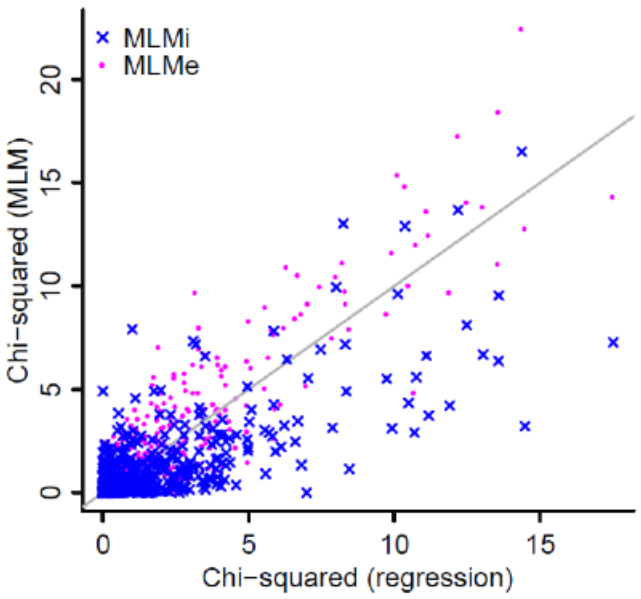
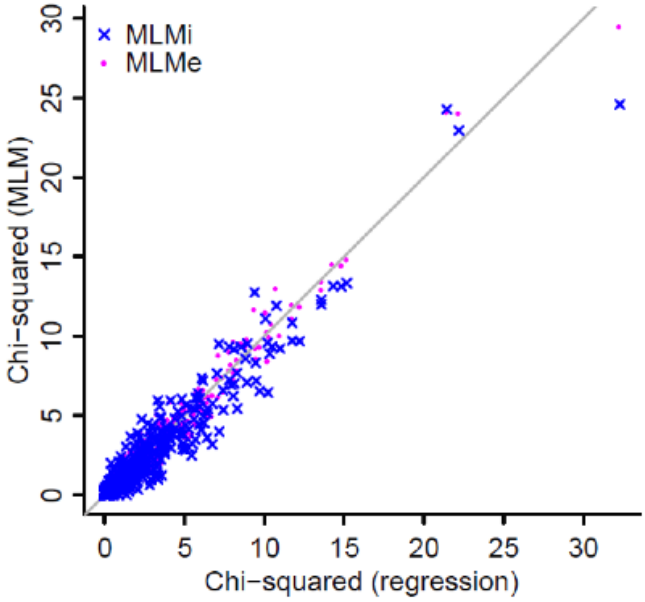
Supplementary Figures

Supplementary Figure 1. MLMe increases power but MLMi reduces power vs. linear regression. We plot χ^2 association statistics for MLMi and MLMe vs. linear regression for a single simulation, for various values of *N* and *M*. Plotted are chi-squared test-statistics at the 500 candidate markers (see Methods section for the simulation design).

(a) *N*=1,000, *M*=1,000



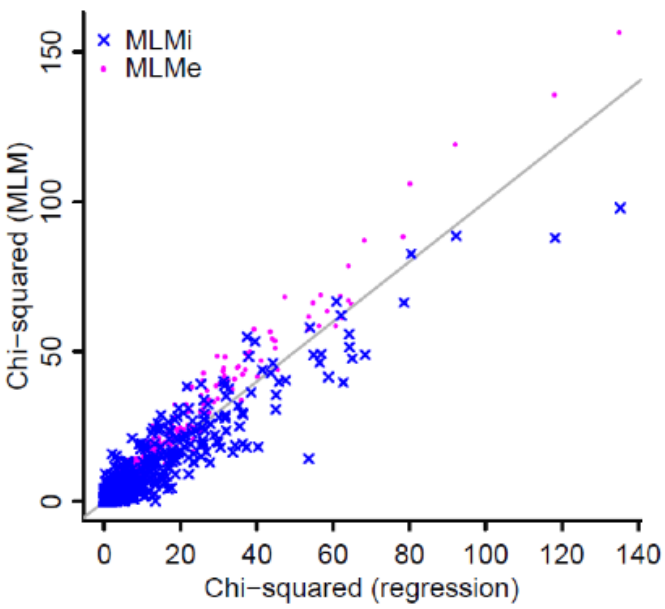
(b) *N*=1,000, *M*=10,000

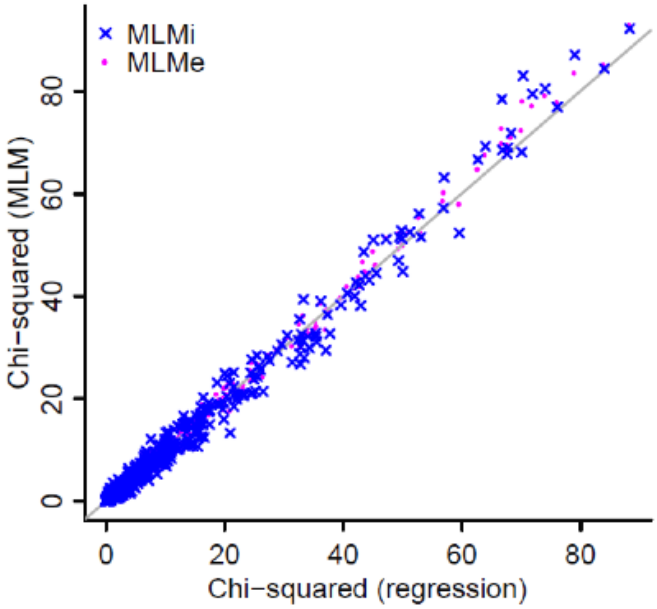


(c) *N*=10,000, *M*=10,000

(d) *N*=10,000, *M*=100,000

Nature Genetics: doi:10.1038/ng.2876





Nature Genetics: doi:10.1038/ng.2876

Supplementary Tables

Supplementary Table 1. Running time and memory usage of GCTA. For each value of *N* and *M*, we report median running time and memory usage across 10 simulations for GCTA with all markers included in the GRM (GCTA-MLMi) and GCTA with leave-one-chromosome-out analysis (as described in the Discussion section; GCTA-LOCO).

# samples (*N*)	# markers (*M*)	GCTA-MLMi	GCTA-LOCO
5,000	50,000	0.3hr / 2.0GB	1.0hr / 4.1GB
5,000	100,000	0.6hr / 3.0GB	1.4hr / 5.2GB
10,000	50,000	1.3hr / 5.9GB	7.2hr / 14.3GB
10,000	100,000	2.5hr / 7.9GB	8.2hr / 16.3GB

Nature Genetics: doi:10.1038/ng.2876

Supplementary Table 2. Expected mean of test statistics for causal, null and all markers. We list these derivations for linear regression, MLMi, MLMe, and also list the ratios of these means for MLMe vs. MLMi.

	Linear regression	MLMi	MLMe	MLMe / MLMi
Causal markers (M_q)	$1+Nh_g^2/M_q$	$\frac{Nh_g^2/M_q+1-r^2h_g^2}{Nh_g^2/M+1-r^2h_g^2}$	$1+\frac{Nh_g^2/M_q}{1-r^2h_g^2}$	$1+\frac{Nh_g^2/M}{1-r^2h_g^2}$
Null markers (M–M_q)	1	$\frac{1-r^2h_g^2}{Nh_g^2/M+1-r^2h_g^2}$	1	$1+\frac{Nh_g^2/M}{1-r^2h_g^2}$
All markers (M)	$1+Nh_g^2/M$	1	$1+\frac{Nh_g^2/M}{1-r^2h_g^2}$	$1+\frac{Nh_g^2/M}{1-r^2h_g^2}$

M is the number of independent markers, M_q of which explain h_g^2 of phenotypic variance.

$r^2=[(1+\theta)-\sqrt{(1+\theta)^2-4h_g^2\theta}]/2h_g^2$ with $\theta=Nh_g^2/M$. If $M>N$, $r^2\approx Nh_g^2/M$.

Nature Genetics: doi:10.1038/ng.2876

Supplementary Table 3. MLMe increases power but MLMi decreases power vs. linear regression. (a) There are 500 candidate causal markers that explain half of the heritability, and M non-candidate markers including 0.05*M causal markers that explain the other half of heritability and 0.95*M null markers. Thus, there are M + 500 markers in total. We report average χ^2 association statistics (± standard errors) at these markers for linear regression, MLMi and MLMe, averaged across 100 simulations. In the MLMe analysis, for ease of computation, we excluded the 500 candidate markers from the GRM and tested each of the 500 candidate markers based on the GRM computed from the M non-candidate markers (test statistics at the M non-candidate markers were not calculated). For consistency, we also report statistics for the 500 candidate causal markers for linear regression and MLMe. The first and second table are based on simulations with heritability (h_g^2) = 100% and 50%, respectively. (b) We report average χ^2 association statistics (± standard errors) at 200 causal markers (h_g^2 = 50%) on chromosomes 1 and 2 (100 on each chromosome), 39,668 null markers on chromosome 3 and all 133,036 markers on chromosomes 1-3 for ref. [1] data with simulated phenotypes. For both (a) and (b), expected results based on theoretical derivations are given in parentheses (see footnotes and Table S1). (c) We report power to detect significant associations at different P-value thresholds for results from (a) (which includes all results from Table 2a). (d) We report power to detect significant associations at different P-value thresholds for results from (b) (which includes all results from Table 2b).

Nature Genetics: doi:10.1038/ng.2876

(a)

*50% variance explained by 500 candidate causal markers and 50% by 0.05*M non-candidate causal markers:*

# samples (*N*)	# markers (*M*)	markers tested	Linear regression (Expected value [$])	MLMi (Expected value [$$])	MLMe [&] (Expected value [$$$])	Expected MLMe / MLMi
1,000	1,000	500 causal	1.982 ± 0.010 (2.00)	1.322 ± 0.004 (1.33*)	2.173 ± 0.010 (2.24)	1.68
		950 null	1.003 ± 0.003 (1.00)	0.327 ± 0.002 (0.33)	n/a	n/a
		1,500 all	1.692 ± 0.004 (1.67)	0.997 ± 0.001 (1.00)	n/a	n/a
1,000	10,000	500 causal	1.962 ± 0.007 (2.00)	1.856 ± 0.007 (1.90*)	1.971 ± 0.007 (2.02)	1.06
		9,500 null	1.000 ± 0.001 (1.00)	0.903 ± 0.001 (0.90)	n/a	n/a
		10,500 all	1.095 ± 0.001 (1.095)	0.997 ± 0.001 (1.00)	n/a	n/a
10,000	10,000	500 causal	10.929 ± 0.027 (11.00)	9.809 ± 0.006 (10.05*)	13.268 ± 0.029 (13.36)	1.33
		9,500 null	1.005 ± 0.001 (1.00)	0.072 ± 0.0003 (0.05)	n/a	n/a
		10,500 all	1.957 ± 0.002 (1.95)	1.0001 ± 0.00001 (1.00)	n/a	n/a
10,000	100,000	500 causal	11.195 ± 0.026 (11.00)	10.973 ± 0.022 (10.09*)	11.40 ± 0.025 (11.25)	1.03
		95,000 null	1.002 ± 0.0004 (1.0)	0.897 ± 0.0003 (0.90)	n/a	n/a
		100,500 all	1.102 ± 0.0004 (1.10)	0.997 ± 0.0003 (1.00)	n/a	n/a

[*] For MLMi, the h_g^2 of markers included in the GRM is 100%, and the derivation is much less accurate. However, the derivation is much more accurate at lower values of h_g^2 (see below).

[$] Expected mean test statistic is $1+0.5Nh_g^2/500$ for causal markers, 1 for null markers, and $1+Nh_g^2/(M+500)$ for all markers.

[$$] Expected mean test statistic is $[0.5Nh_g^2/500+(1-r^2h_g^2)]/[Nh_g^2/(M+500)+(1-r^2h_g^2)]$ for causal markers, $(1-r^2h_g^2)/[Nh_g^2/(M+500)+(1-r^2h_g^2)]$ for null markers, and 1 for all markers, where $r^2=[(1+\theta)-\sqrt{(1+\theta)^2-4h_g^2\theta}]/2h_g^2$ with $\theta=Nh_g^2/(M+500)$.

[$$$] Expected mean test statistic is $1+(0.5Nh_g^2/500)/[1-0.5r^2h_g^2)]$ for causal markers, where $r^2=[(1+\theta)-\sqrt{(1+\theta)^2-2h_g^2\theta}]/h_g^2$ with $\theta=0.5Nh_g^2/M$.

[&] If we had included all non-candidate markers in the GRM in the MLMe analysis of each candidate marker (instead of excluding all 500 candidate markers from the GRM for ease of computation), the mean test statistics at the 500 candidate markers would be even higher, since the GRM would capture more heritability. For example, for *M* = 1000 and *N*=1000, the expected mean test statistics at the 500 candidate markers would increase from 2.24 to 4.00.

