Factorizing polygenic epistasis improves prediction and uncovers biological pathways in complex traits

David Tang[1,2,*], Jerome Freudenberg[1], and Andy Dahl[1,**]

[1]Section of Genetic Medicine, University of Chicago
[2]Program in Bioinformatics and Integrative Genomics, Harvard Medical School
[*]davidtang@g.harvard.edu
[**]andywdahl@uchicago.edu

Abstract

Epistasis is central in many domains of biology, but it has not yet proven useful for complex traits. This is partly because complex trait epistasis involves polygenic interactions that are poorly captured in current models. To address this gap, we develop a new model called Epistasis Factor Analysis (EFA). EFA assumes that polygenic epistasis can be factorized into interactions between a few Epistasis Factors (EFs), which represent latent polygenic components of the observed complex trait. The statistical goals of EFA are to improve polygenic prediction and to increase power to detect epistasis, while the biological goal is to unravel genetic effects into more-homogeneous units. We mathematically characterize EFA and use simulations to show that EFA outperforms current epistasis models when its assumptions approximately hold. Applied to predicting yeast growth rates, EFA outperforms the additive model for several traits with large epistasis heritability and uniformly outperforms the standard epistasis model. We replicate these prediction improvements in a second dataset. We then apply EFA to four previously-characterized traits in the UK Biobank and find statistically significant epistasis in all four, including two that are robust to scale transformation. Moreover, we find that the inferred EFs partly recover pre-defined biological pathways for two of the traits. Our results demonstrate that more realistic models can identify biologically and statistically meaningful epistasis in complex traits, indicating that epistasis has potential for precision medicine and characterizing the biology underlying GWAS results.

Introduction

Epistasis refers to interactions between genetic effects on a trait. Epistasis is central in many domains of biology, including rare human disorders [1–3], protein evolution [4, 5], natural selection [6, 7], and functional genomics [8]. In model systems, statistical models of epistasis can be useful for characterizing genetic architecture [9–16], improving genomic selection [17, 18], and unbiasedly screening for unknown genomic mechanisms [19–22]. In striking contrast with the others domains, it remains debated if epistasis matters in complex traits [23–25]. This is primarily because complex trait biology is poorly understood, which severely limits our ability to study epistatic interactions between causal genetic mechanisms.

Unbiased genome-wide statistical tests for epistasis could quantify and characterize the nature of epistasis in complex traits. However, modelling epistasis in complex traits is challenging because they are affected by a large number of genetic variants, i.e., they are polygenic (**Figure 1a**). Although studies have identified some specific examples of epistasis in complex traits [26–28], they have not explained significant missing heritability [29, 30]. We hypothesize this is partly due to shortcomings in current models of complex trait epistasis. Specifically, prior studies have assumed that each epistatic interaction effect is completely random–independent of all other additive effects and interaction effects (**Figure 1b**, [31–33]). This model is mathematically simple but biologically unrealistic and statistically under-powered.

Here, we develop a new complex trait epistasis model called Epistasis Factor Analysis (EFA). EFA assumes a "coordinated" form of epistasis where SNP-SNP interactions are structured by interactions between a few latent components, which we call Epistasis Factors (EFs) (**Figure 1c**, [33, 34]). EFA simultaneously partitions GWAS effects into latent EFs and estimates the interactions between the EFs. The inferred EFs can be validated with known genomic annotations and used to discover novel biological pathways underlying GWAS results. A key feature of EFA is leveraging additive effects to improve estimates of epistasis effects, which is important because additive effects are usually much stronger in complex traits. We illustrate this

advantage of EFA over standard models in simulations where heritability is mostly additive. We also show EFA improves prediction over the additive model for traits with modest epistasis heritability in two multi-trait yeast datasets. We apply EFA to the four UK Biobank traits characterized by [35] and find significant epistasis in all four, including two that are robust to phenotype scale. Finally, we find that the inferred EFs partly recover predefined biological pathways. Our results show that structured models of polygenic epistasis can improve statistical performance and biological understanding in complex traits.

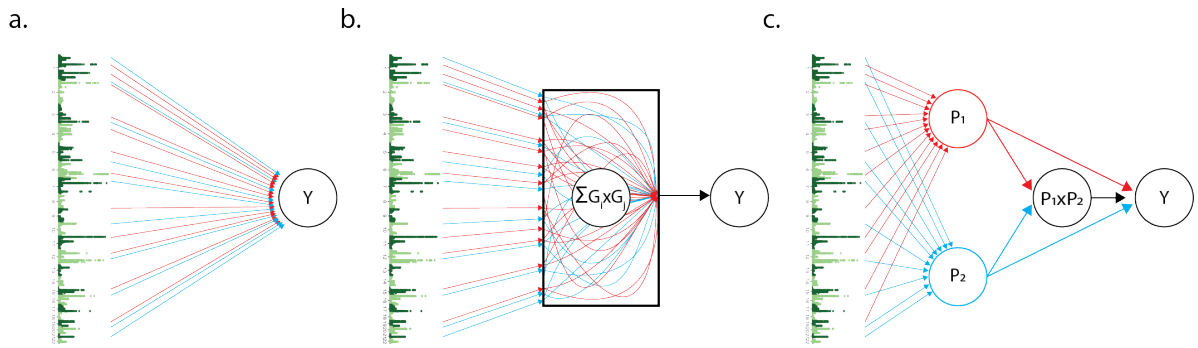

Figure 1: Three genetic architectures consistent with a typical GWAS. (a) Additive: Each SNP has an independent effect on the phenotype, Y. (b) Uncoordinated: Each SNP interacts randomly with all other SNPs. (c) Coordinated: SNPs can be grouped into two interacting factors. EFA uses the P1xP2 interaction to (1) improve statistical power to detect epistasis and (2) sort GWAS loci into factors enriched in distinct biological pathways.

Epistasis Factor Analysis

Nearly all complex trait studies assume the additive model, and it remains debated whether epistasis matters in complex traits [24, 25]. However, this debate has largely been predicated on an "uncoordinated" model of epistasis that assumes the epistasis effects are completely random–independent of both additive effects and other epistasis effects [30, 31, 36]. Uncoordinated epistasis is not interpretable, realistic, or statistically supported [30]. In contrast, we recently proposed a more realistic "coordinated" model of epistasis where epistasis and additive effects are structured by latent factors (**Figure 1c**, [33]). Coordination is motivated by known forms of epistasis in simpler traits, such as genetic modifiers of a Mendelian gene [2].

Here, we introduce a new coordinated epistasis model to identify these latent factors called Epistasis Factor Analysis (EFA). EFA makes two key assumptions on the nature of epistasis in complex traits:

A1: SNP effects are mediated through a few distinct polygenic factors

A2: These polygenic factors interact

A1 and A2 are biologically motivated. For example, the factors in A1 could reflect distinct causal tissues [37, 38], core genes [39, 40], or heritable exposures like smoking [41]. In accordance with A2, causal tissues might signal each other, core genes might encode proteins that physically interact, or an exposure's effect may depend on underlying genetic risk.

The key idea in EFA is to share information between additive effects and epistasis effects. On one hand, leveraging additive signals greatly improves power to detect epistasis. On the other hand, leveraging epistasis helps unbiasedly partition SNPs into distinct factors, which is impossible in the additive model without external data or biological annotations.

EFA applies to a quantitative trait measured on N individuals, y, and a matrix of genotypes measured on N individuals at M loci, G. The EFA model is:

$$y_i = P_{i1} + P_{i2} + 2\lambda P_{i1} P_{i2} + \varepsilon_i; \qquad P_{ik} := \sum_{j=1}^{M} G_{ij} U_{jk} \tag{1}$$

where P_{ik} is the level of EF k for sample i, U_{jk} is the weight of SNP j on EF k, and λ is the interaction between EFs. We call this Epistasis Factor Analysis because it implicitly makes a low-rank factorization to the genome-wide epistasis matrix (**Methods**). We generalize to $K > 2$ EFs in the **Methods**, and we theoretically characterize EFA in the **Supplement**. Our full model also allows quadratic effects of individual EFs, but we conservatively exclude this term in our main analyses because our goal is to find interactions across biologically distinct pathways.

