

Supplementary File 1

Table S1: Detailed MR and sensitivity analyses between microbiota and inflammatory bowel disease using FinnGen GWAS data.

Table S2: Detailed MR and sensitivity analyses between microbiota and inflammatory bowel disease using IIBDGC GWAS data.

Table S3: Detailed MR and sensitivity analyses between microbiota and inflammatory bowel disease using Mbatchou J et al.’s GWAS data.

Table S4: Detailed MR analyses between inflammatory bowel disease and colorectal cancer.

Table S5: Detailed MR analyses between microbiota and colorectal cancer using Huyghe JR et al.’s GWAS data.

Table S6: Detailed MR analyses between microbiota and colorectal cancer using Ishigaki K et al.’s GWAS data.

Table S7: The Mendelian randomization results of the microbiota and IBD which had mediating effects in CRC.

Table S8: Expression analysis and LASSO regression.

Table S9: Detailed results of drug target MR analyses between seven candidate genes and colorectal cancer using Huyghe JR et al.’s GWAS data.

Table S10: Detailed results of drug target MR analyses between seven candidate genes and colorectal cancer using Ishigaki K et al.’s GWAS data.

Table S11: Detailed results of SMR and HEIDI tests between the expression of seven candidate genes and colorectal cancer using Ishigaki K et al.’s GWAS data.

Figure S1: Parameter tuning, variable importance, and model performance evaluation of six machine learning algorithms used for CRC classification.

Table S1:
Detailed MR and sensitivity analyses between microbiota and inflammatory bowel disease using FinnGen GWAS data.

Trait	OR(IVW)	95%CI(I VW)	Pval(IV W)	OR(MR-E gger)	95%CI(MR- Egger)	Pval(MR- Egger)	OR(Weight ed median)	95%CI(We ighted median)	Pval(We ighted median)	OR(Wei ghted mode)	95%CI(We ighted mode)	Pval(We ighted mode)	OR(Sim ple mode)	95%CI(Simple mode)	Pval(Sim ple mode)	Rucker Q(MR-Egg er)	Pval_Q(MR -Egger)	Cochran Q(IVW)	Pval_Q(I VW)	Intercept(M R-Egger)	Pval_Inter cept	Power	Steiger P-value	Number of SNP	
																								Not Passing MR-Steiger	P-adj(IV W)
class.Actinobacteria.id. 419	0.830114185	0.810-0.851	2.9865E-48	0.690994703	0.613-0.778	2.1619E-09	0.821630886	0.794-0.850	8.35922E-30	0.825231235	0.740-0.921	0.000612473	0.825231235	0.732-0.930	0.001704942	452.8329148	0.999958597	462.3627164	0.999863028	0.015741907	0.002118659	1	NA	0	1.575377E-46
family.Alcaligenaceae.id.2875	1.278980433	1.185-1.381	3.19309E-10	0.428524404	0.241-0.762	0.004938608	1.398965304	1.268-1.543	1.9569E-11	1.412117798	1.101-1.811	0.007858339	1.412117798	1.109-1.797	0.006180176	16.55133177	1	30.64004547	0.99999999	0.063476366	0.000310412	0.999699913974686	0	0	3.368705E-09
family.Bifidobacteriaceae.id.433	0.781641809	0.761-0.803	1.45307E-74	1.072143858	0.952-1.207	0.24962162	0.830182878	0.801-0.861	6.18111E-24	0.832590551	0.749-0.926	0.000778142	0.832590551	0.741-0.936	0.002275238	486.7639447	0.494495	515.4595642	0.188284115	-0.027322842	1.30855E-07	1	NA	0	1.532988E-72
genus.Eubacteriumcoprostanoligenesgroup.id.11375	1.161357197	1.027-1.313	0.016677897	0.679804855	0.305-1.516	0.350615793	1.198258154	1.021-1.406	0.026909628	1.236377439	0.908-1.684	0.184885066	1.242803009	0.907-1.704	0.183421353	10.03699249	0.9999999991	11.79077906	0.9999999926	0.041312018	0.192088528	0.961749067304915	5.21111284895923e-220	0	5.324683E-02
genus.Bifidobacterium.id.436	0.789139645	0.768-0.811	2.30783E-63	1.157571278	1.020-1.313	0.023277861	0.824332467	0.794-0.855	1.12287E-24	0.833493401	0.756-0.919	0.000272625	0.829938184	0.739-0.932	0.001805521	602.2684234	0.012652304	644.7474445	0.000365505	-0.032717356	2.09416E-09	1	NA	0	1.623176E-61
genus.Lachnoclostridium.id.11308	0.507598756	0.460-0.560	6.74382E-42	1.904478222	0.853-4.253	0.121203609	0.445089595	0.386-0.513	1.02072E-28	0.4162189	0.298-0.582	3.13225E-06	0.414909216	0.306-0.563	4.14962E-07	36.56443644	0.994487623	47.12808715	0.919110441	-0.072563762	0.001879085	0.99002993386255	5.25248431963958e-247	0	2.371577E-40
genus.LachnospiraceaeND3007group.id.11317	1.554014627	1.207-2.000	0.000622449	0.603650241	1.497-243475.231	0.941048678	1.546101809	1.129-2.117	0.00656180106899353	1.51357243607311	0.956-2.397	0.115178122932628	1.51357243607311	1.009-2.271	0.0801983243181022	0.0719936197477242	0.999999260174171	0.0926193252833435	0.999999815329275	0.0516021221675589	0.889850266363259	1.129-2.117	0.00656180106899353	0	2.626734E-03
genus.Peptococcus.id.2037	0.911836931	0.885-0.939	6.76529E-10	1.247928773	1.023-1.522	0.029934736	0.844706213	0.811-0.879	2.45206E-16	0.844772467	0.765-0.933	0.001045464	0.844772467	0.759-0.941	0.002418547	100.3725139	0.999999978	110.1914013	0.999999911	-0.036964485	0.002004123	1	0	0	6.797502E-09
genus.RuminococcaceaeUCG011.id.11368	1.070770632	1.044-1.098	1.39456E-07	0.33471752	0.266-0.420	3.54369E-17	1.145886928	1.106-1.187	3.82432E-14	1.146857159	1.033-1.274	0.01123225	1.146857159	1.035-1.270	0.009350782	14.87997836	1	115.9932305	0.999571781	0.144356661	5.84044E-19	1	NA	0	9.808399E-07
genus.Streptococcus.id.1853	0.795489014	0.752-0.841	9.0014E-16	0.548241022	0.427-0.703	4.8003E-06	0.788019677	0.733-0.848	1.55569E-10	0.789046231	0.643-0.968	0.024651704	0.789046231	0.634-0.982	0.03484855	38.03422663	1	47.06000117	1	0.04011826	0.003072089	0.999999999937151	0	0	1.726632E-14
genus.unknowngenus.id.1868	0.881000639	0.822-0.944	0.000316869	2.190787339	1.299-3.695	0.004014713	0.682185686	0.628-0.741	1.12599E-19	0.67447734	0.578-0.787	2.06769E-06	0.67312999	0.565-0.801	2.10204E-05	155.2711774	0.001967043	172.3073877	0.000106566	-0.072451005	0.000821479	0.9999999990267841	0	0	1.364476E-03
order.Bifidobacteriales.id.432	0.781641809	0.761-0.803	1.45307E-74	1.072143858	0.952-1.207	0.24962162	0.830182878	0.801-0.861	6.46247E-24	0.832590551	0.757-0.916	0.000188936	0.832590551	0.744-0.931	0.001415213	486.7639447	0.494495	515.4595642	0.188284115	-0.027322842	1.30855E-07	1	NA	0	1.532988E-72
phylum.Actinobacteria.id.400	0.764931281	0.734-0.797	2.0622E-36	0.445225824	0.334-0.593	7.62367E-08	0.768167834	0.726-0.813	5.53746E-20	0.770290459	0.653-0.908	0.002126723	0.770290459	0.655-0.906	0.00174271	130.802303	1	144.762702	1	0.037917887	0.000227509	0.999999999998	0	0	6.216051E-35

Abbreviation: OR, odds ratio; IVW, inverse variance weighted method; CI, confidential interval; a Pval of 0 means Pval<0.001; P-adj, P-value after P-value adjustment

Table S2:
Detailed MR and sensitivity analyses between microbiota and inflammatory bowel disease using IIBDGC GWAS data.

