

Influence of genetic interactions on polygenic prediction

Zhijun Dai^{1,#}, Nanye Long², Wen Huang^{1,@}

¹Department of Animal Science, ²Institute for Cyber-Enabled Research

Michigan State University, East Lansing, MI 48824

[#]Present address: Hunan Engineering and Technology Research Center for Agricultural Big Data Analysis and Decision-Making, Hunan Agricultural University, Changsha, China 410128

[@]Correspondence: huangw53@msu.edu

9 **Abstract**

10 Prediction of phenotypes from genotypes is an important objective to fulfill the promises of
11 genomics, precision medicine and agriculture. Although it's now possible to account for the
12 majority of genetic variation through model fitting, prediction of phenotypes remains a
13 challenge, especially across populations that have diverged in the past. In this study, we
14 designed simulation experiments to specifically investigate the role of genetic interactions in
15 failure of polygenic prediction. We found that non-additive genetic interactions can
16 significantly reduce the accuracy of polygenic prediction. Our study demonstrated the
17 importance of considering genetic interactions in genetic prediction.

18 Introduction

19 The problem of “missing heritability” has attracted much attention and controversy in
20 quantitative genetics (Manolio et al., 2009), yet its definition remains ambiguous in the
21 literature. A widely used definition is that genetic associations identified in large-scale
22 genome-wide association studies (GWAS) cannot fully account for heritability estimates (e.g.
23 from twin studies) in the sense that the model fitting can only capture a fraction of the total
24 variance. As sample sizes for GWAS increase from thousands to hundreds of thousands, and
25 advanced statistical methods are developed to fit all DNA variants in the model
26 simultaneously, including those not significantly associated with the trait, the variance that
27 can be explained by DNA variants also increase. For example, adult human height is a
28 classical quantitative trait with a narrow sense heritability (h^2) of approximately 0.8 based on
29 twin studies (Silventoinen et al., 2003). However, early GWAS studies identified common
30 variants explaining only a total of 2-4% phenotypic variance (Gudbjartsson et al., 2008;
31 Lettre et al., 2008; Weedon et al., 2008) with sample sizes in the order of 20,000. In 2010, a
32 landmark study increased this proportion to about 45% by fitting ~300,000 SNP markers in
33 the model for ~4,000 individuals with the covariance among individuals determined by
34 genome-wide SNP similarity (Yang et al., 2010). Importantly, applying the same idea, the
35 most recent study using whole genome sequences of ~20,000 individuals in the TOPMed
36 almost entirely closed the gap between the genomic heritability and the presumed
37 heritability (Wainschtein et al., 2019). The progress has been remarkable and it can be
38 cautiously expected that the combination of large sample size and full genome sequences
39 may finally capture all heritability. Perhaps more importantly, it also suggests that the failure
40 to explain all heritability in early GWAS was due largely to low statistical power and
41 incomplete variant coverage thus those with smaller effects, lower minor allele frequencies,
42 and non-SNP variants were missed from the model fitting.

43 A second, more implicit but more practical definition of missing heritability, is that the
44 prediction accuracy of quantitative phenotypes based on genotypes (polygenic scores) is far
45 less than the heritability of the trait. A perfect genetic model with precise effects and model
46 specification should be able to predict unobserved phenotypes with an accuracy (measured
47 by r^2) equal to the heritability. But that’s not always the case. For example, a large GWAS on
48 adult human height with almost 200,000 individuals identified over 180 loci, capturing only
49 ~10% of the phenotypic variation (Lango Allen et al., 2010). This proportion of variance was
50 measured based on “leave-one-out” out-of-sample prediction (International Schizophrenia
51 Consortium et al., 2009), *i.e.*, the effects of the genetic loci were estimated in one subset of
52 the sample and polygenic scores (genetic effects summed over all significant loci) was
53 computed to predict phenotypes in another subset. The partition between the subsets
54 conveniently followed sample origin from different European countries (Lango Allen et al.,
55 2010). In contrast to the mixed model genomic heritability approach, this method of
56 estimating explained heritability was more akin to genomic prediction widely used in animal
57 and plant breeding (Meuwissen et al., 2001; VanRaden, 2008), in which effects of genetic

58 markers across the whole genome, regardless of their statistical significance, are summed to
59 compute genetic prediction.

60 Both the genomic heritability and the prediction accuracy of polygenic scores (or polygenic
61 breeding values) take the form of variance proportion, but have vastly different properties.
62 One of the most contrasting differences is that prediction accuracy can be small even when
63 the genomic heritability is large (Makowsky et al., 2011). Our discussion on the two
64 definitions of missing heritability above is a clear example of this distinction. The
65 implications of this distinction are profound. Most notably, even if there was no missing
66 heritability based on genomic heritability, the utility of polygenic score would be very limited
67 if prediction accuracy is low.

68 Recently, there has been renewed interest in the application of polygenic score
69 (International Schizophrenia Consortium et al., 2009) with the advent of large public data
70 sets such as the UK biobank (e.g. Khera et al., 2018). In particular, many studies have
71 observed poor prediction by polygenic scores across different ancestry groups (Martin et al.,
72 2019) or even within an ancestry group but with variable characteristics (Mostafavi et al.,
73 2019). In fact, earlier studies with smaller sample sizes observed similar patterns, but were
74 interpreted as missing heritability (Lango Allen et al., 2010; Makowsky et al., 2011). In animal
75 breeding, similar observations have also been made. Although genomic prediction works
76 exceedingly well within a breed, cross-breed prediction generally fails (Hayes et al., 2009).
77 The explanation is obvious, genetic effects are context dependent and heterogeneous
78 between groups. Variable linkage disequilibrium (LD) patterns, environments, and other
79 factors can all contribute to the variable genetic effects, manifesting as variable accuracy of
80 polygenic prediction.

