Supplementary Table 3. Cohort-level quality control

Cohort	Platform	Starting samples[c]	Starting SNVs[d]	Samples CR > 0.95, deduplicated	Samples no sex mismatch[e]	SNVs pass cohort QC[f]	Samples unrelated, het |F| < 0.2
Discovery							
cPRG[a]	IlluminaGSA	9863	695905	9469	9431	555560	9296
cSWE1	IlluminaGSA	4194	691954	4079	4022	566458	3891
cNOR	IlluminaGSA	542	691859	542	542	525995	529
cBEL1[b]	IlluminaGSA	347	498725	347	347	466914	347
Replication							
cBEL2	Illumina670	174	561353	174	174	549684	174
cDEN	Illumina660	328	559414	328	328	546859	328
cUKC	Illumina670	666	553258	666	n.a.	536876	666
cGER1	Illumina670	1053	567693	1053	1050	549926	1045
cGER2	IlluminaOmniExpress	2681	712283	2659	2637	607925	2611
cIT1	Illumina550	370	493211	370	n.a.	492476	370
cIT2.1	Illumina550	320	493210	320	n.a.	492277	320
cIT2.2	IlluminaOmniExpress	1146 (774)	720169	1135	1116	641463	240
cIT2.3	IlluminaOmni2.5	317	2235110	291	269	1806623	1107
cNOR1	Illumina670	131	545752	131	n.a.	537003	131
cNOR2	IlluminaGSA	825	548000	825	822	491820	809
cNOR3	IlluminaOmniExpress	138 (98)	658234	138	137	536865	64
cNOR4	IlluminaOmniExpressExome	64 (31)	670362	64	64	640833	137
cSP2	IlluminaMEGA	288	978314	288	278	793673	277
cSWE2	IlluminaGSA	5743 (4194)	583755	5710	5590	498003	5473
cSWE3	IlluminaOmniExpress	13651 (7258)	704358	13651	13648	641875	13496
cSWE4	Illumina660	1362 (732)	509975	1362	1362	506050	1337

[a] Samples from this cohort (internally referred to as 'Progression GWAS [PRG]') were genotyped at the Hussman Institute for Human Genomics Genotyping Core, University of Miami (n = 5430) and the Institute for Molecular Medicine Finland (FIMM), University of Helsinki (n = 4433); genoyptes were called jointly.

[b] This dataset was subjected to local quality control prior to inclusion in this study (PMID: 33704824). To ensure consistency between cohorts, the same quality control filters were still applied.

[c] Where the sample count included controls (only for quality control and removed prior to analysis), the multiple sclerosis case count is indicated in parenthesis

[d] Count excludes mitochondrial variants and indels. Genomic positions were assigned on the same build and alleles oriented to the forward strand

[e] n.a. indicates that the cohort did not include sex chromosome data, and therefore a mismatch between genetic and reported sex could not be assessed

[f] See Supplementary Note for the details of tests performed for cohort-level variant QC

CR, call rate; het |F|, absolute inbreeding coefficient estimate F