Supplementary Information

A mixed-model approach for genome-wide association studies of correlated traits in structured populations

Arthur Korte[1,*], Bjarni J. Vilhjálmsson[1,2,*], Vincent Segura[1,3,*], Alexander Platt[1,2], Quan Long[1], Magnus Nordborg[1,2]

1 Gregor Mendel Institute, Austrian Academy of Sciences, Vienna, Austria.

2 Molecular and Computational Biology, University of Southern California, Los Angeles, California, United States of America.

3 Institut National de la Recherche Agronomique, UR0588, F-45075 Orléans, France.

* These authors contributed equally to this work.

Contents

Supplementary Tables

Supplementary Table 1 SNPs detected in the analysis of HDL and CRP.

		MTMM (p-value[a])			EMMAX (p-value[a])	
SNP	Position	full test	interaction	common	HDL	CRP
CRP region, chromosome 1						
rs1811472	157908973	1.86×10^{-12}	1.46×10^{-11}	1.62×10^{-3}	0.05	1.28×10^{-15}
rs12093699	157914612	8.39×10^{-13}	1.19×10^{-8}	1.53×10^{-6}	0.92	1.02×10^{-14}
rs2592887	157919563	1.55×10^{-14}	2.45×10^{-13}	1.62×10^{-3}	0.03	7.44×10^{-18}
rs2794520	157945440	1.37×10^{-19}	9.77×10^{-16}	2.23×10^{-6}	0.18	2.97×10^{-23}
rs11265260	157966663	1.06×10^{-7}	5.65×10^{-8}	0.10	0.04	1.13×10^{-8}
PPP1R3B region, chromosome 8						
rs983309	9215142	2.28×10^{-7}	0.12	1.09×10^{-7}	1.62×10^{-3}	1.61×10^{-4}
rs2126259	9222556	4.04×10^{-9}	0.27	9.37×10^{-10}	3.68×10^{-5}	1.50×10^{-4}
LEF1 region, chromosome 12						
rs2650000	119873345	2.07×10^{-10}	5.65×10^{-9}	1.62×10^{-3}	0.30	1.99×10^{-12}
rs7953249	119888107	1.09×10^{-9}	1.20×10^{-8}	1.62×10^{-3}	0.24	1.67×10^{-11}
rs1169300	119915608	2.56×10^{-9}	2.26×10^{-7}	1.62×10^{-3}	0.71	1.07×10^{-10}
rs2464196	119919810	2.55×10^{-9}	1.87×10^{-7}	1.62×10^{-3}	0.67	1.05×10^{-10}
rs735396	119923227	6.86×10^{-8}	2.11×10^{-6}	1.62×10^{-3}	0.74	5.22×10^{-9}
LIPC region, chromosome 15						
rs1532085	56470658	4.92×10^{-11}	1.62×10^{-3}	4.60×10^{-10}	9.98×10^{-12}	0.71
rs415799	56478046	4.40×10^{-7}	1.62×10^{-3}	1.28×10^{-5}	4.92×10^{-8}	0.50
CETP region, chromosome 16						
rs9989419	55542640	2.57×10^{-8}	1.62×10^{-3}	2.76×10^{-7}	4.48×10^{-9}	0.76
rs3764261	55550825	6.03×10^{-31}	1.42×10^{-10}	4.11×10^{-23}	3.86×10^{-32}	0.33
rs1532624	55562980	2.14×10^{-23}	1.08×10^{-8}	2.50×10^{-17}	1.54×10^{-24}	0.30
rs7499892	55564091	2.32×10^{-16}	4.93×10^{-6}	8.62×10^{-13}	3.50×10^{-17}	0.40
LCAT region, chromosome 16						
rs255049	66570972	8.37×10^{-8}	1.62×10^{-3}	1.55×10^{-6}	1.36×10^{-8}	0.78
HNF4A region, chromosome 20						
rs1800961	42475778	8.71×10^{-9}	0.10	4.33×10^{-9}	1.84×10^{-8}	0.17

[a] P-values below the Bonferroni-corrected 5% cut-off of 1.5×10^{-7} are highlighted in red.

Supplementary Table 2 SNPs detected in the analysis of TG and HDL.

		MTMM (p-value[a])			EMMAX (p-value[a])	
SNP	Position	full test	interaction	common	TG	HDL
APOB region, chromosome 2						
rs6728178	21047434	4.57×10^{-9}	6.84×10^{-10}	0.56	1.81×10^{-7}	4.54×10^{-7}
rs6754295	21059688	9.42×10^{-9}	1.52×10^{-9}	0.50	4.12×10^{-7}	5.59×10^{-7}
rs676210	21085029	2.97×10^{-9}	4.11×10^{-10}	0.65	9.21×10^{-8}	5.21×10^{-7}
rs673548	21091049	1.95×10^{-9}	2.66×10^{-10}	0.66	6.43×10^{-8}	4.13×10^{-7}
GCKR region, chromosome 2						
rs1260326	27584444	6.84×10^{-10}	5.57×10^{-7}	3.45×10^{-5}	1.87×10^{-10}	0.28
rs780094	27594741	6.57×10^{-9}	8.42×10^{-6}	2.40×10^{-5}	3.15×10^{-9}	0.54
LPL region, chromosome 8						
rs10096633	19875201	2.79×10^{-9}	3.47×10^{-10}	0.96	1.93×10^{-8}	2.70×10^{-6}
LIPC region, chromosome 15						
rs166358	56468097	4.39×10^{-8}	0.08	2.97×10^{-8}	0.34	5.79×10^{-7}
rs1532085	56470658	2.12×10^{-15}	0.03	2.00×10^{-15}	0.09	9.98×10^{-12}
rs415799	56478046	2.89×10^{-10}	0.10	1.35×10^{-10}	0.15	4.92×10^{-8}
rs473224	56524633	4.17×10^{-8}	0.28	1.00×10^{-8}	3.49×10^{-4}	3.01×10^{-3}
rs261336	56529710	7.54×10^{-10}	0.27	1.71×10^{-10}	9.09×10^{-5}	8.01×10^{-4}
CETP region, chromosome 16						
rs9989419	55542640	7.18×10^{-9}	1.19×10^{-3}	2.03×10^{-7}	0.58	4.48×10^{-9}
rs3764261	55550825	7.29×10^{-33}	5.63×10^{-12}	1.16×10^{-23}	0.12	3.86×10^{-32}
rs1532624	55562980	2.40×10^{-24}	3.70×10^{-10}	7.59×10^{-17}	0.07	1.54×10^{-24}
rs7499892	55564091	2.99×10^{-20}	5.79×10^{-4}	9.92×10^{-19}	0.34	3.50×10^{-17}
LCAT region, chromosome 16						
rs6499137	66229305	1.47×10^{-7}	0.03	2.40×10^{-7}	0.67	5.69×10^{-7}
rs255049	66570972	6.66×10^{-8}	1.76×10^{-4}	1.32×10^{-5}	0.18	1.36×10^{-8}
HNF4A region, chromosome 20						
rs1800961	42475778	8.14×10^{-8}	9.03×10^{-7}	3.51×10^{-3}	1.47×10^{-3}	1.84×10^{-8}

[a] P-values below the Bonferroni-corrected 5% cut-off of 1.5×10^{-7} are highlighted in red.

Supplementary Table 3 SNPs detected in the analysis of LDL and CRP.

		MTMM (p-value[a])			EMMAX (p-value[a])	
SNP	Position	full test	interaction	common	LDL	CRP
CELSR2 region, chromosome 1						
rs611917	109616775	1.02×10^{-7}	3.17×10^{-3}	1.26×10^{-6}	1.80×10^{-8}	0.86
rs646776	109620053	4.12×10^{-15}	3.68×10^{-7}	2.08×10^{-10}	3.92×10^{-15}	0.14
CRP region, chromosome 1						
rs1811472	157908973	1.27×10^{-12}	3.12×10^{-12}	0.01	0.59	1.28×10^{-15}
rs12093699	157914612	2.47×10^{-12}	8.48×10^{-11}	7.71×10^{-4}	0.63	1.02×10^{-14}
rs2592887	157919563	1.32×10^{-14}	1.52×10^{-13}	2.20×10^{-3}	0.87	7.44×10^{-18}
rs2794520	157945440	6.02×10^{-20}	1.04×10^{-18}	1.18×10^{-3}	0.61	2.97×10^{-23}
rs11265260	157966663	2.99×10^{-8}	5.98×10^{-9}	0.37	0.12	1.13×10^{-8}
APOB region, chromosome 2						
rs10198175	20997364	1.24×10^{-6}	4.67×10^{-3}	1.18×10^{-5}	9.48×10^{-8}	0.92
rs6728178	21047434	4.96×10^{-7}	0.29	1.27×10^{-7}	7.95×10^{-8}	0.05
rs6754295	21059688	4.17×10^{-7}	0.31	1.01×10^{-7}	7.10×10^{-8}	0.05
rs693	21085700	4.05×10^{-10}	0.02	6.80×10^{-10}	2.84×10^{-11}	0.18
rs1429974	21154275	9.03×10^{-7}	9.34×10^{-3}	4.41×10^{-6}	7.69×10^{-8}	0.48
rs754524	21165046	5.41×10^{-8}	1.43×10^{-3}	1.38×10^{-6}	7.83×10^{-9}	0.80
rs754523	21165196	8.04×10^{-7}	7.26×10^{-3}	4.94×10^{-6}	7.15×10^{-8}	0.53
LEF1 region, chromosome 12						
rs2650000	119873345	7.31×10^{-11}	1.61×10^{-10}	0.02	0.71	1.99×10^{-12}
rs7953249	119888107	4.29×10^{-10}	6.69×10^{-10}	0.02	0.64	1.67×10^{-11}
rs1169300	119915608	1.32×10^{-9}	7.09×10^{-10}	0.09	0.33	1.07×10^{-10}
rs2464196	119919810	1.70×10^{-9}	1.26×10^{-9}	0.06	0.44	1.05×10^{-10}
rs735396	119923227	3.07×10^{-8}	1.03×10^{-8}	0.18	0.25	5.22×10^{-9}
LDLR region, chromosome 19						
rs11668477	11056030	5.00×10^{-8}	3.72×10^{-3}	5.15×10^{-7}	3.89×10^{-9}	0.87
rs2228671	11071912	4.33×10^{-7}	7.61×10^{-3}	2.48×10^{-6}	4.47×10^{-8}	0.62

[a] P-values below the Bonferroni-corrected 5% cut-off of 1.5×10^{-7} are highlighted in red.