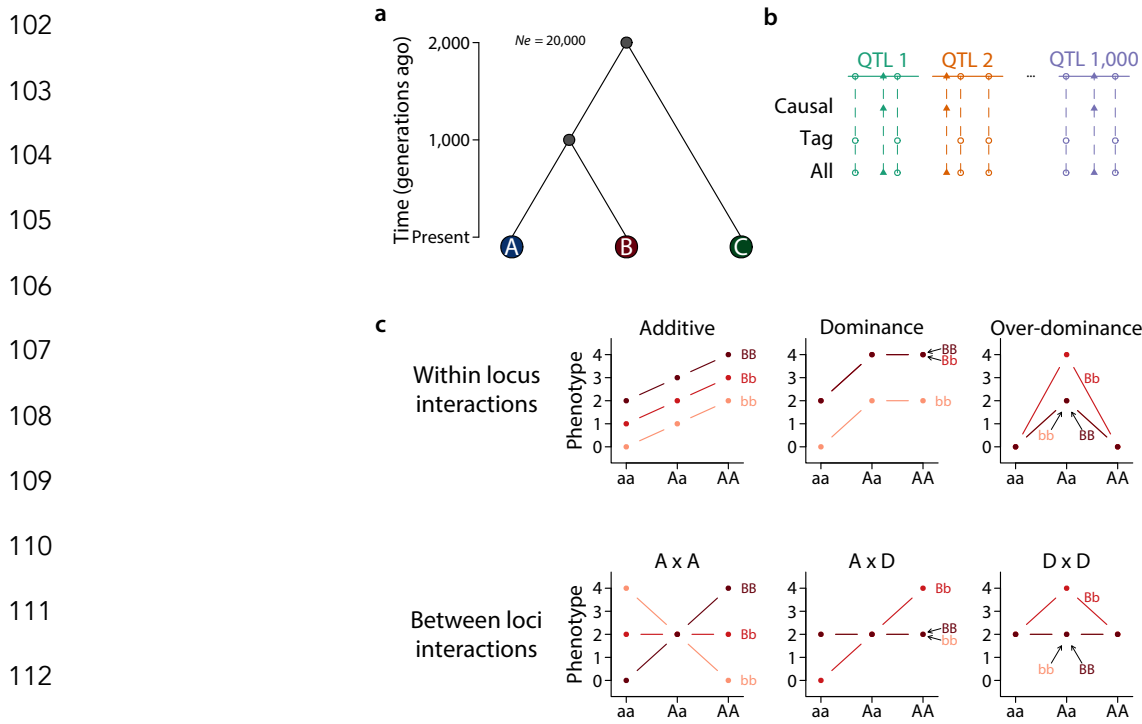
81 Genetic interactions are pervasive, and an important type of context dependent effects
82 (Mackay, 2014; Mackay and Moore, 2014). It has been previously shown that the presence of
83 genetic interactions does not have a strong effect on genomic heritability (Hill et al., 2008;
84 Huang and Mackay, 2016), therefore the magnitude of genomic heritability offers no
85 indication of the genetic architecture. However, genetic interactions may influence genomic
86 prediction accuracy and models taking into account the complexity improves prediction
87 (Morgante et al., 2018). This clearly suggests that the simplification of genetic architecture to
88 the additive infinitesimal model when the true model is not, although convenient and no
89 comparable alternatives exist, can be risky. In this study, we specifically investigate the
90 influence of genetic interactions on prediction of polygenic scores, with an emphasis on
91 polygenic prediction across diverged populations.

92 **Results**

93 **Experimental design**

94 Because it's not yet possible to unambiguously know the true genetic architecture of a
95 quantitative trait, all experiments in this study were performed using simulated data instead

96 of real data. This allows us to specifically ask simple questions while eliminating influence
 97 from other factors. We simulated a sample of 75,000 diploid individuals from three ancestry
 98 groups, where population A and B diverged 1,000 generations ago and their ancestors
 99 diverged from population C an additional 1,000 generations ago (Figure 1a). This
 100 specification is qualitatively similar to the global human population history where the
 101 ancestral population that went out of Africa were further split into multiple populations.



114 **Figure 1. Simulation of genome sequences, population structure, and genetic**
 115 **architecture.** (a) Three populations (A, B, C) were simulated with an effective population
 116 size of 20,000 each. A and B diverged 1,000 generations before present and A and C
 117 diverged 2,000 generations ago. (b) 1,000 independently inherited chromosomes were
 simulated, each containing one QTL. Three sets of variants were considered, including
 "causal", "tag", and "all" as illustrated. (c) Six different genetic architecture were
 simulated, each illustrated by one of the panels.

118 We considered three possible variant sets (Figure 1b), 1) causal: all and only causal variants;
 119 2) tag: all variants except causal variants; and 3) all: all variants including causal variants.
 120 These represent three simplified scenarios 1) a best case scenario where causal variants have
 121 been identified, 2) a realistic scenario where causal variants are tagged by genotyped
 122 variants, and 3) an achievable scenario in the near future with whole genome sequences. We
 123 did not consider variants that were rare ($MAF < 0.01$) in all three populations as they led to
 124 gross overestimation of genomic heritability approaching one, similar to findings in a
 125 simulation study using real genotypes (Evans et al., 2018). The three variant sets were used
 126 to compute genomic heritability and perform polygenic prediction. When performing
 127 polygenic prediction, we did not select variants based on association tests. This choice was

128 based on the consideration that selection of markers introduced another variable in the
129 experiment to complicate the design and interpretation. Instead, we draw from the
130 distinction between causal and all variants to represent the extreme scenarios where a
131 perfect selection or no selection was performed.

132 We simulated a quantitative trait controlled by 1,000 independently inherited QTLs (Figure
133 1b) of broad sense heritability $H^2 = 0.8$ but different types of genetic architecture. When the
134 genetic architecture is strictly additive, the narrow sense heritability $h^2 = H^2 = 0.8$, whereas in
135 other cases $h^2 < 0.8$. Six simple models of genetic architecture were simulated, including
136 additive, dominance, over-dominance, and pairwise additive by additive (A x A), additive by
137 dominance (A x D), and dominance by dominance (D x D) (Figure 1c). No higher order
138 interaction was simulated and effects across loci or across pairs were additive.

