

SUPPLEMENTARY MATERIALS

Santantonio, N., Jannink, J., Sorrells, M. 2018. Prediction of sub-genome additive and interaction effects in allohexaploid wheat. G3.

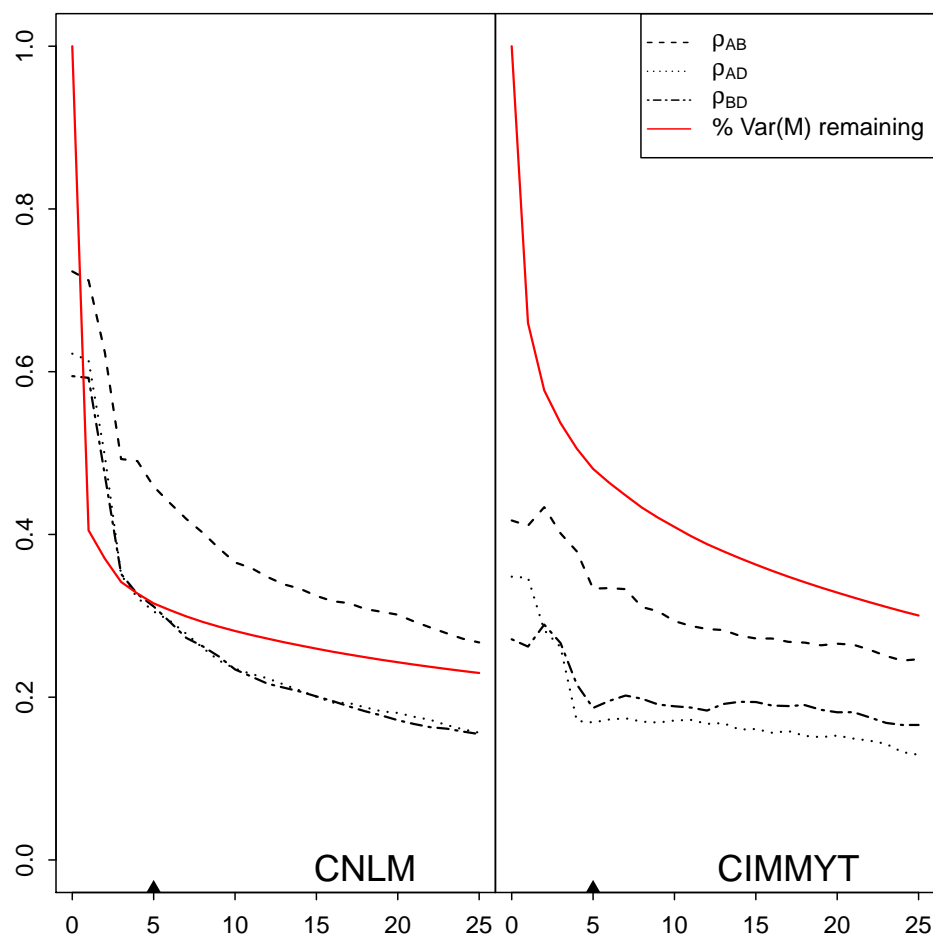


Figure S1 Correlation coefficient, ρ , of off-diagonal elements of estimated additive covariance matrices \mathbf{K}_A , \mathbf{K}_B and \mathbf{K}_D . The percent genotype marker variance remaining in the marker matrix after removing k dimensions is shown in red. The chosen population structure dimension $k = 5$, is indicated by a \blacktriangle .

■ **Table S1** Table of model fit statistics for whole genome and subgenome genomic prediction models in the CNLM population.

Model	Terms	GY (0.30)	PH (0.73)	TW (0.53)	HD (0.79)
G (2) ^a	log \mathcal{L}	-48	2237	1547	6343
	parameters	28	26	28	27
	AIC	153	-4423	-3037	-12631
	G	0.268 ^b (12.59) ^c	3.823 (20.75)	1.067 (16.66)	3.900 (21.16)
	R	0.324 (61.86) ^d	0.135 (56.17)	0.200 (60.12)	0.054 (58.76)
ABD (4)	log \mathcal{L}	-48	2242	1549	6366
	parameters	30	28	30	29
	AIC	155	-4428	-3037	-12673
	A	0.098 (5.86)	1.280 (8.41)	0.503 (8.24)	0.907 (7.41)
	B	0.153 (6.88)	1.585 (8.7)	0.380 (7.10)	1.534 (9.04)
	D	0.021 (2.47)	0.892 (6.24)	0.181 (4.38)	1.387 (8.03)
	R	0.324 (61.87)	0.135 (56.16)	0.199 (60.12)	0.054 (58.75)
G×G (5)	log \mathcal{L}	-43	2360	1630	6432
	parameters	29	27	29	28
	AIC	144	-4665	-3203	-12808
	G	0.203 (7.86)	0.889 (6.46)	0.194 (4.47)	1.121 (7.30)
	H	0.018 (3.04)	0.478 (11.95)	0.184 (11.33)	0.451 (11.13)
	R	0.322 (61.39)	0.132 (56.50)	0.195 (60.25)	0.053 (58.98)
ABD×ABD (7)	log \mathcal{L}	-41	2375	1634	6451
	parameters	33	31	33	32
	AIC	149	-4687	-3201	-12839
	A	0.076 (4.66)	0.292 (3.7)	0.079 (2.79)	0.104 (1.71)
	B	0.119 (5.40)	0.429 (4.21)	0.114 (3.44)	0.587 (5.38)
	D	0.015 (1.95)	0.279 (3.31)	0.007 (0.50)	0.664 (5.26)
	AB	0.005 (0.49)	0.007 (0.11)	0.073 (2.38)	0.223 (3.29)
	AD	0.012 (1.36)	0.276 (4.16)	0.073 (2.51)	0.167 (2.49)
	BD	0.000	0.149 (2.19)	0.034 (1.15)	0.005 (0.08)
	R	0.322 (61.40)	0.132 (56.51)	0.195 (60.26)	0.053 (58.98)

^a Equation

^b Variance component estimates reported for additive main effects (G, A, B and D) and epistatic interactions (H, A×B, A×D, B×D) are the ratios of the actual variance component to the residual variance component for ease of comparison.