Supplementary Table 4 SNPs detected in the analysis of TG and CRP.

		MTMM (p-value[a])			EMMAX (p-value[a])	
SNP	Position	full test	interaction	common	TG	CRP
CRP region, chromosome 1						
rs1811472	157908973	1.25×10^{-12}	1.11×10^{-9}	2.58×10^{-5}	0.89	1.28×10^{-15}
rs12093699	157914612	9.81×10^{-13}	2.65×10^{-10}	8.77×10^{-5}	0.85	1.02×10^{-14}
rs2592887	157919563	1.51×10^{-14}	2.59×10^{-10}	1.14×10^{-6}	0.59	7.44×10^{-18}
rs2794520	157945440	1.06×10^{-20}	2.40×10^{-16}	6.54×10^{-7}	0.67	2.97×10^{-23}
rs11265260	157966663	2.11×10^{-8}	9.80×10^{-8}	8.46×10^{-3}	0.43	1.13×10^{-8}
APOB region, chromosome 2						
rs676210	21085029	6.02×10^{-7}	9.41×10^{-3}	2.87×10^{-6}	9.21×10^{-8}	0.09
rs673548	21091049	4.31×10^{-7}	7.98×10^{-3}	2.36×10^{-6}	6.43×10^{-8}	0.09
GCKR region, chromosome 2						
rs1260326	27584444	1.04×10^{-9}	6.13×10^{-4}	5.22×10^{-8}	1.87×10^{-10}	0.05
rs780094	27594741	1.72×10^{-8}	8.90×10^{-4}	6.65×10^{-7}	3.15×10^{-9}	0.11
LPL region, chromosome 8						
rs10096633	19875201	1.04×10^{-7}	5.61×10^{-4}	6.80×10^{-6}	1.93×10^{-8}	0.38
LEF1 region, chromosome 12						
rs2650000	119873345	9.46×10^{-11}	1.17×10^{-7}	2.11×10^{-5}	0.58	1.99×10^{-12}
rs7953249	119888107	5.12×10^{-10}	1.78×10^{-7}	8.12×10^{-5}	0.74	1.67×10^{-11}
rs1169300	119915608	5.11×10^{-10}	1.94×10^{-8}	8.00×10^{-4}	0.69	1.07×10^{-10}
rs2464196	119919810	5.72×10^{-10}	2.69×10^{-8}	6.45×10^{-4}	0.75	1.05×10^{-10}
rs735396	119923227	6.15×10^{-9}	2.93×10^{-8}	7.88×10^{-3}	0.34	5.22×10^{-9}
APO cluster region, chromosome 19						
rs2075650	50087459	1.64×10^{-8}	2.84×10^{-9}	0.45	6.93×10^{-5}	3.35×10^{-4}

[a] P-values below the Bonferroni-corrected 5% cut-off of 1.5×10^{-7} are highlighted in red.

Supplementary Table 5 SNPs detected in the analysis of HDL and LDL.

		MTMM (p-value[a])			EMMAX (p-value[a])	
SNP	Position	full test	interaction	common	HDL	LDL
CELSR2 region, chromosome 1						
rs611917	109616775	1.29×10^{-7}	2.21×10^{-5}	2.10×10^{-4}	0.33	1.80×10^{-8}
rs646776	109620053	3.58×10^{-14}	1.12×10^{-9}	6.33×10^{-7}	0.14	3.92×10^{-15}
APOB region, chromosome 2						
rs10198175	20997364	5.29×10^{-7}	1.40×10^{-5}	1.54×10^{-3}	0.23	9.48×10^{-8}
rs3923037	21011755	3.42×10^{-8}	1.37×10^{-8}	0.14	5.01×10^{-4}	2.72×10^{-7}
rs6728178	21047434	4.80×10^{-11}	6.27×10^{-12}	0.60	4.54×10^{-7}	7.95×10^{-8}
rs6754295	21059688	4.81×10^{-11}	6.39×10^{-12}	0.58	5.59×10^{-7}	7.10×10^{-8}
rs676210	21085029	3.40×10^{-10}	4.13×10^{-11}	0.83	5.21×10^{-7}	7.23×10^{-7}
rs693	21085700	5.36×10^{-11}	1.63×10^{-9}	9.48×10^{-4}	0.01	2.84×10^{-11}
rs673548	21091049	2.38×10^{-10}	2.86×10^{-11}	0.83	4.13×10^{-7}	5.97×10^{-7}
rs1429974	21154275	2.25×10^{-7}	1.44×10^{-6}	0.01	0.05	7.69×10^{-8}
rs754524	21165046	2.57×10^{-8}	3.51×10^{-7}	2.69×10^{-3}	0.05	7.83×10^{-9}
rs754523	21165196	2.19×10^{-7}	1.54×10^{-6}	0.01	0.05	7.15×10^{-8}
intergenic between FRMD1 and DACT2, chromosome 6						
rs2171981	168269089	2.73×10^{-7}	0.721	4.11×10^{-8}	1.04×10^{-4}	4.76×10^{-4}
PPP1R3B region, chromosome 8						
rs2126259	9222556	8.45×10^{-8}	0.53	1.41×10^{-8}	3.68×10^{-5}	6.29×10^{-4}
LIPC region, chromosome 15						
rs1532085	56470658	3.04×10^{-11}	1.83×10^{-5}	4.16×10^{-8}	9.98×10^{-12}	0.53
rs415799	56478046	2.98×10^{-7}	5.07×10^{-4}	2.25×10^{-5}	4.92×10^{-8}	0.71
CETP region, chromosome 16						
rs9989419	55542640	2.15×10^{-8}	8.28×10^{-5}	8.51×10^{-6}	4.48×10^{-9}	0.86
rs3764261	55550825	9.99×10^{-32}	4.53×10^{-19}	1.93×10^{-15}	3.86×10^{-32}	0.31
rs1532624	55562980	7.43×10^{-24}	1.94×10^{-14}	4.45×10^{-12}	1.54×10^{-24}	0.41
rs7499892	55564091	2.33×10^{-16}	2.61×10^{-8}	1.51×10^{-10}	3.50×10^{-17}	0.72
LCAT region, chromosome 16						
rs255049	66570972	1.03×10^{-8}	2.93×10^{-3}	1.26×10^{-7}	1.36×10^{-8}	0.15
LDLR region, chromosome 19						
rs11668477	11056030	9.37×10^{-9}	9.93×10^{-8}	3.39×10^{-3}	0.02	3.89×10^{-9}
rs2228671	11071912	2.75×10^{-7}	2.88×10^{-4}	3.61×10^{-5}	0.80	4.47×10^{-8}
HNF4A region, chromosome 20						
rs1800961	42475778	1.55×10^{-7}	4.41×10^{-6}	1.34×10^{-3}	1.84×10^{-8}	0.26

[a] P-values below the Bonferroni-corrected 5% cut-off of 1.5×10^{-7} are highlighted in red.

Supplementary Table 6 Correlation- and heritability-estimates for the *A. thaliana* data

	Genetic[a]	Phenotypic (MTMM)	Phenotypic (Pearson)	Heritability[b]
"Spain-Spring/Spain-Summer"	0.90	0.85	0.91	0.93/0.95
"Spain-Spring/Sweden-Spring"	0.88	0.84	0.90	0.93/0.98
"Spain-Spring/Sweden-Summer"	0.84	0.78	0.87	0.93/0.92
"Spain-Summer/Sweden-Spring"	0.92	0.89	0.95	0.95/0.98
"Spain-Summer/Sweden-Summer"	0.96	0.9	0.94	0.95/0.92
"Sweden-Spring/Sweden-Summer"	0.93	0.88	0.94	0.98/0.92

[a] estimated SE are all below ≤ 0.02 , only the SE for Spain-Summer/Sweden-Spring is 0.05

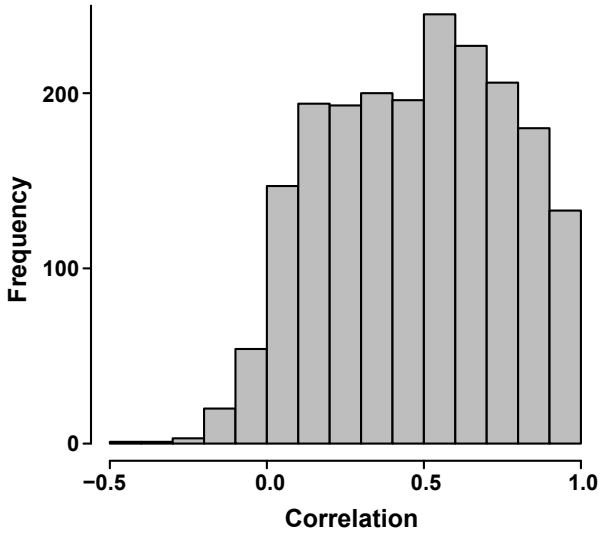
[b] estimated SE are ≤ 0.02; heritability estimates from marginal mixed models are all ≥ 0.99.