139 Genomic heritability misses little heritability

140

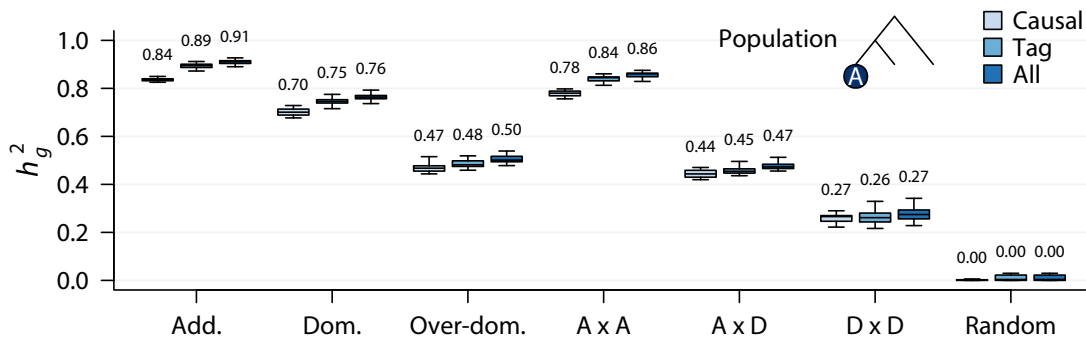
141

142

143

144

145



146

147

148

149

150

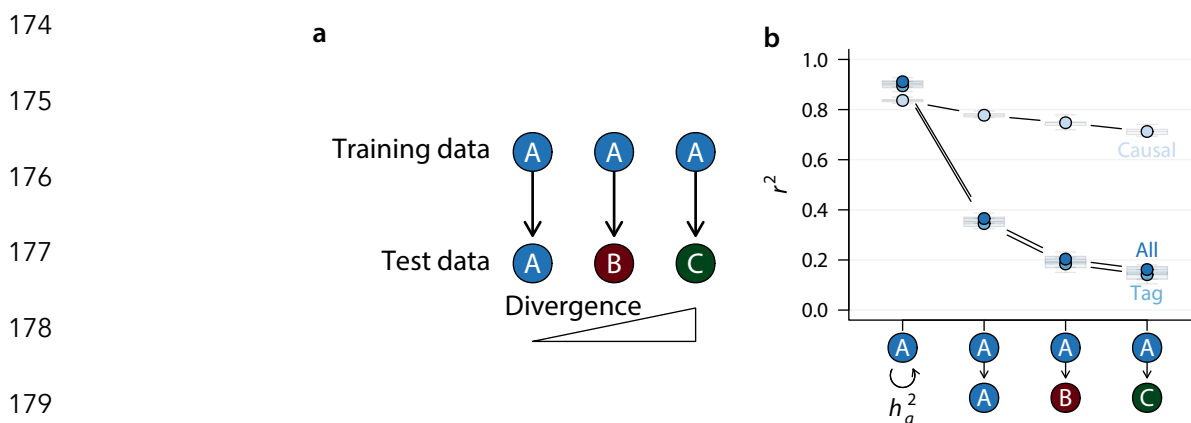
Figure 2. Genomic heritability in the simulated populations. Box plot (median indicated on top) showing the genomic heritability (h_g^2) estimated using GREML under different genetic architecture, where Add. = additive, Dom. = dominance, Over-dom. = over-dominance, A x A = additive by additive, A x D = additive by dominance, D x D = dominance by dominance, and random is a non-genetic model where the phenotypic variation was entirely due to random environmental variation. The population in which the genomic heritability was estimated was indicated in the top right corner. Genomic heritabilities in all populations were given in Figure S1.

151 We first recapitulated a result that has been consistently shown (Hill et al., 2008; Huang and
152 Mackay, 2016). We fitted a linear mixed model in each of the three populations or combined
153 samples using GREML implemented in the GCTA (Yang et al., 2011) using 20,000
154 individuals. We found that h_g^2 were uniformly high when the genetic architecture was
155 additive, dominance, or additive by additive, accounting for nearly all heritability (Figure 2,
156 Figure S1). Whether or not the variant sets included causal variants appeared to have little
157 effects on h_g^2 ; variant sets excluding causal variants performed as well as causal variants only
158 and there was a slight tendency of upward bias (Figure 2). Similar results were obtained
159 regardless of whether the samples were from a homogeneous population or a mixture of
160 samples from two diverged populations (Figure S1). When the genetic architecture was
161 entirely overdominance, additive by dominance, or dominance by dominance, h_g^2 was lower,

162 but still consistently explained > 50% of the heritability (Figure 2, Figure S1). Taken together,
163 these results suggest that as long as a large number of genome-wide markers were fitted,
164 little heritability was missed, regardless of the genetic architecture. In other words, the
165 magnitude of genomic heritability offers no discrimination of the underlying genetic
166 architecture (Huang and Mackay, 2016).

167 Accuracy of polygenic prediction with an additive genetic architecture

168 We then asked a simple question. If genome-wide variants are able to capture the majority
169 of heritability, are they able to predict phenotypes accurately? This question directly
170 addresses the distinction between the two definitions of missing heritability as we outlined in
171 the introduction. If there is no missing heritability based on mixed model fitting, is there
172 missing heritability in polygenic prediction? Many illuminating results could be obtained by
173 comparing different scenarios of simulations (Figure S2).