^c The variance component divided by their respective standard errors are shown in parentheses.

^d The residual variance components are the actual estimates from the centered and scaled data (refer to Table ?? for scaling coefficients) with their associated standard errors in parentheses.

■ **Table S2** Table of model fit statistics for whole genome and subgenome genomic prediction models in the W-GY population.

Model	Terms	E1	E2	E3	E4
G (2) ^a	$\log \mathcal{L}$	-243	-243	-264	-248
	parameters	2	2	2	2
	AIC	489	490	533	500
	G	0.550 ^b (5.47) ^c	0.456 (5.10)	0.295 (4.23)	0.369 (4.69)
	R	0.541 (11.88) ^d	0.568 (12.17)	0.669 (12.73)	0.606 (12.5)
ABD (4)	$\log \mathcal{L}$	-242	-242	-264	-247
	parameters	4	4	4	4
	AIC	492	492	536	503
	A	0.241 (3.38)	0.090 (1.90)	0.062 (1.66)	0.097 (2.10)
	B	0.215 (2.88)	0.267 (3.48)	0.222 (3.23)	0.188 (2.89)
	D	0.067 (1.72)	0.093 (1.92)	0.022 (0.93)	0.064 (1.86)
	R	0.547 (11.97)	0.571 (12.28)	0.668 (12.79)	0.613 (12.59)
G×G (5)	$\log \mathcal{L}$	-222	-242	-249	-233
	parameters	3	3	3	3
	AIC	451	491	504	472
	G	0.426 (3.25)	0.440 (4.39)	0.303 (2.73)	0.314 (3.04)
	H	0.272 (4.59)	0.033 (1.02)	0.325 (4.81)	0.231 (4.34)
	R	0.343 (7.31)	0.532 (9.34)	0.368 (6.61)	0.391 (7.68)
ABD×ABD (7)	$\log \mathcal{L}$	-222	-241	-247	-232
	parameters	7	7	7	7
	AIC	458	495	509	478
	A	0.124 (1.56)	0.070 (1.52)	0.061 (1.15)	0.080 (1.47)
	B	0.269 (2.43)	0.273 (3.25)	0.239 (2.23)	0.138 (1.79)
	D	0.070 (1.26)	0.093 (1.81)	0.041 (0.94)	0.087 (1.74)
	AB	0.219 (3.39)	0.04 (1.33)	0.215 (2.69)	0.167 (2.88)
	AD	0.029 (0.62)	0.000	0.097 (1.52)	0.007 (0.14)
	BD	0.000	0.000	0.000	0.027 (0.57)
	R	0.349 (7.38)	0.526 (9.57)	0.366 (6.74)	0.409 (7.92)

^a Equation

^b Variance component estimates reported for additive main effects (G, A, B and D) and epistatic interactions (H, A×B, A×D, B×D) are the ratios of the actual variance component to the residual variance component for ease of comparison.

^c The variance component divided by their respective standard errors are shown in parentheses.

^d The residual variance components are the actual estimates from the centered and scaled data (refer to [Crossa et al. \(2010\)](#) for scaling coefficients).

