Supplementary Information

Multiple apical plasma membrane constituents are associated with susceptibility to meconium ileus in individuals with cystic fibrosis

Lei Sun*[1,2], Johanna M Rommens*[3,4], Harriet Corvol[5,6], Weili Li[1,7], Xin Li[7], Theodore Chiang[7], Fan Lin[3], Ruslan Dorfman[3], Pierre-François Busson[8], Rashmi V Parekh[3], Diana Zelenika[9], Scott Blackman[10], Mary Corey[7], Vishal Doshi[11], Lindsay Henderson[11], Kathleen Naughton[11], Wanda K O'Neal[12], Rhonda G Pace[12], Jaclyn R Stonebraker[12], Sally D Wood[12], Fred A Wright[13], Julian Zielenski[3], Annick Clement[5,6], Mitchell L Drumm[14], Pierre-Yves Boëlle[6,8], Garry R Cutting[10,11], Michael R Knowles[12], Peter R Durie[15,16], Lisa J Strug[7,1]

Correspondence should be addressed to Lisa Strug: lisa.strug@utoronto.ca

* These authors contributed equally to this work

[1]Biostatistics Division, Dalla Lana School of Public Health, University of Toronto, Toronto, Ontario, Canada.
[2]Department of Statistics, University of Toronto, Toronto, Ontario, Canada.
[3]Program in Genetics and Genome Biology, the Hospital for Sick Children, Toronto, Ontario, Canada.
[4]Department of Molecular Genetics, University of Toronto, Toronto, Ontario, Canada.
[5]AP-HP, Hospital Trousseau, Pediatric Pulmonary Department, Inserm U938, Paris, France
[6]Pierre et Marie Curie University-Paris 6, Paris, France.
[7]Child Health Evaluative Sciences, the Hospital for Sick Children, Toronto, Ontario, Canada.
[8]AP-HP, St Antoine Hospital, Biostatistics Department, Inserm UMR-S707, Paris, France.
[9]Centre National de Génotypage, CEA, Evry, France.
[10]Department of Pediatrics, Johns Hopkins University School of Medicine, Baltimore, Maryland, USA.
[11]McKusick-Nathans Institute of Genetic Medicine, Johns Hopkins University School of Medicine, Baltimore, Maryland, USA.
[12]Cystic Fibrosis-Pulmonary Research and Treatment Center, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina, USA.
[13]Department of Biostatistics, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina, USA.
[14]Departments of Pediatrics and Genetics, Case Western Reserve University, Cleveland, Ohio, USA.
[15]Program in Physiology and Experimental Medicine, the Hospital for Sick Children, Toronto, Ontario, Canada.
[16]Department of Pediatrics, University of Toronto, Toronto, Ontario, Canada.

Table of Contents

Supplementary Tables:

Supplementary Figures:

Supplementary Table 1. Sex-specific results for rs3788766 and rs5905283 in *SLC6A14*; meconium ileus (MI).

	Risk Allele	Risk Allele Frequency		Sex (Coding of the Risk Allele)	Sample Size (by Genotype)	*P* value	OR	OR 95%CI	
		MI	Non-MI						
rs3788766	T	0.737	0.591						
				Female (1 vs. 2)	816 vs. 671	3.20×10^{-5}	1.79	1.36	2.37
				Female (0 vs. 2)	255 vs. 671	2.80×10^{-3}	1.89	1.25	2.87
				Male (0 vs. 2)	814 vs. 1207	2.10×10^{-9}	2.31	1.75	3.04
				All (0 vs. 1 vs. 2)	1069 vs. 816 vs. 1878	1.28×10^{-12}	1.49	1.34	1.67
rs5905283	C	0.605	0.505						
				Female (1 vs. 2)	849 vs. 488	5.50×10^{-4}	1.66	1.25	2.22
				Female (0 vs. 2)	405 vs. 488	8.96×10^{-3}	1.61	1.13	2.29
				Male (0 vs. 2)	1014 vs. 1007	3.80×10^{-7}	1.91	1.49	2.45
				All (0 vs. 1 vs. 2)	1419 vs. 849 vs. 1495	1.69×10^{-8}	1.33	1.21	1.49

Supplementary Table 2. Association results for SNPs in *SLC6A14* and *SLC26A9* with and without adjustment for *CFTR*. Without adjustment for *CFTR* indicates the results of original GWAS with the site covariate only as in **Figure 1a** and **Table 2**. With adjustment for *CFTR* indicates the results with both the site covariate and the *CFTR* covariate for which Phe508del/Phe508del genotype is coded as 1 and Phe508del/Other or Other/Other genotypes are coded as 0.