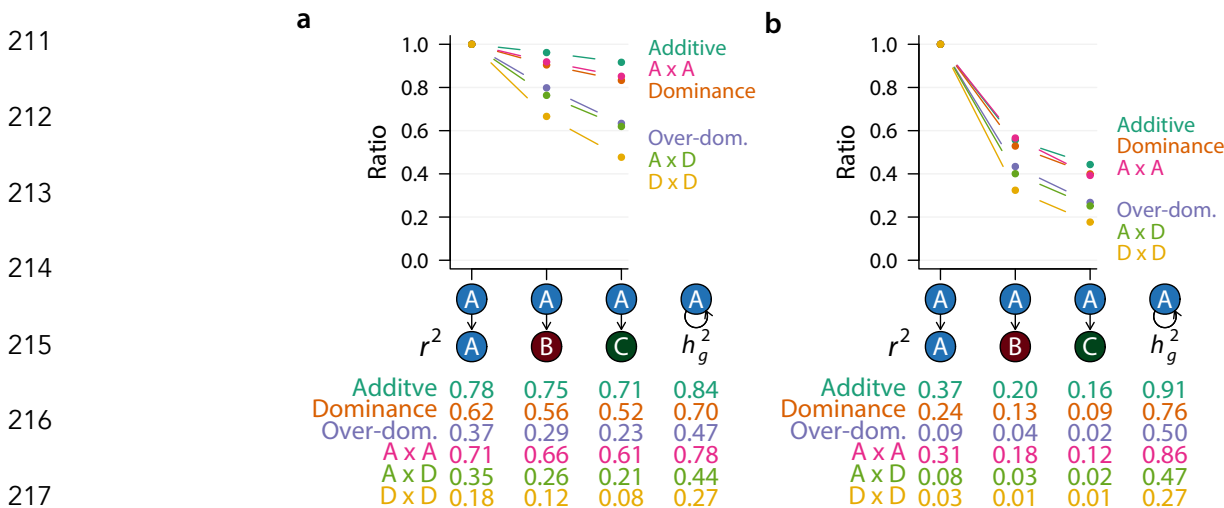
174
175
176
177
178
179
180 **Figure 3. Polygenic prediction under additive genetic architecture.** (a) Polygenic
181 prediction was performed according to the diagram, where the model was trained in
182 population A and tested in populations A, B, and C at increasing divergence. (b)
183 Prediction accuracy was plotted according to the training – test population relationships.
184 For comparison, genomic heritability was also plotted along side. Only the additive
185 genetic architecture was considered in this plot.

183 We first considered the simplest and best scenario, in which the genetic architecture was
184 fully additive, and all causal and only causal variants were known. In this case, the statistical
185 model took the form of the true model and only model parameters needed to be estimated.
186 We trained the model in one population ($n = 20,000$, training population) and computed
187 polygenic scores of new individuals ($n = 5,000$, test population) either in the same
188 population or a different population. To test the performance of cross-population prediction,
189 we considered three possible relationships between the training and test populations,
190 representing a gradient of divergence between training and test data (Figure 3a).

191 As expected, the accuracy of polygenic prediction was very high in this best case scenario,
192 approaching the true heritability ('causal' in Figure 3b). There was a small decline in accuracy
193 when cross-population prediction was performed and the degree of population divergence

194 negatively affected prediction accuracy. However, when non-causal variants were included
 195 to make predictions, accuracy plummeted from ~0.8 to ~0.4 (Figure 3b) even when training
 196 and test samples were from the same population. This was likely due to the inclusion of
 197 independent predictors whose number vastly exceeded that of the causal variants. As
 198 populations become more divergent, prediction accuracy further dropped, the rate of which
 199 was much more pronounced when tag or all variants were used. These results (in the cases
 200 of tag or all variant sets) largely agreed with the large body of empirical work that accuracy
 201 of polygenic prediction was substantially lower than genomic heritability and cross-
 202 population prediction was poor (Lango Allen et al., 2010; Makowsky et al., 2011; Martin et
 203 al., 2019).

204 One important lesson could be learned in this simple experiment. The facts that simply
 205 adding non-causal variants to the model drastically reduced prediction accuracy, and that
 206 the rate of decay in the accuracy of cross-population prediction was much greater in the
 207 presence of non-causal variants indicated that the agreement between model and true
 208 genetic architecture mattered. This is in sharp contrast to genomic heritability estimation,
 209 where including more variants generally improves model fit (compare (Yang et al., 2010) with
 210 (Wainschtein et al., 2019)).



218 **Figure 4. Polygenic prediction with different genetic architecture.** (a) Polygenic
 219 prediction was performed using causal variants only for six different genetic
 220 architecture. The median prediction accuracy (r^2) across 20 replicates in each
 221 scenario was listed below the graph, as well as genomic heritability (h_g^2). Each point
 222 on the graph represents a normalized median r^2 , dividing each prediction accuracy
 223 by its counterpart in the within population (A -> A) prediction. (b) Polygenic
 224 prediction with all variants. Data are presented the same way as in (a). Data in these
 225 graphs were summarized from Figure S2.

222 Accuracy of polygenic prediction in the presence of genetic interactions

223 We then tested the influence of genetic interactions on the accuracy of polygenic prediction,
 224 which fits an additive model. In a favorable condition when all causal variants were known

225 (but not their effects or interactions) and prediction was performed within the same
226 homogenous population, polygenic prediction accuracy was highly dependent on the
227 genetic architecture (A → A in [Figure 4a](#)). In general, prediction accuracy was higher for
228 genetic architecture with higher h_g^2 , such as additive, dominance, and additive by additive.
229 In contrast, under overdominance, additive by dominance, and dominance by dominance
230 genetic architecture, polygenic prediction performed substantially worse (A → A in [Figure](#)
231 [4a](#)). When all variants were used, including non-causal ones, the prediction accuracies
232 decreased dramatically and their dependency on genetic architecture appeared to be
233 stronger (A → A in [Figure 4b](#)).

234 We then asked how genetic interactions influence the rate of decay in prediction accuracies
235 when the training and test populations diverge. We set the accuracy of within-population
236 prediction as the baseline and compared cross-population prediction accuracies to this
237 baseline. When all variants were used for polygenic prediction, the accuracy of cross-
238 population prediction dropped to about 40-60% of the accuracy of within-population
239 prediction, depending on genetic architecture ([Figure 4b](#)). Additive, additive by additive,
240 and dominance genetic architecture, those with the highest h_g^2 and r^2 retained the most
241 prediction accuracy while over-dominance, additive by dominance, and dominance by
242 dominance lost the most ([Figure 4b](#)). The more diverged the populations were, the more
243 predictive ability of polygenic scores was lost ([Figure 4b](#)).

