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

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

Understanding the nature and extent of horizontal pleiotropy, where one genetic variant has independent effects on multiple observable traits, is vitally important for our understanding of the genetic architecture of human phenotypes, as well as the design of genome-wide association studies (GWASs) and Mendelian randomization (MR) studies. Many recent studies have pointed to the existence of horizontal pleiotropy among human phenotypes, but the exact extent remains unknown, largely due to difficulty in disentangling the inherently correlated nature of observable traits. Here, we present a statistical framework to isolate and quantify horizontal pleiotropy in human genetic variation using a two-component pleiotropy score computed from summary statistic data derived from published GWASs. This score uses a statistical whitening procedure to remove correlations between observable traits and normalize effect sizes across all traits, and is able to detect horizontal pleiotropy under a range of different models in our simulations. When applied to real human phenotype data using association statistics for 1,564 traits measured in 337,119 individuals from the UK Biobank, our score detects a significant excess of horizontal pleiotropy. This signal of horizontal pleiotropy is pervasive throughout the human genome and across a wide range of phenotypes and biological functions, but is especially prominent in regions of high linkage disequilibrium and among phenotypes known to be highly polygenic and heterogeneous. Using our pleiotropy score, we identify thousands of loci with extreme levels of horizontal pleiotropy, a majority of which have never been previously reported in any published GWAS. This highlights an under-recognized class of genetic variation that has weak effects on many distinct phenotypes but no specific marked effect on any one phenotype. We show that a large fraction of these loci replicate using independent datasets of GWAS summary statistics. 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 (Zhan et al., 2018), of the genetic architectures of traits (Chesmore et al., 2017; Socrates et al., 2017), of the biology underlying common diseases (Solovieff et al., 2013) and of the dynamics of natural selection (Keightley and Hill, 1990). 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 (Paaby and Rockman, 2013; Tyler et al., 2009). 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. For example, a genetic variant that increases LDL cholesterol and has an additional effect on an unrelated trait, such as schizophrenia risk, may be an example of horizontal pleiotropy, as there is currently no known intrinsic causal relationship between these traits. 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 (Pickrell et al., 2016; Chesmore et al., 2017; Socrates et al., 2017; Kanai et al., 2018). 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 idea that horizontal pleiotropy may be ubiquitous has recently been advanced in the form of “network pleiotropy” (Boyle et al., 2017). According to this theory, traits are controlled by densely connected networks of biological regulators, and any perturbation in the network will affect all traits connected to that network. These effects fall off with increasing distance in the network, but can still be several significant steps away. This can lead to extreme polygenicity in many traits, since any one trait is controlled by all genes in the network, with a wide range of effect sizes depending on the position of the gene in the network. More importantly for our purposes, it has been suggested that this can also lead to widespread pleiotropy, since any variant in the network influences every trait connected to the network, even distantly. However,

there has been very little concrete evidence supporting this hypothesis, as most recent studies investigating pleiotropy have reported genetic correlations between specific pairs of traits or diseases, or have focused on genome-wide significant disease-associated single nucleotide variants (SNVs) only, rather than this kind of ubiquitous horizontal pleiotropy.

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 (Pasaniuc and Price, 2017).

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 the input phenotypes originate from 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 whitened (decorrelated) association Z-scores, we measure two related but distinct components of pleiotropy: the total magnitude of effect on independent traits ("total magnitude" score, denoted P_m) and the total number of independent traits affected by a variant ("number of traits" score, denoted P_n), both calculated after accounting for correlation between traits. The total magnitude score P_m is defined as the Euclidean norm of the vector of trait associations, or the square root of the sum of squared trait associations. Meanwhile, the number of traits score P_n is defined as the number of whitened traits with associations stronger than some arbitrarily chosen Z-score cutoff - we use Z-score ≥ 2, approximately corresponding to nominal significance ($P < 0.05$). More formally,

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

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

where z_i are the whitened variant-trait associations, l is the total number of traits, and H is the Heaviside step function. 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 (See **Supplemental Information**). In addition, we can compute *P-values* for the two components of our pleiotropy score based on the theoretically expected distributions of the scores under the null hypothesis of no pleiotropy - that is, that the effect of each variant on each whitened trait obeys an independent standard Gaussian distribution (**Figure 1**; **Supplemental Information**).

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 controlling 100 traits. 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 variants and forced them to have simultaneous associations with multiple traits, varying the number of traits (ν) and the strength of association (μ). The simulation study showed that both components of the pleiotropy score were well-powered to detect horizontal pleiotropy (**Table 1**). Under the null hypothesis of no added horizontal pleiotropy, the false positive rate was well controlled for both scores. In the presence of added horizontal pleiotropy, using a *P-value* cut-off of 0.05, our approach was powered to detect pleiotropy with effect size μ as small as 2 and number of traits ν as small as 10. With a more stringent *P-value* cutoff ($P < 5 \times 10^{-8}$, corresponding to the threshold for genome-wide significance), we could still detect pleiotropic variants with $\mu \geq 3$ and $\nu \geq 20$. As expected, both scores responded in the presence of horizontal pleiotropy, but the magnitude score P_m was much more sensitive to effect size parameter μ, while the number of traits score P_n responded similarly to perturbations of both parameters. Adding a realistic correlation structure between the traits neither caused any inflation of the false positive rate nor any loss of power to detect pleiotropy. Furthermore, the values of the score were very similar regardless of the presence or absence of correlation between the traits (**Table S1**). In contrast, the false positive rate was highly inflated when the decorrelation step was skipped in the presence of correlation (**Table S2**).

Genome-Wide Pleiotropy Study (GWPS) reveals pervasive pleiotropy

To apply our method to real human association data, we used GWAS association statistics for 1,564 medical traits measured in 337,119 individuals from the UK Biobank. We successfully computed our two-component pleiotropy score for 767,095 variants genome-wide. Since we can compute *P-values* for our two-component pleiotropy score for every variant genome-wide, we were able to use our pleiotropy score to conduct a genome-wide pleiotropy study (GWPS), by analogy to a standard GWAS (**Figure 1**; **Supplemental Information**).

Figure S1 shows the resulting quantile-quantile plots (Q-Q plots). We observed significant inflation for both the magnitude score P_m (λ_{GC} = 9.95) and the number of traits score P_n (λ_{GC} = 7.92). 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 S2**). Testing an alternative strategy for computing the phenotype-correlation matrix using a pruned set of SNVs (r^2 < 0.1) indicated comparable results (Pearson r = 0.78 and 0.91 for P_m and P_n respectively) to our strategy of using genetic correlations computed using all variants (**Figure S3**). Taken together, these observations point to a significant inflation of pleiotropy genome-wide, affecting not only most of the functional loci detectable by GWAS, but also a very large number of loci with weak effects on many different traits.

Pleiotropy is driven by linkage disequilibrium

We next investigated the origins of this pervasive signal of pleiotropy. The naive model of pleiotropy is that certain specific variants affect multiple traits due to having multiple biological functions. However, it is likely that pleiotropy can also arise as a consequence of certain features of genetic architecture (**Figure 1)**. One example of this is linkage disequilibrium (LD). Multiple variants in LD with each other that are causal to different traits will be associated with all of these traits, resulting in inflated association test statistics compared to the scenario where these variants are not in LD with each other. This results in the known phenomenon of LD inducing correlation in GWAS association statistics (Bulik-Sullivan et al., 2015a, 2015b). This situation may look indistinguishable from biological pleiotropy, where a single causal variant is associated with two or more traits, each through a separate biological pathway. We therefore hypothesize that our pleiotropy score may detect higher levels of pleiotropy in areas of high LD.