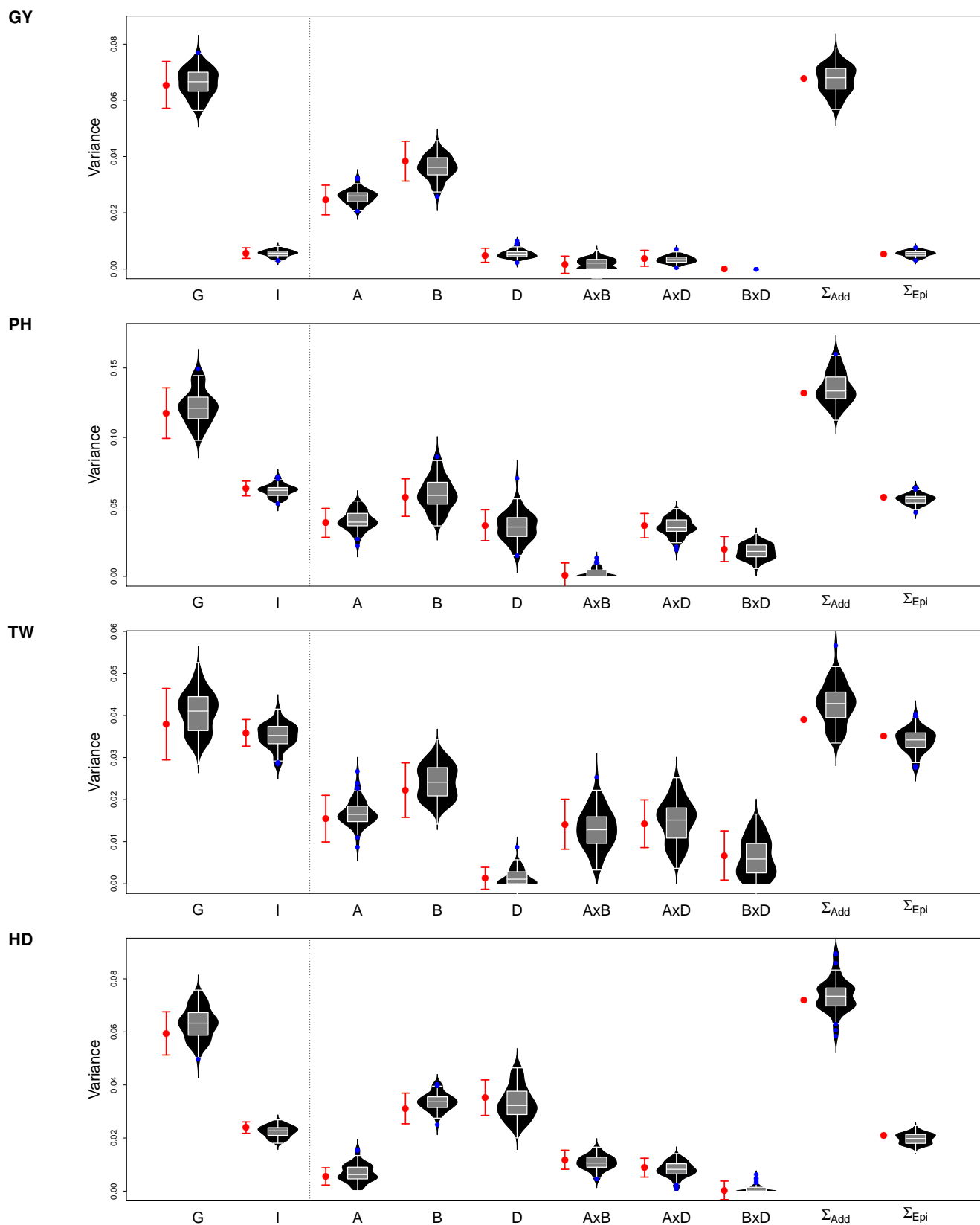


Figure S2 Estimates and standard errors of variance components for four traits in the CNLM populations from the full model (red) compared to the sampling distribution of variance component estimates from the cross-validation scheme (black violins). G×G (5) and ABD×ABD (7) models are shown to the left and right of the dotted line, respectively. The sum of the additive (Σ_{Add}) and epistatic (Σ_{Epi}) variance components is also shown for the ABD×ABD model.

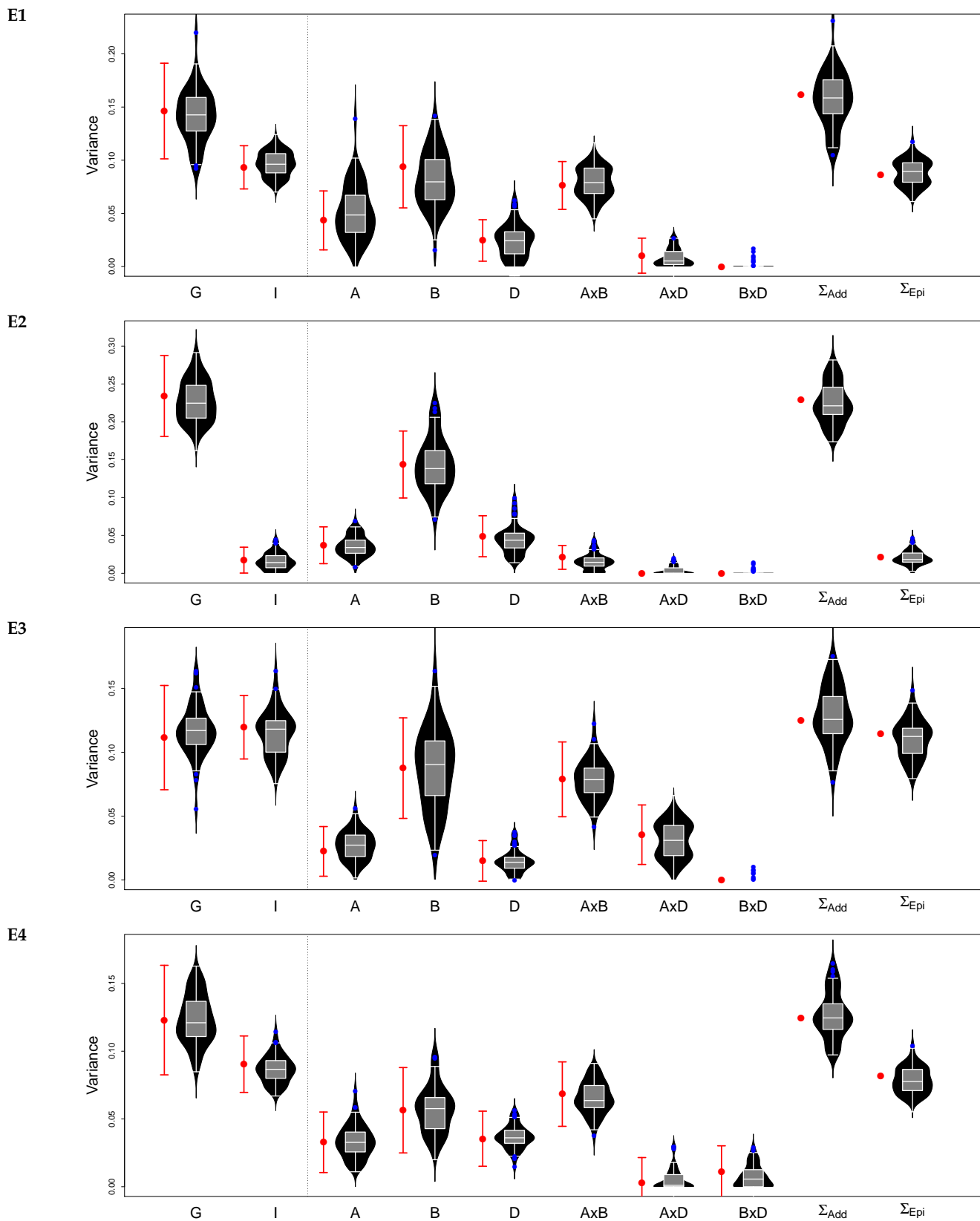


Figure S3 Estimates and standard errors of variance components for four traits in the W-GY populations from the full model (red) compared to the sampling distribution of variance component estimates from the cross-validation scheme (black violins). $G \times G$ (5) and $ABD \times ABD$ (7) models are shown to the left and right of the dotted line, respectively. The sum of the additive (Σ_{Add}) and epistatic (Σ_{Epi}) variance components is also shown for the $ABD \times ABD$ model.