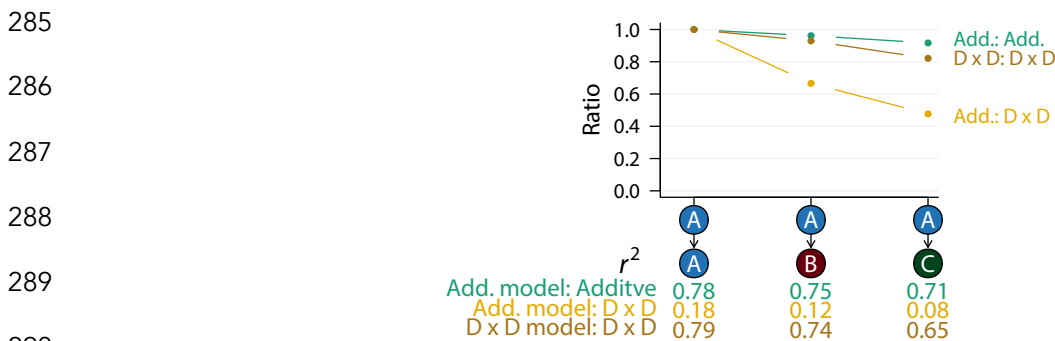
244 There are many reasons why polygenic prediction failed when test population diverged from
245 training population. In our simple simulation setting, genetic effects were the same across
246 populations and were not sensitive to any non-genetic factors. The difference in the linkage
247 disequilibrium structure between populations may in part explain the drop. Importantly,
248 simulations allowed us to directly use causal variants for prediction, thus eliminating the
249 influence of LD ([Figure 4a](#)). Remarkably, while the accuracy of cross-population prediction
250 was lower for all genetic architecture, the rate of decay was much greater when the genetic
251 architecture was over-dominance, additive by dominance, or dominance by dominance
252 ([Figure 4a](#), compare slopes of the different lines). These results clearly suggest that genetic
253 interactions can not only cause cross-population polygenic prediction to fail, but also in a
254 more severe manner compared to an additive genetic architecture.

255 Discussion

256 We demonstrate in this study through simulations that genetic interactions can influence the
257 accuracy of polygenic prediction. In particular, cross-population polygenic prediction
258 performed worse than intra-population prediction in all cases. For traits controlled by
259 genetic interactions, the cross-population decay in prediction accuracy was far greater. The
260 results make intuitive sense. For a statistical model to predict new data accurately, two
261 conditions must be met. First, the model specification must be correct or at least sufficiently
262 accurate to capture variation in the data. Second, parameters in the model must be precise.

263 When genetic interactions are present, the additive polygenic model clearly is not accurate.
264 Previous studies have mostly focused on improving parameter estimation, through
265 increasing sample size and methodological improvement. For example, increasing sample
266 size substantially increased accuracy of polygenic prediction of height within individuals of
267 European ancestry (Lello et al., 2018). Inclusion of samples of different backgrounds in the
268 training data also helped (Martin et al., 2019, Figure S2).

269 However, the complexity of the genetic architecture of a quantitative trait makes it nearly
270 impossible to specify a model prior to modeling. As a consequence, the polygenic
271 infinitesimal model or variants of it (Gianola et al., 2009) has been used as the default model.
272 The infinitesimal model has been instrumental and allowed for many theoretical insights as
273 well as applications to be developed. In particular, prediction of breeding values in animal
274 and plant breeding relying on the infinitesimal model has been very successful (García-Ruiz
275 et al., 2016). However, its limitations are also apparent. Cross-population and cross-breed
276 polygenic prediction was low in accuracy (Hayes et al., 2009; Lango Allen et al., 2010; Martin
277 et al., 2019). Although many factors may contribute to this limitation, our simulation results
278 clearly indicated that genetic interactions unaccounted for was a major contributor. Indeed,
279 if the correct genetic model could be specified, cross-population prediction can achieve very
280 high accuracy (Figure 5). There have been attempts to explicitly model non-additive genetic
281 effects in the context of polygenic prediction; some moderate improvement was observed
282 (Varona et al., 2018). However, these studies modeled non-additive effects using genome-
283 wide markers, which added a large number of independent predictors as noise to the model
284 and may negatively impact the performance.



285
286
287
288
289
290
291
292
293
Figure 5. Match between model and true genetic architecture improves polygenic prediction. Two genetic architecture were considered, the additive and D x D. The prediction was performed with either an additive model (Add. model) as implemented in GREML or a D x D model in which the correct genetic model was presumed to be known and fitted. Only causal variants were used in these analyses.

294 We did not analyze existing large data sets, some of which contained subjects from multiple
295 ancestries. Previous work with real data has consistently shown that cross-population
296 polygenic prediction generally fails (Martin et al., 2019). However, it is difficult to disentangle
297 the different factors that may contribute to effect heterogeneity and the failure of prediction

298 in real data sets. Using simulations, we can focus on specific questions and our results clearly
299 indicated a contribution of genetic interactions to the failure of cross-population polygenic
300 prediction. While the additive infinitesimal model is the most sensible model when no other
301 information is available, our study suggests that the development in the field should be
302 expanded to include efforts to more explicitly model genetic interactions. Although it is
303 challenging, recent advances in modeling (Boyle et al., 2017; Liu et al., 2019) and genomic
304 assays informing regulatory networks (Gerstein et al., 2012) may finally offer new ways to
305 develop biologically sensible models.

306 **Methods**

307 *Population simulation*

308 We used the coalescent simulator MaCS (Chen et al., 2008) to simulate genome sequences
309 of 75,000 individuals, with 25,000 in each of the populations, according to the demographic
310 history in [Figure 1a](#). We simulated 1,000 independently inherited chromosomes of 100,000
311 base pairs in size and set mutation rate as 1.25×10^{-8} per bp and recombination as 1.25×10^{-8}
312 per bp. Effective population size was set to 20,000. The MaCS command for one
313 chromosome was "macs 150000 100000 -s "\$random_seed" -i 1 -h 1000 -t
314 0.001 -r 0.001 -I 3 50000 50000 50000 0 -ej 0.0125 3 2 -ej 0.025 2
315 1". This simulation was performed once but the partition between samples were repeated
316 20 time, which were summarized as box plots in figures.

