Bivariate causal mixture model quantifies polygenic overlap between complex traits beyond genetic correlation

Oleksandr Frei[1*], Dominic Holland[2,3&], Olav B. Smeland[1,4&], Alexey A. Shadrin[1], Chun Chieh Fan[2,5,6], Steffen Maeland[1], Kevin S. O'Connell[1], Yunpeng Wang[1,2,6], Srdjan Djurovic[7,8], Wesley K. Thompson[9,10], Ole A. Andreassen[1,4], Anders M. Dale[2,3,6,11*]

Author affiliations

[1]NORMENT, KG Jebsen Centre for Psychosis Research, Institute of Clinical Medicine, University of Oslo, 0424 Oslo, Norway

[2]Center for Multimodal Imaging and Genetics, University of California at San Diego, La Jolla, CA 92037, USA

[3]Department of Neurosciences, University of California, San Diego, La Jolla, CA 92093, USA

[4]Division of Mental Health and Addiction, Oslo University Hospital, 0407 Oslo, Norway

[5]Department of Cognitive Sciences, University of California at San Diego, La Jolla, CA 92093, USA

[6]Department of Radiology, University of California, San Diego, La Jolla, CA 92093, USA

[7]Department of Medical Genetics, Oslo University Hospital, 0424, Oslo, Norway

[8]NORMENT, KG Jebsen Centre for Psychosis Research, Department of Clinical Science, University of Bergen, 5020 Bergen, Norway

[9]University of California, San Diego, Department of Family Medicine and Public Health

[10]Institute of Biological Psychiatry, Mental Health Center Sct. Hans, Capital Region of Denmark

[11]Department of Psychiatry, University of California, San Diego, La Jolla, CA, USA

[&]These authors contributed equally to this work.

[*] To whom correspondence ought to be addressed:

oleksandr.frei@gmail.com, andersmdale@gmail.com

ABSTRACT

Accumulating evidence from genome wide association studies (GWAS) suggests an abundance of shared genetic influences among complex human traits and disorders, such as mental disorders. While current cross-trait analytical methods focus on genetic correlation between traits, we developed a novel statistical tool (MiXeR), which quantifies polygenic overlap independent of genetic correlation, using summary statistics from GWAS. MiXeR results can be presented as a Venn diagram of unique and shared polygenic components across traits. At 90% of SNP-heritability explained for each phenotype, MiXeR estimates that more than 9K variants causally influence schizophrenia, 7K influence bipolar disorder, and out of those variants 6.9K are shared between these two disorders, which have high genetic correlation. Further, MiXeR uncovers extensive polygenic overlap between schizophrenia and educational attainment. Despite a genetic correlation close to zero, these traits share more than 9K causal variants, while 3K additional variants only influence educational attainment. By considering the polygenicity, heritability and discoverability of complex phenotypes, MiXeR provides a more complete quantification of shared genetic architecture than offered by other available tools.

INTRODUCTION

In recent years, genome-wide association studies (GWASs) have successfully detected genetic variants associated with multiple complex human traits or disorders, providing important insights into human biology[1]. Understanding the degree to which complex human phenotypes share genetic influences is critical for identifying the etiology of phenotypic relationships, which can inform disease nosology, diagnostic practice and improve drug development. Most human phenotypes are known to be influenced by multiple genetic variants, many of which are expected to influence more than one phenotype (i.e. exhibit allelic pleiotropy)[2,3]. This has led to cross-trait analyses, quantifying genetic overlap, becoming a widespread endeavor in genetic research, made possible by the public availability of most GWAS summary statistics (p-values and z-scores)[4,5].

Currently, the prevailing measure to quantify genetic overlap is genetic correlation. The sign of the correlation indicates whether the shared genetic effects predominantly have the same or the opposite effect directions. Available methods can quantify genetic correlation using raw genotypes[6,7] or GWAS summary statistics[8-10]. However, these methods report overall positive, negative or no genetic correlation, but do not capture mixtures of effect directions across shared genetic variants. This scenario is exemplified by the genetic relationship between schizophrenia and educational attainment. Despite consistent estimates of a non-significant genetic correlation[11,12], many genetic loci are found to be jointly associated with both phenotypes[13]. Among 25 shared loci[14], 16 had effects in the opposite direction, while 9 had effects in the same direction. Thus, new statistical tools are needed to improve our understanding of the polygenic architecture of complex traits and their intricate relationships.

Here we developed a statistical tool (MiXeR), which quantifies polygenic overlap independent of genetic correlation, using summary statistics from GWAS. To evaluate

polygenic overlap between two traits MiXeR estimates the total number of shared and trait-specific causal variants (i.e. variants with non-zero additive genetic effect on a trait). MiXeR bypasses the intrinsically difficult problem of detecting the exact location of causal variants, but rather aims at estimating their overall amount. MiXeR builds upon univariate causal mixture[15-18], extending the model to four bivariate normal distributions as illustrated in Figure 1, with two causal components specific to each trait; one causal component of variants affecting both traits; and a null component of variants with no effect on either trait. From the prior distribution of genetic effects, we derive likelihood function of the observed signed test statistics (GWAS z-scores), incorporating effects of linkage disequilibrium (LD) structure, minor allele frequency, sample size, cryptic relationships, and sample overlap. The parameters of the mixture model are estimated from the summary statistics by direct optimization of the likelihood function.

We show in simulations that MiXeR provides accurate estimates of model parameters in the presence of realistic LD structure. Using GWAS summary data, we quantify polygenic overlap of several psychiatric disorders, including schizophrenia and bipolar disorder, with educational attainment and human height, with large implications for understanding how genetic factors overlap between complex human phenotypes.

RESULTS

Simulations studies

In our first set of simulations we generate synthetic GWAS data that follow model assumptions and check validity of MiXeR estimates (polygenic overlap, π_{12}; correlation of effect sizes within the shared polygenic component, ρ_{12}, and genetic correlation, rg) in the presence of realistic LD structure (Figure 2). We observe no bias in the estimates across a wide range of simulation scenarios (Supplementary Figure 1-3), except for a specific scenario with correlated effect sizes (ρ_{12}=0.5) and high polygenicity ($\pi_1^u = 3\times10^{-3}$). In this case polygenic overlap (π_{12}) is underestimated, while correlation of effect sizes ρ_{12} is overestimated by the same factor, so that the estimated genetic correlation remains unbiased. This bias in π_{12} and ρ_{12} estimates is attributed to complete sample overlap. Additionally, the scenario with low heritability ($h_2 = 0.1$) and high polygenicity (π_1^u=3×10^{-3}) shows large variation among estimates, which is due to low GWAS signal. Standard errors estimated by the model are shown in Supplementary Table 1.

Additionally we show that univariate estimates of polygenicity and heritability are correct in all scenarios except when heritability is low and polygenicity is high (Supplementary Figure 4, Supplementary Table 2). This case corresponds to insufficiently powered GWAS, yielding large variation among parameters, which leads to the bias from truncation, as the polygenicity parameter is bound to be non-negative.

Finally, we validate that the model accurately predicts GWAS quantile-quantile (Q-Q) plots (Supplementary Figure 5) and detailed Q-Q plots with SNPs partitioned into disjoint groups according to minor allele frequency (MAF) and LD score (Supplementary Figure 6a,b). Detailed Q-Q plots show a stronger GWAS signal for SNPs with higher MAF and higher LD score. The model's prediction follows the same pattern, indicating that it correctly captures dependency of GWAS association statistics on MAF and LD score.

Conditional Q-Q plots (Supplementary Figure 7) show observed versus expected −log10 p-values in the primary trait as a function of significance of association with a secondary trait at the level of $p \leq 0.1$, $p \leq 0.01$, $p \leq 0.001$, with data Q-Q plots being closely reproduced by the model's predictions. Interestingly, scenarios without polygenic overlap are also showing a minor enrichment, arising because GWAS p-values depend on allele frequency and LD structure, though this effect is generally smaller than enrichment arising due to shared causal variants.

Sensitivity analysis

For sensitivity analysis, we conducted simulations with traits that have a shared pattern of differential enrichment of heritability across genomic categories[19], which is not accounted for by MiXeR model. Simulations were informed by the enrichment pattern of schizophrenia[20], as estimated by stratified LD score regression[21] (Supplementary Figure 8a). In the univariate analysis, polygenicity was underestimated by about 20% (Supplementary Figure 9), indicating that the model may group adjacent causal variants together and interpret them as a single cluster. In the bivariate analysis we observe a small upwards bias in the estimate of polygenic overlap (Supplementary Table 3), but it did not exceed 10% of the polygenicity across all sufficiently powered scenarios.

Another assumption of MiXeR model is that effect sizes are independent of allele frequencies. We run simulations where, in addition to differential enrichment of genomic categories, all causal variants equally contribute to heritability regardless of their allele frequency (as modeled by stratified LDSR), as opposite to MiXeR assumption where causal variants contribute depending on their allele frequency. The results are showing that heritability is underestimated by 10% to 40% of its true value, while polygenicity is underestimated by a factor of 10 (Supplementary Figure 10), i.e. a larger bias than was observed in simulations with genomic annotations alone. The implications are less notable for the bivariate model: the estimated number of shared causal variants is consistent with the actual number arising by chance due to high polygenicity (Supplementary Table 4).