To examine the contribution of LD to our pleiotropy score, we divided the genome into 15 equal-sized bins of about 50,000 variants each based on LD score in European samples from the 1000 Genomes Project (Bulik-Sullivan et al., 2015b) (**Supplemental Information**). We observed a strong relationship between LD score and inflation of the pleiotropy score, with stronger LD corresponding to higher levels of pleiotropy (**Figure 2 a-b**; **Table S3**). This demonstrates that LD drives a portion of our signal of widespread pleiotropy. Using an LD-corrected version of our pleiotropy score (**Supplemental Information**) reduced the level of pleiotropy observed genome-wide, but we still observed substantial inflation in our pleiotropy score genome-wide (λ_{GC} = 5.99 for P_m and 4.94 for P_n).

Pleiotropy is driven by polygenicity

Another feature of genetic architecture that may result in elevated pleiotropy is extreme polygenicity. For example, it has been estimated that approximately 100,000 independent loci are causal for height in humans (Boyle et al., 2017). 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 or omnigenic 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. Using our simulation framework and our pleiotropy score, we tested this model and detected significant inflation of pleiotropy solely from including multiple extremely polygenic traits, without any specifically added pleiotropy (**Table S4**). This demonstrates that widespread horizontal pleiotropy is a consequence of extreme polygenicity amongst multiple phenotypes and that pleiotropy induced in this way is detected by our method.

Similarly to our test for LD, to test whether polygenicity drives pleiotropy, we calculated the polygenicity of the same 1,564 traits from the UK Biobank for which LD score regression could be performed. We measured polygenicity using a version of the genomic inflation factor corrected using LD score λ_{GC}^{c} (Bulik-Sullivan et al., 2015b). We then used λ_{GC}^{c} to divide these 1,564 traits into 15 equal-sized bins of about 100 traits each (**Supplemental Information**), and calculated the two-component pleiotropy score and *P-values* for each component independently for every variant in the genome using each of these bins of traits. We observed substantial inflation in both scores in the 8th through 15th high-polygenicity bins, corresponding to 35.56 ≥ λ_{GC}^{c} > 2.04 (P_m) and 24.25 ≥ λ_{GC}^{c} > 1.09 (P_n), and lower inflation for traits with polygenicity below this level (**Figure 2 c-d**; **Table S5**). This suggests that extreme polygenicity drives pleiotropy as predicted by our simulations, and that this has a profound 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. For a more accurate representation of pleiotropy, we limited our further analyses to only significantly heritable traits within the UK Biobank (heritability significant at Bonferroni-corrected cut-off $P < 0.05/1,564$), since association statistics for these traits are more likely to represent true biological associations. This left a total of 367 traits for analysis. Restricting in this way increased the inflation observed above for the LD-corrected pleiotropy score: λ_{GC} = 14.55 for P_m and λ_{GC} = 10.66 for P_n.

Figure 3a shows the distribution of pleiotropy score for independent SNVs (LD pruned to a threshold of $r^2 < 0.1$). The median variant has an LD-corrected P_n score of 24.3 and an LD-corrected P_m score of 20.3. Under the null hypothesis that effects of SNVs on different traits are drawn from independent Gaussian distributions, we would expect a median P_n score of 17.0 and a P_m score of 19.1. This represents a large excess in the number of traits affected by each variant, and a smaller but still highly significant excess in total magnitude pleiotropic effect. 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 3a**), 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 pleiotropy score **(Table 2; Supplemental Information)**. Using annotations from Ensembl Variant Effect Predictor (McLaren et al., 2016), 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 (Bernstein et al., 2010), 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 indicates higher levels of pleiotropy, whether that regulation is repressive or activating.

We also used data from the Genotype-Tissue Expression (GTEx Consortium et al., 2017) project to measure the connection between transcriptional effects and our pleiotropy score **(Table 2)**. 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: r = 0.027, $P = 5.7 \times 10^{-13}$; P_n: r = 0.029, $P = 1.2 \times 10^{-14}$) and tissues (P_m: r = 0.051, $P = 2.4 \times 10^{-43}$; P_n: r = 0.052, $P = 4.9 \times 10^{-44}$) where the variant was annotated as an eQTL, showing that our pleiotropy score is related to transcriptional measures of pleiotropy. Similarly, we found higher values of our pleiotropy score in eQTLs identified as master regulators (P_m: +1.93 (±0.56), $P = 5.62 \times 10^{-4}$; P_n: +6.47 (±1.92), $P = 7.82 \times 10^{-4}$) (Tong et al., 2017).

Finally, we used model organism phenotypes measured by the International Mouse Phenotyping Consortium (IMPC) (Dickinson et al., 2016) and the Saccharomyces Cerevisiae Morphological Database (SCMD) (Saito et al., 2004) to test whether our pleiotropy score predicts functional genes in model organisms. 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 (P_m: +0.18 (±0.04), $P = 5.65 \times 10^{-5}$; P_n: +0.75 (±0.17), $P = 5.87 \times 10^{-6}$ for mouse; and P_m: +0.26 (±0.04), $P = 1.56 \times 10^{-12}$; P_n: +0.94 (±0.14), $P = 5.31 \times 10^{-12}$ for yeast). This indicates that our pleiotropy score is also related to pleiotropy in model organisms.

Genome-wide pleiotropy study identifies novel biological loci

Much as standard GWAS is used with a *P-value* threshold of 5×10^{-8} to find genome-wide significant associations, our GWPS methodology can identify individual variants that have a genome-wide significant level of horizontal pleiotropy, and therefore, extreme levels of horizontal pleiotropy. Using the LD-corrected number of traits score P_n we identified 43,424 variants in 6,645 independent loci with a genome-wide significant level of horizontal pleiotropy (**Supplemental Information**). The magnitude of pleiotropy score P_m showed a larger number of significant loci (96,364 variants in N = 12,375 loci), indicating that these loci may have a

significant effect on only a few traits without being associated with an unexpectedly high number of traits (**Table S6**). Strikingly, about a third of loci for the number of traits score P_n identified in this way (2,216 of 6,645) do not reach genome-wide significance for any of the 367 component traits used to calculate our pleiotropy score. Furthermore, a majority of these highly pleiotropic loci (4,234 of 6,645) have 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. Given the lack of specific significant associations, it is difficult to characterize precisely which biological processes these may be. Functional enrichment analysis on genes overlapping these loci includes a wide range of biological functions, including genes involved in neurogenesis, synapse formation, and production and regulation of neurotransmitters; components of the cytoskeleton; transcription factors; and components of intracellular signaling cascades (**Table S7)**. 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 scores: previously published GWASs and meta-analyses for 82 human complex traits and diseases, which we collected and curated manually from the literature (**Supplemental Information**) (Verbanck et al., 2018); and a previously published study of 453 blood metabolites measured in 7,824 European adults (Shin et al., 2014). For all variants covered by the UK Biobank, we were able to compute our pleiotropy score independently using these two datasets (**Figure 4**). 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 causes 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. This generates a null expectation that preserves the polygenicity and inflation within each dataset. Under this null model, we did indeed find that an inflated fraction of loci replicated; however, the fraction that replicated in the actual data was still substantially higher. In the traits and diseases dataset, our null model expected that 22% of P_m loci and 15% of loci for P_n should replicate; in the actual data, we found that 51% of P_m loci and 42% of P_n loci replicated. Likewise, in the blood metabolites dataset, our null model expected that 7% of P_m loci and 9% of P_n loci should replicate; in the actual data, we found that 11% of P_m loci and 12% of P_n loci replicated. 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.

