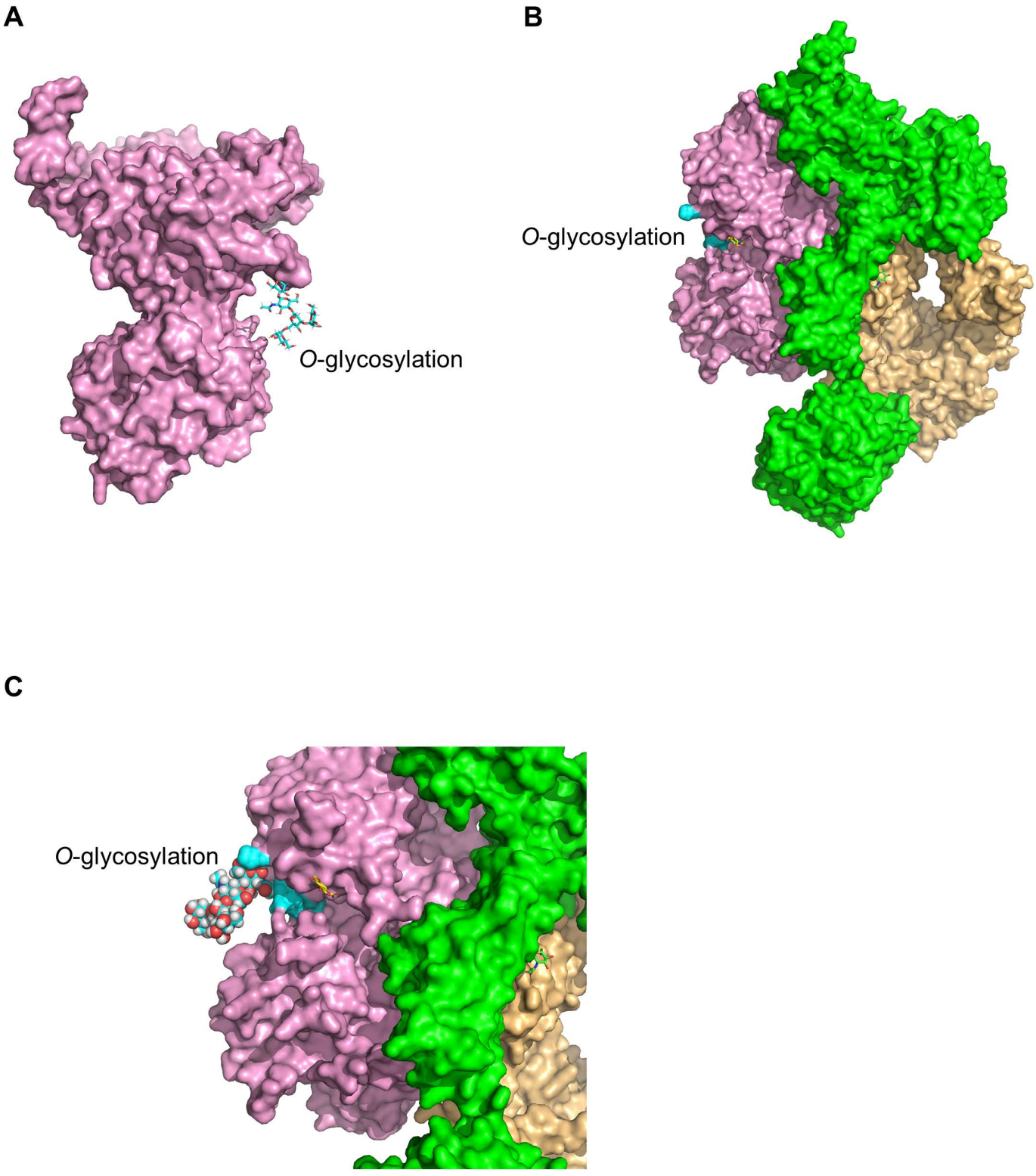
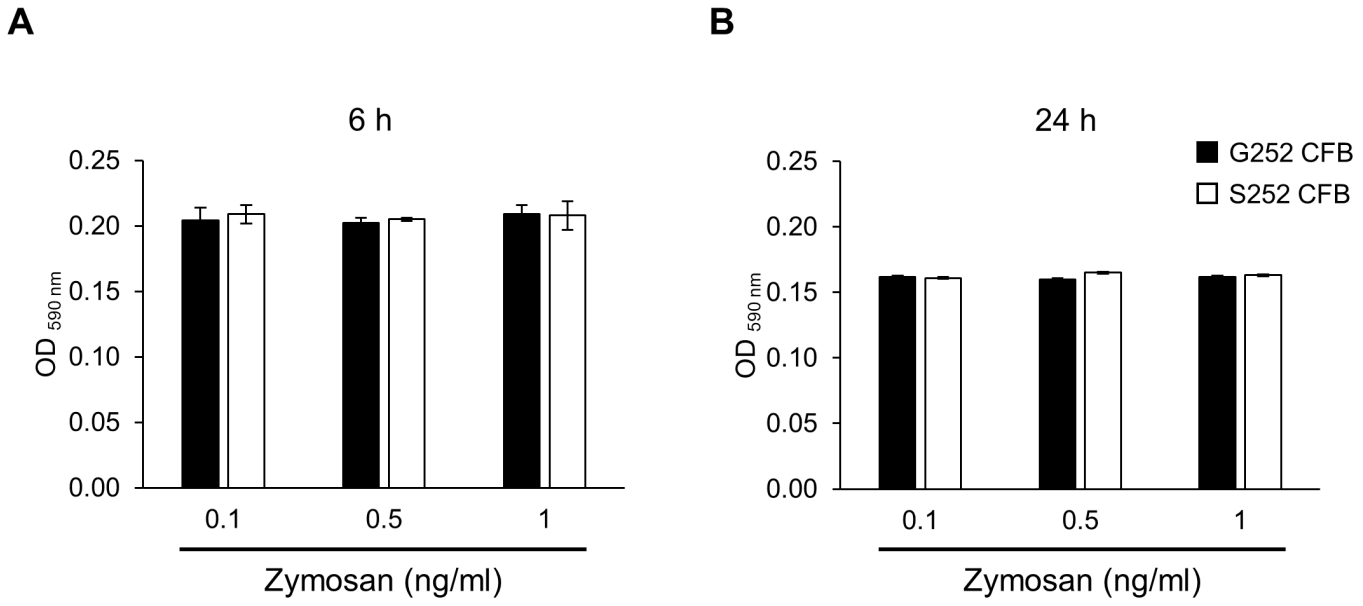
Supplementary Figure 1