In practice, G will typically contain dozens of SNPs that have been pre-screened based on their additive effects. We fit U and λ as fixed effects using maximum likelihood (**Methods**). We develop an efficient block coordinate descent algorithm in the **Supplement**. Using linear algebra techniques, we reduce the computational complexity of EFA to the same cost as ordinary linear regression.

EFA recovers true pathways and improves prediction in simulations

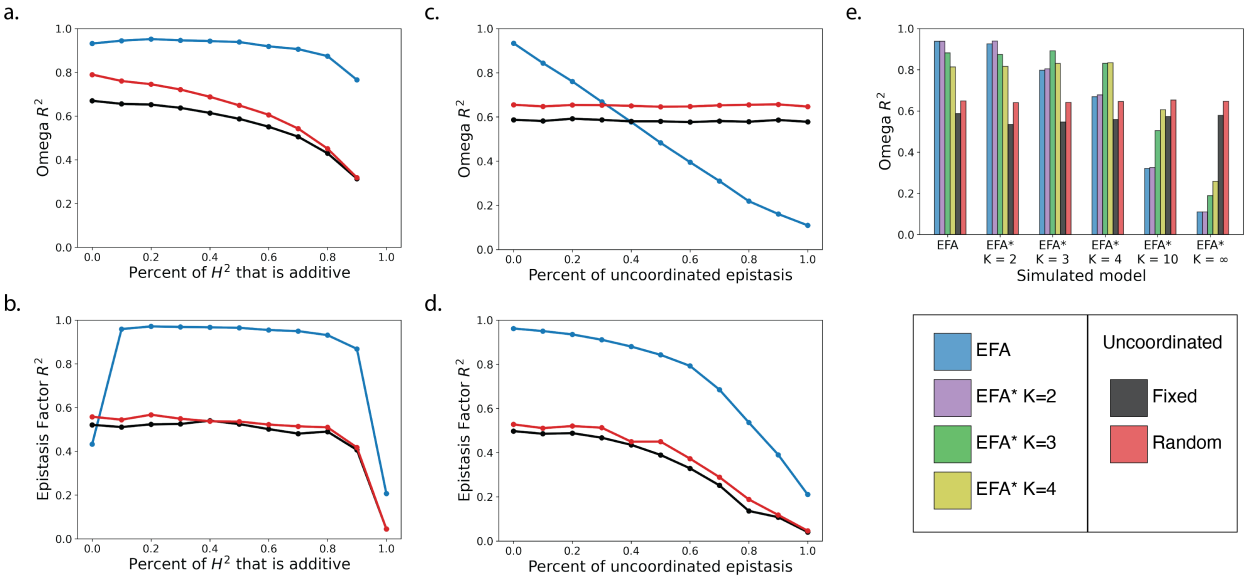

Figure 2: Simulations characterize the accuracy of polygenic epistasis estimates. Accuracy is quantified by the squared correlation between estimated and true pairwise SNP epistasis effects, ω (a,c,e) or between estimated and true EFs, U (b,d). Simulations under the EFA model vary the fraction of heritability due to additive vs epistasis effects (a,b). Partial coordination is simulated by combining epistasis effects from the EFA model with uncoordinated epistasis effects (c,d). Other forms of coordination are simulated using the EFA* model with $K \geqslant 2$ (e), which is a more general version of EFA (**Methods**). In (b,d), the EF R^2 is nonzero even without any coordinated epistasis because the EFs still capture the additive effects.

We performed simulations to characterize EFA as a function of additive heritability (h^2_{add}), epistasis heritability (h^2_{epi}), and the fraction of epistasis heritability due to coordinated vs uncoordinated effects (**Methods**). We center and scale genotypes to remove the ambiguity between variance "attributable to" vs variance "caused by" additive effects (in practice, these effects are partly collinear, which can cause confusion [24, 25]). We define broad-sense heritability as $H^2 := h^2_{add} + h^2_{epi}$, and always use $H^2 = 0.5$. We compared EFA to uncoordinated epistasis estimates fit either by fixed or random effects (**Methods**). We evaluated estimation accuracy for the pairwise SNP-SNP interactions, $\omega_{jj'}$ (see (3) in the **Methods**), and for the EFs, U_{jk}. EFA directly fits EFs and implicitly fits ω (**Methods**). Uncoordinated methods directly fit ω, and we derive uncoordinated EFs *post hoc* using PCs of the estimated ω matrix (**Methods**).

We first simulated from the EFA model in (1) with $N = 1,000$ samples and $M = 20$ SNPs. As expected, EFA outperforms uncoordinated estimates of ω across the range of h^2_{epi} (**Figure 2a**). EFA's gain is greatest when $h^2_{add} \gg h^2_{epi}$, illustrating how EFA's epistasis estimates borrow strength from additive effects. Analogously,

EFA improves additive effect estimates over standard additive models when $h^2_{epi} \gg h^2_{add}$ (**Supplementary Figure S1**), though this scenario is not likely in practice. EFA also substantially outperforms uncoordinated estimates of EFs, which is expected because EFA specifically targets these low-dimensional factors (**Figure 2b**). This also shows how EFA leverages additive signals to improve EF estimates: when $h^2_{add} = 0$, EFA actually underperforms the uncoordinated model, whereas EFA remains accurate even when heritability is 90% additive.

We then fix $h^2_{add} = h^2_{epi} = 25\%$ and vary the level of coordination by adding in uncoordinated epistasis effects (**Methods**). As expected, EFA's performance decays as coordination weakens, while the uncoordinated estimates remain stable (**Figure 2c**). Nonetheless, EFA always outperforms uncoordinated estimates of the EFs because EFA specifically targets the coordinated component of epistasis (**Figure 2d**).

We next simulated and fit $K > 2$ EFs using a generalization of EFA called EFA* (**Methods**). As expected, ω estimates were more accurate when fitting EFA* with the true value of K (**Figure 2e**), though the base EFA model was robust to modest misspecification of K. However, the uncoordinated model performed better as K grows. This is expected, as the $K = \infty$ case is equivalent to uncoordinated epistasis [33]; intuitively, the independent pathways become individually negligible and wash out. We did not consider EF estimation accuracy because EFs are not identified for $K > 2$ (**Supplement**).

We next varied the number of SNPs, M, and sample size, N (**Supplementary Figure S2**). Across all N and M, EFA always outperforms uncoordinated estimates. Performance for all methods decays as M grows because each individual SNP effect is smaller. This illustrates why modelling epistasis in complex traits is difficult. Likewise, all methods improve as N grows.

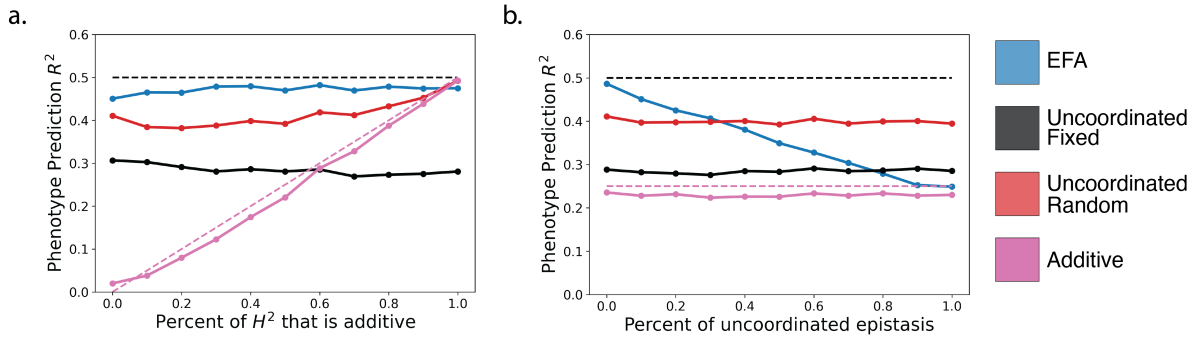

Figure 3: Phenotype prediction accuracy in simulations. (a) varies the fraction of heritability due to additive vs epistasis effects. (b) varies the fraction of epistasis effects due to coordinated pathway-level interactions vs uncoordinated random interactions. Dotted line is the broad-sense heritability, H^2.

