (c), where two linked SNVs have independent effects on different traits, and polygenicity-induced pleiotropy (d), where two highly polygenic traits have an overlap in their polygenic footprint. In panel e, we (i) collect association statistics from the UK Biobank, (ii) process them using Mahalanobis whitening, (iii) compute the two components of our pleiotropy score (P_m and P_n) based on the whitened association statistics, (iv) use LD scores to correct for LD-induced pleiotropy ($P_m^{LD}$and P_n^{LD}), and (v) use permutation-based *P*-values to correct for polygenic architecture (P_m^P and P_n^P).

Figure 2: Simulation study showing false positive rate (a,b,c,d) and power (e,f,g,h) of two-component pleiotropy score.



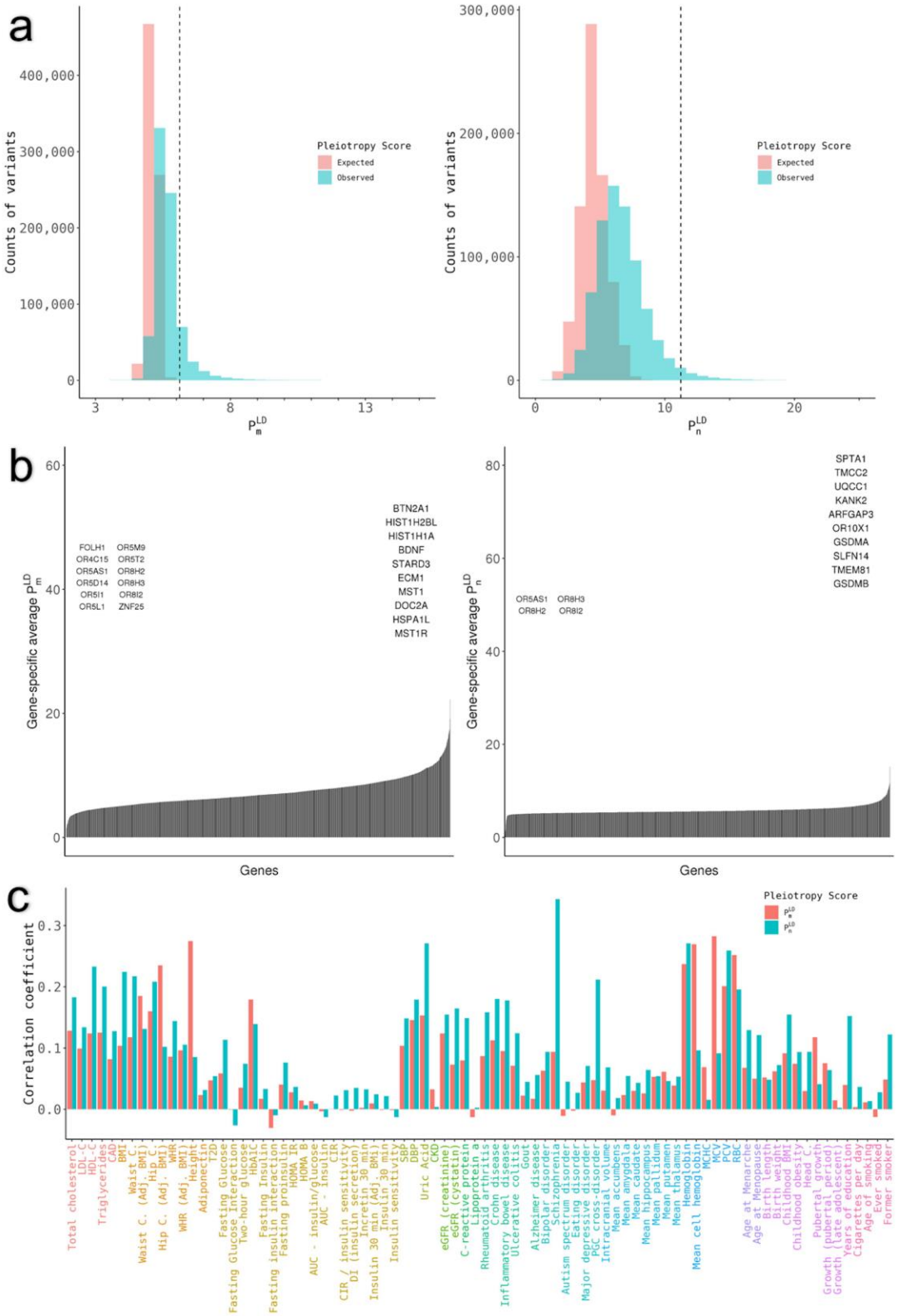
Top row shows performance on non-pleiotropic simulated variants (black line shows 5% false positive rate); bottom row shows performance on pleiotropic variants (black line shows 80% power). Simulations were run for both P_m^{LD} (left) and P_n^{LD} (right), and both without correction for polygenicity (a,c,e,g) and with the correction (b,f,d,h), with per-variant heritability ranging from 0.0002 to 0.2, proportion of non-pleiotropic causal loci ranging from 0 to 1%, and proportion of pleiotropic causal loci ranging from 0.1% to 1%. Our method has good power to detect pleiotropy for highly heritable traits, though its power is reduced by extreme polygenicity. Extreme polygenicity also increases the false positive rate, though this effect is corrected by our polygenicity correction.