Nature Genetics: doi:10.1038/ng.2876

*25% variance explained by 500 candidate causal markers and 25% by 0.05*M non-candidate causal markers:*

# samples (*N*)	# markers (*M*)	markers tested	Linear regression (Expected value)	MLMi (Expected value)	MLMe (Expected value)	Expected MLMe / MLMi
1,000	1,000	500 causal	1.462 ± 0.007 (1.50)	1.124 ± 0.005 (1.14)	1.485 ± 0.007 (1.53)	1.34
		950 null	0.989 ± 0.004 (1.00)	0.723 ± 0.002 (0.72)	n/a	n/a
		1,500 all	1.322 ± 0.004 (1.33)	1.000 ± 0.0004 (1.00)	n/a	n/a
1,000	10,000	500 causal	1.541 ± 0.007 (1.50)	1.470 ± 0.007 (1.44)	1.542 ± 0.007 (1.50)	1.04
		9,500 null	1.006 ± 0.001 (1.00)	0.951 ± 0.001 (0.95)	n/a	n/a
		10,500 all	1.057 ± 0.001 (1.05)	1.000 ± 0.00005 (1.00)	n/a	n/a
10,000	10,000	500 causal	5.977 ± 0.020 (6.00)	4.484 ± 0.012 (4.50)	6.246 ± 0.021 (6.28)	1.39
		9,500 null	0.992 ± 0.001 (1.00)	0.631 ± 0.001 (0.63)	n/a	n/a
		10,500 all	1.470 ± 0.002 (1.48)	1.000 ± 0.00002 (1.00)	n/a	n/a
10,000	100,000	500 causal	6.044 ± 0.018 (6.00)	5.842 ± 0.018 (5.83)	6.062 ± 0.019 (6.03)	1.03
		95,000 null	1.003 ± 0.0004 (1.00)	0.952 ± 0.0001 (0.95)	n/a	n/a
		100,500 all	1.053 ± 0.0004 (1.05)	1.000 ± 0.000004 (1.00)	n/a	n/a

Nature Genetics: doi:10.1038/ng.2876

(b)

# samples (N)	# markers (M)	markers tested	Linear regression (Expected value [$])	MLMi (Expected value [$$])	MLMe [&] (Expected value [$$$])	Expected MLMe / MLMi
10,000	133,036	200 causal	26.987 ± 0.095 (26.00)	21.436 ± 0.053 (19.85*)	28.417±0.102 (27.94*)	1.46
		39,668 null	1.001 ± 0.0002 (1.00)	0.743 ± 0.001 (0.62*)	1.001 ± 0.002 (1.00)	1.61
		133,036 all	1.488 ± 0.005 (1.49*)	1.005 ± 0.001 (1.00)	1.515 ± 0.006 (1.53*)	1.53

* These derivations depend on the value of the effective number of unlinked markers (M_{eff}), which we estimated at M_{eff1} = 3,525, M_{eff2} = 3,666 and M_{eff3} = 3,055 for chromosomes 1, 2 and 3, respectively (based on mean χ^2 test statistics from linear regression analyses; see Supp Note) for a total M_{eff} = 10,246.

& The test statistics were calculated via leave-one-chromosome-out (LOCO) analysis.

$ Expected mean test statistic is $1+Nh_{\text{g}}^2/200$ for causal markers, 1 for null markers, and $1+Nh_{\text{g}}^2/M_{\text{eff}}$ for all markers.

$$ Expected mean test statistic is $[Nh_{\text{g}}^2/200+(1-r^2h_{\text{g}}^2)]/[Nh_{\text{g}}^2/M_{\text{eff}}+(1-r^2h_{\text{g}}^2)]$ for causal markers, $(1-r^2h_{\text{g}}^2)/[Nh_{\text{g}}^2/M_{\text{eff}}+(1-r^2h_{\text{g}}^2)]$ for null markers, and 1 for all markers, where $r^2=[(1+\theta)-\sqrt{(1+\theta)^2-4h_{\text{g}}^2\theta}]/2h_{\text{g}}^2$ with $\theta=Nh_{\text{g}}^2/M_{\text{eff}}$.

$$$ Expected mean test statistic is $1+(0.5Nh_{\text{g}}^2/100)/[1-0.5r^2h_{\text{g}}^2)]$ for the 100 causal markers on chromosome 1 or 2, where $r^2=[(1+\theta)-\sqrt{(1+\theta)^2-2h_{\text{g}}^2\theta}]/h_{\text{g}}^2$ with $\theta=0.5Nh_{\text{g}}^2/M_{\text{eff1}}$ for chromosome 1 and $\theta=0.5Nh_{\text{g}}^2/M_{\text{eff2}}$ for chromosome 2. The expected mean test statistic is $1+(0.5Nh_{\text{g}}^2/M_{\text{eff1}})/[1-0.5r^2h_{\text{g}}^2)]$ with $\theta=0.5Nh_{\text{g}}^2/(M_{\text{eff2}}+M_{\text{eff3}})$ for all markers on chromosome 1, $1+(0.5Nh_{\text{g}}^2/M_{\text{eff2}})/[1-0.5r^2h_{\text{g}}^2)]$ with $\theta=0.5Nh_{\text{g}}^2/(M_{\text{eff1}}+M_{\text{eff3}})$ for all markers on chromosome 2, and 1 for all markers on chromosome 3.

Nature Genetics: doi:10.1038/ng.2876

(c)

*50% variance explained by 500 candidate causal markers and 50% by 0.05*M non-candidate causal markers:*

Method	*N*	*M*	Power			
			$P < 0.05$	$P < 0.001$	$P < 1 \times 10^{-6}$	$P < 5 \times 10^{-8}$
Linear regression	1000	1000	0.161	0.022	0.001	0.000
Linear regression	1000	10000	0.158	0.020	0.001	0.000
Linear regression	10000	10000	0.514	0.302	0.140	0.104
Linear regression	10000	100000	0.526	0.308	0.145	0.106
MLMi	1000	1000	0.092	0.006	0.000	0.000
MLMi	1000	10000	0.148	0.018	0.001	0.000
MLMi	10000	10000	0.488	0.278	0.124	0.089
MLMi	10000	100000	0.519	0.304	0.142	0.104
MLMe	1000	1000	0.178	0.029	0.001	0.000
MLMe	1000	10000	0.161	0.021	0.001	0.000
MLMe	10000	10000	0.548	0.342	0.177	0.137
MLMe	10000	100000	0.528	0.313	0.149	0.110

*25% variance explained by 500 candidate causal markers and 25% by 0.05*M non-candidate causal markers:*

Method	*N*	*M*	Power			
			$P < 0.05$	$P < 0.001$	$P < 1 \times 10^{-6}$	$P < 5 \times 10^{-8}$
Linear regression	1000	1000	0.105	0.007	0.000	0.000
Linear regression	1000	10000	0.115	0.008	0.000	0.000
Linear regression	10000	10000	0.399	0.176	0.050	0.031
Linear regression	10000	100000	0.403	0.179	0.052	0.032
MLMi	1000	1000	0.066	0.003	0.000	0.000
MLMi	1000	10000	0.106	0.007	0.000	0.000
MLMi	10000	10000	0.335	0.123	0.026	0.014
MLMi	10000	100000	0.395	0.173	0.048	0.029
MLMe	1000	1000	0.108	0.008	0.000	0.000
MLMe	1000	10000	0.115	0.008	0.000	0.000
MLMe	10000	10000	0.409	0.184	0.055	0.034
MLMe	10000	100000	0.404	0.180	0.052	0.032

(d)

Method	*N*	*M*	Power			
			$P < 0.05$	$P < 0.001$	$P < 1 \times 10^{-6}$	$P < 5 \times 10^{-8}$
Linear regression	10000	133036	0.673	0.493	0.320	0.271
MLMi	10000	133036	0.646	0.450	0.271	0.225
MLMe	10000	133036	0.683	0.505	0.332	0.284

Nature Genetics: doi:10.1038/ng.2876

Supplementary Table 4. Effectiveness of MLM using M_R random or M_T top associated markers in correcting for stratification. We report the average λ_{median} (± standard error) and proportion of markers with $P < 0.05$ or $P < 0.001$, averaged across 100 simulations (with N=10,000, M=100,000) with two subpopulations with a mean trait difference of 0.25 standard deviations and either (a) F_{ST}=0.005 or (b) F_{ST}=0.0025 between subpopulations.

(a) F_{ST}=0.005

Using M_R random markers:

#markers (M_R)	λ_{median} (± SE)	Power	
		$P < 0.05$	$P < 0.001$
0	1.78 ± 0.02	0.1393	0.0135
100	1.57 ± 0.02	0.1145	0.0091
300	1.40 ± 0.02	0.0945	0.0056
1,000	1.22 ± 0.01	0.0750	0.0028
3,000	1.13 ± 0.01	0.0639	0.0017
10,000	1.07 ± 0.01	0.0563	0.0013
30,000	1.04 ± 0.01	0.0532	0.0011
100,000	1.02 ± 0.01	0.0512	0.0010

Using M_T top associated markers:

#markers (M_T)	λ_{median} (± SE)	Power	
		$P < 0.05$	$P < 0.001$
0	1.78 ± 0.02	0.1393	0.0135
100	1.15 ± 0.01	0.0674	0.0019
300	1.36 ± 0.02	0.0918	0.0048
1,000	1.51 ± 0.02	0.1066	0.0076
3,000	1.49 ± 0.02	0.1043	0.0069
10,000	1.32 ± 0.01	0.0872	0.0044
30,000	1.15 ± 0.01	0.0661	0.0020
100,000	1.02 ± 0.01	0.0512	0.0010

Nature Genetics: doi:10.1038/ng.2876

(b) F_{ST}=0.0025

Using M_R random markers:

#markers (M_R)	λ_{median} (± SE)	Power	
		$P < 0.05$	$P < 0.001$
0	1.41 ± 0.02	0.0977	0.0053
100	1.36 ± 0.02	0.0919	0.0046
300	1.30 ± 0.02	0.0856	0.0039
1,000	1.21 ± 0.01	0.0740	0.0028
3,000	1.14 ± 0.01	0.0649	0.0020
10,000	1.09 ± 0.01	0.0584	0.0016
30,000	1.05 ± 0.01	0.0546	0.0013
100,000	1.03 ± 0.01	0.0520	0.0012

Using M_T top associated markers:

#markers (M_T)	λ_{median} (± SE)	Power	
		$P < 0.05$	$P < 0.001$
0	1.41 ± 0.02	0.0977	0.0053
100	1.07 ± 0.01	0.0567	0.0015
300	1.29 ± 0.01	0.0861	0.0038
1,000	1.51 ± 0.02	0.1120	0.0075
3,000	1.51 ± 0.02	0.1111	0.0080
10,000	1.33 ± 0.01	0.0868	0.0042
30,000	1.13 ± 0.01	0.0646	0.0020
100,000	1.03 ± 0.01	0.0520	0.0012

Nature Genetics: doi:10.1038/ng.2876

Supplementary Table 5. Effectiveness of MLM using M_T top associated markers in increasing study power, for various values of M and N. We report the average $-\log_{10}P$-values (± standard error) and power to detect significant associations at various P-value thresholds, averaged across 100 simulations with fraction p=0.05 or p=0.005 of causal markers, for various values of M and N. In each column, the maximum value is denoted in bold font.

(a) M=100,000, N=10,000, p=0.05 [MLM using causal markers attained a $-\log_{10}P$-value of 3.85]

#markers (M_T)	$-\log_{10}P$-value	Power			
		$P < 0.05$	$P < 0.001$	$P < 10^{-6}$	$P < 5 \times 10^{-8}$
0	2.89 ± 0.01	0.5549	0.3212	0.1403	0.1003
100	2.81 ± 0.01	0.5469	0.3131	0.1347	0.0955
300	2.64 ± 0.01	0.5330	0.2955	0.1199	0.0842
1,000	2.32 ± 0.01	0.5015	0.2595	0.0949	0.0615
3,000	1.97 ± 0.01	0.4610	0.2145	0.0659	0.0399
10,000	1.68 ± 0.01	0.4166	0.1724	0.0425	0.0241
30,000	1.66 ± 0.01	0.4140	0.1713	0.0414	0.0235
100,000 ($M_T = M$)	**2.94 ± 0.01**	**0.5585**	**0.3263**	**0.1443**	**0.1038**

(b) M=100,000, N=10,000, p=0.005 [MLM using causal markers attained a $-\log_{10}P$-value of 4.95]

#markers (M_T)	$-\log_{10}P$-value	Power			
		$P < 0.05$	$P < 0.001$	$P < 10^{-6}$	$P < 5 \times 10^{-8}$
0	2.90 ± 0.01	0.5551	0.3217	0.1425	0.1011
100	3.84 ± 0.01	0.6133	0.3954	0.2105	0.1621
300	**3.94 ± 0.01**	**0.6181**	**0.4031**	**0.2160**	**0.1679**
1,000	3.36 ± 0.01	0.5859	0.3603	0.1747	0.1306
3,000	2.58 ± 0.01	0.5278	0.2901	0.1150	0.0789
10,000	1.91 ± 0.01	0.4536	0.2082	0.0602	0.0360
30,000	1.72 ± 0.01	0.4228	0.1793	0.0454	0.0251
100,000 ($M_T = M$)	2.96 ± 0.01	0.5597	0.3271	`0.1472	0.1050

Nature Genetics: doi:10.1038/ng.2876

(c) M=100,000, N=1,000, p=0.05 [MLM using causal markers attained a $-\log_{10}P$-value of 0.734]

#markers (M_T)	$-\log_{10}P$-value	Power			
		$P < 0.05$	$P < 0.001$	$P < 10^{-6}$	$P < 5 \times 10^{-8}$
0	0.721 ± 0.003	0.1685	0.0205	0.0006	0.0002
100	0.582 ± 0.003	0.1101	0.0078	0.0001	0.0000
300	0.531 ± 0.002	0.0885	0.0045	0.0000	0.0000
1,000	0.494 ± 0.002	0.0728	0.0028	0.0000	0.0000
3,000	0.379 ± 0.002	0.0302	0.0003	0.0000	0.0000
10,000	0.299 ± 0.001	0.0108	0.0000	0.0000	0.0000
30,000	0.376 ± 0.001	0.0299	0.0004	0.0000	0.0000
100,000 ($M_T = M$)	**0.721 ± 0.003**	**0.1696**	**0.0209**	**0.0007**	**0.0002**

