

Supplementary Material

Supplementary Table 1. Summary of QC pipeline for genotype data
[Genotype QC]

	Before sample QC		After sample QC	
	Removed	Retained	Removed	Retained
-		827,783		409,212
Duplicated SNP	567	827,216	0	409,212
Call rate < 95%, MAF < 0.01	418,004	409,212	2,790 (with HWE < 10^{-7})	406,422
Total	418,571	409,212	2,790	406,422

[Sample QC]

	Removed	Retained
-		36 (with 11 out-of-study subjects)
Call rate (<95%), heterozygosity	1 (= out-of-study subject)	35
Multi-dimensional scaling	0	35
Total	1 (= out-of-study subject)	35 (with 10 out-of-study subjects)

* No study subjects were excluded in the QC procedure.

Supplementary Table 2. The collected disease loci associated with inflammatory bowel disease, Crohn's disease and ulcerative colitis to calculate wGRS (n=218)

	rsID	REF	ALT	Effect (ALT)	PMID
	rs3766606	G	T	-0.122	28067908
	rs6426833	G	A	0.113	28067908
	rs12568930	T	C	-0.129	28067908
	rs76418789	G	A	-0.545	29584801
	rs2651244	G	A	-0.003	28067908
	rs11583043	C	T	0.046	28067908
	rs4845604	G	A	-0.139	28067908
	rs3766920	G	A	0.392	27569725
	rs4656958	A	G	0.076	28067908
	rs1801274	A	G	-0.095	28067908
	rs10798069	G	T	-0.028	28067908
	rs2488397	G	C	0.098	28067908
	rs3024505	G	A	0.190	28067908
	rs59043219	G	A	0.074	28067908
	rs13407913	A	G	0.063	28067908
	rs7608910	A	G	0.139	28067908
	rs6740462	C	A	0.049	28067908
	rs13001325	C	T	-0.085	28067908
	rs4664304	A	G	-0.022	28067908
	rs2111485	A	G	-0.056	28067908
	rs6740847	A	G	-0.092	28067908
	rs1517352	A	C	0.048	28067908
	rs3116494	G	A	-0.055	28067908
	rs2382817	A	C	-0.065	28067908
	rs111781203	C	T	0.054	28067908
	rs6752107	G	A	0.137	28067908
	rs2975788	A	G	0.207	29584801
	rs4676408	G	A	0.101	28067908
IBD	rs35320439	T	C	0.020	28067908
	rs3197999	G	A	0.161	28067908
	rs2593855	C	T	-0.083	28067908
	rs503734	A	G	-0.069	28067908
	rs2073505	G	A	0.086	28067908
	rs6856616	T	C	0.093	28067908
	rs11734570	G	A	0.069	28067908
	rs2472649	A	G	0.052	28067908
	rs2930047	T	C	0.058	28067908
	rs395157	C	T	0.078	28067908
	rs11742570	T	C	0.169	28067908
	rs4703855	C	T	-0.059	28067908
	rs1363907	G	A	0.081	28067908
	rs4836519	T	C	-0.047	28067908
	rs254560	G	A	0.060	28067908
	rs6863411	A	T	0.080	28067908
	rs11741861	A	G	0.178	28067908
	rs56167332	C	A	0.173	28067908
	rs359457	C	T	0.043	28067908
	rs4976646	T	C	0.079	28067908
	rs7773324	G	A	0.019	28067908
	rs17309827	T	G	-0.026	28067908
	rs17119	G	A	0.103	28067908
	rs7775228	T	C	-0.386	29584801
	rs1847472	C	A	-0.079	28067908
	rs12199775	A	G	-0.107	28067908
	rs7758080	A	G	0.053	28067908
	rs1819333	T	G	-0.096	28067908
	rs798502	A	C	-0.072	28067908
	rs11768365	A	G	-0.084	28067908