Figure 3: Quantile-quantile (Q-Q) plots showing the inflation of the pleiotropy score as a function of polygenicity.



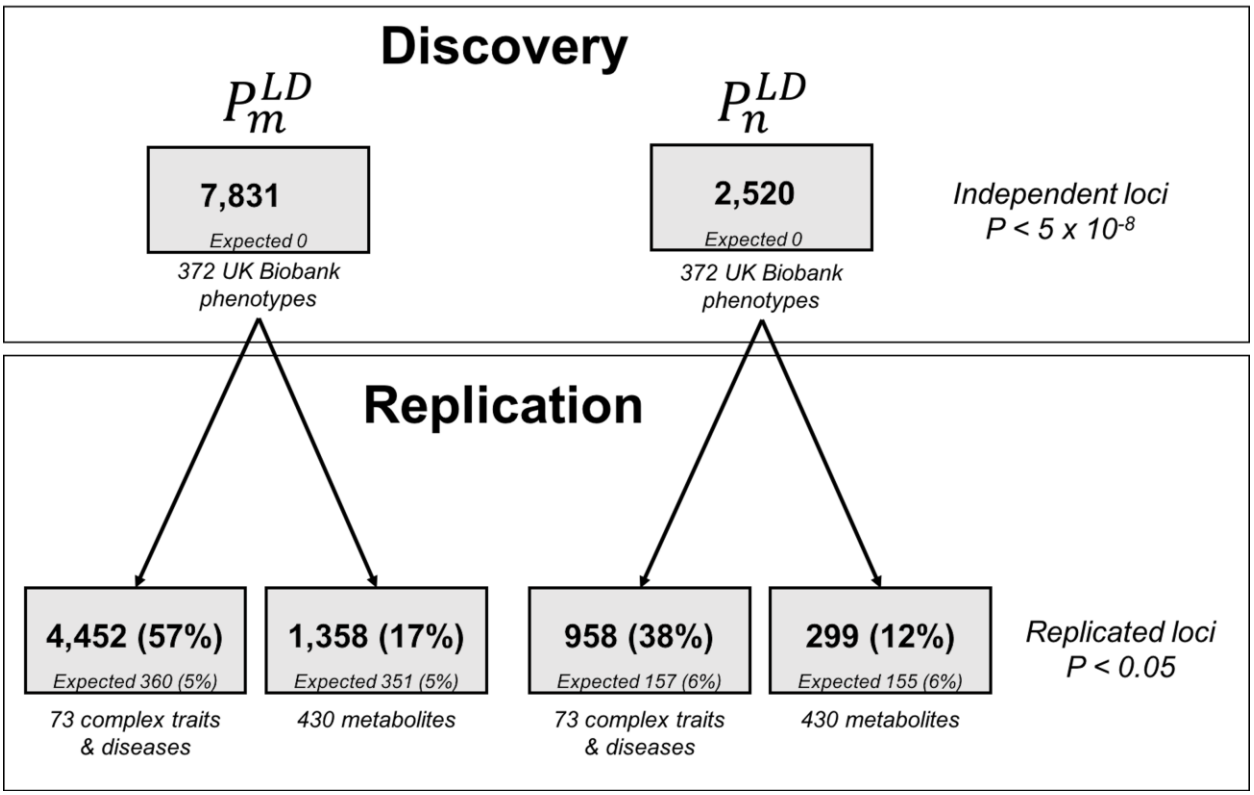
Variants are stratified into 4 batches of about 80 traits each by heritability, and then subdivided into 5 batches of about 20 traits each by polygenicity, as measured by corrected genomic inflation factor λ_{GC}^{c}. Darker shades represent low polygenicity and lighter shades represent high polygenicity. All panels show -$\log_{10}$ transformed *P-values*. The black lines show the expected value under the null hypothesis.

Figure 4: Distribution of the pleiotropy score among variants (a), genes (b), and traits (c).