Finally, we run simulations with incomplete reference, and simulate phenotypes where causal variants are spread across our entire reference of N=11,015,833 variants, but only a fraction (50%, 25% or 12.5%) of the variants enter LD structure estimation and fit procedure. The results (Supplementary Table 5) show that the total number of causal SNPs, as well as the heritability, are estimated correctly, while polygenicity parameter is different from simulated value, because it reflects the fraction of all tagged causal variants with respect to the reference that went into LD structure estimation.

GWAS Summary Statistics

We apply MiXeR to summary statistics from GWAS representing 7 phenotypes[11,20,22-26] (see Supplementary Table 6 for metadata about the studies).

MiXeR estimates of genetic correlation (Table 1, Supplementary Table 7) were generally consistent with those of cross-trait LD Score Regression[8], with the highest genetic correlation observed between schizophrenia and bipolar disorder. Naturally, these disorders also exhibit substantial polygenic overlap, sharing 6.9K out of 9.5K causal variants involved. Here and below the numbers of causal variants are reported as 22.6%

of their total estimate, which jointly accounts for 90% of SNP heritability in each phenotype, to avoid extrapolating model parameters into the area of infinitesimally small effects (Supplementary Figure 11).

Further, MiXeR reveals important differences among traits with low genetic correlation, represented as Venn diagrams of shared and unique polygenic components (Figure 3, Supplementary Figures 12, 13a-g). For example, schizophrenia and educational attainment exhibit substantial polygenic overlap, sharing 9.1K out of 12.1K of causal variants involved. On the contrary, schizophrenia and height share only about 1.1K out of 11.9K causal variants. Intriguingly, educational attainment and height also show low polygenic overlap, sharing 1.8K out of 14.0K causal variants. Nevertheless, these traits have relatively high correlation of effect sizes within the shared component, $\rho_{12} = 0.44$ (0.03), which at genome-wide level is observed as genetic correlation of rg=0.12 (0.01) according to MiXeR, or rg=0.14 (0.01) according to LDSR.

MiXeR estimates of the unique polygenic components provide insight into the trait-specific genomic architecture. For example, while schizophrenia has 2.2K causal variants not shared with bipolar disorder, only 0.1K are not shared with educational attainment, and as many as 8.1K are non-overlapping with height. Also, for the other phenotypes the number of trait-specific causal variants varies across different pairs of traits (Figure 3).

Figure 4 and Supplementary Figures 14a-g visualize bivariate density of the observed GWAS signed test statistics (z_{1j}, z_{2j}), the predicted density $(\hat{z}_{1j}, \hat{z}_{2j})$ from MiXeR model, and estimated bivariate density of additive causal effects (β_{1j}, β_{2j}) that underlie model prediction. Figure 4 gives real examples for the three different scenarios of polygenic overlap (genetically independent traits, polygenic overlap with and without genetic correlation, as previously shown by Figure 1). Finally, we use conditional Q-Q plots[27,28], where a consecutive deflection of the curves indicates polygenic overlap, and shows that MiXeR-based prediction provides accurate estimates of the data Q-Q plots (Figure 5).

DISCUSSION

MiXeR is a novel method for cross-trait analysis of GWAS summary statistics, which enables a more complete quantification of polygenic overlap than provided by other existing tools[8,9,29,30]. In addition to genetic correlation, MiXeR estimates the total number of shared and trait-specific causal variants, providing new information into the genetic relationships between complex traits and disorders.

MiXeR extends cross-trait LD score regression[8] by incorporating a causal mixture model[15-18], thus relying on a biologically more plausible prior distribution of genetic effect sizes compared to the "infinitesimal" model[31,32]. We show that polygenicity, measured as a total number of causal variants, and discoverability[33], measured as variance of individual causal variants, has major implications on the future of GWAS discoveries (Supplementary Figure 15).

Applying MiXeR to real phenotype data, we provide new insights into the genetic relationships between schizophrenia, bipolar disorder, educational attainment and height. In line with the strong clinical relationship[34] between schizophrenia and bipolar disorder,

and prior genetic studies[25,35], we find substantial polygenic overlap between these two disorders. Intriguingly, both schizophrenia and bipolar disorder are estimated to have a small fraction of causal variants conferring individual risk of a specific disorder (Figure 3). Identifying such genetic variants could provide critical knowledge about the distinct genetic risk architectures underlying these psychiatric disorders. Moreover, we find that nearly all causal variants influencing schizophrenia risk also appear to influence educational attainment, despite a genetic correlation close to zero (Table 1). This is in line with recent studies demonstrating shared genetic loci between schizophrenia and educational attainment[14] and a strong genetic dependence between the phenotypes possibly related to different subtypes of schizophrenia[13]. In contrast, 85% of genetic variants influencing bipolar disorder also appear to influence educational attainment, but there is a significant positive genome-wide correlation of 0.20(0.03) in compliance with the cross-trait LD score regression estimate of 0.18(0.03) (Table 1, Supplementary Table 7).

We show that polygenicity is best expressed as a total number of causal variants (Supplementary Table 5). Previous studies presented it as a fraction, which is highly dependent on the used reference (1.1M hapmap in ref[17], or 484K Affymetrix SNPs in ref[18]). When expressed as a total number, or estimates of polygenicity for schizophrenia, bipolar disorder, educational attainment and height are consistent with previously reported results. In addition, we estimate that just 5% of causal variants are needed to explain 50% of heritability, and 22.6% of causal variants are needed to explain 90% of heritability (Supplementary Figure 11). These numbers are expected to be less dependent on modeling assumptions, because with finite GWAS sample it is not possible to distinguish small effects from truly null effects. The actual number of causal variants is, potentially, even higher, as our model tends to clump together variants if they are located too close to each other (Supplementary Tables 3, 4).

Some existing methods can already uncover polygenic overlap in the absence of genetic correlation. For example, conjFDR analysis[27,28] is a non-parametric model-free approach, which detects shared genetic loci regardless of their allelic effect directions, by prioritizing variants with strong associations across more than one GWAS[36]. Other methods, including gwas-pw[37] and HESS[29], also aim at detecting genomic loci jointly associated with two traits. MiXeR complements these methods by providing an easily interpretable high-level overview of the shared and unique genomic architectures underlying complex phenotypes.

MiXeR has some notable advantages compared to the existing methods that implement causal mixture. First, our mathematical model for the likelihood term $p(z_j|\beta_j)$ is conceptually simpler and more flexible, resulting in unbiased estimates of model parameters across a wide range of simulation scenarios (Supplementary Figure 1-3) and providing accurate prediction of GWAS z-scores across varying ranges of MAF and LD (Supplementary Figure 6a,b). Second, MiXeR implementation works well with a large reference of 10M variants, while other methods have reduced it to 1.1M HapMap SNPs (ref [17]) or 484K Affymetrix SNPs (ref [18]). Finally, our model individually processes all SNPs, without grouping them into bins (ref [15]).

MiXeR models causal effects as a single gaussian component, while recent work[17,38] suggests that certain phenotypes, including height, require at least two causal components of small and large effects. We note that MiXeR model still provides good fit for SNPs not reaching GWAS threshold (Supplementary Figure 16) and shows deviations only towards the tail of the distribution. To further investigate the effects of model misspecification we implemented right-censoring of genome-wide significant SNPs (see Online methods). Results (Supplementary Tables 7, 8) are consistent with our main analysis, except for height which received a lower estimate of heritability (65% instead of 70%), a slight increase in polygenicity, and increased polygenic overlap with other traits. We advocate that for a better estimate of height's polygenicity it would be beneficial to run MiXeR on a residualized GWAS, after covarying association statistics for genotypes of all genome-wide significant SNPs.

Recent work suggests the importance of MAF- and LD-dependent genetic architectures[18,39], which are not directly modeled by MiXeR. Our simulations with an extreme case of a different MAF model shows 10-40% underestimation of heritability (Supplementary Figure 10), but less noticeable effect on the relative proportion of shared causal variants (Supplementary Table 4). On real data we observe effects of MAF-dependent architectures by drawing Q-Q plots for subsets of SNPs (Supplementary Figures 17a-g) partitioned into 9 groups according to minor allele frequency (MAF) and LD score, where the model tends to underestimate z-scores in low MAF bins. This effect, however, is quite subtle, and does not manifest itself on the overall Q-Q plots (Supplementary Figure 16).

The MiXeR method requires large GWAS studies. Our recommendation is to apply MiXeR to studies with at least N=50 000 participants, and inspect standard errors reported by MiXeR. Polygenicity estimation requires more GWAS power than heritability estimation, which can be visually explained by GWAS Q-Q plots (Supplementary Figure 16): heritability is determined by the overall departure of the GWAS curve from the null line, while polygenicity is determined by its curvature, i.e. the point where the GWAS curve begins to bend upwards from the null line, which is harder to estimate when GWAS signal is weak. This is captured by MiXeR standard errors, which show that individual parameters of the mixture model have lower estimation accuracy than their combinations – for example, relative errors for π_1 and σ_β^2 are larger than for the heritability estimate $h^2 \propto \pi_1 \sigma_\beta^2$, due to inversely-correlated errors (Supplementary Table 2). Despite these limitations, there is still a clear minimum on the energy landscape of cost function (Supplementary Figures 18, 19, showing log-likelihood as a function of model parameters around the optimum).