■ **Table S3** Correlation of whole genome and subgenome additive effects in the CNLM population from models G (2) and ABD (4). Correlations of additive random effects without correcting for population structure are shown above the diagonal, while correlations of effects correcting for populations structure using the first $k = 5$ PCs is shown below the diagonal.

	GY				PH				TW				HD			
	G	A	B	D	G	A	B	D	G	A	B	D	G	A	B	D
G		0.80	0.82	0.48		0.74	0.68	0.60		0.78	0.74	0.50		0.60	0.74	0.70
A	0.75		0.36	0.22	0.73		0.15	0.25	0.75		0.26	0.25	0.60		0.20	0.14
B	0.83	0.31		0.31	0.70	0.19		0.17	0.73	0.20		0.13	0.71	0.18		0.26
D	0.46	0.18	0.26		0.59	0.23	0.14		0.49	0.23	0.12		0.65	0.14	0.14	

■ **Table S4** Correlation of whole genome and subgenome additive effects in the W-GY population from models G (2) and ABD (4). Correlations of additive random effects without correcting for population structure are shown above the diagonal, while correlations of effects correcting for populations structure using the first $k = 5$ PCs is shown below the diagonal.

	E1				E2				E3				E4			
	G	A	B	D	G	A	B	D	G	A	B	D	G	A	B	D
G		0.83	0.80	0.51		0.69	0.84	0.57		0.68	0.90	0.50		0.73	0.80	0.52
A	0.83		0.41	0.28	0.65		0.31	0.25	0.74		0.36	0.07	0.63		0.34	0.09
B	0.79	0.39		0.20	0.91	0.39		0.27	0.91	0.45		0.37	0.86	0.33		0.21
D	0.57	0.36	0.24		0.57	0.22	0.32		0.48	0.28	0.25		0.63	0.22	0.30	