rs1456896	C	T	0.088	28067908
rs9297145	C	A	-0.080	28067908
rs1734907	A	G	-0.074	28067908
rs2538470	A	G	-0.052	28067908
rs243505	A	G	-0.081	28067908
rs17057051	A	G	-0.044	28067908
rs7011507	G	A	-0.086	28067908
rs7015630	T	C	-0.047	28067908
rs921720	A	G	0.089	28067908
rs1991866	G	C	-0.061	28067908
rs10975003	T	C	-0.211	29584801
rs3731257	G	A	0.157	29063720
rs4743820	C	T	0.064	28067908
rs4366152	T	C	-0.446	29584801
rs10781499	G	A	0.157	28067908
rs1042058	T	C	0.089	28067908
rs34779708	T	G	0.089	28067908
rs1819658	C	T	-0.075	28067908
rs10761659	A	G	0.159	28067908
rs382412	A	C	0.207	29584801
rs2227564	T	C	0.057	28067908
rs1250550	C	A	-0.095	28067908
rs6586030	A	G	0.096	28067908
rs12413565	C	A	-0.211	29584801
rs3740415	G	A	0.041	28067908
rs111456533	G	A	-0.103	28067908
rs907611	G	A	0.066	28067908
rs11229555	G	T	-0.061	28067908
rs11230563	C	T	-0.062	28067908
rs4246215	G	T	0.053	28067908
rs559928	T	C	0.070	28067908
rs11236797	C	A	0.149	28067908
rs6592362	A	G	-0.046	28067908
rs11221332	C	T	0.080	28067908
rs941823	T	C	0.074	28067908
rs3764147	A	G	0.081	28067908
rs9557195	T	C	-0.073	28067908
rs4902642	G	A	-0.047	28067908
rs1569328	C	T	-0.072	28067908
rs8005161	C	T	0.140	28067908
rs16967103	T	C	0.042	28067908
rs17293632	C	T	0.137	28067908
rs7495132	C	T	-0.105	28067908
rs12928665	A	G	-0.174	29584801
rs7404095	T	C	0.066	28067908
rs28449958	G	A	0.095	28067908
rs16953946	T	C	0.207	27569725
rs16940202	T	C	0.113	28067908
rs3091316	G	A	-0.099	28067908
rs12946510	C	T	0.133	28067908
rs4796647	G	A	0.182	29584801
rs3853824	T	C	0.058	28067908
rs1292053	A	G	0.053	28067908
rs17780256	A	C	-0.107	28067908
rs80262450	G	A	0.158	28067908
rs7240004	A	G	-0.073	28067908
rs727088	G	A	-0.038	28067908
rs2024092	G	A	0.128	28067908
rs11879191	G	A	-0.105	28067908
rs17694108	G	A	0.103	28067908
rs11672983	G	A	0.046	28067908
rs4256018	T	G	0.079	28067908