In our future work we are planning to incorporate an additional gaussian component to model small and large effects[17], and explicitly account for MAF-dependent architectures[39]. Further extensions may account for differential enrichment for true associations across genomic annotations[19]. Another limitation to address is that MiXeR model assumes similar LD structure among studies and is not currently applicable for analysis across different ethnicities. We aim to extend the MiXeR modelling framework to be used to improve power for SNP discovery by estimating the posterior effect size of SNPs associated with

one trait given the test statistics in another trait, as well as for improving predictive power of polygenic risk scores.

In conclusion, MiXeR represents a useful addition to the tool-box for cross-trait GWAS analysis. By taking into account the intricate polygenic architectures of complex phenotypes MiXeR allows for measures of polygenic overlap beyond genetic correlation. We expect this to lead to new insights into the pleiotropic nature of human genetic etiology.

ACKNOWLEDGMENTS:

This work was supported by the Research Council of Norway (#223273, #225989, #248778) South-East Norway Health Authority (#2016-064, #2017-004), KG Jebsen Stiftelsen (#SKGJ-Med-008), and National Institutes of Health (R01MH100351, R01GM104400). The simulations were performed on resources provided by UNINETT Sigma2 - the National Infrastructure for High Performance Computing and Data Storage in Norway.

AUTHOR CONTRIBUTIONS:

Conceived and designed the study – AMD, OAA, OF, Method development – AMD, OF, DH, AAS, Analysis and interpretation of results – OF, DH, AAS, OBS, CCF, YW, AW, SD, Drafting manuscript – OF, OAA, OBS, Revision and approval of final manuscript – ALL AUTHORS

URLS:

https://github.com/precimed/ - MiXeR code (MATLAB/Octave), will be released after publication; https://data.broadinstitute.org/alkesgroup/LDSCORE/ - LD scores and reference panel derived from 1000 Genomes phase 3.

COMPETING FINANCIAL INTERESTS: The authors declare no competing financial interests.

CORRESPONDING AUTHORS:

Oleksandr Frei (oleksandr.frei@gmail.com)
Anders M. Dale (andersmdale@gmail.com)

FIGURES AND TABLES

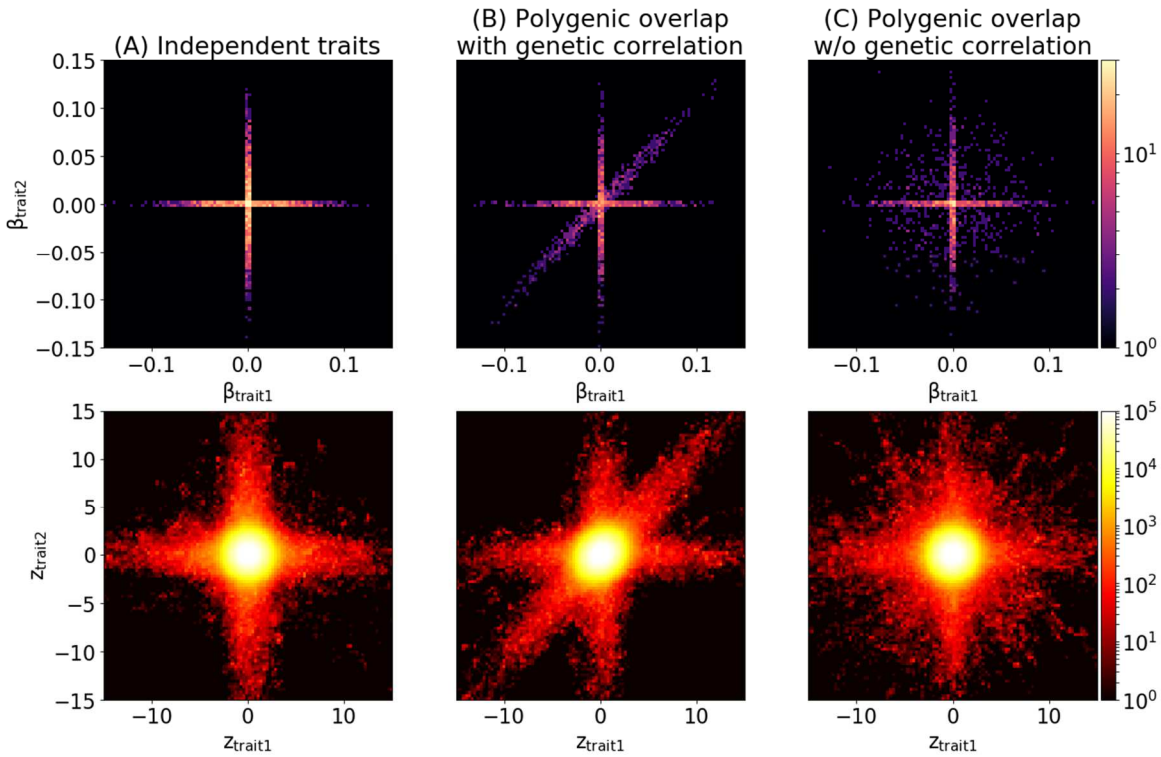

Figure 1. Components of the bivariate mixture in three scenarios of polygenic overlap. All figures are generated from synthetic data, where causal variants were drawn from MiXeR model, total polygenicity in each trait set to 0.01%, SNP heritability set to 0.4, GWAS N=100 000. First column shows two traits where causal variants do not overlap. Second column adds a component of causal variants affecting both traits in the same (concordant) direction. Third column shows scenario of polygenic overlap without genetic correlation. Top row shows simulated bivariate density of additive effects of allele substitution (β_{1j}, β_{2j}), bottom row shows bivariate density of GWAS signed test statistics (z_{1j}, z_{2j}) for GWAS SNPs (genotyped or imputed). Due to Linkage Disequilibrium, GWAS signed test statistic has substantially larger volume of SNPs associated with the phenotype. The aim of MiXeR model is to infer distribution of causal effects (top row), using GWAS data (bottom row) as an input. Figures are generated on a regular grid of 100x100 bins, color histogram indicates log10(N) where N is the number of SNPs projected into a bin.

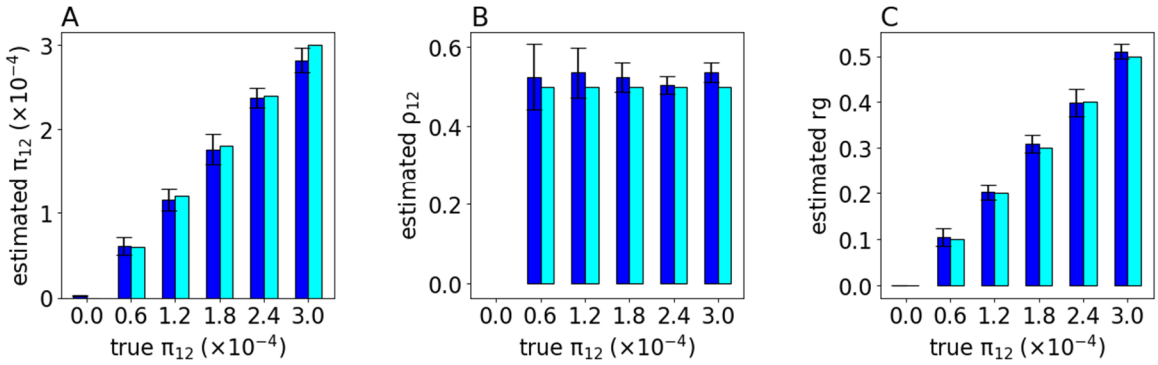

Figure 2. Selected simulations with bivariate model: (A) estimates of polygenic overlap; (B) estimates of correlation of the effect sizes in shared polygenic component; (C) estimates of genetic correlation. The bars in blue indicate an average value of model estimates across 10 simulation runs. The bars in cyan show true (simulated) parameters. Error bars represent standard deviation of the model estimate across 10 simulation runs. Different bars correspond to levels of polygenic overlap: from zero (no overlap) to complete polygenic overlap. Simulated heritability is 0.4, simulated fraction of causal variants is 0.03% in both traits.