(d) M=100,000, N=1,000, p=0.005 [MLM using causal markers attained a $-\log_{10}P$-value of 0.825]

#markers (M_T)	$-\log_{10}P$-value	Power			
		$P < 0.05$	$P < 0.001$	$P < 10^{-6}$	$P < 5 \times 10^{-8}$
0	0.719 ± 0.003	0.1662	0.0203	0.0006	0.0001
100	0.595 ± 0.003	0.1158	0.0084	0.0001	0.0000
300	0.543 ± 0.002	0.0930	0.0047	0.0000	0.0000
1,000	0.504 ± 0.002	0.0782	0.0034	0.0000	0.0000
3,000	0.383 ± 0.002	0.0317	0.0003	0.0000	0.0000
10,000	0.300 ± 0.001	0.0101	0.0000	0.0000	0.0000
30,000	0.376 ± 0.001	0.0298	0.0003	0.0000	0.0000
100,000 ($M_T = M$)	**0.719 ± 0.003**	**0.1670**	**0.0207**	**0.0006**	**0.0001**

(e) M=10,000, N=10,000, p=0.05 [MLM using causal markers attained a $-\log_{10}P$-value of 4.97]

#markers (M_T)	$-\log_{10}P$-value	Power			
		$P < 0.05$	$P < 0.001$	$P < 10^{-6}$	$P < 5 \times 10^{-8}$
0	2.92 ± 0.01	0.5554	0.3233	0.1430	0.1030
100	3.88 ± 0.01	0.6157	0.3981	0.2117	0.1622
300	**4.35 ± 0.01**	**0.6376**	**0.4285**	**0.2408**	**0.1909**
1,000	4.05 ± 0.01	0.6238	0.4113	0.2224	0.1738
3,000	3.57 ± 0.01	0.5982	0.3780	0.1909	0.1445
10,000 ($M_T = M$)	3.46 ± 0.01	0.5924	0.3686	0.1820	0.1368

Nature Genetics: doi:10.1038/ng.2876

(f) M=10,000, N=10,000, p=0.005 [MLM using causal markers attained a $-\log_{10}P$-value of 5.17]

#markers (M_T)	$-\log_{10}P$-value	Power			
		$P < 0.05$	$P < 0.001$	$P < 10^{-6}$	$P < 5 \times 10^{-8}$
0	2.91 ± 0.02	0.5572	0.3232	0.1429	0.1021
100	**5.01 ± 0.01**	**0.6660**	**0.4656**	**0.2792**	**0.2272**
300	4.74 ± 0.01	0.6555	0.4534	0.2632	0.2144
1,000	4.20 ± 0.01	0.6329	0.4229	0.2328	0.1854
3,000	3.61 ± 0.01	0.6053	0.3836	0.1942	0.1488
10,000 ($M_T = M$)	3.45 ± 0.02	0.5954	0.3712	0.1827	0.1387

(g) M=10,000, N=1,000, p=0.05 [MLM using causal markers attained a $-\log_{10}P$-value of 0.831]

#markers (M_T)	$-\log_{10}P$-value	Power			
		$P < 0.05$	$P < 0.001$	$P < 10^{-6}$	$P < 5 \times 10^{-8}$
0	0.723 ± 0.003	0.1673	0.0203	0.0005	0.0001
100	0.653 ± 0.003	0.1391	0.0132	0.0003	0.0000
300	0.610 ± 0.002	0.1217	0.0094	0.0001	0.0000
1,000	0.572 ± 0.002	0.1070	0.0069	0.0001	0.0000
3,000	0.532 ± 0.003	0.0896	0.0045	0.0001	0.0000
10,000 ($M_T = M$)	**0.730 ± 0.003**	**0.1709**	**0.0216**	**0.0006**	**0.0002**

(h) M=10,000, N=1,000, p=0.005 [MLM using causal markers attained a $-\log_{10}P$-value of 0.959]

#markers (M_T)	$-\log_{10}P$-value	Power			
		$P < 0.05$	$P < 0.001$	$P < 10^{-6}$	$P < 5 \times 10^{-8}$
0	0.721 ± 0.004	0.1656	0.0211	0.0005	0.0001
100	**0.775 ± 0.003**	0.1871	0.0274	0.0011	0.0003
300	0.683 ± 0.002	0.1519	0.0164	0.0004	0.0000
1,000	0.604 ± 0.002	0.1205	0.0092	0.0001	0.0000
3,000	0.536 ± 0.003	0.0907	0.0048	0.0000	0.0000
10,000 ($M_T = M$)	0.731 ± 0.004	0.1692	0.0221	0.0005	0.0001

Nature Genetics: doi:10.1038/ng.2876

Supplementary Table 6. Effectiveness of MLM using M_T top associated markers in correcting for stratification, in simulations with both stratification and causal markers. We report the average λ_{median} (± standard error) and proportion of markers with P<0.05 or P<0.001, averaged across 100 simulations (with N=10,000, M=100,000) with two subpopulations with a mean trait difference of 0.25 standard deviations and either F_{ST}=0.005 or F_{ST}=0.0025 between subpopulations, with fraction p=0.05 or p=0.005 of causal markers. Values of M_T selected by the ref. [2] approach of using the M_T top associated markers based on the global maximum of out-of-sample prediction accuracy are denoted in bold font.

(a) F_{ST}=0.005, p=0.05

#markers (M_T)	λ_{median} (± SE)	Power	
		$P < 0.05$	$P < 0.001$
0	2.06 ± 0.07	0.1634	0.0252
100	1.17 ± 0.02	0.0706	0.0025
300	1.41 ± 0.03	0.0981	0.0066
1,000	1.50 ± 0.03	0.1086	0.0084
3,000	1.43 ± 0.02	0.0999	0.0064
10,000	1.25 ± 0.01	0.0812	0.0039
30,000	1.10 ± 0.01	0.0607	0.0016
100,000	**1.01 ± 0.01**	**0.0502**	**0.0011**

(b) F_{ST}=0.005, p=0.005

#markers (M_T)	λ_{median} (± SE)	Power	
		$P < 0.05$	$P < 0.001$
0	1.97 ± 0.11	0.1487	0.0250
100	1.07 ± 0.01	0.0585	0.0017
300	**1.26 ± 0.02**	**0.0804**	**0.0036**
1,000	1.34 ± 0.02	0.0908	0.0049
3,000	1.29 ± 0.02	0.0852	0.0041
10,000	1.20 ± 0.01	0.0726	0.0028
30,000	1.06 ± 0.01	0.0550	0.0013
100,000	1.00 ± 0.01	0.0493	0.0012

Nature Genetics: doi:10.1038/ng.2876

(c) F_{ST}=0.0025, p=0.05

#markers (M_T)	λ_{median} (± SE)	Power	
		$P < 0.05$	$P < 0.001$
0	1.40 ± 0.04	0.0951	0.0070
100	1.05 ± 0.01	0.0554	0.0012
300	1.30 ± 0.02	0.0824	0.0043
1,000	1.50 ± 0.02	0.1065	0.0069
3,000	1.46 ± 0.02	0.0982	0.0060
10,000	1.25 ± 0.02	0.0781	0.0033
30,000	1.06 ± 0.01	0.0549	0.0015
100,000	**1.01 ± 0.01**	**0.0509**	**0.0010**

(d) F_{ST}=0.0025, p=0.005

#markers (M_T)	λ_{median} (± SE)	Power	
		$P < 0.05$	$P < 0.001$
0	1.48 ± 0.03	0.1031	0.0072
100	1.04 ± 0.01	0.0538	0.0012
300	**1.15 ± 0.01**	**0.0671**	**0.0021**
1,000	1.34 ± 0.02	0.0904	0.0043
3,000	1.34 ± 0.02	0.0901	0.0046
10,000	1.23 ± 0.02	0.0782	0.0033
30,000	1.08 ± 0.01	0.0598	0.0015
100,000	1.02 ± 0.01	0.0513	0.0009

Nature Genetics: doi:10.1038/ng.2876

Supplementary Table 7. Effectiveness of MLM using M_T top associated markers in increasing study power, in simulations with both stratification and causal markers. We report the average $-\log_{10}P$-values (± standard error) and power to detect significant associations at various P-value thresholds, averaged across 100 simulations (with N=10,000, M=100,000) with two subpopulations with a mean trait difference of 0.25 standard deviations and either F_{ST}=0.005 or F_{ST}=0.0025 between subpopulations, with fraction p=0.05 or p=0.005 of causal markers. In each column, the maximum value (excluding the M_T=0 run with no stratification correction) is denoted in bold font. In each case, the ref. [2] approach of using the M_T top associated markers based on the global maximum of out-of-sample prediction accuracy selected the value of M_T that maximized power.

(a) F_{ST}=0.005, p=0.05 [MLM using causal markers attained a $-\log_{10}P$-value of 3.83]

#markers (M_T)	$-\log_{10}P$-value	Power			
		$P < 0.05$	$P < 0.001$	$P < 10^{-6}$	$P < 5 \times 10^{-8}$
0	3.07 ± 0.02	0.5686	0.3365	0.1538	0.1120
100	2.89 ± 0.02	0.5548	0.3196	0.1395	0.0999
300	2.93 ± 0.02	0.5587	0.3243	0.1415	0.1026
1,000	2.88 ± 0.02	0.5551	0.3193	0.1387	0.1004
3,000	2.75 ± 0.02	0.5425	0.3050	0.1292	0.0919
10,000	2.53 ± 0.01	0.5198	0.2804	0.1117	0.0770
30,000	2.39 ± 0.01	0.5048	0.2646	0.1001	0.0685
100,000 ($M_T = M$)	**2.94 ± 0.01**	**0.5586**	**0.3246**	**0.1431**	**0.1035**

(b) F_{ST}=0.005, p=0.005 [MLM using causal markers attained a $-\log_{10}P$-value of 4.95]

#markers (M_T)	$-\log_{10}P$-value	Power			
		$P < 0.05$	$P < 0.001$	$P < 10^{-6}$	$P < 5 \times 10^{-8}$
0	3.09 ± 0.02	0.5716	0.3385	0.1549	0.1143
100	3.61 ± 0.02	0.6015	0.3813	0.1944	0.1482
300	**3.78 ± 0.02**	**0.6094**	**0.3941**	**0.2049**	**0.1584**
1,000	3.66 ± 0.02	0.6022	0.3854	0.1965	0.1498
3,000	3.25 ± 0.02	0.5766	0.3526	0.1688	0.1239
10,000	2.71 ± 0.02	0.5377	0.2989	0.1261	0.0877
30,000	2.39 ± 0.01	0.5052	0.2645	0.0994	0.0661
100,000 ($M_T = M$)	2.97 ± 0.02	0.5601	0.3281	0.1462	0.1058

Nature Genetics: doi:10.1038/ng.2876

(c) F_{ST}=0.0025, p=0.05 [MLM using causal markers attained a $-\log_{10}P$-value of 3.84]

#markers (M_T)	$-\log_{10}P$-value	Power			
		$P < 0.05$	$P < 0.001$	$P < 10^{-6}$	$P < 5 \times 10^{-8}$
0	2.94 ± 0.02	0.5609	0.3274	0.1424	0.1022
100	2.83 ± 0.02	0.5520	0.3169	0.1349	0.0956
300	2.83 ± 0.02	0.5511	0.3156	0.1351	0.0955
1,000	2.75 ± 0.02	0.5443	0.3089	0.1283	0.0898
3,000	2.54 ± 0.01	0.5255	0.2847	0.1123	0.0765
10,000	2.24 ± 0.01	0.4920	0.2490	0.0871	0.0564
30,000	2.09 ± 0.01	0.4763	0.2311	0.0750	0.0471
100,000 ($M_T = M$)	**2.94 ± 0.02**	**0.5608**	**0.3268**	**0.1428**	**0.1025**

(d) F_{ST}=0.0025, p=0.005 [MLM using causal markers attained a $-\log_{10}P$-value of 4.95]

#markers (M_T)	$-\log_{10}P$-value	Power			
		$P < 0.05$	$P < 0.001$	$P < 10^{-6}$	$P < 5 \times 10^{-8}$
0	2.99 ± 0.02	0.5601	0.3309	0.1486	0.1078
100	3.70 ± 0.02	0.6046	0.3858	0.2014	0.1550
300	**3.85 ± 0.02**	**0.6137**	**0.3973**	**0.2119**	**0.1636**
1,000	3.73 ± 0.02	0.6095	0.3884	0.2024	0.1553
3,000	3.29 ± 0.02	0.5820	0.3555	0.1708	0.1277
10,000	2.65 ± 0.01	0.5326	0.2949	0.1225	0.0854
30,000	2.28 ± 0.01	0.4971	0.2545	0.0911	0.0594
100,000 ($M_T = M$)	2.98 ± 0.02	0.5575	0.3296	0.1477	0.1066

Nature Genetics: doi:10.1038/ng.2876

Supplementary Table 8. MLMe decreases power vs. linear regression under case-control ascertainment.

We report $-\log_{10}P$-values (± standard error) and power to detect significant associations at various P-value thresholds for linear regression and MLMe, averaged across 100 simulations with N samples, M non-candidate markers, disease prevalence f, fraction p=0.05 or p=0.005 of non-candidate causal markers, and observed-scale effect size (variance explained) of 10/N or 20/N for candidate causal markers.