Supplementary Table 7 SNPs detected in the analysis of flowering data using a genome-wide significance of 0.05

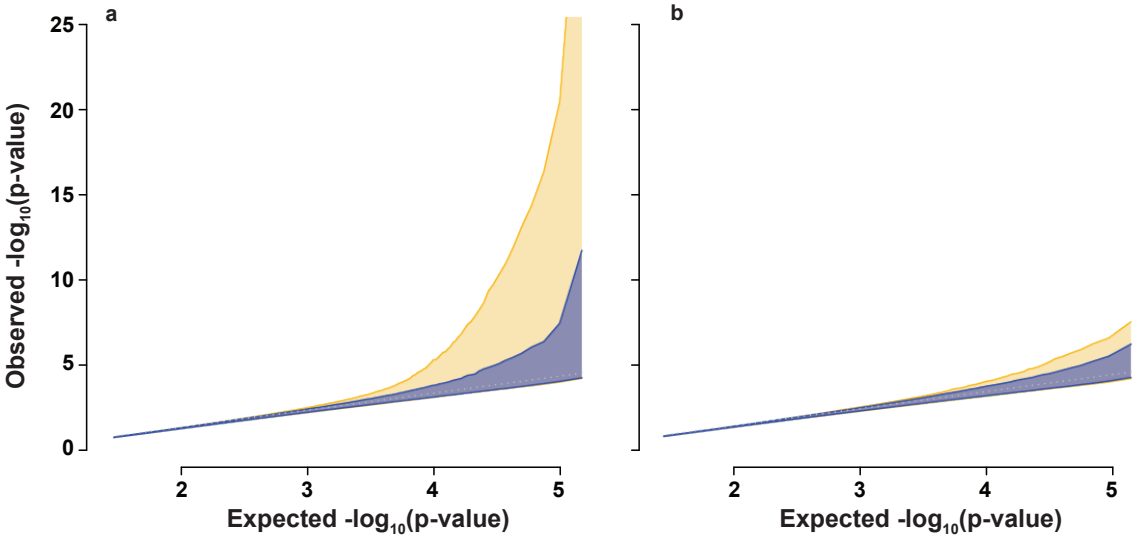
	MTMM (p-value[a])					Marginal models (p-value[a])			
Position	Full model	Interaction	Location	Season	Common	"Spain-Spring"	"Spain-Summer"	"Sweden-Spring"	"Sweden-Summer"
Chromosome 1, around 3.9 Mb									
3978064	2.14×10^{-6}	0.64	0.57	0.61	1.08×10^{-7}	3.61×10^{-7}	1.57×10^{-7}	1.22×10^{-6}	4.58×10^{-7}
Chromosome 1, around 19.5 Mb									
19564595	1.74×10^{-7}	2.07×10^{-6}	0.04	1.56×10^{-4}	4.46×10^{-3}	3.77×10^{-3}	1.02×10^{-4}	0.15	4.52×10^{-4}
19566263	1.74×10^{-7}	2.07×10^{-6}	0.04	1.56×10^{-4}	4.46×10^{-3}	3.77×10^{-3}	1.02×10^{-4}	0.15	4.52×10^{-4}
Chromosome 2, around 9.0 Mb									
9017190	1.33×10^{-6}	3.09×10^{-3}	0.30	6.88×10^{-3}	1.67×10^{-5}	2.39×10^{-4}	1.86×10^{-7}	2.80×10^{-4}	9.08×10^{-6}
Chromosome 5, around 2.0 Mb									
2047036	5.07×10^{-6}	0.42	0.19	0.77	4.22×10^{-7}	2.10×10^{-5}	3.83×10^{-6}	2.99×10^{-7}	2.20×10^{-7}
2050104	3.31×10^{-6}	0.53	0.26	0.73	2.10×10^{-7}	6.51×10^{-6}	2.10×10^{-6}	2.44×10^{-7}	1.04×10^{-7}
Chromosome 5, around 3.2 Mb									
3184569	2.01×10^{-7}	0.18	0.85	0.09	3.34×10^{-8}	3.09×10^{-7}	2.11×10^{-8}	1.21×10^{-6}	4.20×10^{-8}
3185806	3.43×10^{-7}	0.03	0.37	0.04	3.76×10^{-7}	1.94×10^{-6}	6.48×10^{-8}	2.55×10^{-5}	4.68×10^{-7}
3188327	4.77×10^{-11}	0.05	0.14	0.19	1.79×10^{-11}	7.22×10^{-11}	1.66×10^{-12}	2.15×10^{-9}	4.15×10^{-10}
Chromosome 5, around 8.6 Mb									
8642017	9.08×10^{-7}	3.47×10^{-7}	0.03	5.30×10^{-8}	0.29	0.89	0.06	0.71	9.70×10^{-4}
Chromosome 5, around 25.3Mb									
25345918	7.94×10^{-7}	0.31	0.15	0.32	8.04×10^{-8}	3.98×10^{-6}	6.20×10^{-7}	1.84×10^{-7}	4.99×10^{-8}
25346644	7.48×10^{-7}	0.35	0.17	0.40	6.55×10^{-8}	2.88×10^{-6}	5.36×10^{-7}	1.28×10^{-7}	5.60×10^{-8}
25347012	1.65×10^{-6}	0.33	0.15	0.43	1.61×10^{-7}	8.53×10^{-6}	1.18×10^{-6}	2.38×10^{-7}	1.90×10^{-7}
25347575	9.49×10^{-7}	0.33	0.18	0.29	8.99×10^{-8}	3.57×10^{-6}	6.21×10^{-7}	2.85×10^{-7}	5.14×10^{-8}
25348565	7.48×10^{-7}	0.35	0.17	0.40	6.55×10^{-8}	2.88×10^{-6}	5.36×10^{-7}	1.28×10^{-7}	5.60×10^{-8}
25349020	7.48×10^{-7}	0.35	0.17	0.40	6.55×10^{-8}	2.88×10^{-6}	5.36×10^{-7}	1.28×10^{-7}	5.60×10^{-8}
25349890	3.49×10^{-6}	0.24	0.11	0.29	5.04×10^{-7}	2.64×10^{-5}	2.85×10^{-6}	7.48×10^{-7}	2.20×10^{-7}
25351192	7.48×10^{-7}	0.35	0.17	0.40	6.55×10^{-8}	2.88×10^{-6}	5.36×10^{-7}	1.28×10^{-7}	5.60×10^{-8}
25351355	7.48×10^{-7}	0.35	0.17	0.40	6.55×10^{-8}	2.88×10^{-6}	5.36×10^{-7}	1.28×10^{-7}	5.60×10^{-8}
25351688	7.48×10^{-7}	0.35	0.17	0.40	6.55×10^{-8}	2.88×10^{-6}	5.36×10^{-7}	1.28×10^{-7}	5.60×10^{-8}
25352040	7.48×10^{-7}	0.35	0.17	0.40	6.55×10^{-8}	2.88×10^{-6}	5.36×10^{-7}	1.28×10^{-7}	5.60×10^{-8}
25352209	6.74×10^{-7}	0.12	0.10	0.09	1.81×10^{-7}	6.05×10^{-6}	7.66×10^{-7}	7.69×10^{-7}	2.49×10^{-8}
25352490	7.48×10^{-7}	0.35	0.17	0.40	6.55×10^{-8}	2.88×10^{-6}	5.36×10^{-7}	1.28×10^{-7}	5.60×10^{-8}
25352617	7.48×10^{-7}	0.35	0.17	0.40	6.55×10^{-8}	2.88×10^{-6}	5.36×10^{-7}	1.28×10^{-7}	5.60×10^{-8}
25352982	1.63×10^{-6}	0.43	0.21	0.43	1.22×10^{-7}	3.67×10^{-6}	8.65×10^{-7}	2.72×10^{-7}	1.42×10^{-7}
25353765	4.67×10^{-7}	0.31	0.17	0.27	4.49×10^{-8}	2.09×10^{-6}	3.24×10^{-7}	1.42×10^{-7}	2.42×10^{-8}
25353990	2.25×10^{-7}	0.32	0.18	0.27	2.01×10^{-8}	9.33×10^{-7}	1.45×10^{-7}	7.55×10^{-8}	9.89×10^{-9}
25354449	7.48×10^{-7}	0.35	0.17	0.40	6.55×10^{-8}	2.88×10^{-6}	5.36×10^{-7}	1.28×10^{-7}	5.60×10^{-8}
25354528	7.48×10^{-7}	0.35	0.17	0.40	6.55×10^{-8}	2.88×10^{-6}	5.36×10^{-7}	1.28×10^{-7}	5.60×10^{-8}
25359325	5.58×10^{-8}	0.09	0.05	0.13	1.67×10^{-8}	1.82×10^{-6}	2.80×10^{-7}	5.35×10^{-8}	1.90×10^{-9}
25361555	2.03×10^{-6}	0.08	0.03	0.23	8.94×10^{-7}	3.93×10^{-5}	1.58×10^{-5}	1.19×10^{-6}	1.40×10^{-7}
25369739	1.42×10^{-7}	0.16	0.12	0.13	2.47×10^{-8}	2.21×10^{-6}	1.74×10^{-7}	7.88×10^{-8}	4.29×10^{-9}
25369821	1.36×10^{-7}	0.07	0.04	0.09	5.74×10^{-8}	8.28×10^{-6}	6.14×10^{-7}	1.12×10^{-7}	4.97×10^{-9}
25373812	1.51×10^{-7}	0.10	0.09	0.08	4.35×10^{-8}	3.55×10^{-6}	3.15×10^{-7}	1.89×10^{-7}	3.26×10^{-9}
25376120	2.10×10^{-7}	0.06	0.06	0.05	1.05×10^{-7}	1.19×10^{-5}	5.67×10^{-7}	3.50×10^{-7}	5.38×10^{-9}
25379701	4.04×10^{-7}	0.01	0.13	3.60×10^{-3}	1.10×10^{-6}	6.02×10^{-5}	9.72×10^{-7}	1.43×10^{-5}	2.49×10^{-8}
25382600	1.31×10^{-6}	0.07	0.21	0.03	6.23×10^{-7}	5.33×10^{-5}	1.04×10^{-6}	5.57×10^{-6}	5.07×10^{-8}
25386559	3.12×10^{-6}	0.09	0.34	0.03	1.27×10^{-6}	9.08×10^{-5}	8.12×10^{-7}	1.01×10^{-5}	2.16×10^{-7}
25390342	3.17×10^{-7}	0.02	0.24	4.24×10^{-3}	6.28×10^{-7}	2.26×10^{-5}	5.92×10^{-7}	1.26×10^{-5}	2.19×10^{-8}
25390884	2.63×10^{-7}	0.06	0.20	0.02	1.29×10^{-7}	8.50×10^{-6}	1.93×10^{-7}	1.17×10^{-6}	1.11×10^{-8}
25398020	8.01×10^{-7}	0.04	0.01	0.34	6.00×10^{-7}	7.75×10^{-5}	1.13×10^{-5}	2.15×10^{-7}	1.36×10^{-7}

