Supplementary Information

Deep resequencing of GWAS loci identifies independent rare variants associated with inflammatory bowel disease

Manuel A. Rivas, Melissa Beaudoin, Agnes Gardet, Christine Stevens, Yashoda Sharma, Clarence K. Zhang, Gabrielle Boucher, Stephan Ripke, Noel Burtt, Tim Fennell, Andrew Kirby, Anna Latiano, Philippe Goyette, Todd Green, Jonas Halfvarson, Talin Haritunians, Joshua M. Korn, Finny Kuruvilla, Caroline Lagacé, Benjamin Neale, Ken Sin Lo, Phil Schumm, Mark Silverberg, Leif Törkvist, David Ellinghaus, NIDDK IBD Genetics Consortium[¶], International Inflammatory Bowel Disease Genetics Consortium[¶], Marla Dubinsky, Steven R. Brant, Richard H. Duerr, David Altshuler, Stacey Gabriel, Guillaume Lettre, Andre Franke, Mauro D'Amato, Dermot P.B. McGovern, Judy H. Cho, John D. Rioux, Ramnik J. Xavier, Mark J. Daly

Table of Contents

Supplementary Note

Section S1. Study design

In order to test the role of rare coding variants in CD GWAS regions to IBD we sequenced 350 patients with CD and 350 healthy controls of European Ancestry in a pooled fashion for a total of 107.5 kb of target sequence. We identified all nonsynonymous, nonsense, or splice site variants which occurred in 2 or more copies up to a frequency of 5% - a total of 115 variants. Excluding known GWAS associated low-frequency coding variants at NOD2, IL23R, and LRRK2/MUC19 follow-up genotyping was performed for 70 of these markers in nine independent case-control sample collections totaling 16054 CD disease patients, 12153 UC disease patients, and 17575 healthy controls. Given these numbers, we estimated the power to detect association at varying odds ratios and minor allele frequencies with an estimated prevalence of 4/1000 consistent with the additive genetic variance estimates of Franke et al. (2010). We report that of the 70 markers successfully genotyped 22%,60%,79%,88%,91% have at least 80% power to detect association at minor allele frequency odds ratios of 1.5,2,3,4, and 5 respectively and the required number of samples to achieve at least 80% power to detect nominal and experiment-wide association is met by our sample collection (Figure 1b,S3a,S3b), implying that we are well positioned to address the contribution of rare and low frequency polymorphisms in GWAS loci to IBD.

Section S2. DNA preparation and pooling

Crohn's disease patients and Controls from NIDDK consortium were selected with priority given to samples with adequate amounts of DNA and those with GWAS data available. Samples from the NIDDK consortium undergo rigorous clinical phenotyping and control matching for genetic studies. DNA purification methods are also performed on these samples. The case/control samples selected have already been stringently matched in previous GWAS studies. The

baseline concentration of genomic DNA was quantified by Quant-iT™ PicoGreen® dsDNA reagent and detected on the Thermo Scientific Varioskan Flash. All DNAs were normalized to 20ng/µl and repeat quantification was performed to assess accuracy of the normalization step. The quantification and normalization was repeated again to ensure that all samples fell within the desired concentration range. A 10% variance was allowed, as that is the limit of quantitation of PicoGreen® detection system. The normalization steps were done with robotic automation using the Packard Multiprobe II HT EX. Equimolar amounts of each DNA in a pool of samples is essential thus the same robotic automation was used to guarantee a uniform pipetting error across all samples in all steps. Once each individual sample is normalized to 10ng/ul, groups of 50 individuals were pooled together using a Multiprobe or Packard Robotic to total 14 pools (700 samples).

Section S3. Target selection and design

Candidate exonic targets from top GWAS published, confirmed genes along with a sample of other highly significant regions of interest were uploaded against HG17 freeze to an in-house database, which houses PRIMER3 software. Amplicons encompassing each target region (coding exons only) were designed using Illumina parameters including a minimum amplicon length of 150bp and maximum amplicon length of 600bp with no buffer sequence added. Additionally, Not1 tails were added to the primer pairs to provide a recognition site for downstream concatenation and shearing step. The primer design process was reiterative and included a validation step on 3 HapMap CEU samples to ensure oligo synthesis quality. Amplicons were validated by running PCR product on agarose gels to assess clarity of single bands. Amplicons that had 2/3 clear bands were considered validated. Pfu enzyme, used in Illumina sequencing protocol for PCR, was used in the characterization process. In total, 593 primer pairs passed and covered 95% of the 108 kb target. PCRs contained 20 ng of

pooled genomic DNA, 1x HotStar buffer, 0.8 mM dNTPs, 2.5 mM MgCl2, 0.2 units of HotStar Enzyme (Qiagen), and 0.25 μM forward and reverse primers in a 6- or 10-μl reaction volume. PCR cycling parameters were: one cycle of 95°C for 15 min; 35 cycles of 95°C for 20 s, 60°C for 30 s, and 72°C for 1 min; followed by one cycle of 72°C for 3 min. Each PCR product was then treated to similar steps used for the pooling of DNA individuals. The quantification, normalization, and pooling process was again required to ensure that equimolar PCR product went into library construction to have equal representation of all targets. PCR yield was assessed by the same quantification system and the lowest product yield was then used to normalize across PCR plates. Secondary confirmation was ascertained by testing one column of PCR product per plate on 2% agarose E-gel against 1kb DNA ladder to visualize PCR product size. The 593 PCR products were then combined, using the Packard Multiprobe II HT EX, resulting in an amplified target product per sample pool for sequencing.

Section S4. Discovery via Sequencing

The PCR products for each pooled sample were concatenated using *NotI* adapters and sheared into fragments. Libraries were constructed by a modified Illumina single-end library protocol, with 225-275 bp gel size selection and PCR enrichment using 14 cycles of PCR, and then single-end sequenced with 76 cycles on an Illumina Genome Analyzer. Each sample pool was sequenced using a single lane of a Illumina GAII analyzer flowcell. 76bp, 36bp and 52bp reads were aligned to the genome using MAQ algorithm within the Picard analysis pipeline, and further processed using the SAMtools software and custom scripts.

Rare variant discovery with Syzygy

In this study we have developed software called Syzygy to analyze pooled sequencing data. Syzygy is a sequencing analysis tool that allows: 1) rare variant SNP detection, 2) allele frequency estimation, 3) single marker association testing, 4) group-wise marker test association, 5) pooled experiment statistical summary and power evaluation, and 6) annotation of rare variants discovered in the pooled sequencing experiment from primary sequencing data in BAM/SAM format. Algorithm details are available in Syzygy.pdf on the online software site: http://www.broadinstitute.org/software/syzygy

Section S5. Validation

137 high confidence Single Nucleotide Variants (SNVs) were assayed in two phases of genotyping using Sequenom MassARRAY iPLEX GOLD chemistry50. The 137 high confidence SNVs were randomly drawn and chosen based on functional annotation (nonsense, missense, splice). We did not take into account allele counts or allele frequency since we were interested in singleton detection as well. The first phase consisted of 72 SNVs and the second phase of 65 SNVs on 350 NIDDK Crohns samples and 350 NIDDK controls for validation purposes.

In each phase of Sequenom genotyping, oligos were synthesized and mass-spec QCed at Integrated DNA Techologies. All SNVs were genotyped in multiplexed pools of 25–36 assays, designed by AssayDesigner v.3.1 software, starting with 10 ng of DNA per pool. Around 7 nl of reaction was loaded onto each position of a 384-well SpectroCHIP preloaded with 7 nl of matrix (3-hydroxypicolinic acid). SpectroCHIPs were analyzed in automated mode by a MassArray MALDI-TOF Compact system with a solid phase laser mass spectrometer (Bruker Daltonics Inc.). We obtained high quality data (>95% genotype call rate, HWE P> 0.001) in all samples that had at least one SNV.