Finally, we asked if EFA can improve phenotype prediction. We first simulated from EFA's model and found that EFA performed well, with prediction R^2 near the theoretical limit of H^2 (**Figure 3a**). EFA always outperformed the fixed-effect uncoordinated model, and EFA outperformed the random effect uncoordinated model except when epistasis was absent. (We derive the predictions from the uncoordinated random effect model in the **Supplement**.) We then fixed $h^2_{epi} = h^2_{add} = .25$ and varied the fraction of epistasis due to coordination, as above. Results were similar to ω estimation accuracy: when epistasis is mostly coordinated, EFA is optimal, but EFA performs poorly when epistasis is uncoordinated (**Figure 3b**). Overall, EFA improves estimation of latent pathways in all our tests and improves prediction when epistasis is sufficiently strong and coordinated.

EFA improves prediction in complex yeast traits

We next asked if EFA could improve genetic prediction in complex traits. We studied 46 complex yeast traits measured on 1008 samples, where each trait was growth rate on a different medium [13]. These traits have varying levels of epistasis heritability, as defined by the difference between the phenotypic similarity of clones (H^2) and the additive heritability estimated from genotype data (h^2). We measured prediction accuracy by

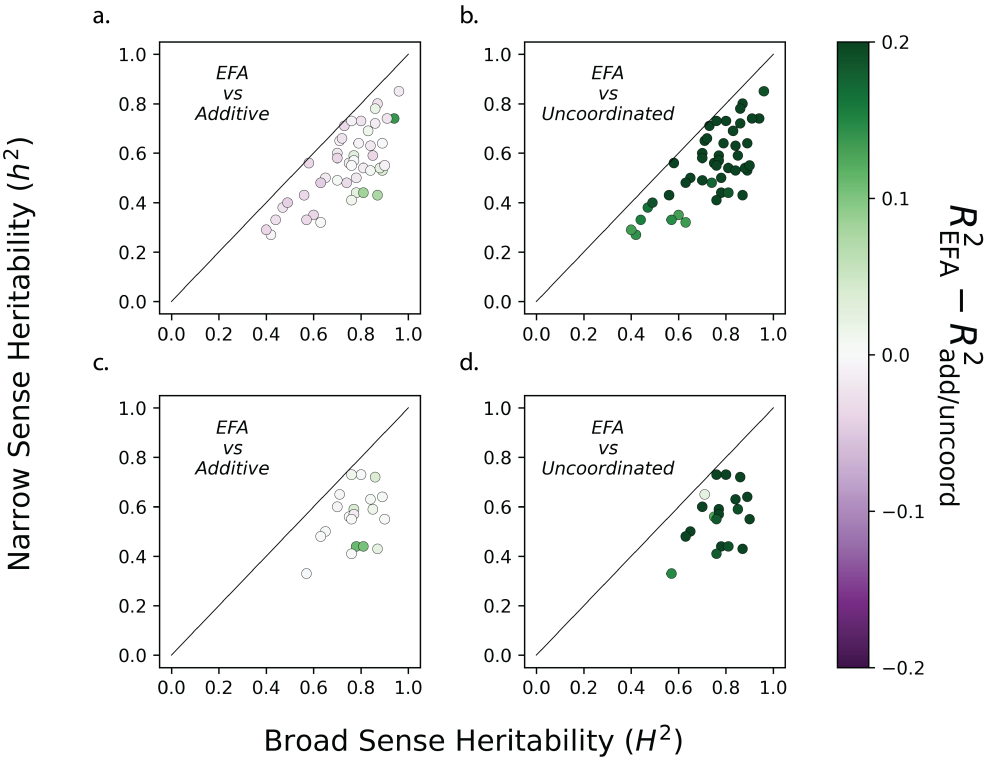

Figure 4: Prediction accuracy in complex yeast traits. Each point is a trait with narrow- and broad-sense heritability estimates from [42]. (a) EFA significantly outperforms the additive model for 6/46 traits from Bloom et al 2013 data. (c) EFA significantly outperforms the additive model for 7/20 traits from Bloom et al 2015 data. (b,d) EFA always outperforms the uncoordinated fixed-effect epistasis model.

the squared correlation between predicted and observed phenotypes using 10-fold cross-validation. For each trait, this yielded 30-70 clumped+thresholded SNPs ($r^2 < 0.2$, **Methods**).

We first compared EFA to the additive model and found that EFA significantly improved prediction accuracy for 6/46 traits ($p < .05$, paired t-test, **Methods**, **Supplementary Table 1**). These 6 traits all had large epistasis heritability (**Figure 4a**). More generally, the difference between R^2_{EFA} and R^2_{add} correlated with the epistasis heritability (Spearman's $\rho = 0.48$, $p < 0.001$). This is consistent with our simulations showing that EFA can outperform additive predictions when h^2_{epi} is higher. By comparison, the additive model significantly improved prediction over EFA for 20/46 traits. Finally, EFA predictions were superior to uncoordinated epistasis predictions using fixed effects for all 46 traits (**Figure 4b**).

Next, we compared EFA to additive and epistasis random effects models. Unlike fixed effect models, including EFA, random effect models can fit the whole genome without LD pruning, which is often a major advantage for prediction in complex traits [18, 43, 44]. Nonetheless, EFA significantly outperforms both additive and uncoordinated epistasis random effect models for 7/46 traits ($p < .05$, **Supplementary Figure S3**). We also compared the two random effect models, and we found that the uncoordinated epistasis model significantly outperformed the additive model for 9/46 traits ($p < .05$). Intriguingly, these 9 traits only overlap 1 of the 7 traits where EFA outperforms the additive random effect model, suggesting that the degree of coordination varies across traits.

Finally, we evaluated replication in a second dataset from the same group with 4390 samples measured on 20 of the above 46 traits [14] (**Supplementary Figure S4**). We replicated the superiority of EFA over the additive model for all 3 traits where EFA beat the additive model in the first dataset ($p < .05$; the other 3/6 traits were not measured in the second dataset). Further, at this larger sample size, we find 4 additional

traits where EFA significantly outperforms the additive fixed effect model (**Figure 4c**, **Supplementary Table 2**). We also replicate that EFA always outperforms the uncoordinated fixed effects model (**Figure 4d**). Overall, EFA improves prediction over the additive model for some complex traits, and the utility of EFA will likely grow with larger sample sizes.