	rs4911259	T	G	0.038	28067908
	rs6088765	T	G	0.033	28067908
	rs1569723	C	A	-0.067	28067908
	rs913678	T	C	-0.069	28067908
	rs259964	A	G	-0.074	28067908
	rs6062496	G	A	0.137	28067908
	rs2823286	G	A	-0.157	28067908
	rs2836878	G	A	-0.164	28067908
	rs7282490	G	A	-0.126	28067908
	rs2266959	G	T	0.122	28067908
	rs5763767	G	A	0.068	28067908
	rs2072711	A	G	0.191	29584801
	rs2413583	C	T	-0.173	28067908
CD	rs488200	A	T	0.270	30500874
	rs3897478	T	C	-0.100	28067908
	rs3180018	C	T	0.095	28067908
	rs7517810	C	T	0.172	28067908
	rs10801047	A	T	-0.041	28067908
	rs11894081	G	T	-0.022	28067908
	rs80078893	C	T	1.556	26891255
	rs1260326	T	C	-0.105	28067908
	rs3749172	A	C	0.372	25489960
	rs4256159	C	T	0.092	28067908
	rs2581828	C	G	-0.094	28067908
	rs724016	A	G	0.086	28067908
	rs6856616	T	C	0.358	23850713
	rs7438704	A	G	0.074	28067908
	rs34804116	C	A	-0.075	28067908
	rs56167332	C	A	0.457	30500874
	rs9358372	G	A	-0.094	28067908
	rs71559680	G	T	0.091	28067908
	rs114985235	T	C	0.811	26891255
	rs9271060	C	G	-1.050	26891255
	rs184950714	G	A	0.718	30500874
	rs7775228	T	C	0.372	27569725
	rs293526	C	T	-0.755	26891255
	rs9491697	A	G	0.096	28067908
	rs212388	C	T	-0.107	28067908
	rs394522	A	C	0.285	27569725
	rs10486483	G	A	0.065	28067908
	rs12677663	G	T	0.030	28067908
	rs3731257	G	A	0.199	29063720
	rs9408254	A	G	0.089	28067908
	rs78898421	G	T	0.742	30500874
	rs224136	C	T	-0.462	30500874
	rs11195128	C	T	0.113	28067908
	rs10734105	G	A	0.017	28067908
	rs11229030	C	T	-0.017	28067908
	rs2231884	C	T	0.083	28067908
	rs11235604	C	T	0.476	23850713
	rs11235667	A	G	0.378	23850713
	rs630923	C	A	-0.104	28067908
	rs7329174	A	G	0.239	23266558
	rs80244186	T	C	0.125	28067908
	rs2062305	G	A	-0.082	28067908
	rs2026029	A	G	0.052	28067908
	rs529866	C	T	-0.108	28067908
	rs10492862	C	A	0.107	28067908
	rs10775412	A	C	-0.095	28067908
	rs6503695	T	C	0.351	27569725

	rs9319943	T	C	0.080	28067908
	rs4802307	G	T	-0.090	28067908
	rs2284553	A	G	0.128	28067908
	rs4821544	T	C	0.097	28067908
	rs2072711	A	G	0.239	29584812
	rs727563	C	T	-0.065	28067908
	rs6426833	G	A	0.217	28067908
	rs10797432	C	T	-0.102	28067908
	rs1801274	A	G	-0.462	19915573
	rs6738825	A	G	0.024	28067908
	rs1811711	C	G	-0.130	28067908
	rs4676408	G	A	0.143	28067908
	rs9847710	T	C	0.063	28067908
	rs616597	A	C	0.039	28067908
	rs2189234	T	G	-0.070	28067908
	rs4957048	G	A	-0.101	28067908
	rs3194051	A	G	0.055	28067908
	rs17656349	C	T	0.090	28067908
	rs7192	T	G	0.451	29733354
	rs9271366	G	A	0.742	23511034
	rs943072	G	T	-0.100	28067908
	rs2858829	A	G	0.072	28067908
UC	rs1077773	G	A	0.072	28067908
	rs10486483	G	A	0.001	28067908
	rs4380874	T	C	-0.151	28067908
	rs38904	T	C	-0.071	28067908
	rs4728142	G	A	0.100	28067908
	rs7911117	T	G	-0.134	28067908
	rs483905	G	A	0.084	28067908
	rs11612508	A	G	0.061	28067908
	rs11168249	T	C	0.084	28067908
	rs17085007	T	C	0.156	28067908
	rs28374715	A	G	-0.094	28067908
	rs1728785	A	C	0.104	28067908
	rs16940186	T	C	0.445	23511034
	rs4802307	G	T	-0.029	28067908
	rs138788	G	A	0.090	28067908
	rs5771069	A	G	0.100	28067908

The effect estimates (beta) of disease variants in Asian or European populations were retrieved from previous genome-wide association studies (PubMed IDs: 19915573, 23266558, 23511034, 23850713, 25489960, 26891255, 27569725, 28067908, 29063720, 29584801, 29584812, 29733354, 30500874). For the loci associated in both Asian and European populations, effect estimates in Asians were used in the wGRS calculations. To eliminate the correlation among the variants, we pruned the collected variants at an r^2 threshold of 0.2, bringing the total number of disease variants to 133 IBD variants, 53 CD variants, and 32 UC variants were used. Thus, the table includes only independent variants with disease association p -values $< 5e-08$.