Panel a shows the global distribution of P_m^{LD} (left) and P_n^{LD} (right) for the 767,057 tested variants. The expected distribution under the null hypothesis of no pleiotropy is shown in red and the observed distribution is shown in blue. The vertical line represents the value of the pleiotropy score corresponding to genome-wide significance ($P < 5 \times 10^{-8}$). 1,769 (P_m^{LD}) and 643 (P_n^{LD}) variants are not represented for the sake of clarity, because they have extreme values for the pleiotropy score. Panel b shows the distribution of the average pleiotropy score for coding variants in each gene for P_m^{LD} (left) and P_n^{LD} (right). The top ten genes are represented on the right side of the plots, whereas genes with a pleiotropy score of 0 are represented on the left side of the plots. Panel c shows correlation coefficients between the absolute value of Z-scores of 82 complex traits and diseases and the pleiotropy score among variants that are genome-wide significant for the pleiotropy score ($P < 5 \times 10^{-8}$ for P_m^{LD} and P_n^{LD} respectively).

Figure 5: Replication analysis for the genome-wide pleiotropy study.



We used 372 UK Biobank heritable medical traits as our discovery dataset, and independent datasets of 73 complex traits and diseases and 430 blood metabolites as replication datasets. In each case, expected fraction of replication was empirically determined using a permutation analysis.

Tables

Table 1: Functional enrichment analysis of pleiotropy score.

			P_m^{LD}	P_n^{LD}
Variant effect predictor		UTR	**+0.24 (±0.01); *P* = 1.72x10^{-234}**	**+0.69 (±0.02); *P* = 2.16x10^{-236}**
		coding synonymous	**+0.24 (±0.01); *P* = 2.49x10^{-99}**	**+0.61 (±0.03); *P* = 1.92x10^{-76}**
		non synonymous	**+0.19 (±0.01); *P* = 3.82x10^{-82}**	**+0.48 (±0.03); *P* = 3.62x10^{-62}**
Roadmap Epigenomics	Promoter	H327ac	**+0.20 (±0.01); *P* < 10^{-308}**	**+0.54 (±0.01); *P* < 10^{-308}**
		H3K27me3	**+0.02 (±0.01); *P* = 1.40x10^{-18}**	+0.01 (±0.01); *P* = 0.4
		Active TSS	**+0.20 (±0.02); *P* = 1.42x10^{-36}**	**+0.54 (±0.04); *P* = 8.56x10^{-34}**
		Promoter Upstream TSS	**+0.16 (±0.01); *P* = 4.44x10^{-130}**	**+0.43 (±0.02); *P* = 4.33x10^{-103}**
		Promoter Downstream TSS 1	**+0.35 (±0.01); *P* = 1.87x10^{-220}**	**+0.92 (±0.03); *P* = 3.59x10^{-197}**
		Promoter Downstream TSS 2	**+0.30 (±0.01); *P* = 2.70x10^{-203}**	**+0.86 (±0.03); *P* = 3.44x10^{-210}**
	Transcription	Transcribed - 5' preferential	**+0.29 (±0.01); *P* < 10^{-308}**	**+0.88 (±0.01); *P* < 10^{-308}**
		Strong transcription	**+0.38 (±0.01); *P* < 10^{-308}**	**+1.10 (±0.01); *P* < 10^{-308}**

Transcription & regulation	Transcribed - 3' preferential	**+0.29 (±0.01); $P < 10^{-308}$**	**+0.82 (±0.01); $P < 10^{-308}$**
	Weak transcription	**+0.21 (±0.01); $P < 10^{-308}$**	**+0.60 (±0.01); $P < 10^{-308}$**
	Transcribed & regulatory (Prom/Enh)	**+0.36 (±0.01); $P < 10^{-308}$**	**+1.00 (±0.02); $P < 10^{-308}$**
	Transcribed 5' preferential and Enh	**+0.35 (±0.01); $P < 10^{-308}$**	**+1.00 (±0.01); $P < 10^{-308}$**
	Transcribed 3' preferential and Enh	**+0.33 (±0.01); $P < 10^{-308}$**	**+0.92 (±0.02); $P < 10^{-308}$**
	Transcribed and Weak Enhancer	**+0.32 (±0.01); $P < 10^{-308}$**	**+0.97 (±0.01); $P < 10^{-308}$**
Active enhancer	Active Enhancer 1	**+0.13 (±0.01); $P = 4.54 \times 10^{-295}$**	**+0.32 (±0.01); $P = 5.1 \times 10^{-216}$**
	Active Enhancer 2	**+0.11 (±0.01); $P = 2.64 \times 10^{-294}$**	**+0.28 (±0.01); $P = 5.63 \times 10^{-238}$**
	Active Enhancer Flank	**+0.11 (±0.01); $P < 10^{-308}$**	**+0.29 (±0.01); $P = 6.06 \times 10^{-270}$**
Weak enhancer	Weak Enhancer 1	**+0.07 (±0.01); $P = 2.79 \times 10^{-89}$**	**+0.16 (±0.01); $P = 6.89 \times 10^{-60}$**
	Weak Enhancer 2	**+0.08 (±0.01); $P < 10^{-308}$**	**+0.23 (±0.01); $P = 6.52 \times 10^{-291}$**
	Primary H3K27ac possible Enhancer	**+0.09 (±0.01); $P = 2.72 \times 10^{-259}$**	**+0.24 (±0.01); $P = 1.53 \times 10^{-187}$**
	Primary DNase	**+0.03 (±0.01); $P = 3.83 \times 10^{-21}$**	**+0.05 (±0.01); $P = 1.11 \times 10^{-7}$**
	ZNF genes & repeats	**+0.08 (±0.01); $P = 1.29 \times 10^{-7}$**	**+0.20 (±0.04); $P = 6.9 \times 10^{-7}$**

