

Rough Sets for Phenotype-Based Prioritization of Causative Variants

Jyoti Sharma, Khadija Sana Hafeez and Sushmita Paul

The association results (top 30 SNPs) computed from PLINK1.9 are displayed in Table 1.

The SNPs obtained through the proposed pipeline additionally show associations with Coronary Artery Disease (CAD), Type 2 Diabetes, hypertension, BMI, smoking addiction problems, and several other complex diseases linked with variations in blood cholesterol levels. The aforementioned details have been obtained from the GRASP database, and additional details are provided in Table S2. The NDEx Integrated Query v1.3 (www.ndexbio.org) is used to identify tumors linked to the genes that have been annotated. As demonstrated in Figure S1, the NDEx tool aids in determining the link between acquired genes and various cancers. Figure S2 shows a comparison of the annotated genes and enriched cancer-related genes identified by all three association methods applied in this study. As can be observed from the graph, RBFS outperforms other approaches.

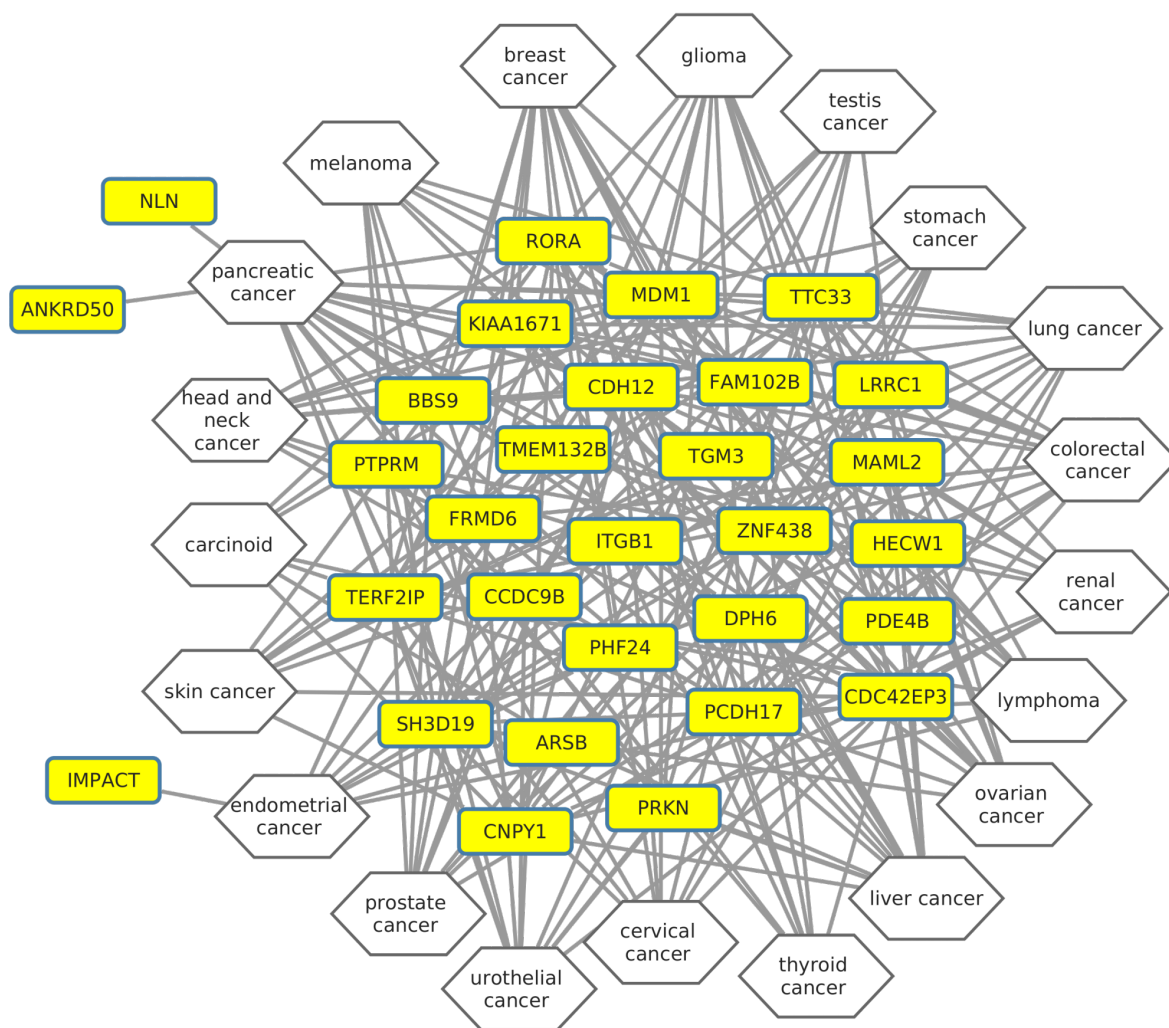


Figure S1: Cancer-related genes identified by the proposed pipeline

Table 1: Association results using PLINK1.9

CHR	SNP	BP	NMISS	BETA	SE	R2	T	P
14	kgp11175112	0	316	-0.4653	0.0925	0.07459	-5.031	0.000000824
2	kgp7149728	0	316	-0.2197	0.0491	0.05993	-4.474	0.00001073
6	kgp8685782	0	316	-0.312	0.0702	0.05918	-4.444	0.00001225
2	exm-rs2278729	0	316	-0.2795	0.06291	0.05914	-4.443	0.00001232