[a] P-values below the Bonferroni-corrected 5% cut-off of 2.3×10^{-7} are highlighted in red.

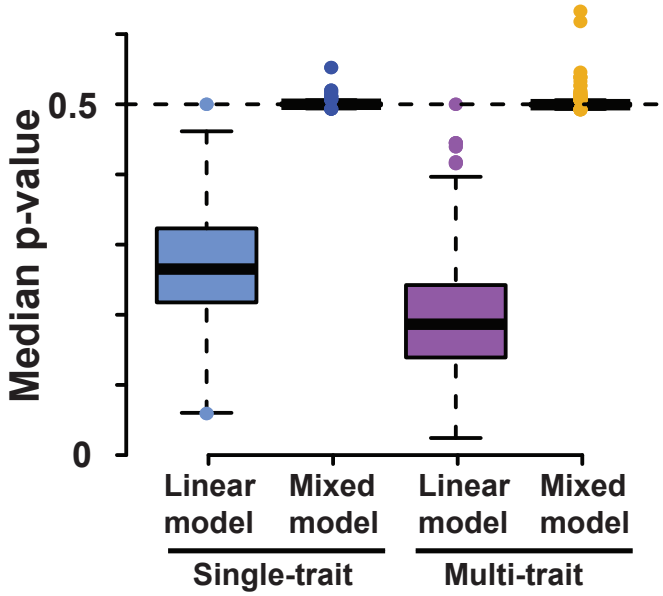
Supplementary Figures



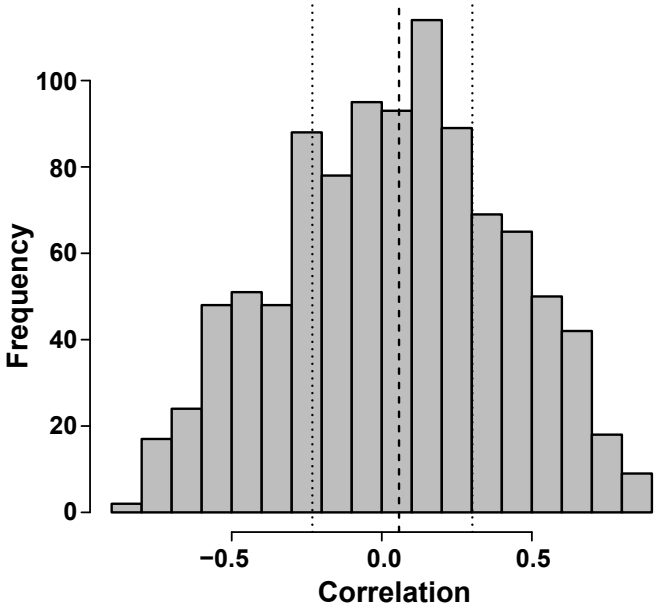
Supplementary Figure 1 The distribution of phenotypic correlations for 2,000 pairs of phenotypes simulated under the 10,000-locus model (mean 0.48, median 0.49).



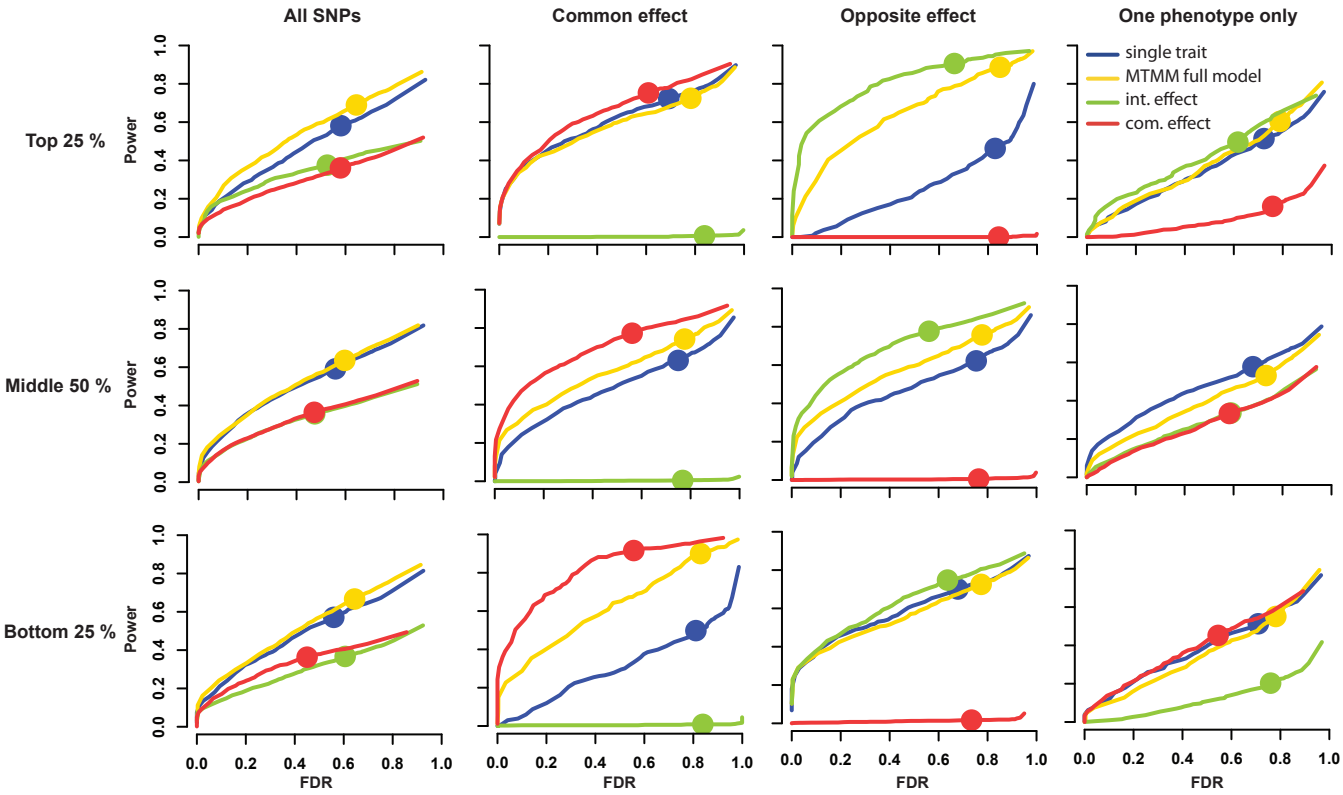
Supplementary Figure 2 Quantile-Quantile plots of 2,000 simulations. Only the lowest 10,000 p-values are considered. Shaded areas indicate the 0.05 and 0.95 point-wise quantile of the ordered p-values across all simulations. Blue areas represent the p-values of the marginal analysis and yellow those for MTMM. Causative markers in a 100 kb window around the simulated causative SNP were either included (**a**) or dropped from the data prior to analysis (**b**).