EFA detects epistasis and recovers known pathways in human traits

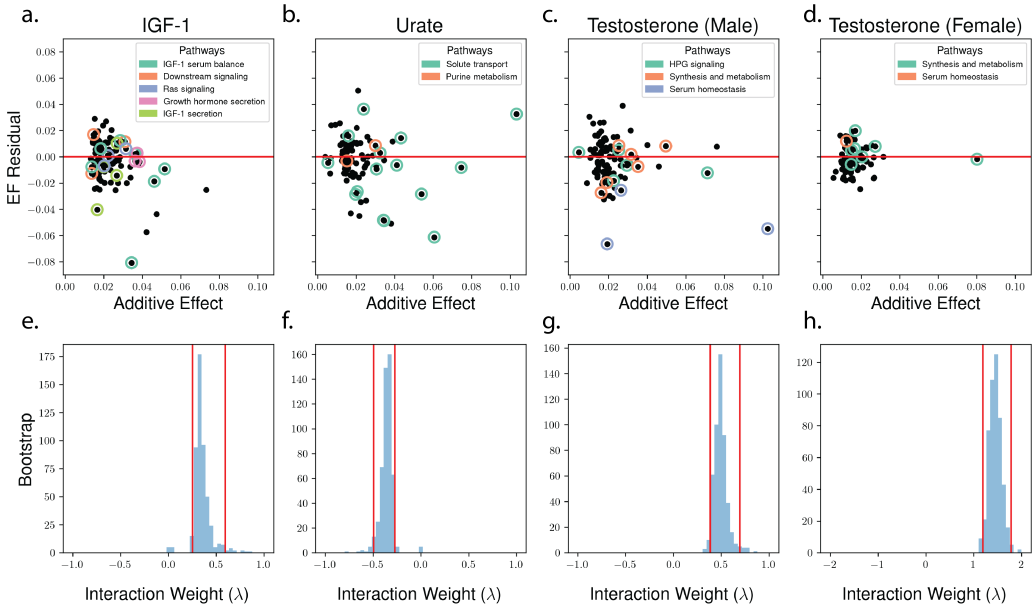

Figure 5: EFA identifies statistically and biologically significant epistasis in 4 complex human traits. (a-d) EFA jointly partitions the additive effects of GWAS-significant SNPs (x-axis) into two EFs (y-axis: difference between the EFs and the additive effects, i.e. $U_{jk} - \beta_j/2$). (e-f) Bootstrap replicates show that EFA identifies significant epistasis in all four traits. For visualization, 14 values of λ are excluded across all plots (out of a total of 6,000). See **Supplementary Figure S5** for comparable plots on quantile-normalized and permuted phenotypes.

		IGF-1	Urate	Testosterone (Male)	Testosterone (Female)
Standard	CI	(0.25, 0.60)	(-0.50, -0.27)	(0.39, 0.70)	(1.19, 1.78)
	Empirical p-val	**.002**	**.006**	**.002**	**.002**
	Gaussian p-val	3.52e-04	1.16e-07	.011	2.75e-04
QNorm	CI	(-0.30, 0.37)	(-0.57, -0.36)	(-0.46, 0.42)	(0.63, 1.04)
	Empirical p-val	.570	**.002**	.826	**.002**
	Gaussian p-val	.689	7.67e-15	.846	3.72e-12
Permuted	CI	(-0.28, 0.32)	(-0.47, 0.00)	(-0.46, 0.41)	(-1.48, 0.75)
	Empirical p-val	.730	.066	.618	.274
	Gaussian p-val	.749	.090	.642	.984

Table 1: EFA finds statistically significant evidence of epistasis in complex traits. Confidence intervals were computed using the 0.025 and 0.975 quantiles of the bootstrap distribution. All p-values are two sided.

We applied EFA to complex human traits in UK Biobank (UKB). We studied the same four traits that were extensively characterized in [35]: urate, insulin-like growth factor 1 (IGF1), testosterone in males, and testosterone in females. We used the top 100 LD-pruned GWAS hits for each trait (or all 77 for female testosterone, **Methods**).

We first asked if EFA detects statistically significant epistasis using bootstrap. Specifically, we resample

individuals to calculate the sampling distribution of λ, the factor-level epistasis, which yields empirical p-values and confidence intervals (**Methods**). We used simulations to confirm that this bootstrap test is calibrated under the null model where SNPs are additive (**Supplementary Figure S6**). Additionally, we performed realistic simulations with true epistasis and found that EFA's λ estimates are unbiased and that EFA's bootstrap test is powerful. A minor caveat is that our bootstrap confidence intervals are liberal in the presence of true epistasis, but this does not cause false positive tests of epistasis.

The EFA bootstrap test finds significant epistasis for all four tested traits (all empirical $p < .05/4$, **Figure 5e-h**, **Table 1**). For IGF-1 and male and female testosterone, λ is positive, and for urate, λ is negative. Because we use 499 bootstrap replicates, the empirical p-values are noisy and lower-bounded by 1/500. If we further assume that our estimate of λ is roughly Gaussian under the additive model, we get much lower and more precise p-values. Our simulations suggest that this Gaussian approximation is conservative (**Supplementary Figure S7**).

To test robustness under the null, we repeated our analysis after permuting phenotypes. As expected, our bootstrap test was null (**Table 1**, **Supplementary Figure S5**).

As epistasis can be scale-dependent [45], we conservatively tested robustness by quantile-normalizing phenotypes (**Supplementary Figure S5**). We found that urate and female testosterone remained highly significant (empirical $p = .002, .002$; Gaussian $p = 7.7 \times 10^{-15}, 3.7 \times 10^{-13}$). Moreover, the sign and magnitude of λ remained stable after this transformation. However, quantile-normalization eliminated the EFA signal for male testosterone and IGF-1, which emphasizes the importance of evaluating scale-dependence in interaction tests. Overall, our results increase confidence that the EFA signals for urate and female testosterone are not scale dependent, while the discrepancies for IGF-1 and male testosterone could either be false positives on the original scale or false negatives on the transformed scale [33].

We next performed simulations to assess EFA's robustness to dominance and un-genotyped causal variants in LD with genotyped variants. First, we simulated a pure-dominance model without any epistasis between distinct SNPs (**Supplement Section 3.2-3**). Under modest levels of dominance ($h^2_{dom} = 0.1\%$, [30, 46]), the bootstrap test is roughly null (**Supplementary Figure S6**). Under strong dominance ($h^2_{dom} = 1\%$), however, the bootstrap test becomes significant (**Supplementary Figure S6**). Both results are expected: dominance is not technically a false positive under EFA (unlike our prior Even-Odd test [33]), but EFA primarily targets inter-SNP epistasis, hence the EFA test has low power if dominance effects are the only form of epistasis.

Second, we simulated an additive model with unobserved causal variants in LD with measured variants. Our simulation meets the necessary condition for 'phantom' epistasis in [47], which can cause severe bias in pairwise SNP-SNP epistasis tests (**Supplement**). For EFA, however, we do not observe any bias from ungenotyped variants (**Supplementary Figure S6**). Nonetheless, in principle, other forms of LD could bias EFA estimates. Overall, the structured and polygenic nature of EFA substantially mitigates the LD biases that can plague other epistasis tests [18, 48–51].

We next asked if EFA captures biologically meaningful epistasis by comparing estimated EFs to the known biological pathways defined in [35]. Specifically, we used a hypergeometric test for enrichment of the top 10 SNPs per EF in each predefined biological pathway (relative to the other 90 GWAS SNPs, **Table 2**).

We find that EF2 for male testosterone is enriched in "serum homeostasis" ($p = 0.0001$). This enrichment is driven by SNPs assigned to the genes *SHBG* and *SLCO1B3*, which live on different chromosomes (17 and 12). In urate, EF2 is enriched in "solute transport" ($p = 0.007$, **Table 2**), which is driven solely by a single locus with large effect (*SLC2A9*). This urate signal could either reflect genuine *cis-trans* interactions or could merely reflect subtle LD with unmeasured additive effects and/or dominance effects [30]. We confirmed these enrichments are robust if we instead use the top 15 or 20 SNPs per EF (**Supplementary Table 3-6**). The results are also robust if we increase the window size used to assign SNPs to genes (**Supplementary Table 3-6**). Altogether, EFA is capable of recovering biologically-meaningful pathways solely using statistical interactions.

Trait	Biological Pathway	# SNPs in EF1	# SNPs in EF2	# SNPs total
IGF-1	Downstream signaling	1 ($p = 0.284$)	0 ($p = 1.000$)	3
	Growth horm. secr.	0 ($p = 1.000$)	1 ($p = 0.361$)	4
	IGF-1 secretion	0 ($p = 1.000$)	1 ($p = 0.361$)	4
	IGF-1 serum balance	0 ($p = 1.000$)	1 ($p = 0.549$)	7
	Ras signaling	0 ($p = 1.000$)	1 ($p = 0.430$)	5
Urate	Solute transport	3 ($p = 0.175$)	5 ($p = 0.007$) $\star$	15
Testosterone (male)	Steroid synthesis & metabolism	0 ($p = 1.000$)	1 ($p = 0.478$)	6
	HPG signaling	0 ($p = 1.000$)	0 ($p = 1.000$)	5
	Serum homeostasis	0 ($p = 1.000$)	3 ($p = 0.001$) $\star$	3
Testosterone (female)	Steroid synthesis & metabolism	1 ($p = 0.638$)	0 ($p = 1.000$)	7

Table 2: Epistasis Factors are enriched in predefined biological pathways from [35]. We show pathways that have been assigned > 2 SNPs used in EFA. p-values are from a hypergeometric test of the top 10 EF SNPs in each biological pathway relative to the other 90 SNPs.

