SUPPLEMENTARY INFORMATION

Prediction of Colorectal Cancer Risk Using a Genetic Risk Score: The Korean Cancer Prevention Study-II (KCPS-II)

Jaeseong Jo[1,2], Chung Mo Nam[2,3], Jae Woong Sull[4], Ji Eun Yun[1], Sang Yeun Kim[5], Sun Ju Lee[5], Yoon Nam Kim[2], Eun Jung Park[2,5], Heejin Kimm[1,5], Sun Ha Jee[1,5]*

[1]Institute for Health Promotion and Department of Epidemiology and Health Promotion, Graduate School of Public Health, Yonsei University, Seoul 120-752, Korea, [2]Department of Public Health, Graduate School of Yonsei University, Seoul 120-752, Korea, [3]Department of Preventive Medicine, Yonsei University College of Medicine, Seoul 120-752, Korea, [4]Metabolic Syndrome Research Initiatives, Seoul 120-752, Korea, [5]Department of Biomedical Laboratory Science, College of Health Sciences, Eulji University, Seongnam 461-713, Korea

Supplementary Table 1. Association of SNPs with colorectal cancer based on logistic regression model in 687 Korean men in the KCPS-II

Chr	rs no.	Position	Gene/Locus[a]	A1	A2	MAF[b]	OR	95% CI	p-value
10	rs17391002	44175933	CXCL12	G	A	0.138	2.5	1.7-3.8	5.24E-06
13	rs9549448	111883255	SOX1	T	C	0.436	2.1	1.5-2.9	7.17E-06
5	rs254833	16933460	MYO10	T	C	0.418	0.5	0.3-0.7	9.47E-06
8	rs2553614	133827396	TMEM71	T	G	0.345	2.1	1.5-2.8	1.53E-05
5	rs13153032	6629580	NSUN2	C	G	0.363	2	1.5-2.7	1.90E-05
2	rs2288073	24266802	FLJ30851	G	A	0.153	2.3	1.6-3.4	1.93E-05
13	rs9604214	111891087	SOX1	G	A	0.071	3	1.8-4.9	2.22E-05
3	rs9865670	194946907	OPA1	G	A	0.445	0.5	0.3-0.7	2.44E-05
6	rs17186320	84954566	KIAA1009	A	T	0.014	6.6	2.7-15.8	2.45E-05
2	rs1509497	101504616	RFX8	T	C	0.493	0.5	0.3-0.7	2.46E-05
8	rs235428	133884908	PHF20L1	A	G	0.26	2	1.5-2.9	2.53E-05
3	rs9845920	194960582	OPA1	A	C	0.443	0.5	0.3-0.7	2.58E-05
3	rs9846212	194660628	OPA1	A	G	0.443	0.5	0.3-0.7	2.58E-05
3	rs6763744	194964794	OPA1	G	T	0.439	0.5	0.3-0.7	3.12E-05
2	rs4128317	29513358	ALK	C	A	0.039	4.1	2.1-7.8	3.17E-05
3	rs7646304	194965291	OPA1	A	C	0.439	0.5	0.3-0.7	3.21E-05
1	rs17047306	216310320	SPATA17	T	C	0.283	2	1.4-2.8	3.28E-05
1	rs1490338	216218974	SPATA17	T	A	0.189	2.2	1.5-3.2	3.62E-05
1	rs902351	216271369	SPATA17	A	G	0.281	2	1.4-2.8	3.82E-05
2	rs2543662	24296775	ITSN2	C	A	0.131	2.3	1.6-3.5	3.86E-05

p-values were calculated under a logistic regression model under an additive mode, incorporating age as a covariate
SNP, single nucleotide polymorphism; KCPS-II, Korean Cancer Prevention Study-II; Chr, chromosome; A1, major allele; A2, minor allele; MAF, minor allele frequencies; OR, odds ratio; 95% CI, 95% confidence interval; GWAS, genomewide association study.
[a]Nearest gene or locus for each SNP; [b]MAF based on controls of GWAS studies.