SNP	CHR	POS	GENE	Risk Allele	Without adjustment for *CFTR*			With adjustment for *CFTR*		
					Estimate	SE	*P*	Estimate	SE	*P*
rs4077468	1	204181380	*SLC26A9*	T	0.372	0.065	9.88×10^{-9}	0.373	0.065	9.44×10^{-9}
rs7512462	1	204166218	*SLC26A9*	T	0.371	0.066	2.14×10^{-8}	0.371	0.066	2.03×10^{-8}
rs7419153	1	204183932	*SLC26A9*	T	0.347	0.065	1.01×10^{-7}	0.345	0.065	1.22×10^{-7}
rs12047830	1	204183322	*SLC26A9*	C	0.293	0.063	3.72×10^{-6}	0.295	0.063	3.26×10^{-6}
rs3788766	X	115480867	*SLC6A14*	T	0.402	0.057	1.28×10^{-12}	0.405	0.057	1.02×10^{-12}
rs5905283	X	115479909	*SLC6A14*	C	0.294	0.052	1.69×10^{-8}	0.296	0.052	1.42×10^{-8}
rs12839137	X	115479578	*SLC6A14*	C	0.331	0.068	1.20×10^{-6}	0.335	0.068	1.02×10^{-6}

Supplementary Table 3. *SLC6A14* **and 157 apical genes with gene-based and Lasso association results.** A list of 157 genes was annotated as described in **Figure 2**.

[a] The number of GWAS SNPS within ±10 kb of the boundaries of indicated gene.

[b] Permutation based *P* value for the gene. See Online Methods for details.

[c] North American CF Gene Modifier Consortium (NACFGMC) discovery cohort (n=3,763).

[d] French replication cohort (n=1,232); For *SLC6A14* and *SLC26A9*, the combined replication results were used when appropriate, combining the French sample with the independent North American sample (n=1,140, **Table 2**).

[e] Indicating if the gene was in the multivariate model selected by Lasso in the discovery sample; 1=yes, 0=no.

[f] Two apical genes (*MSN* and *SLC34A3* at the end of the table) that were not tagged by GWAS SNPs passing the QC criteria.

[g] *CFTR* showed some evidence for association with meconium ileus possibly reflecting the CF mutations that are present.

[h] For *MUC20*, there were no GWAS SNPs passing the QC criteria in the French data.

[i] Evidence for association of *SLC26A9* to meconium ileus was apparent in the North American portion of the replication sample, but was not in the French cohort portion. It remains unclear if this difference is a reflection of chance and limited power due to the smaller sample size of the French cohort, or if there are actual differences in the studied populations.

[j] *SLC6A14* was not annotated by AmiGO on March 28, 2010 as localized to the apical plasma membrane, thus it was not included in the analysis restricted to the apical gene list

GENE	CHR	# SNPs[a]	Gene *P* value[b]		Lasso[e]
			Discovery[c]	Replication[d]	
SLC6A14[j]	X	15	<0.0001	<0.001	0
ABCA7	19	18	0.9832	0.03	0
ABCB4	7	43	0.5761	0.344	1
ABCC2	10	34	0.2613	0.835	0
ABCC6	16	31	0.0565	0.628	1
ABCG5	2	26	0.2233	0.676	0
ABCG8	2	41	0.0261	0.23	1
ACY3	11	7	0.6527	0.305	0
ADAM17	2	9	0.3567	0.228	0
ADRB2	5	15	0.4085	0.304	0
AJAP1	1	50	0.788	0.859	0
AKAP7	6	29	0.6997	0.286	0
AKR1A1	1	5	0.5602	0.168	0
ANK2	4	116	0.6156	0.85	1
ANXA6	5	53	0.4416	0.369	1
AQP1	7	10	0.8664	0.489	0
AQP2	12	6	0.5037	0.596	0
AQP5	12	4	0.4272	0.402	0
AQP8	16	6	0.9199	0.56	0
ATP1B1	1	22	0.8975	0.992	0
ATP2B2	3	62	0.0006	0.283	1