Discussion

Despite strong evidence for epistasis in many domains of biology, it remains unclear if modelling epistasis adds biological or statistical value in complex traits. We introduced a model to address this question called Epistasis Factor Analysis (EFA). In contrast to existing uncoordinated models of polygenic epistasis, which lack biological motivation, EFA assumes a structured form of polygenic epistasis driven by latent factors (**Figure 1C**). We showed that EFA outperforms additive predictions in several yeast traits. Additionally, we found that EFA can identify epistasis in complex human traits where standard models fail and that the EFs can partly recover known biological pathways without prior information.

We found that EFA can improve prediction in model organisms and simulations, and it is likely that epistasis will eventually improve prediction in natural populations. However, the within-population prediction gain will likely be modest because current additive models partly capture epistasis effects–especially when causal variants have been driven to lower frequencies by natural selection [24, 52]). It is possible that epistasis predictions will improve portability across ancestries, but epistasis has not yet been shown to contribute to this important problem [53, 54]. Therefore, the greater promise of epistasis models in complex human traits is not better statistical prediction, but rather better biological understanding of GWAS results. Our results showing that EFs are enriched in predefined genomic annotations is a step in this direction.

Many polygenic epistasis models have been previously developed. The standard model, which we call "uncoordinated" [33], assumes that all SNP-SNP epistasis effects are independent of each other and of additive effects [31]. While uncoordinated models are mathematically simple and rigorous, they are biologically implausible and statistically under-powered [30]. More recent polygenic epistasis models make parsimonious assumptions to improve power and interpretation. For example, autosome-sex interactions represent a particular form of structured epistasis that is strongly supported in diverse complex traits [55–58] (though these interactions are at least partly due to gender, not sex). As another example, interactions between a single locus and a polygenic background can identify epistasis hubs [59–64]. Finally, BANN models complex traits as epistatic compositions of simpler pathways, like EFA, though BANN pathways are more numerous, less polygenic, and require prior biological annotations [65].

There are several important limitations to our study. First, like linear regression, EFA can fit dozens to thousands of genetic variants (**Supplementary Figure S3**) but cannot jointly fit the entire genome. In practice, we pre-screen SNPs based on additive signals, such as selecting only GWAS-significant SNPs in UK Biobank, which enriches for SNPs with epistasis effects [24, 66]. Second, our polygenic epistasis tests do not establish interactions between any given SNP pair. In the future, sparse models (Bayesian or frequentist) could test which SNPs affect which EFs. Third, following [35], we have only studied four relatively simple complex human traits, and it will be important in the future to evaluate EFA more broadly across more complex traits such as height and BMI.

Another set of concerns are statistical false positives from phenotype scale and/or complex LD patterns. First, it is obvious that scale transformations of an additive trait can induce interactions [33, 45, 67, 68].

Nonetheless, 2/4 of our human EFA signals are qualitatively identical on the natural and quantile-normalized scales, and coordination is theoretically robust to modest rescaling [33]. Second, LD with unmeasured causal variants can cause dramatic and replicating false positives, especially in small *cis*-windows with large effects [18, 48–51]. We partly address this by LD-pruning SNPs, and we also use simulations to show that EFA's polygenic nature reduces sensitivity to locus-specific LD biases. More broadly, EFA's utility for prediction and biological characterization persists even if EFA signals reflect scaling and/or subtle LD patterns. Overall, these caveats are crucial for interpreting interactions, but they do not impact our central conclusions.

There are several extensions to EFA that may prove useful. First, EFA could jointly model multiple traits that partly share EFs. This would add power to detect shared EFs and provide a rich decomposition of pleiotropy that goes beyond genetic correlation, which is overly simplistic [69–73]. Second, EFA could model complex diseases in a generalized linear model framework, which is closely related to the idea of limiting disease pathways [34] and disease subtyping [57]. However, binary disease traits will have lower power and also have subtle scale dependence issues [74]. Third, we have focused on SNP interactions, but in principle, EFA is capable of modeling more powerful and/or interpretable genetic features such as imputed gene expression [75, 76], copy number variants [27, 28, 77], or polygenic scores for secondary traits [78] or exposures [41]. EFA could also incorporate non-genetic variables such as epigenomic marks, medical image-derived features [79], disease symptoms, or comorbidities from electronic health records. Ultimately, we hope that the EFA model and its strong results in real data provide a solid step toward unravelling epistasis in complex human traits.

Code availability

A python implementation of EFA is available at: `https://github.com/tangdavid/efa`. This link also contains the code for all analyses in this paper, though the UKB analysis is not fully reproducible because the raw data is not public. An R implementation is available at: `https://github.com/andywdahl/EFA`.

Data availability

We downloaded the first yeast data set from `http://genomics-pubs.princeton.edu/YeastCross_BYxRM/data/cross.Rdata` and the second yeast dataset from `https://static-content.springer.com/esm/art%3A10.1038%2Fncomms9712/MediaObjects/41467_2015_BFncomms9712_MOESM729_ESM.zip`

This research has been conducted using the UK Biobank Resource under application number 30397.

Author contributions

D.T. developed statistical methodology, performed analysis, and wrote the manuscript. J.F. performed analysis. A.D. conceived and supervised the project and wrote the manuscript.

Acknowledgements

We thank Matthew Stephens, Sasha Gusev, Sriram Sankararaman, and Noah Zaitlen for helpful feedback. We also thank the participants in UKB for making this study possible. Finally, we thank the Center for Research Informatics and the Research Computing Center for providing the compute resources necessary for carrying out this project. A.D. is supported by K25HL157603.

Methods

Mathematical description of Epistasis Factor Analysis

Epistasis Factor Analysis (EFA) aims to find a low-dimensional representation of polygenic epistasis that is both statistically powerful and biologically interpretable. The EFA model is:

$$y = G\,U\,\mathbf{1}_K + ((GU) \bullet (GU))\,\mathrm{vec}(\Lambda) + \varepsilon, \tag{2}$$

where y is a quantitative trait measured on N samples, G is an $N \times M$ matrix of genotypes at M SNPs, U is an $M \times K$ matrix of SNP effects on each of K factors, Λ is a symmetric $K \times K$ matrix of factor-level interactions, $\bullet$ is the face-splitting product ($A \bullet B$ is a matrix where each column is an element-wise product of a column of A and a column of B), and ε is independent and identically distributed (i.i.d.) Gaussian noise.

Intuitively, each column of U represents the SNP weights on a different latent epistasis factor (or EF), and $P_{,k} := GU_{,k}$ is an individual's weight on the k-th EF. $\Lambda_{kk'}$ represents the factor-level interaction between EFs k and k'. When $\Lambda_{kk'} > 0$, factors k and k' interact synergistically, meaning that the their combined phenotypic effect is greater than the sum of their individual effects. The reverse is true for antagonistic factors, when $\Lambda_{kk'} < 0$. When $\Lambda_{kk'} = 0$, then factors k and k' do not interact. Finally, Λ_{kk} refers to the quadratic effect of pathway k, which can help identify the pathway weights in U and which we do consider *bona fide* statistical epistasis. Nonetheless, our primary interest is in the $\Lambda_{kk'}$ terms where $k \neq k'$, as we are motivated by unravelling interactions across distinct biological factors.