317 *Simulation of quantitative phenotypes*

318 We simulated quantitative phenotypes according to the genetic architecture depicted in
319 [Figure 1c](#). For each of the three possible genotypes for a biallelic locus with alleles A and a,
320 we used the additive coding aa = -1, Aa = 0, and AA = 1 and the dominance coding aa = 0,
321 Aa = 1, AA = 0 to code genotypes. The simulation of phenotypes consisted of two steps. In
322 the first step, the corresponding genotype coding for an individual or multiplication of
323 genotype codings (in the case of between-loci interactions) were multiplied by a genetic
324 effect randomly drawn from the standard normal distribution and summed over all loci or all
325 pairs of loci to obtain the genetic values. In the second step, an environmental effect was
326 added by drawing from a normal distribution with a computed variance such that the broad
327 sense heritability $H^2 = 0.8$. We performed this simulation in each of the 20 random partitions
328 of populations and independently sampled causal variants and genetic effects.

329 *Fitting GREML*

330 We fitted the GREML model using GCTA (Yang et al., 2011) with 20,000 individuals from
331 each of the A, B, and C populations and A + B and A + C. The GREML partitioned
332 phenotypic variance into a genomic (σ_g^2) and an environmental component (σ_e^2). Genomic
333 heritability was computed as $h^2_g = \sigma_g^2 / (\sigma_g^2 + \sigma_e^2)$.

334 *Polygenic score prediction*

335 The BLUP estimates of SNP effects were obtained using GCTA and provided to PLINK2
336 (<https://www.cog-genomics.org/plink/2.0/credits>) to compute a polygenic score in 5,000
337 new individuals either from the same population as the fitted model or from a different
338 population. Prediction accuracy of polygenic score was computed as the r^2 of correlating
339 predicted polygenic scores and the simulated true phenotypes. In the case of prediction
340 using causal variants with the correct dominance by dominance model (Figure 5), we
341 constructed pseudo-variants using the relevant genotype coding (for D x D, double
342 heterozygotes were coded as one genotype class and all others another) and ran GREML
343 and polygenic score prediction the same way as an additive model.

344 Reference

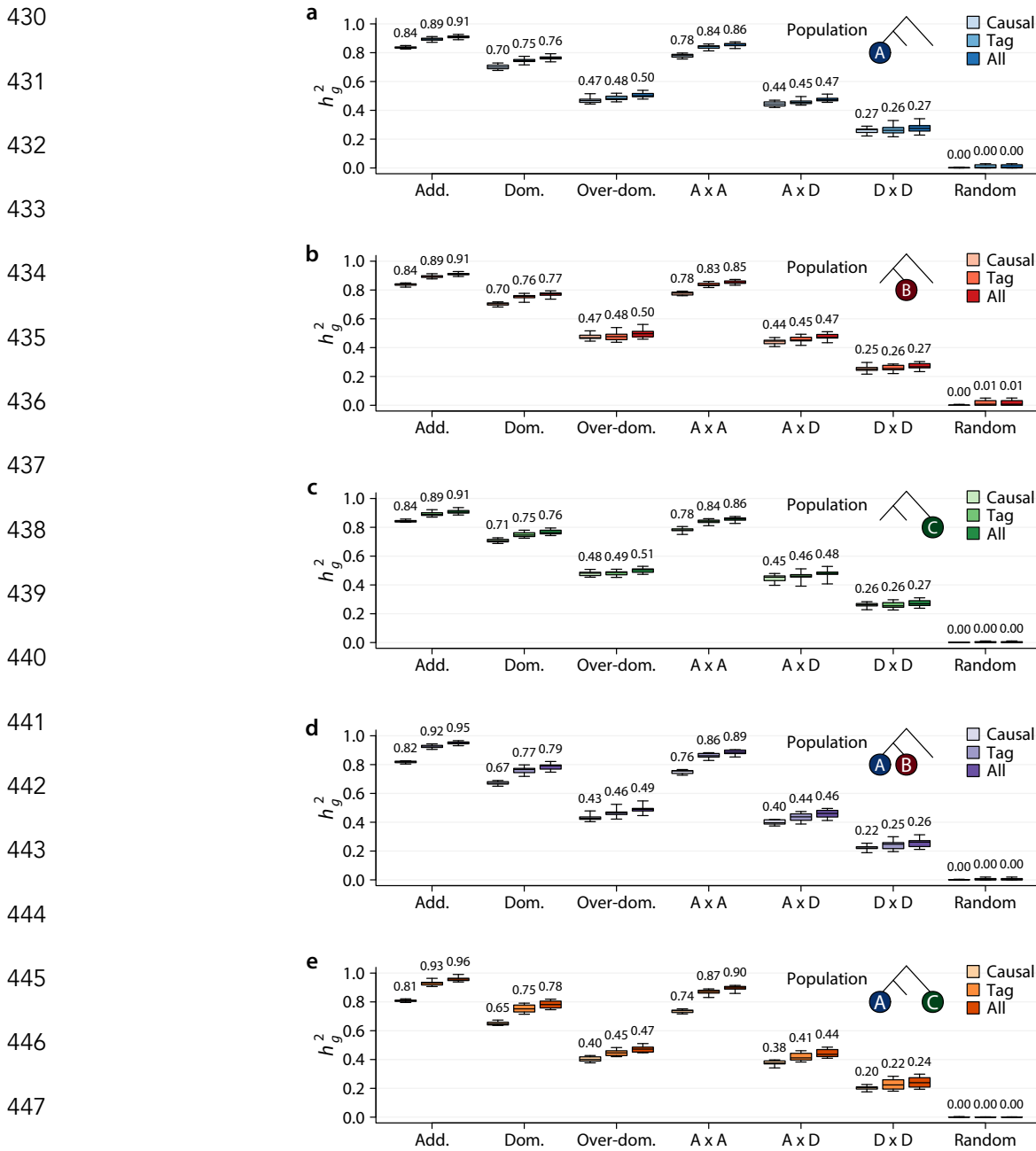
- 345 Boyle, E.A., Li, Y.I., and Pritchard, J.K. (2017). An Expanded View of Complex Traits: From
346 Polygenic to Omnigenic. *Cell* 169, 1177–1186.
- 347 Chen, G.K., Marjoram, P., and Wall, J.D. (2008). Fast and flexible simulation of DNA
348 sequence data. *Genome Res.* 19, 136–142.
- 349 Evans, L.M., Tahmasbi, R., Vrieze, S.I., Abecasis, G.R., Das, S., Gazal, S., Bjelland, D.W., de
350 Candia, T.R., Goddard, M.E., Neale, B.M., et al. (2018). Comparison of methods that use
351 whole genome data to estimate the heritability and genetic architecture of complex traits.
352 *Nat. Genet.* 50, 737–745.
- 353 García-Ruiz, A., Cole, J.B., VanRaden, P.M., Wiggans, G.R., Ruiz-López, F.J., and Van Tassell,
354 C.P. (2016). Changes in genetic selection differentials and generation intervals in US Holstein
355 dairy cattle as a result of genomic selection. *Proc. Natl. Acad. Sci.* 113, E3995–E4004.
- 356 Gerstein, M.B., Kundaje, A., Hariharan, M., Landt, S.G., Yan, K.-K., Cheng, C., Mu, X.J.,
357 Khurana, E., Rozowsky, J., Alexander, R., et al. (2012). Architecture of the human regulatory
358 network derived from ENCODE data. *Nature* 489, 91–100.
- 359 Gianola, D., de los Campos, G., Hill, W.G., Manfredi, E., and Fernando, R. (2009). Additive
360 Genetic Variability and the Bayesian Alphabet. *Genetics* 183, 347–363.
- 361 Gudbjartsson, D.F., Walters, G.B., Thorleifsson, G., Stefansson, H., Halldorsson, B. V,
362 Zusmanovich, P., Sulem, P., Thorlacius, S., Gylfason, A., Steinberg, S., et al. (2008). Many
363 sequence variants affecting diversity of adult human height. *Nat. Genet.* 40, 609–615.
- 364 Hayes, B.J., Bowman, P.J., Chamberlain, A.C., Verbyla, K., and Goddard, M.E. (2009).
365 Accuracy of genomic breeding values in multi-breed dairy cattle populations. *Genet. Sel.*
366 *Evol.* 41, 51.
- 367 Hill, W.G., Goddard, M.E., and Visscher, P.M. (2008). Data and theory point to mainly
368 additive genetic variance for complex traits. *PLoS Genet.* 4, e1000008.
- 369 Huang, W., and Mackay, T.F.C. (2016). The Genetic Architecture of Quantitative Traits