Trait	OR(IVW)	95%CI(IVW)	Pval(IVW)	OR(MR-Egger)	95%CI(MR-Egger)	Pval(MR-Egger)	OR(Weighted median)	95%CI(Weighted median)	Pval(Weighted median)	OR(Weighted mode)	95%CI(Weighted mode)	Pval(Weighted mode)	OR(Simple mode)	95%CI(Simple mode)	Pval(Simple mode)	Rucker Q(MR-Egger)	Pval_Q(MR-Egger)	Cochran Q(IVW)	Pval_Q(IVW)	Intercept(MR-Egger)	Pval_Intercept	Power	Steiger P-value	Number of SNPs	P-adj(Ig MR-Steiger)	P-adj(IVW)
class.Actinobacteria.id.419	0.805460076	0.788-0.823	2.48419E-82	0.497560277	0.447-0.554	9.40785E-33	0.855928113	0.828-0.885	3.82273E-20	0.903719169	0.818-0.998	0.046355982	0.622685824	0.547-0.709	2.98049E-12	431.576586	0.9999999816	511.4799875	0.991937341	0.041015844	5.01816E-18	1	NA	0	1.308200E-80	
family.Alcaligenaceae.id.2875	0.886988674	0.832-0.946	0.000272588	1.109917885	0.711-1.734	0.64790858	0.880934426	0.809-0.960	0.003664272	0.886004929	0.735-1.069	0.208870415	0.886004929	0.723-1.085	0.245091555	25.49217104	1	26.48347508	1	-0.013031274	0.32212317	0.9999998129659	1.66209684158975e-291	0	1.437902E-03	
family.Bifidobacteriaceae.id.433	0.783964512	0.766-0.802	6.75641E-99	0.669620278	0.603-0.744	2.83944E-13	0.831250397	0.804-0.859	5.98714E-28	0.902923802	0.804-1.014	0.085008356	0.643811418	0.562-0.738	4.86619E-10	312.5265726	1	321.6561534	1	0.013605627	0.002644215	1	NA	0	7.131800E-97	
genus.Eubacteriumcoprostanoligenesgroup.id.11375	1.319727304	1.201-1.451	9.07258E-09	2.114084009	1.081-4.135	0.033968068	1.402057008	1.229-1.599	4.78967E-07	1.44515067	1.124-1.858	0.006111729	1.448892628	1.101-1.906	0.010986265	16.93839991	0.999954493	18.87199879	0.999868145	-0.036399725	0.171204373	0.9982086383226	5.3069998132714e-124	0	8.320739E-08	
genus.Bifidobacterium.id.436	0.805665752	0.788-0.824	1.04127E-82	0.645639334	0.581-0.718	3.38496E-15	0.853074961	0.826-0.881	3.37519E-22	0.903766716	0.814-1.003	0.05811452	0.628498666	0.551-0.716	9.59778E-12	325.371617	1	342.9882854	1	0.018823437	3.16033E-05	1	NA	0	7.314667E-81	
genus.Lachnoclostridium.id.11308	0.885260553	0.816-0.960	0.003312507	1.394571881	0.598-3.251	0.444188851	0.882315545	0.793-0.981	0.021076063	0.881745793	0.705-1.110	0.273104911	0.880119035	0.692-1.119	0.302132179	8.348884467	1	9.466891684	1	-0.024554417	0.294586594	0.9998407775226	6.25701442638981e-187	0	1.270798E-02	
genus.LachnospiraceaeND3007group.id.11317	1.526049397	1.209-1.927	0.000378189	0.614930502	0.005-82.010	0.850426075	1.515877194	1.145-2.007	0.003644433	1.510862786	0.984-2.320	0.092030841	1.511929559	0.977-2.339	0.09632106	0.044705495	0.999999999	0.177564563	0.999999671	0.050339363	0.724932556	0.6503244702184	1.55197931133598e-25	0	1.946290E-03	
genus.Peptococcus.id.2037	1.165715926	1.137-1.195	1.79085E-33	0.816350485	0.689-0.967	0.020197589	1.191416239	1.153-1.231	1.08259E-25	1.181592651	1.074-1.300	0.000717025	1.181592651	1.079-1.294	0.00042621	35.02898291	1	52.30890766	1	0.041775845	4.84675E-05	1	0	0	4.196556E-32	
genus.RuminococcaceaeUCG011.id.11368	1.043095251	1.020-1.067	0.000197942	0.752471304	0.610-0.928	0.008655226	1.06098418	1.031-1.092	5.37026E-05	1.079748642	0.997-1.170	0.061780275	1.079748642	0.994-1.172	0.069239327	15.80760814	1	25.20837234	1	0.040293387	0.002509171	1	NA	0	1.099099E-03	
genus.Streptococcus.id.1853	1.416790672	1.359-1.477	1.37564E-59	2.017181004	1.683-2.418	1.92109E-12	1.514926649	1.431-1.604	5.53529E-46	1.584509684	1.364-1.841	1.03892E-08	1.590147471	1.369-1.847	7.66711E-09	118.4927094	0.999209562	133.9449339	0.985723134	-0.038228336	0.00012271	1	0	0	4.853000E-58	
genus.unknowngenus.id.1868	1.0690961	1.022-1.119	0.003808806	1.334135717	0.931-1.911	0.118772565	1.216884457	1.142-1.297	1.49253E-09	1.266486479	1.049-1.529	0.015636919	1.269815003	1.064-1.515	0.009075242	97.83456659	0.844454952	99.31630786	0.834720685	-0.017709114	0.226038302	1	0	0	1.339430E-02	
order.Bifidobacteriales.id.432	0.783964512	0.766-0.802	6.75641E-99	0.669620278	0.603-0.744	2.83944E-13	0.831250397	0.805-0.858	6.1979E-30	0.902923802	0.803-1.015	0.086923184	0.643811418	0.564-0.734	1.33566E-10	312.5265726	1	321.6561534	1	0.013605627	0.002644215	1	NA	0	7.131800E-97	
phylum.Actinobacteria.id.400	0.720972989	0.695-0.748	1.62935E-68	0.59690824	0.465-0.766	6.77103E-05	0.778984303	0.738-0.822	1.03807E-19	0.867422029	0.742-1.014	0.076056653	0.867422029	0.730-1.030	0.10599125	178.4730745	0.999999967	180.7124472	0.999999497	0.013106682	0.135657861	1	0	0	6.878600E-67	

Abbreviation: OR, odds ratio; IVW, inverse variance weighted method; CI, confidential interval; a Pval of 0 means Pval<0.001; P-adj, P-value after P-value adjustment

Table S3:
Detailed MR and sensitivity analyses between microbiota and inflammatory bowel disease using Mbatchou J et al.’s GWAS data.

Trait	OR(IVW)	95%CI(IVW)	Pval(IVW)	OR(MR-Egger)	95%CI(MR-Egger)	Pval(MR-Egger)	OR(Weighted median)	95%CI(Weighted median)	Pval(Weighted median)	OR(Weighted mode)	95%CI(Weighted mode)	Pval(Weighted mode)	OR(Simple mode)	95%CI(Simple mode)	Pval(Simple mode)	Rucker Q(MR-Egger)	Pval_Q(MR-Egger)	Cochran Q(IVW)	Pval_Q(IVW)	Intercept(MR-Egger)	Pval_Intercept	Power	Steiger P-value	Number of SNP Not Passing MR-Steiger	P-adj(IVW)
class.Actinobacteria.id.419	0.996544241	0.996-0.997	1.7847E-17	9.95E-01	0.993-0.996	3.36083E-14	9.97E-01	0.997-0.997	6.3418E-48	0.99750545	0.996-0.999	1.9271E-05	9.97E-01	0.996-0.999	0.00040252	767.6862441	2.63889E-06	778.7814063	8.07393E-07	0.000171535	0.003437502	1	NA	0	1.249276E-115
family.Alcaligenaceae.id.2875	0.99863805	0.998-0.999	0.000205467	0.993899525	0.989-0.999	0.018056152	0.998944529	0.998-1.000	0.035730913	0.998964631	0.997-1.001	0.351864206	0.998964631	0.997-1.001	0.353616609	56.88363547	0.997506653	60.46308573	0.99431191	0.000276912	0.061713742	0.996170574438701	0	0	9.922914E-04
family.Bifidobacteriaceae.id.433	0.996247737	0.996-0.999	1.097E-155	9.94E-01	0.993-0.999	5.37942E-20	9.97E-01	0.997-0.997	1.91502E-48	0.997824043	0.997-0.999	0.00051575	9.98E-01	0.996-0.999	0.004844426	535.7794105	0.129979011	550.1502798	0.063560071	0.000191423	0.000276789	1	NA	0	1.151881E-153
genus.Eubacterium coprostanoligenesgroup.id.11375	1.005838873	1.005-1.007	5.57818E-26	1.00798246	1.000-1.01	0.055628757	1.01E+00	1.006-1.009	2.92683E-21	1.007703786	1.005-1.011	1.40284E-05	1.007738071	1.005-1.011	3.82069E-06	52.20530819	0.245480599	52.52509659	0.268571422	-0.00016465	0.598092424	0.896066305134508	3.53412455182894e-217	0	1.301574E-24
genus.Bifidobacterium.id.436	0.996660004	0.996-0.999	4.6159E-110	9.96E-01	0.994-0.999	2.01505E-10	9.97E-01	0.997-0.998	8.71976E-48	0.997806247	0.997-0.999	0.000324698	9.96E-01	0.994-0.997	1.50096E-08	714.5127389	1.29166E-06	718.323652	9.26085E-07	9.73468E-05	0.088776455	1	NA	0	2.423351E-108
genus.Lachnoclostridium.id.11308	1.002007232	1.001-1.003	1.85252E-05	0.995353928	0.986-1.00	0.349590457	1.002061157	1.001-1.003	0.000439785	1.002372513	1.000-1.005	0.05879437	1.002372513	1.000-1.005	0.068778765	9.140893504	110.97684	110.97684102	0.000359454	0.180592166	0.950070006006836	2.73773018312391e-297	0	1.080639E-04	
genus.LachnospiraceaeND3007group.id.11317	1.002530758	1.000-1.005	0.0213831	0.976604603	0.932-1.02	0.348626907	1.002852063	1.000-1.006	0.035161468	1.002855505	0.999-1.007	0.202134909	1.002855505	0.999-1.007	0.19286303	0.165302338	0.99999818	1.381938708	0.997931518	0.001449449	0.302093655	0.33523523076676	7.10006190385631e-46	0	5.756989E-02
genus.Peptococcus.id.2037	0.999327556	0.999-1.000	1.83108E-06	1.001288975	0.999-1.000	0.173456107	9.99E-01	0.999-0.999	4.66262E-07	0.999119471	0.998-1.000	0.102496887	0.999119471	0.998-1.000	0.102890052	32.81177968	1	37.23524373	1	-0.000230501	0.036724063	1	0	0	1.13096E-05
genus.RuminococcaceaeUCG011.id.11368	1.001244635	1.001-1.003	2.06004E-22	1.00923661	1.007-1.01	1.70717E-12	1.00E+00	1.000-1.001	5.08742E-06	1.000681988	1.000-1.002	0.166676605	1.000681988	1.000-1.002	0.162917032	34.0023509	1	77.58464008	1	-0.000980408	4.5148E-10	1	NA	0	4.326084E-21
genus.Streptococcus.id.1853	1.002538852	1.002-1.003	4.33492E-26	1.01E+00	1.007-1.01	3.39594E-16	1.00E+00	1.001-1.002	0.000145365	1.00E+00	0.999-1.003	0.374424112	1.00E+00	0.999-1.003	0.3857454	74.69525957	1	120.3876916	0.999437444	-0.000734477	2.00889E-10	0.999999953698397	0	0	1.137916E-24
genus.unknowngenus.id.1868	1.000738489	1.000-1.001	0.003952954	1.002969547	0.999-1.000	0.115100563	1.00E+00	1.001-1.003	3.14618E-07	1.001955925	1.000-1.004	0.042870672	1.001955925	1.000-1.004	0.042871632	88.98489531	0.960040829	90.43421263	0.956114038	-0.000179307	0.231131314	0.999998545866264	0	0	1.431242E-02
order.Bifidobacteriales.id.432	0.996247737	0.996-0.999	1.097E-155	9.94E-01	0.993-0.999	5.37942E-20	9.97E-01	0.997-0.997	8.40591E-46	0.997824043	0.997-0.999	0.000330908	9.98E-01	0.996-0.999	0.006670869	535.7794105	0.129979011	550.1502798	0.063560071	0.000191423	0.000276789	1	NA	0	1.151881E-153
phylum.Actinobacteria.id.400	0.995366467	0.995-0.996	1.3E-82	0.992394503	0.989-0.999	0.00000278	0.996994512	0.996-0.998	1.85E-21	0.997450582	0.996-0.999	0.00058597	0.997354455	0.996-0.999	0.001442321	338.1872456	0.01091321	342.5099316	0.007907117	0.000206869	0.059094266	0.99999999966559	0	0	5.470782E-81