REF, reference allele; ALT, alternative allele

Supplementary Table 3. Summary of sequencing coverage statistics in coding regions

Sample ID	Sequenced bases (Gb)	Median depth	Mean depth	Bases with coverages of		
				≥ 10x (%)	≥ 15x (%)	≥ 30x (%)
F1-I-1	6.47	68	83.00	94.64	93.37	85.50
F1-II-1	5.43	54	64.45	94.43	92.68	80.67
F1-I-2	5.71	56	67.19	94.70	93.50	83.40
F2-II-1	6.17	63	73.23	94.75	93.59	85.65
F2-II-2	6.34	63	72.82	94.69	93.60	85.82
F2-I-1	6.45	66	76.55	94.71	93.68	86.76
F3-I-1	5.49	55	63.73	94.30	92.63	81.29
F3-II-1	6.38	63	71.66	94.81	93.73	86.08
F3-I-2	5.87	58	67.75	94.57	93.22	83.79
F4-I-2	5.62	54	63.89	94.27	92.55	80.80
F4-II-1	6.44	63	76.07	94.87	93.75	85.58
F4-I-1	5.12	49	59.10	94.35	92.33	77.53
F5-II-1	6.50	64	77.07	94.86	93.72	85.73
F5-II-2	5.71	58	69.53	94.58	93.07	82.53
F5-I-1	6.32	63	74.80	94.70	93.63	85.97
F6-I-2	6.69	69	81.45	94.73	93.78	87.38
F6-II-1	6.28	64	76.60	94.73	93.60	85.65
F6-I-1	6.43	65	77.80	94.81	93.72	86.09
F7-II-1	5.81	55	65.82	94.45	92.77	81.34
F7-II-3	5.55	53	64.64	94.51	92.75	79.93
F7-II-2	5.82	55	64.05	94.44	92.83	81.42
F8-II-1	5.28	51	61.61	94.54	92.82	79.39
F8-II-2	5.71	57	66.47	94.55	93.19	83.48
F8-I-1	6.68	66	75.76	94.86	93.93	87.19
F8-I-2	5.20	50	57.57	94.29	92.48	79.41

Supplementary Table 4. Genes with rare deleterious alleles that were detected in multiple families with familial Crohn's disease and shared among affected FDRs only.

Gene	Number of families with any variants	Variant name	Variation location (hg38)	Allele	Most affected transcript	Amino acid position	Amino acids	SIFT	PolyPhen-2
<i>AC113554.1</i>	2	rs373319603	chr17:63494457	A	ENST00000290866	1123	V/M	deleterious	probably damaging
		rs538659872	chr17:63496827	T	ENST00000290866	1178	P/L	deleterious	possibly damaging
<i>ACE</i>	2	rs373319603	chr17:63494457	A	ENST00000577647	549	V/M	deleterious	possibly damaging
		rs538659872	chr17:63496827	T	ENST00000577647	604	P/L	deleterious	possibly damaging
<i>AKAP9</i>	2	rs201048693	chr7:92096722	G	ENST00000394534	1101	R/G	deleterious	benign
		rs777536948	chr7:92058368	C	ENST00000491695	67	I/T	deleterious	benign
<i>ANK2</i>	2	chr4_113335848_C/G	chr4:113335848	G	ENST00000357077	1128	L/V	deleterious	probably damaging
		rs759647520	chr4:113323785	G	ENST00000506722	955	D/G	deleterious low confidence	probably damaging
<i>ASIC3</i>	2	rs777980557	chr7:151051247	G	ENST00000490540	12	M/R	deleterious	probably damaging
<i>DUS3L</i>	2	rs375298428	chr19:5785332	A	ENST00000590343	112	T/M	deleterious low confidence	benign
		rs757677248	chr19:5791086	G	ENST00000590110	19	G/E	deleterious low confidence	probably damaging
<i>NUTM2F</i>	2	rs12237784	chr9:94318485	G	ENST00000253262	751	K/Q	deleterious	benign
		rs773143235	chr9:94318511	A	ENST00000253262	742	D/V	deleterious	possibly damaging
<i>ZC3H4</i>	2	rs144221330	chr19:47086476	T	ENST00000253048	260	G/S	tolerated	probably damaging
		rs373777449	chr19:47066888	A	ENST00000601973	734	A/V	deleterious	probably damaging

The effect of the amino acid substitution on protein function was predicted using SIFT and PolyPhen-2. In Polyphen-2, the predicted damaging impact is increasing in the order of 'benign', 'possibly damaging', and 'probably damaging'.