Variants were called by real-time SpectroCaller algorithm, analyzed by SpectroTyper v.4.0 software and manually reviewed for rare variants.

In Phase 1, 72 SNVs passed QC (HWE and genotype call rate). Of the 72 SNVs that passed QC 68 were identified as True Positives (TP) and 4 were monomorphic (FP). In Phase 2, 65 SNVs passed QC. Of the 65 SNVs 57 were true positives and 8 monomorphic. The validation data suggests a high true positive rate of approximately 91.2%.

Section S6. Follow-up Genotyping

We identified all nonsynonymous, nonsense, or splice site variants which occurred in 2 or more copies up to a frequency of 5% - a total of 115 variants. Excluding known GWAS associated low-frequency coding variants at NOD2, IL23R, and LRRK2/MUC19 follow-up genotyping was performed for 70 of these markers.

Follow-up genotyping was successfully performed for 70 of these markers in nine independent case control samples totaling 16054 CD patients, 12153 UC patients, 17575: 1) samples from the MGH-PRISM study (456 CD, 305 UC, 1009 controls), 2) samples assembled from throughout North America and Australia by the NIDDK IBDGC (2291 CD, 1921 UC, 1337 controls), 3) an Italian-Dutch case-control sample (696 CD, 1560 UC, 900 controls), 4) CCFA Repository Collection (366 CD, 77 UC, 525 controls), 5) Swedish samples (0 CD, 529 UC, 501 controls), 6) Cedars samples (1369 CD, 1009 UC, 0 controls), 7) German samples (2887 CD, 0 UC, 2244 controls), as well as Immunochip genotype data provided by 8) the International IBD Genetics Consortium (5264 CD, 3771 UC, 6575 controls) and 9) UK IBD Genetics Consortium (2507 CD, 2847 UC, 4484 controls) (n.b., rare coding variants discovered in this study were contributed to Immunochip design). Samples 1, 3, 4, 5, and 7 were genotyped for sets of

markers using Sequenom iPLEX, Sample 2 genotyping was done as part of a larger IBDGC Illumina GoldenGate study - because of design constraints and assay failures not all markers were examined in all eight follow-up sample sets.

Additional Sequenom Genotyping was carried out for 9 SNVs in 2887 CD cases and 2244 healthy controls from the German popgen biobank collection. German patients were recruited either at the Department of General Internal Medicine of the Christian-Albrechts-University Kiel, the Charité University Hospital Berlin, through local outpatient services, or nationwide with the support of the German Crohn and Colitis Foundation. German healthy control individuals were obtained from the popgen biobank. Written, informed consent was obtained from all study participants and all protocols were approved by the institutional ethical review committees of the participating centres.

Genotyping of the Italian cohort was performed at the Laboratory for Genetics and Genomic Medicine of Inflammation (www.inflammgen.org) of the Université de Montreal as described above.

Swedish ulcerative colitis patients and controls were from a large case-control cohort of individuals described in detail in previous publications.(1-3) Briefly, 529 UC patients (aged 47.5 ± 15.8 SD, 57.1% males) and 501 healthy controls (aged 44.6 ± 14.6 SD, 52.7% males) were recruited at the Karolinska University Hospital, Stockholm, and at the Örebro University Hospital, Örebro, Sweden. Diagnosis of ulcerative colitis was established based on standard clinical, endoscopic, radiologic and histologic criteria. Control individuals were healthy blood donors free of inflammatory disease. Informed consent was obtained from all participants, and local ethics committees approved the study.

Genotyping of Swedish UC cases and controls was performed at Karolinska Institutet's Mutational Analysis core facility (MAF) using matrix-assisted laser desorption/ionization time-of-flight (MALDI-TOF) mass spectrometry and the

iPLEX Gold chemistry on a MassARRAY Platform (SEQUENOM Inc., San Diego, CA, USA). Average genotyping success rate was 99.9% per marker, and none of the individuals included in the study had > 5% missing genotypes. Typer 4.0 Software Package (SEQUENOM) was used for post-run data QC, allele calling and generation of genotypes, and all rare variant "moderate" calls were manually inspected to avoid false-positives.

NIDDK IBDGC subjects were recruited by the Genetic Research Centers (GRC) comprising the NIDDK IBD Genetics Consortium: Cedars Sinai, Johns Hopkins University, University of Chicago and Yale, University of Montreal, University of Pittsburgh and University of Toronto. Additional samples were obtained from the Queensland Institute for Medical Research, Emory University and the University of Utah.

CD and UC diagnoses were confirmed by standard clinical, endoscopic, histologic and radiologic criteria. Medical history was collected with standardized NIDDK IBD Genetics Consortium (IBDGC) phenotype forms. Demographics, including smoking behavior were also collected. Healthy controls are defined as those with no personal or family history of IBD.

Recruitment protocols were approved by Institutional Review Boards at each GRC, and all participants provided signed consents for data sharing through dbGap.

The Illumina GoldenGate BeadExpress® 384 plex OPA assay was used. All genotyping was performed by the Laboratory of Genomics and Human Genetics at the Feinstein Institute for Medical Research. A total of 416 variants: 308 registered SNPs and 108 novel variants were submitted for SNP design. For those variants with poor design scores, surrogate SNPs having the highest LD and r-square value greater than 0.8 were chosen from the adjacent 150kb

flanking region, from the HapMap 3 CEU population. The final 384 plex included 282 SNPs and 102 novel variants.

Grouping of samples

Samples were grouped based on informed ancestry. Samples from the Immunochip datasets were grouped into 8 separate clusters based on MDS estimates from GWAS data. Across all samples self-reported Jewish individuals were grouped together.

Quality Control of Immunochip datasets

All Immunochip datasets have been quality-controlled separately with the following common parameters in order to achieve a high standard data-quality while keeping as much as possible data:

i. Missing rate per SNP <0.05 (prior to sample removal below).
ii. Missing rate per individual <0.02
iii. Missing rate per SNP <0.02 (after sample removal above)
iv. Missing rate per SNP difference cases-controls <0.02
v. SNP frequency difference to HapMap <0.15
vi. Hardy-Weinberg Equilibrium (controls) p<10-6

To assess relatedness between samples and population stratification, we used the chip-wide SNP data passing stringent quality control . We removed SNPs in the highly stratified the Major Histocompatability Complex (chr 6, 25-35 MB) and the chromosome 8 inversion region (chr 8, 7-13 MB). We then pruned SNPs to insure that there was little linkage disequilibrium between SNPs (r2<0.2). We used the resulting 17,413 SNPs to assess recent common ancestry and population stratification with PLINK and EIGENSTRAT.

Supplementary Tables

Supplementary Table 1. List of sequenced CD candidate genes listed with GWAS associated SNP.