16	kgp1635892	0	316	0.1996	0.04508	0.05874	4.427	0.00001321
6	kgp10258598	0	316	-0.3886	0.08781	0.05869	-4.425	0.00001333
8	rs11786580	0	316	-0.3323	0.07536	0.05831	-4.41	0.00001424
9	rs17611535	0	316	-0.3651	0.08357	0.05731	-4.369	0.00001698
12	rs4831955	0	316	-0.1979	0.04622	0.05515	-4.281	0.00002473
8	kgp22756774	0	316	0.1955	0.04634	0.05366	4.22	0.00003205
12	kgp4378118	0	316	-0.1956	0.04675	0.0528	-4.184	0.00003725
1	rs6676664	0	316	0.2865	0.06924	0.0517	4.137	0.00004514
3	rs11914552	0	316	0.3244	0.07878	0.05124	4.118	0.00004892
3	rs1523060	0	316	0.1878	0.04564	0.05119	4.116	0.00004938
4	rs7670760	0	316	-0.2128	0.0524	0.04991	-4.061	0.00006166
1	rs11209464	0	316	-0.1815	0.04477	0.04973	-4.054	0.00006361
12	rs2393073	0	316	-0.3369	0.08338	0.04942	-4.04	0.00006718
2	rs834128	0	316	0.1781	0.04455	0.04841	3.997	0.00008003
9	rs7038474	0	316	0.1712	0.04289	0.04829	3.992	0.00008177
10	kgp9119184	0	316	-0.3614	0.0907	0.04813	-3.985	0.00008407
5	rs12659858	0	316	-0.293	0.07354	0.04811	-3.984	0.00008439
12	kgp1947877	0	316	-0.4031	0.1013	0.04798	-3.978	0.0000863
13	rs9285165	0	316	0.1968	0.04946	0.04798	3.978	0.00008631
5	kgp7304412	0	316	-0.2334	0.05878	0.04783	-3.971	0.00008863
8	kgp4486162	0	316	0.1769	0.04456	0.0478	3.97	0.00008907
9	kgp10441784	0	316	0.1681	0.04243	0.04762	3.962	0.00009187
13	kgp8797368	0	316	-0.3683	0.09338	0.04719	-3.944	0.00009905
3	kgp17674859	0	316	0.3089	0.07842	0.04709	3.939	0.0001008
20	kgp5617879	0	316	0.3354	0.08535	0.04686	3.929	0.0001049
18	kgp610372	0	316	-0.1705	0.04343	0.04679	-3.926	0.0001063