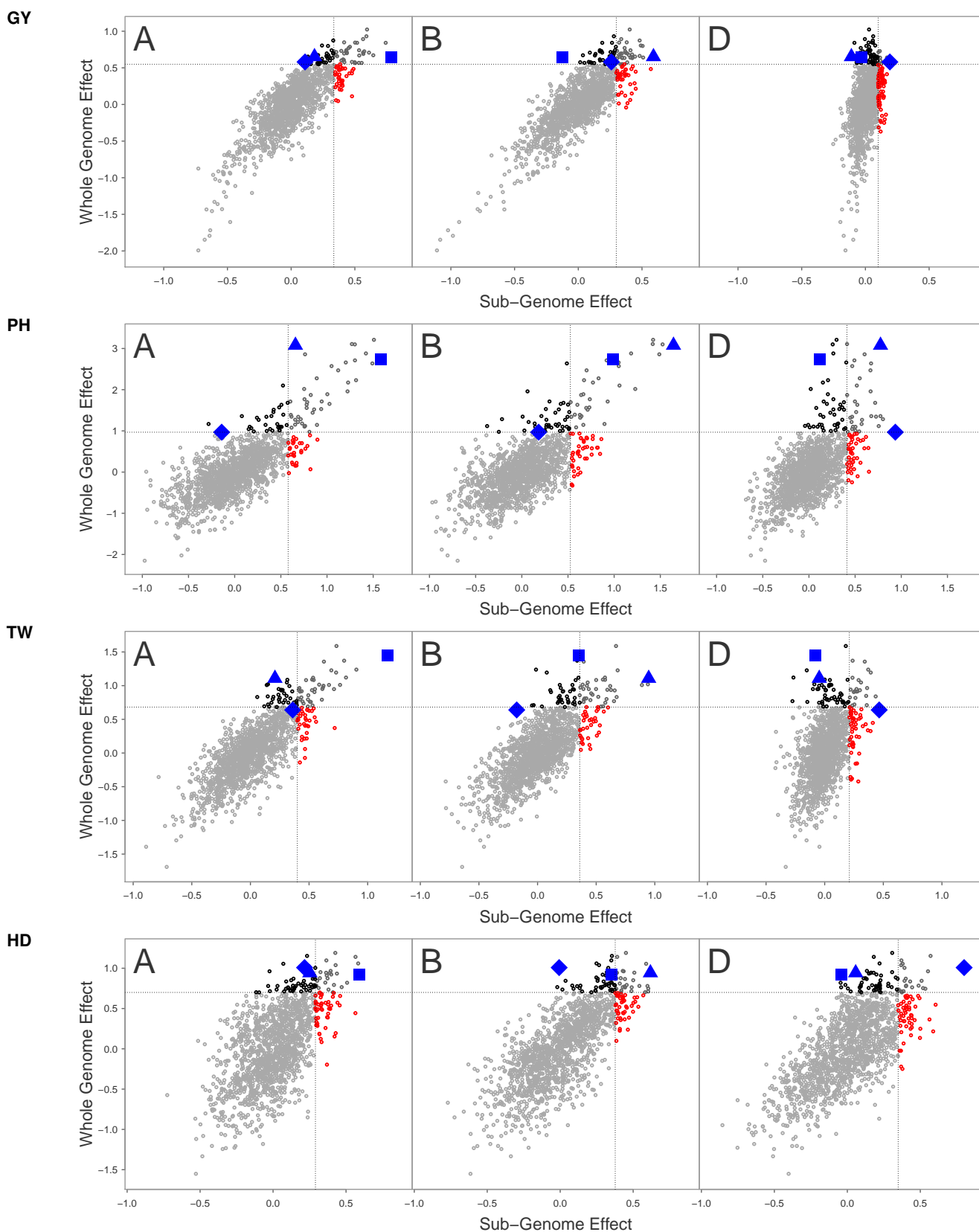
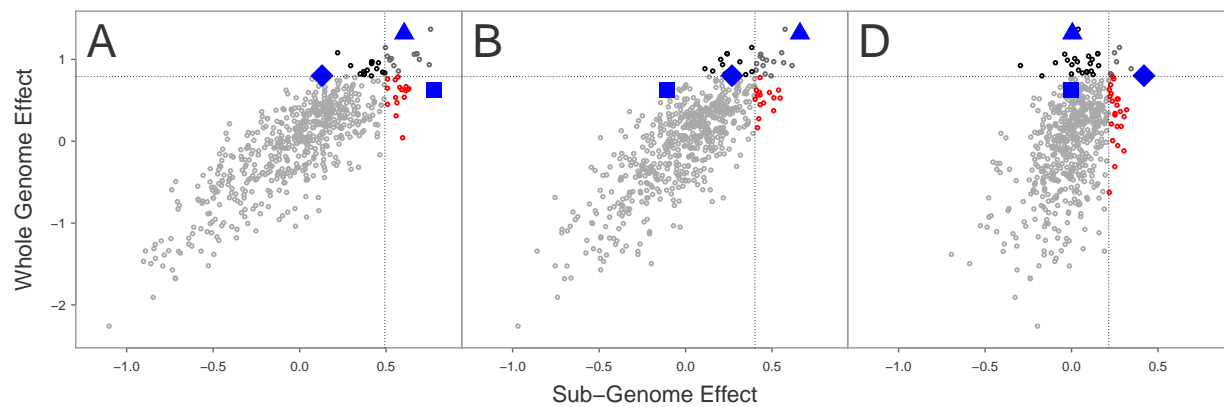
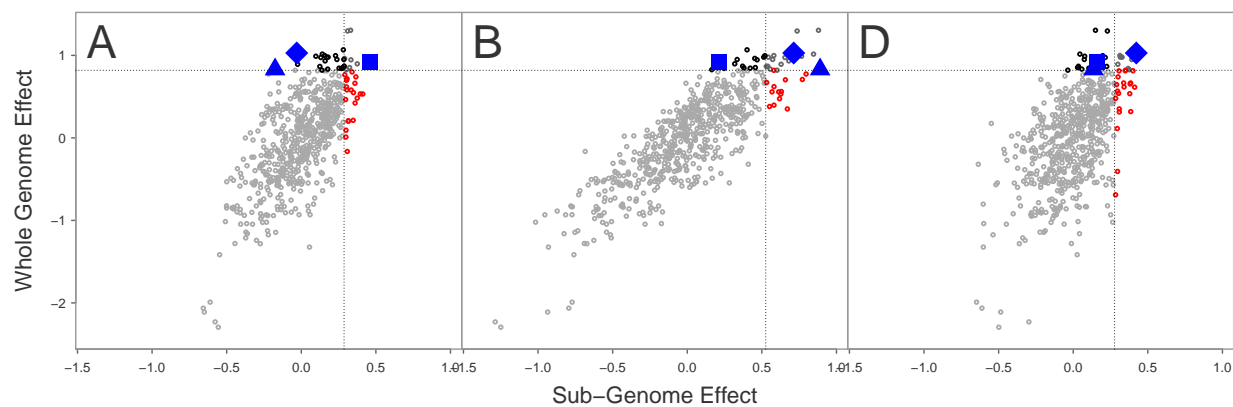


Figure S4 Plot of whole genome additive effects (GEBV) by subgenome additive effects (SGEBV) for four traits in the CNLM populations. The dotted line indicates the 95% quantiles for whole or subgenome effects. Blue squares, triangles and diamonds indicate the line with the highest SGEV for each of the A, B and D subgenomes, respectively.