Abbreviation: OR, odds ratio; IVW, inverse variance weighted method; CI, confidential interval; a Pval of 0 means Pval<0.001; P-adj, P-value after P-value adjustment

Table S4:
Detailed MR analyses between inflammatory bowel disease and colorectal cancer.

Mediator	Outcome	Method	Number.of. SNPs	Beta	Se	Pval	LO_CI	UP_CI	OR	OR_LCI9 5	OR_UCI 95	Rucker Q(MR-Egger)	Pval_Q(MR-Eg ger)	Cochran Q(IVW)	Pval_Q(IV W)	Intercept(MR-E gger)	Pval_Interc ept	Power	Steiger P-value	Number of SNP Not Passing MR-Steiger
Inflammatory bowel disease (FinnGen)	Colorectal cancer (Huyghe JR et al.)	MR Egger	37	0.02150 5427	0.1233 01312	0.8625 45406	-0.2201 65145	0.26317 5998	1.0217 38335	0.80238 6277	1.30105 5682	22.51420319	0.949177135	23.80436596	0.9407064 57	-0.020465476	0.2637361 19			
Inflammatory bowel disease (FinnGen)	Colorectal cancer (Huyghe JR et al.)	Inverse variance weighted	37	-0.1071 0114	0.0488 20055	0.0282 50229	-0.2027 88448	-0.0114 13832	0.8984 34801	0.81645 0945	0.98865 1059	22.51420319	0.949177135	23.80436596	0.9407064 57	-0.020465476	0.2637361 19			
Inflammatory bowel disease (FinnGen)	Colorectal cancer (Huyghe JR et al.)	Weighted median	37	-0.0979 91552	0.0671 34831	0.1443 93677	-0.2295 7582	0.03359 2716	0.9066 56563	0.79487 0699	1.03416 3323	22.51420319	0.949177135	23.80436596	0.9407064 57	-0.020465476	0.2637361 19	0.551262560457 053	0.26302582664 6883	9
Inflammatory bowel disease (FinnGen)	Colorectal cancer (Huyghe JR et al.)	Weighted mode	37	-0.1045 54947	0.1131 42721	0.3615 88329	-0.3263 14679	0.11720 4786	0.9007 25304	0.72157 8086	1.12434 9657	22.51420319	0.949177135	23.80436596	0.9407064 57	-0.020465476	0.2637361 19			
Inflammatory bowel disease (FinnGen)	Colorectal cancer (Huyghe JR et al.)	Simple mode	37	-0.0931 96526	0.1322 34134	0.4854 79544	-0.3523 75429	0.16598 2377	0.9110 14444	0.70301 614	1.18055 2296	22.51420319	0.949177135	23.80436596	0.9407064 57	-0.020465476	0.2637361 19			

Mediator	Outcome	Method	Number.of. SNPs	Beta	Se	Pval	LO_CI	UP_CI	OR	OR_LCI9 5	OR_UCI 95	Rucker Q(MR-Egger)	Pval_Q(MR-Eg ger)	Cochran Q(IVW)	Pval_Q(IV W)	Intercept(MR-E gger)	Pval_Interc ept	Power	Steiger P-value	Number of SNP Not Passing MR-Steiger
Inflammatory bowel disease (IIBDGC)	Colorectal cancer (Huyghe JR et al.)	MR Egger	108	-0.0376 69021	0.0623 30897	0.5469 09328	-0.1598 37579	0.08449 9537	0.9630 31632	0.85228 2206	1.08817 2341	81.87799006	0.960414379	82.05510433	0.9650640 1	-0.00379059	0.6747170 54			
Inflammatory bowel disease (IIBDGC)	Colorectal cancer (Huyghe JR et al.)	Inverse variance weighted	108	-0.0615 64989	0.0257 12558	0.0166 49632	-0.1119 61603	-0.0111 68375	0.9402 91835	0.89407 8586	0.98889 376	81.87799006	0.960414379	82.05510433	0.9650640 1	-0.00379059	0.6747170 54			
Inflammatory bowel disease (IIBDGC)	Colorectal cancer (Huyghe JR et al.)	Weighted median	108	-0.0487 20981	0.0413 08497	0.2382 22177	-0.1296 85635	0.03224 3674	0.9524 46844	0.87837 1517	1.03276 9133	81.87799006	0.960414379	82.05510433	0.9650640 1	-0.00379059	0.6747170 54	0.972401462356 547	1.56918906830 179e-128	0
Inflammatory bowel disease (IIBDGC)	Colorectal cancer (Huyghe JR et al.)	Weighted mode	108	-0.0612 28917	0.0625 69323	0.3299 96255	-0.1838 64791	0.06140 6956	0.9406 07894	0.83204 8297	1.06333 1556	81.87799006	0.960414379	82.05510433	0.9650640 1	-0.00379059	0.6747170 54			
Inflammatory bowel disease (IIBDGC)	Colorectal cancer (Huyghe JR et al.)	Simple mode	108	-0.1530 98401	0.0941 70901	0.1069 44545	-0.3376 73366	0.03147 6565	0.8580 45286	0.71342 828	1.03197 7191	81.87799006	0.960414379	82.05510433	0.9650640 1	-0.00379059	0.6747170 54			

Mediator	Outcome	Method	Number.of. SNPs	Beta	Se	Pval	LO_CI	UP_CI	OR	OR_LCI9 5	OR_UCI 95	Rucker Q(MR-Egger)	Pval_Q(MR-Eg ger)	Cochran Q(IVW)	Pval_Q(IV W)	Intercept(MR-E gger)	Pval_Interc ept	Power	Steiger P-value	Number of SNP Not Passing MR-Steiger
Inflammatory bowel disease (Mbatchou J et al.)	Colorectal cancer (Ishigaki K et al)	MR Egger	33	40.0420 0337	16.367 10957	0.0202 95391	7.96246 8606	72.1215 3814	2.4548 3E+17	2871.15 1953	2.1E+31	24.15463174	0.804184132	27.51197218	0.6932711 68	-0.034690447	0.0765283 49	0.537179036646 65	1.11887369862 568e-27	0
Inflammatory bowel	Colorectal cancer	Inverse	33	10.9579	3.9916	0.0060	3.13443	18.7815	57411.	22.9756	1434581	24.15463174	0.804184132	27.51197218	0.6932711	-0.034690447	0.0765283			

disease (Mbatchou J et al.)	(Ishigaki K et al.)	variance weighted		9397	1207	46385	4309	5362	15871	3505	08.5					68	49
Inflammatory bowel disease (Mbatchou J et al.)	Colorectal cancer (Ishigaki K et al.)	Weighted median	33	14.4329	5.6286	0.0103	3.40078	25.4651	185419	29.9876	1.14649	24.15463174	0.804184132	27.51197218	0.6932711	-0.034690447	0.0765283
				6332	61443	4166	6893	3975	8.936	8786	E+11				68		49
Inflammatory bowel disease (Mbatchou J et al.)	Colorectal cancer (Ishigaki K et al.)	Weighted mode	33	15.7340	9.0012	0.0900	-1.9084	33.3765	681100	0.14831	3.12782	24.15463174	0.804184132	27.51197218	0.6932711	-0.034690447	0.0765283
				4964	64046	57281	27893	2717	0.856	3368	E+14				68		49
Inflammatory bowel disease (Mbatchou J et al.)	Colorectal cancer (Ishigaki K et al.)	Simple mode	33	15.3447	10.819	0.1657	-5.8616	36.5511	461468	0.00284	7.48125	24.15463174	0.804184132	27.51197218	0.6932711	-0.034690447	0.0765283
				5436	60319	87577	67902	7662	5.883	6492	E+15				68		49

Abbreviation: beta, effect size; Se, Standard Error; Pval, P-value; CI, confidential interval; LO_CI, lower bound of the confidence interval; UP_CI, upper bound of the confidence interval; OR: odds ratio; OR_LCI95: lower 95% confidence interval for the odds ratio; OR_UCI95: upper 95% confidence interval for the odds ratio; P-adj, P-value after P-value adjustment

Table S5:
Detailed MR analyses between microbiota and colorectal cancer using Huyghe JR et al.’s GWAS data.