Gene	SNP	Chr	Critical Region	Risk-Allele	RAF	OR (Case/Control)	OR (TDT)
Confirmed evidence of association to CD (Barrett et al. 2008)							
MAST3	**						
NCF4	**						
IL12B	rs10045431	5q33	158.69-158.76	C	0.708	1.11	1.36
JAK2	rs10758669	9p24	4.94-5.26	C	0.348	1.12	1.21
CARD9	rs10870077	UC**					
C10orf22	rs10995271	10q21	64.05-64.12	C	0.387	1.25	1.53
EGR2	rs10995271	10q21	64.05-64.12	C	0.387	1.25	1.53
ZNF365	rs10995271	10q21	64.05-64.12	C	0.387	1.25	1.53
LRRK2	rs11175593	12q12	38.61-39.31	T	0.017	1.54	1.44
MUC19	rs11175593	12q12	38.61-39.31	T	0.017	1.54	1.44
NKX2-3	rs11190140	10q24	101.26-101.32	T	0.478	1.2	1.28
IL23R	rs11465804	1p31	67.4	T	0.933	2.5	2.77
C1orf106	rs11584383	1q32	197.60-197.77	T	0.697	1.18	1.2
KIF21B	rs11584383	1q32	197.60-197.77	T	0.697	1.18	1.2
IRGM	rs11747270	5q33	150.15-150.32	G	0.09	1.33	1.31
TRIB1	rs1551398	8q24	126.60-126.62	A	0.619	1.08	1.25
IRF8	rs17445836						
CREM	rs17582416	10p11	35.30-35.60	G	0.345	1.16	1.26
CUL2	rs17582417	10p12	35.30-35.61	G	0.345	1.16	1.26
NOD2	rs2066847	16q12	49.3	C	0.018	3.99	2.57
IRF1	rs2188962	5q31	131.44-131.90	T	0.425	1.25	1.26
LOC44110[illegible]	rs2188962	5q31	131.44-131.90	T	0.425	1.25	1.26
P4HA2	rs2188962	5q31	131.44-131.90	T	0.425	1.25	1.26
PDLIM4	rs2188962	5q31	131.44-131.90	T	0.425	1.25	1.26
SLC22A4	rs2188962	5q31	131.44-131.90	T	0.425	1.25	1.26
SLC22A5	rs2188962	5q31	131.44-131.90	T	0.425	1.25	1.26
ZNF300	rs2188962	5q31	131.44-131.90	T	0.425	1.25	1.26
CD244	rs2274910	1q23	157.65-157.72	C	0.682	1.14	1.62
ITLN1	rs2274910	1q23	157.65-157.72	C	0.682	1.14	1.62
ITLN2	rs2274910	1q23	157.65-157.72	C	0.682	1.14	1.62
CCR6	rs2301436	6q27	167.32-167.52	T	0.463	1.21	1.16
PTPN22	rs2476601	1p13	113.79-114.17	G	0.899	1.31	1.17
PTPN2	rs2542151	18p11	12.73-12.88	G	0.152	1.35	1.14
ORMDL3	rs2872507	17q21	34.63-35.34	A	0.473	1.12	1.24
C13orf31	rs3764147	13q14	43.13-43.54	G	0.221	1.25	1.19
ATG16L1	rs3828309	2q37	320.9	G	0.533	1.28	1.3
SP110	rs3828309	2q37	320.9	G	0.533	1.28	1.3
SP140	rs3828309	2q37	320.9	G	0.533	1.28	1.3
TNFSF15	rs4263839	9q32	114.61-114.78	G	0.677	1.22	1.07
TNFSF8	rs4263839	9q32	114.61-114.78	G	0.677	1.22	1.07
PTGER4	rs4613763	5p13	40.32-40.48	C	0.125	1.32	1.28
SBNO2	rs4807569	19p13	1.05-1.15	C	0.217	1.02	1.26
CDKAL1	rs6908425	6p22	20.63-20.84	C	0.78	1.21	1.09
STAT3	rs744166	17q21	37.74-37.95	A	0.565	1.18	1.25
ICOSLG	rs762421	21q22	44.43-44.48	G	0.389	1.13	1.21
PRDM1	rs7746082	6q21	106.52-106.62	C	0.289	1.17	1.19
GCKR	rs780094	2p23	27.30-27.77	T	0.397	1.08	1.13
C11orf30	rs7927894	11q13	75.80-76.02	T	0.386	1.16	1.07
LRRC32	rs7927894	11q13	75.80-76.02	T	0.386	1.16	1.07
IL18R1	rs917997	2q11	102.31-102.64	T	0.222	1.05	1.11
IL18RAP	rs917997	2q11	102.31-102.65	T	0.222	1.05	1.11

Supplementary Table 2. Primers designed for CD pooled sequencing experiment.

Amplicon	Gene	Forward Primer	Reverse Primer	Amplicon Length
SBNO2_e30_a001_0-100	SBNO2	GTAGCCACTCTCCCAGGGCT	CTCTCCTACAAGGTGGGGTC	348
SBNO2_e14_a001_0-100	SBNO2	CACATTCAGCACCTCTGGGT	CAACTTAGTCGTCCCCTGAC	301
SBNO2_e07_a001_0-100	SBNO2	TTAGGTCCCATGAGAGGGCA	GAACAGCAGGAGAGCAAGACC	289
SBNO2_e04_a001_0-100	SBNO2	GAGGTCACTGAGTCCCAGGC	GTTGAGCACCCCATCTCACC	268
GCKR_e04_a001_0-100	GCKR	AATTTTGTGATGAGAGGAGGG	GCCACCTGGGAGCTTAGTTA	287
GCKR_e19_cds_a001_0-100	GCKR	TGATGACCTCATTCCCTCAG	AATATTTCTCCCTACCCTGGGCT	364
LRRK2_e01_cds_a001_0-100	LRRK2	CTTCCTCATAAACAGGCGGG	GTTTGCAAAAATAAAAGCACAGTTTA	360
LRRK2_e41_a001_0-100	LRRK2	CATTTGAATAAGATTTCCTGTGC	TCACATCTGAGGTCAGTGGTTATC	261
MUC19_e12_a001_0-100	MUC19	GGCTCAGAGGGTTCTATTTCAG	TCCTCAACAGAGGACTCAATTTT	291
MUC19_e17_a001_0-100	MUC19	GGGGCTTCAGTCATGGAATA	TGAGGAAAATGTCCACTAAGTTC	277
ICOSLG_e07_cds_a001_0-100	ICOSLG	TCTGGGAGTCCATGCTCAAG	CAGCTTGCTAAAAACACGGC	285
IL18R1_e08_a001_0-100	IL18R1	GTGATGATGGACAACACTGGTA	AGAACATAAGGTAAGAAATTCCAGC	268
PTPN22_e13_a001_25-59	PTPN22	GATGAAAAATAAGGATTTTCCACTAA	AAGGTGCCAATAACACGGAC	296
P4HA2_e14_a001_0-100	P4HA2	GATGGGGGTGGGATATAAGG	AGCCTGTTGTAAACTTGGGTGA	290
P4HA2_e12_a001_0-100	P4HA2	TTCCAGATTTGGGATTCACC	GGCTGCTACAGAGCTGTTCAT	283
P4HA2_e10_a001_0-100	P4HA2	TTGAGCAGCTACGAACCCCT	CAGACTCTTGCCCTCAACTG	280
P4HA2_e06_a001_0-100	P4HA2	TAAAGAGACACATGCGGGACA	CTCTTTGTGGCTGCCTCTCT	268
P4HA2_e03_cds_a001_0-100	P4HA2	TCTGCCAGTTCCACTCTCCT	CTTTTTCACTTTGTTGATCCAGTG	288
CARD9_e04_a001_0-100	CARD9	CAGAAGGTCCAGTGAGCCTG	AACACGGCACCATCTGACACT	476
KIF21B_e27_a001_0-100	KIF21B	CAAGGACCATCCACAGGCTC	GAAGCCCAGCACCAAGTAGG	258
KIF21B_e22_a001_0-100	KIF21B	ACTGGAATGCCAAGGAGCTG	GGTAACTCCAGGCTTTGTTCC	244
KIF21B_e12_a001_0-100	KIF21B	ATCACCATGGTAACCAGTCAGAG	CTAATCCAGCTTCCCTGTCCT	278
KIF21B_e11_a001_0-100	KIF21B	TCCTAAACACTGGTGGGCAG	GCTACGGTGAGTGAGTGAGC	430
SP110_e20_cds_a001_0-100	SP110	AGTGTCTGGGTTTGGGTCCT	TCCTTCAATGTGTTTGTTTGCT	276
SP110_e17_a001_0-100	SP110	TGAGTTTGGGCATTAAATCGT	CCCTCTGCTCACCTTGTGTC	279
SP140_e01_cds_a001_0-100	SP140	CTCAGAGCTGCAGGAAGGAA	ACAGCCAATCCCTGCAATTT	288
SP140_e21_a001_0-100	SP140	TGGTGAGTGGCCATAGTCTG	TTGGTGAATAAACAGTCTCCTCAT	286