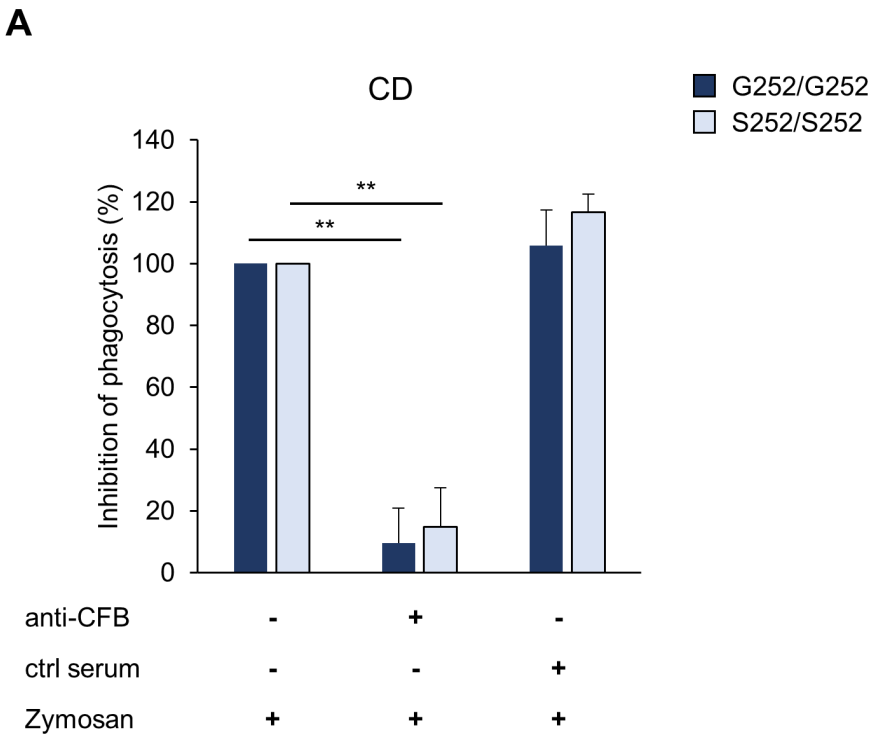
Supplementary Figure 1. G252S CFB does not impact cell viability compared to WT CFB. (A) Model of S252 CFB with *O*-glycosylation at serine 252. **(B)**. Model of glycosylated S252 CFB (light magenta) in complex with complement factor C3b. The beta and alpha chains of C3b are depicted in green and beige, respectively. The site of S252 is depicted in teal. **(C)** Zoomed view of *O*-glycosylation site S252 (teal) and attached sugar. Figures were generated using Pymol.

Supplementary Figure 2



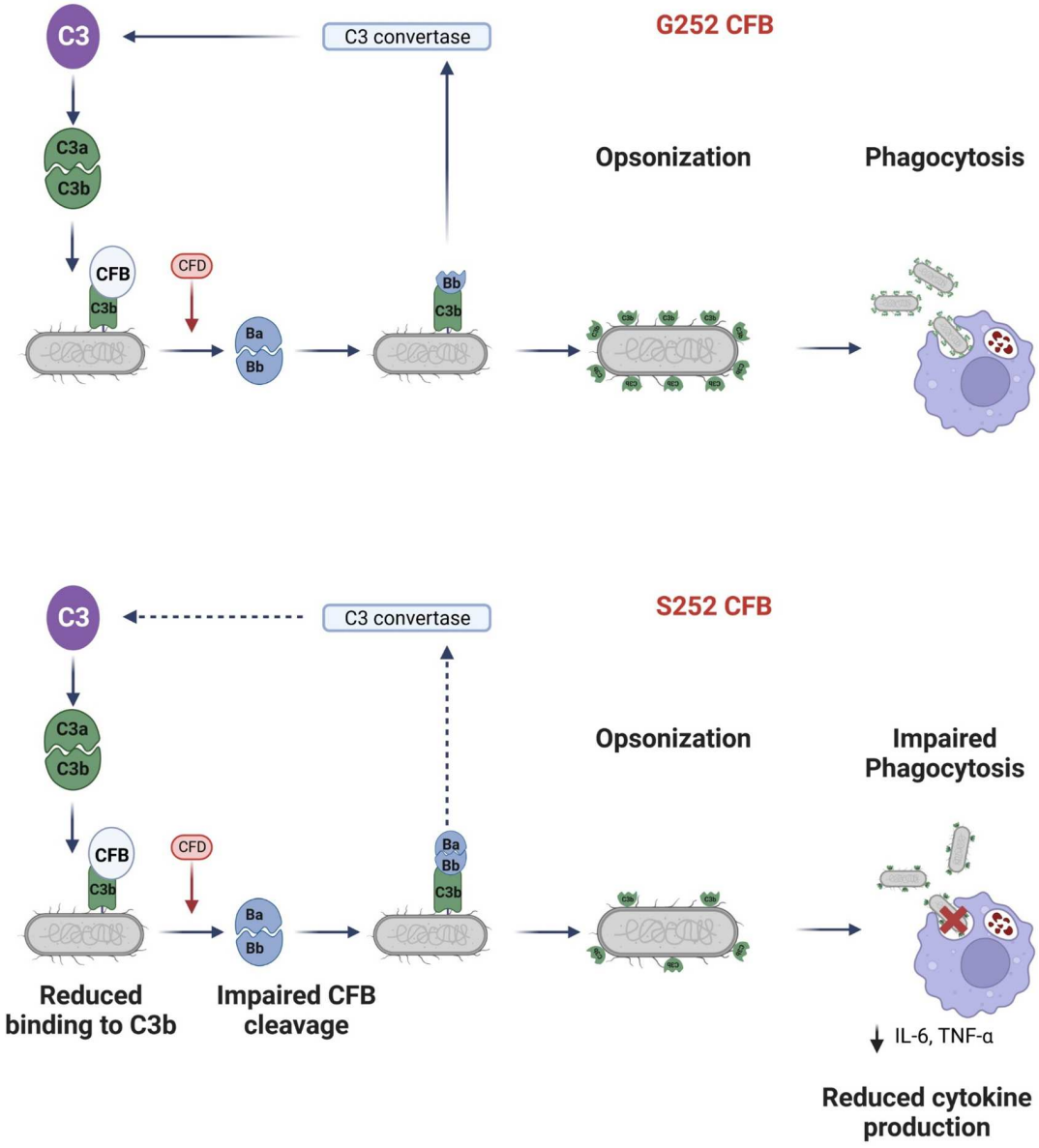
Supplementary Figure 2. G252S CFB does not impact cell viability compared to WT CFB. PBMC-derived macrophages were incubated with WT CFB or G252S CFB (250 ng/ml) in the presence of zymosan (0.1, 0.5, 1 ng/ml) for 6 (**A**) or 24 h (**B**). MTT assays were performed to determine cell viability. Data are presented as means ± SD. Data are representative of three independent experiments.