Supplementary Table 2. Association of SNPs with colorectal cancer based on logistic regression model in 476 Korean women in the KCPS-II

Chr	rs no.	Position	Gene/Locus[a]	A1	A2	MAF[b]	OR	95% CI	p-value
16	rs10083736	57413317	GOT2	T	C	0.266	4.3	2.3-8.1	4.54E-06
20	rs16987827	37230200	DHX35	G	T	0.07	6.7	3.0-15.1	4.56E-06
16	rs8046516	57422045	GOT2	C	G	0.271	4.4	2.3-8.3	4.62E-06
16	rs9926182	57407072	GOT2	A	G	0.269	4.4	2.3-8.3	4.77E-06
15	rs17523778	90957863	FAM174B	G	A	0.121	5.4	2.6-11.2	9.08E-06
3	rs4974411	128763249	TPRA1	A	G	0.109	5.8	2.6-12.9	1.38E-05
5	rs1834902	134616780	H2AFY	A	G	0.431	3.9	2.1-7.3	1.40E-05
5	rs16895308	65944570	MAST4	G	A	0.334	3.6	2.0-6.3	1.79E-05
15	rs8032832	90959099	FAM174B	A	G	0.073	4.9	2.3-10.3	2.67E-05
6	rs6901560	37039972	PI16	C	G	0.042	7.4	2.9-18.9	3.37E-05
11	rs11025480	20282667	PRMT3	A	G	0.083	4.3	2.1-8.7	4.39E-05
9	rs3814110	16427728	BNC2	C	G	0.048	6.6	2.6-14.6	4.87E-05
5	rs16895307	65944488	MAST4	A	G	0.402	3.4	1.9-6.2	4.97E-05
10	rs7089063	45266395	MARCH8	A	C	0.173	3.8	2.0-7.3	4.99E-05
6	rs16893688	82706672	IBTK	T	G	0.082	6.1	2.5-14.6	5.61E-05
5	rs6861487	65927298	MAST4	T	C	0.324	3.3	1.8-5.8	5.97E-05
22	rs9613463	26235926	MN1	G	C	0.527	0.2	0.1-0.5	5.98E-05
5	rs11242237	134618811	H2AFY	A	C	0.493	0.2	0.1-0.5	6.09E-05
16	rs11150094	77272068	WWOX	T	G	0.484	3.5	1.9-6.6	6.28E-05
22	rs9625253	26242970	MN1	C	G	0.341	3.5	1.9-6.3	6.37E-05

p-values were calculated under a logistic regression model under an additive mode, incorporating age as a covariate.
SNP, single nucleotide polymorphism; KCPS-II, Korean Cancer Prevention Study-II; Chr, chromosome; A1, major allele; A2, minor allele; MAF, minor allele frequencies; OR, odds ratio; 95% CI, 95% confidence interval; GWAS, genomewide association study.
[a]Nearest gene or locus for each SNP; [b]MAF based on controls of GWAS studies.

Supplementary Table 3. Cross-validation results for counted GRS

Group		Actual no. of cases	Predicted group		Correctly classified (%)	AUC	95% CI
			Cases	Controls			
Men							
	Cases	187	98	89	78.25	0.697	0.652-0.743
	Controls	976	164	812			
Women							
	Cases	187	91	96	77.77	0.600	0.561-0.639
	Controls	976	161	815			

GRS, genetic risk score; AUC, area under the curve; CI, confidence interval.

Supplementary Table 4. Cross-validation results for weighted GRS

Group		Actual no. of cases	Predicted group		Correctly classified (%)	AUC	95% CI
			Cases	Controls			
Men							
	Cases	187	103	84	74.46	0.689	0.643-0.734
	Controls	976	213	763			
Women							
	Cases	187	110	77	78.16	0.583	0.545-0.623
	Controls	976	177	799			

GRS, genetic risk score; AUC, area under the curve; CI, confidence interval.