SP140_e24_a001_0-100	SP140	TGTTTTTCCTTCATATCCCAGG	ACAAAAAGAGAGACGCCTGC	288
SBNO2_e15_a001_0-100	SBNO2	GTCACCTGATGACGCCTGTG	GAATGTGGGTCGCTCGTTTC	486
SBNO2_e13_a001_0-100	SBNO2	GGACAGCCAGGCAGAAAGAC	CTTCAGAATACTCCTTAACCCACG	212
SBNO2_e09_a001_0-100	SBNO2	AGGACCTCGTGTTGCTGTTG	CATACCCTTTACCTTGCCCTCC	306
GCKR_e01_cds_a001_0-100	GCKR	CTGCCAACTGGAAGCAGAGT	GTGCTCTGGGCATATTAGGG	280
GCKR_e15_a001_0-100	GCKR	GGACGGGGTGAATATCCTGA	GAAAGAGGAGCAGGAGAGGC	277
GCKR_e17_a001_0-100	GCKR	TTTCCTGTCTGATAACAGGGC	CTAAGCATTGAGGCCAGGTC	286
LRRK2_e16_a001_0-100	LRRK2	AGGGCATTAGCTGGTGACTG	TCGCACAAGTTCCCAAATTC	266
MUC19_e09_a001_0-100	MUC19	CTGCAGGAGTCAAGACTGGC	ACCCTCAAATGTGAGACAAAAT	266
MUC19_e22_a001_0-100	MUC19	GGCCAACCTAGATCAATTCTATGA	TCCAGGGAAAGAAAACCAGA	293
ICOSLG_e05_a001_0-100	ICOSLG	AAACACTAACATCCACAATATGGG	TGCAACACTTTCTACTCAAGGA	285
ICOSLG_e02_a001_0-100	ICOSLG	TCCGGAGTTTGCAGAAGTGT	CTCACCTGCCTTTCTGGTTG	283
C11orf30_e18_a001_0-100	C11orf30	TGCTGTGGTGATTTCTGGAG	TCTTATTTCAAGCATGGGCA	255
IRF8_e05_a001_0-100	IRF8	TGAAACTTGACTTTTGTCTCCTG	CTGGTGACAATTCTAGGCCAC	279
IL18R1_e03_a001_0-100	IL18R1	TTTGCAAAGTTTCTGTAGCATTAT	TAGCTCTCAGCCCCTCATTT	293
PRDM1_e06_a001_38-65	PRDM1	GAATACGGTGTCCCCTGTGG	AGAAGTGGGGTTGAGCATGG	325
PTPN22_e11_a001_0-100	PTPN22	CCCTTGACCAAAAGCAAGTA	TCGTCCTTTTGGGTCTGCTA	268
PTPN22_e10_a001_0-100	PTPN22	GCACGTAATGGTAAATTGGTGTTA	AGCAGATGCCATGGAAGTACA	288
PTPN22_e08_a001_0-100	PTPN22	GGTTCATTTTGGGTAGAAGGTTTAT	acccagccTCAACTTTAATCTTATG	353
KIF21B_e23_a001_0-100	KIF21B	TTGAAATCGTCATGGCTGGT	CTTGGACAGGGCTTCTCAAC	291
SP110_e19_a001_0-100	SP110	CCCAGGCTGAAACTGAGTGTA	TAACGTCACCTCTGGTGGCT	272
SP110_e11_a001_0-100	SP110	GGCCCTCCTACATTGAGCTA	ACTCCTGAAAGGGTAGGGGC	274
C11orf30_e05_a001_0-100	C11orf30	ACCCTTGTAGCAGCGCTTTC	AGAAAAGTAATCTCTAACACAGGCATA	248
C11orf30_e07_a001_0-100	C11orf30	AAGTGTATAATGCAAGATTAGTTTGGT	AATTGGCAGGCAGAGACTCC	352
C1orf106_e03_a001_0-100	C1orf106	ATCTTCAGCCTCCCTGAGCA	GAATTCCCTCAAGAATGGGG	275
CARD9_e13_cds_a001_0-100	CARD9	CACCAGATTCCTCGTTCCAG	GTTTTGAGAACTACCGCAGG	330
IRF8_e06_a001_0-100	IRF8	GCATTTGTGAAACAATGGCTC	CTTTGTCTTCAAACAGTGCCC	290
KIF21B_e16_a001_0-100	KIF21B	TAGGACAGTGGAGACTGAGGC	CCCTGTCTCACCACCATAAT	254
KIF21B_e25_a001_0-100	KIF21B	ACTCCAACCCACATTCCTGC	CAAGTTGATCCACCCTGTTCTT	270
KIF21B_e34_cds_a001_0-100	KIF21B	AAGAAGCAGCTCAGATCGAA	CTTTCTTTCTTGTTCCGACTTTG	281
LRRK2_e35_a001_0-100	LRRK2	GGGTGTTTTGTGAGGCTGTATAAC	TGCCATCTCCCTAATTTCTCTAAA	285
MUC19_e08_a001_0-100	MUC19	CTGGAATAGGCCCTCTGAGT	TAAGGAAGCTTTGGCCTCTG	270