ATP4A	19	11	0.5505	0.763	0
ATP6V0A4	7	39	0.8663	0.95	0
ATP6V0D1	16	4	0.1689	0.956	0
ATP6V0D2	8	20	0.1898	0.918	0
ATP6V1B1	2	13	0.6323	0.338	0
ATP6V1E1	22	12	0.8157	0.045	0
ATP8B1	18	27	0.0123	0.108	1
CA4	17	6	0.1653	0.267	0
CACNB3	12	4	0.8725	0.417	0
CAV1	7	15	0.1856	0.604	1
CD300LG	17	3	0.6072	0.611	0
CD44	11	53	0.3061	0.168	1
CDHR2	5	9	0.167	0.596	0
CFTR[g]	7	49	0.0053	0.001	0
CIB1	15	6	0.5588	0.607	0
CLCA4	1	8	0.3229	0.466	0
CRB1	1	27	0.3027	0.179	0
CRB3	19	6	0.3403	0.763	0
CSPG4	15	5	0.1625	0.43	1
CTSB	21	6	0.0021	0.201	1
CUBN	10	144	0.2361	0.393	1
DPEP1	16	2	0.0946	0.755	0
DPP4	2	12	0.3709	0.591	1
DUOX1	15	13	0.7363	0.317	0
DUOX2	15	6	0.7511	0.267	0
EGFR	7	78	0.8477	0.041	1
ENPEP	4	26	0.5939	0.757	0
ERBB2	17	2	0.0701	0.853	1
ERBB3	12	4	0.3871	0.68	0
EZR	6	10	0.0317	0.023	0
F2RL2	5	19	0.7075	0.188	0
GIF	11	1	0.4834	0.039	0
GJB6	13	6	0.8792	0.939	0
GNAT3	7	16	0.0365	0.957	0
GPR64	23	20	0.6066	0.51	0
IGFBP2	2	7	0.1872	0.586	0
IL6R	1	9	0.2712	0.031	0
INADL	1	100	0.6673	0.76	0
KCNA1	12	8	0.1868	0.638	0
KCNE1	21	28	0.9999	0.981	0
KCNMA1	10	225	0.0955	0.129	1
KIAA1919	6	5	0.5382	0.037	0
KNCN	1	4	0.9284	0.664	0
LCT	2	11	0.799	0.036	0
LHFPL5	6	4	0.3254	0.551	0
LMO7	13	56	0.0267	0.782	1
LRP2	2	65	0.2323	0.964	1
LZTS1	8	10	0.1894	0.072	0