Trait	OR(IVW)	95%CI(I VW)	Pval(I VW)	OR(MR-E gger)	95%CI(MR -Egger)	Pval(MR- Egger)	OR(Wei ghted median)	95%CI(W eighted median)	Pval(W eighted median)	OR(Wei ghted mode)	95%CI(W eighted mode)	Pval(W eighted mode)	OR(Sim ple mode)	95%CI(Simple mode)	Pval(Si mple mode)	Rucker Q(MR-Egg er)	Pval_Q(M R-Egger)	Cochran Q(IVW)	Pval_Q (IVW)	Intercept(M R-Egger)	Pval_Inter cept	Power	Steiger P-value	Number of SNP Not Passing MR-Steiger	P-adj(IVW)
class.Actinobacteria.id .419	0.8425815 91	0.810-0. 876	8.02E- 18	0.524617 494	0.432-0.63 7	1.46E-10	0.83538 6014	0.794-0.8 79	4.38427 E-12	0.7827 21879	0.658-0.9 31	5.76E-0 3	0.9790 43962	0.809- 1.185	0.8281 18617	287.45582 31	1 5761	311.395 5761	1 1	0.04026105 6	1.28187E- 06	1	NA	0	1.880244E-16
family.Alcaligenaceae.i d.2875	0.4998519 35	0.439-0. 569	9.7315 6E-26	9.017912 364	4.103-19.8 23	3.95626E -07	0.40395 5307	0.341-0.4 79	1.38995 E-25	0.3922 45201	0.274-0.5 62	1.79277 E-06	0.3901 93776	0.274- 0.556	1.2313 E-06	54.705178 17	0.9987874 84	107.736 1215	0.1111 41203	-0.1692774 71	1.19659E- 10	0.963884 3395543 22	2.72032921 14425e-57	0	4.106060E-24
family.Bifidobacteriac eae.id.433	0.8890102 42	0.854-0. 925	8.23E- 09	1.101375 913	0.913-1.32 9	3.14E-01	0.89644 0696	0.850-0.9 46	5.9974E -05	0.9850 80171	0.829-1.1 70	8.64E-0 1	0.9850 80171	0.824- 1.178	0.8691 07937	152.74626 36	1 5761	157.988 6142	1 1	-0.0185606 37	0.0224617 75	1	NA	0	5.988034E-08
genus..Eubacteriumco prostanoligenesgroup. id.11375	1.6102062 36	1.358-1. 910	4.3759 9E-08	5.874841 011	1.708-20.2 06	7.26E-03	1.78370 811	1.421-2.2 38	5.8333E -07	1.7942 19508	1.139-2.8 28	0.01522 6659	1.7942 19508	1.159- 2.778	0.0117 45154	20.884399 29	0.9994705 02	25.1837 7561	0.9961 88764	-0.1000993 34	0.0437531 85	0.743674 5123592 2	3.72367542 287884e-43	0	2.981226E-07
genus.Bifidobacterium .id.436	0.8640139 83	0.831-0. 898	1.47E- 13	1.099438 845	0.910-1.32 8	3.26E-01	0.83991 041	0.798-0.8 84	1.79251 E-11	0.7924 66079	0.664-0.9 45	9.94E-0 3	0.9798 69497	0.809- 1.187	0.8350 67449	156.44226 5	1 5761	162.956 4645	1 1	-0.0205485 42	0.0109752 72	1	NA	0	2.385923E-12
genus.Lachnoclostridi um.id.11308	0.6037590 37	0.518-0. 704	1.1168 5E-10	2.175147 627	0.422-11.2 20	0.357002 299	0.52838 4365	0.433-0.6 44	2.73448 E-10	0.4920 71727	0.325-0.7 46	0.00143 8731	0.4896 91285	0.322- 0.744	0.0014 27779	29.897830 16	0.9994283 67	32.2629 4109	0.9987 22144	-0.0691232 29	0.1294212 73	0.833545 1113676 07	3.13921574 856696e-50	0	1.074182E-09
genus.Lachnospiracea eND3007group.id.113 17	1.7810454 22	1.237-2. 564	0.0019 05531	0.497274 584	0.000-143 8.000	0.867851 505	1.83505 5662	1.189-2.8 33	0.00615 0304	1.8345 02968	1.013-3.3 23	0.07630 0914	1.8345 02968	0.962- 3.498	0.0984 49733	0.0878213 06	0.9999998 5	0.18647 6508	0.9999 99591	0.07053390 9	0.7614838 95	0.238203 8469483 18	1.29930151 184086e-10	0	6.775551E-03
genus.Peptococcus.id. 2037	0.9243875 39	0.882-0. 969	0.0009 81388	0.567210 125	0.416-0.77 3	4.10E-04	0.90091 2629	0.848-0.9 57	0.00077 6072	0.9000 19848	0.750-1.0 81	0.26006 481	0.8979 93607	0.746- 1.082	0.2581 93909	46.888436 17	1 5761	56.7014 5913	1 1	0.05740556 6	0.0019984 29	0.999999 9999999 97	0	0	3.907035E-03
genus.Ruminococcace aeUCG011.id.11368	1.3708797 86	1.314-1. 430	6.3933 1E-49	0.185559 769	0.125-0.27 6	1.89E-14	1.54199 9391	1.452-1.6 37	1.12745 E-45	1.5460 113	1.298-1.8 42	2.39685 E-06	1.5460 113	1.316- 1.816	3.2487 5E-07	32.887332 37	1 5761	132.228 2065	0.9964 43092	0.2469147 19	7.14474E- 19	1	NA	0	4.494300E-47
genus.Streptococcus.i d.1853	0.4151552 64	0.384-0. 449	4.85E- 109	0.217810 745	0.156-0.30 5	7.97E-16	0.40082 692	0.362-0.4 44	1.30E-6 7	0.4151 42424	0.314-0.5 48	3.85E-0 9	0.4151 42424	0.312- 0.552	9.3670 6E-09	99.493144 24	0.9999988 93	114.418 7101	0.9998 8123	0.06964033 8	0.0001576 13	0.999963 9582804 87	2.40946899 567747e-16 2	0	1.023350E-106
genus.unknowngenus. id.1868	0.8750601 96	0.804-0. 952	0.0019 13126	1.831098 814	0.945-3.54 7	7.56E-02	1.02673 1587	0.919-1.1 47	0.64165 5519	1.0366 99774	0.770-1.3 95	0.81254 3463	1.0366 99774	0.766- 1.403	0.8160 05776	50.669583 22	0.9999999 56	55.5410 2398	0.9999 99462	-0.0590982 17	0.0293080 37	0.999698 7144626 16	4.56291277 433065e-24 4	0	6.775551E-03
order.Bifidobacteriales .id.432	0.8890102 42	0.854-0. 925	8.23E- 09	1.101375 913	0.913-1.32 9	3.14E-01	0.89644 0696	0.850-0.9 46	6.1557E -05	0.9850 80171	0.822-1.1 80	8.71E-0 1	0.9850 80171	0.816- 1.189	0.8754 02067	152.74626 36	1 5761	157.988 6142	1 1	-0.0185606 37	0.0224617 75	1	NA	0	5.988034E-08
phylum.Actinobacteria .id.400	0.8444396 14	0.791-0. 901	3.2869 9E-07	0.839356 902	0.536-1.31 3	0.443920 901	0.82337 6792	0.757-0.8 96	6.30241 E-06	0.7329 93173	0.562-0.9 56	0.02253 8859	0.9697 60409	0.738- 1.273	0.8253 3623	42.548999 2	1 5761	42.5497 1279	1 1	0.00042060 6	0.9787075 22	0.999999 5872300 7	0	0	2.103606E-06

Abbreviation: OR, odds ratio; IVW, inverse variance weighted method; CI, confidential interval; a Pval of 0 means Pval<0.001; P-adj, P-value after P-value adjustment

Table S6:
Detailed MR analyses between microbiota and colorectal cancer using Ishigaki K et al.’s GWAS data.