Pleiotropic loci are enriched for specific complex traits and diseases

To characterize the phenotypic associations of these loci, we used our replication dataset of published GWAS summary statistics for 82 human quantitative traits and diseases (**Supplementary Information**). **Figure 3c** shows the correlation between our pleiotropy score 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; childhood obesity, pubertal growth, schizophrenia, educational attainment, and age at menarche. 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, eating disorder, lipoprotein(a), and gout. 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 S8**).

Potential application to Genomic Medicine

Our observation that horizontal pleiotropy is widespread has important clinical implications for genomic medicine, particularly in the areas of genome editing and drug target discovery and validation (Visscher and Yang, 2016). In genome editing, understanding the pleiotropic effects of corrected mutations can help to avoid unexpected secondary phenotypic effects. Furthermore, genetic associations have been shown to provide predictive evidence for main indications and adverse side effects for therapeutics in clinical trials (Nelson et al., 2015; Nguyen et al., 2018). To this end, we produced a catalogue of pleiotropic effects by gene using a combination of variants localized to each gene and variants controlling expression of each gene in any tissue. Using this catalogue, we observed a wide range of pleiotropy across all of the genes. **Figure 3b** lists the top ten genes and the genes scoring zero for each component of the pleiotropy score.

As a case example of the clinical utility of such a catalogue, we restricted to a selected list of genes for coronary artery disease (CAD) (Dewey et al., 2017; Khera and Kathiresan, 2017; Stitziel et al., 2017) and type 2 diabetes (T2D) (Thomsen and Gloyn, 2017) that are current drug targets and have been shown to have prior human genetic evidence for these traits. We observed that these genes exhibited a range of levels of pleiotropy, with exceptionally high pleiotropy in the T2D drug targets *PPARG*, *KCNJ11*, and *GCKR* and substantially elevated pleiotropy in the CAD drug targets *NPC1L1*, *ANGPTL4*, and *LPA*, in addition to generally high horizontal pleiotropy across all genes **(Table 3).**

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. This means that loci identified as pleiotropic by our analyses have effects on multiple biological processes which are themselves largely independent. After applying both components of our pleiotropy score to 367 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 LD; 3) horizontal pleiotropy is driven by extreme polygenicity of traits; 4) horizontal pleiotropy occurs predominantly amongst variants with weak effects on multiple traits, but no singular strong effect on any one trait; 5) 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; 6) there are thousands of loci that exhibit extreme levels of horizontal pleiotropy, a majority of which have no previously reported associations; 7) pleiotropic loci are enriched in specific complex traits including body mass index, height and educational attainment; and 8) a number of drug target genes for coronary artery disease and type 2 diabetes have significant elevated levels of horizontal pleiotropy.

Our findings are in keeping with several recent studies that have found abundant pleiotropy in the genome (Wang et al., 2010; Sivakumaran et al., 2011; Pickrell et al., 2016; Chesmore et al., 2017; Kanai et al., 2018). 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 (Bowden et al., 2017; Pickrell et al., 2016; Verbanck et al., 2018). Furthermore, the current study has evaluated horizontal pleiotropy in human genetic variation at the genome-wide level 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 (Bowden et al., 2015, 2017; Verbanck et al., 2018); our results reinforce the importance of these methods. Indeed, 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 (Boyle et al., 2017), 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. 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. Connected to this idea of network pleiotropy, Boyle, Li, and Pritchard propose an "omnigenic" model of genetic architecture, wherein traits are controlled by an extremely large number of genetic loci spread throughout the whole genome, with each trait being controlled by as many as 100,000 independent loci.

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 Boyle, Li, and Pritchard 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 our signal of widespread horizontal pleiotropy is most pronounced in the most highly polygenic traits, and restricting to oligogenic traits significantly attenuates our signal. It may be the case that horizontal pleiotropy is only truly widespread among the most complex and polygenic subset of human traits.

Third, we have identified thousands of novel pleiotropic loci, with extreme levels of horizontal pleiotropy but no previously known primary associations with any individual trait. Individually, each of these loci represents an understudied but biologically important gene or region, and functional investigation of the most highly pleiotropic of these novel is likely to reveal important unknown biology. As a group, they represent an under-recognized class of genetic variation exhibiting broad pleiotropy but weak primary effects. These loci are largely invisible to standard GWAS methodology, pointing to the need for GWAS methods that account for effects on multiple traits simultaneously (Maier et al., 2018; Turley et al., 2018). They also demonstrate an underappreciated fact about fine mapping of causal genes and variants, in both GWAS and traditional genetics: most currently-available tools for variant and gene prioritization, including CADD (Kircher et al., 2014), SIFT (Kumar et al., 2009), PolyPhen-2 (Adzhubei et al., 2013), phyloP (Pollard et al., 2010), and the ExAC and gnoMAD constraint scores (Lek et al., 2016), prioritize using signals of natural selection, under the assumption that natural selection is a good proxy for overall biological function. However, our results demonstrate that overall biological function is not necessarily a good proxy for any specific trait. On the contrary, many of the specific traits and diseases we examined show little correlation with pleiotropic biological function, and a majority of loci with high levels of pleiotropic biological function show no association with any specific trait. This demonstrates the need for caution when interpreting the output of these prioritization tools, and the need for trait-specific prioritization tools.

Fourth and finally, it is now becoming common to use genetic associations, particularly from GWAS, to identify potential drug targets. There is substantial evidence that drug trials supported by genetic associations are more likely to succeed (Nelson et al., 2015; Nguyen et al., 2018); however, the fact that we are able to detect widespread horizontal pleiotropy using the same GWAS association summary statistics indicates that many drug targets supported by GWAS may have multiple off-target associations. Our score could potentially be useful as an additional screening step in drug development, to select drug targets with a smaller chance of having undesirable off-target effects.

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, LD 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.

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

None.

References

Adzhubei, I., Jordan, D.M., and Sunyaev, S.R. (2013). Predicting Functional Effect of Human Missense Mutations Using PolyPhen-2. Curr Protoc Hum Genet *0 7*, Unit7.20.

Bernstein, B.E., Stamatoyannopoulos, J.A., Costello, J.F., Ren, B., Milosavljevic, A., Meissner, A., Kellis, M., Marra, M.A., Beaudet, A.L., Ecker, J.R., et al. (2010). The NIH Roadmap Epigenomics Mapping Consortium. Nature Biotechnology *28*, 1045.

Bowden, J., Smith, G.D., and Burgess, S. (2015). Mendelian randomization with invalid instruments: effect estimation and bias detection through Egger regression. Int. J. Epidemiol. *44*, 512–525.

Bowden, J., Del Greco M, F., Minelli, C., Davey Smith, G., Sheehan, N., and Thompson, J. (2017). A framework for the investigation of pleiotropy in two-sample summary data Mendelian randomization. Statist. Med. n/a-n/a.

Boyle, E.A., Li, Y.I., and Pritchard, J.K. (2017). An Expanded View of Complex Traits: From Polygenic to Omnigenic. Cell *169*, 1177–1186.

Bulik-Sullivan, B., Finucane, H.K., Anttila, V., Gusev, A., Day, F.R., Loh, P.-R., ReproGen Consortium, Psychiatric Genomics Consortium, Genetic Consortium for Anorexia Nervosa of the Wellcome Trust Case Control Consortium 3, Duncan, L., et al. (2015a). An atlas of genetic correlations across human diseases and traits. Nat Genet *47*, 1236–1241.