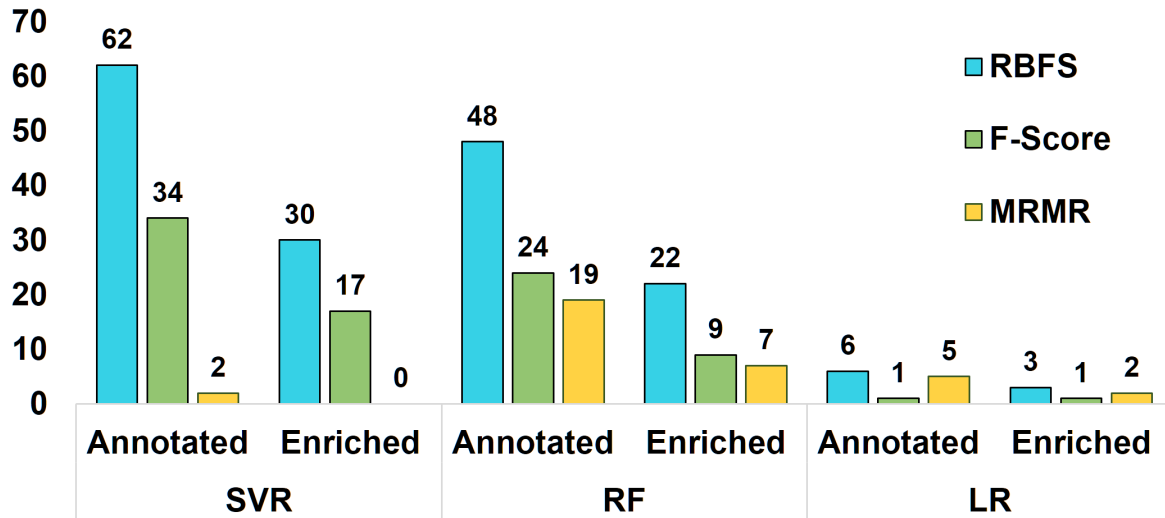


Figure S2: Comparative analysis of annotated genes and enriched cancer-related genes identified by RF, LR, and SVR methods, respectively

Table S2: Details of SNPs associated with various cholesterol-related complex diseases.

SNP	p-value	PMID	Phenotype	Gene
rs567279	4.10E-21	23063622	LDL cholesterol	(PDE4B)
rs6840033	7.10E-21	17554300	Hypertension, gender-differentiated	(SCOC)
rs16906340	2.40E-11	23063622	Triglycerides	(PRMT3)
rs486438	3.40E-11	23063622	LDL cholesterol	(PDE4B)
rs11809491	8.00E-11	23063622	LDL cholesterol	(PDE4B)
rs16942951	8.90E-10	17554300	Hypertension	(RORA)
rs567279	1.30E-09	23063622	Triglycerides	(PDE4B)
rs1992567	5.90E-09	23063622	Triglycerides	(PRMT3)
rs4765066	6.60E-09	17554300	Hypertension, gender-differentiated	
rs2226284	3.40E-08	22322875	Hypertension in premenopausal women(short sleep)	(LRRC7)
rs17128817	3.70E-08	23063622	LDL cholesterol	(PDE4B)
rs12756253	4.30E-08	22322875	Hypertension in premenopausal women(short sleep)	(LRRC7)
rs6691577	5.20E-08	22322875	Hypertension in premenopausal women(short sleep)	(LRRC7)
rs567279	3.10E-07	23063622	Total cholesterol	(PDE4B)
rs726914	7.10E-07	20838585	Systolic blood pressure (SBP)	(RORA)
rs12591650	7.30E-07	22589738	Subcutaneous adipose tissue (female)	(RORA)
rs559477	1.00E-06	21799836	Type 2 diabetes	(MAML2)
rs341408	1.70E-06	22184326	Essential hypertension	(RORA)
rs486438	1.80E-06	23063622	Triglycerides	(PDE4B)