Trait	OR(IVW)	95%CI(I VW)	Pval(I VW)	OR(MR-E gger)	95%CI(MR -Egger)	Pval(MR- Egger)	OR(Wei ghted median)	95%CI(W eighted median)	Pval(W eighted median)	OR(Wei ghted mode)	95%CI(W eighted mode)	Pval(W eighted mode)	OR(Sim ple mode)	95%CI(Simple mode)	Pval(Si mple mode)	Rucker Q(MR-Egg er)	Pval_Q(M R-Egger)	Cochran Q(IVW)	Pval_Q (IVW)	Intercept(M R-Egger)	Pval_Inter cept	Power	Steiger P-value	Number of SNP Not Passing MR-Steiger	P-adj(IVW)
class.Actinobacteria.id .419	0.9025493 78	0.876-0. 930	9.8392 2E-12	1.077623 79	0.912-1.27 3	0.380675 89	0.94632 2204	0.913-0.9 81	0.00250 8774	0.9308 75644	0.847-1.0 23	0.13570 1783	0.9308 75644	0.839- 1.032	0.1759 00877	751.11906 5	2.73E-12 3456	757.818 3456	1.0667 6E-12	-0.0154147 68	0.0350198 23	1 0.999815 2156420 94	NA 1.00028142 758826e-23 7	0	2.306933E-10 4.388800E-21
family.Alcaligenaceae.i d.2875	1.5076985 92	1.389-1. 636	8.3167 E-23	0.376914 677	0.213-0.66 7	0.001199 3	1.67763 2117	1.499-1.8 77	1.67329 E-19	1.7116 06538	1.339-2.1 87	4.37733 E-05	1.7176 04257	1.344- 2.195	3.9242 9E-05	34.621689 62	0.9999999 65	57.7057 6616	0.9967 81624	0.07984749 3	6.24442E- 06	0.999815 2156420 94	1.00028142 758826e-23 7	0	4.388800E-21
family.Bifidobacteriac eae.id.433	0.9137615 67	0.891-0. 937	1.7922 4E-12	1.325693 074	1.156-1.52 0	6.12702E -05	0.94995 9821	0.919-0.9 82	0.00270 0189	0.9393 61005	0.860-1.0 26	0.16484 4567	0.9244 57269	0.830- 1.029	0.1523 53309	479.00928 45	0.2827441 95	509.534 164	0.0664 81811	-0.0344240 41	9.32049E- 08	1 0.950653 4778961 98	NA 5.97187785 645653e-11 2	0	4.721125E-11 7.712414E-05
genus..Eubacteriumco prostanoligenesgroup. id.11375	1.2217839 12	1.118-1. 336	1.0629 6E-05	0.771490 391	0.314-1.89 3	0.574278 69	1.37138 9214	1.217-1.5 45	2.17193 E-07	1.4037 00167	1.109-1.7 77	0.00734 4196	1.4037 00167	1.095- 1.800	0.0107 25136	20.943670 78	0.9943683 55	21.9612 1767	0.9934 98	0.03564603 4	0.3191649 7	0.950653 4778961 98	5.97187785 645653e-11 2	0	7.712414E-05
genus.Bifidobacterium .id.436	0.8879775 9	0.865-0. 912	1.4696 2E-18	1.308711 09	1.128-1.51 9	0.000435 27	0.94647 7028	0.912-0.9 82	0.00346 3361	0.9263 44743	0.847-1.0 13	0.09478 3079	0.9112 15674	0.826- 1.006	0.0649 62308	604.73178 53	1.63E-03 5484	636.842 5484	7.4055 8E-05	-0.0347979 68	3.15343E- 07	1 0.988992 3787827 42	NA 8.86915171 742314e-16 9	0	6.203400E-17 7.186831E-04
genus.Lachnoclostridiu m.id.11308	0.8086270 52	0.725-0. 901	0.0001 26025	2.052019 352	0.652-6.46 0	0.224303 728	0.79998 784	0.697-0.9 18	0.00148 2784	0.8232 09994	0.618-1.0 97	0.18983 8981	0.8151 45246	0.607- 1.094	0.1789 40665	9.2313558 08	1 8115	11.7872 8115	1 1	-0.0500629 46	0.1154105 98	0.988992 3787827 42	8.86915171 742314e-16 9	0	7.186831E-04
genus.Lachnospiracea eND3007group.id.113 17	2.3683816 62	1.846-3. 038	1.1409 9E-11	0.000748 132	0.000-30.9 15	0.221040 983	2.43170 9716	1.734-3.4 10	2.56481 E-07	2.4322 09797	1.588-3.7 25	0.00273 2675	2.4322 09797	1.619- 3.655	0.0020 55177	4.0509611 89	0.8524965 23	6.26119 0883	0.7135 19294	0.44039683 9	0.1754091 6	0.451543 4423522 09	9.54909078 153761e-17	0	2.405400E-10
genus.Peptococcus.id. 2037	1.2614203 61	1.224-1. 300	1.3680 8E-52	0.870102 21	0.645-1.17 4	0.364176 283	1.26163 1199	1.212-1.3 14	1.23842 E-29	1.2622 13957	1.143-1.3 94	8.2026E -06	1.2622 13957	1.143- 1.394	8.4277 1E-06	40.606419 23	1.00E+00 4201	46.5615 4201	1 1	0.04296131 1	0.0156543 64	1 0.798040 8062700 3	0 2.12387204 981615e-59	0	2.890700E-50 2.837838E-02
genus.Ruminococcace aeUCG011.id.11368	1.2490418 03	1.203-1. 297	2.7457 7E-31	1.631085 751	1.238-2.14 8	0.000629 966	1.23558 1931	1.174-1.3 00	3.73569 E-16	1.2207 82986	1.073-1.3 89	0.00283 5237	1.2207 82986	1.064- 1.401	0.0050 07309	13.853102 75	1 0487	17.5256 0487	1 1	-0.0342663 19	0.0569283 34	1 0.798040 8062700 3	NA 2.12387204 981615e-59	0	2.901250E-29 2.837838E-02
genus.Streptococcus.i d.1853	0.8451662 04	0.745-0. 959	0.0091 4564	0.536098 971	0.326-0.88 0	0.022051 524	0.77468 1121	0.655-0.9 17	0.00293 4932	0.7656 12471	0.647-0.9 06	0.00484 6802	0.7579 39935	0.603- 0.952	0.0259 87427	14.773002 01	0.8717940 27	18.2337 0033	0.7446 92355	0.03894446 4	0.0762652 12	0.798040 8062700 3	2.12387204 981615e-59	0	2.837838E-02
genus.unknowngenus. id.1868	1.0822719 06	1.027-1. 141	0.0031 44005	3.741654 142	1.610-8.69 8	0.002736 902	1.15505 4915	1.077-1.2 39	5.91343 E-05	1.1781 10394	0.980-1.4 17	0.08457 5251	1.1781 10394	0.980- 1.416	0.0839 53545	35.989476 08	1 2591	44.3285 2591	0.9999 99997	-0.0970788 74	0.0046805 74	0.999999 9958491 01	0 2.12387204 981615e-59	0	1.184616E-02
order.Bifidobacteriales .id.432	0.9137615 67	0.891-0. 937	1.7922 4E-12	1.325693 074	1.156-1.52 0	6.12702E -05	0.94995 9821	0.918-0.9 83	0.00308 0105	0.9393 61005	0.859-1.0 27	0.16953 3148	0.9244 57269	0.834- 1.025	0.1366 69591	479.00928 45	0.2827441 95	509.534 164	0.0664 81811	-0.0344240 41	9.32049E- 08	1 0.999999 9999999 8	NA 0	0	4.721125E-11 2.250667E-05
phylum.Actinobacteria .id.400	0.9127191 7	0.879-0. 948	2.5641 5E-06	1.778464 868	1.365-2.31 8	2.86942E -05	0.88998 5361	0.845-0.9 38	1.2194E -05	0.9124 69529	0.760-1.0 95	0.32608 2091	0.8979 02415	0.751- 1.073	0.2373 98868	159.97953 47	0.9999993 97	184.862 9822	0.9997 33533	-0.0478086 97	1.12768E- 06	0.999999 9999999 8	0	0	2.250667E-05

Abbreviation: OR, odds ratio; IVW, inverse variance weighted method; CI, confidential interval; a Pval of 0 means Pval<0.001; P-adj, P-value after P-value adjustment

Table S7:
The Mendelian randomization results of the microbiota and IBD which had mediating effects in CRC.