Supplementary Figure 3



Supplementary Figure 3. Neutralization of CFB decreases macrophage phagocytosis. Macrophages from healthy donors were incubated with either anti-CFB or control serum and fluorescently labeled zymosan particles for 15 min. in the presence of 10 % serum from CD patients (S252/S252 n = 3, G252/G252 n = 3) **(A)**. Phagocytosis was measured by flow cytometry. The percentage of phagocytic cells for zymosan treated macrophages was set to 100 %. Data are presented as means ± SD. ** $p < 0.01$.

Supplementary Figure 4



Supplementary Figure 4. Perianal CD-associated risk variant rs4151651 leads to glycine to serine amino acid substitution in CFB. Serine 252 CFB confers significantly impaired binding to complement factor 3b and subsequent cleavage of CFB into Ba and Bb subunits. Reduced opsonization of bacteria with C3bBb complexes leads to reduced phagocytosis and subsequent cytokine release.

Supplementary Table 1. Pre- and Post-QC Numbers of Study Samples and SNPs Across Cohorts

Cohorts	Pre-QC			Post-QC		
	pCD+	pCD-	SNPs	pCD+	pCD-	SNPs
CSMC	780	1679	181,570	719	1596	108,231
IIBDGC[10 12]	3351	8653	156,446	2974	7764	122,607
SHARE[13]	431	1983	253,702	363	1728	208,151
Total	**4,562**	**12,315**	**n/a**	**4,056**	**11,088**	**n/a**

pCD+: CD patients with perianal involvement; pCD-: CD patients without perianal involvement.

Supplementary Table 2. Top genetic associations with pCD

						Meta-Analysis				CSMC					IIBDGC					SHARE				
CHR	Position (MB)	SNP	Gene	Location	Novel (Y/N)	Allele	*P*-value	OR	95% CI	Allele	*P*-value	OR	95% CI	MAF	Allele	*P*-value	OR	95% CI	MAF	Allele	*P*-value	OR	95% CI	MAF
6	32.41	rs17496549	*HLA-DRA*	intronic	N	A	4.68E-07	1.24	1.14-1.35	A	5.64E-05	1.54	1.25-1.89	0.09	A	8.15E-04	1.18	1.07-1.30	0.11	A	4.09E-02	1.29	1.01-1.65	0.11
6	31.92	rs4151651	*CFB*	exonic	N	A	9.35E-06	1.32	1.17-1.49	A	2.59E-04	1.75	1.30-2.37	0.04	A	5.90E-03	1.22	1.06-1.41	0.04	A	4.13E-02	1.39	1.01-1.91	0.05
6	32.37	rs2294884	*BTNL2*	intronic	N	A	1.79E-05	0.86	0.80-0.92	C	6.42E-03	1.28	1.07-1.52	0.15	C	2.62E-03	1.13	1.04-1.22	0.17	C	2.42E-02	1.27	1.03-1.55	0.17
4	47.28	rs13127214	*GABRB1*	intronic	Y	C	5.04E-05	1.2	1.10-1.31	G	2.15E-01	0.88	0.71-1.08	0.10	G	6.84E-05	0.82	0.74-0.90	0.11	G	4.48E-01	0.90	0.69-1.18	0.10
1	70.23	rs17325887	*LRRC7*	intronic	Y	A	8.57E-05	0.73	0.63-0.86	G	7.32E-02	1.42	0.97-2.08	0.02	G	3.65E-03	1.30	1.09-1.56	0.03	G	1.39E-02	1.73	1.12-2.69	0.03
8	126.53	rs6470362	*TRIB1*	intergenic	N	A	9.55E-05	0.89	0.84-0.94	G	1.40E-01	1.11	0.97-1.27	0.34	G	9.43E-04	1.12	1.05-1.20	0.26	G	7.09E-02	1.18	0.99-1.42	0.29
19	55.22	rs1325156	*LILRB4, LILRP2*	intergenic	N	A	1.16E-04	1.15	1.07-1.23	A	6.23E-01	1.04	0.88-1.24	0.16	A	1.58E-04	1.17	1.08-1.27	0.16	A	9.44E-02	1.20	0.97-1.49	0.15
11	71.14	rs1790353	*DHCR7*	intergenic	Y	A	1.44E-04	0.88	0.83-0.94	A	2.89E-01	0.92	0.79-1.07	0.23	A	5.90E-04	0.88	0.81-0.94	0.22	A	8.70E-02	0.84	0.69-1.03	0.22
21	46.08	rs7275203	*TSPEAR*	intronic	N	A	1.66E-04	0.89	0.84-0.95	G	2.13E-01	1.09	0.95-1.26	0.25	G	7.56E-04	1.13	1.05-1.21	0.29	G	1.37E-01	1.15	0.96-1.37	0.27
4	103.62	rs227375	*MANBA*	intronic	N	A	1.91E-04	0.9	0.86-0.95	A	4.71E-02	0.88	0.78-1.00	0.48	A	7.79E-03	0.92	0.87-0.98	0.48	A	1.97E-02	0.82	0.70-0.97	0.50
1	98.83	rs4542263	*LOC729987 SNX7*	intergenic	Y	A	2.50E-04	1.19	1.08-1.31	C	4.21E-02	0.80	0.64-0.99	0.10	C	8.84E-03	0.87	0.78-0.96	0.09	C	2.98E-02	0.71	0.53-0.97	0.09
4	123.26	rs6851362	*KIAA1109*	intronic	N	A	2.58E-04	0.89	0.84-0.95	G	9.64E-02	1.13	0.98-1.31	0.24	G	1.11E-03	1.12	1.05-1.20	0.25	G	4.01E-01	1.08	0.90-1.31	0.24
12	123.65	rs34484751	*MPHOSPH9*	intronic	Y	A	2.67E-04	0.76	0.65-0.88	C	3.85E-01	1.20	0.8-1.79	0.02	C	4.18E-04	1.35	1.14-1.59	0.03	C	2.94E-01	1.29	0.8-2.07	0.03
4	123.15	rs6848868	*KIAA1109*	exonic	N	A	2.74E-04	0.85	0.78-0.93	G	4.31E-02	1.22	1.01-1.48	0.10	G	2.78E-03	1.17	1.05-1.29	0.09	G	3.26E-01	1.15	0.87-1.53	0.09
21	44.5	rs2850146	*CBS, U2AF1*	intergenic	Y	C	2.94E-04	0.84	0.77-0.92	C	1.39E-02	0.75	0.59-0.94	0.09	C	2.51E-03	0.85	0.77-0.94	0.09	C	7.16E-01	0.95	0.71-1.26	0.09
11	113.42	rs10736469	*DRD2, TMPRSS5*	intergenic	Y	A	3.04E-04	0.91	0.86-0.96	G	1.32E-02	1.17	1.03-1.33	0.41	G	3.70E-03	1.09	1.03-1.16	0.39	G	5.80E-01	1.05	0.89-1.23	0.41
6	32.38	rs3763313	*BTNL2, HLA-DRA*	intergenic	N	A	3.06E-04	0.89	0.83-0.95	C	6.37E-02	1.15	0.99-1.33	0.21	C	2.67E-03	1.12	1.04-1.20	0.20	C	2.59E-01	1.12	0.92-1.36	0.20
4	123.52	rs975403	*IL2,IL21*	intergenic	N	A	3.39E-04	1.10	1.05-1.16	A	1.14E-01	1.11	0.98-1.26	0.40	A	1.16E-03	1.11	1.04-1.17	0.41	A	4.48E-01	1.07	0.9-1.26	0.39
2	241.58	rs34228697	*GPR35, AQP12B*	intergenic	N	A	3.88E-04	1.14	1.06-1.22	G	1.62E-02	0.80	0.67-0.96	0.16	G	9.14E-03	0.90	0.83-0.97	0.19	G	1.69E-01	0.86	0.69-1.07	0.18
1	63.18	rs4409689	*DOCK7, ATG4C*	intergenic	N	A	4.34E-04	1.11	1.05-1.17	G	5.73E-02	0.88	0.76-1.00	0.33	G	1.84E-02	0.93	0.87-0.99	0.34	G	8.82E-03	0.79	0.66-0.94	0.33
5	66.35	rs468938	*MAST4*	intronic	Y	A	4.36E-04	0.88	0.81-0.94	G	1.46E-02	1.23	1.04-1.45	0.17	G	3.60E-03	1.13	1.04-1.23	0.14	G	8.36E-01	1.02	0.82-1.28	0.15
1	92.15	rs913059	*TGFBR3*	intronic	N	A	4.69E-04	1.11	1.05-1.18	C	4.20E-01	0.94	0.82-1.09	0.25	C	3.66E-04	0.88	0.82-0.95	0.26	C	4.98E-01	0.94	0.78-1.13	0.27
7	3.51	rs17133379	*SDK1*	intronic	Y	A	4.83E-04	0.87	0.81-0.94	A	3.19E-02	0.82	0.68-0.98	0.15	A	7.13E-03	0.89	0.81-0.97	0.14	A	2.31E-01	0.87	0.68-1.10	0.14
6	82.84	rs13211186	*LINC01526, IBTK*	intergenic	Y	A	4.94E-04	1.1	1.04-1.16	A	3.74E-03	1.21	1.06-1.38	0.35	A	1.44E-02	1.08	1.02-1.15	0.38	A	4.08E-01	1.07	0.91-1.27	0.38