MAL	2	3	0.9152	0.869	0
MAL2	8	7	0.6686	0.789	0
MGAM	7	27	0.2256	0.191	1
MIP	12	5	0.914	0.768	0
MPDZ	9	31	0.788	0.781	0
MREG	2	37	0.819	0.719	0
MUC1	1	3	0.3971	0.887	0
MUC20[h]	3	1	0.909	NA	0
MUC3A	2	92	0.1207	0.498	1
MUC3B	7	5	0.4523	0.23	0
MYO1A	12	5	0.338	0.632	0
MYO7B	2	31	0.7972	0.575	0
NAALADL1	11	6	0.456	0.878	0
NPC1L1	7	6	0.4787	0.014	0
NRG1	8	260	0.8421	0.781	1
OTOA	16	8	0.2995	0.02	0
OTOG	11	25	0.5511	0.285	0
OXTR	3	30	0.8667	0.65	0
P2RY1	3	2	0.2739	0.569	0
P2RY2	11	11	0.5776	0.869	1
P2RY4	23	5	0.7478	0.535	0
P2RY6	11	11	0.6034	0.327	0
PDPN	1	25	0.2071	0.244	0
PFKM	12	8	0.2429	0.616	0
PKHD1	6	84	0.0475	0.356	0
PLB1	2	56	0.8802	0.972	0
PRKCI	3	8	0.639	0.756	0
PROM1	4	59	0.2262	0.39	1
PROM2	2	2	0.4913	0.325	0
PTK2	8	38	0.079	0.088	1
RAB14	9	7	0.3595	0.214	0
RAB27A	15	25	0.3257	0.583	0
RHCG	15	5	0.1909	0.581	1
S100G	23	2	0.3144	0.916	0
SCNN1A	12	16	0.3261	0.288	0
SCNN1B	16	18	0.4818	0.826	0
SCNN1G	16	7	0.8408	0.911	0
SHROOM2	23	42	0.5546	0.178	1
SHROOM3	4	92	0.4466	0.224	1
SHROOM4	23	13	0.5186	0.778	1
SI	3	10	0.7507	0.201	0
SLC10A2	13	35	0.9229	0.433	0
SLC11A2	12	8	0.3498	0.971	0
SLC12A2	5	18	0.6973	0.905	0
SLC12A3	16	31	0.6003	0.702	0
SLC14A2	18	57	0.1743	0.724	1
SLC22A11	11	4	0.7337	0.533	0
SLC22A12	11	4	0.0451	0.863	0

|					
SLC22A18	11	14	0.9952	0.154	0
SLC22A4	5	47	0.1902	0.663	0
SLC22A5	5	24	0.1561	0.375	1
SLC23A1	5	1	0.2446	0.887	0
SLC23A2	20	48	0.0884	0.218	1
SLC26A3	7	19	0.5823	0.613	0
SLC26A4	7	21	0.9859	0.997	0
SLC26A9[i]	1	39	<0.0001	0.235	1
SLC29A4	7	3	0.0703	0.754	1
SLC30A5	5	12	0.8699	0.907	0
SLC34A2	4	12	0.5698	0.764	0
SLC39A4	8	2	0.1793	0.492	0
SLC3A2	11	7	0.115	0.978	0
SLC46A1	17	3	0.5825	0.145	0
SLC4A5	2	40	0.5031	0.729	0
SLC4A7	3	22	0.5563	0.05	0
SLC5A1	22	15	0.8685	0.435	0
SLC5A12	11	45	0.2398	0.081	1
SLC5A8	12	24	0.5075	0.044	0
SLC6A20	3	24	0.2724	0.564	0
SLC7A5	16	16	0.6744	0.607	0
SLC9A3	5	10	0.0001	0.017	1
SLC9A3R1	17	8	0.5838	0.625	0
SLC9A3R2	16	10	0.2069	0.091	0
SLC9A4	2	55	0.8562	0.012	0
STK39	2	64	0.9003	0.112	0
STX3	11	5	0.4133	0.71	0
STX4	16	2	0.2168	0.632	0
STXBP3	1	12	0.6591	0.502	0
TACSTD1	2	7	0.2121	0.594	0
TCIRG1	11	11	0.6047	0.718	0
TF	3	13	0.7929	0.498	0
TGFBR1	9	10	0.4879	0.746	0
TLR9	3	4	0.0715	0.313	0
TRPM6	9	39	0.4276	0.723	0
TRPV5	7	6	0.3986	0.231	0
UMOD	16	14	0.2544	0.009	0
UPK3A	22	10	0.5462	0.024	0
VANGL2	1	10	0.5457	0.284	0
MSN [f]	23	0	NA	NA	NA
SLC34A3 [f]	9	0	NA	NA	NA