In the main text, we call this general model EFA* and focus instead on the special case where (i) $K = 2$ and (ii) $\Lambda_{11} = \Lambda_{22} = 0$. We focus on this special case for simplicity and also because it is identified, i.e., the estimated EFs and their interactions are meaningful (**Supplement Section 1.5.1**). In contrast, the general EFA* model in (2) is not identified without further assumptions, such as that U is sparse (Note that most factor models, including PCA, are not identified without additional assumptions.) We prove these facts and comprehensively characterize the theoretical properties of EFA in the **Supplement**. We also formally characterize the equivalence class of solutions to EFA*, proving that the sign of $\Lambda_{kk'}$ is well-identified and providing a canonical representative of the equivalence class for $K = 2$ (**Supplement Section 1.5.2**). We also calculate the level of coordinated epistasis under EFA in the **Supplement** [33].

EFA is a special case of the standard pairwise polygenic epistasis model that underlies many epistasis models for complex traits:

$$y_i = \sum_j G_{ij}\beta_j + \sum_{j,j'} G_{ij}G_{ij'}\omega_{jj'} + \varepsilon_i \tag{3}$$

where β_j is the additive effect of SNP j, and $\omega_{jj'}$ is the epistatic interaction between SNPs j and j'. The factor-level interaction model in EFA can be directly connected with this SNP-level epistasis model by the linked factorization of the SNP-level additive and epistasis effects:

$$\beta = \sum_k U_{,k}; \quad \omega = U\Lambda U^T$$

Fitting EFA parameters

We always estimate the parameters of EFA using maximum likelihood (**Supplement Section 1.4**). In general, we fit the EFA* model using gradient descent based on automatic differentiation. We evaluate multiple random restarts, each initialized with $U_{,k}$ drawn from a neighborhood around 0. We use this approach in simulations and in the yeast analyses. This generic, flexible estimation method can easily accommodate future model extensions, e.g., introducing sparsity in U with ℓ_1 penalties or "anchoring" EFs to specific SNPs (**Supplement Section 1.5.3**).

To scale our method to UKB sample size, we derive a block coordinate ascent algorithm (**Supplement Section 1.4**). The updates are remarkably simple, requiring only iterative regressions that scale linearly in sample size and quadratically in SNPs (i.e., $O(NM^2)$). The key to this efficiency is that we only evaluate EF-EF interactions and never explicitly evaluate SNP-SNP interactions, which would cost $O(NM^4)$. With 100 SNPs, as in our UKB analyses, this is a 10,000-fold speedup.

Uncoordinated models of polygenic epistasis

We compare EFA with other polygenic models in simulations and real data analyses. All comparison methods are special cases of the overarching polygenic pairwise epistasis model in (3).

First, we compare to the standard additive model, which ignores epistasis by setting $\omega_{jj'} = 0$ for all j, j'. Because we mostly study scenarios with $M < N$, we primarily fit the additive model with fixed effects (by ordinary least squares). The predictions from this model are then essentially Polygenic Risk Scores (PRS). We also fit the additive model with random effects (assuming each β_j is i.i.d. Gaussian) in the real yeast analysis, enabling us to fit the full genome-wide data without any pruning.

Second, we compare to the standard pairwise epistasis model in (3) fit with fixed-effects. Because this model has $O(M^2)$ parameters, jointly fitting all pairwise epistasis effects is often noisy or even impossible. Instead, we fit each pairwise interaction $\omega_{jj'}$ one-at-a-time (for $j \leqslant j'$) while controlling for additive fixed effects (note that only $\omega_{jj'} + \omega_{j'j}$ is identified).

Third, we compare to the uncoordinated random effect model for pairwise epistasis [31]. This model assumes that $\omega_{jj'}$ are drawn i.i.d. Gaussian. This assumption simplifies computation because the ω can be easily marginalized out of the likelihood. We fit the variance components (one for β and one for ω) with maximum likelihood. We derive the best linear unbiased predictors (BLUPs) for both ω and for phenotype prediction in the **Supplement**.

Finally, we also develop a novel approach to estimate latent pathways from uncoordinated estimates by applying PCA to the estimated ω matrix. This is a naive baseline approach that only involves post-processing uncoordinated estiamtes. In contrast, EFA directly learns pathways, linking the low dimensional components of ω with the additive effects.

Simulations

We use a polygenic simulation framework that is fully described in the **Supplement Section 3**. In brief, we simulate under the EFA model (2) with independent SNPs. The latent pathways, $U_{,k}$, are drawn from i.i.d. Gaussians with variance chosen to give the appropriate additive heritability (recall that the additive effect is related to the latent pathways by $\beta = \sum_k U_{,k}$). The upper triangular entries of Λ are also drawn fro i.i.d. Gaussians with variance chosen to give the appropriate epistasis heritability. The pairwise polygenic epistasis effects are then generated by $\omega = \sum_{k,k'} \Lambda_{kk'} U_{,k} U_{,k'}^T$.

We also simulate partly coordinated epistasis by adding i.i.d. Gaussian random variables to each entry of a coordinated ω. By varying the relative contribution of these effects, we can interpolate between the EFA model (coordinated) and the classical random effect model (uncoordinated). By choosing the relative variances of the different components appropriately, we are able to simulate the standard additive model, the standard uncoordinated model, and the coordinated EFA model.

Finally, we use two alternative simulation frameworks for evaluating the calibration of our bootstrap test with respect to LD and dominance that we describe fully in the **Supplement**.

Yeast data

We downloaded genotype and phenotype data for prototrophic haploid segregants from a cross between a laboratory strain and a wine strain of yeast [13]. This dataset consisted of 46 quantitative traits for 1,008 samples genotyped at 11,623 unique markers. We standardized each phenotype and SNP to mean 0 and variance 1. To account for linkage disequilibrium in fixed effect models, we clumped+thresholded variants using threshold of $r^2 < 0.2$, minimum distance of 250 kb, and marginal additive $p < 0.05$ using PLINK 1.9 [80]. We performed clumping+thresholding separately on each cross-validated fold to avoid bias from overfitting [81].

Additionally, we downloaded a second yeast growth trait dataset from the same group with a larger sample size ($N = 4390$) [14]. This second dataset contained 20/46 of the growth phenotypes in the first dataset, and it included 28,220 unique genotype markers for each sample. We preprocessed this dataset in the same way as above.

We evaluated phenotype prediction accuracy with 10-fold cross validation, using the same 10 folds across all methods so we could test for differential prediction accuracy with paired t-tests.

UK Biobank data

We studied the same four traits in UK Biobank (UKB) that were extensively characterized in [35]: urate, IFG1, testosterone in males, and testosterone in females. We analyzed unrelated "white British" individuals as defined by UKB. We used the top 100 LD-pruned GWAS hits from [35], or all 77 GWAS hits for female testosterone. Prior to fitting EFA, we regressed out sex, age, batch, assessment center, and the top 10 genotype PCs from the phenotype.

We assess statistical significance using 499 bootstrap samples to estimate standard errors and confidence intervals on the pathway level interaction. We compute empirical p-values as $2 \cdot \min(B_{\lambda<0}, B_{\lambda>0})$ where $B_{\lambda<0}$ is the number of bootstrap samples with $\lambda < 0$ and $B_{\lambda>0}$ is the number of bootstrap samples with $\lambda > 0$. We also report Gaussian p-values obtained from z-scores computed as the bootstrap mean divided by the bootstrap standard deviation.

To test biological significance, we used the annotations of GWAS hits to biological pathways provided in [35]. We assigned SNPs to the epistasis factors by (1) polarizing SNPs to have positive additive effect sizes and (2) choosing the largest 15 effects on each pathway after subtracting out additive effects from the pathways. To test robustness, we also evaluate results using 15 or 20 SNPs per pathway (**Supplementary Table 3-7**). We calculate p-values for enrichment of biological pathways in epistasis factors with a hypergeometric test.

References

[1] Cutting, G. R. Modifier genes in mendelian disorders: the example of cystic fibrosis. Ann. N. Y. Acad. Sci. **1214**, 57–69 (2010).