(a) p=0.05, each candidate marker explains 10/N of observed-scale variance

# samples (N)	#markers (M)	Disease prevalence (f)	$-\log_{10}P$-values ± SE	
			Linear regression	MLMe
1,000	1,000	0.001	3.85 ± 0.22	2.91 ± 0.19
1,000	1,000	0.01	3.39 ± 0.21	3.01 ± 0.20
1,000	1,000	0.1	3.22 ± 0.21	3.20 ± 0.21
1,000	10,000	0.001	3.04 ± 0.15	2.77 ± 0.14
1,000	10,000	0.01	3.26 ± 0.17	3.14 ± 0.17
1,000	10,000	0.1	3.10 ± 0.16	3.11 ± 0.17
10,000	10,000	0.001	3.06 ± 0.15	2.22 ± 0.12
10,000	10,000	0.01	3.04 ± 0.16	2.64 ± 0.14
10,000	10,000	0.1	3.04 ± 0.17	3.06 ± 0.17
10,000	100,000	0.001	2.96 ± 0.16	2.78 ± 0.16
10,000	100,000	0.01	2.66 ± 0.14	2.54 ± 0.13
10,000	100,000	0.1	3.24 ± 0.16	3.26 ± 0.16

Power to detect significant associations:

(N)	(M)	(f)	Power			
			$P < 0.05$ LR / MLMe	$P < 0.001$ LR / MLMe	$P < 10^{-6}$ LR / MLMe	$P < 5 \times 10^{-8}$ LR / MLMe
1,000	1,000	0.001	0.88 / 0.77	0.61 / 0.43	0.12 / 0.06	0.07 / 0.03
1,000	1,000	0.01	0.86 / 0.84	0.52 / 0.35	0.10 / 0.07	0.06 / 0.03
1,000	1,000	0.1	0.84 / 0.86	0.43 / 0.42	0.10 / 0.09	0.04 / 0.05
1,000	10,000	0.001	0.92 / 0.85	0.44 / 0.35	0.03 / 0.02	0.01 / 0.00
1,000	10,000	0.01	0.90 / 0.86	0.49 / 0.48	0.05 / 0.03	0.02 / 0.02
1,000	10,000	0.1	0.88 / 0.90	0.49 / 0.47	0.05 / 0.08	0.02 / 0.02
10,000	10,000	0.001	0.88 / 0.75	0.54 / 0.21	0.03 / 0.01	0.01 / 0.00
10,000	10,000	0.01	0.90 / 0.80	0.45 / 0.32	0.04 / 0.01	0.01 / 0.01
10,000	10,000	0.1	0.85 / 0.85	0.46 / 0.47	0.03 / 0.04	0.02 / 0.03
10,000	100,000	0.001	0.82 / 0.81	0.43 / 0.38	0.06 / 0.03	0.01 / 0.02
10,000	100,000	0.01	0.83 / 0.81	0.41 / 0.39	0.01 / 0.01	0.00 / 0.00
10,000	100,000	0.1	0.92 / 0.92	0.54 / 0.55	0.05 / 0.06	0.02 / 0.02

Nature Genetics: doi:10.1038/ng.2876

(b) p=0.005, each candidate marker explains 10/N of observed-scale variance

# samples (N)	#markers (M)	Disease prevalence (f)	$-\log_{10}P$-values ± SE	
			Linear regression	MLMe
1,000	1,000	0.001	3.26 ± 0.15	2.58 ± 0.15
1,000	1,000	0.01	3.34 ± 0.18	2.94 ± 0.15
1,000	1,000	0.1	3.12 ± 0.15	3.19 ± 0.16
1,000	10,000	0.001	3.07 ± 0.18	2.90 ± 0.17
1,000	10,000	0.01	2.83 ± 0.13	2.78 ± 0.13
1,000	10,000	0.1	2.90 ± 0.16	2.90 ± 0.16
10,000	10,000	0.001	2.90 ± 0.16	2.16 ± 0.13
10,000	10,000	0.01	2.96 ± 0.13	2.58 ± 0.12
10,000	10,000	0.1	3.29 ± 0.17	3.26 ± 0.16
10,000	100,000	0.001	3.09 ± 0.15	2.80 ± 0.15
10,000	100,000	0.01	2.88 ± 0.14	2.80 ± 0.14
10,000	100,000	0.1	3.17 ± 0.16	3.20 ± 0.16

Power to detect significant associations:

(N)	(M)	(f)	Power			
			$P < 0.05$ LR / MLMe	$P < 0.001$ LR / MLMe	$P < 10^{-6}$ LR / MLMe	$P < 5 \times 10^{-8}$ LR / MLMe
1,000	1,000	0.001	0.92 / 0.77	0.54 / 0.55	0.05 / 0.06	0.01 / 0.01
1,000	1,000	0.01	0.85 / 0.87	0.52 / 0.34	0.04 / 0.03	0.03 / 0.01
1,000	1,000	0.1	0.90 / 0.90	0.56 / 0.39	0.07 / 0.04	0.01 / 0.01
1,000	10,000	0.001	0.85 / 0.80	0.56 / 0.53	0.04 / 0.05	0.03 / 0.01
1,000	10,000	0.01	0.87 / 0.86	0.47 / 0.43	0.05 / 0.04	0.00 / 0.01
1,000	10,000	0.1	0.83 / 0.82	0.41 / 0.39	0.01 / 0.02	0.01 / 0.01
10,000	10,000	0.001	0.84 / 0.72	0.40 / 0.43	0.04 / 0.04	0.01 / 0.00
10,000	10,000	0.01	0.90 / 0.85	0.41 / 0.24	0.05 / 0.01	0.00 / 0.00
10,000	10,000	0.1	0.91 / 0.90	0.46 / 0.40	0.02 / 0.00	0.05 / 0.03
10,000	100,000	0.001	0.93 / 0.92	0.56 / 0.52	0.07 / 0.07	0.01 / 0.01
10,000	100,000	0.01	0.84 / 0.86	0.48 / 0.39	0.04 / 0.04	0.01 / 0.01
10,000	100,000	0.1	0.91 / 0.92	0.47 / 0.46	0.01 / 0.02	0.02 / 0.03

Nature Genetics: doi:10.1038/ng.2876

(c) p=0.05, each candidate marker explains 20/N of observed-scale variance

# samples (N)	#markers (M)	Disease prevalence (f)	$-\log_{10}P$-values ± SE	
			Linear regression	MLMe
1,000	1,000	0.001	6.75 ± 0.33	4.95 ± 0.27
1,000	1,000	0.01	6.07 ± 0.35	5.39 ± 0.34
1,000	1,000	0.1	5.42 ± 0.30	5.47 ± 0.29
1,000	10,000	0.001	5.33 ± 0.21	4.95 ± 0.20
1,000	10,000	0.01	5.75 ± 0.23	5.63 ± 0.23
1,000	10,000	0.1	5.62 ± 0.23	5.64 ± 0.23
10,000	10,000	0.001	5.36 ± 0.19	3.99 ± 0.16
10,000	10,000	0.01	5.46 ± 0.23	4.67 ± 0.21
10,000	10,000	0.1	5.36 ± 0.23	5.41 ± 0.22
10,000	100,000	0.001	5.20 ± 0.27	4.96 ± 0.25
10,000	100,000	0.01	5.00 ± 0.20	4.78 ± 0.19
10,000	100,000	0.1	5.62 ± 0.22	5.66 ± 0.22

Power to detect significant associations:

(N)	(M)	(f)	Power			
			$P < 0.05$ LR / MLMe	$P < 0.001$ LR / MLMe	$P < 10^{-6}$ LR / MLMe	$P < 5 \times 10^{-8}$ LR / MLMe
1,000	1,000	0.001	0.99 / 0.96	0.87 / 0.73	0.57 / 0.28	0.40 / 0.17
1,000	1,000	0.01	0.95 / 0.95	0.83 / 0.76	0.45 / 0.31	0.28 / 0.21
1,000	1,000	0.1	0.96 / 0.96	0.80 / 0.82	0.33 / 0.31	0.25 / 0.23
1,000	10,000	0.001	0.98 / 0.98	0.90 / 0.82	0.33 / 0.27	0.23 / 0.17
1,000	10,000	0.01	0.99 / 1.00	0.88 / 0.88	0.42 / 0.42	0.24 / 0.19
1,000	10,000	0.1	0.99 / 0.99	0.90 / 0.91	0.41 / 0.39	0.25 / 0.28
10,000	10,000	0.001	1.00 / 0.99	0.88 / 0.67	0.37 / 0.13	0.13 / 0.04
10,000	10,000	0.01	0.99 / 0.99	0.87 / 0.79	0.34 / 0.24	0.24 / 0.10
10,000	10,000	0.1	0.99 / 0.99	0.84 / 0.84	0.42 / 0.39	0.14 / 0.19
10,000	100,000	0.001	1.00 / 1.00	0.80 / 0.78	0.34 / 0.31	0.18 / 0.18
10,000	100,000	0.01	1.00 / 0.98	0.82 / 0.81	0.32 / 0.24	0.14 / 0.09
10,000	100,000	0.1	0.99 / 0.99	0.89 / 0.89	0.41 / 0.38	0.21 / 0.20

Nature Genetics: doi:10.1038/ng.2876

(d) p=0.005, each candidate marker explains 20/N of observed-scale variance

# samples (N)	#markers (M)	Disease prevalence (f)	$-\log_{10}P$-values ± SE	
			Linear regression	MLMe
1,000	1,000	0.001	5.81 ± 0.20	4.34 ± 0.20
1,000	1,000	0.01	5.86 ± 0.27	5.24 ± 0.22
1,000	1,000	0.1	5.55 ± 0.22	5.66 ± 0.22
1,000	10,000	0.001	5.42 ± 0.22	5.11 ± 0.22
1,000	10,000	0.01	5.22 ± 0.19	5.13 ± 0.19
1,000	10,000	0.1	5.24 ± 0.22	5.24 ± 0.22
10,000	10,000	0.001	5.26 ± 0.23	3.90 ± 0.18
10,000	10,000	0.01	5.30 ± 0.18	4.62 ± 0.17
10,000	10,000	0.1	5.64 ± 0.22	5.68 ± 0.21
10,000	100,000	0.001	5.49 ± 0.26	5.01 ± 0.23
10,000	100,000	0.01	5.14 ± 0.19	5.01 ± 0.19
10,000	100,000	0.1	5.55 ± 0.22	5.61 ± 0.23

Power to detect significant associations:

(N)	(M)	(f)	Power			
			$P < 0.05$ LR / MLMe	$P < 0.001$ LR / MLMe	$P < 10^{-6}$ LR / MLMe	$P < 5 \times 10^{-8}$ LR / MLMe
1,000	1,000	0.001	1.00 / 0.97	0.90 / 0.71	0.46 / 0.19	0.21 / 0.09
1,000	1,000	0.01	0.97 / 0.99	0.85 / 0.89	0.43 / 0.34	0.25 / 0.21
1,000	1,000	0.1	0.98 / 0.98	0.91 / 0.91	0.43 / 0.43	0.15 / 0.23
1,000	10,000	0.001	1.00 / 0.99	0.87 / 0.87	0.35 / 0.33	0.19 / 0.13
1,000	10,000	0.01	1.00 / 1.00	0.85 / 0.85	0.32 / 0.27	0.10 / 0.12
1,000	10,000	0.1	1.00 / 1.00	0.83 / 0.83	0.36 / 0.34	0.17 / 0.15
10,000	10,000	0.001	0.99 / 0.92	0.84 / 0.64	0.36 / 0.15	0.20 / 0.06
10,000	10,000	0.01	1.00 / 0.98	0.90 / 0.85	0.34 / 0.26	0.12 / 0.02
10,000	10,000	0.1	0.98 / 0.98	0.90 / 0.94	0.36 / 0.42	0.22 / 0.21
10,000	100,000	0.001	0.99 / 0.99	0.89 / 0.81	0.37 / 0.33	0.21 / 0.09
10,000	100,000	0.01	1.00 / 0.99	0.85 / 0.86	0.38 / 0.32	0.25 / 0.21
10,000	100,000	0.1	0.99 / 0.99	0.93 / 0.93	0.32 / 0.34	0.15 / 0.23

Nature Genetics: doi:10.1038/ng.2876

Supplementary Table 9 . Mis-estimation of h_g^2 in simulations with case-control ascertainment. We report the true and estimated values of heritability explained by genotyped markers, in simulations with liability-scale h_g^2=0.5. Values are reported on the observed scale, denoted $h_{g,obs}^2$. Estimates of $h_{g,obs}^2$ (± standard error) are based on 100 simulations with the search space for $h_{g,obs}^2$ constrained to the interval [0.0,1.0].