Supplementary Figure 1. **QQ-Plot of GWAS results prior to prioritization**. The original GWAS results without prioritization from the GEE model adjusting for the site covariate using the full 3,763 North American samples (see **Figure 1a** for the corresponding Manhattan plot). No violation of model assumptions is evident, and there is significant evidence of association.
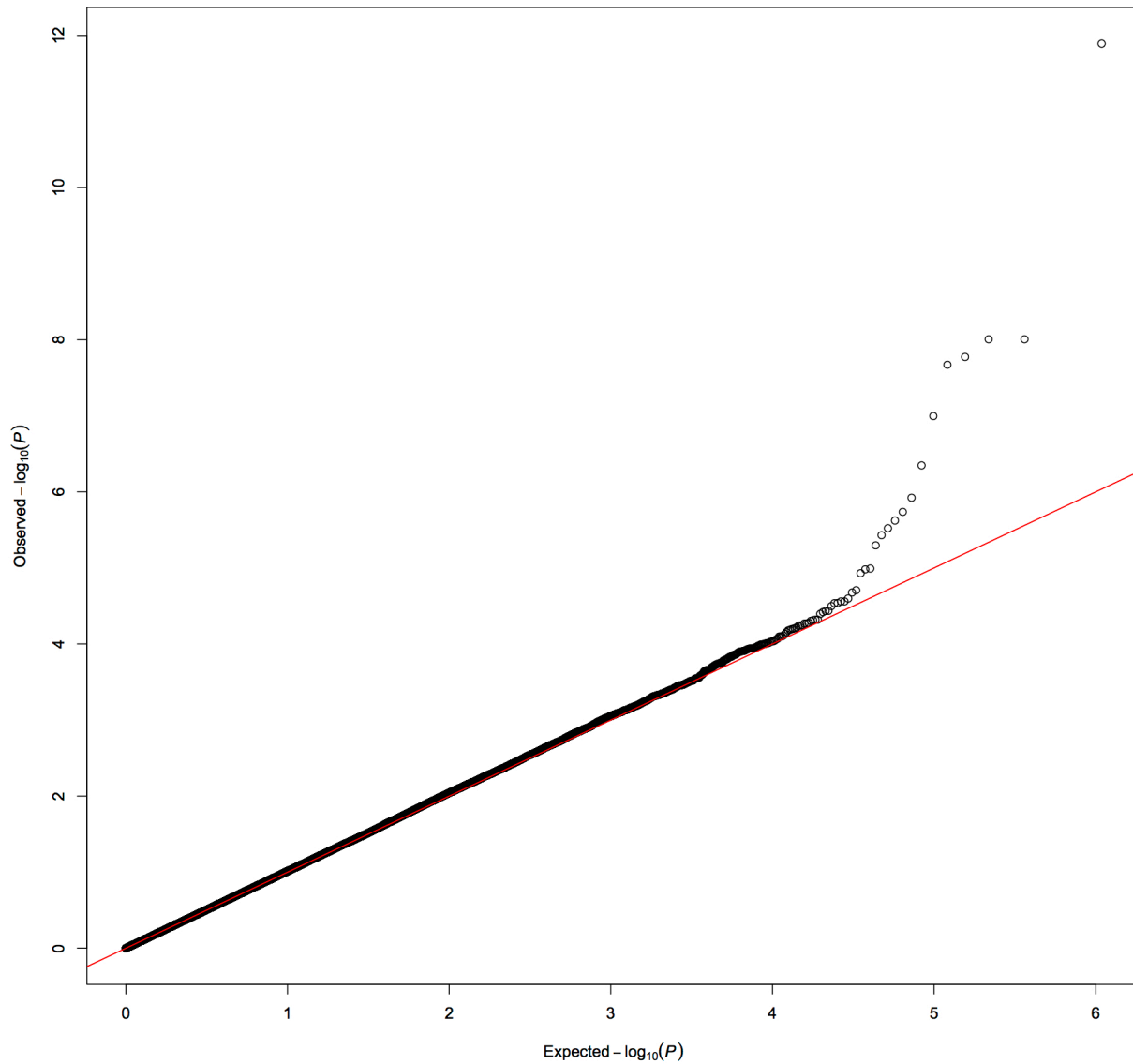
Supplementary Figure 2. **GWAS results with and without adjusting for the effect of** **_CFTR_.