Supplementary Table 5. Tissue specificity and relative expression levels in relevant tissues of identified candidate genes.

Candidate genes	Tissue-specificity	mean TPM (colon)	mean TPM (small intestine)
AC113554.1	Not-expressed	0	0
ACE	Tissue-Enhanced (small intestine)	3.7	56
AKAP17A	Expressed-In-All	62	77
AKAP9	Expressed-In-All	18	14.5
ANK2	Tissue-Enhanced (Brain)	11	4
ASB16	Tissue-Enhanced (Muscle)	2.5	3
ASIC3	Expressed-In-All	9.5	6
DNPH1	Expressed-In-All	46.5	61
DUS3L	Expressed-In-All	21	27
FAM200A	Expressed-In-All	5.5	5
FZD10	Tissue-Enhanced (spleen)	0.2	1
LAMA5	Expressed-In-All	109	35
NUTM2F	Not-expressed	0	0
PKN1	Expressed-In-All	71.5	58
PRR26	Mixed	3	3
WDR66	Tissue-Enhanced (Testis)	0.25	0.3
ZC3H4	Expressed-In-All	18	19

All information was from TissueEnrich (<https://tissueenrich.gdcb.iastate.edu/>) using the relative expression level in GTEx data sets. The tissue-specificity was defined according to an algorithm used in the Human Protein Atlas (Uhlén et al. 2015).

TPM, Transcripts Per Million mapped reads; Not-expressed, Genes with an expression level less than 1 (TPM or FPKM) across all the tissues; Tissue-Enhanced, Genes with an expression level greater than or equal to 1 (TPM or FPKM) that also have at least five-fold higher expression levels in a particular tissue compared to the average levels in all other tissues, but not always higher in pairwise comparison; Expressed in all, Genes with an expression level greater than or equal to 1 (TPM or FPKM) across all of the tissues that are not in tissue-specific groups; Mixed, Genes that are not assigned to any of other groups.

Supplementary Table 6. Relative abundance of gut microbiota at different taxonomic levels between the patients with inflammatory bowel disease and the unaffected first-degree relatives

ID	Taxon	Level	LFC	LFCSE	P-value	False discovery rate-corrected P-value
1	<i>Deltaproteobacteria</i>	Class	-2.170	0.832	9.1.E-03	4.1.E-02
2	<i>Gammaproteobacteria</i>	Class	2.471	0.896	5.8.E-03	4.1.E-02
3	<i>Turicibacterales</i>	Order	2.630	0.894	3.3.E-03	1.8.E-02
4	<i>Pasteurellales</i>	Order	3.471	1.136	2.2.E-03	1.8.E-02
5	<i>Desulfovibrionales</i>	Order	-2.274	0.839	6.7.E-03	2.5.E-02
6	<i>Turicibacteraceae</i>	Family	3.001	0.890	7.5.E-04	8.6.E-03
7	<i>Veillonellaceae</i>	Family	2.724	0.788	5.4.E-04	8.6.E-03
8	<i>Pasteurellaceae</i>	Family	3.615	1.129	1.4.E-03	1.0.E-02
9	<i>Peptostreptococcaceae</i>	Family	2.353	0.798	3.2.E-03	1.8.E-02
10	<i>Desulfovibrionaceae</i>	Family	-2.316	0.855	6.8.E-03	3.1.E-02
11	<i>Veillonella</i>	Genus	4.354	1.193	2.6.E-04	1.0.E-02
12	<i>Turicibacter</i>	Genus	2.873	0.899	1.4.E-03	1.9.E-02
13	<i>Haemophilus</i>	Genus	3.744	1.137	9.9.E-04	1.9.E-02
14	<i>unclassified Bacteroides</i>	Species	-4.443	0.900	8.0.E-07	4.6.E-05
15	<i>Ruminococcus torques</i>	Species	-3.960	0.844	2.7.E-06	7.9.E-05
16	<i>Veillonella dispar</i>	Species	4.264	1.243	6.0.E-04	1.2.E-02
17	<i>Bacteroides uniformis</i>	Species	-3.510	1.057	9.0.E-04	1.3.E-02
18	<i>Haemophilus parainfluenzae</i>	Species	3.711	1.155	1.3.E-03	1.5.E-02
19	<i>unclassified Turicibacter</i>	Species	2.747	0.909	2.5.E-03	2.4.E-02
20	<i>Blautia producta</i>	Species	2.679	0.943	4.5.E-03	3.7.E-02
21	<i>unclassified Bilophila</i>	Species	-2.385	0.851	5.1.E-03	3.7.E-02
22	<i>unclassified Roseburia</i>	Species	-2.145	0.788	6.5.E-03	4.2.E-02