Meta-analysis results shown alongside individual cohorts CSMC, IIBDGC, and SHARE.
CHR: chromosome; Position: GRCh37; Allele: effect allele; OR: Odds Ratio; CI: Confidence Interval; MAF: Minor Allele Frequency;
SNPs in the same genomic region are r^2<0.5; Variants annotated as potentially novel fell outside of 1 MB of previously reported variants in an IBD-associated genomic region [9 10].

Supplementary Table 2. Top genetic associations with pCD

						Meta-Analysis				CSMC					IIBDGC					SHARE				
CHR	Position (MB)	SNP	Gene	Location	Novel (Y/N)	Allele	*P*-value	OR	95% CI	Allele	*P*-value	OR	95% CI	MAF	Allele	*P*-value	OR	95% CI	MAF	Allele	*P*-value	OR	95% CI	MAF
6	32.41	rs17496549	*HLA-DRA*	intronic	N	A	4.68E-07	1.24	1.14-1.35	A	5.64E-05	1.54	1.25-1.89	0.09	A	8.15E-04	1.18	1.07-1.30	0.11	A	4.09E-02	1.29	1.01-1.65	0.11
6	31.92	rs4151651	*CFB*	exonic	N	A	9.35E-06	1.32	1.17-1.49	A	2.59E-04	1.75	1.30-2.37	0.04	A	5.90E-03	1.22	1.06-1.41	0.04	A	4.13E-02	1.39	1.01-1.91	0.05
6	32.37	rs2294884	*BTNL2*	intronic	N	A	1.79E-05	0.86	0.80-0.92	C	6.42E-03	1.28	1.07-1.52	0.15	C	2.62E-03	1.13	1.04-1.22	0.17	C	2.42E-02	1.27	1.03-1.55	0.17
4	47.28	rs13127214	*GABRB1*	intronic	Y	C	5.04E-05	1.2	1.10-1.31	G	2.15E-01	0.88	0.71-1.08	0.10	G	6.84E-05	0.82	0.74-0.90	0.11	G	4.48E-01	0.90	0.69-1.18	0.10
1	70.23	rs17325887	*LRRC7*	intronic	Y	A	8.57E-05	0.73	0.63-0.86	G	7.32E-02	1.42	0.97-2.08	0.02	G	3.65E-03	1.30	1.09-1.56	0.03	G	1.39E-02	1.73	1.12-2.69	0.03
8	126.53	rs6470362	*TRIB1*	intergenic	N	A	9.55E-05	0.89	0.84-0.94	G	1.40E-01	1.11	0.97-1.27	0.34	G	9.43E-04	1.12	1.05-1.20	0.26	G	7.09E-02	1.18	0.99-1.42	0.29
19	55.22	rs1325156	*LILRB4, LILRP2*	intergenic	N	A	1.16E-04	1.15	1.07-1.23	A	6.23E-01	1.04	0.88-1.24	0.16	A	1.58E-04	1.17	1.08-1.27	0.16	A	9.44E-02	1.20	0.97-1.49	0.15
11	71.14	rs1790353	*DHCR7*	intergenic	Y	A	1.44E-04	0.88	0.83-0.94	A	2.89E-01	0.92	0.79-1.07	0.23	A	5.90E-04	0.88	0.81-0.94	0.22	A	8.70E-02	0.84	0.69-1.03	0.22
21	46.08	rs7275203	*TSPEAR*	intronic	N	A	1.66E-04	0.89	0.84-0.95	G	2.13E-01	1.09	0.95-1.26	0.25	G	7.56E-04	1.13	1.05-1.21	0.29	G	1.37E-01	1.15	0.96-1.37	0.27
4	103.62	rs227375	*MANBA*	intronic	N	A	1.91E-04	0.9	0.86-0.95	A	4.71E-02	0.88	0.78-1.00	0.48	A	7.79E-03	0.92	0.87-0.98	0.48	A	1.97E-02	0.82	0.70-0.97	0.50
1	98.83	rs4542263	*LOC729987 SNX7*	intergenic	Y	A	2.50E-04	1.19	1.08-1.31	C	4.21E-02	0.80	0.64-0.99	0.10	C	8.84E-03	0.87	0.78-0.96	0.09	C	2.98E-02	0.71	0.53-0.97	0.09
4	123.26	rs6851362	*KIAA1109*	intronic	N	A	2.58E-04	0.89	0.84-0.95	G	9.64E-02	1.13	0.98-1.31	0.24	G	1.11E-03	1.12	1.05-1.20	0.25	G	4.01E-01	1.08	0.90-1.31	0.24
12	123.65	rs34484751	*MPHOSPH9*	intronic	Y	A	2.67E-04	0.76	0.65-0.88	C	3.85E-01	1.20	0.8-1.79	0.02	C	4.18E-04	1.35	1.14-1.59	0.03	C	2.94E-01	1.29	0.8-2.07	0.03
4	123.15	rs6848868	*KIAA1109*	exonic	N	A	2.74E-04	0.85	0.78-0.93	G	4.31E-02	1.22	1.01-1.48	0.10	G	2.78E-03	1.17	1.05-1.29	0.09	G	3.26E-01	1.15	0.87-1.53	0.09
21	44.5	rs2850146	*CBS, U2AF1*	intergenic	Y	C	2.94E-04	0.84	0.77-0.92	C	1.39E-02	0.75	0.59-0.94	0.09	C	2.51E-03	0.85	0.77-0.94	0.09	C	7.16E-01	0.95	0.71-1.26	0.09