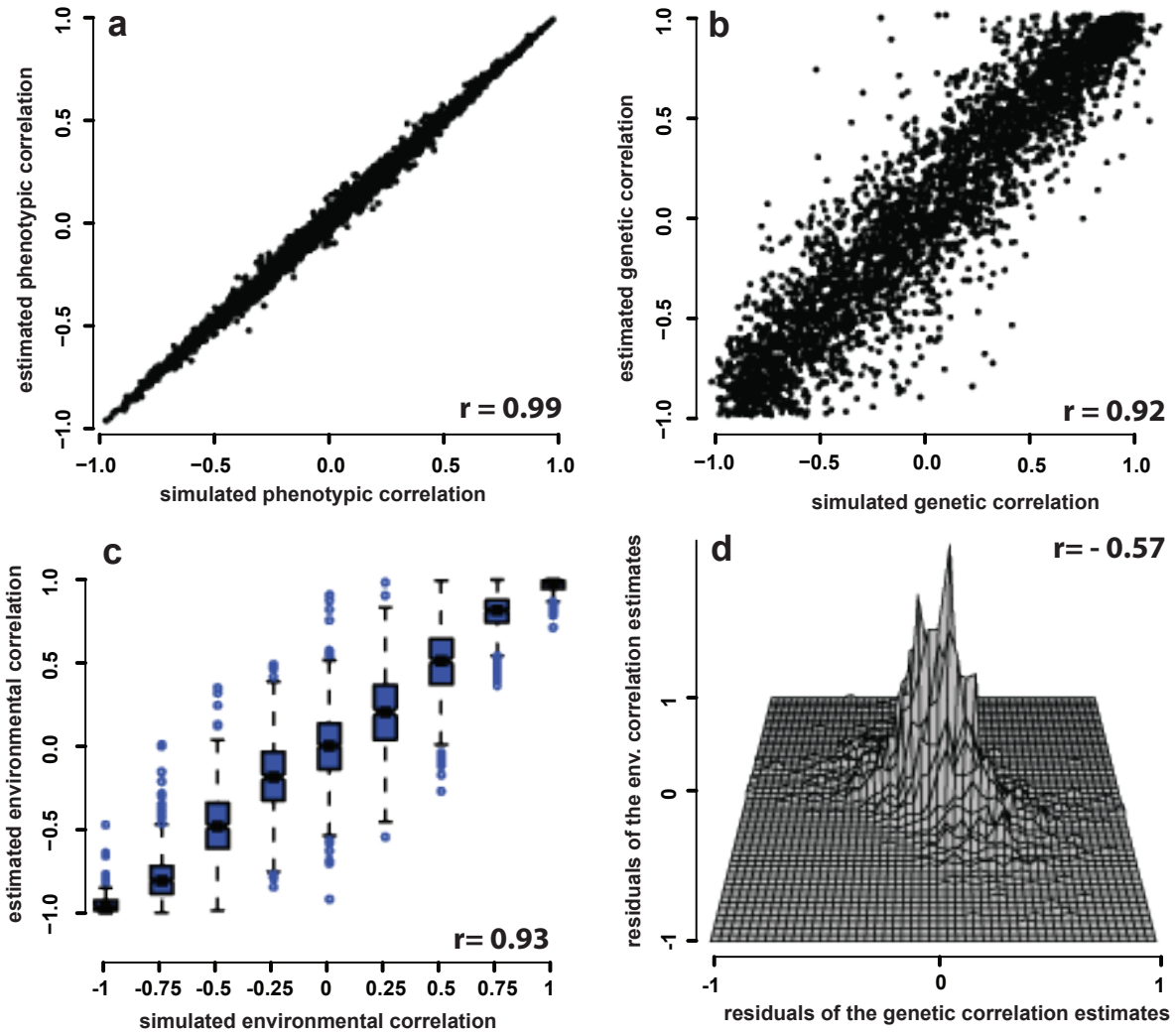
Supplementary Figure 3 Performance of mixed-models compared to a standard linear model that does not correct for genetic background. Boxplot of the distribution of median p-values across 2,000 simulations. The dotted line marks the expected value for uniformly distributed p-values.



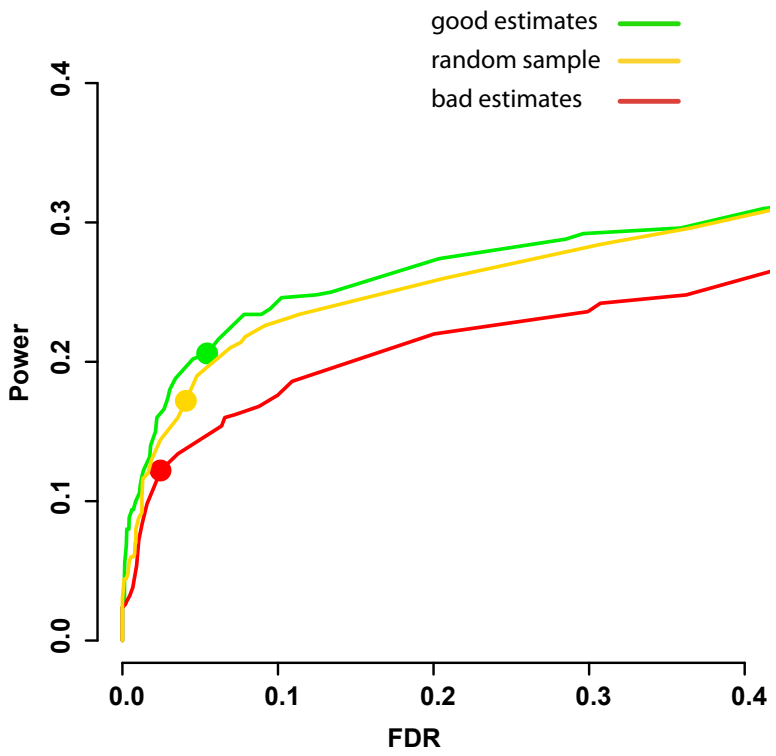
Supplementary Figure 4 The distribution of phenotypic correlations for 1,000 pairs of phenotypes simulated under the 20-locus model. The dashed line marks the mean and the dotted lines the 25% and 75% quantiles, respectively.



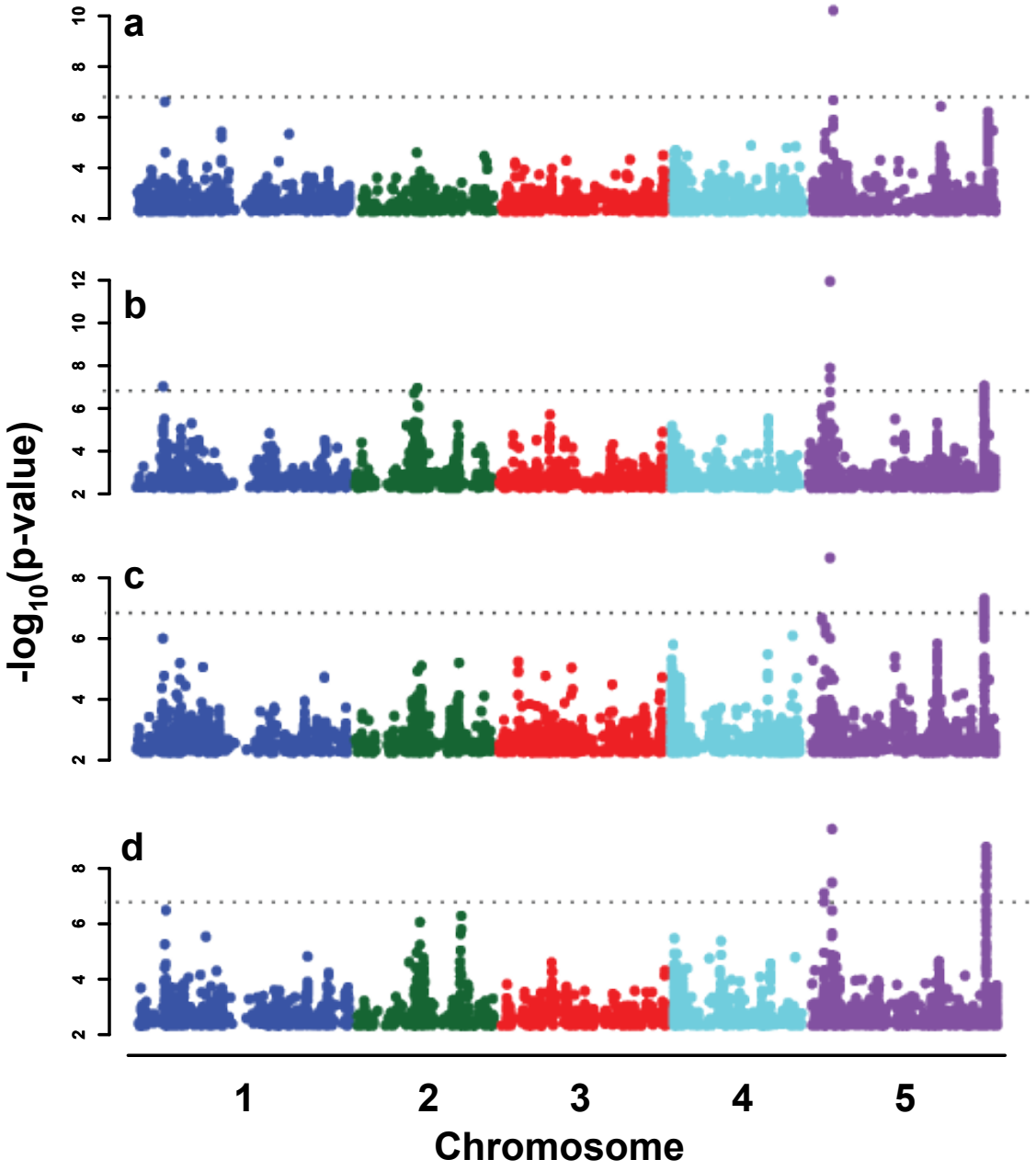
Supplementary Figure 5 Comparison of power and FDR for the 20-locus simulations. The results are divided by phenotypic correlation (*cf.* **Supplementary Fig.4**) top row, top 25%; middle row, middle 50%; bottom row, lowest 25%; as well as by phenotypic effect (from left to right: all SNPs; SNPs with common effect; SNPs with opposite effect; and SNPs with effect in one phenotype only). Dots denote Bonferroni-corrected 5% significance thresholds.