rs11021499	1.90E-06	23192594	Waist hip ratio and alcohol intake	(MAML2)
rs886374	2.40E-06	17463248	Type 2 diabetes	(SORCS2)
rs912062	2.60E-06	23251661	Diet carbohydrate (g/d) in children	
rs7694661	2.70E-06	23251661	Fat mass change (kg/y) in children	(SORCS2)
rs4655595	2.70E-06	17554300	Type 2 diabetes	(PDE4B)
rs4655595	2.70E-06	17463249	Type 2 diabetes	(PDE4B)
rs10273260	3.00E-06	21037115	Body mass index (BMI)	
rs1867511	3.00E-06	21980299	Type 1 diabetes	(ATP10A)
rs880976	3.60E-06	21037115	Body mass index (BMI)	
rs4370013	3.70E-06	17903302	Diastolic blood pressure (DBP)	(CNTN4)
rs1407808	4.20E-06	23251661	Diet carbohydrate (g/d) in children	(DMRT1)
rs734526	5.30E-06	20031538	HDL cholesterol	(SORCS2)
rs11635314	5.50E-06	22479346	Hypertension (early onset hypertension)	(RORA)
rs2024125	5.50E-06	23251661	Diet carbohydrate (%energy) in children	(HECW1)
rs1023008	5.60E-06	23251661	Light activity by accelerometry in children	(LRRC7)
rs10519080	5.80E-06	22184326	Essential hypertension	(RORA)
rs2165412	6.30E-06	17463246	APOA1 (apolipoprotein AI)	(PTPRM)
rs7819143	6.40E-06	21347282	Type 2 diabetes	(ZMAT4)
rs12546688	6.60E-06	21347282	Type 2 diabetes	(ZMAT4)
rs6576426	7.50E-06	21874001	Type 2 diabetes	
rs1498024	7.90E-06	17846124	Type 2 diabetes	
rs3823763	8.10E-06	20066028	Triglycerides	(BBS9)
rs6459707	8.50E-06	20158304	Nicotine dependence (smoking)	(CNPY1)
rs7694661	8.70E-06	23251661	Fat mass deposition (kcal/d) in children	(SORCS2)
rs6576426	9.40E-06	21874001	Type 2 diabetes (males)	
rs6118476	9.50E-06	23400010	Fasting blood glucose change	(PLCB4)
rs11996523	1.00E-05	19151714	Adult obesity (body mass index (BMI))	
rs2253959	1.00E-05	23263444	HDL cholesterol particle diameter	(SORCS2)
rs13324783	1.10E-05	22566498	Blood pressure response to candesartan treatment	(CNTN4)
rs12490128	1.10E-05	17463248	Type 2 diabetes	(TMEM108)
rs10044817	1.10E-05	23263444	HDL cholesterol particle diameter	(ADAMTS16)
rs6971850	1.20E-05	20158304	Nicotine dependence (smoking)	(CNPY1)
rs1995626	1.20E-05	20418889	Nicotine dependence, ever smokers versus never smokers	
rs11058304	1.30E-05	17463246	Hypertension	
rs6691577	1.30E-05	22322875	Hypertension in women with short sleep duration	(LRRC7)
rs2913645	1.30E-05	23263444	HDL cholesterol particle diameter	(ADAMTS16)
rs1914009	1.30E-05	23669352	BMI in adolescence/young adulthood	(CNTN4)