- 370 Cannot Be Inferred from Variance Component Analysis. *PLOS Genet.* 12, e1006421.
- 371 International Schizophrenia Consortium, Purcell, S.M., Wray, N.R., Stone, J.L., Visscher, P.M.,
372 O'Donovan, M.C., Sullivan, P.F., and Sklar, P. (2009). Common polygenic variation
373 contributes to risk of schizophrenia and bipolar disorder. *Nature* 460, 748–752.
- 374 Khera, A. V., Chaffin, M., Aragam, K.G., Haas, M.E., Roselli, C., Choi, S.H., Natarajan, P.,
375 Lander, E.S., Lubitz, S.A., Ellinor, P.T., et al. (2018). Genome-wide polygenic scores for
376 common diseases identify individuals with risk equivalent to monogenic mutations. *Nat.*
377 *Genet.* 50, 1219–1224.
- 378 Lango Allen, H., Estrada, K., Lettre, G., Berndt, S.I., Weedon, M.N., Rivadeneira, F., Willer,
379 C.J., Jackson, A.U., Vedantam, S., Raychaudhuri, S., et al. (2010). Hundreds of variants
380 clustered in genomic loci and biological pathways affect human height. *Nature* 467, 832–
381 838.
- 382 Lello, L., Avery, S.G., Tellier, L., Vazquez, A.I., de los Campos, G., and Hsu, S.D.H. (2018).
383 Accurate Genomic Prediction of Human Height. *Genetics* 210, 477–497.
- 384 Lettre, G., Jackson, A.U., Gieger, C., Schumacher, F.R., Berndt, S.I., Sanna, S., Eyheramendy,
385 S., Voight, B.F., Butler, J.L., Guiducci, C., et al. (2008). Identification of ten loci associated
386 with height highlights new biological pathways in human growth. *Nat. Genet.* 40, 584–591.
- 387 Liu, X., Li, Y.I., and Pritchard, J.K. (2019). Trans Effects on Gene Expression Can Drive
388 Omnigenic Inheritance. *Cell* 177, 1022-1034.e6.
- 389 Mackay, T.F.C. (2014). Epistasis and quantitative traits: using model organisms to study
390 gene-gene interactions. *Nat. Rev. Genet.* 15, 22–33.
- 391 Mackay, T.F., and Moore, J.H. (2014). Why epistasis is important for tackling complex human
392 disease genetics. *Genome Med.* 6, 125.
- 393 Makowsky, R., Pajewski, N.M., Klimentidis, Y.C., Vazquez, A.I., Duarte, C.W., Allison, D.B.,
394 and de los Campos, G. (2011). Beyond Missing Heritability: Prediction of Complex Traits.
395 *PLoS Genet.* 7, e1002051.
- 396 Manolio, T.A., Collins, F.S., Cox, N.J., Goldstein, D.B., Hindorff, L.A., Hunter, D.J.,
397 McCarthy, M.I., Ramos, E.M., Cardon, L.R., Chakravarti, A., et al. (2009). Finding the missing
398 heritability of complex diseases. *Nature* 461, 747–753.
- 399 Martin, A.R., Kanai, M., Kamatani, Y., Okada, Y., Neale, B.M., and Daly, M.J. (2019). Clinical
400 use of current polygenic risk scores may exacerbate health disparities. *Nat. Genet.* 51, 584–
401 591.
- 402 Meuwissen, T.H., Hayes, B.J., and Goddard, M.E. (2001). Prediction of total genetic value
403 using genome-wide dense marker maps. *Genetics* 157, 1819–1829.