11	113.42	rs10736469	*DRD2, TMPRSS5*	intergenic	Y	A	3.04E-04	0.91	0.86-0.96	G	1.32E-02	1.17	1.03-1.33	0.41	G	3.70E-03	1.09	1.03-1.16	0.39	G	5.80E-01	1.05	0.89-1.23	0.41
6	32.38	rs3763313	*BTNL2, HLA-DRA*	intergenic	N	A	3.06E-04	0.89	0.83-0.95	C	6.37E-02	1.15	0.99-1.33	0.21	C	2.67E-03	1.12	1.04-1.20	0.20	C	2.59E-01	1.12	0.92-1.36	0.20
4	123.52	rs975403	*IL2,IL21*	intergenic	N	A	3.39E-04	1.10	1.05-1.16	A	1.14E-01	1.11	0.98-1.26	0.40	A	1.16E-03	1.11	1.04-1.17	0.41	A	4.48E-01	1.07	0.9-1.26	0.39
2	241.58	rs34228697	*GPR35, AQP12B*	intergenic	N	A	3.88E-04	1.14	1.06-1.22	G	1.62E-02	0.80	0.67-0.96	0.16	G	9.14E-03	0.90	0.83-0.97	0.19	G	1.69E-01	0.86	0.69-1.07	0.18
1	63.18	rs4409689	*DOCK7, ATG4C*	intergenic	N	A	4.34E-04	1.11	1.05-1.17	G	5.73E-02	0.88	0.76-1.00	0.33	G	1.84E-02	0.93	0.87-0.99	0.34	G	8.82E-03	0.79	0.66-0.94	0.33
5	66.35	rs468938	*MAST4*	intronic	Y	A	4.36E-04	0.88	0.81-0.94	G	1.46E-02	1.23	1.04-1.45	0.17	G	3.60E-03	1.13	1.04-1.23	0.14	G	8.36E-01	1.02	0.82-1.28	0.15
1	92.15	rs913059	*TGFBR3*	intronic	N	A	4.69E-04	1.11	1.05-1.18	C	4.20E-01	0.94	0.82-1.09	0.25	C	3.66E-04	0.88	0.82-0.95	0.26	C	4.98E-01	0.94	0.78-1.13	0.27
7	3.51	rs17133379	*SDK1*	intronic	Y	A	4.83E-04	0.87	0.81-0.94	A	3.19E-02	0.82	0.68-0.98	0.15	A	7.13E-03	0.89	0.81-0.97	0.14	A	2.31E-01	0.87	0.68-1.10	0.14
6	82.84	rs13211186	*LINC01526, IBTK*	intergenic	Y	A	4.94E-04	1.1	1.04-1.16	A	3.74E-03	1.21	1.06-1.38	0.35	A	1.44E-02	1.08	1.02-1.15	0.38	A	4.08E-01	1.07	0.91-1.27	0.38

Meta-analysis results shown alongside individual cohorts CSMC, IIBDGC, and SHARE.
CHR: chromosome; Position: GRCh37; Allele: effect allele; OR: Odds Ratio; CI: Confidence Interval; MAF: Minor Allele Frequency;
SNPs in the same genomic region are r^2<0.5; Variants annotated as potentially novel fell outside of 1 MB of previously reported variants in an IBD-associated genomic region [9 10].

Supplemental Methods

Study Population

The Cedars-Sinai Medical Center (CSMC) IBD Research Repository (MIRIAD) contains clinical, serologic, and genetic data on IBD patients followed at CSMC as described[1]. Details on the International IBD Genetics Consortium (IIBDGC) and the Sinai-Helmsley Alliance for Research Excellence (SHARE) consortium studies have been previously reported[2-4].

Serological Phenotyping

Following study consent and enrollment of CSMC subjects, sera were collected at first blood collection appointment and subsequently analyzed for expression of anti-Saccharomyces cerevisiae antibodies (ASCA IgG and IgA), perinuclear antinuclear cytoplasmic antibody (ANCA), anti-flagellin (CBir-1), anti-outer membrane porin C (OmpC), and anti-Pseudomonas fluorescence-associated sequence I2 (I2) as described[5]. Antibody levels were measured, and results expressed as ELISA units (EU/ml) relative to a Cedars-Sinai or Prometheus Laboratory standards (San Diego, CA) derived from a pool of patient sera with well-characterized CD found to have reactivity to these antigens. We have previously shown these IBD-related serologies to be stable over time and independent of disease activity[5].