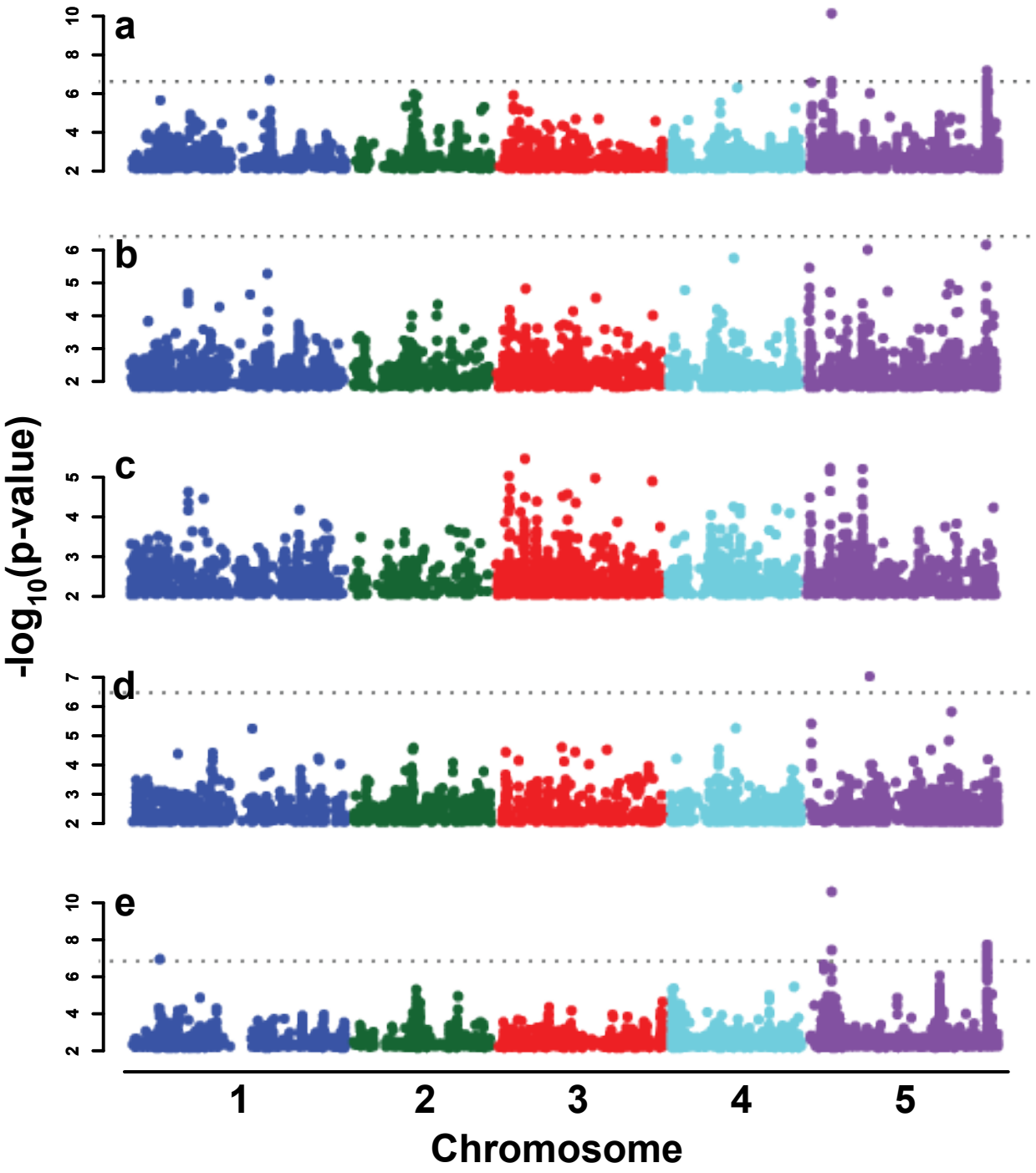
Supplementary Figure 6 Accuracy of the correlation estimates in the MTMM: (**a**) Comparison of the simulated and estimated phenotypic correlation of the two traits. (**b**) Comparison of the simulated and estimated genetic correlation of the two traits. (**c**) Comparison of the simulated and estimated environmental correlation of the two traits. (**d**) Comparison of the residuals of the genetic and environmental correlation estimates.



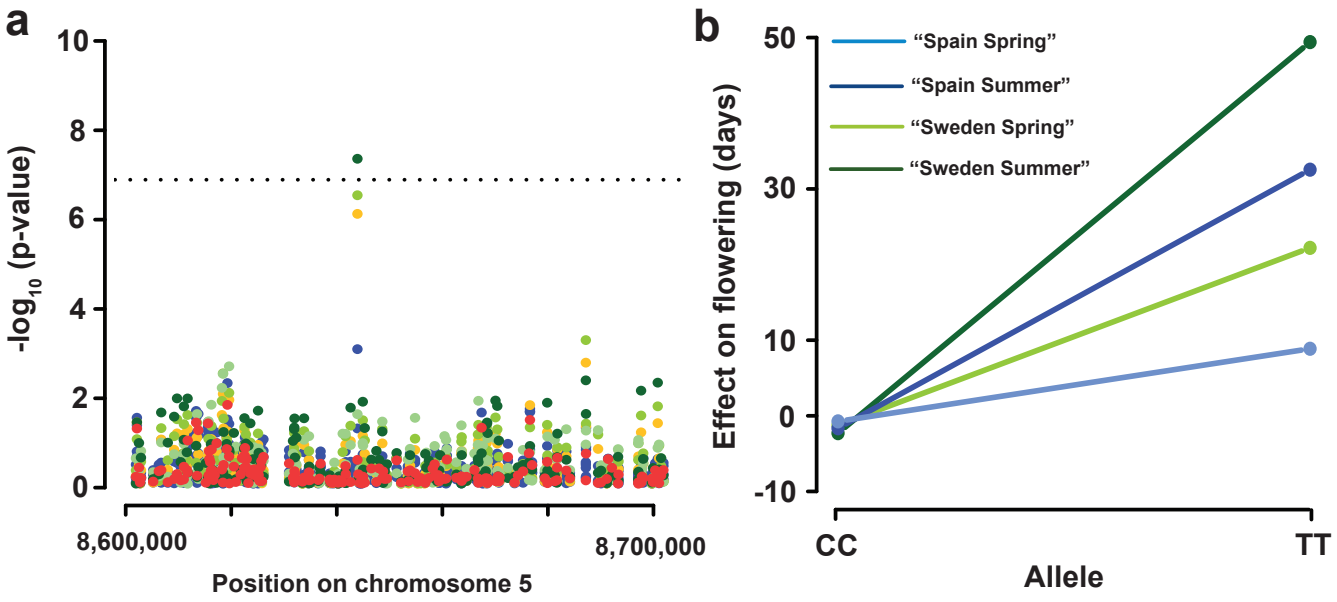
Supplementary Figure 7 Comparison of the power and FDR of the MTMM analysis using simulations with different accuracy of the correlation estimates. The red line represents shows results for the 500 simulations with the least accurate estimates, the yellow line results for 500 randomly chosen simulations, and the green line the results for the 500 simulations with the most accurate estimates. The big dot denotes the genome-wide Bonferroni-corrected 5% significance threshold.