Bulik-Sullivan, B.K., Loh, P.-R., Finucane, H.K., Ripke, S., Yang, J., Consortium, S.W.G. of the P.G., Patterson, N., Daly, M.J., Price, A.L., and Neale, B.M. (2015b). LD Score regression distinguishes confounding from polygenicity in genome-wide association studies. Nature Genetics *47*, 291.

Chesmore, K., Bartlett, J., and Williams, S.M. (2017). The ubiquity of pleiotropy in human disease. Hum Genet 1–6.

Dewey, F.E., Gusarova, V., Dunbar, R.L., O'Dushlaine, C., Schurmann, C., Gottesman, O., McCarthy, S., Van Hout, C.V., Bruse, S., Dansky, H.M., et al. (2017). Genetic and Pharmacologic Inactivation of ANGPTL3 and Cardiovascular Disease. N. Engl. J. Med. *377*, 211–221.

Dickinson, M.E., Flenniken, A.M., Ji, X., Teboul, L., Wong, M.D., White, J.K., Meehan, T.F., Weninger, W.J., Westerberg, H., Adissu, H., et al. (2016). High-throughput discovery of novel developmental phenotypes. Nature *537*, 508–514.

GTEx Consortium, Laboratory, Data Analysis &Coordinating Center (LDACC)—Analysis Working Group, Statistical Methods groups—Analysis Working Group, Enhancing GTEx (eGTEx) groups, NIH Common Fund, NIH/NCI, NIH/NHGRI, NIH/NIMH, NIH/NIDA,

Biospecimen Collection Source Site—NDRI, et al. (2017). Genetic effects on gene expression across human tissues. Nature *550*, 204–213.

Kanai, M., Akiyama, M., Takahashi, A., Matoba, N., Momozawa, Y., Ikeda, M., Iwata, N., Ikegawa, S., Hirata, M., Matsuda, K., et al. (2018). Genetic analysis of quantitative traits in the Japanese population links cell types to complex human diseases. Nat. Genet. *50*, 390–400.

Keightley, P.D., and Hill, W.G. (1990). Variation maintained in quantitative traits with mutation–selection balance: pleiotropic side-effects on fitness traits. Proc. R. Soc. Lond. B *242*, 95–100.

Khera, A.V., and Kathiresan, S. (2017). Genetics of coronary artery disease: discovery, biology and clinical translation. Nat. Rev. Genet. *18*, 331–344.

Kircher, M., Witten, D.M., Jain, P., O'Roak, B.J., Cooper, G.M., and Shendure, J. (2014). A general framework for estimating the relative pathogenicity of human genetic variants. Nat. Genet. *46*, 310–315.

Kumar, P., Henikoff, S., and Ng, P.C. (2009). Predicting the effects of coding non-synonymous variants on protein function using the SIFT algorithm. Nat Protoc *4*, 1073–1081.

Lek, M., Karczewski, K.J., Minikel, E.V., Samocha, K.E., Banks, E., Fennell, T., O'Donnell-Luria, A.H., Ware, J.S., Hill, A.J., Cummings, B.B., et al. (2016). Analysis of protein-coding genetic variation in 60,706 humans. Nature *536*, 285–291.

Maier, R.M., Zhu, Z., Lee, S.H., Trzaskowski, M., Ruderfer, D.M., Stahl, E.A., Ripke, S., Wray, N.R., Yang, J., Visscher, P.M., et al. (2018). Improving genetic prediction by leveraging genetic correlations among human diseases and traits. Nature Communications *9*, 989.

McLaren, W., Gil, L., Hunt, S.E., Riat, H.S., Ritchie, G.R.S., Thormann, A., Flicek, P., and Cunningham, F. (2016). The Ensembl Variant Effect Predictor. Genome Biology *17*, 122.

Nelson, M.R., Tipney, H., Painter, J.L., Shen, J., Nicoletti, P., Shen, Y., Floratos, A., Sham, P.C., Li, M.J., Wang, J., et al. (2015). The support of human genetic evidence for approved drug indications. Nat Genet *47*, 856–860.

Nguyen, P.A., Deaton, A.M., Nioi, P., and Ward, L.D. (2018). Phenotypes associated with genes encoding drug targets are predictive of clinical trial side effects. BioRxiv 285858.

Paaby, A.B., and Rockman, M.V. (2013). The many faces of pleiotropy. Trends in Genetics *29*, 66–73.

Pasaniuc, B., and Price, A.L. (2017). Dissecting the genetics of complex traits using summary association statistics. Nature Reviews Genetics *18*, 117–127.

Pickrell, J.K., Berisa, T., Liu, J.Z., Ségurel, L., Tung, J.Y., and Hinds, D.A. (2016). Detection and interpretation of shared genetic influences on 42 human traits. Nat. Genet. *48*, 709–717.

Pollard, K.S., Hubisz, M.J., Rosenbloom, K.R., and Siepel, A. (2010). Detection of nonneutral substitution rates on mammalian phylogenies. Genome Res *20*, 110–121.

Saito, T.L., Ohtani, M., Sawai, H., Sano, F., Saka, A., Watanabe, D., Yukawa, M., Ohya, Y., and Morishita, S. (2004). SCMD: Saccharomyces cerevisiae Morphological Database. Nucleic Acids Res. *32*, D319-322.

Shin, S.-Y., Fauman, E.B., Petersen, A.-K., Krumsiek, J., Santos, R., Huang, J., Arnold, M., Erte, I., Forgetta, V., Yang, T.-P., et al. (2014). An atlas of genetic influences on human blood metabolites. Nat Genet *46*, 543–550.

Sivakumaran, S., Agakov, F., Theodoratou, E., Prendergast, J.G., Zgaga, L., Manolio, T., Rudan, I., McKeigue, P., Wilson, J.F., and Campbell, H. (2011). Abundant pleiotropy in human complex diseases and traits. Am. J. Hum. Genet. *89*, 607–618.

Socrates, A., Bond, T., Karhunen, V., Auvinen, J., Rietveld, C., Veijola, J., Jarvelin, M.-R., and O'Reilly, P. (2017). Polygenic risk scores applied to a single cohort reveal pleiotropy among hundreds of human phenotypes. BioRxiv 203257.

Solovieff, N., Cotsapas, C., Lee, P.H., Purcell, S.M., and Smoller, J.W. (2013). Pleiotropy in complex traits: challenges and strategies. Nat. Rev. Genet. *14*, 483–495.

Stitziel, N.O., Khera, A.V., Wang, X., Bierhals, A.J., Vourakis, A.C., Sperry, A.E., Natarajan, P., Klarin, D., Emdin, C.A., Zekavat, S.M., et al. (2017). ANGPTL3 Deficiency and Protection Against Coronary Artery Disease. J. Am. Coll. Cardiol. *69*, 2054–2063.

Thomsen, S.K., and Gloyn, A.L. (2017). Human genetics as a model for target validation: finding new therapies for diabetes. Diabetologia *60*, 960–970.

Tong, P., Monahan, J., and Prendergast, J.G.D. (2017). Shared regulatory sites are abundant in the human genome and shed light on genome evolution and disease pleiotropy. PLOS Genetics *13*, e1006673.

Turley, P., Walters, R.K., Maghzian, O., Okbay, A., Lee, J.J., Fontana, M.A., Nguyen-Viet, T.A., Wedow, R., Zacher, M., Furlotte, N.A., et al. (2018). Multi-trait analysis of genome-wide association summary statistics using MTAG. Nature Genetics *50*, 229–237.