A positive value of log₂ fold change represents an increase of a certain taxon, and vice versa.

De novo strain was depicted as “unclassified”.

LFC, log₂ (fold change); LFCSE, standard error of log₂ (fold change).

Supplementary Table 7. Relative abundance of gut microbiota at different taxonomic levels between the patients with ulcerative colitis and the unaffected first-degree relatives.

ID	Taxon	Level	LFC	LFCSE	P-value	False discovery rate-corrected P-value
1	<i>Turicibacterales</i>	Order	3.861	1.038	2.0.E-04	1.8.E-03
2	<i>Peptostreptococcaceae</i>	Family	-4.304	1.247	5.6.E-04	5.0.E-03
3	<i>Paraprevotellaceae</i>	Genus	3.384	0.944	3.4.E-04	5.0.E-03
4	<i>unclassified Rikenellaceae</i>	Genus	-4.262	1.266	7.6.E-04	2.4.E-02
5	<i>unclassified Bacteroides</i>	Species	-3.251	0.992	1.0.E-03	2.4.E-02
6	<i>unclassified Clostridium</i>	Species	-4.021	1.355	3.0.E-03	3.9.E-02
7	<i>Eubacterium dolichum</i>	Species	3.332	1.136	3.4.E-03	3.9.E-02

A positive value of log₂ fold change represents an increase of a certain taxon, and vice versa.

De novo strain was depicted as “unclassified”.

LFC, log₂ (fold change); LFCSE, standard error of log₂ (fold change).

Supplementary Table 8. Relative abundance of gut microbiota at different taxonomic levels between the patients with Crohn’s disease and the unaffected first-degree relatives.

ID	Taxon	Level	LFC	LFCSE	P-value	False discovery rate-corrected P-value
1	<i>Gammaproteobacteria</i>	Class	3.558	1.094	1.1.E-03	9.1.E-03
2	<i>Enterobacteriales</i>	Order	4.772	1.229	1.0.E-04	8.3.E-04
3	<i>Enterobacteriaceae</i>	Family	4.931	1.273	1.1.E-04	1.8.E-03
4	<i>Veillonellaceae</i>	Family	3.072	0.921	8.5.E-04	7.2.E-03
5	<i>Shigella</i>	Genus	6.157	1.340	4.3.E-06	1.4.E-04
6	<i>Shigella sonnei</i>	Species	6.131	1.362	6.7.E-06	3.2.E-04
7	<i>Bacteroides fragilis</i>	Species	4.561	1.134	5.8.E-05	1.4.E-03
8	<i>unclassified Faecalibacterium</i>	Species	-3.260	0.944	5.5.E-04	8.9.E-03
9	<i>unclassified Faecalibacterium</i>	Species	-3.673	1.274	3.9.E-03	4.7.E-02

A positive value of log₂ fold change represents an increase of a certain taxon, and vice versa.

De novo strain was depicted as “unclassified”.

LFC, log₂ (fold change); LFCSE, standard error of log₂ (fold change).