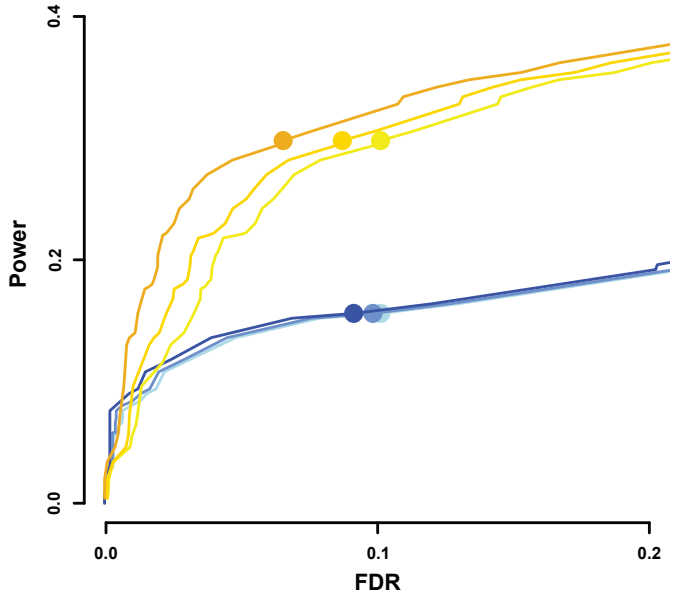
Supplementary Figure 8 Manhattan plots for the four marginal GWAS for the *A. thaliana* flowering data: (**a**) "Spain-Spring", (**b**) "Spain-Summer"; (**c**) "Sweden-Spring", and; (**d**) "Sweden-Summer". The dashed line denotes the 5% Bonferroni-corrected genome-wide significance threshold.



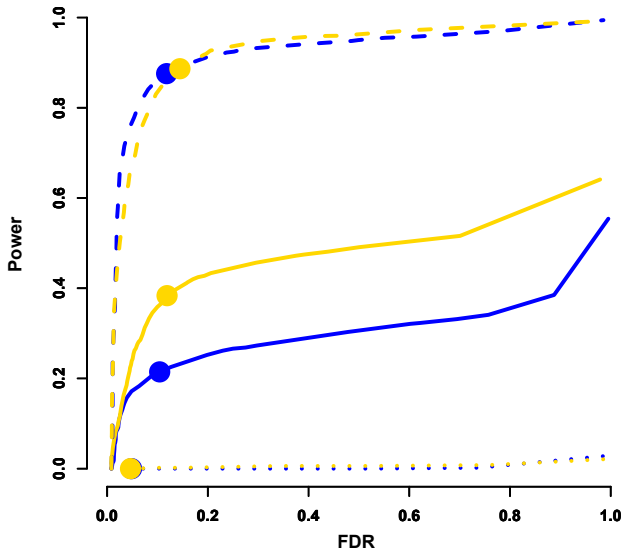
Supplementary Figure 9 Manhattan plot for the five different MTMM GWAS for the *A. thaliana* flowering data: (**a**) full test; (**b**) three-way interaction test; (**c**) genotype-by-location interaction test; (**d**) genotype-by-season interaction test; (**e**) common effect test. The dashed line represents the 5% Bonferroni-corrected genome-wide significance threshold.



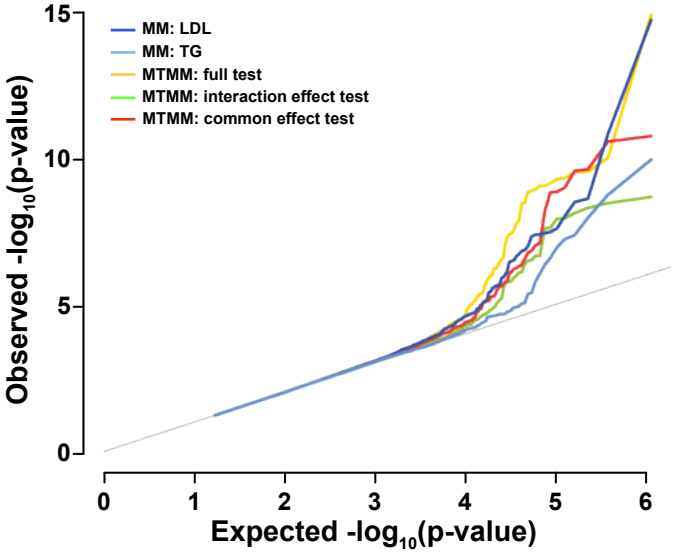
Supplementary Figure 10 A putative interaction effect in the *A. thaliana* data. (**a**) Close-up of Manhattan plot showing results of marginal tests (blue dots) and five different MTMM tests: full model (orange); three-way interaction (light green); genotype-by-location (green), genotype-by-season (dark green), and common effect (red). (**b**) Mean phenotype for the two allelic classes in each environment.



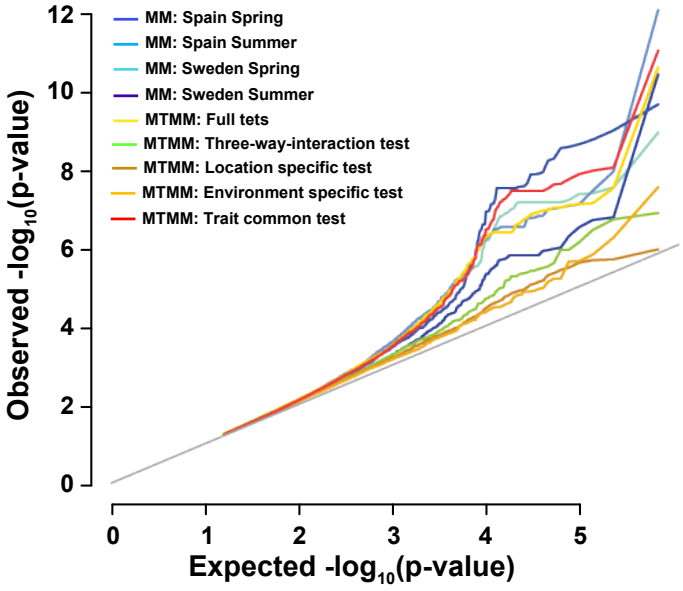
Supplementary Figure 11 Comparison of the power and FDR of marginal mixed-model (blue) and MTMM analyses (yellow) using different window sizes to distinguish true and false positives. The window sizes used were 25, 50 and 100 kb, with lighter colors representing smaller window sizes. The big dot denotes the genome-wide Bonferroni-corrected 5% significance threshold.



Supplementary Figure 12 Comparison of the power and FDR of marginal mixed-model (blue) and MTMM analysis (yellow) using different effect sizes of the simulated causative SNP. In these simulations, the causal locus explains either 0.5% of the overall phenotypic variance (dotted lines), 2% (solid lines) or 10% (dashed lines). The big dot denotes the genome-wide Bonferroni-corrected 5% significance threshold.



Supplementary Figure 13 Quantile-quantile plot for the analyses of the TG and LDL data.



Supplementary Figure 14 Quantile-quantile plot for the *A. thaliana* flowering data.

Supplementary note

Extending the covariance model for multiple traits with different individuals

The covariance model can be generalized for more that two traits and for partially overlapping or disjoint sets of individuals. The covariance for the first phenotype value of the i'th individual and the second phenotype value of the j'th individual is as follows:

$$\mathrm{cov}\left(\mathbf{y}_{1i}, \mathbf{y}_{2j}\right) = \sigma_{g1}\sigma_{g2}\rho_g K_{(ij)} + \sigma_{e1}\sigma_{e2}\rho_e I_{(ij)}, \quad (1)$$

where $K_{(ij)}$ is the kinship between the i'th and the j'th individuals, and similarly $I_{(ij)}$is the value at the i'th column and the j'th row in the identity matrix. Finally, the covariance equations can be generalized for arbitrary number of traits as follows:

$$\mathrm{var}\left(\mathbf{y}_{ki}\right) = \sigma^2_{g_k} K_{(ij)} + \sigma^2_{e_k} I_{(ij)}, \quad (2)$$

$$\mathrm{cov}\left(\mathbf{y}_{ki}, \mathbf{y}_{hj}\right) = \sigma_{g_k}\sigma_{g_h}\rho_{g_{kh}} K_{(ij)} + \sigma_{e_k}\sigma_{e_h}\rho_{e_{kh}} I_{(ij)}, \quad (3)$$

where k and h denote the k'th and h'th trait.

Estimating the variance parameters

The direct way to estimate the variance components is to obtain the maximum likelihood estimates. If we write the covariance matrix as $\mathrm{cov}(Y) = V_{\sigma_g,\sigma_e,\rho_g,\rho_e} = V_{\sigma,\rho}$ where σ_g, σ_e,are vectors containing all genetic and error variance scalars in the model, ρ_g, ρ_e are matrices with all the genetic and error trait correlations in the model, and $\theta = \left(\sigma_g, \sigma_e, \rho_g, \rho_e\right)$. The general likelihood can then be written as

$$ll(\beta, \theta) = C - \frac{1}{2}\left[\log|V_\theta| + (Y - X\beta)V_\theta^{-1}(Y - X\beta)\right], \quad (4)$$

where C is a constant. Similarly, we can write the restricted likelihood as

$$ll_{re}\left(\hat{\beta},\theta\right)=C-\frac{1}{2}\left[\log|V_\theta|+\left(Y-X\hat{\beta}\right)V_\theta^{-1}\left(Y-X\hat{\beta}\right)+\log|X'V_\theta^{-1}X|\right], \quad (5)$$

where $\hat{\beta}$ is the generalized least square (GLS) estimate, and thus intrinsically dependent on the covariance matrix structure[1,2]. In practice, maximizing the likelihood is not trivial for large datasets, especially when the number of variance component parameters in θ is large. An approach used by both ASReml[3] and GCTA[4] is the average information algorithm[5], which makes use of the first and second derivative of the likelihood to rapidly maximize the likelihood using a hill climbing approach.