rs13277268	1.50E-05	20066028	Triglycerides	(ZMAT4)
rs2214959	1.50E-05	17903303	Common carotid artery intima media thickness	
rs2359630	1.60E-05	23263444	HDL cholesterol particle diameter	(SORCS2)
rs10846960	1.70E-05	17463246	Hypertension	
rs1038164	1.70E-05	20686565	HDL cholesterol	(MDM1)
rs9872038	1.80E-05	22566498	Blood pressure response to candesartan treatment	(CNTN4)
rs10502365	1.90E-05	17463246	APOA1 (apolipoprotein AI)	(PTPRM)
rs644399	2.00E-05	17463246	Body mass index (BMI) in Type 2 diabetes	(PTPRM)
rs17584923	2.10E-05	22566498	Blood pressure response to candesartan treatment	(CNTN4)
rs2226284	2.10E-05	22322875	Hypertension in women with short sleep duration	(LRRC7)
rs12318567	2.20E-05	17463246	Hypertension	
rs11161183	2.20E-05	23209189	Type 2 diabetes	
rs7679784	2.20E-05	20418889	Nicotine dependence, ever smokers versus never smokers	
rs7639356	2.30E-05	22566498	Blood pressure response to candesartan treatment	(CNTN4)
rs4798615	2.30E-05	22961080	Type 2 diabetes	(PTPRM)
rs1950918	2.30E-05	22384028	Hypertension	(FRMD6)
rs1197316	2.30E-05	17463248	Type 2 diabetes	(TMEM108)
rs11875152	2.40E-05	22961080	Type 2 diabetes	(PTPRM)
rs1708373	2.50E-05	17463248	Type 2 diabetes	(TMEM108)
rs9990572	2.60E-05	20418889	Nicotine dependence, ever smokers versus never smokers	
rs9429549	2.60E-05	17554300	Type 1 diabetes, Rheumatoid arthritis	
rs543759	2.60E-05	17554300	Type 2 diabetes	
rs10512896	2.70E-05	17463248	Type 2 diabetes	(TMEM108)
rs11707897	2.80E-05	22377092	Nicotine dependence (smoking)	(CNTN4)
rs7934164	2.80E-05	22479346	Hypertension (early onset hypertension)	(MAML2)
rs7631834	2.80E-05	22566498	Blood pressure response to candesartan treatment	(CNTN4)
rs17297332	2.90E-05	17463248	Type 2 diabetes	(TMEM108)
rs9869344	2.90E-05	22566498	Blood pressure response to candesartan treatment	(CNTN4)
rs7625110	2.90E-05	17463248	Type 2 diabetes	(TMEM108)
rs10772719	3.00E-05	18519826	Nicotine dependence (smoking), smoking cessation	(GRIN2B)
rs16840161	3.10E-05	17463248	Type 2 diabetes	(TMEM108)

rs1920021	3.10E-05	17463248	Type 2 diabetes	(TMEM108)
rs576192	3.10E-05	20158304	Nicotine dependence (smoking)	
rs10518442	3.20E-05	17846124	Type 2 diabetes	(ANKRD50)
rs16875358	3.30E-05	17463246	Body mass index (BMI)	
rs2584808	3.30E-05	23209189	Type 2 diabetes	
rs10060539	3.40E-05	20339536	Total cholesterol change with statins	(CDH12)
rs2913634	3.40E-05	17554300	Type 2 diabetes, combined control dataset	(ADAMTS16)
rs12648772	3.50E-05	17554300	Hypertension, gender differentiated	
rs7599888	3.50E-05	17554300	Type 1 diabetes, gender differentiated	(HDLBP)
rs6857392	3.50E-05	20418889	Nicotine dependence, ever smokers versus never smokers	
rs268972	3.50E-05	21701570	Body mass index (BMI)	(CDH12)
rs1444123	3.60E-05	20339536	LDL cholesterol change with statins	(CDH12)
rs17760889	3.70E-05	22566498	Blood pressure response to candesartan treatment	(GRIN2B)
rs11745606	3.70E-05	20339536	Total cholesterol change with statins	(CDH12)
rs1938576	3.70E-05	22322875	Hypertension in women with short sleep duration	(LRR7)
rs567338	3.70E-05	20158304	Nicotine dependence (smoking)	
rs12756253	3.90E-05	22322875	Hypertension in women with short sleep duration	(LRR7)
rs11767191	3.90E-05	21935397	Body mass index (BMI)	
rs4723264	3.90E-05	23152477	carotid artery intima media thickness	(BBS9)
rs1506547	3.90E-05	20418889	Nicotine dependence, ever smokers versus never smokers	
rs1373370	3.90E-05	20339536	LDL cholesterol change with statins	(CDH12)
rs10013239	4.00E-05	21701565	Body mass index (BMI) (within family tests)	(SORCS2)
rs2920764	4.10E-05	20339536	LDL cholesterol change with statins	(CDH12)
rs1245058	4.10E-05	17903299	Mean LDL cholesterol	
rs13185430	4.20E-05	20339536	Total cholesterol change with statins	(CDH12)
rs6055905	4.20E-05	21909115	Systolic blood pressure (SBP)	(PLCB1)
rs2968277	4.20E-05	20339536	LDL cholesterol change with statins	(CDH12)
rs533682	4.20E-05	20158304	Nicotine dependence (smoking)	
rs16979683	4.30E-05	17463246	Systolic blood pressure (SBP)	(KIAA1671)
rs9283772	4.30E-05	20339536	Total cholesterol change with statins	(CDH12)
rs1395327	4.30E-05	20418889	Nicotine dependence, ever smokers versus never smokers	
rs12188645	4.40E-05	20158304	Nicotine dependence (smoking)	
rs7613278	4.40E-05	17463246	Weight in controls	(CNTN4)