[2] Cutting, G. R. Cystic fibrosis genetics: from molecular understanding to clinical application. Nat. Rev. Genet. **16**, 45–56 (2015).

[3] Timberlake, A. T. et al. Two locus inheritance of non-syndromic midline craniosynostosis via rare SMAD6 and common BMP2 alleles. Elife **5** (2016).

[4] Starr, T. N. & Thornton, J. W. Epistasis in protein evolution. Protein Sci. **25**, 1204–1218 (2016).

[5] Bakerlee, C. W., Nguyen Ba, A. N., Shulgina, Y., Rojas Echenique, J. I. & Desai, M. M. Idiosyncratic epistasis leads to global fitness-correlated trends. Science **376**, 630–635 (2022).

[6] Corbett-Detig, R. B., Zhou, J., Clark, A. G., Hartl, D. L. & Ayroles, J. F. Genetic incompatibilities are widespread within species. Nature **504**, 135–137 (2013).

[7] Barton, N. H. How does epistasis influence the response to selection? Heredity (Edinb.) **118**, 96–109 (2017).

[8] Lin, X. et al. Nested epistasis enhancer networks for robust genome regulation. Science **377**, 1077–1085 (2022).

[9] Brem, R. B., Storey, J. D., Whittle, J. & Kruglyak, L. Genetic interactions between polymorphisms that affect gene expression in yeast. Nature **436**, 701–703 (2005).

[10] Shao, H. et al. Genetic architecture of complex traits: large phenotypic effects and pervasive epistasis. Proc. Natl. Acad. Sci. U. S. A. **105**, 19910–19914 (2008).

[11] Sittig, L. J. et al. Genetic background limits generalizability of genotype-phenotype relationships. Neuron **91**, 1253–1259 (2016).

[12] Huang, W. et al. Epistasis dominates the genetic architecture of drosophila quantitative traits. Proc. Natl. Acad. Sci. U. S. A. **109**, 15553–15559 (2012).

[13] Bloom, J. S., Ehrenreich, I. M., Loo, W. T., Lite, T.-L. V. & Kruglyak, L. Finding the sources of missing heritability in a yeast cross. Nature **494**, 234–237 (2013).

[14] Bloom, J. S. et al. Genetic interactions contribute less than additive effects to quantitative trait variation in yeast. Nature communications **6**, 8712 (2015).

[15] Mackay, T. F. C. Epistasis and quantitative traits: using model organisms to study gene–gene interactions. Nature Reviews Genetics **15**, 22–33 (2014).

[16] Forsberg, S. K. G., Bloom, J. S., Sadhu, M. J., Kruglyak, L. & Carlborg, Ö. Accounting for genetic interactions improves modeling of individual quantitative trait phenotypes in yeast. Nature Genetics **49**, 497–503 (2017).

[17] Jiang, Y. & Reif, J. C. Modeling epistasis in genomic selection. Genetics **201**, 759–768 (2015).

[18] Schrauf, M. F. et al. Phantom epistasis in genomic selection: On the predictive ability of epistatic models. G3 (Bethesda) **10**, 3137–3145 (2020).

[19] Hickey, K. L. et al. GIGYF2 and 4EHP inhibit translation initiation of defective messenger RNAs to assist ribosome-associated quality control. Mol. Cell **79**, 950–962.e6 (2020).

[20] Norman, T. M. et al. Exploring genetic interaction manifolds constructed from rich single-cell phenotypes. Science **365**, 786–793 (2019).

[21] Horlbeck, M. A. et al. Mapping the Genetic Landscape of Human Cells. Cell **174**, 953–967.e22 (2018).

[22] Dixit, A. et al. Perturb-Seq: Dissecting Molecular Circuits with Scalable Single-Cell RNA Profiling of Pooled Genetic Screens. Cell **167**, 1853–1866.e17 (2016).

[23] Carlborg, Ö. & Haley, C. S. Epistasis: too often neglected in complex trait studies? Nature Reviews Genetics **5**, 618–625 (2004).

[24] Hill, W. G., Goddard, M. E. & Visscher, P. M. Data and Theory Point to Mainly Additive Genetic Variance for Complex Traits. PLoS Genetics **4**, e1000008 (2008).

[25] Huang, W. & Mackay, T. F. C. The genetic architecture of quantitative traits cannot be inferred from variance component analysis. PLoS Genet. **12**, e1006421 (2016).

[26] Schrode, N. et al. Synergistic effects of common schizophrenia risk variants. Nat. Genet. **51**, 1475–1485 (2019).

[27] Weiner, D. J. et al. Polygenic transmission disequilibrium confirms that common and rare variation act additively to create risk for autism spectrum disorders. Nature Genetics **49**, 978–985 (2017).

[28] Bergen, S. E. et al. Joint contributions of rare copy number variants and common SNPs to risk for schizophrenia. Am. J. Psychiatry **176**, 29–35 (2019).

[29] Manolio, T. A. et al. Finding the missing heritability of complex diseases. Nature **461**, 747–753 (2009).

[30] Hivert, V. et al. Estimation of non-additive genetic variance in human complex traits from a large sample of unrelated individuals. Am. J. Hum. Genet. **108**, 786–798 (2021).

[31] Cockerham, C. C. An Extension of the Concept of Partitioning Hereditary Variance for Analysis of Covariances among Relatives When Epistasis Is Present. Genetics **39**, 859 (1954).

[32] Henderson, C. R. Best linear unbiased prediction of nonadditive genetic merits in noninbred populations. Journal of Animal Science (1985).

[33] Sheppard, B. et al. A model and test for coordinated polygenic epistasis in complex traits. Proceedings of the National Academy of Sciences **118**, e1922305118 (2021). URL `https://www.pnas.org/doi/abs/10.1073/pnas.1922305118`. `https://www.pnas.org/doi/pdf/10.1073/pnas.1922305118`.

[34] Zuk, O., Hechter, E., Sunyaev, S. R. & Lander, E. S. The mystery of missing heritability: Genetic interactions create phantom heritability. Proceedings of the National Academy of Sciences of the United States of America **109**, 1193–1198 (2012).

[35] Sinnott-Armstrong, N., Naqvi, S., Rivas, M. & Pritchard, J. K. Gwas of three molecular traits highlights core genes and pathways alongside a highly polygenic background. eLife **10**, e58615 (2021). URL `https://doi.org/10.7554/eLife.58615`.

[36] Young, A. I. et al. Estimating heritability without environmental bias. BioRxiv 218883 (2017).

[37] Finucane, H. K. et al. Partitioning heritability by functional annotation using genome-wide association summary statistics. Nature Genetics (2015).

[38] Finucane, H. et al. Heritability enrichment of specifically expressed genes identifies disease-relevant tissues and cell types. BioRxiv 103069 (2017).

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

[40] Liu, X., Li, Y. I. & Pritchard, J. K. Trans Effects on Gene Expression Can Drive Omnigenic Inheritance. Cell **177**, 1022–1034.e6 (2019).

[41] Ma, Y., Patil, S., Zhou, X., Mukherjee, B. & Fritsche, L. G. ExPRSweb: An online repository with polygenic risk scores for common health-related exposures. Am. J. Hum. Genet. **109**, 1742–1760 (2022).

[42] Blair, D. R. et al. A nondegenerate code of deleterious variants in Mendelian loci contributes to complex disease risk. Cell **155**, 70–80 (2013).

[43] Vilhjálmsson, B. J. et al. Modeling Linkage Disequilibrium Increases Accuracy of Polygenic Risk Scores. The American Journal of Human Genetics **97**, 576–592 (2015).

[44] Márquez-Luna, C. et al. Incorporating functional priors improves polygenic prediction accuracy in UK biobank and 23andme data sets. Nat. Commun. **12**, 6052 (2021).

[45] Sverdlov, S. & Thompson, E. A. The epistasis boundary: Linear vs. nonlinear genotype-phenotype relationships. bioRxiv (2018). URL `https://www.biorxiv.org/content/early/2018/12/21/503466`. `https://www.biorxiv.org/content/early/2018/12/21/503466.full.pdf`.