Genotyping and Quality Control (QC)

Genotyping using the Illumina Immunochip array was performed at CSMC from DNA isolated from EBV-transformed lymphoblastoid cell lines and using in-house clustering algorithm and QC pipeline, or as previously described for IIBDGC and SHARE[3 4]. Average genotyping call rates for CSMC, IIBDGC and SHARE samples that passed QC measures were 99.9%, 99.8%, and 99.5%,

respectively. Inter- and intra-cohort duplicate and related samples (PI_HAT > 0.1875) were identified, and the sample with greater proportion of missing data within each pair was removed. Samples with gender discrepancies, >3% missing data and those without a clear and definitive CD phenotype were removed. SNPs with >2% missing data, minor allele frequency (MAF) < 1%, or failing Hardy-Weinberg equilibrium in controls ($p < 1 \times 10^{-5}$) were excluded. Post-QC sample and SNP numbers are shown in **Supplementary Table 1.** We performed supervised Admixture analysis to determine the degree of European ancestry using reference samples from the Human Genome Diversity Project to identify CSMC subjects ≥ 50% European ancestry [6-8]. Subjects with ≤ 50% European ancestry in the CSMC cohort were excluded. Principal Components Analysis (PCA), a dimensionality reduction technique commonly used to infer population structure and variation in genetic data, was used to confirm European ancestry in IIBDGC and SHARE[9]. We utilized the top principal components (PC) as covariates to correct for population sub-structure. Genotype clustering was manually reviewed where possible to ensure accurate allele calling.

Statistical Analyses

Clinical, demographic, and serological parameters

Serological data included both categorical (seropositivity) or continuous variables (antibody level). Quartile Sum Score (QSS) for anti-I2, anti-OmpC, anti CBir-1 and IgA and IgG ASCA was calculated as previously described [10]. In brief, for each antigen, patients whose antibody levels were in the 1st, 2nd, 3rd, and 4th quartile of the distribution were given a quartile score of 1, 2, 3, or 4, respectively. By adding individual quartile sum scores for each microbial antigen, a quartile sum score (range from 5-20) was generated to represent the cumulative quantitative immune response toward all 5 antigens for each patient.

Genetic analysis

Uncontrolled confounding from population sub-structure was evaluated by calculating the genomic control inflation factor using R [11]. Genomic inflation estimates the amount of inflation by comparing observed test statistics across all genetic variants to those expected under the null hypothesis[9]. A genomic inflation factor (λ) close to 1 reflects no evidence of genomic inflation. Minimal genomic inflation was observed for genetic association study: $\lambda_{GC\text{-}meta}$= 1.008; $\lambda_{GC\text{-}CSMC}$=1.006; $\lambda_{GC\text{-}IIBDGC}$= 1.06; and $\lambda_{GC\text{-}SHARE}$= 1.017.

Variants were annotated as "known" if they fell within one megabase of previously reported variant in an IBD-associated genomic region [12]. Linkage disequilibrium (r^2) for European population was evaluated using LDlink [13]; SNPs displayed in **Table 3** in the same genomic region are r^2<0.5. RegulomeDBv2.0 [14], which integrates data from ENCODE and Roadmap, was utilized to investigate DNA features and regulatory elements overlapping associated SNPs of interest and expression quantitative trait loci (eQTL) were examined using data available from the Genotype-Tissue Expression project (GTEx) [15].

Mutagenesis of human CFB

A mammalian expression plasmid containing His-tagged human CFB was purchased from Sino Biological, Inc (HG17064-NH). The S252 mutation was introduced using the QuickChange site-directed mutagenesis kit (Agilent) according to the manufacturer's instructions using the following primers: S252-forward: 5'-GTTCCCCTGGGCTGTGCCCATCCTC-3' and S252-reverse: 5'-

GAGGATGGGCACAGCCCAGGGGAAC-3'. Successful mutagenesis was confirmed by sequencing.

Cell lines, transfection, and purification of recombinant G252 and S252 CFB

HEK293 cells were maintained in DMEM supplemented with 10% FBS, penicillin/streptomycin, and fungizone. G252 and S252 CFB plasmid DNA were transfected into HEK293 cells using polyethyleneimine-branched (PEI) transfection reagent (Sigma Aldrich) according to manufacturer's instructions. In brief, HEK293 cells were transfected with 500 ng of plasmid DNA mixed with 1 mg/ml PEI and cells were incubated for 48 h. His-tagged CFB proteins were purified from cell culture supernatants by affinity chromatography using HisPur Ni-NTA resin (ThermoFisher). G252 and S252 CFB proteins were dialyzed against PBS and purity of the proteins were confirmed by Coomassie stain and Western blot using anti-human complement factor B antibody (Complement Tech). Bands were visualized on an Odyssey imaging system (LI-COR). Protein concentrations were determined using BCA protein assay (Pierce).

CFB cleavage assay

G252 and S252 CFB (0.5 μg/ml) were incubated with complement factor D (R&D Systems, 0.2 μg/ml), and C3b (Sigma Aldrich, 2 μg/ml) in PBS supplemented with 10 mM $MgCl_2$, and 25 mM NaCl for 30 minutes at 37 °C [16 17]. Cleavage of CFB was analyzed by Western blot using goat anti-human complement factor B antibody (Complement Tech, 178 μg/ml) and secondary IRDye 800CW donkey anti-goat antibody (LI-COR, 2 μg/ml). Bands were visualized on an Odyssey imaging system (LI-COR).

CFB binding assay

A cell-free C3b binding assay was utilized [17]. 96-well plates were coated with recombinant C3b (5 μg/ml) in PBS overnight at 4 ºC. Plates were blocked with PBS supplemented with 25 mM NaCl, 4 % BSA, and 0.05 % Tween 20 for 1 h at 37 °C. G252 and S252 CFB (0.2 μg/ml), properdin (1 μg/ml), and factor D (0.025 μg/ml) were incubated in binding buffer (11 mM Na_2HPO_4, 1.8 mM NaH_2PO_4, 75 mM NaCl, 10 mM $MgCl_2$, 4% BSA, 0.05% Tween 20) for 2 h at 37 ºC. Reactions containing EDTA (20 mM) served as no-binding controls. Plates were washed and incubated with goat anti-human factor B antibody (Complement Tech) for 1 h at 37 ºC, followed by HRP-conjugated anti-goat antibody (Abcam) for 1 h at 37 ºC. Plates were developed using TMB substrate and read at 450 nm using the Emax microplate reader (Molecular Devices). In separate experiments, 96-well plates were coated with recombinant G252 or S252 CFB (5 μg/ml) in PBS overnight at 4 ºC. C3b (0.2 μg/ml), properdin (1 μg/ml), and factor D (0.025 μg/ml) were incubated in binding buffer for 2 h at 37 ºC. Plates were incubated with FITC-conjugated anti-human C3 antibody (Cedarlane) for 1 h at 37 ºC, and read at 492 nm using the CLARIOstar Plus microplate reader (BMG Labtech).