Tyler, A.L., Asselbergs, F.W., Williams, S.M., and Moore, J.H. (2009). Shadows of complexity: what biological networks reveal about epistasis and pleiotropy. BioEssays *31*, 220–227.

Verbanck, M., Chen, C.-Y., Neale, B., and Do, R. (2018). Detection of widespread horizontal pleiotropy in causal relationships inferred from Mendelian randomization between complex traits and diseases. Nature Genetics 1.

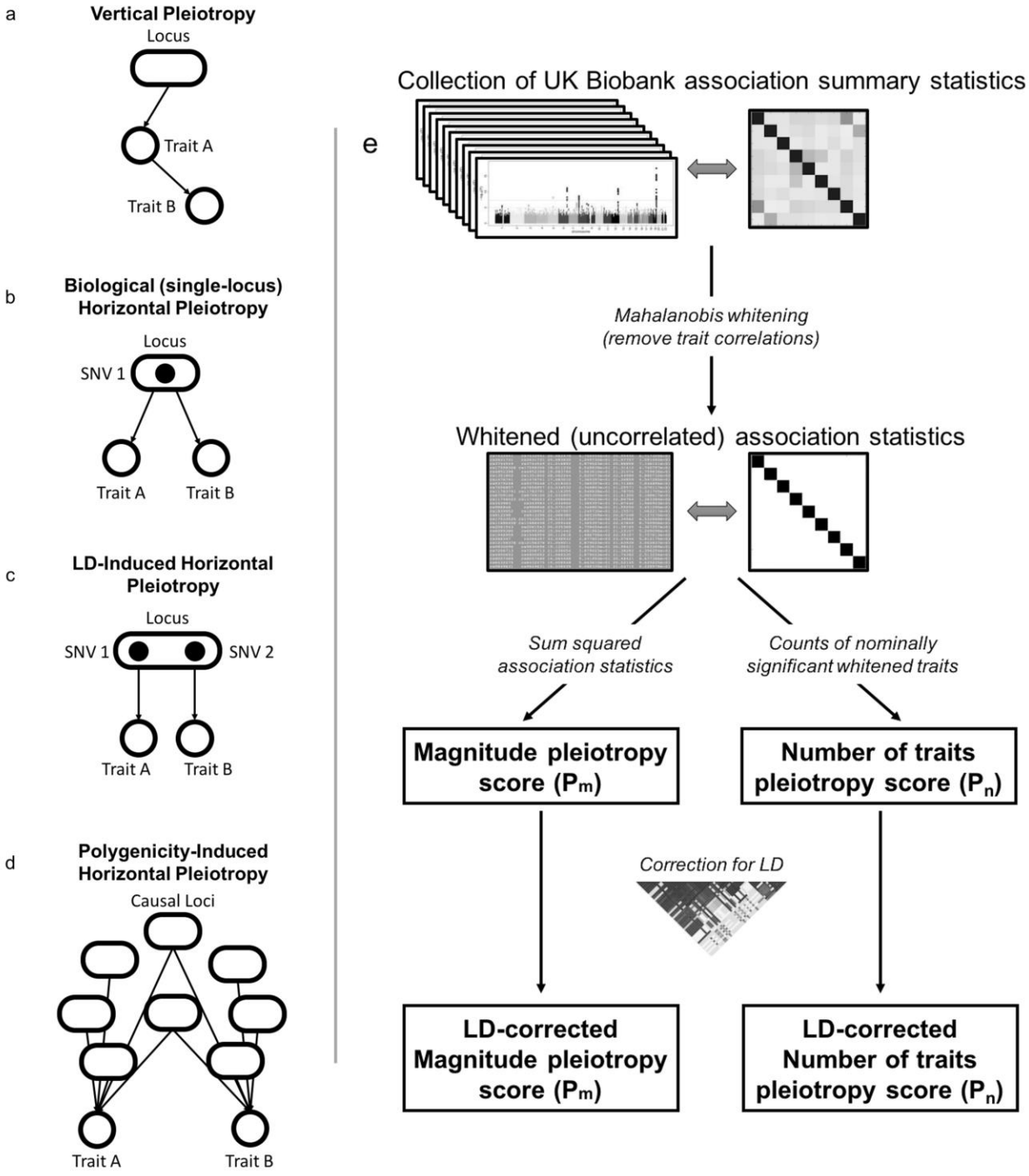
Visscher, P.M., and Yang, J. (2016). A plethora of pleiotropy across complex traits. Nat. Genet. *48*, 707–708.

Wang, Z., Liao, B.-Y., and Zhang, J. (2010). Genomic patterns of pleiotropy and the evolution of complexity. PNAS *107*, 18034–18039.

Zhan, J., Arking, D.E., and Bader, J.S. (2018). Discovering patterns of pleiotropy in genome-wide association studies. BioRxiv 273540.