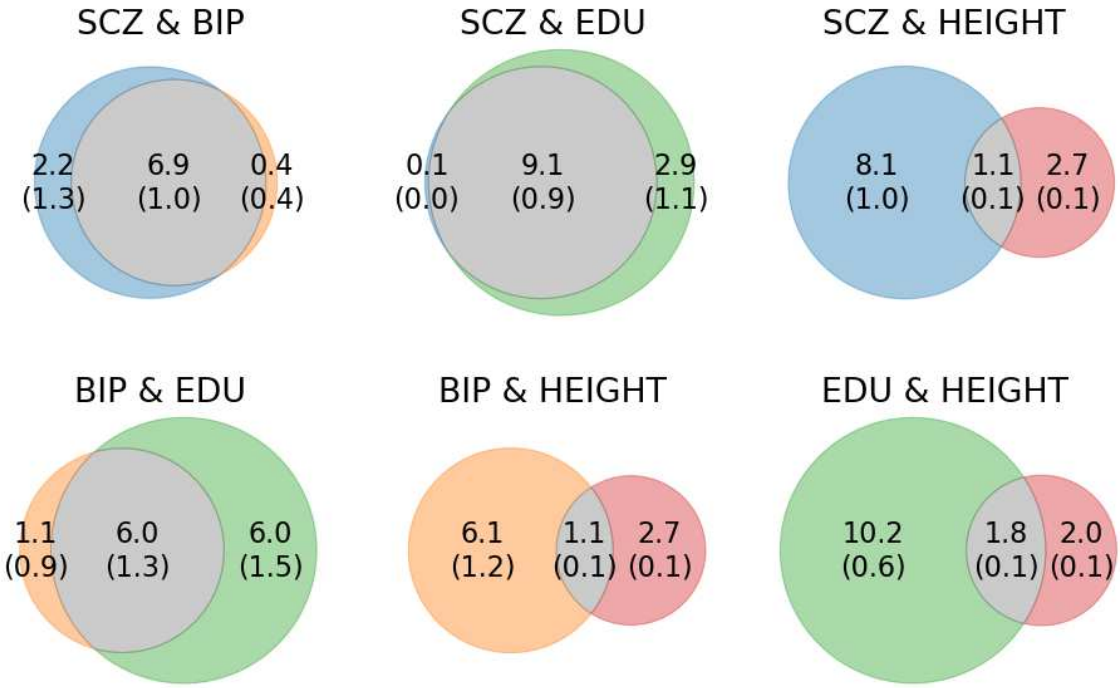

Figure 3. Venn diagrams of unique and shared polygenic component at the causal level, showing polygenic overlap (grey) between schizophrenia (SCZ, blue), bipolar disorder (BIP, orange), educational attainment (EDU, green) and height (red). The numbers indicate estimated quantity of causal variants (in 1,000) per component, explaining 90% of SNP heritability in each phenotype, followed by the standard error. The size of circles reflects the polygenicity.

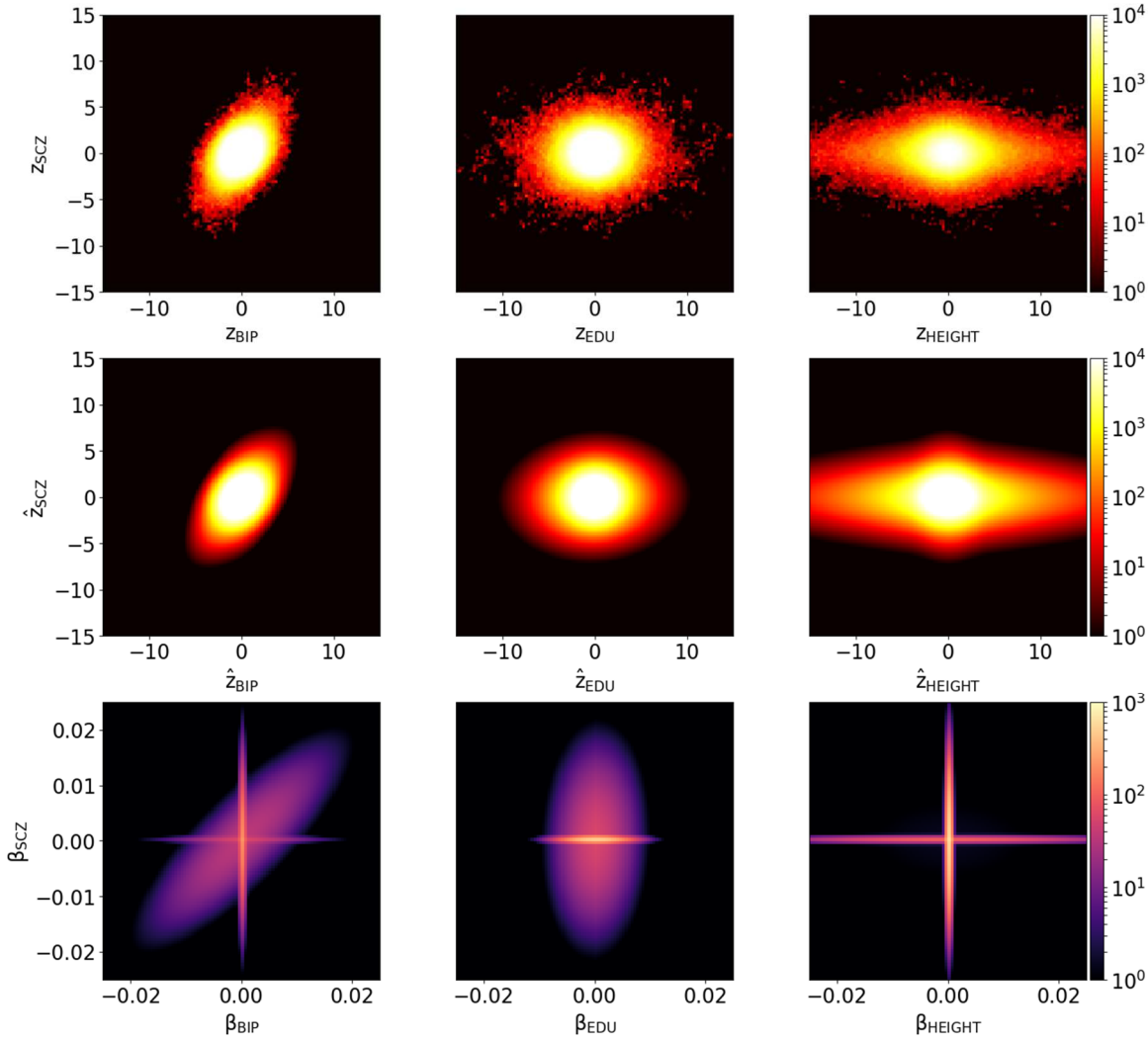

Figure 4. Top row shows bivariate density of the observed GWAS signed test statistics (z_{1j}, z_{2j}), middle row shows predicted density $(\hat{z}_{1j}, \hat{z}_{2j})$ from MiXeR model. Bottom row shows estimated bivariate density of additive causal effects (β_{1j}, β_{2j}) that underlie model prediction. Three columns represent schizophrenia (SCZ) versus bipolar disorder (BIP), educational attainment (EDU) and height GWAS. Density is visualized using regular grid of 100x100 bins, color indicates log10(N) where N is the observed number (for the top row) or the expected number (for the middle and bottom rows) of SNPs projected into a bin.

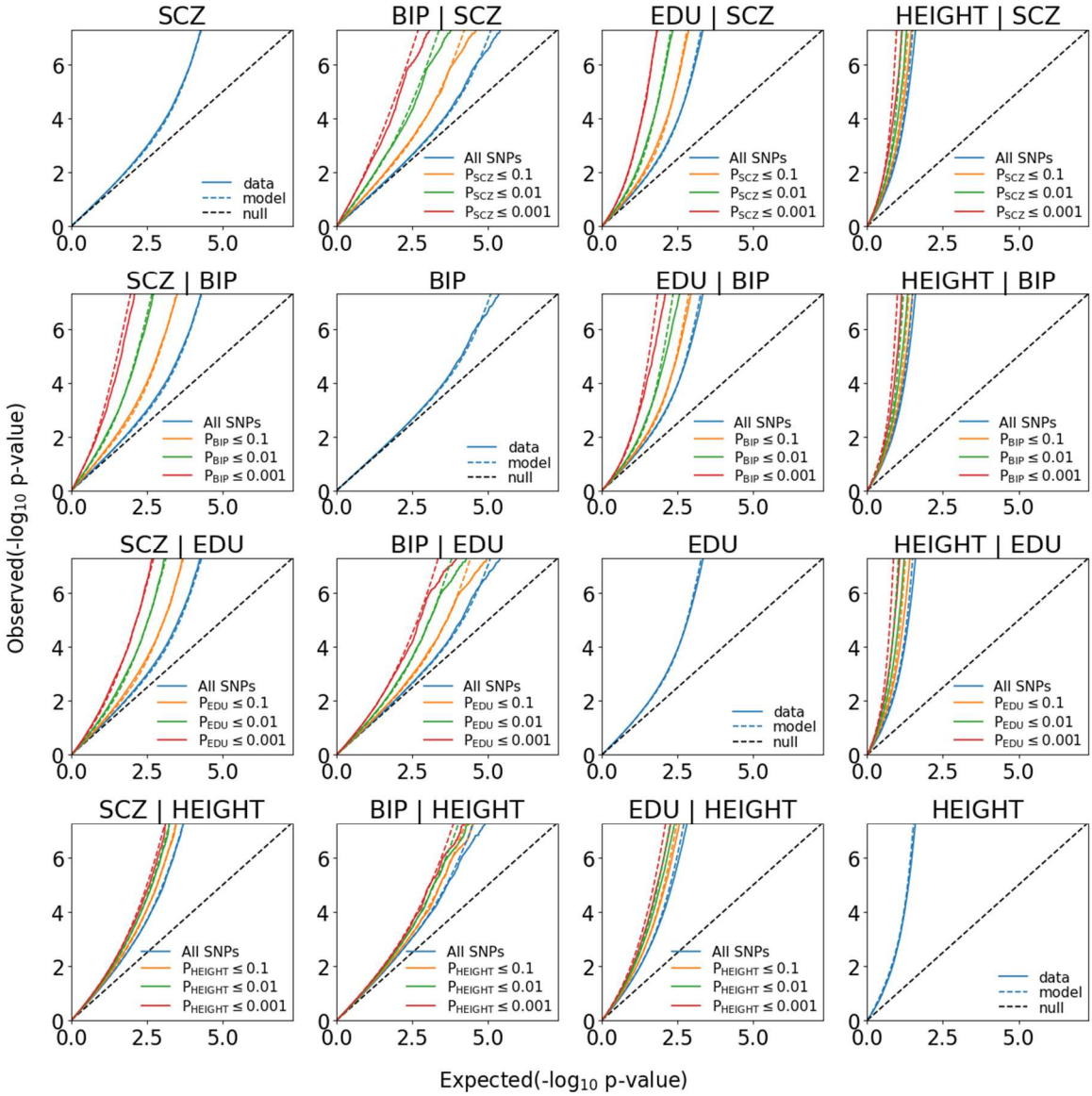

Figure 5. Conditional Q-Q plots of observed versus expected $-\log_{10}$ p-values in the primary trait as a function of significance of association with a secondary trait at the level of p≤0.1, p≤0.01, p≤0.001. Blue line indicates all SNPs. Dotted lines indicate model predictions for each stratum. Points on the QQ plot are weighted according to LD structure, using n=64 iterations of random pruning at LD threshold r2=0.1.