MUC19_e26_a001_0-100	MUC19	GGGAGTTTTTGGTTCAGGCT	TGTACATGCCCAAATGATGG	277
MUC19_e27_a001_0-100	MUC19	AATTTGTGCCATCATTTGGG	TGCTGCATCTTGCTGATGTT	269
ORMDL3_e04_cds_a001_0-100	ORMDL3	TGTCCAGAGGCTTCTTCTTTCT	GGGTCCTCCCAACAGTCTTT	282
P4HA2_e05_a001_0-100	P4HA2	AGACACAACAACCCAGGCAT	AAGACCTATGTGGGTGGACG	273
P4HA2_e09_a001_0-100	P4HA2	TTCAGTGACCCGAGAAGGAC	CAGACTGGGACCGAGGATTA	274
P4HA2_e17_cds_a001_0-100	P4HA2	GTCTGTCACGTTGACATGGG	AATTCTTCAGCAAATTGAGGC	278
PTPN22_e01_cds_a001_0-100	PTPN22	GGAAAATTGGACCCCCATAG	ACAGCCTTCAGCATGCTCTG	281
PTPN22_e16_cds_a001_0-100	PTPN22	CCATTCTTCCAATCTTAGGGC	TGCCATGGGATATAAAAATCAGTG	290
PTPN22_e21_cds_a001_0-100	PTPN22	GGTGTACTTGCAGCCCATATTA	TGGTCAAGATGCTGCCTAAC	238
SBNO2_e03_a001_0-100	SBNO2	CTGCACCCAGAACTCCGTTT	CAGGTCCTCTGCTTCCCACT	200
SP110_e06_a001_0-100	SP110	ACACAGAAACAAAGGCAAGC	TCCCACTCTATGCCATCCTT	274
C11orf30_e19_a001_35-85	C11orf30	AAAACTCTGCAGTGCTTCCA	GATTTTTGGTTGGTTCGGAGC	303
C1orf106_e01_a001_0-100	C1orf106	GGACAGATGTCTTGGGATCTT	CTAAGGCTGAAGACCGTGGGT	277
GCKR_e13_a001_0-100	GCKR	CCACAAGGGCTACTCCTCAC	TTTTCTTCCACCCTCAGCAC	275
LRRK2_e06_a001_0-100	LRRK2	GATTACACTTGATGTATCTCACACAAC	ATGAATGGGTTGAGCATCCA	366
LRRK2_e20_a001_0-100	LRRK2	TGATTTCTAAGTTGCTGGTGTATCT	TGGGTCCTATTGTTCAATGTCA	292
LRRK2_e50_a001_0-100	LRRK2	GAATTGTGAATTCAGTTCCAAGG	TCATTCACATCATTGCCCTG	241
MUC19_e19_a001_0-100	MUC19	CTGGGATAACTACTGGCACGA	AGCAAAGAACTGTTGAGAACCC	285
MUC19_e35_cds_a001_0-100	MUC19	AAAAGAATAAATTGGTTACAAGGCT	AGGCCTGTTAAAACTGGGAAGG	333
PRDM1_e06_a002_77-100	PRDM1	ATGAAGGACAAGGCCTGTAGC	CCCCTTGGACTGCTCTCTCT	291
PTPN22_e04_a001_0-100	PTPN22	CCTGAACCCTGAAGTTCCAC	aaaGGAAGCGGAGGACTAGG	260
C1orf106_e10_cds_a001_0-100	C1orf106	ATTTACCCACCTGTCCTTTCTT	CTTAGGTCAGGGCTTCTGAAA	300
GCKR_e08_a001_0-100	GCKR	TCATCATGCCCTTCTCTCTC	GGGGCACCCCATATTTACTT	266
GCKR_e10_a001_0-100	GCKR	TCTTCATCCTCTCCCAATTCC	CTGCAGTCTTACTGCCACCA	276
GCKR_e12_a001_0-100	GCKR	GATTGCATGTTGAAGGGTCA	ACAGTGTTTGGAGAGCAGCC	280
IL18R1_e07_a001_0-100	IL18R1	CGAACAGAGCCTACTGCTAAA	TGGGTTCTGGTGTGGAAAAT	279
SBNO2_e18_a002_0-100	SBNO2	CTCTTCCCTAAGCGTTCCAT	CTCCACACAGGAGACCCTCT	423
SBNO2_e16_a002_0-100	SBNO2	CAGGAAGGACCTGGAAGAAG	CCCATCCTGCCCTAAATTCC	383
SBNO2_e02_cds_a002_0-100	SBNO2	AAAAGCTGGGGCCATGAGTAG	CGTCTCCCTCTGCAGATCAT	456
GCKR_e11_a002_0-100	GCKR	TCTAAGGGAGCTGTGCCTTC	CCCTGCTCTCAGTATTCTCCA	259
LRRK2_e10_a002_0-100	LRRK2	GCACGTGCAGGTAGGACTCT	TTCTAACAATAAAAGTTACGGTTAAGG	273
LRRK2_e36_a002_0-100	LRRK2	CTTGTGTTGTGTGCAGTAGATTT	AAACCTCATACTGTCGCAAAG	322

LRRK2_e49_a002_0-100	LRRK2	GGTGGTGGTGTCATGTTTTAAT	GGAAAGTTCAGAGAAAATGTAAACG	411
MUC19_e30_a003_0-100	MUC19	TGGTTTGTTTATGTAGCCTGGAG	ACACCTGCCTAGTATTATGTTCGAC	280
PTGER4_e02_cds_a005_0-27	PTGER4	GCAGCTTTGTCTCTCTTCTACC	CTTCATGTACGTGGCGATGG	319
IRF8_e08_a002_0-100	IRF8	GAGTTGGAGGTCATCTCGGC	CCAAAACTTGATCCTTACATGAC	550
IL18R1_e02_a002_0-100	IL18R1	AAAAAGTTACCTTGTCATTTTGGT	AAGGACAAGAGGGCTGTGAA	356
CARD9_e02_cds_a002_0-100	CARD9	CTCCGAGCTGACTCTGTGGT	CACTGTGTCTCTGCTGCCTC	457
C1orf106_e04_a002_0-100	C1orf106	TCTATGGAAGGCAGGCACAT	GATCAGCTTTGGGGATCGAG	258
C1orf106_e09_a004_36-63	C1orf106	CTCCCAAGACCCATCGTCAC	GAGTGCCCTCCTCCTCGTAAG	201
SP110_e13_a002_0-100	SP110	TGCAGTCTAGCTAGCAGGGTC	TCAAAAGGAAGATGGAGTGAAG	272
SP140_e12_a002_0-100	SP140	AGCATTTTCCCATCTTGGAT	GCGAATGTGAAGAAGAGAAGG	282
SP140_e17_a002_0-100	SP140	TGTACTCAACCTCCACGCTG	TGGACAGTGCCCTAATAAACTC	280
SBNO2_e27_a002_0-100	SBNO2	TCTCACCCAGTACCCTCGCT	CAAAGCAGGGAGAGGGATTC	502
C11orf30_e11_a002_0-100	C11orf30	CTTGACACCTTGAAATTCCATT	TCCTCTCAATGAAGGCCACC	352
PTPN22_e03_a002_0-100	PTPN22	CTCACCTAGTCCTCCGCTTCC	actcatGAGCATACACTATTCACAG	551
CARD9_e11_a002_0-100	CARD9	CTGTGATCGGTCACCCTGAG	GAAAGCAGAGTCTACACTGGACCC	283
CARD9_e03_a002_0-100	CARD9	AGCTCAGGGCTCCTAGGGAT	CGCAGGGTACCTCATCCAAG	383
KIF21B_e28_a002_0-100	KIF21B	AAAGGAAGCAGGGGTGAGTAG	ACACACGTGCCTTGATCAGC	274
KIF21B_e19_a002_0-100	KIF21B	CCAGAGAAGTTGGAGGTGTCA	GGTTAAAAGGAGTCCCATCCA	264
KIF21B_e13_a002_0-100	KIF21B	GGGCCTGGACACCATTATCT	CCCTCCTAGAGGGGTTCTGC	417
SP140_e23_a002_0-100	SP140	GGTTCCTGAATCCCTTCAAA	ACACAGCCACATACTCCAGC	267
SBNO2_e30_a005_0-100	SBNO2	AGGCAGTCCTTACCCTCCTG	CTCTCCTACAAGGTGGGGTC	590
SBNO2_e21_a002_0-100	SBNO2	CACCTCCTCACCCACTAGGC	GACCACTGTAGGACATAGGGAAG	449
LRRK2_e30_a002_0-100	LRRK2	AATATGCTAAAAGTGGTCAATCCTA	TTGTTTGCCCAAAATAAATTAAC	406
MUC19_e07_cds_a002_0-100	MUC19	CCAAGTAAAAATGGGAAGTTGG	ATATCAGGCAGCTATTCATACTTTC	262
MUC19_e34_a002_0-100	MUC19	CCTCAATAATGTTGTGAAGTAGTTGC	CCATTTTCTCCACAAGGTTACAA	276
GCKR_e06_a002_0-100	GCKR	TACGCCATAGGCTTCTGCTT	GGATTCCAGGGCACAAATCT	276
LRRK2_e12_a002_0-100	LRRK2	GAATTATCTTTAAGCTGTCAATGAACT	TGATATTCTACCTGGCCCAAT	340
C11orf30_e20_a005_55-100	C11orf30	CACTGGTGAAGCAGGATCATTA	AAACCCACAACCCCCAAGAT	289
IRF8_e07_a003_6-100	IRF8	GTGATCAGCTTCTACTATGGGG	GATAGGAGGCGTGATCTCTAAG	428
IRF8_e09_cds_a002_0-100	IRF8	GTCAAAGACAGTGCCCACCC	CATTCTTTAATCATGATGCGGG	328
C1orf106_e05_a002_0-100	C1orf106	ATCTCCCTGCCTCTCTCTAGCC	GAGCCAAATGCAAAGGCAGT	363
KIF21B_e04_a002_0-100	KIF21B	AAGTGAGGGCTTGGGACTACC	GATTCAAAACCCAGAGGGGAC	368
GCKR_e07_a002_0-100	GCKR	CACTGACATTGACCAGAGACC	TCCATTCCCACTCACTTCCT	290