** The x-axis shows the association *P* values (on the -log10 scale) of the original GWAS with the site covariate but without adjusting for the effect of *CFTR* as in **Figure 1a**; the y-axis shows the association *P* values with both the site covariate and the *CFTR* covariate for which Phe508del/Phe508del genotype is coded as 1 and Phe508del /Other or Other/Other genotypes are coded as 0. SNPs within 155 kb of *CFTR* have been removed from this figure, and the SNPs at the bottom-left that have some noticeable discrepancy between the two sets of analyses are the SNPs that are in LD with *CFTR*.
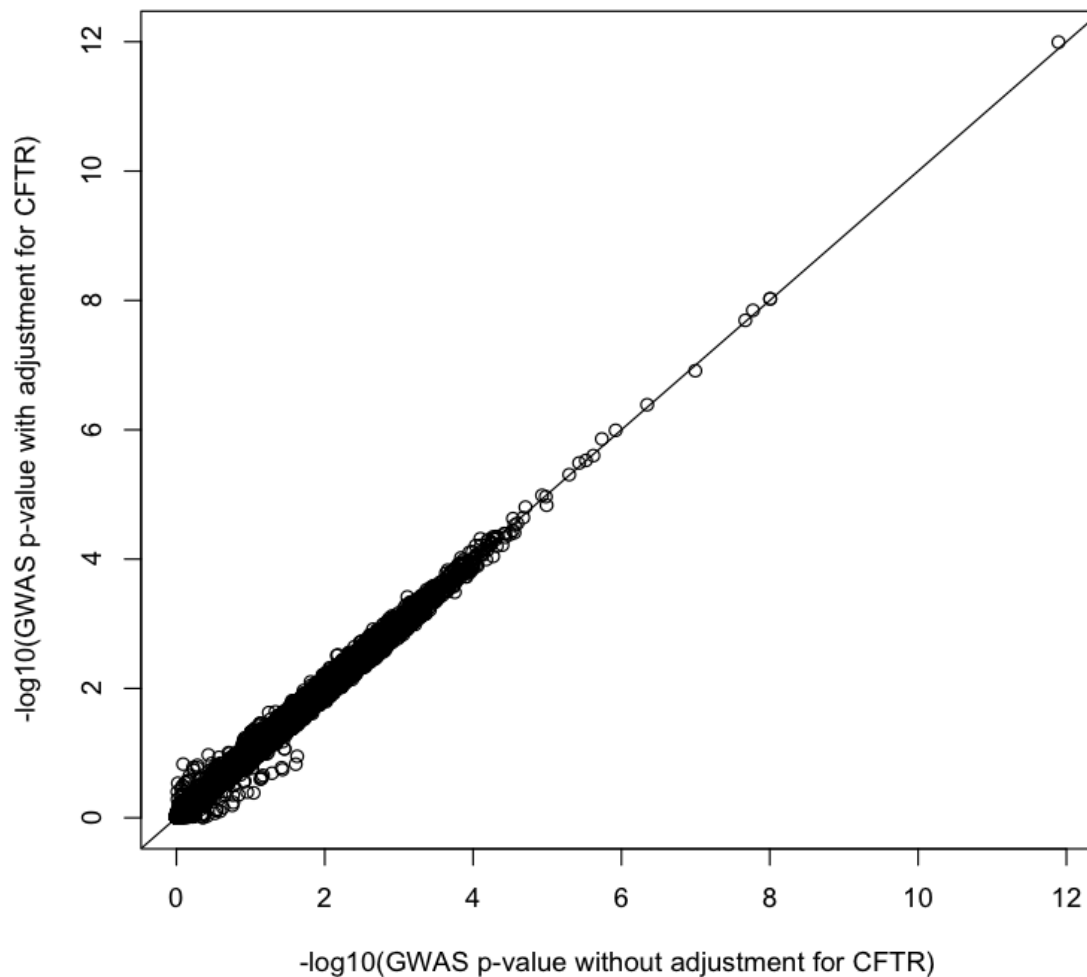
Supplementary Figure 3. Manhattan plot of GWAS with 1000 genomes imputation of 7,245,292 SNPs. The reference sample was the 87 CEU subjects extracted from the EUR continental group of the 1000 genomes November 2010 release provided in the four-site (Broad Institute, Michigan University, Boston College and NCBI) merged dataset. The figure only include 78,352 SNPs with *P* values<0.01, with solid circles representing 5,953 genotyped SNPs and empty circles representing 72,399 imputed SNPs.