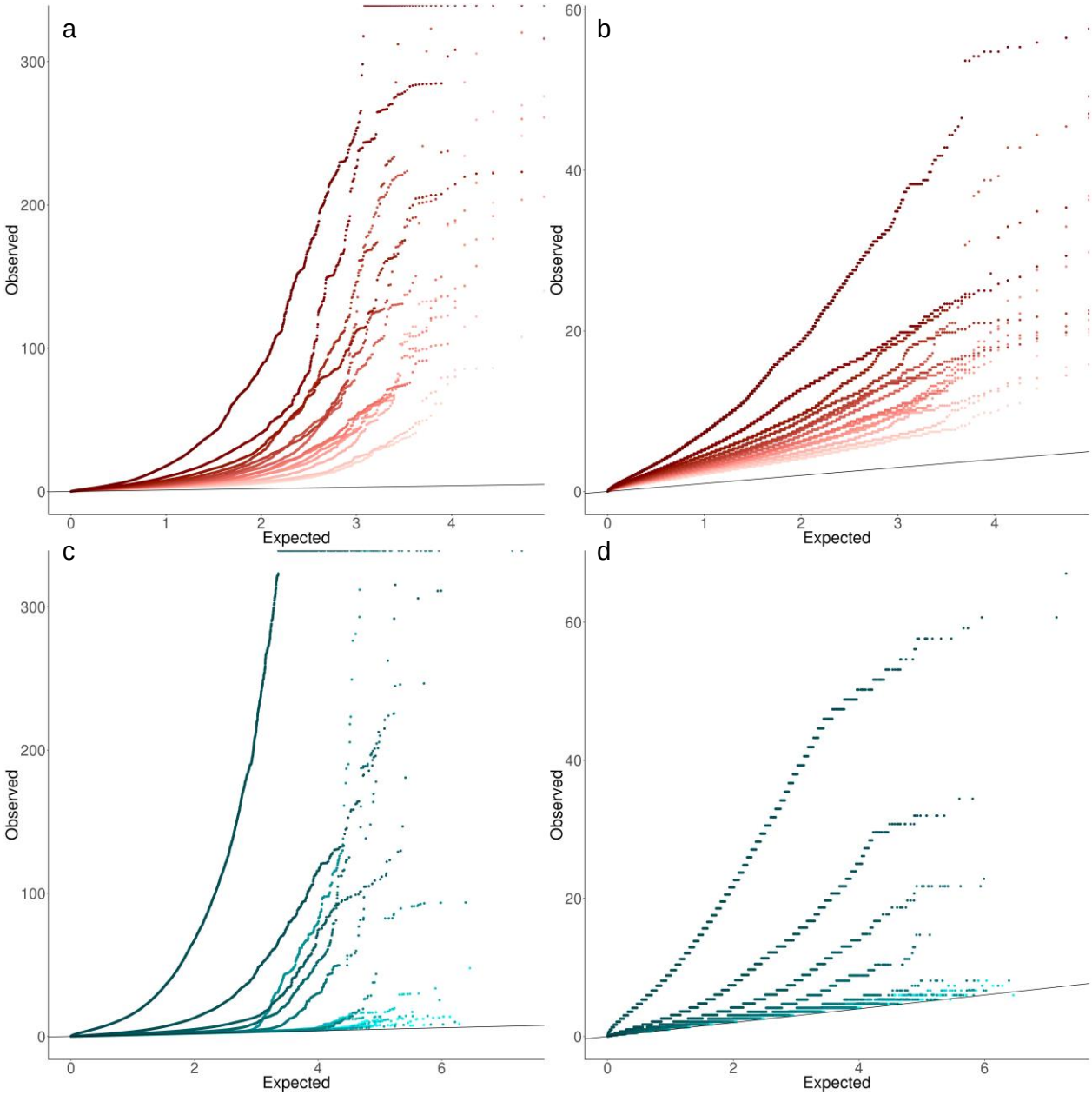
Figures

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



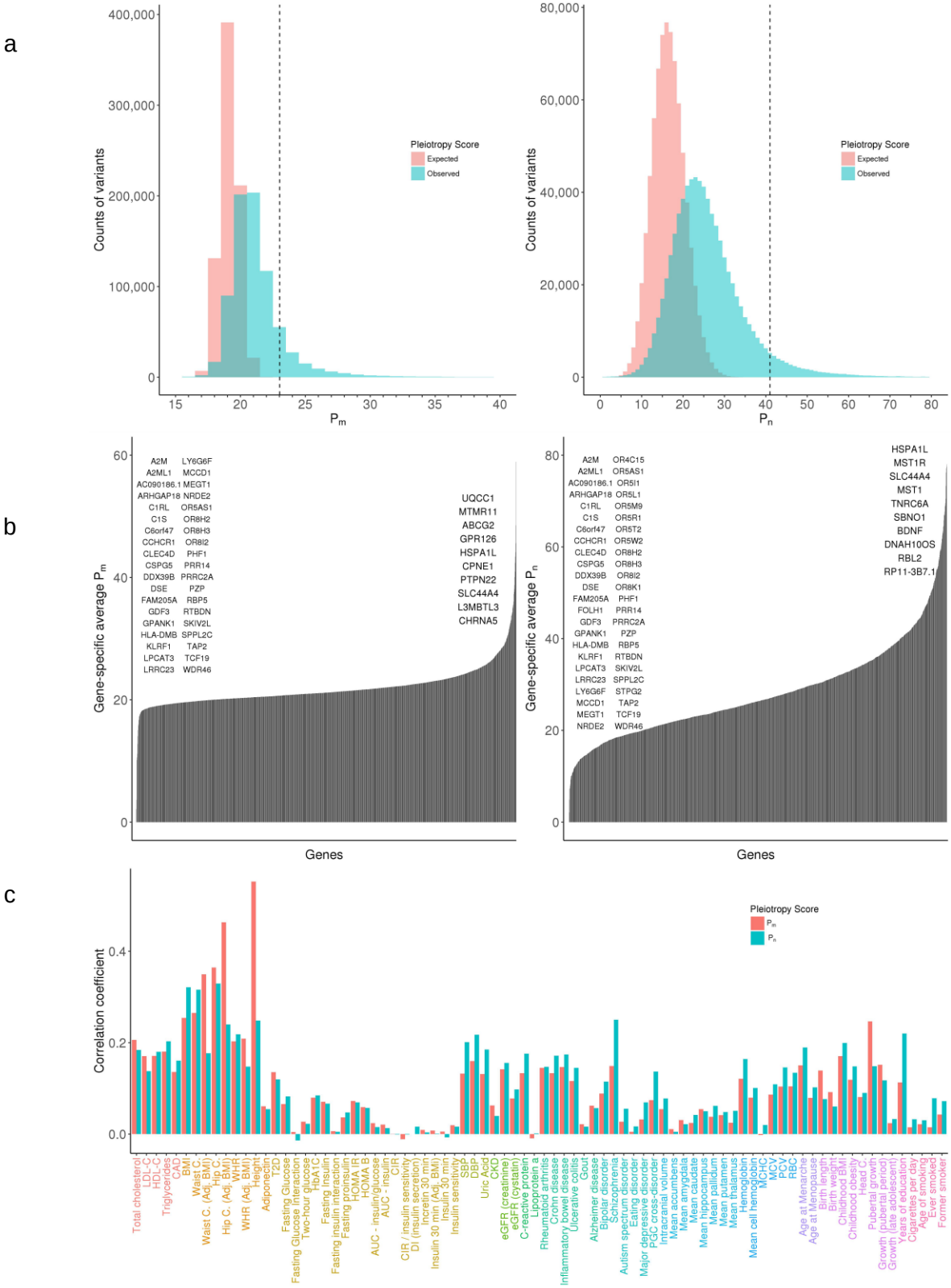
Previous studies distinguish between vertical pleiotropy (a), where effects on one trait are mediated through effects on another trait, and horizontal pleiotropy (b), where effects on multiple traits are independent. In this study, we also discuss linkage disequilibrium (LD)-induced pleiotropy (c), where 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, and (iv) use LD scores to correct for LD-induced pleiotropy.

Figure 2: Quantile-quantile (Q-Q) plots showing the inflation of the pleiotropy score as a function of linkage disequilibrium (LD score) and polygenicity.