CUL2_e21_cds_a001_0-100	CUL2	AACAGCAGAAAACAGGGGCT	TGGCAGTAGTTTGACTTCCAC	290
CUL2_e20_a001_0-100	CUL2	GCCCTCTACCCTAAGAAGCC	TGCACTCACTGATAAAATGAATTG	243
CUL2_e18_a001_0-100	CUL2	CTGATGATGCTTAAATTCTTTGGA	GAATGGTCAATTCACTGTCAGTTT	280
CUL2_e15_a001_0-100	CUL2	CGTACTGAAATGATTAGTAAATGCTCA	AGACCTAGCAATACCGACACAT	290
CUL2_e14_a001_0-100	CUL2	TTTCTAGGACAATACCACATTAAAGA	AGTGGAAGACAGGCTCACGA	297
CUL2_e10_a001_0-100	CUL2	TGGCATGGTAAGACTTTTTCTTT	TGGGAGTTAACCCTCAGATGTC	269
CUL2_e08_a001_0-100	CUL2	AAACCAGGCCAGATACAAAAT	TCCCCTTAAAGGTAAACAGCC	271
CUL2_e06_a001_0-100	CUL2	CAGTAAAACCAGCCCCAAGG	TCAGCCCTGATGATTTTCGT	226
CUL2_e04_a001_0-100	CUL2	CCCCACAGCAGAGAAAAGTC	TGACAGTGAAAATGATTATTGCTT	297
CREM_e05_cds_a001_0-100	CREM	TGCTGTGTTACAACACTGTGAG	ACCCCTCTTTTCACAAGGCT	284
CREM_e06_cds_a001_0-100	CREM	CAGAAAATGACTGTTCCAGGAC	TGGCTGTGGTGTCTGAATTAC	298
CREM_e10_cds_a001_0-100	CREM	ATAGGGCTTGGGTTTCAGTG	TGAAACACTCAGATGTTCCTGAAT	289
CREM_e11_cds_a001_0-100	CREM	CCGCTTTGTAAAATGCTGCT	CCCGCCAGCTTATACAATACAT	284
CREM_e12_a001_0-100	CREM	TAAGGAGGGGAAGAGGAAGGG	ACTACATATCTCTCTTTGCAAGCTTTA	262
CREM_e14_cds_a001_0-100	CREM	GGAAGGCTGTTCCTGTAGTCAT	ACGTGTCCTTCCCACAAGTC	278
CREM_e14_cds_a001_0-100	CREM	CTTTGTTTGCCCTTTGCTTC	CCACCATTCTTGCCAGCTAT	260
TNFRSF6B_e06_a001_0-100	TNFRSF6B	TTCCCCTGACCCTGTTCTTC	AGGAGTGAGGGCACAAGTGG	291
TNFSF8_e04_cds_a001_0-100	TNFSF8	GCTTTCTTCCTGAAGGCCAA	TGCGTACCTACAATCTTCTCAATAAT	477
TNFSF8_e03_a001_0-100	TNFSF8	CATTTGGCCATGGGACTCTA	AGAACGGTAATTGTCTCCATCAA	278
TNFSF8_e02_a001_0-100	TNFSF8	GTGGGGATGAGTTCAGGAGC	TGTCACACCTCATGCTGCTT	280
TNFSF8_e01_cds_a001_0-100	TNFSF8	AAAAGGAAAGGGAGGAAGAGG	CCCAGAACTGAATCAGATGAAGA	286
TRIB1_e03_cds_a001_0-100	TRIB1	GGAAATAATGGCTTGGTTCCTC	GGTGTTAAGAATTATGAGGTTTCTGA	535
FASLG_e01_cds_a001_0-100	FASLG	AAGTAAAACCGTTTGCTGGGG	GAAAAGCACTTTGCAAGCCAG	478
PRDM1_e07_a001_0-100	PRDM1	AGAGCCAGCTTGAGAGCAGA	TGTTGGCTTTAACTACGGGC	270
PRDM1_e02_cds_a001_0-100	PRDM1	TCCAGCACTGTGAGGTTTCA	ACCAGGTCCCCAATCTTCTT	272
P4HA2_e15_a001_0-100	P4HA2	GCCTTGCTTCAACCTCTGAC	CATGGGAAGAGCCTTTCTTG	276
P4HA2_e14_a002_0-100	P4HA2	AAACAGAGGCCAACTCTCCC	CGTAAGTACTGGGTCCAGGC	282
P4HA2_e08_a001_0-100	P4HA2	AGAGTCCAGGTCCACAGCTT	CCTTTTCCTTCTCATTCCTTG	274
P4HA2_e07_a001_0-100	P4HA2	ATTCCCATCTTACCTTCCCCTA	GGAGCCTGACAATTTACTTCC	305
ORMDL3_e03_a001_0-100	ORMDL3	CTGGTGTCTTCCTGTAGCCC	TTGTGTCTGTGACTGGGTGG	262
MUC19_e32_a001_0-100	MUC19	CAAAGTGCCTCTGGAAAATCA	TGAATCTCGGCGTTCATTAG	267