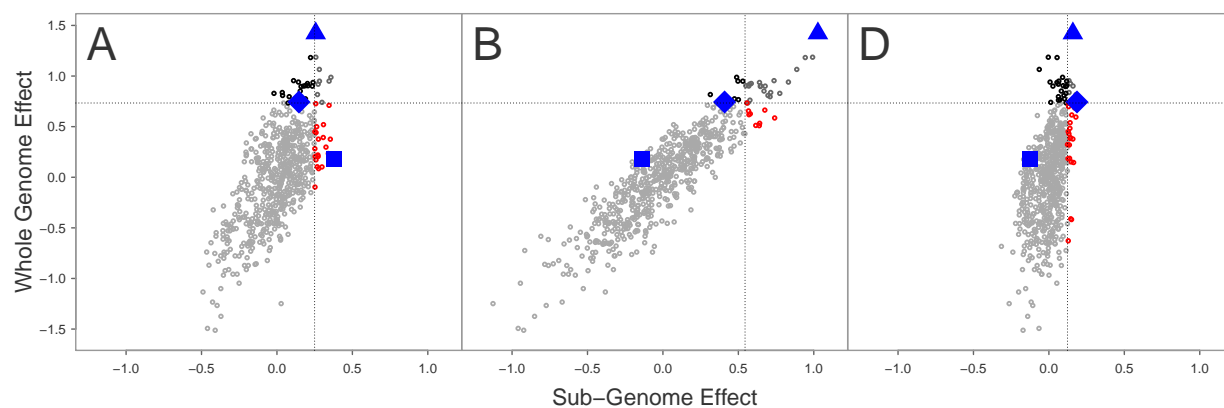
E1



E2



E3



E4

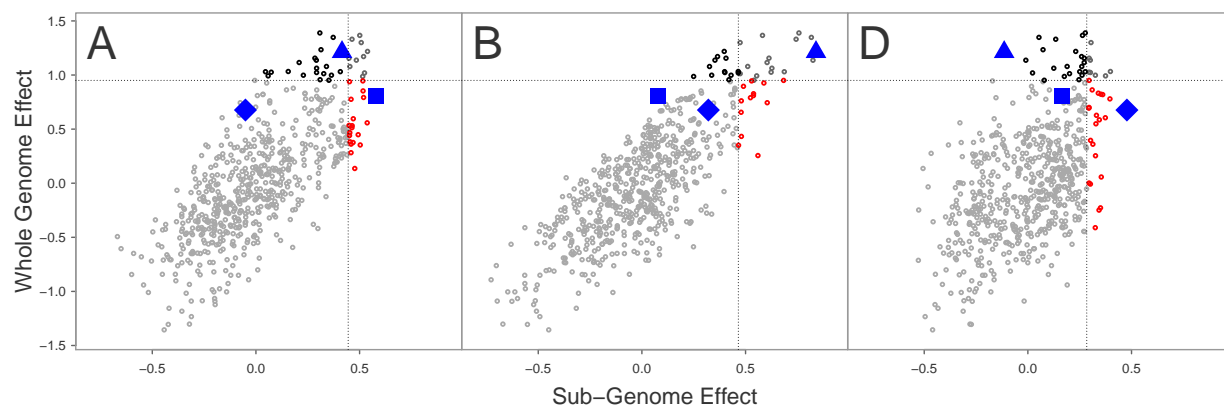
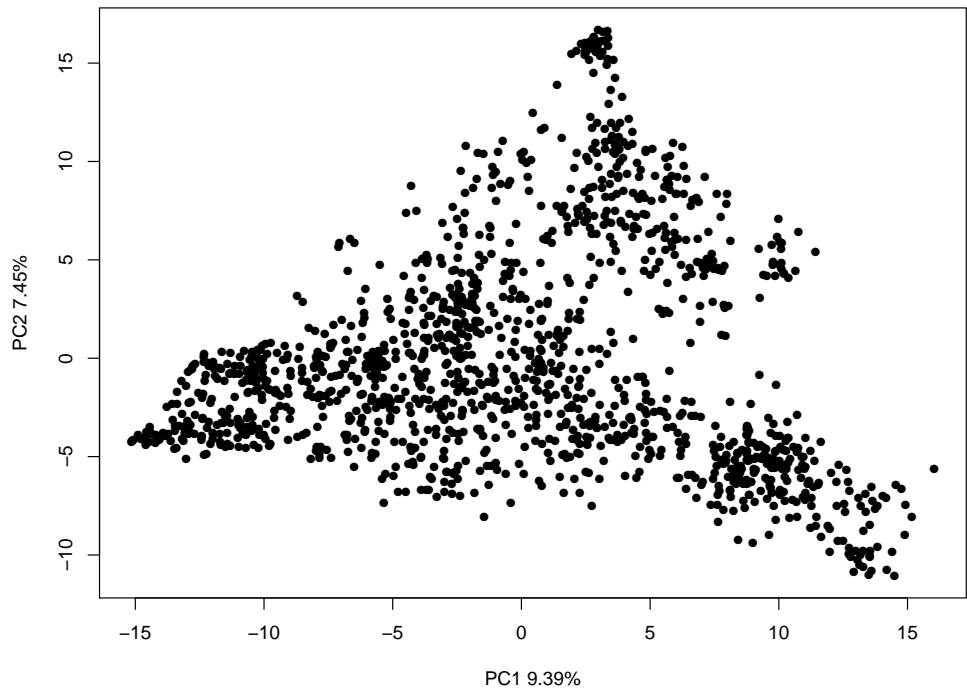


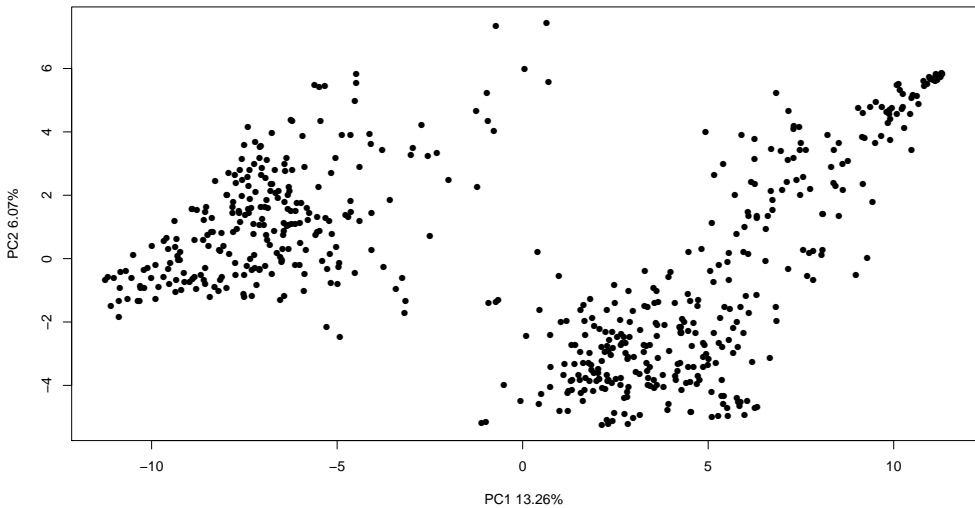
Figure S5 Plot of whole genome additive effects (GEBV) by subgenome additive effects (SGEBV) for four traits in the W-GY populations. The dotted line indicates the 95% quantiles for whole or subgenome effects. Blue squares, triangles and diamonds indicate the line with the highest SGEV for each of the A, B and D subgenomes, respectively.