Picking good starting points can drastically increase the computational efficiency. We propose using the marginal (single trait) genetic and error variance estimates as starting values for the corresponding variances, σ_{gi} and σ_{ei}, in the combined model.

Extending the infinitesimal model for multiple traits

By extending Fisher's classical infinitesimal model[6] for multiple traits, a covariance model for two traits can be derived (Online Methods, equation (3)). Assume the infinitesimal model holds for a pair of traits, and furthermore assume that the traits share all causal markers but with correlated effects (and also assume that different causal markers are independent between traits). Let ρ_g be the correlation of the genetic effects between traits and ρ_e the error correlation, then the between trait covariance can be written out for the identity by state (IBS) kinship matrix as fraction of shared alleles[1]. It can also be derived for an unbiased estimate for the identity by descent (IBD) matrix used by Yang *et al.*[7]. First, let's consider the IBD matrix, and let Z_i be the normalized genotype matrix for the individuals in the i'th trait, then the covariance is as follows:

$$\mathrm{cov}\left(\mathbf{y}_1,\mathbf{y}_2\right)=Z_1'Z_2\sigma_{u1}\sigma_{u2}\rho_g+T_{12}\sigma_{r1}\sigma_{r2}\rho_e, \quad (6)$$

where σ_{ui}^2 is the variance of the effects within the i'th trait, and T_{12} is the $n_1\times n_2$ incidence matrix, i.e. $T_{12(ij)}$ is 1 if i = j and 0 otherwise. Now summing up over all loci we obtain

$$\begin{aligned}\mathrm{cov}\left(\mathbf{y}_1,\mathbf{y}_2\right)&=m\left(\sigma_{u1}\sigma_{u2}\rho_gK_{12}+\sigma_{r1}\sigma_{r2}\rho_eT_{12}\right)\\&=\sigma_{g1}\sigma_{g2}\rho_gK_{12}+\sigma_{e1}\sigma_{e2}\rho_eT_{12},\end{aligned} \quad (7)$$

where m is the number of causal loci, and K_{12} is a $n_1 \times n_2$ matrix where $K_{12(ij)}$ is the kinship between the i'th individual in the first trait $(\mathbf{y}_1)$ and the j'th individual in the second trait $(\mathbf{y}_2)$.

If we instead modify the assumptions, arguing that the correlation between the two traits stems from a fraction of loci with same effect, whereas the effects of remaining loci are not correlated, then we still obtain the model as follows:

$$\text{cov}(\mathbf{y}_1, \mathbf{y}_2) = Z_{s1}{}' Z_{s2} \sigma_{u1} \sigma_{u2} + T_{12} \sigma_{e1} \sigma_{e2} \rho_e = \sigma_{g1} \sigma_{g2} \rho_g K_{12} + \sigma_{e1} \sigma_{e2} \rho_e T_{12}, \quad (8)$$

where Z_{si} represents only the loci which have the same effects in both traits, and ρ_g is now the fraction of common effects among all effects. With analogous arguments we can show that regardless of which extension of the infinitesimal model is used, the same hold for the IBS matrix. These derivations give insight into how to interpret the parameter ρ_g, which Henderson defines as the genetic correlation between traits. It follows from these equations that the phenotypic correlation ρ_{y_1,y_2} (*i.e.*, the correlation of the two traits when measured in the same individual) can be written as follows:

$$\rho_{y1,y2} = \frac{\sigma_{g1} \sigma_{g2} \rho_g + \sigma_{e1} \sigma_{e2} \rho_e}{\sqrt{\left(\sigma_{g1}^2 + \sigma_{e1}^2\right)\left(\sigma_{g2}^2 + \sigma_{e2}^2\right)}}. \quad (9)$$

If there is no environmental correlation, or if there is no overlap between individuals for which the two traits were measured in, then the following equality is obtained:

$$\rho_g = \frac{\rho_{y1,y2} \sqrt{\left(\sigma_{g1}^2 + \sigma_{e1}^2\right)\left(\sigma_{g2}^2 + \sigma_{e2}^2\right)}}{\sigma_{g1} \sigma_{g2}} = \frac{\rho_{y1,y2}}{h_1 h_2}, \quad (10)$$

where h_i^2 is the heritability of the i'th trait. Equations (9) and (10) can also be derived using other models than the infinitesimal model[8].

If multiple environmental variables are of interest, it may not only make sense to modify the fixed effects but also to write the trait correlations as a function of the environmental levels, *e.g.*,

$$\rho_{ij} = \gamma_0 \mathbf{1} + \gamma_1 \Delta c_1 + ... \gamma_p \Delta c_p, \quad (11)$$

where c_k denotes the k'th environmental variable and Δc_k denotes the difference in the k'th environment between the i'th and the j'th trait. This way it is possible to reduce the number of parameters in the model.

Analysis of flowering time in a factorial setting

Extending the MTMM for the analysis of flowering time in plants grown in four different environments, consisting of a factorial setting with two simulated seasons ("Spring" and "Summer") and two simulated locations ("Spain" and "Sweden"). Stacking the four phenotype vectors together we can write the model as follows:

$$\mathbf{y} = \begin{bmatrix} \mathbf{y}_1 \\ \mathbf{y}_2 \\ \mathbf{y}_3 \\ \mathbf{y}_4 \end{bmatrix} = \sum_{i=1}^{4} \mathbf{s}_i \mu_i + \mathbf{x}\beta + (\mathbf{x} \times \mathbf{l})\alpha_1 + (\mathbf{x} \times \mathbf{f})\alpha_2 + \boldsymbol{\upsilon} \qquad (12)$$

,

where $\mathbf{x}$ is the vector of SNPs and $\mathbf{s}_i$ is a vector with 1 for all values belonging to the i'th trait and 0 otherwise. $\mathbf{l}$ is a vector with 1 for all the values measured in the same location, $\mathbf{f}$ is a vector with 1 for all the values measured in the same season, and $\boldsymbol{\upsilon} \sim N(0, \mathrm{cov}(\mathbf{y}))$ is a random variable capturing both the error and genetic random effects. We applied the following five F-tests genome-wide:

- The full model tested against a null model where $\beta = 0$ and $\alpha_1 = 0$ and $\alpha_2 = 0$. This identifies both loci with common and differing effects in one model, but suffers in power from the extra degree of freedom.

- The common genetic effect model test the genetic model $\alpha_1 = 0$ and $\alpha_2 = 0$ against a null model where $\beta = 0$ and $\alpha_1 = 0$ and $\alpha_2 = 0$.

- To identify gene × location interactions, we propose to test the full model against a null model where $\alpha_1 = 0$.

- Likewise, to identify gene × season interactions, we propose to test the full model against a null model where $\alpha_2 = 0$.

- Finally, to test for any $G \times E$ interaction, a 'three-way-interaction' test, where we test the full model against a null model where $\alpha_1 = 0$ and $\alpha_2 = 0$ is performed.

Estimating the error of correlation estimates

We extracted the sampling covariances of the random effects using[9] and then applied the delta method using[10] to estimate the standard error (SE) for the correlation estimates. We used a likelihood ratio test to calculate p-values for the correlation estimates being different from zero.

Supplementary References

1. Kang, H. M., Zaitlen, N. A., Wade, C. M., Kirby, A., Heckerman,D., Daly, M. J., and Eskin, E. Efficient control of population structure in model organism association mapping. Genetics, 178:1709–23, 2008.

2. Harville, D. Bayesian inference for variance components using only error contrasts. Biometrika, 61:383–385, 1974.

3. Gilmour, A., Gogel, B., Cullis, B., Welham, S. J., and Thompson,R. ASReml User Guide Release 1.0, 2002.

4. Yang, J., Lee, S. H., Goddard, M. E., and Visscher, P. M. GCTA:A tool for genome-wide complex trait analysis. Am J Hum Genet,88:76–82, 2011.

5. Gilmour, A., Thompson, R., and Cullis, B. Average Information REML: An efficient algorithm for variance parameter estimation in linear mixed models. Biometrics, 51:1440–1450, 1995.

6. Fisher, R. The correlation between relatives on the supposition of Mendelian inheritance. Transactions of the Royal Society ofEdinburgh, 52:399–433, 1918.

7. Yang, J., Benyamin, B., McEvoy, B. P., Gordon, S., Henders,A. K., Nyholt, D. R., Madden, P. A., Heath, A. C., Martin, N. G.,Montgomery, G. W., Goddard, M. E., and Visscher, P. M. Common snps explain a large proportion of the heritability for human height. NatureGenet., 42:565–9, 2010.

8. Lynch, M. and Walsh, B. Genetics and Analysis of Quantitative Traits. Sinauer Associates, Inc., Sunderland, Massachusetts,1997.

9. Wolak, M. nadiv: Functions to construct non-additive genetic relatedness matrices, 2012. R package version 2.8.

10. Jackson, C. H. Multi-state models for panel data: The msm package for R. Journal of Statistical Software, 38:1–29, 2011.