# samples (*N*)	#markers (*M*)	Disease prevalence (*f*)	True $h_{g,obs}^2$	Estimated $h_{g,obs}^2$ ± SE
1,000	1,000	0.001	1.42	0.671 ± 0.003
1,000	1,000	0.01	0.91	0.579 ± 0.003
1,000	1,000	0.1	0.48	0.418 ± 0.005
1,000	10,000	0.001	1.42	0.899 ± 0.015
1,000	10,000	0.01	0.91	0.803 ± 0.011
1,000	10,000	0.1	0.48	0.436 ± 0.013
10,000	10,000	0.001	1.42	0.667 ± 0.001
10,000	10,000	0.01	0.91	0.581 ± 0.001
10,000	10,000	0.1	0.48	0.422 ± 0.002
10,000	100,000	0.001	1.42	0.862 ± 0.000
10,000	100,000	0.01	0.91	0.812 ± 0.004
10,000	100,000	0.1	0.48	0.469 ± 0.004

Nature Genetics: doi:10.1038/ng.2876

Supplementary Table 10. Additional empirical results in MS and UC data sets. (a) We report average χ^2 association statistics for all markers (λ_{median} in parentheses) and for published associated markers, for PCA with 10 or 20 PC covariates. (b) For FaST-LMM using the top M_T markers for various values of M_T, we report average χ^2 association statistics for all markers (λ_{median} in parentheses), log likelihoods (logL) for out-of-sample prediction, and average χ^2 association statistics for published markers. Results are reported for values up to M_T=3,000 for MS and M_T=1,000 for UC. FaST-Top selected M_T=2,000 top markers for MS and M_T=400 top markers for UC, based on the first local minimum in λ_{median}. Fast-TopX selected M_T=2,800 top markers for MS (logL=−871.16) and M_T=3 top markers for UC (logL=−545.87), based on the first local maximum in logL.

(a)

	PCA10	PCA20
MS, 360,557 SNPs (λ_{median})	1.25 (1.24)	1.20 (1.18)
MS, 75 published SNPs	10.11	9.77
UC, 458,560 SNPs (λ_{median})	1.10 (1.10)	1.10 (1.10)
UC, 24 published SNPs	13.57	13.48

(b)

	100	200	300	400	500	1,000	1,500	2,000	2,500	3,000
MS, 360,557 SNPs	2.22 (2.16)	1.76 (1.71)	1.63 (1.58)	1.55 (1.50)	1.53 (1.49)	1.46 (1.43)	1.42 (1.39)	1.42 (1.39)	1.42 (1.40)	1.42 (1.39)
logL	-980.24	-952.65	-938.12	-928.29	-922.00	-896.40	888.45	-884.28	-888.45	-871.62
MS, 75 published SNPs	14.75	13.62	13.06	12.50	12.31	11.70	11.18	10.99	10.73	10.49
UC, 458,560 SNPs	1.12 (1.12)	1.11 (1.09)	1.10 (1.08)	1.08 (1.09)	1.09 (1.08)	1.10 (1.10)				
logL	-561.74	-580.03	-588.65	-600.34	-613.61	-649.62				
UC, 24 published SNPs	12.94	12.21	11.19	10.75	10.67	10.48				

Nature Genetics: doi:10.1038/ng.2876

Supplementary Table 11. Results for published associated markers in MS and UC data sets. We list results for each published associated marker used to compute the averages reported in Table 4 and Table S5.

MS

NHGRI SNP	Tag SNP	Chr.	Position	Linear	PCA	PCA10	PCA20	MLMi	MLMe	FaST-4K	FaST-Top	FaST-TopX
rs4648356	rs4648356	1	2699024	90.80	36.04	35.44	38.02	42.28	54.17	61.48	47.30	42.11
rs233100	rs233100	1	85544597	7.79	9.67	9.56	6.91	5.71	8.03	7.63	12.34	12.58
rs6604026	rs6604026	1	93076191	18.04	19.65	17.08	14.69	7.41	13.11	17.75	21.61	20.27
rs11581062	rs11581062	1	101180107	11.84	25.27	22.90	19.64	18.55	24.23	16.38	27.07	26.44
rs2300747	rs1335532	1	116902480	15.19	37.39	36.70	35.13	23.68	25.63	27.30	36.70	36.56
rs3761959	rs3761959	1	155935902	60.34	19.28	18.66	16.67	12.92	21.37	35.34	22.25	20.27
rs1323292	rs1323292	1	190807644	14.84	5.61	5.10	5.47	6.39	7.87	8.78	2.17	2.22
rs12466022	rs12466022	2	43212565	9.08	2.86	3.30	4.38	2.57	2.92	6.70	1.68	1.78
rs7595037	rs7595037	2	68500599	17.39	1.02	1.06	0.67	4.36	6.28	7.94	1.36	1.86
rs17174870	rs17174870	2	112381672	15.76	13.97	13.35	12.85	11.16	19.61	21.23	18.99	18.53
rs10201872	rs10201872	2	230814968	3.85	12.52	13.09	13.32	6.33	7.21	11.43	14.60	16.88
rs9821630	rs9821630	3	16945942	15.17	1.97	1.73	1.55	4.29	4.49	6.73	2.94	2.51
rs11129295	rs11129295	3	27763784	50.42	12.72	13.82	14.99	11.18	11.68	27.10	10.66	10.30
rs669607	rs669607	3	28046448	32.97	14.92	15.56	13.59	8.14	10.28	21.74	14.71	12.49
rs771767	rs771767	3	103231328	2.12	4.00	3.39	4.66	4.80	5.39	4.74	8.64	8.14
rs2293370	rs2293370	3	120702624	61.47	19.26	16.95	15.21	18.78	27.85	41.68	16.35	15.58
rs4285028	rs4285028	3	123143354	3.83	5.56	5.06	7.90	1.19	1.93	3.63	5.72	5.57
rs4308217	rs4308217	3	123275877	3.05	1.54	1.42	2.45	1.02	1.48	1.77	0.44	0.16
rs9282641	rs9282641	3	123279458	9.31	16.55	15.61	14.56	12.72	14.88	12.93	11.38	12.87
rs908821	rs908821	3	142023408	5.06	0.64	0.92	1.23	1.28	1.68	3.55	0.43	0.33
rs1841770	rs1841770	3	149239376	23.29	0.28	0.25	0.78	0.51	0.42	2.28	1.65	1.98
rs2243123	rs2243123	3	161192345	2.01	4.57	4.47	5.78	7.14	5.30	6.81	8.87	8.00
rs10936599	rs10936599	3	170974795	12.36	17.67	17.83	17.05	15.11	20.63	17.71	13.20	11.84
rs228614	rs228614	4	103797685	0.65	2.32	1.94	1.98	1.09	1.48	0.21	3.68	4.11
rs12644284	rs12644284	4	154373450	8.78	0.29	0.40	0.46	0.59	0.59	1.55	0.03	0.03
rs7672826	rs7672826	4	182636689	0.12	0.66	0.91	0.64	0.09	0.14	0.01	1.01	1.06
rs6897932	rs6897932	5	35910332	31.44	14.55	14.49	17.74	17.74	26.31	26.32	18.43	15.76
rs756699	rs756699	5	133474474	1.28	7.94	8.53	7.41	4.15	3.97	1.24	9.51	11.57
rs1062158	rs1062158	5	141503184	33.02	5.81	5.24	6.16	8.76	10.55	14.28	4.41	3.30
rs2546890	rs2546890	5	158692478	13.88	7.75	8.33	11.99	12.85	16.22	9.88	12.24	12.54
rs4075958	rs4075958	5	176717118	18.19	9.31	9.80	9.30	7.68	9.93	13.89	8.04	9.45
rs11755724	rs11755724	6	7063989	3.34	3.04	2.90	1.72	3.64	4.53	1.54	0.33	0.12
rs11962089	rs11962089	6	105718913	0.26	0.29	0.46	0.77	0.13	0.53	0.83	1.25	1.88
rs802734	rs802734	6	128320491	32.26	3.83	4.66	3.27	4.70	5.66	19.58	4.58	3.72
rs9321490	rs9321490	6	135536568	6.15	11.38	11.06	12.82	8.43	8.90	8.87	10.04	8.29
rs11154801	rs11154801	6	135781048	27.00	27.68	28.75	27.68	17.76	25.71	29.81	37.26	35.39
rs17066096	rs17066096	6	137494601	17.81	16.21	15.44	14.88	12.10	13.23	17.29	14.29	12.91
rs13192841	rs13192841	6	138008907	7.07	7.23	7.25	5.45	3.27	4.40	5.74	4.04	3.14
rs1738074	rs1738074	6	159385965	56.40	12.17	12.29	12.22	17.60	24.42	40.40	20.70	18.64
rs6952809	rs6952809	7	2415019	16.34	12.10	11.49	11.43	9.67	10.63	12.70	9.84	8.67
rs758944	rs758944	7	75791233	3.88	1.46	1.95	2.65	4.84	4.76	1.54	3.21	3.35
rs354033	rs354033	7	148920397	8.39	7.81	7.62	6.96	5.85	7.07	8.12	10.14	10.95

Nature Genetics: doi:10.1038/ng.2876

rs1520333	rs1520333	8	79563593	32.00	6.25	5.33	5.57	3.60	6.10	15.95	2.95	3.10
rs2019960	rs2019960	8	129261453	0.22	1.33	0.92	1.24	0.86	0.69	0.09	0.11	0.46
rs2150702	rs2150702	9	5883861	7.86	0.95	1.04	0.77	1.58	1.67	3.45	3.78	2.67
rs1755289	rs1755289	9	17928351	1.74	0.83	0.74	0.58	1.50	2.02	1.33	1.92	1.71
rs290986	rs290986	9	92603357	12.84	15.09	16.29	18.48	12.42	14.17	18.47	18.62	16.54
rs3780792	rs3780792	9	135825164	2.47	0.01	0.00	0.01	0.13	0.19	1.05	0.07	0.01
rs3118470	rs3118470	10	6141719	43.11	17.58	18.97	12.74	19.90	25.05	32.51	25.55	23.07
rs1250550	rs1250550	10	80730323	1.60	11.86	12.43	10.50	12.26	14.09	7.00	14.37	14.47
rs7923837	rs7923837	10	94471897	21.49	9.47	7.94	5.95	5.99	7.04	18.16	12.07	10.97
rs650258	rs650258	11	60588858	24.88	14.44	14.37	13.00	13.82	16.42	15.31	12.71	10.73
rs4409785	rs4409785	11	94951070	18.66	9.10	8.73	9.00	8.34	9.10	8.45	8.38	8.76
rs630923	rs630923	11	118259563	0.79	6.86	7.61	6.90	6.77	8.77	2.47	8.95	8.27
rs1458175	rs1458175	12	40252128	5.53	1.41	1.87	1.95	0.00	0.24	0.20	2.67	3.69
rs703842	rs703842	12	56449006	44.12	30.15	28.06	22.17	17.52	26.00	35.58	19.95	19.34
rs9523762	rs9523762	13	92129887	0.60	0.01	0.17	0.33	0.01	0.18	0.10	0.02	0.03
rs4902647	rs4902647	14	68323944	15.50	5.55	5.64	4.12	3.48	5.23	10.51	3.99	4.49
rs2300603	rs2300603	14	75075310	28.46	8.95	9.87	9.94	12.34	16.78	14.11	8.79	7.29
rs2744148	rs2744148	16	1013553	16.75	9.99	10.31	10.58	7.52	8.50	13.87	5.27	4.45
rs7200786	rs7200786	16	11085302	73.22	28.13	27.79	25.30	25.30	38.07	57.77	36.28	35.78
rs386965	rs386965	16	78210042	31.93	26.01	25.13	21.59	20.85	24.29	27.22	24.35	25.48
rs13333054	rs13333054	16	84568534	5.60	14.12	15.77	14.13	11.91	12.59	8.67	10.35	11.55
rs4792814	rs4792814	17	40758788	7.47	5.44	5.29	3.86	1.39	3.62	5.56	5.16	4.67
rs180515	rs180515	17	55379057	11.98	12.66	13.29	18.67	15.43	16.19	19.10	17.76	16.17
rs12456021	rs12456021	18	54364370	11.83	2.56	2.52	5.35	4.22	4.60	6.81	1.59	0.94
rs7238078	rs7238078	18	54535172	11.36	7.56	6.85	6.80	4.00	4.79	7.75	4.80	5.50
rs1077667	rs1077667	19	6619972	33.19	24.83	24.96	24.61	27.38	29.93	33.68	33.32	32.05
rs874628	rs874628	19	18165700	26.94	13.24	14.65	11.30	16.23	22.26	19.89	23.37	24.00
rs7255066	rs7255066	19	49837943	10.87	5.45	5.36	4.58	5.17	5.82	9.17	4.53	4.27
rs307896	rs307896	19	52353333	18.90	11.94	11.41	12.67	9.52	11.34	13.33	9.41	7.86
rs281380	rs281380	19	53906282	62.93	7.50	7.77	8.15	10.83	14.43	30.18	13.39	13.13
rs397020	rs397020	20	1153886	0.34	6.42	6.86	6.56	4.86	5.95	3.43	5.46	6.17
rs2283792	rs2283792	22	20461125	11.66	6.64	5.74	5.08	5.39	6.70	9.66	2.54	3.33
rs140522	rs140522	22	49318132	18.07	8.14	8.27	7.45	8.17	9.16	11.00	7.59	7.05

UC

NHGRI SNP	Tag SNP	Chr	Position	Linear	PCA	PCA10	PCA20	MLMi	MLMe	FaST-4K	FaST-Top	FaST-TopX
rs734999	rs10797432	1	2491198	15.19	15.92	15.69	14.81	13.00	13.59	16.10	13.42	15.23
rs1317209	rs1317209	1	20012623	11.01	8.63	8.61	8.66	8.77	9.26	9.93	6.95	11.04
rs10889677	rs2201841	1	67466790	17.67	16.43	16.87	16.61	16.02	17.81	18.00	19.98	17.75
rs11209026	rs11209026	1	67478546	49.80	49.13	49.31	48.39	46.31	48.24	49.51	33.80	51.37
rs7554511	rs11584383	1	199202489	23.63	20.69	20.29	19.50	17.36	20.55	22.64	16.87	23.76
rs7608910	rs7608910	2	61058360	14.29	10.95	11.14	11.24	9.77	11.36	13.53	6.72	14.30
rs2310173	rs2310173	2	102030060	7.54	7.36	6.69	6.86	2.74	4.00	7.57	3.19	7.47
rs1016883	rs1016883	2	198589913	8.42	9.67	10.16	10.37	5.99	8.03	8.13	8.07	8.11
rs11676348	rs11676348	2	218718391	12.48	10.90	11.15	10.59	11.69	13.07	12.88	10.10	12.88
rs17388568	rs17388568	4	123548812	2.96	3.84	3.10	3.23	2.12	2.44	3.38	2.97	3.69