trait1	trait2	n_{12} (se)	n_1 (se)	n_2 (se)	ρ_{12} (se)	rg (se)	rg_{LDSR} (se)
SCZ	BIP	6.88 (1.04)	2.25 (1.34)	0.36 (0.43)	0.837 (0.019)	0.708 (0.062)	0.717 (0.024)
SCZ	EDU	9.09 (0.92)	0.08 (0.04)	2.94 (1.12)	0.056 (0.015)	0.048 (0.014)	0.079 (0.022)
SCZ	Height	1.09 (0.12)	8.12 (0.97)	2.72 (0.14)	-0.007 (0.057)	-0.001 (0.011)	-0.008 (0.019)
BIP	EDU	6.04 (1.34)	1.07 (0.94)	5.99 (1.49)	0.308 (0.054)	0.201 (0.030)	0.176 (0.026)
BIP	Height	1.08 (0.13)	6.05 (1.24)	2.71 (0.15)	0.000 (0.064)	0.000 (0.013)	-0.011 (0.023)
EDU	Height	1.83 (0.11)	10.19 (0.64)	2.02 (0.11)	0.442 (0.034)	0.119 (0.009)	0.141 (0.012)

Table 1. Results of cross-trait analysis with MiXeR model for schizophrenia (SCZ), bipolar disorder (BIP), educational attainment (EDU) and height GWAS. Columns: n_{12} – estimated number of shared causal variants, reported in 1,000; n_1 (n_2)– estimated number of causal variants, unique to trait1 (trait2), expressed in 1,000; ρ_{12} – correlation of effect sizes in shared polygenic component; rg – genetic correlation ($r_g = \rho_{12}\pi_{12}/\sqrt{\pi_1^u \pi_2^u}$, see Online Methods); rg_{LDSR} - estimate of genetic correlation from LD Score Regression. Number of variants (n_{12}, n_1 and n_2) are adjusted to explain 90% of heritability in the corresponding component. Parameters are fitted using ca. 1.1M HapMap3 SNPs.

REFERENCES

1. Visscher, P.M. *et al.* 10 Years of GWAS Discovery: Biology, Function, and Translation. *Am J Hum Genet* **101**, 5-22 (2017).
2. Sivakumaran, S. *et al.* Abundant pleiotropy in human complex diseases and traits. *Am J Hum Genet* **89**, 607-18 (2011).
3. Solovieff, N., Cotsapas, C., Lee, P.H., Purcell, S.M. & Smoller, J.W. Pleiotropy in complex traits: challenges and strategies. *Nat Rev Genet* **14**, 483-95 (2013).
4. Schork, A.J., Wang, Y., Thompson, W.K., Dale, A.M. & Andreassen, O.A. New statistical approaches exploit the polygenic architecture of schizophrenia--implications for the underlying neurobiology. *Curr Opin Neurobiol* **36**, 89-98 (2016).
5. Pasaniuc, B. & Price, A.L. Dissecting the genetics of complex traits using summary association statistics. *Nat Rev Genet* **18**, 117-127 (2017).
6. Dudbridge, F. Power and Predictive Accuracy of Polygenic Risk Scores. *PLOS Genetics* **9**, e1003348 (2013).
7. Yang, J., Lee, S.H., Goddard, M.E. & Visscher, P.M. GCTA: a tool for genome-wide complex trait analysis. *Am J Hum Genet* **88**, 76-82 (2011).
8. Bulik-Sullivan, B. *et al.* An atlas of genetic correlations across human diseases and traits. *Nat Genet* **47**, 1236-41 (2015).
9. Verhulst, B., Maes, H.H. & Neale, M.C. GW-SEM: A Statistical Package to Conduct Genome-Wide Structural Equation Modeling. *Behav Genet* **47**, 345-359 (2017).
10. Davey Smith, G. & Hemani, G. Mendelian randomization: genetic anchors for causal inference in epidemiological studies. *Hum Mol Genet* **23**, R89-98 (2014).
11. Lee, J.J. *et al.* Gene discovery and polygenic prediction from a genome-wide association study of educational attainment in 1.1 million individuals. *Nat Genet* **50**, 1112-1121 (2018).
12. Okbay, A. *et al.* Genome-wide association study identifies 74 loci associated with educational attainment. *Nature* **533**, 539-42 (2016).
13. Bansal, V. *et al.* Genome-wide association study results for educational attainment aid in identifying genetic heterogeneity of schizophrenia. *Nature Communications* **9**, 3078 (2018).
14. Le Hellard, S. *et al.* Identification of Gene Loci That Overlap Between Schizophrenia and Educational Attainment. *Schizophr Bull* **43**, 654-664 (2017).
15. Holland, D. *et al.* Estimating Degree Of Polygenicity, Causal Effect Size Variance, And Confounding Bias In GWAS Summary Statistics. *bioRxiv* (2017).
16. Zhu, X. & Stephens, M. Bayesian large-scale multiple regression with summary statistics from genome-wide association studies. *Ann. Appl. Stat.* **11**, 1561-1592 (2017).
17. Zhang, Y., Qi, G., Park, J.-H. & Chatterjee, N. Estimation of complex effect-size distributions using summary-level statistics from genome-wide association studies across 32 complex traits. *Nature Genetics* **50**, 1318-1326 (2018).
18. Zeng, J. *et al.* Signatures of negative selection in the genetic architecture of human complex traits. *Nature Genetics* **50**, 746-753 (2018).
19. Schork, A.J. *et al.* All SNPs are not created equal: genome-wide association studies reveal a consistent pattern of enrichment among functionally annotated SNPs. *PLoS Genet* **9**, e1003449 (2013).
20. Schizophrenia Working Group of the Psychiatric Genomics, C. Biological insights from 108 schizophrenia-associated genetic loci. *Nature* **511**, 421-7 (2014).

21. Finucane, H.K. *et al.* Partitioning heritability by functional annotation using genome-wide association summary statistics. *Nat Genet* **47**, 1228-35 (2015).
22. Grove, J. *et al.* Common risk variants identified in autism spectrum disorder. *bioRxiv* (2017).
23. Demontis, D. *et al.* Discovery Of The First Genome-Wide Significant Risk Loci For ADHD. *bioRxiv* (2017).
24. Wray, N.R. *et al.* Genome-wide association analyses identify 44 risk variants and refine the genetic architecture of major depression. *Nature Genetics* **50**, 668-681 (2018).
25. Stahl, E. *et al.* Genomewide association study identifies 30 loci associated with bipolar disorder. *bioRxiv* (2017).
26. Yengo, L. *et al.* Meta-analysis of genome-wide association studies for height and body mass index in approximately 700000 individuals of European ancestry. *Hum Mol Genet* **27**, 3641-3649 (2018).
27. Andreassen, O.A. *et al.* Improved detection of common variants associated with schizophrenia by leveraging pleiotropy with cardiovascular-disease risk factors. *Am J Hum Genet* **92**, 197-209 (2013).
28. Andreassen, O.A. *et al.* Improved detection of common variants associated with schizophrenia and bipolar disorder using pleiotropy-informed conditional false discovery rate. *PLoS Genet* **9**, e1003455 (2013).
29. Shi, H., Mancuso, N., Spendlove, S. & Pasaniuc, B. Local Genetic Correlation Gives Insights into the Shared Genetic Architecture of Complex Traits. *Am J Hum Genet* **101**, 737-751 (2017).
30. Turley, P. *et al.* Multi-trait analysis of genome-wide association summary statistics using MTAG. *Nature Genetics* **50**, 229-237 (2018).
31. Park, J.H. *et al.* Estimation of effect size distribution from genome-wide association studies and implications for future discoveries. *Nat Genet* **42**, 570-5 (2010).
32. Gibson, G. Rare and common variants: twenty arguments. *Nat Rev Genet* **13**, 135-45 (2012).
33. Fan, C.C. *et al.* Beyond heritability: improving discoverability in imaging genetics. *Hum Mol Genet* **27**, R22-r28 (2018).
34. Keshavan, M.S., Nasrallah, H.A. & Tandon, R. Schizophrenia, "Just the Facts" 6. Moving ahead with the schizophrenia concept: from the elephant to the mouse. *Schizophr Res* **127**, 3-13 (2011).
35. Ruderfer, D.M. *et al.* Genomic Dissection of Bipolar Disorder and Schizophrenia, Including 28 Subphenotypes. *Cell* **173**, 1705-1715.e16 (2018).
36. Smeland, O.B. *et al.* Identification of Genetic Loci Jointly Influencing Schizophrenia Risk and the Cognitive Traits of Verbal-Numerical Reasoning, Reaction Time, and General Cognitive Function. *JAMA Psychiatry* (2017).
37. Pickrell, J.K. *et al.* Detection and interpretation of shared genetic influences on 42 human traits. *Nat Genet* **48**, 709-17 (2016).
38. Moser, G. *et al.* Simultaneous Discovery, Estimation and Prediction Analysis of Complex Traits Using a Bayesian Mixture Model. *PLOS Genetics* **11**, e1004969 (2015).
39. Speed, D. *et al.* Reevaluation of SNP heritability in complex human traits. *Nature genetics* **49**, 986-992 (2017).