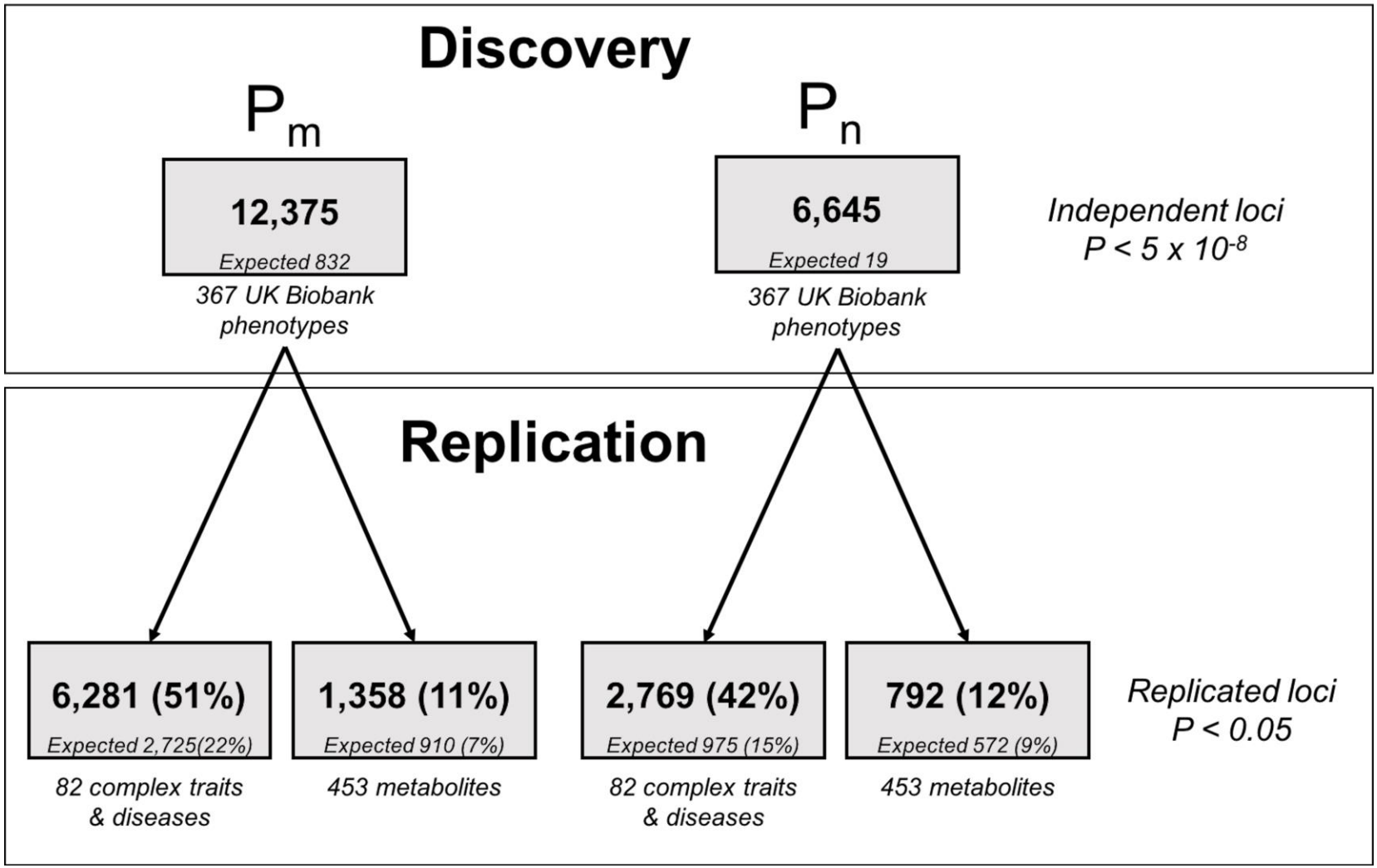
In panels a (P_m) and b (P_n), variants are stratified into 15 batches of about 50,000 variants each by LD score, with light red representing low LD score and dark red representing high LD score. In panels c (P_m) and d (P_n), traits are stratified into 15 batches of about 100 traits each by polygenicity, as measured by corrected genomic inflation factor λ_{GC}^{c}, with light blue representing low polygenicity and dark blue representing high polygenicity. All panels show -$\log_{10}$ transformed *P-values*. The black lines show the expected value under the null hypothesis.

Figure 3: Distribution of the pleiotropy score among variants (a), genes (b), and traits (c).



Panel a shows the global distribution of P_m (left) and P_n (right) for the 762,847 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) and 643 (P_n) 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 (left) and P_n (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 and P_n respectively).

Figure 4: Replication analysis for the genome-wide pleiotropy study.



We used 367 UK Biobank heritable medical traits as our discovery dataset, and independent datasets of 82 complex traits and diseases and 453 blood metabolites as replication datasets. In each case, expected fraction of replication was empirically determined using a permutation analysis.

Tables

Table 1: False positive rate and power to detect horizontal pleiotropy in simulations.

Correlation	% Pleiotropy	Pleiotropy magnitude (μ)	Pleiotropy number traits (ν)	False positive rate				Power			
				0.05 cut-off		$5x10^{-8}$ cut-off		0.05 cut-off		$5x10^{-8}$ cut-off	
				P_m	P_n	P_m	P_n	P_m	P_n	P_m	P_n
No	0	0	0	5.00	8.59	$6.25x10^{-6}$	$8.75x10^{-6}$	-	-	-	-
No	0.1	2	10	5.00	8.59	$6.26x10^{-6}$	$1.25x10^{-5}$	79.08	72.81	0.45	0.04
No	0.1	3	10	5.00	8.59	$5.01x10^{-6}$	$7.51x10^{-6}$	99.91	99.14	41.27	0.81
No	1	2	20	5.00	8.59	$3.79x10^{-6}$	$1.14x10^{-5}$	99.64	98.65	25.79	5.04
No	1	3	20	5.00	8.59	$3.79x10^{-6}$	$1.39x10^{-5}$	100	100	99.90	78.83
Yes	0	0	0	5.00	8.59	$5.00x10^{-6}$	$8.75x10^{-6}$	-	-	-	-
Yes	0.1	2	10	5.00	8.59	$3.75x10^{-6}$	$1.00x10^{-5}$	79.40	71.82	0.44	0.06
Yes	0.1	3	10	5.00	8.59	$1.25x10^{-5}$	$1.88x10^{-5}$	99.91	99.03	41.20	3.77
Yes	1	2	20	5.00	8.60	$3.79x10^{-6}$	$5.05x10^{-6}$	99.63	98.24	25.74	5.63
Yes	1	3	20	5.00	8.60	$5.05x10^{-6}$	$1.39x10^{-5}$	100	100	99.89	83.83

We simulate ten scenarios, varying (i) the presence or absence of a correlation structure, (ii) the fraction of variants with pleiotropic effects, (iii) the number of traits v affected by a pleiotropic variant, and (iv) the mean magnitude μ of pleiotropic effects. Using both nominal significance ($P < 0.05$) and genome-wide significance ($P < 5 \times 10^{-8}$) cutoffs, our two-component pleiotropy score has controlled false positive rates and is well-powered to detect horizontal pleiotropy under all scenarios.

Table 2: Functional enrichment analysis of pleiotropy score.

			P_m	P_n
Variant effect predictor		UTR	**+0.75 (±0.03); P = 2.82x10^{-178}**	**+2.86 (±0.1); P = 1.87x10^{-186}**
		coding synonymous	**+0.61 (±0.04); P = 3.25x10^{-50}**	**+2.26 (±0.15); P = 2.11x10^{-49}**
		non synonymous	**+0.51 (±0.04); P = 2.37x10^{-44}**	**+1.83 (±0.14); P = 5.59x10^{-41}**
Roadmap	Promoter	H327ac	**+0.62 (±0.01); $P < 10^{-308}$**	**+2.3 (±0.02); $P < 10^{-308}$**
		H3K27me3	**+0.06 (±0.01); P = 9.98x10^{-15}**	+0.09 (±0.03); P = 2.6x10^{-3}
		Active TSS	**+0.64 (±0.06); P = 9.32x10^{-27}**	**+2.46 (±0.21); P = 8.9x10^{-31}**
		Promoter Upstream TSS	**+0.51 (±0.02); P = 1.51x10^{-97}**	**+1.87 (±0.09); P = 7.45x10^{-92}**
		Promoter Downstream TSS 1	**+1.07 (±0.04); P = 4.24x10^{-155}**	**+4 (±0.14); P = 6.3x10^{-169}**
		Promoter Downstream TSS 2	**+0.95 (±0.04); P = 2.45x10^{-160}**	**+3.61 (±0.13); P = 8.86x10^{-173}**
	Transcription	Transcribed - 5' preferential	**+0.95 (±0.02); $P < 10^{-308}$**	**+3.73 (±0.06); $P < 10^{-308}$**
		Strong transcription	**+1.11 (±0.02); $P < 10^{-308}$**	**+4.37 (±0.08); $P < 10^{-308}$**
		Transcribed - 3' preferential	**+0.88 (±0.01); $P < 10^{-308}$**	**+3.44 (±0.05); $P < 10^{-308}$**
		Weak transcription	**+0.68 (±0.01); $P < 10^{-308}$**	**+2.59 (±0.03); $P < 10^{-308}$**
	Transcription & regulation	Transcribed & regulatory (Prom/Enh)	**+1.07 (±0.02); $P < 10^{-308}$**	**+4.16 (±0.08); $P < 10^{-308}$**
		Transcribed 5' preferential and Enh	**+1.06 (±0.02); $P < 10^{-308}$**	**+4.09 (±0.07); $P < 10^{-308}$**