Id.exposure	Exposure	Id.outcome	Outcome	Method	Nsnp	β	Se	Pval	LO_CI	UP_CI	OR	OR_LCI95	or_UCI95
MiBioGen	family.Alcaligenaceae.id.2875	ebi-a-GCST012880	Colorectal cancer (Huyghe JR et al.)	Inverse variance weighted	92	-0.693443355	0.066113111	9.73156E-26	-0.823025054	-0.563861657	0.499851935	0.439101338	0.569007504
MiBioGen	family.Alcaligenaceae.id.2875	finn-b-K11_IBD	Inflammatory bowel disease (FinnGen)	Inverse variance weighted	91	0.246063224	0.039125338	3.19E-10	0.169377561	0.322748887	1.278980433	1.184567301	1.380918541
finn-b-K11_IBD	Inflammatory bowel disease (FinnGen)	ebi-a-GCST012880	Colorectal cancer (Huyghe JR et al.)	Inverse variance weighted	37	-0.10710114	0.048820055	0.028250229	-0.202788448	-0.011413832	0.898434801	0.816450945	0.988651059
MiBioGen	genus.Streptococcus.id.1853	ebi-a-GCST012880	Colorectal cancer (Huyghe JR et al.)	Inverse variance weighted	176	-0.879102699	0.039626881	4.847457298E-109	-0.956771384	-0.801434013	0.415155264	0.384131098	0.448685082
MiBioGen	genus.Streptococcus.id.1853	ieu-a-31	Inflammatory bowel disease (IIBDGC)	Inverse variance weighted	173	0.348394223	0.021400544	1.375635820E-59	0.306449158	0.390339289	1.416790672	1.358592392	1.477482003
MiBioGen	genus.unknowngenus.id.1868	ebi-a-GCST012880	Colorectal cancer (Huyghe JR et al.)	Inverse variance weighted	116	-0.1334626	0.043005322	0.001913126	-0.217753032	-0.049172168	0.875060196	0.80432406	0.952017209
MiBioGen	genus.unknowngenus.id.1868	ieu-a-31	Inflammatory bowel disease (IIBDGC)	Inverse variance weighted	115	0.066813525	0.023090286	0.003808806	0.021556565	0.112070485	1.069096100	1.021790587	1.118591701
MiBioGen	genus.Peptococcus.id.2037	ebi-a-GCST012880	Colorectal cancer (Huyghe JR et al.)	Inverse variance weighted	198	-0.07862388	0.023855717	0.000981388	-0.125381086	-0.031866675	0.924387539	0.882160659	0.968635717
MiBioGen	genus.Peptococcus.id.2037	ieu-a-31	Inflammatory bowel disease (IIBDGC)	Inverse variance weighted	195	0.153335427	0.012717996	1.790845880E-33	0.128408154	0.178262700	1.165715926	1.137016986	1.195139243
ieu-a-31	Inflammatory bowel disease (IIBDGC)	ebi-a-GCST012880	Colorectal cancer (Huyghe JR et al.)	Inverse variance weighted	108	-0.061564989	0.025712558	0.016649632	-0.111961603	-0.011168375	0.940291835	0.894078586	0.98889376
MiBioGen	class.Actinobacteria.id.419	bbj-a-76	Colorectal Cancer (Ishigaki K et al.)	Inverse variance weighted	503	-0.102531878	0.015058652	9.83922E-12	-0.132046836	-0.07301692	0.902549378	0.876299952	0.929585102
MiBioGen	class.Actinobacteria.id.419	ebi-a-GCST90038683	Inflammatory bowel disease (Mbatchou J et al.)	Inverse variance weighted	599	-0.003461744	0.000150239	1.7847E-117	-0.003756211	-0.003167276	0.996544241	0.996250834	0.996837735
MiBioGen	family.Bifidobacteriaceae.id.433	bbj-a-76	Colorectal Cancer (Ishigaki K et al.)	Inverse variance weighted	464	-0.090185609	0.012792717	1.79224E-12	-0.115259334	-0.065111884	0.913761567	0.891135012	0.936962626
MiBioGen	family.Bifidobacteriaceae.id.433	ebi-a-GCST90038683	Inflammatory bowel disease (Mbatchou J et al.)	Inverse variance weighted	502	-0.00375932	0.000141424	1.097E-155	-0.00403651	-0.00348213	0.996247737	0.995971625	0.996523926
MiBioGen	genus.Eubacteriumcoprostanoligen esgroup.id.11375	bbj-a-76	Colorectal Cancer (Ishigaki K et al.)	Inverse variance weighted	42	0.200312014	0.045484592	1.06296E-05	0.111162214	0.289461814	1.221783912	1.117576178	1.335708434
MiBioGen	genus.Eubacteriumcoprostanoligen esgroup.id.11375	ebi-a-GCST90038683	Inflammatory bowel disease (Mbatchou J et al.)	Inverse variance weighted	48	0.005821893	0.000552299	5.57818E-26	0.004739387	0.006904398	1.005838873	1.004750636	1.006928289
MiBioGen	genus.Bifidobacterium.id.436	bbj-a-76	Colorectal Cancer (Ishigaki K et al.)	Inverse variance weighted	508	-0.118808772	0.01351333	1.46962E-18	-0.145294899	-0.092322646	0.88797759	0.864767237	0.91181091
MiBioGen	genus.Bifidobacterium.id.436	ebi-a-GCST90038683	Inflammatory bowel disease (Mbatchou J et al.)	Inverse variance weighted	547	-0.003345587	0.000150093	4.6159E-110	-0.00363977	-0.003051404	0.996660004	0.996366846	0.996953247
MiBioGen	genus.RuminococcaceaeUCG011.id. 11368	bbj-a-76	Colorectal Cancer (Ishigaki K et al.)	Inverse variance weighted	179	0.2223767	0.019113242	2.74577E-31	0.184914745	0.259838654	1.249041803	1.203115864	1.296720849
MiBioGen	genus.RuminococcaceaeUCG011.id. 11368	ebi-a-GCST90038683	Inflammatory bowel disease (Mbatchou J et al.)	Inverse variance weighted	180	0.001243861	0.000127722	2.06004E-22	0.000993525	0.001494197	1.001244635	1.000994019	1.001495314

MiBioGen	genus.unknowngenus.id.1868	bbj-a-76	Colorectal Cancer (Ishigaki K et al.)	Inverse variance weighted	111	0.079062448	0.02677092	0.003144005	0.026591444	0.131533452	1.082271906	1.026948152	1.140576061
MiBioGen	genus.unknowngenus.id.1868	ebi-a-GCST90038683	Inflammatory bowel disease (Mbatchou J et al.)	Inverse variance weighted	116	0.000738217	0.000256157	0.003952954	0.000236149	0.001240285	1.000738489	1.000236177	1.001241054
MiBioGen	order.Bifidobacteriales.id.432	bbj-a-76	Colorectal Cancer (Ishigaki K et al.)	Inverse variance weighted	464	-0.090185609	0.012792717	1.79224E-12	-0.115259334	-0.065111884	0.913761567	0.891135012	0.936962626
MiBioGen	order.Bifidobacteriales.id.432	ebi-a-GCST90038683	Inflammatory bowel disease (Mbatchou J et al.)	Inverse variance weighted	502	-0.00375932	0.000141424	1.097E-155	-0.00403651	-0.00348213	0.996247737	0.995971625	0.996523926
MiBioGen	phylum.Actinobacteria.id.400	bbj-a-76	Colorectal Cancer (Ishigaki K et al.)	Inverse variance weighted	257	-0.091327036	0.019419049	2.56415E-06	-0.129388372	-0.0532657	0.91271917	0.878632663	0.948128061
MiBioGen	phylum.Actinobacteria.id.400	ebi-a-GCST90038683	Inflammatory bowel disease (Mbatchou J et al.)	Inverse variance weighted	283	-0.004644301	0.00024121	1.30257E-82	-0.005117073	-0.004171529	0.995366467	0.994895997	0.99583716
ebi-a-GCST90038 683	Inflammatory bowel disease (Mbatchou J et al.)	bbj-a-76	Colorectal Cancer (Ishigaki K et al.)	Inverse variance weighted	33	10.95799397	3.99161207	0.006046385	3.134434309	18.78155362	57411.15871	22.97563505	143458108.5

Abbreviation: β , effect size; Se, Standard Error; Pval, P-value; CI, confidential interval; LO_CI, lower bound of the confidence interval; UP_CI, upper bound of the confidence interval; OR: odds ratio; OR_LCI95: lower 95% confidence interval for the odds ratio; OR_UCI95: upper 95% confidence interval for the odds ratio; P-adj, P-value after P-value adjustment

Table S8:
Expression analysis and LASSO regression.

Gene	P-val	P-adj	LASSO Coefficient
ADCYAP1R1	2.031873365E-28	3.11E-27	-0.72041028
AHRR	0.008265042	0.010560887	.
ARHGAP26	2.678206030E-12	6.85E-12	.
ATP11A	6.315772822E-25	4.15E-24	-0.500913326
AXDND1	1.331285748E-15	5.10E-15	.
BCAS3	9.358943149E-10	2.15E-09	.
BVES	1.036509562E-13	3.19E-13	.
DEPDC4	3.644700496E-26	2.79E-25	.
DOCK2	2.855396332E-08	5.98E-08	.
DOCK4	5.132016487E-15	1.69E-14	.
FAM120A	6.172990050E-27	5.68E-26	-0.666156045
GAS2L3	6.027631421E-10	1.46E-09	.
GBE1	6.726785945E-32	3.10E-30	-0.925015451
KIF14	9.621553077E-08	1.77E-07	.
LCT	0.000795597	0.001180563	.
LIN28B	0.001470550	0.001989568	.
MAP3K13	4.009357076E-13	1.15E-12	.
MBLAC2	0.000327083	0.000518821	.
MCM6	1.210432791E-29	2.78E-28	-0.473108172
MEIS2	8.862220358E-07	1.57E-06	.
MGA	2.865189635E-08	5.74E-08	.
MSRA	4.259842744E-20	1.96E-19	0.754643932
NR1H4	8.899476455E-05	0.000146214	.
PMS1	1.526567843E-15	5.41E-15	.
RPTOR	1.2727603145E-27	1.46E-26	.
SLC17A8	0.001101233	0.001535052	.
SLC9A3	3.889350112E-08	7.46E-08	.
SOAT1	2.952278631E-24	1.70E-23	0.226702616
ST8SIA2	0.000792786	0.001215605	.
STIM1	0.000828942	0.001191604	.
TBC1D5	4.745091713E-06	8.09E-06	.
THSD7B	0.001924538	0.002529393	.
TMTC1	1.235057525E-08	2.72E-08	.
UBXN4	1.064349637E-18	4.43E-18	.
WAC	2.315678573E-12	6.28E-12	.
ZDHHC4	2.594442230E-23	1.32E-22	0.02477055
ZRANB3	0.019642490	0.024420393	.