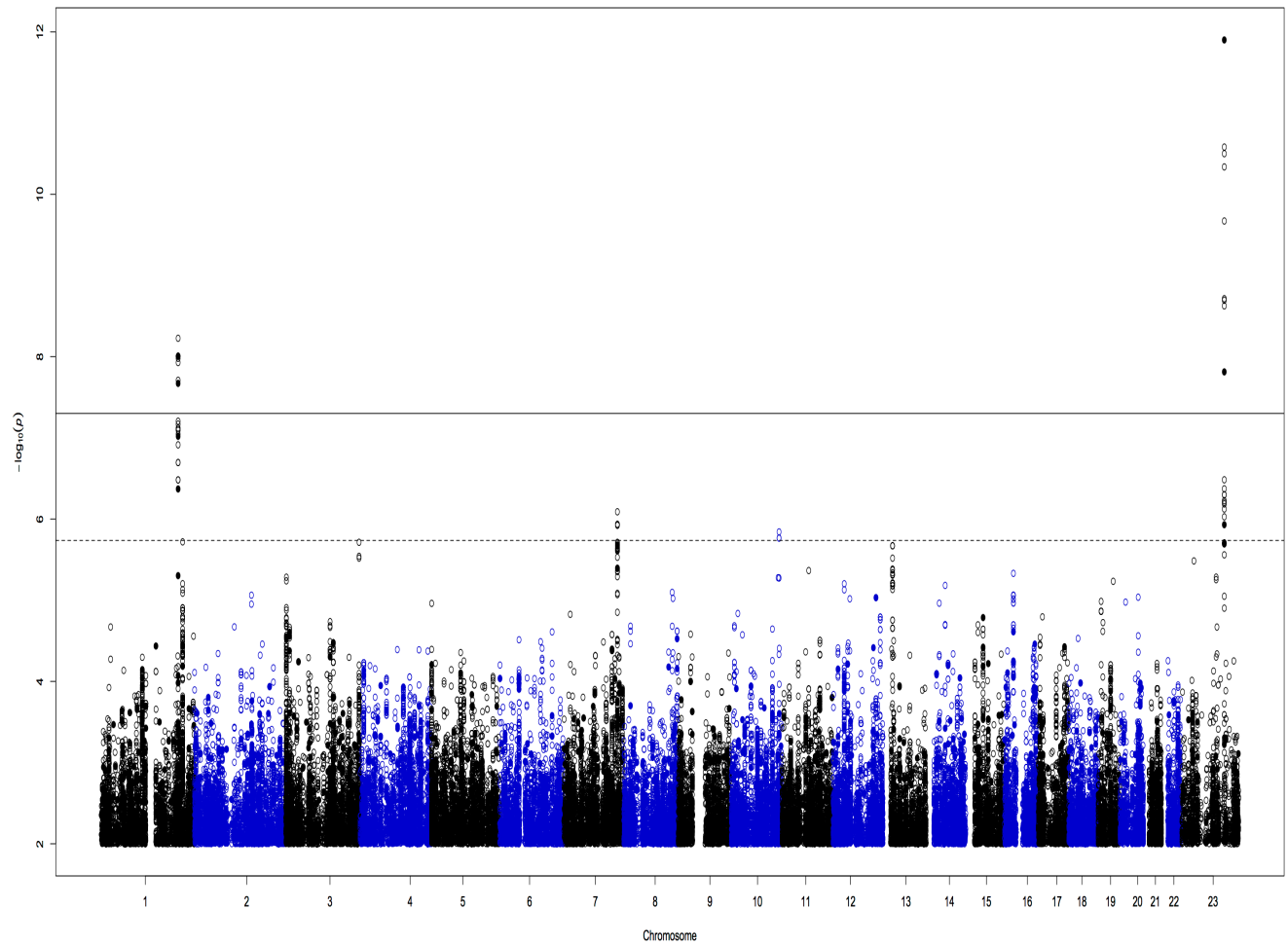
Supplementary Figure 4. GWAS-HD flow chart. See Online Methods for details on the single-SNP and multi-SNP/gene analyses.
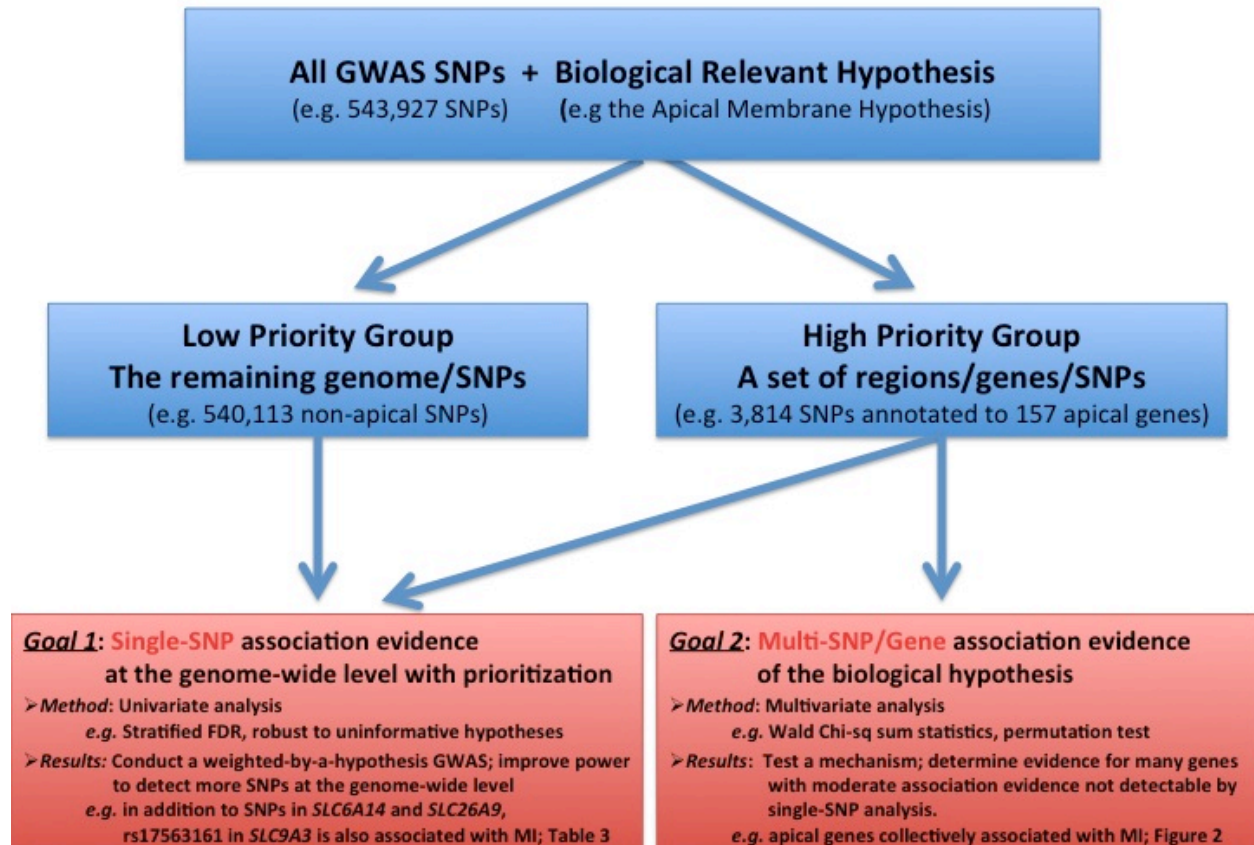
Supplementary Figure 5. SNP rankings from GWAS vs. GWAS-HD. GWAS ranks are based on the original association evidence alone and GWAS-HD ranks are based on SFDR weight-adjusted *q* values after incorporating the apical plasma membrane hypothesis (Online Methods).