- 404 Morgante, F., Huang, W., Maltecca, C., and Mackay, T.F.C. (2018). Effect of genetic
405 architecture on the prediction accuracy of quantitative traits in samples of unrelated
406 individuals. *Heredity (Edinb)*. 120, 500–514.
- 407 Mostafavi, H., Harpak, A., Conley, D., Pritchard, J.K., and Przeworski, M. (2019). Variable
408 prediction accuracy of polygenic scores within an ancestry group. *BioRxiv* 629949.
- 409 Silventoinen, K., Sarmalisto, S., Perola, M., Boomsma, D.I., Cornes, B.K., Davis, C., Dunkel,
410 L., de Lange, M., Harris, J.R., Hjelmberg, J.V.B., et al. (2003). Heritability of Adult Body
411 Height: A Comparative Study of Twin Cohorts in Eight Countries. *Twin Res.* 6, 399–408.
- 412 VanRaden, P.M. (2008). Efficient methods to compute genomic predictions. *J. Dairy Sci.* 91,
413 4414–4423.
- 414 Varona, L., Legarra, A., Toro, M.A., and Vitezica, Z.G. (2018). Non-additive Effects in
415 Genomic Selection. *Front. Genet.* 9, 78.
- 416 Wainschtein, P., Jain, D.P., Yengo, L., Zheng, Z., TOPMed Anthropometry Working Group,
417 T.-O. for P.M.C., Cupples, L.A., Shadyab, A.H., McKnight, B., Shoemaker, B.M., Mitchell,
418 B.D., et al. (2019). Recovery of trait heritability from whole genome sequence data. *BioRxiv*
419 588020.
- 420 Weedon, M.N., Lango, H., Lindgren, C.M., Wallace, C., Evans, D.M., Mangino, M., Freathy,
421 R.M., Perry, J.R.B., Stevens, S., Hall, A.S., et al. (2008). Genome-wide association analysis
422 identifies 20 loci that influence adult height. *Nat. Genet.* 40, 575–583.
- 423 Yang, J., Benyamin, B., McEvoy, B.P., Gordon, S., Henders, A.K., and others (2010).
424 Common SNPs explain a large proportion of the heritability for human height. *Nat Gen* 42,
425 565–569.
- 426 Yang, J., Lee, S.H., Goddard, M.E., and Visscher, P.M. (2011). GCTA: A tool for genome-
427 wide complex trait analysis. *Am. J. Hum. Genet.* 88, 76–82.

428

429 Supplemental Figures



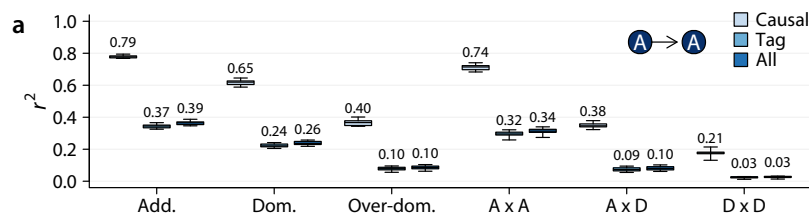
445 Figure S1. Genomic heritability in different simulated populations. Genomic
 446 heritability was plotted for different population samples. (a) population A; (b)
 447 population B; (c) population C; (d) population A + B; and (e) population A + C.
 448
 449
 450

451

452

453

454

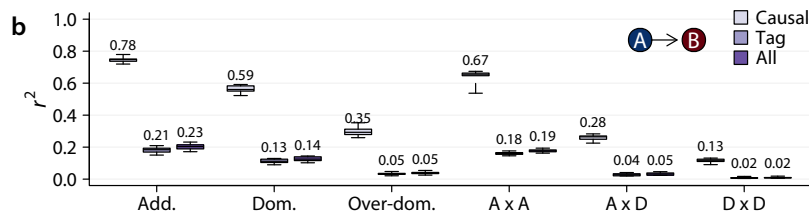


455

456

457

458

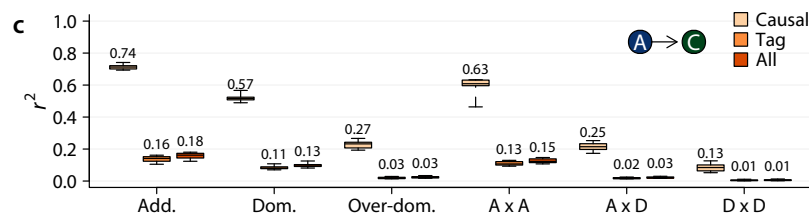


459

460

461

462

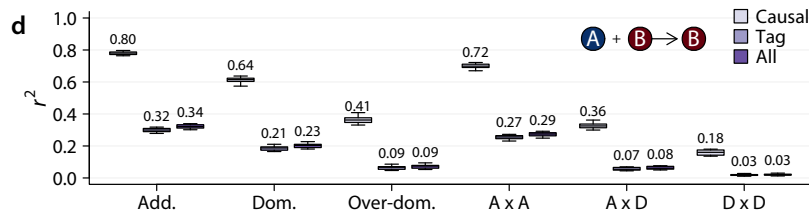


463

464

465

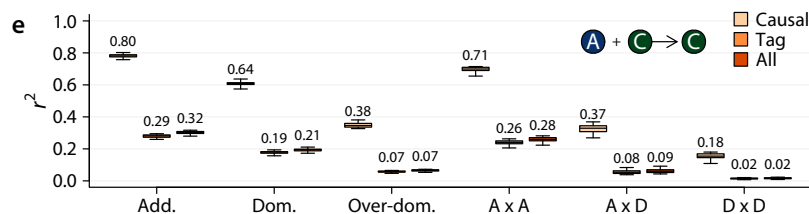
466



467

468

469



470

471

472

Figure S2. Polygenic prediction under different genetic architecture in different populations. Accuracies of polygenic prediction under different genetic architecture in (a) fit model in A, predict in A; (b) fit in A, predict in B; (c) fit in A, predict in C; (d) fit in A + B, predict in B; (e) fit in A + C, predict in C.