Abbreviation: LASSO, least absolute shrinkage and selection operator; P-val, P-value; P-adj, P-value after P-value adjustment

Table S9:
Detailed results of drug target MR analyses between seven candidate genes and colorectal cancer using Huyghe JR et al.’s GWAS data.

Gene	OR(IV W)	95%CI(IVW)	Pval(IV W)	OR(MR-Eg ger)	95%CI(MR-E gger)	Pval(MR-E gger)	OR(Weighted median)	95%CI(Wei ghted median)	Pval(Wei ghted median)	OR(Weig hted mode)	95%CI(Wei ghted mode)	Pval(Weight ed mode)	OR(Simp le mode)	95%CI(Si mple mode)	Pval(Simp le mode)	Rucker Q(MR-Egger)	Pval_Q(MR-Egg er)	Cochran Q(IVW)	Pval_Q(I VW)	Intercept(MR -Egger)	Pval_Inte rcept	Power	Steiger P-value	Numbe r of SNP Not Passing MR-Ste iger	P-adj(I VW)
ADCYAP 1R1	N/A																								
ATP11A	1.3611 90184	1.173- 1.580	5.0693 6E-05	1.028230 705	0.540-1.958	0.9327543 16	1.441496654	1.179-1.76 3	0.000372 243	1.63270 7759	1.080-2.46 7	0.022893531	1.64146 2443	1.014-2.6 58	0.047805 364	24.62137339	0.99999 9726	25.390887 9	0.99999 9676	0.016497806	0.383456 013	0.862119 38816766 1	2.0696151 513488e-7 6	0	5.0693 60E-05
FAM120 A	0.7521 99959	0.716- 0.791	5.2608 6E-29	1.063850 71	0.913-1.239	0.4271148 44	0.727164415	0.679-0.77 8	4.32525E -20	0.75422 9789	0.632-0.90 1	0.001981141	0.79954 9964	0.649-0.9 84	0.035657 347	147.8554711	1	170.05677 47	1	-0.025196653	3.45551E -06	0.999999 99999988 1	0	0	1.2275 34E-28
GBE1	1.0821 1322	1.052- 1.113	5.5802 7E-08	0.767236 352	0.705-0.835	1.59645E- 09	1.071122207	1.029-1.11 5	0.000873 91	1.05487 1904	0.975-1.14 2	0.186437057	1.06460 4488	0.963-1.1 77	0.222452 882	694.6686905	0.93974 023	765.37895 55	0.38858 1463	0.034515593	2.06047E -16	1	NA	2	6.5103 15E-08
MCM6	0.9689 12025	0.963- 0.975	7.9024 8E-23	0.989283 338	0.978-1.001	0.0751399 88	0.958223515	0.950-0.96 7	3.4972E- 20	0.95834 3078	0.939-0.97 8	2.74846E-05	0.95834 3078	0.934-0.9 84	0.001315 725	1786.909287	1	1803.3570 56	1	-0.00655054	5.12395E -05	1	NA	0	1.3829 34E-22
MSRA	0.9412 17558	0.935- 0.947	1.6245 4E-74	0.990292 47	0.973-1.007	0.2653386 4	0.940962413	0.932-0.95 0	7.64282E -35	0.91388 5368	0.888-0.94 0	6.57851E-10	0.96634 5123	0.929-1.0 05	0.087293 325	3959.601457	1	3998.9321 62	1	-0.008874538	3.82999E -10	1	NA	1	1.1371 78E-73
SOAT1	1.0717 22014	1.063- 1.081	1.8851 6E-57	1.127084 491	1.111-1.143	1.43426E- 59	1.067621907	1.056-1.07 9	7.281E-3 2	1.09501 205	1.072-1.11 9	1.48149E-16	1.06139 8181	1.033-1.0 91	1.89579E- 05	3738.638878	2.44415 E-32	3841.9501 57	2.32643E -38	-0.013473517	3.77085E -18	1	NA	7	6.5980 60E-57
ZDHHC4	0.9498 5642	0.939- 0.961	9.0960 8E-18	0.924521 842	0.907-0.942	4.23124E- 16	0.943841502	0.928-0.96 0	1.79043E -11	0.94559 1717	0.911-0.98 2	0.003393716	0.94559 1717	0.900-0.9 94	0.026884 641	1279.138727	1	1292.2540 11	1	0.006637313	0.000299 956	1	NA	0	1.2734 51E-17

Abbreviation: OR, odds ratio; IVW, inverse variance weighted method; CI, confidential interval; a Pval of 0 means P-value <0.001; P-adj, P-value after P-value adjustment

Table S10:
Detailed results of drug target MR analyses between seven candidate genes and colorectal cancer using Ishigaki K et al.’s GWAS data.

Gene	OR(IV W)	95%CI(IVW)	Pval(IV W)	OR(MR-Eg ger)	95%CI(M R-Egger)	Pval(MR-E gger)	OR(Weighted median)	95%CI(Wei ghted median)	Pval(Wei ghted median)	OR(Weig hted mode)	95%CI(Wei ghted mode)	Pval(Weight ed mode)	OR(Simp le mode)	95%CI(Si mple mode)	Pval(Simp le mode)	Rucker Q(MR-Egger)	Pval_Q(MR-Egg er)	Cochran Q(IVW)	Pval_Q(I VW)	Intercept(MR -Egger)	Pval_Inte rcept	Power	Steiger P-value	Number of SNP Not Passing MR-Ste iger	P-adj(I VW)
ADCYAP1R1													N/A												
ATP11A	0.9445 97882	0.855- 1.044	0.26295 6003	0.9368275 12056681	0.595-1.4 74	0.7787179 61604112	1.0289373931 2216	0.890-1.18 9	0.699388 8730769	1.18394 6448080	0.833-1.68 2	0.349349941 722174	1.17368 7906458	0.815-1.6 91	0.393178 26172341	47.61938352 92017	0.93743 2933999	47.620724 3023931	0.94805 9860580	0.0004902057 46277876	0.970904 7045712	0.99265965 0897038	2.387904 8401672	0 5E-205	3.0678 20E-01
FAM120A	0.8408 63471	0.813- 0.869	1.01822 9135E-2 4	0.9730286 81378124	0.852-1.1 12	0.6878145 10920196	0.9073465245 93624	0.866-0.95 1	4.316398 8589726	0.91206 2142151	0.831-1.00 1	0.052252267 9805121	0.93679 6029836	0.831-1.0 56	0.284822 56767996	270.6944313 83211	0.99172 8750495	275.60861 0087265	0.98678 4584040	-0.010029436 5186095	0.027320 5186095	1	0 0	0 68E-24	2.3758 68E-24
GBE1	0.9286 67517 805	0.912- 0.946	1.17202 424477E -15	0.9357058 68211616	0.884-0.9 90	0.0209580 03191836	0.9134204220 97156	0.888-0.93 9	1.801821 3535058	0.84797 6503757	0.792-0.90 8	3.175959524 11839E-06	0.99725 2542714	0.913-1.0 89	0.951099 25236438	523.6392428 23224	0.99999 9998993	523.71633 3733508	0.99999 9999137	-0.000793695 7400584	0.781357 7400584	1	NA	0 0	2.0510 42E-15
MCM6	0.9687 43043	0.965- 0.972	7.48709 2450E-6 4	0.9699147 88001067	0.963-0.9 77	1.3814437 8494444E- 16	0.9660052359 60148	0.962-0.97 0	5.324713 0064423	0.97004 8295411	0.957-0.98 4	2.065552420 53737E-05	0.97004 8295411	0.956-0.9 84	2.344363 82439955	3345.772192 41229	5.46885 0267568	3345.9578 5138174	6.03414 0850899	-0.000411209 5650029	0.701347 5650029	1	NA	0 0	5.2409 65E-63
MSRA	1.0298 55315	1.025- 1.035	2.10572 7205E-2 9	1.0537129 7358826	1.039-1.0 69	1.5907455 2177676E- 12	1.0326426296 063	1.024-1.04 1	2.244026 7627685	1.04000 4826032	1.002-1.07 9	0.036415202 8330431	0.99910 7315159	0.957-1.0 43	0.967819 90116350	3850.554071 15046	0.99999 9919138	3861.5564 3957649	0.99999 9848781	-0.003712278 4776461	0.000917 4776461	1	NA	0 0	7.3700 45E-29
SOAT1	1.0057 61443	0.997- 1.014	0.17955 8459	1.0207627 3568425	1.005-1.0 37	0.0094250 50460277	0.9746107911 71137	0.964-0.98 5	1.451899 1322963	0.92467 2641794	0.900-0.95 0	1.096817700 71163E-08	1.08996 4600887	1.050-1.1 32	6.985786 14515246	3438.203767 47688	2.90087 0192004	3445.9553 6805253	8.89430 5090209	-0.002906296 8740835	0.026142 8740835	1	NA	0 0	2.5138 18E-01
ZDHHC4	1.0026 68506	0.992- 1.014	0.63212 7021	0.9508015 79640441	0.933-0.9 69	1.4407278 7724616E- 07	0.9855514958 95111	0.970-1.00 1	0.074870 1735430	1.38778 7859936	0.137-14.0 41	0.781405509 707096	1.38778 7859936	0.127-15. 158	0.788233 60674644	1508.650373 94629	6.98132 8053037	1562.0859 0682162	1.12782 6250855	0.0105997573 791103	1.602044 6228059	1	NA	0 0	6.3212 70E-01