rs7035993	4.40E-05	22377092	Nicotine dependence (smoking)	
rs2719791	4.40E-05	19268276	Nicotine dependence (smoking)	(PLCB1)
rs582683	4.40E-05	21935397	Body mass index (BMI)	(PTPRM)
rs8123323	4.50E-05	21082022	Hypertension	(PLCB1)
rs6965061	4.50E-05	19304780	Diastolic blood pressure (DBP)	(HECW1)
rs3099034	4.50E-05	20339536	LDL cholesterol change with statins	(CDH12)
rs10010676	4.60E-05	20418889	Nicotine dependence, ever smokers versus never smokers	
rs1395325	4.70E-05	20418889	Nicotine dependence, ever smokers versus never smokers	
rs11915941	4.80E-05	20339536	Triglycerides change with statins	(CNTN4)
rs9968229	4.80E-05	17554300	Type 2 diabetes, gender differentiated	(CNTN4)
rs4906719	4.80E-05	23209189	Type 2 diabetes	
rs2968280	4.90E-05	20339536	LDL cholesterol change with statins	(CDH12)
rs2251671	4.90E-05	17903303	Ankle brachial index	
rs6534452	5.00E-05	20418889	Nicotine dependence, ever smokers versus never smokers	
rs6004454	5.10E-05	17463246	Systolic blood pressure (SBP)	(KIAA1671)
rs1483597	5.10E-05	20418889	Nicotine dependence, ever smokers versus never smokers	
rs12595506	5.30E-05	23209189	Type 2 diabetes	
rs6047163	5.30E-05	19060906	HDL cholesterol	
rs12728294	5.40E-05	19268276	Nicotine dependence (smoking)	(PDE4B)
rs463	5.40E-05	18519826	Nicotine dependence (smoking), smoking cessation	(HECW1)
rs12652874	5.50E-05	17463246	Diastolic blood pressure (DBP)	
rs11931821	5.60E-05	20339536	HDL cholesterol change with statins	(SORCS2)
rs1106701	5.60E-05	22961080	Type 2 diabetes	(PTPRM)
rs41346644	5.70E-05	17554300	Type 1 diabetes	(RORA)
rs10222899	5.70E-05	20418889	Nicotine dependence, ever smokers versus never smokers	
rs1444127	5.70E-05	20339536	LDL cholesterol change with statins	(CDH12)
rs789451	5.70E-05	21935397	Body mass index (BMI)	(PTPRM)
rs7231577	6.00E-05	20066028	HDL cholesterol	(PTPRM)
rs7512836	6.10E-05	17554300	Type 1 diabetes, gender differentiated	
rs1867511	6.10E-05	17554300	Type 1 diabetes, combined control dataset	(ATP10A)
rs6889081	6.20E-05	20339536	LDL cholesterol change with statins	(CDH12)
rs16979684	6.30E-05	17463246	Systolic blood pressure (SBP)	(KIAA1671)
rs10519077	6.40E-05	20339536	LDL cholesterol change with statins	(RORA)