ONLINE METHODS

This article is accompanied by a Supplementary Note with further details.

Bivariate causal mixture model. Consider simple additive model of genetic effects, ignoring gene-environment interactions, epistasis and dominance effects. Under these assumptions, the contribution of the genotype to the phenotype is modelled as a sum of individual contributions from genetic variants: $y_k = \sum_j g_{jk}\,\beta_j$, where y_k is a quantitative phenotype or disease liability of k-th individual, g_{jk} is 0,1,2-coded number of reference alleles for j-th variant, and β_j is additive genetic effect of allele substitution. We say that genetic variant is causal for a trait if it has non-zero effect on that trait ($\beta_j \neq 0$).

MiXeR builds upon univariate causal mixture model[15], $\beta_j \sim \pi_0 N(0,0) + \pi_1 N\left(0, \sigma_\beta^2\right)$, which assumes that only a small fraction (π_1) of variants has an effect on the trait, while the effect of the remaining variants is zero. In a joint analysis of two traits we expect some variants to affect both traits; some variants to affect one trait but not the other; and most variants to have no effect on either trait. Based on these assumptions, MiXeR models additive genetic effects β_{1j}, β_{2j} of variant j on the two traits as a mixture of four bivariate Gaussian components (Figure 1):

$$\left(\beta_{1j}, \beta_{2j}\right) \sim \pi_0 N(0,0) + \pi_1 N(0, \Sigma_1) + \pi_2 N(0, \Sigma_2) + \pi_{12} N(0, \Sigma_{12}), \qquad (1)$$

$$\Sigma_1 = \begin{bmatrix} \sigma_1^2 & 0 \\ 0 & 0 \end{bmatrix},\ \Sigma_2 = \begin{bmatrix} 0 & 0 \\ 0 & \sigma_2^2 \end{bmatrix},\ \Sigma_{12} = \begin{bmatrix} \sigma_1^2 & \rho_{12}\sigma_1\sigma_2 \\ \rho_{12}\sigma_1\sigma_2 & \sigma_2^2 \end{bmatrix}, \qquad (2)$$

where π_1 and π_2 are weights of the unique components (variants with an effect on the first only, and on the second trait only); π_{12} is a weighting of the component affecting both traits; and π_0 is a fraction of variants that are non-causal for both traits, $\pi_0 + \pi_1 + \pi_2 + \pi_{12} = 1$; σ_1^2 and σ_2^2 control expected magnitudes of per-variant effect sizes; and ρ_{12} is the correlation coefficient of the effect sizes in the shared component. All parameters are assumed to be the same for all genetic variants.

The effects $(\hat{\beta}_{1j}, \hat{\beta}_{2j})$ estimated by a GWAS, represent only proxies of the true causal effects (β_{1j}, β_{2j}), which are distorted by limited sample size (poor statistical power), cryptic relatedness within a GWAS sample as well as LD between variants. To disentangle these effects we derive the likelihood term for observed GWAS signed test statistics (z_{1j}, z_{2j}), incorporating effects of LD structure (allelic correlation r_{ij} between variants i and j); heterozygosity $H_j = 2p_j(1 - p_j)$ where p_j is the minor allele frequency of the j-th variant; number of subjects genotyped per variant (N_{1j} and N_{2j}); and variance distortion parameters σ_{01}^2, σ_{02}^2 , and ρ_0. Specifically (see Supplementary Note),

$$\left(z_{1j}, z_{2j}\right) = \left(\delta_{1j}, \delta_{2j}\right) + N\left((0,0), \begin{bmatrix} \sigma_{01}^2 & \rho_0\sigma_{01}\sigma_{02} \\ \rho_0\sigma_{01}\sigma_{02} & \sigma_{02}^2 \end{bmatrix}\right), \qquad (3)$$

$$\delta_{\cdot j} = \sqrt{N_{\cdot j}} \sum_i \sqrt{H_i} r_{ij} \beta_{\cdot j}$$

The nine parameters of the model $(\pi_1, \pi_2, \pi_{12}, \sigma_1^2, \sigma_2^2, \rho_{12}, \sigma_{01}^2, \sigma_{02}^2, \rho_0)$ are fit by direct optimization of the weighted log likelihood, with standard errors estimated from the Observed Fishers Information matrix.

Forcing $\pi_{12} = 1$ (so that $\pi_0 = \pi_1 = \pi_2 = 0$) reduces our model to an "infinitesimal" assumption that underlie cross-trait LD score regression[8]. Under this constraint our model predicts that GWAS signed test statistics follow bivariate Gaussian distribution with zero mean and variance-covariance matrix

$$\Sigma_j = \ell_j \begin{bmatrix} N_{1j}\sigma_1^2 & \sqrt{N_{1j}N_{2j}}\rho_{12}\sigma_1\sigma_2 \\ \sqrt{N_{1j}N_{2j}}\rho_{12}\sigma_1\sigma_2 & N_{2j}\sigma_2^2 \end{bmatrix} + \begin{bmatrix} \sigma_{01}^2 & \rho_0\sigma_{01}\sigma_{02} \\ \rho_0\sigma_{01}\sigma_{02} & \sigma_{02}^2 \end{bmatrix},$$

i.e., $(z_{1j}, z_{2j}) \sim N(0, \Sigma_j)$, where $\ell_j = \sum_i H_i r_{ij}^2$ is the LD score (adjusted for heterozygosity). This model is consistent with cross-trait LD score regression, with expected chi square statistics $E(z_{1j}^2)$, $E(z_{2j}^2)$ and cross-trait covariance $E(z_{1j}z_{2j})$ being proportional to the LD score of j-th SNP, and parameters ρ_0, σ_{01}, σ_{02} playing the role of LD score regression intercepts[40]. The only distinction here is that we choose to model effect sizes that are independent of allele frequency, leading to the incorporation of H_i in our model; this factor is absent from the LD score regression model due to the assumption there of effect sizes that are inversely proportional to H_i. Thus, MiXeR is a direct extension to cross-trait LD score regression, which relaxes the "infinitesimal" assumption.

Model for bivariate distribution of GWAS z-scores. We derive two models for GWAS z-scores, which we call "fast model" and "full model". "fast model" is quicker to run, and we use it to perform initial search in the space of model's parameters. "full" model is slower but more accurate, and we use it for a final tuning of model estimates.

"full" model for GWAS z-scores approximates (z_{1j}, z_{2j}) distribution of a given GWAS SNP as a mixture of K=20000 bivariate normal distributions, all having equal weight in the mixture. For each $k = 1, \dots, K$ we randomly draw the location of causal variants ($\pi_1 N$ causal variants specific to the first trait, $\pi_2 N$ specific to the second trait, and $\pi_{12} N$ shared causal variants, where N denotes the total number of variants in the reference panel), and calculate the variance-covariance matrix Σ'_{kj} from (3), using estimated LD r^2 correlations between the assumed causal variants and the GWAS SNP. Then

$$(z_{1j}, z_{2j}) = (\delta_{1j}, \delta_{2j}) + N\left((0,0), \begin{bmatrix} \sigma_{01}^2 & \rho_0\sigma_{01}\sigma_{02} \\ \rho_0\sigma_{01}\sigma_{02} & \sigma_{02}^2 \end{bmatrix}\right), \tag{4}$$

$$(\delta_{1j}/\sqrt{N_{1j}}, \delta_{2j}/\sqrt{N_{2j}}) \sim \frac{1}{K}\sum_{k=1..K} N((0,0), \Sigma'_{kj}).$$

The "fast" model is derived from the method of moments (see Supplementary note):

$$\begin{aligned}(\delta_{1j}/\sqrt{N_{1j}}, \delta_{2j}/\sqrt{N_{2j}}) \sim\ & [(1-\pi'_{1j})N(0,0) + \pi'_{1j}N(0, \Sigma'_{1j})] \oplus \\ & [(1-\pi'_{2j})N(0, \Sigma'_{2j}) + \pi'_{2j}N(0, \Sigma'_{2j})] \oplus \\ & [(1-\pi'_{12,j})N(0, \Sigma'_{12,j}) + \pi'_{12,j}N(0, \Sigma'_{12,j})],\end{aligned} \tag{5}$$

where $\oplus$ denotes convolution of probabilistic distribution functions (so that right-hand size evaluates to a mixture of 8 components), $\pi'_{cj} = \ell_j \pi_c / \eta_j$ is adjusted weight of mixture component ($c \in 1,2,12$); $\Sigma'_{cj} = \eta_j \Sigma_c$ is adjusted variance-covariance matrix; $\ell_j = \sum_i H_i r_{ij}^2$ is the LD score, adjusted for heterozygosity[41]; and $\eta_{cj} = \left(\pi_c \ell_j + (1-\pi_c)\frac{\sum_i H_i^2 r_{ij}^4}{\sum_i H_i r_{ij}^2}\right)$ can be interpreted as shape parameter that affects fourth and higher moments of the distribution. This model explains second moments $E[Z_{1j}^2]$, $E[Z_{1j}Z_{2j}]$, $E[Z_{2j}^2]$ and fourth moments $E[Z_{1j}^4]$, $E[Z_{1j}^2 Z_{2j}^2]$, $E[Z_{2j}^4]$ of z score distribution, and forms a theoretical basis for the mixture model of sparse and ubiquitous effects[42,43]. Of interest is that the "fast" model involves the forth power of allelic correlation r_{ij}^4, which is directly proportional to kurtosis (measure of heavy tails) of z-score distribution.