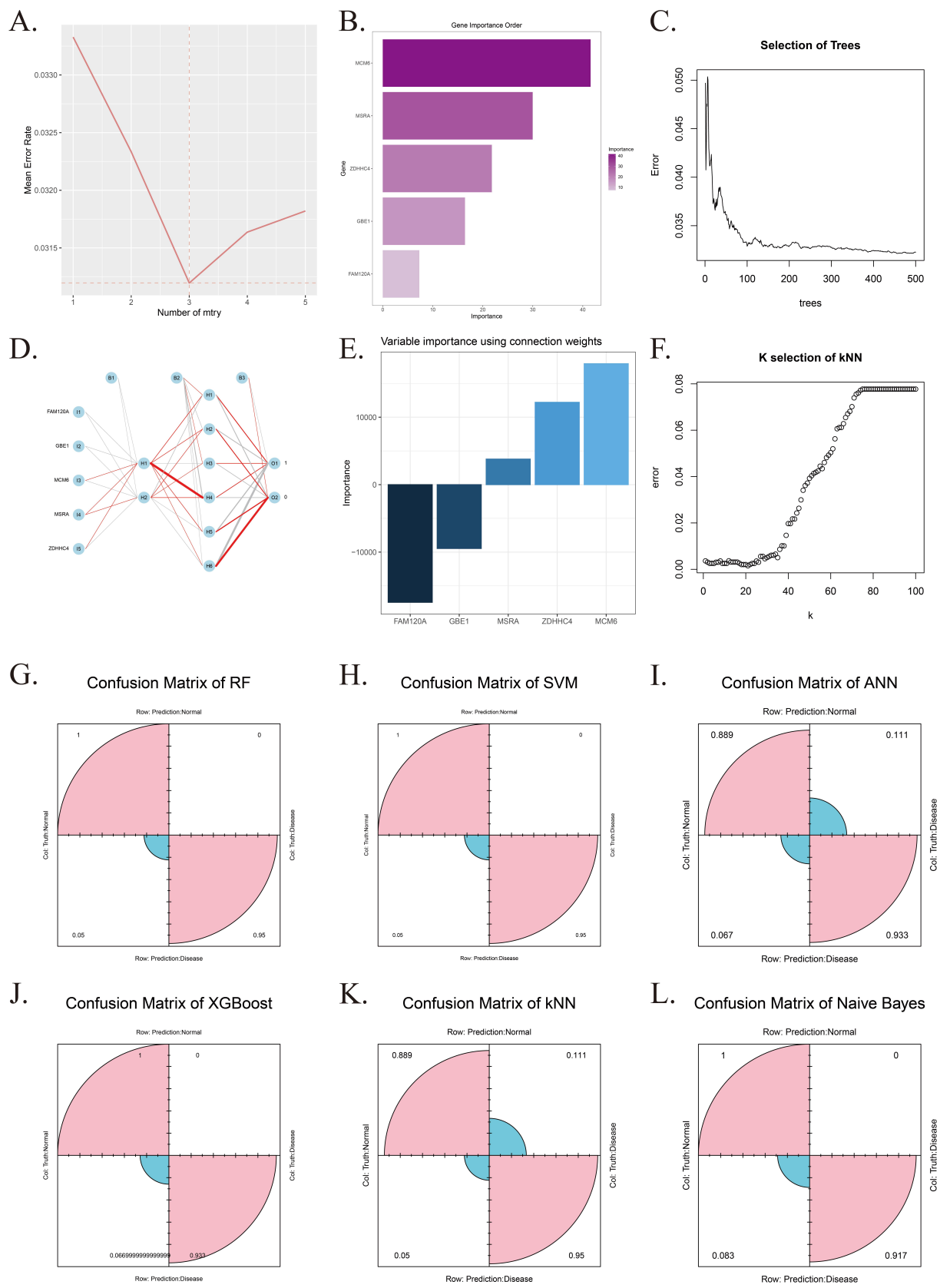
Abbreviation: OR, odds ratio; IVW, inverse variance weighted method; CI, confidential interval; a Pval of 0 means P-value <0.001; P-adj, P-value after P-value adjustment

Table S11:
Detailed results of SMR and HEIDI tests between the expression of seven candidate genes and colorectal cancer using Ishigaki K et al.’s GWAS data.

Gene	ProbeID	ProbeChr	Probe_bp	topSNP	topSNP_chr	topSNP_bp	A1	A2	Freq	Beta_GWAS	Se_GWAS	P_GWAS	Beta_eQTL	Se_eQTL	P_eQTL	Beta_SMR	Se_SMR	P_SMR	P_HEIDI	Nsnp_HEIDI	OR	LO_CI	UP_CI
ADCYAP1R1													N/A										
ATP11A	ENSG00000068650	13	113443062	rs7998551	13	113540649	O	G	0.214712	-0.02	0.026	0.4463998	-0.0931302	0.0116186	1.095995e-15	0.214753	0.280462	0.4438472	0.5920688	18	1.23955568889576	0.715372051411224	2.14783105216759
*FAM120A	ENSG00000165233	9	95867032	rs10821082	9	95943304	C	G	0.385686	0.04	0.021	0.05573012	0.160459	0.0120975	3.75377E-40	0.249286	0.132218	0.05937317	0.2275653	20	1.28310894975785	0.990187182988429	1.6626841926795
GBE1	ENSG00000114480	3	81675081	rs11715625	3	81875938	T	A	0.331014	0.008	2.00E-02	0.6906993	-0.155264	0.00824703	4.571817e-79	-0.0515253	0.128842	0.689223	0.06243605	20	0.949779620233833	0.737819827770496	1.22263090941509
*MCM6	ENSG00000231890	2	136761317	rs2278682	2.00E+00	136743143	C	G	0.209742	-0.045	0.021	0.03522979	-0.208136	0.013788	1.74E-51	0.216205	0.101907	0.03387238	0.1146984	20	1.24135683106881	1.01660361297165	1.51579904141476
*MSRA	ENSG00000079459	8	11674950	rs6601572	8.00E+00	11093337	C	G	0.417495	-0.04	2.00E-02	0.04537013	0.239012	8.71E-03	8.5615E-166	-1.67E-01	0.0838999	4.61E-02	0.2590751	20	0.845898417891167	0.717630975325375	0.997092040329433
SOAT1	ENSG00000186283	1	179058835	rs1128952	1	179065125	T	C	0.2833	-0.035	2.20E-02	0.1191999	-0.218185	1.27E-02	7.646896e-66	1.60E-01	0.101265	1.13E-01	0.5002422	20	1.17399680507382	0.962649981523659	1.43174416950804
*ZDHHC4	ENSG00000164535	7	6486289	rs13235365	7.00E+00	6456091	T	C	0.238569	-0.075	0.027	0.006440951	0.625471	0.00872303	0	-1.20E-01	0.0431998	5.51E-03	0.250002	2.00E+01	0.887000263148598	0.814988229644134	0.965375251087032

Abbreviation: ProbeID, Probe Identifier; ProbeChr, Chromosome on which the probe is located; A1, Effect Allele; A2, Non-effect Allele; Freq, Frequency of the Effect Allele (A1); Se, Standard Error; P_SMR, P-value from Summary data-based Mendelian Randomization analysis; P_HEIDI, P-value from Heterogeneity in Dependent Instruments test; OR: odds ratio; LO_CI, lower bound of the confidence interval; UP_CI, upper bound of the confidence interval. * means that its corresponding gene pass the SMR (p_SMR < 0.1) and HEIDI test (p_HEIDI > 0.05).

Figure S1:



Parameter tuning, variable importance, and model performance evaluation of six machine learning algorithms used for CRC classification.

(A) Selection of the mtry parameter in the random forest (RF) model. The lowest mean error rate occurred at mtry = 3, indicating this as the optimal number of variables randomly sampled at each tree node split. (B) Variable importance in the RF model. MCM6 showed the highest predictive importance, followed by ZGHH4, MRSA, GBE1, and FAM120A, highlighting their contributions to CRC classification. (C) Selection of the number of trees (ntree) in the RF model. The classification error rate began to plateau around 500 trees, indicating model convergence. The final model was trained with ntree = 500 to ensure stability and prevent underfitting. (D) Schematic diagram of the artificial neural network (ANN) model, illustrating the connections between input gene features and output classes. Red and blue lines represent positive and negative connection weights, respectively, with line thickness reflecting the magnitude of influence. (E) Variable importance in the ANN model based on connection weights. MCM6 and ZDHHC4 contributed most positively to CRC prediction, whereas FAM120A and GBE1 showed negative associations. (F) Optimization of the k parameter in the k-nearest neighbors (kNN) model. The lowest error rate was achieved when k = 19. (G–L) Confusion matrix of the six machine learning models: RF, support vector machine (SVM), ANN, XGBoost, kNN, and Naive Bayes. Each matrix shows the distribution of true positives, true negatives, false positives, and false negatives. All models demonstrated high classification accuracy, particularly in discriminating CRC from normal samples.