	Heterochromatin	**-0. 20 (±0.01); *P* < 10^{-308}**	**-0.61 (±0.01); *P* < 10^{-308}**
	Poised Promoter	**+0.05 (±0.01); *P* = $1.03x10^{-35}$**	**+0.09 (±0.01); *P* = $2.27x10^{-16}$**
	Bivalent Promoter	**+0.17 (±0.01); *P* = $1.28x10^{-93}$**	**+0.51 (±0.03); *P* = $6.29x10^{-88}$**
	Repressed Polycomb	**+0.04 (±0.01); *P* = $5.77x10^{-42}$**	**+0.06 (±0.01); *P* = $1.48x10^{-11}$**
	Quiescent/Low	**-0.41 (±0.01); *P* < 10^{-308}**	**-1.20 (±0.01); *P* < 10^{-308}**
GTEx - number of genes the variant is an eQTL for	eGenes<10	**+0.11 (±0.01); *P* = $6.78x10^{-186}$**	**+0.28 (±0.01); *P* = $1.04x10^{-140}$**
	eGenes>10 & <15	**+0.19 (±0.01); *P* = $4.72x10^{-114}$**	**+0.52 (±0.02); *P* = $6.84x10^{-99}$**
	eGenes>15 & <20	**+0.31 (±0.02); *P* = $7.98x10^{-52}$**	**+0.88 (±0.06); *P* = $5.38x10^{-47}$**
	eGenes>20	**+0.66 (±0.06); *P* = $3.40x10^{-27}$**	**+2.07 (±0.18); *P* = $1.35x10^{-30}$**
GTEx - number of tissues the variant is an eQTL for	eTissue<30	**+0.10 (±0.01); *P* = $1.84x10^{-151}$**	**+0.26 (±0.01); *P* = $1.26x10^{-114}$**
	eTissue>30 & <35	**+0.21 (±0.01); *P* = $3.70x10^{-187}$**	**+0.54 (±0.02); *P* = $6.80x10^{-147}$**
	eTissue>35 & <40	**+0.36 (±0.02); *P* = $1.11x10^{-82}$**	**+1.13 (±0.06); *P* = $4.24x10^{-92}$**
	eTissue>40	**+0.35 (±0.05); *P* = $2,42x10^{-13}$**	**+0.97 (±0.14); *P* = $7.08x10^{-12}$**
International Mouse Phenotyping Consortium	Phenotypes > 1	**+0.06 (±0.01); *P* = $1.91x10^{-6}$**	**+0.19 (±0.04); *P* = $2.70x10^{-7}$**
Saccharomyces cerevisiae	Phenotypes > 1	**+0.09 (±0.01); *P* = $4.48x10^{-17}$**	**+0.26 (±0.03); *P* = $1.53x10^{-18}$**

Morphological Database

We grouped variants by (i) molecular function as annotated by Ensembl, (ii) predicted chromatin state as annotated by the NIH Roadmap Epigenomics Project, (iii) transcriptional effects as annotated by the NIH Genotype-Tissue Expression (GTex) Project, and (iv) effects on model organism phenotypes as annotated by the International Mouse Phenotyping Consortium (IMPC) and Saccharomyces Cerevisiae Morphological Database (SCMD). For each grouping, we computed the mean LD-corrected pleiotropy score and used a two-sample Student's t-test to determine whether the mean was significantly different from the baseline. We found (i) that coding regions have higher pleiotropy scores than noncoding regions, (ii) that active promoters and enhancers have the highest pleiotropy scores and quiescent and heterochromatin have the lowest, (iii) that variants that control expression of more genes in more tissues have higher pleiotropy scores, and (iv) that genes associated with more than one model organism phenotype have higher pleiotropy scores.