		Transcribed 3' preferential and Enh	**+0.99 (±0.02); $P < 10^{-308}$**	**+3.83 (±0.08); $P < 10^{-308}$**
		Transcribed and Weak Enhancer	**+1.02 (±0.02); $P < 10^{-308}$**	**+4.04 (±0.07); $P < 10^{-308}$**
	Active enhancer	Active Enhancer 1	**+0.39 (±0.01); $P = 5.32 \times 10^{-209}$**	**+1.38 (±0.05); $P = 5.7 \times 10^{-182}$**
		Active Enhancer 2	**+0.33 (±0.01); $P = 8.97 \times 10^{-220}$**	**+1.16 (±0.04); $P = 1.29 \times 10^{-184}$**
		Active Enhancer Flank	**+0.34 (±0.01); $P = 1.86 \times 10^{-247}$**	**+1.23 (±0.04); $P = 5.28 \times 10^{-226}$**
	Weak enhancer	Weak Enhancer 1	**+0.19 (±0.01); $P = 3.42 \times 10^{-56}$**	**+0.63 (±0.04); $P = 1.04 \times 10^{-44}$**
		Weak Enhancer 2	**+0.28 (±0.01); $P = 8.45 \times 10^{-305}$**	**+1.01 (±0.03); $P = 4.32 \times 10^{-268}$**
		Primary H3K27ac possible Enhancer	**+0.31 (±0.01); $P = 5.83 \times 10^{-204}$**	**+1.08 (±0.04); $P = 2.72 \times 10^{-172}$**
		Primary DNase	**+0.12 (±0.01); $P = 1.08 \times 10^{-27}$**	**+0.35 (±0.04); $P = 2.04 \times 10^{-18}$**
		ZNF genes & repeats	**+0.2 (±0.05); $P = 4.03 \times 10^{-5}$**	**+0.79 (±0.19); $P = 2.6 \times 10^{-5}$**
		Heterochromatin	**-0.72 (±0.02); $P < 10^{-308}$**	**-2.89 (±0.07); $P < 10^{-308}$**
		Poised Promoter	**+0.13 (±0.01); $P = 4.38 \times 10^{-22}$**	**+0.41 (±0.05); $P = 1.51 \times 10^{-15}$**
		Bivalent Promoter	**+0.61 (±0.03); $P = 1.45 \times 10^{-85}$**	**+2.29 (±0.12); $P = 1.16 \times 10^{-83}$**
		Repressed Polycomb	**+0.12 (±0.01); $P = 2.53 \times 10^{-26}$**	**+0.34 (±0.04); $P = 8.18 \times 10^{-15}$**
		Quiescent/Low	**-1.25 (±0.01); $P < 10^{-308}$**	**-5.01 (±0.04); $P < 10^{-308}$**
GTEx - number of genes the		eGenes<10	**+0.31 (±0.01); $P = 4.61 \times 10^{-119}$**	**+1.11 (±0.05); $P = 9.55 \times 10^{-104}$**
		eGenes>10 & <15	**+0.48 (±0.03); $P = 2.16 \times 10^{-57}$**	**+1.77 (±0.11); $P = 2.24 \times 10^{-53}$**

variant is an eQTL for	eGenes>15 & <20	**+0.91 (±0.07); $P = 2.12 \times 10^{-35}$**	**+3.59 (±0.28); $P = 2.95 \times 10^{-37}$**
	eGenes>20	**+2.37 (±0.22); $P = 1.31 \times 10^{-27}$**	**+8.53 (±0.83); $P = 1.32 \times 10^{-24}$**
GTEx - number of tissues the variant is an eQTL for	eTissue<30	**+0.28 (±0.01); $P = 7.87 \times 10^{-92}$**	**+0.98 (±0.05); $P = 5.42 \times 10^{-79}$**
	eTissue>30 & <35	**+0.59 (±0.03); $P = 6.98 \times 10^{-118}$**	**+2.17 (±0.1); $P = 6.6 \times 10^{-110}$**
	eTssue>35 & <40	**+1.1 (±0.07); $P = 8.56 \times 10^{-61}$**	**+4.31 (±0.26); $P = 1.26 \times 10^{-63}$**
	eTissue>40	**+1.2 (±0.17); $P = 1.65 \times 10^{-12}$**	**+3.79 (±0.65); $P = 6.43 \times 10^{-9}$**
Tong et al.	eVariant	**+0.87 (±0.15); $P = 7.01 \times 10^{-9}$**	**+2.86 (±0.51); $P = 3.29 \times 10^{-8}$**
	Master eVariant	**+1.93 (±0.56); $P = 5.62 \times 10^{-4}$**	**+6.47 (±1.92); $P = 7.82 \times 10^{-4}$**
International Mouse Phenotyping Consortium	Phenotypes > 1	**+0.18 (±0.04); $P = 5.65 \times 10^{-5}$**	**+0.75 (±0.17); $P = 5.87 \times 10^{-6}$**
Saccharomyces cerevisiae Morphological Database	Phenotypes > 1	**+0.26 (±0.04); $P = 1.56 \times 10^{-12}$**	**+0.94 (±0.14); $P = 5.31 \times 10^{-12}$**

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.

Table 3: Pleiotropy score for a selection of medically relevant genes for coronary artery disease and type 2 diabetes.

Gene	Original Trait	# of variants	Average P_m Score	*P-value* for P_m	Average P_n Score	*P-value* for P_n
PCSK9	Coronary Artery Disease	11	22.23	1.0×10^{-5}	31.13	4.2×10^{-4}
NPC1L1	Coronary Artery Disease	7	22.70	5.0×10^{-7}	31.43	4.2×10^{-4}
APOC3	Coronary Artery Disease	1	21.57	3.7×10^{-4}	29.87	1.5×10^{-3}
ANGPTL4	Coronary Artery Disease	2	22.92	1.1×10^{-7}	33.86	1.0×10^{-4}
LPA	Coronary Artery Disease	27	23.02	5.1×10^{-8}	34.13	4.8×10^{-5}
ANGPTL3	Coronary Artery Disease	5	26.18	1.4×10^{-21}	30.82	8.2×10^{-4}
PPARG	Type 2 Diabetes	76	26.13	2.6×10^{-21}	44.62	4.7×10^{-9}
KCNJ11	Type 2 Diabetes	14	26.00	1.4×10^{-20}	46.24	5.4×10^{-10}
ABCC8	Type 2 Diabetes	60	22.45	2.6×10^{-6}	32.50	2.1×10^{-4}
SLC30A8	Type 2 Diabetes	75	21.43	7.4×10^{-4}	27.54	4.7×10^{-3}
GCK	Type 2 Diabetes	17	22.30	6.5×10^{-6}	32.52	2.1×10^{-4}
GCKR	Type 2 Diabetes	10	32.96	2.7×10^{-72}	62.30	5.8×10^{-19}
SGLT2	Type 2 Diabetes	5	21.44	7.0×10^{-4}	24.36	1.9×10^{-2}
PTEN	Type 2 Diabetes	9	23.64	3.8×10^{-10}	37.78	4.1×10^{-6}

We selected genes for coronary artery disease and type 2 diabetes that are current drug targets and have been shown to have prior human genetic evidence for these traits. We measured the pleiotropy score of variants located in these genes and/or annotated as

eQTLs for these genes. These genes exhibit a range of pleiotropy scores, but all are higher than expected under the null hypothesis of no pleiotropy.