Nature Genetics: doi:10.1038/ng.2876

rs11739663	rs11739663	5	647083	17.18	15.91	15.96	15.85	17.44	18.91	16.88	15.32	16.76
rs6451493	rs6451493	5	40446692	16.18	17.49	16.65	16.19	14.04	15.93	15.93	7.92	14.91
rs6920220	rs6920220	6	138048197	15.94	19.58	19.69	18.60	16.28	16.61	16.31	13.91	14.98
rs4380874	rs4510766	7	107280025	18.26	17.22	16.61	16.60	12.95	16.46	17.56	9.71	18.57
rs4728142	rs4728142	7	128361203	10.53	8.83	9.39	9.75	8.04	8.95	10.61	9.05	10.51
rs4246905	rs4246905	9	116593070	20.08	19.81	19.68	21.45	18.27	19.42	19.88	17.86	20.78
rs10781500	rs10781500	9	138389159	16.64	18.45	18.16	19.59	14.77	16.33	17.83	11.93	17.25
rs678170	rs561722	11	113892040	7.31	8.53	8.37	7.99	9.11	10.85	8.03	9.90	6.89
rs2870946	rs2870946	12	66882928	0.98	0.79	0.75	0.57	0.46	0.59	1.07	1.28	0.97
rs941823	rs941823	13	39911977	8.21	6.84	7.04	6.17	6.64	6.93	7.82	6.25	8.69
rs2872507	rs2872507	17	35294289	16.13	18.02	18.38	17.89	17.40	19.97	17.02	12.81	17.63
rs2297441	rs2297441	20	61798026	14.16	12.17	11.82	12.03	10.58	11.09	12.94	9.45	12.41
rs1297265	rs1736135	21	15727091	5.71	3.61	3.84	3.97	3.72	4.11	5.24	4.18	5.71
rs2838519	rs2838519	21	44439451	7.23	6.40	6.37	6.55	7.34	7.91	7.04	6.41	6.66

Nature Genetics: doi:10.1038/ng.2876

Supplementary Note

1. Methods

1.1. GCTA implementation of MLMi (GCTA-MLMi). The phenotype **y** is modeled as

$$\mathbf{y} = \mathbf{Kc} + \mathbf{g} + \mathbf{e} \qquad \text{[Model 1]}$$

where **c** is a vector of fixed covariates (including the affine term) with corresponding coefficient matrix **K**; **g** is a vector of genetic effects with $\mathbf{g} \sim N(0, \mathbf{A}\sigma_g^2)$; **A** is the GRM defined by $A_{jk} = \frac{1}{M}\sum_{i=1}^{M}\frac{(x_{ij}-2p_i)(x_{ik}-2p_i)}{2p_i(1-p_i)}$, where $x_{ij} = 0$, 1 or 2 and M is the total number of autosomal markers; and **e** is a vector of non-genetic effects with $\mathbf{e} \sim N(0, \mathbf{I}\sigma_e^2)$. We note that this is the same GRM used in previous work on principal components analysis[3]. The variance explained by the GRM (σ_g^2), the noise variance (σ_ε^2) and the heritability explained by genotyped markers ($h_g^2 = \sigma_g^2/(\sigma_g^2 + \sigma_e^2)$) can be estimated via restricted maximum likelihood (REML)[4,5]. Association statistics are computed by comparing the causal model with an effect at the candidate marker to the null model with no effect at the candidate marker, for example via a generalized least squares (GLS) F-test or χ^2 score test[6].

We then test the effect of SNP i based on the model

$$\mathbf{y} = \mathbf{Kc} + \mathbf{w}_i b_i + \mathbf{g} + \mathbf{e} \qquad \text{[Model 2]}$$

where $\mathbf{w}_i$ is a vector of mean-adjusted genotypes, i.e. $w_{ij} = x_{ij} - 2p_i$. The fixed effects, including both the effects of the covariates and the effect of SNP i, are estimated by GLS, i.e. $\hat{\mathbf{q}} = (\mathbf{Q}^T\mathbf{V}^{-1}\mathbf{Q})^{-1}\mathbf{Q}^T\mathbf{V}^{-1}\mathbf{y}$ with $\mathrm{var}(\hat{\mathbf{q}}) = (\mathbf{Q}^T\mathbf{V}^{-1}\mathbf{Q})^{-1}$, where $\mathbf{q} = [\mathbf{c}^T : b_i]^T$, $\mathbf{Q} = [\mathbf{K} : \mathbf{w}_i]$ and $\mathbf{V} = \mathbf{A}\sigma_g^2 + \mathbf{I}\sigma_e^2$. The test-statistic is calculated as $\chi^2 = \hat{b}_i^2 / \mathrm{var}(\hat{b}_i)$, where $\mathrm{var}(\hat{b}_i)$ is the last diagonal entry of $(\mathbf{Q}^T\mathbf{V}^{-1}\mathbf{Q})^{-1}$. Assuming that the proportion of variance explained by a single SNP is small, the estimates of σ_g^2 and σ_e^2 from [Model 1] will be very similar to those from [Model 2]. To decrease the computational burden, we estimate σ_g^2 and σ_e^2 once based on [Model 1], without any SNPs included in the fixed effect vector **c,** and use them for effect size estimation and significance

Nature Genetics: doi:10.1038/ng.2876

testing of each SNP based on [Model 2]; this is an approximate approach similar to that implemented in EMMAX[6]. If there are no covariates, the vector **c** will become a scalar (i.e. the affine term). In this case, for ease of computation, we can adjust the phenotype as $\mathbf{y}^* = \mathbf{y} - \mathbf{1}\hat{c}$ with $\hat{c} = (\mathbf{1}^T\mathbf{V}^{-1}\mathbf{1})^{-1}\mathbf{1}^T\mathbf{y}$ and simplify the estimation of SNP effect as $\hat{b}_i = \mathbf{w}_i^T\mathbf{V}^{-1}\mathbf{y}^* / (\mathbf{w}_i^T\mathbf{V}^{-1}\mathbf{w}_i)$ with $\text{var}(\hat{b}_i) = 1/(\mathbf{w}_i^T\mathbf{V}^{-1}\mathbf{w}_i)$. However, when there are covariates fitted in the models, we estimate the effects of the covariates and the SNP jointly rather than using the simplified approach. This is because if the covariates and the SNP genotype are correlated and the phenotype is pre-adjusted by the covariates, the power of detecting the SNP effect can be reduced. To improve computational efficiency, the efficient computational libraries EIGEN, BLAS and LAPACK are used for linear algebra calculation, and the parallel computing technique OpenMP is used for multi-thread computing.

1.2. GCTA implementation of MLMe via LOCO analysis (GCTA-LOCO). This implementation is identical to GCTA-MLMi, except that markers on a given autosome are evaluated using a GRM constructed from the remaining autosomes, via pre-computing and storing the GRM constructed from all autosomes. GCTA-LOCO attains running time and memory usage only 2-3x higher than GCTA-MLMi (Supplementary Table 1).

1.3. Simulations of a quantitative trait with no sample structure. We simulated phenotypes for *N* samples using *M* non-candidate markers plus 500 candidate markers. All simulations used *N*=10,000 and *M*=100,000 unless otherwise specified. (We also included simulations at smaller values of *N* and *M*, in order to understand how the relative performance of different methods varies with *N* and *M*.) Allele frequencies were uniformly distributed on [0.1,0.9]. The *M* non-candidate markers included *Mp* causal markers (*p*=0.05 unless otherwise specified) explaining 50% of the variance of the trait, with normalized effect sizes ~ N(0,0.5/*Mp*). The 500 candidate markers explained an additional 50% of the variance of the trait, with normalized effect sizes ~ N(0,0.5/500). We ran MLMe by including the *M* non-candidate markers in the GRM and testing only the 500 candidate markers. We ran MLMi by the including the *M* non-candidate markers plus the 500 candidate

Nature Genetics: doi:10.1038/ng.2876

markers in the GRM. Thus, the heritability explained by markers included in the GRM was 50% for MLMe and 100% for MLMi.

1.4. Yang et al. 2011 data and simulated phenotypes. We analyzed data from 14,347 individuals from the ARIC, HPFS and NHS cohorts that were genotyped using Affymetrix 6.0 arrays at 565,040 autosomal markers after quality control, as described previously[1]. Informed consent was obtained from all subjects. We excluded one of each pair of individuals with genetic relatedness >0.025 in the GRM, leaving 11,586 unrelated individuals. We selected a random subset of 10,000 unrelated individuals. We restricted to 45,772 markers on chr1, 47,596 markers on chr2 and 39,668 markers on chr3, for a total of M=133,036 markers. We randomly selected 200 causal markers (100 each on chr1 and chr2) explaining 50% of the variance of the trait, with normalized effect sizes ~ N(0,0.5/200). We ran MLMi by including all markers in the GRM (GCTA-MLMi), and ran MLMe using a GRM estimated from the remaining chromosomes (GCTA-LOCO). We included causal markers on two different chromosomes in order to assess the benefit of running MLMe using a GRM containing causal markers on a different chromosome, and included a third chromosome with no causal markers in order to assess association statistics at markers that are not causal and not in LD with a causal marker.

1.5. Simulations of a quantitative trait with population stratification. We simulated phenotypes for N/2 samples from each of two discrete subpopulations, based on a mean trait difference of 0.25 standard deviations between subpopulations with no causal marker effects. We simulated M markers for the GRM plus 500 additional candidate markers, based on F_{ST}=0.005 or F_{ST}=0.0025 between subpopulations. Ancestral allele frequencies x were uniformly distributed on [0.1,0.9], and subpopulation allele frequencies were sampled from a beta distribution with parameters $x(1-F_{ST})/F_{ST}$ and $(1-x)(1-F_{ST})/F_{ST}$, which has mean x and variance $F_{ST}x(1-x)$. Causal markers in simulations with both stratification and causal markers were simulated in the same way as in simulations with no sample structure (see above).

Nature Genetics: doi:10.1038/ng.2876

1.6. Simulations of ascertained case-control traits. We simulated normally distributed liabilities, transformed liabilities to case-control status by defining individuals with liability $>T$ to be cases and others to be controls (where the liability threshold T is chosen to achieve a specified disease prevalence f), and continued in this fashion until $N/2$ cases and $N/2$ controls were generated. Liabilities were simulated as in the simulations of a quantitative trait with no sample structure, except that liability-scale effect sizes of the 5 candidate markers were chosen so that each candidate marker explains the proportion $10/N$ of observed-scale variance, after accounting for the transformation between variance explained on the liability scale vs. the observed scale with correction for case-control ascertainment[7]. We used only 5 candidate markers so as to limit the liability-scale variance explained by the candidate markers, which exceeds observed-scale variance when prevalence is low. We also included M non-candidate markers, of which the proportion p were causal (p=0.05 or p=0.005) and explained 50% of the liability-scale variance. To examine results for candidate markers of larger effect, we repeated all simulations with each candidate marker explaining the proportion $20/N$ of observed-scale variance.

1.7. MS and UC datasets. We analyzed data from 10,204 MS cases and 5,429 controls (from NBS and 1958BC) genotyped on Illumina arrays made available to researchers via WTCCC2 (see Web Resources). Although ref. [8] analyzed UK and non-UK samples separately followed by meta-analysis in most of their analyses, the data made available to researchers includes both UK and non-UK cases but only UK controls. We retained all samples in order to maximize sample size. We considered markers that were present in each of MS, NBS and 1958 BC datasets and removed markers with >0.5% missing data, P<0.01 for allele frequency difference between NBS and 1958BC, P<0.05 for deviation from Hardy-Weinberg equilibrium, P<0.05 for differential missingness between cases and controls, or MAF<0.1% in any dataset, leaving 360,557 markers. We employed filters more stringent than in a standard GWAS so as to minimize the impact of assay artifacts on our results[7]. The 75 known associated markers were defined by including, for each MS-associated marker listed

Nature Genetics: doi:10.1038/ng.2876

in the NHGRI GWAS catalogue (see Web Resources), a single best tag at r^2>0.4 from the set of 360,557 markers if available.

We also analyzed data from 2,697 UC cases and 5,652 controls (from NBS and 1958BC) genotyped on Affymetrix arrays[9] and made available to researchers via WTCCC2. The MS and UC datasets contain overlapping control samples. We employed stringent QC filters as described above, leaving 458,560 markers. The 24 known associated markers were defined by, for each UC-associated marker listed in the NHGRI GWAS catalogue, a single best tag at r^2>0.4 from the set of 458,560 if available.