rs2968271	6.40E-05	20339536	LDL cholesterol change with statins	(CDH12)
rs2887738	6.40E-05	19609347	Systolic blood pressure (SBP)	(ZNF438)
rs638251	6.60E-05	21935397	Body mass index (BMI)	(PTPRM)
rs17708673	6.70E-05	17463246	LDL cholesterol	
rs7313149	6.70E-05	19268276	Nicotine initiation (smoking)	(GRIN2B)
rs4600815	6.80E-05	17554300	Type 2 diabetes	(TMEM108)
rs163574	6.80E-05	20031582	LDL cholesterol change with statins	(CNTN4)
rs2367844	6.90E-05	19060906	LDL cholesterol	
rs11760611	7.00E-05	20339536	LDL cholesterol	(CNPY1)
rs1261370	7.00E-05	20339536	LDL cholesterol change with statins	(CDH12)
rs6055905	7.10E-05	21909115	Diastolic blood pressure (DBP)	(PLCB1)
rs2920761	7.10E-05	20339536	LDL cholesterol change with statins	(CDH12)
rs1703317	7.20E-05	20339536	LDL cholesterol change with statins	(CDH12)
rs16951807	7.30E-05	17463246	Hypertension	
rs769097	7.30E-05	17554300	Type 2 diabetes	(TMEM108)
rs502936	7.30E-05	19609347	Diastolic blood pressure (DBP)	
rs16840358	7.40E-05	20031538	HDL cholesterol	(SORCS2)
rs11831620	7.40E-05	22566498	Blood pressure response to candesartan treatment	(GRIN2B)
rs6494228	7.40E-05	20339536	LDL cholesterol change with statins	(RORA)
rs1703311	7.40E-05	20339536	LDL cholesterol change with statins	(CDH12)
rs16994779	7.50E-05	20339536	Triglycerides change with statins	(PLCB1)
rs7612595	7.50E-05	17463248	Type 2 diabetes	(TMEM108)
rs10766671	7.60E-05	17554300	Type 1 diabetes, gender differentiated	(PRMT3)
rs17661843	7.70E-05	20339536	LDL cholesterol change with statins	(ABCA13)
rs13172440	7.70E-05	20339536	Total cholesterol change with statins	(CDH12)
rs1975193	7.80E-05	17463246	Systolic blood pressure (SBP)	(ABCA13)
rs16906290	7.90E-05	17554300	Type 1 diabetes, gender differentiated	(PRMT3)
rs11630108	7.90E-05	20339536	LDL cholesterol change with statins	(RORA)
rs9302216	7.90E-05	20339536	LDL cholesterol change with statins	(RORA)
rs1861931	8.10E-05	19060906	Triglycerides	
rs7650741	8.20E-05	17463248	Type 2 diabetes	(TMEM108)
rs10512891	8.30E-05	17463248	Type 2 diabetes	(TMEM108)
rs2227492	8.30E-05	20686565	Triglycerides	(IL22)
rs1979788	8.40E-05	20418889	Nicotine dependence, ever smokers versus never smokers	
rs1663322	8.40E-05	20339536	LDL cholesterol change with statins	(CDH12)
rs929713	8.40E-05	19060906	Triglycerides	
rs341378	8.40E-05	20339536	LDL cholesterol change with statins	(RORA)

rs13072106	8.70E-05	17463248	Type 2 diabetes	(TMEM108)
rs1395326	8.90E-05	20418889	Nicotine dependence, ever smokers versus never smokers	
rs341380	8.90E-05	20339536	LDL cholesterol change with statins	(RORA)
rs341381	9.00E-05	20339536	LDL cholesterol change with statins	(RORA)
rs7664871	9.10E-05	21701565	Body mass index (BMI) (within family tests)	(SORCS2)
rs6442737	9.20E-05	20066028	Triglycerides	(CNTN4)
rs12624339	9.30E-05	20877124	Systolic blood pressure (SBP)	(PLCB1)
rs7694823	9.40E-05	23502781	Pulmonary arterial hypertension (PAH)	(SORCS2)
rs2028722	9.40E-05	20339536	LDL cholesterol change with statins	(CDH12)
rs9296736	9.50E-05	17463246	APOA1 (apolipoprotein AI)	(MLIP)
rs3112474	9.50E-05	20339536	LDL cholesterol change with statins	(CDH12)
rs731177	9.50E-05	17554300	Type 2 diabetes	(PLCB4)
rs1979015	9.60E-05	19060906	LDL cholesterol	
rs10489532	9.70E-05	17554300	Type 1 diabetes, Rheumatoid arthritis	