Glycoprotein staining of G252 and S252 CFB

Recombinant G252 or S252 CFB were separated by SDS-PAGE and stained for protein glycosylation using a Glycoprotein Staining Kit (Thermo Scientific) according to the manufacturer's protocol. Horseradish peroxidase and soybean trypsin inhibitor were used as positive and negative control, respectively. Protein G plus agarose (Santa Cruz Biotechnology) was pre-incubated with anti-human factor B antibody (Complement Tech, 2 μg) for 1 h at 4 ºC. Next, beads were incubated with 40 μg of pooled CD patient serum (genotypes: S252/S252,

G252/S252, G252/G252) overnight at 4 ºC. Samples were separated by SDS-PAGE and staining as described above.

Purification of CFB from human serum

CFB was purified from human serum by immunoprecipitation. In brief, 200 µl human serum was diluted in 800 µl PBS containing protease inhibitors (Pierce) and pre-cleared using protein G plus/protein A agarose beads (EMD Millipore) and rabbit IgG isotype control antibody (Abcam, 2 µg) at 4 ºC for 30 minutes. Precleared serum was incubated with anti-human CFB antibody (Thermo Scientific, 2 µg) at 4 ºC for 1 h. Next, 40 µl protein G plus/protein A agarose beads were added and incubated at 4 ºC for 1 h. Beads were pelleted by centrifugation, washed five times with washing buffer (1% NP40, 1mM EDTA pH 8, 50 mM HEPES, 250 mM NaCl, and protease inhibitors), and once with PBS containing protease inhibitors. After washing, bound CFB was eluted with 100 µl elution buffer (0.2 M glycine-HCL, 2.5 pH) and immediately neutralized by adding 10 µl neutralization buffer (1 M Tris-HCL, 8.5 pH). Eluted CFB was dialyzed against PBS.

Isolation of peripheral blood monocytes and generation of human macrophages

Peripheral blood mononuclear cells (PBMCs) were isolated from whole blood of healthy donors using Lymphocyte Separation Medium (Corning). Next, monocytes were enriched from PBMCs by plating PBMCs in tissue culture dishes for 2 h. Adherent cells were washed with media and cultured in DMEM supplemented with 10% fetal bovine serum (FBS), penicillin/streptomycin, fungizone, and human granulocyte-macrophage colony-stimulating factor (GM-CSF, 50 ng/ml) (BioLegend). On Day 4, media was replaced with fresh DMEM containing GM-CSF. On day 7, macrophages were stimulated with 25 ng/ml human IFN-γ (Peprotech) and 10 ng/ml LPS (Sigma).

Macrophages were used on day 9. The purity of macrophages was determined by flow cytometry using anti-human CD68 antibody (BioLegend) and was routinely greater than 78%.

Phagocytosis assay

Tetramethylrhodamine isothiocyanate (TRITC)-conjugated zymosan particles (50 μg/ml) [18] were incubated in presence of 10% serum of CD patients or non-IBD controls for 1 hour at 37°C. In separate experiments 10 % serum was replaced with recombinant G252 or S252 CFB (final concentration: 100 – 500 ng/ml), CFD (200 ng/ml), and C3b (2 μg/ml) and incubated in PBS plus 10 mM $MgCl_2$ for 1 hour at 37°C. Reaction mixtures were added to macrophages and briefly centrifuged to ensure that particles were in contact with macrophages. Macrophages incubated without TRITC-zymosan served as negative controls. Cells were incubated for 15 - 60 minutes at room temperature, washed with PBS, resuspended in PBS containing 2 mM EDTA, and 2.5 U/ml proteinase K (Sigma) and incubated at room temperature for 15 minutes to remove all non-internalized zymosan particles followed by fixation with 1% paraformaldehyde. Cells were acquired by flow cytometry using an LSR II analyzer (BD Biosciences). Macrophages were gated by forward/side scatter (FSC/SSC) and the percentage of TRITC-positive macrophages was analyzed. Sera from each CD patient or healthy subject was analyzed in triplicates and blinded to the genotypes. For CFB neutralization experiments, macrophages were incubated with anti-human CFB (Complement Tech), or normal goat serum at 37°C for 30 minutes prior to addition of zymosan.

Cytokine secretion of zymosan stimulated human macrophages

PBMC-derived macrophages from healthy donors were incubated with recombinant 250 ng/ml G252 or S252 CFB (250 ng/ml), CFD (0.2 µg/ml), C3b (2 µg/ml) in the presence of zymosan (1 ng/ml) for 6 h. Concentrations of human TNF-a or IL-6 in supernatants were determined by ELISA (eBiosciences).

MTT Cell proliferation

Macrophages were stimulated for 6 or 24 h as described above. Cell proliferation was determined using the MTT Cell proliferation assay kit (BiVision) according to the manufacturer's protocol.

Molecular modelling of S252 CFB

To understand the significance of glycosylated CFB at the linker location (AA 252), we used molecular modelling. Briefly, the crystal structure of CFB (PDB code: 2OK5) was used as a template for generating the glycosylated CFB. The sugar structure, DGlcpNAcb1-4DGlcpAb1-3DGlcpNAcb1-4DGlcpAb1-3DGlcpNAcb1-4DGlcpAb1-OH was generated using the web server Glycosylator [19]. Then to understand whether glycosylation at the linker region affect CFB complex with C3, we used the crystal structure of CFB-C3b complex (PDB code: 2XWJ) as a template. The glycosylated CFB was superimposed to the CFB chain in the crystal structure. Figures are generated using Pymol (Schrodinger, Inc).

Statistical analysis for functional assays

Samples were analyzed in duplicates or triplicates and experiments were repeated at least three times. Data are presented as means ± SD. Differences between groups were calculated using one-way ANOVA with Bonferroni correction or Student's *t*-test. $p < 0.05$ was considered significant.