PCA refers to PC correction using 5 PCs. Correction using 10 PCs (PCA10) or 20 PCs (PCA20) was also evaluated (Supplementary Table 10). FaST-4K was run by including M_R=4,000 random markers in the GRM. FaST-Top was run by including the top M_T markers based on first local minimum of λ_{median}[10], using a grid search from 100 to 3,000 with a step size of 100. FaST-TopX was run using the command "fastlmmc -autoSelect <outfile> -randomSeed 1 -autoSelectFolds 10 -bfilesim <indata> -pheno <inpheno> -mpheno 1 -autoSelectSearchValues ASvalues.txt -topKbyLinReg 10000 -memoryFraction 0.2", as described on p.10 of the FaST-LMM version 2.05 user manual. (The -topKbyLinReg 10000 option was not used in our simulations, which selected the global optimum of M_T.) FaST-Top selected M_T=2,000 top markers for MS and M_T=400 top markers for UC, and FaST-TopX selected M_T=2,800 top markers for MS and M_T=3 top markers for UC.

Nature Genetics: doi:10.1038/ng.2876

2. Benchmarking the running time and memory usage of GCTA

To quantify the computational cost in datasets of realistic size, we benchmarked the running time and memory usage of GCTA using simulations of a quantitative trait without sample structure (see Methods section for details of the methods and for simulation design). We conducted simulations with either N=5,000 or N=10,000 samples and M=50,000 or M=100,000 markers, using all markers to compute the GRM. Running times ranged from 0.3hr to 2.5hr, consistent with $O(MN^2 + N^3)$ (Supplementary Table 1). We observed similar running time and memory usage for other implementations (e.g. differing by a factor of 4 or less for GCTA-MLMi, EMMAX, FaST-LMM and GEMMA), but we caution that running time comparisons may vary as a function of computing environment. The pairwise correlation of log P-values across different methods was >0.99999 for GCTA-MLMi, FaST-LMM and GEMMA, but only >0.99 for EMMAX. The high correlation (>0.99999) between log P-values of GCTA-MLMi and the exact methods (i.e. FaST-LMM and GEMMA), in simulations with very small effect sizes, suggests that the lower correlation (>0.99) between EMMAX and each other method is due to EMMAX computing a different association statistic.

Nature Genetics: doi:10.1038/ng.2876

3. Effective number of independent markers

Equation (10) of ref. [11] states that $\lambda_{mean} \approx 1 + Nh_g^2 r_{mean}^2 s_{mean} / M$, where N is the number of samples, h_g^2 is the heritability explained by M genotyped and/or imputed markers, r^2 for a given causal marker is the average of r^2 values with all s markers in LD with that causal marker, and r^2_{mean} and s_{mean} denote means across markers (which approximate the means across causal markers). Figure 4 of ref. [11] states that for a specific data set with M=294,831, values of $r^2_{\text{mean}} = 0.0261$ and $s_{\text{mean}} = 188.04$ were obtained, i.e. $r^2_{\text{mean}} s_{\text{mean}} = 4.91$. This implies that $\lambda_{\text{mean}} \approx 1 + Nh_g^2/60{,}073$ in this data set. We define the effective number of independent markers (M_{eff}) as the value such that $\lambda_{\text{mean}} = 1 + Nh_g^2/M_{\text{eff}}$, where λ_{mean} is the mean of Linear Regression χ^2 association statistics. Thus, it follows from ref. [11] that $M_{\text{eff}} \approx 60{,}000$ in this data set. We note that, under the approximations assumed, M_{eff} is independent of phenotype, and independent of the randomly chosen set of markers used to compute λ_{mean}.

We note that M_{eff} should <u>not</u> be used to choose significance thresholds for multiple hypotheses tested[12]. As a simple demonstration of this, we consider the case of a pair of markers with pairwise r^2=0.5. It follows that $M_{\text{eff}} = 2/1.5 = 1.33$. However, $M_{\text{eff}} = 1.33$ is not an appropriate multiple hypothesis testing correction. Indeed, simulations show that the proportion of the time that at least one marker attains $P<0.01/1.33$ is 0.013, which is larger than 0.01. The proportion of the time that at least one marker attains $P<0.01/1.8$ is roughly 0.01, implying that 1.8 (not $M_{\text{eff}} = 1.33$) is the appropriate multiple hypothesis testing correction in this example. Thus, the effective number of independent markers quantifying inflation in association statistics is a different quantity than the effective number of independent statistical tests.

Nature Genetics: doi:10.1038/ng.2876

4. Average χ^2 statistics for linear regression, MLMi and MLMe

4.1. Statistical model. The model for a MLM association analysis can be written as

$$\mathbf{y} = \mathbf{1}\mu + \mathbf{x}_t b_t + \mathbf{g} + \mathbf{e} \quad [S1]$$

where $\mathbf{y}$ is a vector of phenotypes, $\mathbf{1}$ is a vector of ones, μ is the mean term, $\mathbf{x}_t$ is a vector of genotype indicator variables of the candidate SNP, b_t is the fixed effect of the candidate SNP, $\mathbf{g} \sim N(\mathbf{0}, \mathbf{A}\sigma_g^2)$ is a vector of polygenic effects with each element being the aggregated effect of all SNPs for an individual, $\mathbf{A}$ is the genetic relationship matrix (GRM) estimated from SNP data[4] and $\mathbf{e}$ is a vector of residual effects $\mathbf{e} \sim N(\mathbf{0}, \mathbf{I}\sigma_e^2)$. For ease of discussion, we can drop the mean term from the model by standardizing the phenotype and the genotype indicator variable, i.e. $y \sim N(0,1)$ and $w_i = (x_i - 2p_i)/\sqrt{2p_i(1-p_i)}$ with p being the allele frequency.

4.2. Linear regression analysis. Let N be the sample size, M be the number of markers and M_q be the number of causal markers. Assuming no population structure or other artifacts and all markers are independent, the mean of the χ^2 assocation statistics (λ_{mean}) from linear regression (LR) analyses is[11]

λ_{mean} (LR) = $1 + Nh_g^2 / M$ at all SNP markers,

λ_{mean} (LR) = $1 + Nh_g^2 / M_q$ at causal markers and

λ_{mean} (LR) = 1 at null markers.

regardless of the genetic architecture of the trait . On the other hand, λ_{median} (the median of χ^2 association statistics divided by the expected median under the null hypothesis of no association[13]) depends on the genetic architecture of the trait. For highly polygenic traits, λ_{median} may often be only slightly lower than λ_{mean} (ref.[11]), but λ_{median} may be much lower than λ_{mean} for less polygenic traits or very large sample sizes. Other inflation measures (e.g. the truncated mean, excluding the upper tail of the distribution) are also possible.

Nature Genetics: doi:10.1038/ng.2876

4.3. Mixed linear model association analysis including the candidate SNP (MLMi). If the candidate SNP is calculating the GRM, we can also express $\mathbf{g} = \mathbf{w}_t u_t + \mathbf{W}\mathbf{u}$ where u_t is the random effect of the candidate SNP with $\mathbf{w}_t$ being the corresponding genotype vector and $\mathbf{u} \sim N(0, \mathbf{I}\sigma_u^2)$ is a vector of the random effects of all the other SNPs with $\mathbf{W}$ being the corresponding genotype matrix. If all the SNPs are independent and assuming the variance explained by the candidate SNP is small, then $\sigma_g^2 = h_g^2$, $\sigma_e^2 = 1 - h_g^2$ and $\sigma_u^2 = h_g^2 / M$ where h_g^2 is the variance explained by all SNPs. Therefore, we can re-write model [S1] as

$$\mathbf{y} = \mathbf{x}_t b_t + \mathbf{w}_t u_t + \mathbf{W}\mathbf{u} + \mathbf{e} \qquad [S2]$$

This equation clearly shows that the effect of the candidate SNP is fitted twice once as fixed (b_t) and once as random (u_t).

The analysis to estimate b_t from Equation [S2] is equivalent to an analysis of the candidate SNP and the phenotype correcting for the effects of all the other SNPs. If we define $\mathbf{g}^* = \mathbf{W}\mathbf{u}$ then

$$\mathbf{y}^* = \mathbf{y} - \hat{\mathbf{g}}^* = \mathbf{x}_t b_t + \mathbf{w}_t u_t + (\mathbf{g}^* - \hat{\mathbf{g}}^*) + \mathbf{e} = \mathbf{x}_t b_t + \mathbf{w}_t u_t + \boldsymbol{\varepsilon} \qquad [S3]$$

where $\boldsymbol{\varepsilon} \sim N(\mathbf{0}, \mathbf{I}\sigma_\varepsilon^2)$, $\text{var}(\boldsymbol{\varepsilon}) = \text{var}(\mathbf{e}) + \text{var}(\mathbf{g}^* - \hat{\mathbf{g}}^*) = \mathbf{I}\sigma_e^2 + \mathbf{I}(1 - r^2)h_g^2 = \mathbf{I}(1 - r^2 h_g^2)$ i.e. $\sigma_\varepsilon^2 = (1 - r^2 h_g^2)$, r^2 is the accuracy squared of predicting $\mathbf{g}^*$, $r^2 = \theta / (\theta + 1 - h_g^2 r^2)$ with $\theta = N h_g^2 / M$ (see ref. [14]).

In Equation [S3], b_t and u_t can be estimated by the mixed model equation (MME)

$$\begin{bmatrix} \mathbf{w}_t'\mathbf{w}_t & \mathbf{w}_t'\mathbf{w}_t \\ \mathbf{w}_t'\mathbf{w}_t & \mathbf{w}_t'\mathbf{w}_t + \mathbf{I}\dfrac{\sigma_\varepsilon^2}{\sigma_u^2} \end{bmatrix} \begin{bmatrix} \hat{b}_t \\ \hat{u}_t \end{bmatrix} = \begin{bmatrix} \mathbf{w}_t'\mathbf{y}^* \\ \mathbf{w}_t'\mathbf{y}^* \end{bmatrix} \qquad [S4]$$

By solving the MME, we can have $\hat{b}_t = \mathbf{w}_t'\mathbf{y}^* / \mathbf{w}_t'\mathbf{w}_t$ which is in the same form as the least squares estimate of regressing y on w_t, $\hat{u}_t = 0$ because of fitting the same SNP as a fixed effect, and $\sigma^2(\hat{b}_t) = \sigma_u^2 + \sigma_\varepsilon^2 / N$.

Considering a Wald test, $\chi^2 = \hat{b}_t^2 / \text{var}(\hat{b}_t)$. Therefore,

$$\text{E}(\chi^2) = \text{E}(\hat{b}_t^2) / \sigma^2(\hat{b}_t) = [b_t^2 + \text{var}(\hat{b}_t)] / \sigma^2(\hat{b}_t) = (b_t^2 + \sigma_\varepsilon^2 / N) / (\sigma_u^2 + \sigma_\varepsilon^2 / N) \qquad [S5]$$

Nature Genetics: doi:10.1038/ng.2876

It is notable that $\text{var}(\hat{b}_t)$ is the variance of $\hat{b}_t = \mathbf{w}_t'\mathbf{y}^* / \mathbf{w}_t'\mathbf{w}_t$ while $\sigma^2(\hat{b}_t)$ is the sampling variance obtained from the mixed linear model analysis and used for significance test.

If we take the average of the χ^2 statistics across all the SNPs, $\sum_i b_i^2 / M = h_g^2 / M = \sigma_u^2$ so that the mean of the χ^2 assocation statistics from MLMi analyses is

λ_{mean} (MLMi) = $(\sigma_u^2 + \sigma_\varepsilon^2 / N) / (\sigma_u^2 + \sigma_\varepsilon^2 / N) = 1$ at all SNPs,

λ_{mean} (MLMi) = $(h_g^2 / M_q + \sigma_\varepsilon^2 / N) / (\sigma_u^2 + \sigma_\varepsilon^2 / N) = [Nh_g^2 / M_q + (1 - r^2 h_g^2)] / [Nh_g^2 / M + (1 - r^2 h_g^2)]$ at the causal markers, and

λ_{mean} (MLMi) = $(\sigma_\varepsilon^2 / N) / (\sigma_u^2 + \sigma_\varepsilon^2 / N) = (1 - r^2 h_g^2) / [Nh_g^2 / M + (1 - r^2 h_g^2)]$ at the null markers.

We know from above that $r^2 = \theta / (\theta + 1 - h_g^2 r^2)$ so that there is no explicit solution for r^2 because $r^2 = [(1+\theta) \pm \sqrt{(1+\theta)^2 - 4h_g^2\theta}] / 2h_g^2$. We used simulations to validate that the simulation results were consistent with

$$r^2 = [(1+\theta) - \sqrt{(1+\theta)^2 - 4h_g^2\theta}] / 2h_g^2 \text{ with } \theta = Nh_g^2 / M \quad [S6]$$

If $h_g^2 = 1$ or $h_g^2 \to 0$, $r^2 = Nh_g^2 / M$, and if $M > N$ which is usually the case in practice, $r^2 \approx Nh_g^2 / M$.

Since λ_{mean} is expected to be 1 for at all the SNPs that are used to construct the GRM, λ_{median} at these SNPs will be smaller than 1. This indicates the dangers of using the genomic inflation factor (λ_{mean} or λ_{median}) to assess the presence of population stratification or other artifacts. A researcher observes lower λ_{mean} (or λ_{median}) for MLMi than that for LR might conclude that this is due to correction for confounding, but in fact this result is expected even in the absence of any confounding.