MUC19_e30_a002_0-45	MUC19	GCCAATTACAGTACTGGCCACC	AGGGCACTCCTCAGGTTTACA	299
MUC19_e23_a001_0-100	MUC19	TCCATGAATTCTCATCTGGC	TGAGCATGTAATCTTCATTTGAGAC	278
MUC19_e18_a002_0-100	MUC19	GGAATTCATAAATTGATTATCTCCTCT	AGTTGCCATTCCAACTCCAG	290
MUC19_e16_a001_0-100	MUC19	CATGGTGACAGGATAGCACAA	TCCAGTGGTATTCCCAGGGT	286
MUC19_e10_a001_0-100	MUC19	GGAGCAAAAAGGTCAGTGGA	AAAGGGAGCTTTCTTCTCAGC	283
LRRK2_e33_a001_0-100	LRRK2	TCAGTGTCTCAAATATGCTTAAATATG	CCCCGTGCCACAAAAATAAA	351
LRRK2_e31_a001_0-100	LRRK2	GCAAACACAAGAGGGTTTTG	AATAAATTACATTTCTCTACCAGCCT	290
LRRK2_e19_a001_0-100	LRRK2	GCCAGTCTCCTAAAAGGAAGAAA	TGGCATGAATAACCACTAAAAGACTA	410
LRRK2_e13_a001_0-100	LRRK2	GGTTCTGCCCTCCTGTACTT	TTGTTATCTCTTCCCACAATCAAC	252
LRRC32_e03_cds_a004_91-100	LRRC32	GCTGATCTCATTGGTGCTCA	GCAGGGTTCCTGGATAAACA	259
LRRC32_e03_cds_a005_55-75	LRRC32	ATTCAAGAGCTGGGAAAGGG	CTGATGGACATCGAGGATGG	405
LRRC32_e03_cds_a006_34-49	LRRC32	ATTATCCACCAGGCTCAGGC	AGCAGAAACTGCTTGCGGAC	321
KIF21B_e33_a001_0-100	KIF21B	AGCAGAGGCAGACCTCTCTGA	GTTCTCTCCCTCATCACCTCC	285
KIF21B_e31_a001_0-100	KIF21B	CAAAGATTGGGGAGAGTGGA	CCTATTTGTTTTGGGGCAGA	260
KIF21B_e29_a001_0-100	KIF21B	ATCTGTACCCAACACAGCCT	CTAGTGGCTGACATCCCTCC	292
KIF21B_e26_a001_0-100	KIF21B	GGTGAACCTGAATGGCTCAG	GAATGTGGGTTGGAGTGGAG	278
KIF21B_e09_a001_0-100	KIF21B	ATACTCTGGGGTGGATGTCG	AGCTCTGCTGCCTGACTGCT	308
KIF21B_e06_a001_0-100	KIF21B	ACTGAGCTGTGTGGGTGCTG	GAGGGCTGGGTGAGTCAGAG	250
KIF21B_e02_a001_0-100	KIF21B	ATTGGATGCAGGGAGACCAA	CCTTTTCCTCTGGCTCATGG	395
IRF8_e07_a002_0-91	IRF8	AGGGACCCTGTATGTCTCCC	GACCTGGACCACCTCATCAC	416
IRF8_e02_cds_a001_0-100	IRF8	GGGTTGCTGTGATGAATGAGA	AGGGAAAGGAGACACTGTGC	289
IL18R1_e06_a001_0-100	IL18R1	CAAAAATCAGAACTTTAGTTGCC	TCAGCCACCCAGTGTTTTTA	248
IL18R1_e05_a001_0-100	IL18R1	GGTGCTTTCAACATCTGTGACC	TGGAAGGCTCTATTATGTCCC	294
ICOSLG_e04_a001_0-100	ICOSLG	TACCTCAGAGGCTCTCCCAAG	CATCTCTCACAATCTGTTTCCC	399
ICOSLG_e03_a001_0-100	ICOSLG	GTGAGGCTGGGATGGACCTA	GAAAACAGATTTTGCACCTTTT	430
GCKR_e09_a001_0-100	GCKR	GACCTCTGACCTCAATCCCA	CCAAGGAAGTGGCAAGCAGT	283
GCKR_e05_a001_0-100	GCKR	CCACCGACTAATCATCTCCA	AGGTCAAACAAAATATTCATAGGCTAA	285
CARD9_e10_a001_0-100	CARD9	CCTCTGGGTGTACTGCTGTC	CTCTGAGAAGCTCTGTGCAGG	210
CARD9_e05_a001_0-100	CARD9	ATGTAGGGGCTGCTCCTGTC	CACCAGAGTGAGGAGAAGGG	492
C1orf106_e08_a001_0-100	C1orf106	ATTCCTGCCACTGACCTGTCT	CACCCTCCAACGATAAGCTC	270
C1orf106_e06_a001_0-100	C1orf106	ATTTAGCACACAGAGGGTATTTCC	CACTTAAATTCCACACCAGCAA	290
C1orf106_e02_a001_0-100	C1orf106	TGCTCTCAAGCATCTATTCAGC	GACTAGGGCAGGGGGATTTC	266
C11orf30_e21_cds_a001_0-100	C11orf30	CTTACCATAAAGATTCTCCTGTGTCT	AGGGTTTCCCTGGACACTTG	292

C11orf30_e20_a002_40-96	C11orf30	GGAAGCTCAGATTGATACAAATG	AATTTCTGCTGGACTCTCTTCA	301
C11orf30_e19_a002_0-44	C11orf30	CTTATTTCATTCTTACTTTGTGCCT	TTTGTGCTGGAGCTGGTCTG	297
C11orf30_e17_a001_0-100	C11orf30	GCTAATTCAAGCAGAGGCCA	CCAGATGAGAGTAGAGCGCA	256
C11orf30_e12_a001_0-100	C11orf30	CCTTAACTCAGGCCTCCTTTT	TTGTGGGCTTCTCAAGTCCT	269
C11orf30_e03_a001_0-100	C11orf30	CCCTGAAGTTTCACTATACATAGCC	TGAACACCTCTGTTCTTCAACTT	292
C11orf30_e19_a003_70-100	C11orf30	CCTACCTTAATGGCACAGCC	TGCCAGGAGAAAGCACTGAT	279
CARD9_e08_a001_0-100	CARD9	AGGGGTTTGGTTAGGTTGGG	GAGCCTCTGCTTGGAGTTGG	413
GCKR_e03_a001_0-100	GCKR	ACCAGGTAACCAAGACCCAA	GGGAAACAGAAAAGTTTAGGCA	281
IL18R1_e01_cds_a001_0-100	IL18R1	AGCTTTGGCTGAATCTGTTTT	TGCAAGTTTTTCACTCTGCTC	285
KIF21B_e07_a001_0-100	KIF21B	ACTATTTTCCAGGAACCTCCC	GCCAAAGAGGCAGGTACAGA	276
KIF21B_e14_a001_0-100	KIF21B	CTCAGGGAAAAGGGAGGGTC	ACATCTGCTCCCCAATTTGC	309
KIF21B_e17_a001_0-100	KIF21B	CAAAAGAGAAGGCAAGCAATG	GTTTGGCTGCATGAGAATCA	277
KIF21B_e18_a001_0-100	KIF21B	CTTCCCTCTTCCTCACCCTG	GAGCAGCCTCTGTGTTCCTG	421
KIF21B_e30_a001_0-100	KIF21B	AACCAAGACCTGGGGACCTT	GATTGTTTTGACCTCTGGCCT	272
LRRC32_e03_cds_a007_71-88	LRRC32	GGAGATGCAGGTGAGGGAAT	ACCTCAGCCTGGCTCACAAC	338
LRRC32_e03_cds_a008_9-30	LRRC32	CTCAGGGACACCTCCTCCTG	CTTTCTTCCAATCCTGGGCT	416
LRRK2_e08_a001_0-100	LRRK2	CCATGGATGAGAATTCAGCTA	AGGACAGAAAAGGCATTGAA	275
MUC19_e15_a001_0-100	MUC19	TGCAGTTGAATACCTTTACACAAAT	TCCCAAGAAGAGCATGGACA	300
MUC19_e21_a001_0-100	MUC19	CCCAGAGGACTGAAGCTATCC	TGAGAGCCATTTAAAGGAGAAT	291
MUC19_e24_a001_0-100	MUC19	CCACGTGGACGTGTCTTAGG	TCCTGTCTTCTGAGGTGCCA	354
MUC19_e25_a001_0-100	MUC19	GCTGTAGCTCCTCTAGTCCCC	AGCCTGAACCAAAAACTCCC	281
MUC19_e28_a001_0-100	MUC19	GTGAGGTCCAAAGGAAGCAC	AGAGGAACCATTAATGTGGAACA	283
ORMDL3_e02_cds_a001_0-100	ORMDL3	TCCTCCTATCCCTCAGCCCT	CTTGTTTGCAGAACAGCAGAGAG	288
P4HA2_e11_a001_0-100	P4HA2	CCCTGTGAGACCATCTAGCC	GTACTGCCTGTTCATGCTGG	270
PRDM1_e06_a005_58-85	PRDM1	ATCAACAACTTTGGCCTCTTCC	CATCGCTGCTGAGGTAGCTT	315
PRDM1_e08_cds_a003_76-100	PRDM1	GAAACATGGGGAATGGACTC	TTCCAGAGGTGAGGGGAAAC	281
PTGER4_e02_cds_a004_51-93	PTGER4	TATGCGTCCAACGTGCTCTT	GTAAGATGACCATCTGGATCTCG	391
PTGER4_e03_cds_a003_0-43	PTGER4	GATGAAATCTAAATGTGCGATCT	GAGCGAGAGTGGCCTGACAT	303
PTPN22_e06_a002_0-100	PTPN22	GTGATTGCAAAGACAAGCTCTC	TGAAGCCCAAAGGGGTTAGA	279
PTPN22_e13_a004_53-87	PTPN22	CCTTGGTTTCTCTCTGCTGTATCA	CAAAAGAAGAGCTAGTTTTGCAC	297
PTPN22_e17_a001_0-100	PTPN22	CCTTTCCATTTAGGTCCTACTTT	TTGGTGATTGTCGGCTAAGAT	283
SBNO2_e26_a001_0-100	SBNO2	AACAGTCTGGTAGGGGAGGG	GAAAAGGAGCTAGGGGCTTC	283
SP110_e05_a001_0-100	SP110	GAGCTAAGCGGTATCAGCCC	GGGGGTGGGTAAGACCTAGA	286