[46] Pazokitoroudi, A., Chiu, A. M., Burch, K. S., Pasaniuc, B. & Sankararaman, S. Quantifying the contribution of dominance deviation effects to complex trait variation in biobank-scale data. Am. J. Hum. Genet. **108**, 799–808 (2021).

[47] de Los Campos, G., Sorensen, D. A. & Toro, M. A. Imperfect linkage disequilibrium generates phantom epistasis (& perils of big data). G3 (Bethesda) **9**, 1429–1436 (2019).

[48] Hemani, G. et al. Detection and replication of epistasis influencing transcription in humans. Nature **508**, 249–253 (2014).

[49] Brown, A. A. et al. Genetic interactions affecting human gene expression identified by variance association mapping. eLife **3**, e01381 (2014).

[50] Fish, A. E., Capra, J. A. & Bush, W. S. Are interactions between cis-regulatory variants evidence for biological epistasis or statistical artifacts? Am. J. Hum. Genet. **99**, 817–830 (2016).

[51] Hemani, G. et al. Phantom epistasis between unlinked loci. Nature **596**, E1–E3 (2021).

[52] Sackton, T. B. & Hartl, D. L. Genotypic context and epistasis in individuals and populations. Cell **166**, 279–287 (2016). URL `https://www.sciencedirect.com/science/article/pii/S0092867416308558`.

[53] Saitou, M., Dahl, A., Wang, Q. & Liu, X. Allele frequency differences of causal variants have a major impact on low cross-ancestry portability of prs. medRxiv 2022–10 (2022).

[54] Hou, K. et al. Causal effects on complex traits are similar for common variants across segments of different continental ancestries within admixed individuals. Nature genetics 1–10 (2023).

[55] Rawlik, K., Canela-Xandri, O. & Tenesa, A. Evidence for sex-specific genetic architectures across a spectrum of human complex traits. Genome Biol. **17**, 166 (2016).

[56] Khramtsova, E. A., Davis, L. K. & Stranger, B. E. The role of sex in the genomics of human complex traits. Nat. Rev. Genet. **20**, 173–190 (2019).

[57] Dahl, A. & Zaitlen, N. Genetic influences on disease subtypes. Annu. Rev. Genomics Hum. Genet. **21**, 413–435 (2020).

[58] Oliva, M. et al. The impact of sex on gene expression across human tissues. Science **369** (2020).

[59] Jannink, J.-L. Identifying Quantitative Trait Locus by Genetic Background Interactions in Association Studies. Genetics **176**, 553–561 (2007).

[60] Crawford, L., Zeng, P., Mukherjee, S. & Zhou, X. Detecting epistasis with the marginal epistasis test in genetic mapping studies of quantitative traits. PLoS Genetics **13**, e1006869 (2017).

[61] Akdemir, D. & Jannink, J.-L. Locally epistatic genomic relationship matrices for genomic association and prediction. Genetics **199**, 857–871 (2015).

[62] Rau, C. D. et al. Modeling epistasis in mice and yeast using the proportion of two or more distinct genetic backgrounds: Evidence for "polygenic epistasis". PLoS Genet. **16**, e1009165 (2020).

[63] Darnell, G., Smith, S. P., Udwin, D., Ramachandran, S. & Crawford, L. Partitioning tagged non-additive genetic effects in summary statistics provides evidence of pervasive epistasis in complex traits. bioRxiv (2022). URL `https://www.biorxiv.org/content/early/2022/09/11/2022.07.21.501001`. `https://www.biorxiv.org/content/early/2022/09/11/2022.07.21.501001.full.pdf`.

[64] Turchin, M. C., Darnell, G., Crawford, L. & Ramachandran, S. Pathway analysis within multiple human ancestries reveals novel signals for epistasis in complex traits. bioRxiv (2020). URL `https://www.biorxiv.org/content/early/2020/09/25/2020.09.24.312421`. `https://www.biorxiv.org/content/early/2020/09/25/2020.09.24.312421.full.pdf`.

[65] Demetci, P. et al. Multi-scale inference of genetic trait architecture using biologically annotated neural networks. PLOS Genetics **17**, 1–53 (2021). URL `https://doi.org/10.1371/journal.pgen.1009754`.

[66] Marchini, J., Donnelly, P. & Cardon, L. R. Genome-wide strategies for detecting multiple loci that influence complex diseases. Nature Genetics **37**, 413–417 (2005).

[67] Domingue, B. W., Kanopka, K., Mallard, T. T., Trejo, S. & Tucker-Drob, E. M. Distinguishing between interaction and dispersion effects in the analysis of gene-environment interaction. bioRxiv (2020). URL `https://www.biorxiv.org/content/early/2020/10/16/2020.09.08.287888`. `https://www.biorxiv.org/content/early/2020/10/16/2020.09.08.287888.full.pdf`.

[68] Domingue, B. W., Kanopka, K., Trejo, S., Rhemtulla, M. & Tucker-Drob, E. M. Ubiquitous bias and false discovery due to model misspecification in analysis of statistical interactions: The role of the outcome's distribution and metric properties. Psychol. Methods (2022).

[69] Ballard, J. L. & O'Connor, L. J. Shared components of heritability across genetically correlated traits. Am. J. Hum. Genet. **109**, 989–1006 (2022).

[70] Border, R. et al. Cross-trait assortative mating is widespread and inflates genetic correlation estimates. bioRxiv (2022). URL `https://www.biorxiv.org/content/early/2022/03/23/2022.03.21.485215`. `https://www.biorxiv.org/content/early/2022/03/23/2022.03.21.485215.full.pdf`.

[71] Udler, M. S. et al. Type 2 diabetes genetic loci informed by multi-trait associations point to disease mechanisms and subtypes: A soft clustering analysis. PLoS medicine **15**, e1002654 (2018).

[72] Liley, J., Todd, J. A. & Wallace, C. A method for identifying genetic heterogeneity within phenotypically defined disease subgroups. Nature Genetics **49**, 310–316 (2016).

[73] Werme, J., van der Sluis, S., Posthuma, D. & de Leeuw, C. A. An integrated framework for local genetic correlation analysis. Nat. Genet. **54**, 274–282 (2022).

[74] Kendler, K. S. & Gardner, C. O. Interpretation of interactions: guide for the perplexed. Br. J. Psychiatry **197**, 170–171 (2010).

[75] Gusev, A. et al. Quantifying Missing Heritability at Known GWAS Loci. PLoS Genetics **9**, e1003993–19 (2013).

[76] Gamazon, E. R. et al. A gene-based association method for mapping traits using reference transcriptome data. Nature Genetics **47**, 1091–1098 (2015).

[77] Mukamel, R. E. et al. Protein-coding repeat polymorphisms strongly shape diverse human phenotypes. Science **373**, 1499–1505 (2021).

[78] LaBianca, S. et al. Polygenic profiles define aspects of clinical heterogeneity in adhd. medRxiv (2021). URL `https://www.medrxiv.org/content/early/2021/07/15/2021.07.13.21260299`. `https://www.medrxiv.org/content/early/2021/07/15/2021.07.13.21260299.full.pdf`.

[79] Elliott, L. T. et al. Genome-wide association studies of brain imaging phenotypes in UK biobank. Nature **562**, 210–216 (2018).

[80] Purcell, S. et al. Plink: A tool set for whole-genome association and population-based linkage analyses. The American Journal of Human Genetics **81**, 559–575 (2007). URL `https://www.sciencedirect.com/science/article/pii/S0002929707613524`.

[81] Moscovich, A. & Rosset, S. On the cross-validation bias due to unsupervised preprocessing. Journal of the Royal Statistical Society: Series B (Statistical Methodology) **84**, 1474–1502 (2022). URL `https://rss.onlinelibrary.wiley.com/doi/abs/10.1111/rssb.12537`. `https://rss.onlinelibrary.wiley.com/doi/pdf/10.1111/rssb.12537`.