References

1. Kaur M, Panikkath D, Yan X, et al. Perianal Crohn's Disease is Associated with Distal Colonic Disease, Stricturing Disease Behavior, IBD-Associated Serologies and Genetic Variation in the JAK-STAT Pathway. *Inflamm Bowel Dis* 2016;22(4):862-9.
2. Cleynen I, Boucher G, Jostins L, et al. Inherited determinants of Crohn's disease and ulcerative colitis phenotypes: a genetic association study. *Lancet* 2016;387(10014):156-67.
3. Jostins L, Ripke S, Weersma RK, et al. Host-microbe interactions have shaped the genetic architecture of inflammatory bowel disease. *Nature* 2012;491(7422):119-24.
4. Peloquin JM, Goel G, Kong L, et al. Characterization of candidate genes in inflammatory bowel disease-associated risk loci. *JCI Insight* 2016;1(13):e87899.
5. Landers CJ, Cohavy O, Misra R, et al. Selected loss of tolerance evidenced by Crohn's disease-associated immune responses to auto- and microbial antigens. *Gastroenterology* 2002;123(3):689-99.
6. Cann HM, de Toma C, Cazes L, et al. A human genome diversity cell line panel. *Science* 2002;296(5566):261-2.
7. Alexander DH, Novembre J, Lange K. Fast model-based estimation of ancestry in unrelated individuals. *Genome Res* 2009;19(9):1655-64.
8. Stoneking M, Krause J. Learning about human population history from ancient and modern genomes. *Nat Rev Genet* 2011;12(9):603-14.
9. Price AL, Zaitlen NA, Reich D, et al. New approaches to population stratification in genome-wide association studies. *Nat Rev Genet* 2010;11(7):459-63.
10. Mow WS, Landers CJ, Steinhart AH, et al. High-level serum antibodies to bacterial antigens are associated with antibiotic-induced clinical remission in Crohn's disease: a pilot study. *Dig Dis Sci* 2004;49(7-8):1280-6.
11. Team RDC. R: A Language and Environment for Statistical Computing [program]. Vienna, Austria: R Foundation for Statistical Computing, 2020.
12. de Lange KM, Moutsianas L, Lee JC, et al. Genome-wide association study implicates immune activation of multiple integrin genes in inflammatory bowel disease. *Nat Genet* 2017;49(2):256-61.
13. Machiela MJ, Chanock SJ. LDlink: a web-based application for exploring population-specific haplotype structure and linking correlated alleles of possible functional variants. *Bioinformatics* 2015;31(21):3555-7.
14. Boyle AP, Hong EL, Hariharan M, et al. Annotation of functional variation in personal genomes using RegulomeDB. *Genome Res* 2012;22(9):1790-7.
15. Consortium GT. The Genotype-Tissue Expression (GTEx) project. *Nat Genet* 2013;45(6):580-5.
16. Hourcade DE, Mitchell LM, Oglesby TJ. A conserved element in the serine protease domain of complement factor B. *J Biol Chem* 1998;273(40):25996-6000.
17. Hourcade DE, Mitchell LM, Oglesby TJ. Mutations of the type A domain of complement factor B that promote high-affinity C3b-binding. *J Immunol* 1999;162(5):2906-11.
18. Underhill DM, Ozinsky A, Hajjar AM, et al. The Toll-like receptor 2 is recruited to macrophage phagosomes and discriminates between pathogens. *Nature* 1999;401(6755):811-5.
19. Lemmin T, Soto C. Glycosylator: a Python framework for the rapid modeling of glycans. *BMC Bioinformatics* 2019;20(1):513.

Supplemental Information

Members of the International Inflammatory Bowel Disease Genetics Consortium

Clara Abraham, Jean-Paul Achkar, Tariq Ahmad, Leila Amininejad, Ashwin N Ananthakrishnan, Vibeke Andersen, Carl A Anderson, Jane M Andrews, Vito Annese, Guy Aumais, Leonard Baidoo, Robert N Baldassano, Peter A Bampton, Murray Barclay, Jeffrey C Barrett, Theodore M Bayless, Johannes Bethge, Alain Bitton, Gabrielle Boucher, Stephan Brand, Berenice Brandt, Steven R Brant, Carsten Büning, Angela Chew, Judy H Cho, Isabelle Cleynen, Ariella Cohain, Anthony Croft, Mark J Daly, Mauro D'Amato, Silvio Danese, Dirk De Jong, Martine De Vos, Goda Denapiene, Lee A Denson, Kathy L Devane, Olivier Dewit, Renata D'Inca, Marla Dubinsky, Richard H Duerr, Cathryn Edwards, David Ellinghaus, Jonah Essers, Lynnette R Ferguson, Eleonora A Festen, Philip Fleshner, Tim Florin, Denis Franchimont, Andre Franke, Karin Fransen, Richard Gearry, Michel Georges, Christian Gieger, Jürgen Glas, Philippe Goyette, Todd Green, Anne M Griffiths, Stephen L Guthery, Hakon Hakonarson, Jonas Halfvarson, Katherine Hanigan, Talin Haritunians, Ailsa Hart, Chris Hawkey, Nicholas K Hayward, Matija Hedl, Paul Henderson, Xinli Hu, Hailiang Huang, Jean-Pierre Hugot, Ken Y Hui, Marcin Imielinski, Andrew Ippoliti, Laimas Jonaitis, Luke Jostins, Tom H Karlsen, Nicholas A Kennedy, Mohammed Azam Khan, Gediminas Kiudelis, Krupa Krishnaprasad, Subra Kugathasan, Limas Kupcinskas, Anna Latiano, Debby Laukens, Ian C Lawrance, James C Lee, Charlie W Lees, Marcis Leja, Johan Van Limbergen, Paolo Lionetti, Jimmy Z Liu, Edouard Louis, Gillian Mahy, John Mansfield, Dunecan Massey, Christopher G Mathew, Dermot PB McGovern, Raquel Milgrom, Mitja Mitrovic, Grant W Montgomery, Craig Mowat, William Newman, Aylwin Ng , Siew C Ng, Sok Meng Evelyn Ng, Susanna Nikolaus, Kaida Ning, Markus Nöthen, Ioannis Oikonomou, Orazio Palmieri, Miles

Parkes, Anne Phillips, Cyriel Y Ponsioen, Urõs Potocnik, Natalie J Prescott, Deborah D Proctor, Graham Radford-Smith, Jean-Francois Rahier, Soumya Raychaudhuri, Miguel Regueiro, Florian Rieder, John D Rioux, Stephan Ripke, Rebecca Roberts, Richard K Russell, Jeremy D Sanderson, Miquel Sans, Jack Satsangi, Eric E Schadt, Stefan Schreiber, Dominik Schulte, L Philip Schumm, Regan Scott, Mark Seielstad, Mark S Silverberg, Lisa A Simms, Jurgita Skieceviciene, Sarah L Spain, A. Hillary Steinhart, Joanne M Stempak, Laura Stronati, Jurgita Sventoraityte, Stephan R Targan, Kirstin M Taylor, Anje ter Velde, Emilie Theatre, Leif Torkvist, Mark Tremelling, Andrea van der Meulen, Suzanne van Sommeren, Eric Vasiliauskas, Severine Vermeire, Hein W Verspaget, Thomas Walters, Kai Wang , Ming-Hsi Wang, Rinse K Weersma, Zhi Wei, David Whiteman, Cisca Wijmenga, David C Wilson, Juliane Winkelmann, Ramnik J Xavier, Bin Zhang, Clarence K Zhang, Hu Zhang, Wei Zhang, Hongyu Zhao, Zhen Z Zhao