LD structure estimation. To estimate LD structure we use 489 individuals of 1000 Genome project[44] (phase 3 data), obtained from LD score regression website[8,21,45] (see URLs). 14 individuals were excluded due to relatedness[46]. For simulations LD scores were estimated from the genotypes that we use to produce synthetic GWAS data. LD r^2 coefficients were calculated using PLINK[47] with LD r2 cutoff of 0.05 and fixed window size of 50 000 SNPs, corresponding on average to a window of 16 centimorgans. We deliberately choose a larger LD window compared to LDSR-recommended window of 1 centimorgan because the later appears to truncate a noticeable part of LD structure. At the same time, we did not observe an effect of using unbiased estimate[48] of r^2, thus fall back to the standard Pearson correlation coefficient. In simulations LD structure was re-estimated from the actual genotypes used to generate synthetic summary statistics. We employ small integer compression[49] for efficient storage of the LD matrix.

Fit procedure. We fit the model by direct optimization of weighted log likelihood

$$F(\theta) = \sum_j w_j \log(pdf(z_j|\theta)), \qquad (6)$$

where $\theta = (\pi_1, \pi_2, \pi_{12}, \sigma_1^2, \sigma_2^2, \rho_{12}, \sigma_{01}^2, \sigma_{02}^2, \rho_0)$ is a vector of all parameters being optimized, and weights w_j chosen by random pruning (64 iterations at LD r2 0.1). Optimization is done by Nelder-Mead Simplex Method[50] as implemented in MATLAB's `fminsearch`. First, we fit univariate parameters separately for each trait, i.e. $\pi_1^u, \sigma_1^2, \sigma_{01}^2$ for the first trait, and similarly for the second trait. Univariate fit employs a sequence of optimizations to ensure robust convergence: first, we use "fast" model under constraint $\pi_1^u = 1$ to find $\sigma_{1,inf}^2$ and to initialize σ_{01}^2; second, we use constraint $\pi_1^u \sigma_1^2 = \sigma_{1,inf}^2$ to find initial values of π_1^u and σ_1^2, again with "fast" model. Finally, we use full model and unconstrained optimization to jointly fit $\pi_1^u, \sigma_1^2, \sigma_{01}^2$ parameters. The same procedure is repeated for the second trait, to find $\pi_2^u, \sigma_2^2, \sigma_{02}^2$. To improve convergence, we parametrize univariate log-likelihood as a function of $\log(\pi_1^u \sigma_1^2)$ and $\log(\pi_1^u / \sigma_1^2)$, which represent almost independent dimensions of the energy landscape. In bivariate optimization use "fast" model and constraint $\pi_{12} = 1$ to initialize ρ_{12} and ρ_0. Then proceed with "full" model optimization of the parameters specific to bivariate model (π_{12}, ρ_{12}, ρ_0), constraining all other parameters to their univariate estimates. The additional analysis (Supplementary Tables 7, 8) uses right-censoring[51] of z-scores exceeding $z_t = 5.45$, by using cumulative distribution function[52] in the log likelihood:

$$F(\theta) = \sum_{j:|z_j|\le z_t} w_j \log(pdf(z_j|\theta)) + \sum_{j:|z_j|>z_t} w_j \log(cdf(z_{max}|\theta)). \quad (7)$$

Standard error estimation. We estimate standard errors of all parameters from the observed Fisher's information, based on the "fast" model. It is known from the likelihood optimization theory that the observed Fisher's information may not be suitable for a parameter near its boundary, which is applicable to the mixture weights π_1, π_2, π_{12} and the correlation of effect sizes ρ_{12}. To mitigate this problem we apply transformations — MATLAB's `logit()` for π_1, π_2, π_{12}, `exp()` for $\sigma_1^2, \sigma_2^2, \sigma_{01}^2, \sigma_{02}^2$, and `erf()` for ρ_0, ρ_{12}, and estimated variance-covariance matrix of errors in the transformed parameter space. We validated that our estimates based on the observed Fisher's information are in good agreement with block jack-knife estimates. To estimate standard errors for a function of the parameters, such as r_g or h^2, we incorporate linear correlation among parameter errors in the transformed space. We sample N=1000 realizations of the parameter vector, calculating the function (e.g., r_g or h^2) on each of them, and report the standard deviations. In cases when joint hessian was not positive definite, we estimate marginal errors of fitted parameters.

Large LD blocks. The log-likelihood cost function and the Q-Q plots apply a weighting scheme to SNPs to avoid overcounting evidence from large LD blocks. As an alternative to weighting by inverse LD score, we choose to infer the weights by random pruning. This technique is a stochastic procedure which averages log likelihood function across repeatedly selected subsets of variants, such that for each pair of variants i, j in a subset J the squared allelic correlation r_{ij}^2 falls below certain threshold. Given T iterations of random pruning the log-likelihood function can be calculated as follows:

$$F(\theta) = \frac{1}{T}\sum_{t=1}^{T}\sum_{j\in J_t} \log(pdf(z_j|\theta)) \quad (8)$$

which is equivalent to weighted log-likelihood $F(\theta) = \sum_j w_j \log(pdf(z_j|\theta))$ with weights $w_j = |\{t: j \in J_t\}|/T$, $|S|$ denotes cardinality of set S. Random pruning with stringent threshold r2=0.1 justify independent modeling of the residuals in (3) across SNPs, which otherwise would be correlated.

Heritability estimates. In additive model, SNP heritability is defined as a sum across all causal variants: $\sigma_\beta^2 \sum_{j:\beta_j\neq 0} 2p_j(1-p_j)$, which we approximate from an average heterozygosity of all variants in the reference: $\pi_1 H_{total}\sigma_\beta^2$, where $H_{total} = \sum_j 2p_j(1-p_j)$. To estimate the proportion of causal variants that explain certain fraction of heritability (Supplementary Figure 11) we randomly sample N=10000 causal effects from the reference, draw their effects β_j from normal distribution, sort according to $\beta_j^2 p_j(1-p_j)$, and report fraction of variants that cumulatively account 90% of heritability.

Genetic correlation. Parameter ρ_{12} in MiXeR defines correlation of effect sizes within shared polygenic component. Genome-wide genetic correlation, calculated across all SNPs, is related to ρ_{12} by the following formula that involves polygenicity $\pi_1^u = \pi_1 + \pi_{12}$ and $\pi_2^u = \pi_2 + \pi_{12}$ of the traits, and polygenic overlap π_{12}:

$$r_g = \rho_{12}\pi_{12}/\sqrt{\pi_1^u \pi_2^u} \quad (9)$$

For traits with K-fold difference in polygenicity ($\pi_1^u = K\pi_2^u$) the formula predicts an upper bound on genome-wide genetic correlation: $r_g \leq \rho_{12}/\sqrt{k}$, where equality holds if causal variants of the less polygenic trait form a subset of the higher-polygenic trait.

Quantile-quantile plots. Univariate Q-Q plots and stratified Q-Q plots for the model were constructed from $pdf_j(z)$ density as defined by (3), given fitted parameters of the model and LD structure of j-th SNP, calculated across a fine grid of z-scores ranging from 0 to 38 with 0.05 step. We average $pdf_j(z)$ across 1% of randomly sampled SNPs, and numerically integrate the resulting probability density function to convert it into cumulated distribution function. Error bars on data Q-Q plots represent the 95% binomial confidence interval $q \pm 1.96\sqrt{q(1-q)/n_{total}}$, where q is portability of observing a p-value as extreme as, or more extreme then the chosen p-value, and n_{total} is the effective number of SNPs after controlling for LD structure, which in our case was calculated as a sum of random pruning weights across all SNPs.

GWAS power curves. Causal mixture model can project the future of GWAS discoveries, by estimating proportion $S(N)$ of narrow-sense heritability captured by genome-wide significant SNPs at a given sample size N. The $S(N)$ is defined as follows:

$$S(N) = \frac{\sum_j \int_{z:|z|\geq z_t} C(z,N,j)dz}{\sum_j \int_z C(z,N,j)dz}, \quad (10)$$

where $z_t = 5.45$ gives z-score corresponding to the standard genome-wide significance threshold $5{\cdot}10^{-8}$, and $C(z,N,j) \equiv P(z,N,j) \cdot E(\delta^2|z,N,j)$ denotes a posterior effect size $E(\delta^2|z,N,j)$ of the non-centrality parameter δ^2 for a GWAS SNP j, given certain z-score, multiplied by a prior probability of observing that z-score. Probability density function $P(z,N,j)$ is given by (4), and $E(\delta^2|z,N,j)$ can be calculated from the Bayesian rule. Thus, $C(z,N,j) = \int \delta^2 P(z|\delta)P(\delta,j)d\delta$ where $z|\delta \sim N(\delta,\sigma_0^2)$. Analytical expression for $C(z,N,j)$ and $\int_{z:|z|\geq z_t} C(z,N,j)dz$ is given in the Supplementary Note.