Figure S6 Plot of the first two principal components of the additive covariance matrix **K** in the CNLM and W-GY populations.

CNLM



W-GY



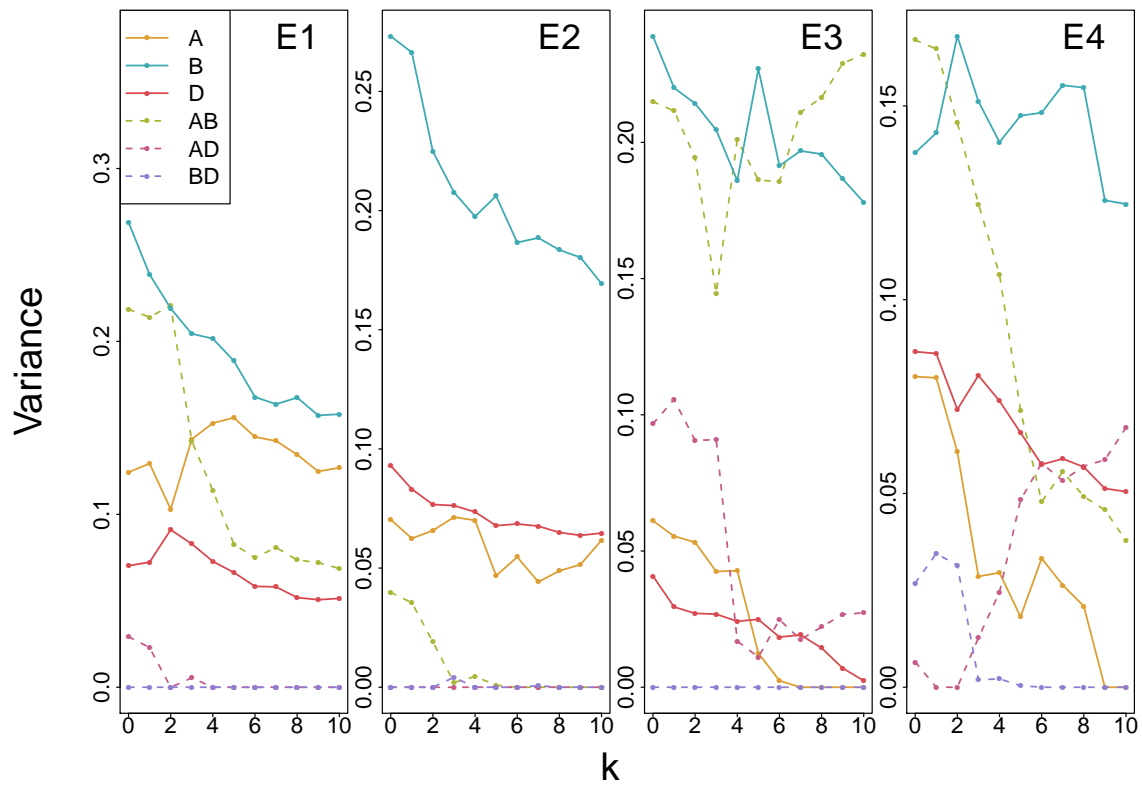


Figure S7 Subgenome additive and interaction variance parameter estimates from the ABD \times ABD model correcting for population structure with $k \in \{0, 1, \dots, 10\}$ principal components as fixed effects. Models were fit with four traits for the W-GY population.