4.4. Mixed linear model association analysis excluding the candidate SNP. If the candidate SNP is not included in calculating the GRM, Equation [S3] can be re-written as

$$\mathbf{y}^* = \mathbf{y} - \hat{\mathbf{g}}^* = \mathbf{x}_t b_t + (\mathbf{g}^* - \hat{\mathbf{g}}^*) + \mathbf{e} = \mathbf{x}_t b_t + \boldsymbol{\varepsilon} \quad [S7]$$

Nature Genetics: doi:10.1038/ng.2876

which is analogue to a linear model of regressing $\mathbf{y}^*$ on $\mathbf{x}_t$. Therefore, the expected chi-squared test-statistic is simply

$$\mathrm{E}(\chi^2) = \mathrm{E}(\hat{b}_t^2) / \mathrm{var}(\hat{b}_t) = [b_t^2 + \mathrm{var}(\hat{b}_t)] / \mathrm{var}(\hat{b}_t) = 1 + b_t^2 / (\sigma_\varepsilon^2 / N) \quad [S8]$$

Therefore, the mean of the χ^2 assocation statistics from MLMe analyses is

λ_{mean} (MLMe) = $1 + (h_g^2 / M) / (\sigma_\varepsilon^2 / N) = 1 + (Nh_g^2 / M) / (1 - r^2 h_g^2)$ at all markers,

λ_{mean} (MLMe) = $1 + (h_g^2 / M_q) / (\sigma_\varepsilon^2 / N) = 1 + (Nh_g^2 / M_q) / (1 - r^2 h_g^2)$ at causal markers and

λ_{mean} (MLMe) = 1 at null markers.

where $r^2 \approx Nh_g^2 / M$ when $M > N$ (see above).

The ratio of λ_{mean} between MLMe and MLMi is also $1 + (Nh_g^2 / M) / (1 - r^2 h_g^2)$, which is consistent for causal, null and all markers. If $M >> N$ (i.e. r^2 is small), this is only slightly larger than $1 + Nh_g^2 / M$. Thus, λ_{mean} for MLMe (but not MLMi) are similar to that for LR and exceed 1 by an amount proportional to N / M, even in the absence of any confounding. The difference between MLMe and MLMi is that MLMe is testing the null hypothesis that the candidate marker has no effect, whereas MLMi is testing the null hypothesis that the candidate marker has an effect size drawn from a normal distribution $N(0, h_g^2/M)$. This issue with MLMi affects causal, null and all markers equally, and can be viewed as a miscalibration issue. However, if a fixed *P*-value threshold (e.g. 5×10^{-8}) is used (as in most GWAS), MLMi will report fewer truly associated markers as statistically significant as a consequence.

4.5. Linked markers. We note that the derivations above assume *M* unlinked markers. However, the derivations can be generalized to linked markers by replacing *M* with the effective number of independent markers (M_{eff}) (see above).

Nature Genetics: doi:10.1038/ng.2876

5. Using a small subset of markers in the GRM

5.1. Correcting for stratification using a subset of random markers. To investigate the number of random markers needed to correct for stratification, we conducted simulations of a quantitative trait with population stratification (see Methods section for the simulation strategy). We assumed N=10,000, M=100,000, two discrete subpopulations with F_{ST}=0.005, and a mean trait difference of 0.25 standard deviations between subpopulations. We applied MLMA methods with various values of the number M_R of random markers included in the GRM. Results displayed in Figure 2 and Supplementary Table 4 indicate that when there is subtle population stratification, a few thousand random markers are not sufficient to provide the best possible correction for stratification (Figure 2, Supplementary Table 4), consistent with previous studies[3,15]. Similar results were obtained under subtler stratification at F_{ST}=0.0025 (Supplementary Table 4). Although Lippert et al.[16] conducted analyses of WTCCC data[17] showing that a few thousand equally spaced markers was close to optimal when using MLMA to correct for inflation in test statistics, we caution that inflation in WTCCC data is likely to be primarily due to cryptic relatedness, not population stratification, as ref. [17] reported that inclusion of principal components as covariates reduced the inflation only slightly. Due to the large length of segments shared identical-descent, correcting for cryptic relatedness requires far fewer markers than correcting for population stratification. Although correcting for cryptic relatedness is an important goal, WTCCC data may not be suitable for drawing conclusions about correcting for population stratification, which may often require a very large number of markers.

5.2 Correcting for stratification using a subset of top associated markers. A separate question is whether using a subset of M_T markers with most significant linear regression P-values[2,10] is effective in correcting for population stratification. Listgarten et al.[10] state that the set of markers included in the GRM should be "exactly those SNPs that are associated with the phenotype, including causal SNPs, tag SNPs, and SNPs that are associated by way of confounding (population structure)". They state that "there is evidence that many SNPs

Nature Genetics: doi:10.1038/ng.2876

are undifferentiated (e.g. the fact that Ancestry Informative Marker panels typically number in the hundreds)". However, this may not be the case. As we previously noted, "A common misconception is that AIMs should be used to infer genetic ancestry even when genome-wide data are available, but in fact the best ancestry estimates are obtained using a large number of random markers"[18]. In human genetic data sets, all markers are differentiated, with population differentiation approximately following a normal distribution according to the value of F_{ST}[19]. Thus, while the simulations of ref.[10] in which only a fraction of markers are differentiated are of theoretical interest, they may have limited relevance to human genetic studies. Our simulations (in which all markers are differentiated) show that using the top M_T associated markers based on the first local minimum of the genomic control factor λ_{median} may not be effective in correcting for stratification, and can lead to a local minimum in λ_{median} that is different from the global minimum[2] (Figure 2). This limitation does not apply to data sets in which inflation is primarily due to cryptic and/or family relatedness, such as the WTCCC, Finnish and GAW14 data sets analyzed by ref. [10]. Thus, the effectiveness of the ref. [10] approach of using the M_T top associated markers achieving the first local minimum in λ_{median} may vary across different data sets. On the other hand, the ref. [2] approach of using the M_T top associated markers based on the global maximum of out-of-sample prediction accuracy selected M_T=M and thus provided an effective correction for stratification in these simulations (but see below).

5.3. Assessment of power in simulations without sample structure. MLMA can increase power even in studies without sample structure, by implicitly conditioning on associated loci other than the candidate locus that are not genome-wide significant in the data being analyzed. We generalized our simulations without sample structure, with fraction p=0.05 or p=0.005 of causal markers, to consider the impact of using the top M_T associated markers in the GRM, for various values of M_T. These simulations were motivated both by the ref. [10] approach of choosing M_T based on the first local minimum in λ_{median}, and by the ref. [2] approach of choosing M_T based on the global maximum of out-of-sample prediction accuracy. We also considered the impact of using

Nature Genetics: doi:10.1038/ng.2876

only the causal markers in the GRM, a cheating computation that is not applicable to real studies in which the set of causal markers is unknown. Results are displayed in Figure 3 and Supplementary Table 5. Including only causal markers in the GRM always performed best. Among the non-cheating computations, the optimal strategy varied with p. When p=0.05, there are a large number of causal markers with small effect sizes, so that the top M_T associated markers do not correspond to the true set of causal markers, and including all markers in the GRM ($M_T = M$) performed best. When p=0.005, there are a smaller number of causal markers with larger effect sizes, so that the top M_T associated markers more closely reflect the true set of causal markers, and including only a small subset of top markers in the GRM performed best. (We note that in each case, there was a severe loss in power when using the top M_T associated markers for large values of $M_T<M$. This is because top associated markers defined in-sample can explain most of the phenotypic variance under the null model—i.e. overestimate the true variance explained by these markers—making it difficult for the causal model including a candidate marker to outperform the null model[2].) Results for other values of M and N are displayed in Supplementary Table 5. Consistent with Figure 3, including all markers in the GRM is the optimal strategy when Mp/N is large and there is lower power to infer the true set of causal markers, and including only a small subset of markers in the GRM is the optimal strategy when Mp/N is small and there is higher power to infer the true set of causal markers. Thus, the optimal strategy depends on both the sample size and the genetic architecture of the trait. The ref. [2] approach of using the M_T top associated markers based on the global maximum of out-of-sample prediction accuracy selected the value of M_T that maximized power in each of these simulations, achieving the optimal strategy.

5.4. Simulations with both stratification and causal markers. see main text.

Nature Genetics: doi:10.1038/ng.2876

6. Mis-estimation of heritability explained by genotyped markers

MLM are also used to estimate the heritability explained by SNPs on a genotyping platform[4], and ref. [7] provides a formula to convert from heritability explained on the observed scale to the liability scale for case-control phenotypes. We conducted an additional set of experiments to examine the effect of ascertainment on estimates of heritability explained. We estimated the heritability explained on the observed scale for case-control data sets generated as previously described in our simulations of case-control traits with p=0.05. For each parameter setting we randomly selected 100 data sets and estimated the heritability explained using the whole-genome estimation approach implemented in GCTA software[4]. The results shown in Supp Table 9 demonstrate that there is a significant downward bias in heritability estimate as a result of ascertainment when N/M is large. The infinitesimal model of MLM assumes independent SNPs so this bias could be due to the induced LD between associated genetic variants in ascertained data[20,21]. The drop in power is not due to the use of the incorrect value of heritability estimated from the data. We reran the experiments using the expected heritability on the observed scale and found a similar loss in power.

Nature Genetics: doi:10.1038/ng.2876

7. Acknowledgments for data from dbGaP

Funding support for the GWAS of Gene and Environment Initiatives in Type 2 Diabetes was provided through the NIH Genes, Environment and Health Initiative [GEI] (U01HG004399). The human subjects participating in the GWAS derive from The Nurses' Health Study (NHS) and Health Professionals' Follow-up Study (HPFS) and these studies are supported by National Institutes of Health grants CA87969, CA55075, and DK58845. Assistance with phenotype harmonization and genotype cleaning, as well as with general study coordination, was provided by the Gene Environment Association Studies, GENEVA Coordinating Center (U01 HG004446). Assistance with data cleaning was provided by the National Center for Biotechnology Information. Funding support for genotyping, which was performed at the Broad Institute of MIT and Harvard, was provided by the NIH GEI (U01HG004424).

The Atherosclerosis Risk in Communities Study (ARIC) is carried out as a collaborative study supported by National Heart, Lung, and Blood Institute contracts (HHSN268201100005C, HHSN268201100006C, HHSN268201100007C, HHSN268201100008C, HHSN268201100009C, HHSN268201100010C, HHSN268201100011C, and HHSN268201100012C), R01HL087641, R01HL59367 and R01HL086694; National HumanGenome Research Institute contract U01HG004402; and National Institutes of Health contract HHSN268200625226C. The authors thank the staff and participants of the ARIC study for their important contributions. Infrastructure was partly supported by Grant Number UL1RR025005, a component of the National Institutes of Health and NIH Roadmap for Medical Research.

Nature Genetics: doi:10.1038/ng.2876

References

1. Yang, J. *et al.* Genome partitioning of genetic variation for complex traits using common SNPs. *Nat Genet* **43**, 519-25 (2011).
2. Lippert, C. *et al.* The benefits of selecting phenotype-specific variants for applications of mixed models in genomics. *Sci Rep* **3**, 1815 (2013).
3. Price, A.L. *et al.* Principal components analysis corrects for stratification in genome-wide association studies. *Nat Genet* **38**, 904-9 (2006).
4. Yang, J. *et al.* Common SNPs explain a large proportion of the heritability for human height. *Nat Genet* **42**, 565-9 (2010).
5. 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).
6. Kang, H.M. *et al.* Variance component model to account for sample structure in genome-wide association studies. *Nat Genet* **42**, 348-54 (2010).
7. Lee, S.H., Wray, N.R., Goddard, M.E. & Visscher, P.M. Estimating missing heritability for disease from genome-wide association studies. *Am J Hum Genet* **88**, 294-305 (2011).
8. Sawcer, S. *et al.* Genetic risk and a primary role for cell-mediated immune mechanisms in multiple sclerosis. *Nature* **476**, 214-9 (2011).
9. Jostins, L. *et al.* Host-microbe interactions have shaped the genetic architecture of inflammatory bowel disease. *Nature* **491**, 119-24 (2012).
10. Listgarten, J. *et al.* Improved linear mixed models for genome-wide association studies. *Nat Methods* **9**, 525-6 (2012).
11. Yang, J. *et al.* Genomic inflation factors under polygenic inheritance. *Eur J Hum Genet* **19**, 807-12 (2011).
12. Pe'er, I., Yelensky, R., Altshuler, D. & Daly, M.J. Estimation of the multiple testing burden for genomewide association studies of nearly all common variants. *Genet Epidemiol* **32**, 381-5 (2008).
13. Devlin, B. & Roeder, K. Genomic control for association studies. *Biometrics* **55**, 997-1004 (1999).
14. Daetwyler, H.D., Villanueva, B. & Woolliams, J.A. Accuracy of predicting the genetic risk of disease using a genome-wide approach. *PLoS One* **3**, e3395 (2008).
15. Zhou, X. & Stephens, M. Genome-wide efficient mixed-model analysis for association studies. *Nat Genet* **44**, 821-4 (2012).
16. Lippert, C. *et al.* FaST linear mixed models for genome-wide association studies. *Nat Methods* **8**, 833-5 (2011).
17. Wellcome Trust Case Control Consortium. Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. Nature 2007; 447:661-78.
18. Price, A.L., Zaitlen, N.A., Reich, D. & Patterson, N. New approaches to population stratification in genome-wide association studies. *Nat Rev Genet* **11**, 459-463 (2010).
19. Bhatia, G. *et al.* Genome-wide comparison of African-ancestry populations from CARe and other cohorts reveals signals of natural selection. *Am J Hum Genet* **89**, 368-81 (2011).
20. Zaitlen, N. *et al.* Analysis of case-control association studies with known risk variants. *Bioinformatics* **28**, 1729-1737 (2012).
21. Zaitlen, N. & Kraft, P. Heritability in the genome-wide association era. *Hum Genet* **131**, 1655-64 (2012).

Nature Genetics: doi:10.1038/ng.2876