SNPs in the analysis. To enable direct comparison of our model with LD score regression we use the same set of SNPs in our log likelihood optimization, which consist of approx. 1.1 million variants, subset of 1000 Genomes and HapMap3[53], with MAF above 0.05, ambiguous SNPs excluded, imputation INFO above 0.9, MHC and other long-range LD regions excluded. Calculation of the LD structure, LD scores ℓ_j and shape parameter η_j are based on the approx. 10 million SNPs from 1000 Genomes Phase 3 data, downloaded from LD score regression website (see URLs). In simulations we generate GWAS and estimate LD structure on a subset of 11,015,833 SNPs from 1000 Genomes Phase 3, with MAF above 0.002, call rate above 90%, excluding duplicated RS numbers; the fit procedure was constrained to approximately 130K GWAS SNPs, keeping only HapMap3 SNPs, and pruning SNPs at LD r2 threshold of 0.1.

LD score regression estimates. For dichotomous phenotypes we used an effective sample size of $N_{eff} = 4/(1/N_{case} + 1/N_{cont})$ to account for imbalanced number of cases

and controls, both in MiXeR and in LD score regression. Additionally, we run LDSR using MiXeR MAF model (using `–per-allele` flags in LD score estimation), and show the results alongside with original LDSR estimates (Supplementary Tables 7, 8). For case/control phenotypes heritability is reported on the observed scale.

Simulations. In our simulations we use a panel of N = 100,000 samples, generated by HapGen2[54] using 1000 Genomes[44] data to approximate the LD structure for European ancestry. For each simulation run we use PLINK to obtain GWAS summary statistics of two synthesized quantitative phenotypes, with complete sample overlap between GWAS samples. Quantitative phenotype y_k of k-th sample is calculated via simple additive genetic model, $y_k = \sum_j g_{kj}\beta_j + \epsilon_k$, where g_{kj} is the number of reference alleles for j-th SNP on k-th sample, β_j is causal effect size, and ϵ is the residual vector drawn from normal distribution with zero mean and variance chosen in a way that sets heritability $h^2 = var(\boldsymbol{G}\beta)/var(y)$ to a predefined level.

For the simulations shown in Figure *2*, we draw effect sizes (β_{1j}, β_{2j}) from the four-component mixture model (1), varying polygenicity of each phenotype ($\pi_1^u = \pi_1 + \pi_{12}$ and $\pi_2^u = \pi_2 + \pi_{12}$), and polygenic overlap (π_{12}). We choose total polygenicity in both traits to be 3×10^{-3} or 3×10^{-4} and include an additional scenario of uneven polygenicity ($\pi_1^u = 3\times10^{-3}$, $\pi_2^u = 3\times10^{-4}$). For each combination π_1^u, π_2^u and h_2 we set polygenic overlap to be a fraction of total polygenicity $\pi_{12} = f\,\pi_1^u$, choosing the fraction f from six equally spaced values (0.0 to 1.0 with a step of 0.2). Correlation of effect sizes ρ_{12} set to 0.0 or 0.5. Heritability was set to 0.1, 0.4 or 0.7, which let us keep GWAS sample size constant (N=100,000) because the distribution of GWAS z-scores depends on N and h_2 only through their product, $h_2 \times N$ (thus, simulations with N=700,000 and $h_2 = 0.1$ would be equivalent to our scenario with N=100,000 and h_2=0.7). Finally, for each combination of heritability, polygenicity, polygenic overlap and correlation of effect sizes, we repeat simulations 10 times.

For the simulations with differential enrichment, we simulate three levels of polygenicity (3K, 30K and 300K causal variants), three levels of heritability (0.1, 0.4 and 0.7), and for each combination generate 20 pairs of genetically independent traits (except for having shared pattern of enrichment). To simulate the enrichment, we keep a constant variance of effect sizes across all SNPs but modulate the probability of having causal variant proportionally to LDSR regression coefficient. We cover two scenarios: first, LDSR with MiXeR MAF model (using `–per-allele` flags in LD score estimation), and secondly with the original LDSR MAF model.

ONLINE METHODS REFERENCES

40. Holland, D. *et al.* Estimating inflation in GWAS summary statistics due to variance distortion from cryptic relatedness. *bioRxiv* (2017).
41. Thompson, W.K. *et al.* An Empirical Bayes Mixture Model for Effect Size Distributions in Genome-Wide Association Studies. *PLoS Genet* **11**, e1005717 (2015).
42. Holland, D. *et al.* Estimating Effect Sizes and Expected Replication Probabilities from GWAS Summary Statistics. *Frontiers in Genetics* **7**, 15 (2016).

43. Wang, Y. *et al.* Leveraging Genomic Annotations and Pleiotropic Enrichment for Improved Replication Rates in Schizophrenia GWAS. *PLOS Genetics* **12**, e1005803 (2016).
44. Genomes Project, C. *et al.* A global reference for human genetic variation. *Nature* **526**, 68-74 (2015).
45. Bulik-Sullivan, B.K. *et al.* LD Score regression distinguishes confounding from polygenicity in genome-wide association studies. *Nat Genet* **47**, 291-5 (2015).
46. Gazal, S., Sahbatou, M., Babron, M.C., Genin, E. & Leutenegger, A.L. High level of inbreeding in final phase of 1000 Genomes Project. *Sci Rep* **5**, 17453 (2015).
47. Purcell, S. *et al.* PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet* **81**, 559-75 (2007).
48. Wherry, R.J. A New Formula for Predicting the Shrinkage of the Coefficient of Multiple Correlation. *Ann. Math. Statist.* **2**, 440-457 (1931).
49. Lemire, D., Boytsov, L. & Kurz, N. SIMD compression and the intersection of sorted integers. *Softw. Pract. Exper.* **46**, 723-749 (2016).
50. Lagarias, J.C., Reeds, J.A., Wright, M.H. & Wright, P.E. Convergence properties of the Nelder--Mead simplex method in low dimensions. *SIAM Journal on optimization* **9**, 112-147 (1998).
51. Duckworth, W.M. & Stephenson, W.R. Beyond Traditional Statistical Methods. *The American Statistician* **56**, 230-233 (2002).
52. Genz, A. Numerical computation of rectangular bivariate and trivariate normal and t probabilities. *Statistics and Computing* **14**, 251-260 (2004).
53. International HapMap, C. *et al.* Integrating common and rare genetic variation in diverse human populations. *Nature* **467**, 52-8 (2010).
54. Su, Z., Marchini, J. & Donnelly, P. HAPGEN2: simulation of multiple disease SNPs. *Bioinformatics* **27**, 2304-5 (2011).

SUPPLEMENTARY NOTE

Separate PDF document.

SUPPLEMENTARY FIGURES

1. Simulations: polygenic overlap estimates in bivariate analysis

2. Simulations: estimates of the correlation of effect sizes in bivariate analysis

3. Simulations: estimates of genetic correlation in bivariate analysis

4. Simulations: polygenicity and heritability estimates in univariate analysis

5. Simulations: Q-Q plots with simulated data

6. Simulations: Q-Q plots of SNPs partitioned into a grid of MAF and LD score

7. Simulations: stratified QQ plots for traits with and without polygenic overlap

8. Sensitivity analysis: differential enrichment profile of synthesized data

9. Sensitivity analysis: polygenicity and heritability estimates under differential genomic enrichment

10. Sensitivity analysis: polygenicity and heritability estimates under differential genomic enrichment and mis-specified MAF architecture

11. Cumulated fraction of explained heritability versus proportion of causal variants

12. Venn diagrams across all traits.

13a-f. Venn diagrams and conditional cross-trait QQ plots

14a-f. Observed and predicted bivariate density of GWAS association statistics

15. Projected power plot for future GWAS sample sizes

16. Univariate Q-Q plots

17a-f. Q-Q plots of SNPs partitioned by MAF and LD score

18. Univariate likelihood as a function of polygenicity parameter

19. Bivariate likelihood as a function of polygenic overlap parameter

SUPPLEMENTARY TABLES

1. Simulations: parameter estimates in bivariate analysis

2. Simulations: polygenicity and heritability estimates in univariate analysis

3. Sensitivity analysis: polygenic overlap estimates under differential genomic enrichment

4. Sensitivity analysis: polygenic overlap estimates under genomic enrichment and mis-specified MAF architecture

5. Sensitivity analysis: simulations with incomplete reference

6. Summary statistics Metadata

7. Results of bivariate analysis with MiXeR, with and without right-censoring

8. Results of univariate analysis with MiXeR, with and without right-censoring

EQUATIONS

1 MiXeR (beta1, beta) prior distribution

2 MiXeR Sigma1, Sigma2, Sigma12

3 MiXeR (z1, z2) exact formula

4 MiXeR (z1, z2) full model (sampling)

5 MiXeR (z1, z2) fast model (two gaussian)

6 MiXeR log likelihood function

7 Right censoring

8 MiXeR random pruning

9 MiXeR genetic correlation

10 MiXeR power curves S(N)