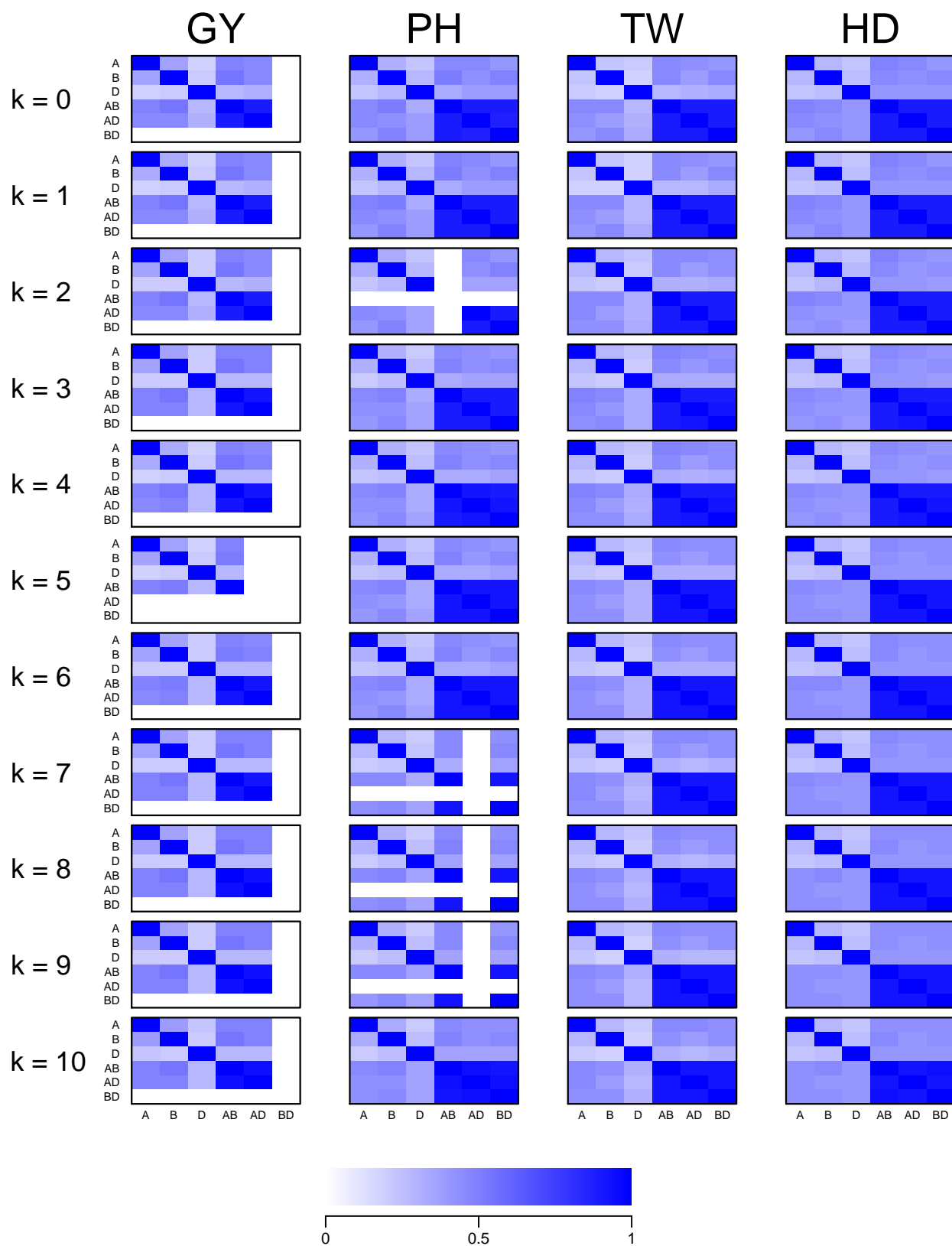


Figure S8 Correlation of variance component estimates derived from the average information from the model fit for models correcting for population structure with $k \in \{0, 1, \dots, 10\}$ principal components for four traits in the CNLM population.

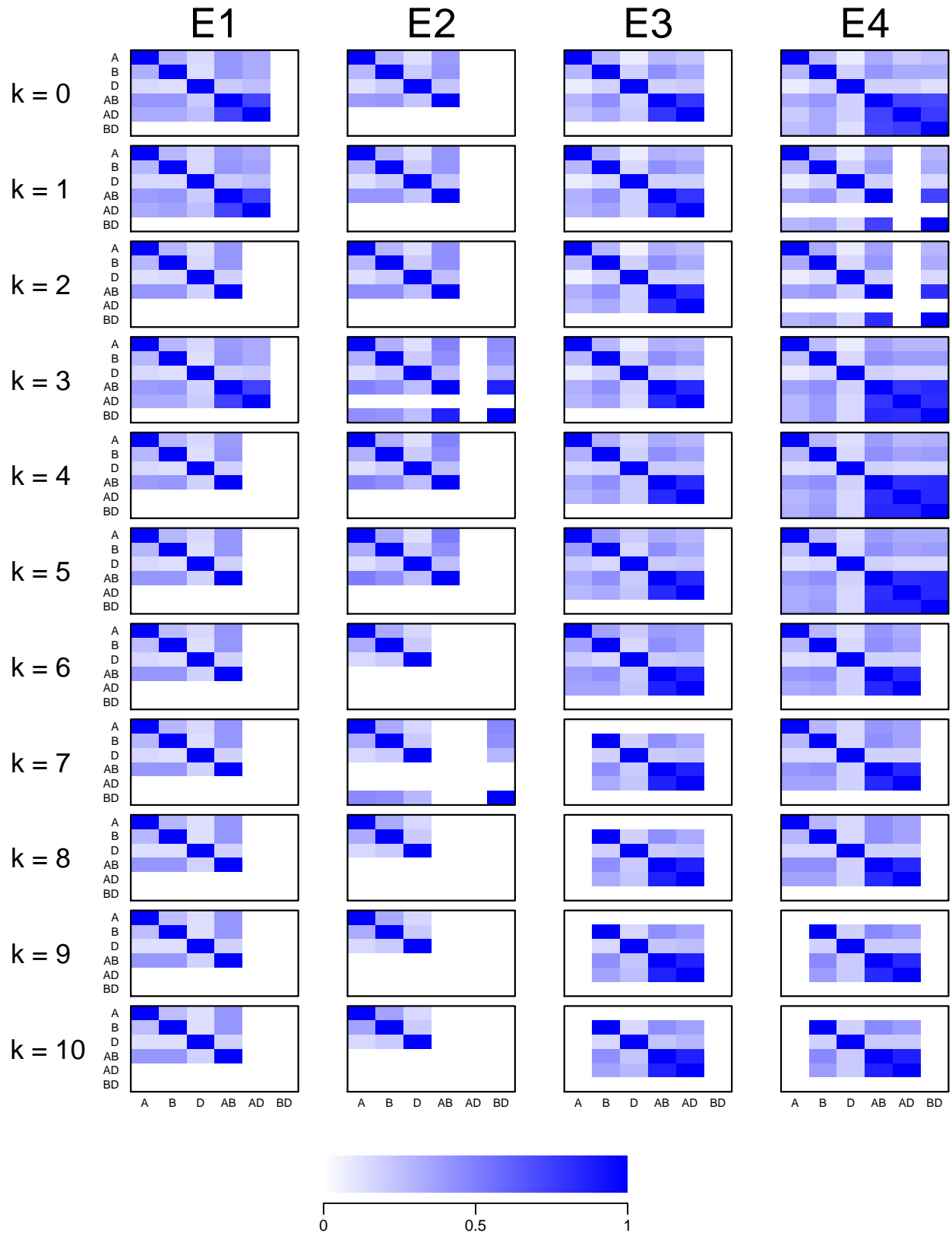


Figure S9 Correlation of variance component estimates derived from the average information from the model fit for models correcting for population structure with $k \in \{0, 1, \dots, 10\}$ principal components for four traits in the W-GY population.