SP110_e07_a001_0-100	SP110	GAAAGCAACTCTGAATTCACCC	TCAAGACAGAGCCAGGGAAT	284
SP110_e12_a001_0-100	SP110	CCAATTCAGCTGGTACCCAT	TGAAGATTGTTAACCTATGAGGC	283
SP110_e15_cds_a001_0-100	SP110	GGAGGTGACTCAGAGCTTGG	TGGGTTTGAGATCCAGTGTG	274
SP110_e16_a001_0-100	SP110	ATCTGGAGGTGAGTGCTGTGT	CTGGCACACACAAAAATGATG	262
SP140_e04_cds_a001_0-100	SP140	AAAGGAGGAAGCCTTCTCTGA	AGTCAGCTCTCACACCCTGC	272
LRRC32_e03_cds_a009_0-15	LRRC32	CCAACTTCTGGCTTCCACAG	GCCTCTACCTGCAGGGGAAT	504
SP140_e28_cds_a001_0-100	SP140	GAGGAGCATTGTTTTTGTCTTTAT	TAGGGCTTCTGCATGGAGGT	541
CARD9_e05_a002_0-100	CARD9	ACCTTCCAGGCACGGAGT	GACCTTCTGCAGATTGACCA	506
PTPN22_e13_a005_0-33	PTPN22	gagaagaagcagaggagaagagg	TGCCTCTAATGTAAAGCACCAT	380
SP110_e08_a001_0-100	SP110	ccctccacagcttgacctac	agtccttcttcccctattactcc	306
C11orf30_e10_a001_0-100	C11orf30	GTGTGACTACTGCCCTCTTGG	tcccaagtgtcaagcatattg	322
LRRK2_e43_a001_0-100	LRRK2	CTTTGCAATGTCTGGACCTTTT	tgcctgggatggtgaaaat	355
SP140_e03_a001_0-100	SP140	GTCTTTTGTAACCCATAAATCTCTAGC	GGCTTGAATCCCCCAAAG	263
SBNO2_e29_a004_0-100	SBNO2	ATGTTGGGCTTGTACACCGT	GACGGGCAGGTGGTCTTCTA	360
IL18R1_e10_cds_a001_0-100	IL18R1	TGTTCCCCCTTTCAGTTATCTT	AGTTCTTTTTGGCGTTCACTG	439
IRF8_e03_a001_0-100	IRF8	GGCCATGAATTTAATGTGCTTC	TCTGTGGTGAAGGCTACAGG	286
KIF21B_e15_a001_0-100	KIF21B	ATTCTAAGAAGATGTCACAGGCACT	CTGTATTTAGTGGGAGCCCTTAG	318
KIF21B_e32_a001_0-100	KIF21B	CGAGTCTACCCTCTCCCTTTC	CACTTGGCCTCTCCCTACTG	241
LRRC32_e02_cds_a001_0-100	LRRC32	gCAAGGTACATTCCCTTTCCT	GGGCTTACAGTGTTCCCCTT	251
LRRC32_e03_cds_a001_47-62	LRRC32	TTAAGTCAAGGAGCATCAGGCAG	CCTGAACTTGTCCAACAACC	305
LRRK2_e18_a002_0-100	LRRK2	CAGTGTTTTTCACTTGCATCC	GAGCAAAGACAAAAGTGAGGC	276
MUC19_e11_a001_0-100	MUC19	CCCCGTCCTTCTAAGTCCTC	AGGAAATCTGCCCAAGTTAATTC	289
MUC19_e20_a001_0-100	MUC19	CAGGATCAGCCAGAATTTGC	TATCATGAACGAGGGCCTTT	274
P4HA2_e16_a001_0-100	P4HA2	ACATTACATGGCTGGATGGTAA	GAATGTGTTCTGAAAGGCCA	290
PRDM1_e05_a001_0-100	PRDM1	TCCTTTTACATGCCTGTCTTCT	ACCGACATTACTGGCATTTTT	334
PTGER4_e02_cds_a002_24-58	PTGER4	TTGGGCACTTTGTTGGTGAG	CACCAGGTGTCTGGGTACTGC	311
PTGER4_e02_cds_a003_70-100	PTGER4	CTCTGCAACGTGCTTGTGTG	GAGAAAAAGGCCGAGTAGGG	307
SBNO2_e06_a001_0-100	SBNO2	ACTGAGACCTCACTGGACCGT	GTCTCCCGTACCACCCTCAC	292
SP140_e08_a001_0-100	SP140	CCCAGGGAGCAGTTCTACAC	GGAATGGGAAGGGAGATGAA	260
KIF21B_e08_a001_0-100	KIF21B	TAGCTCCAAGGGGAGTGCTG	CACCCTCACAACAGCCTTTCT	271
LRRC32_e03_cds_a002_22-41	LRRC32	GTTCTCGGCAAGATTGAGCC	CCCCATACACCTTTGCCAAT	368
LRRK2_e40_a001_0-100	LRRK2	GTTGAATTACTCTTACATGATTTTGG	CTCAGAACCTGAGAGCAAGTCT	297
PRDM1_e08_cds_a001_0-45	PRDM1	TTCCCTGCTGTCTCTCTCCC	GTCAGATCTTCCAAGGGCAG	300

PTPN22_e02_a001_0-100	PTPN22	CCAAGATCAGGATCAGTAAGCAA	TGCCAGTGGAAAAGGAAAGC	288
SBNO2_e25_a001_0-100	SBNO2	ACACAACCGGCCTACAAAGC	CTAGGAGTCTAGGGGAGGGCAG	332
SP110_e09_a001_0-100	SP110	CCTGCTTCAGGAGAGACCCT	TTGGGACTGTAGGAGGCTCA	281
SP110_e10_a001_0-100	SP110	gCCTCAGTGTAAGACAGCTCTGA	AGGAGGGTAAGACCTGGAGC	278
SP140_e07_a001_0-100	SP140	CAGGAGCATCCTGAGGTCTG	ACACACCTGGGGAAAACAAC	282
SP140_e10_a001_0-100	SP140	AAAGAGGAGCTGCAACAAGG	TCCATGCCAACTAGAAAACAA	288
SP140_e19_a001_0-100	SP140	CAGCTGCTCATAAATGACTGTGA	TTTCCTTGACACATTTTGGAA	286
C11orf30_e02_cds_a001_0-100	C11orf30	AGGGAGGACAAGCTCTTTGG	GTAGCCCTCAATAAATCTTAGCTACT	342
CARD9_e12_a002_0-100	CARD9	TGAAGACAGGTGTCTCAGGC	GTGAGAAGGACGGGTTCCCT	216
KIF21B_e03_a001_0-100	KIF21B	ATGGTCCTGTTGTCCACCAA	CAGGGTTACACCCTCCTATCT	271
LRRC32_e03_cds_a003_83-99	LRRC32	TACAGGCTGTTCCCAGACAGG	GAAGGTCTCGTGCCAGGTTC	318
MUC19_e13_a001_0-100	MUC19	GCAACACATCTAAACCTGGGA	AAAAGTGAGCAGGTTAGGAAGC	270
PTGER4_e03_cds_a001_63-100	PTGER4	GAATTTGCTTCCAGGTGTGC	CAGGATTTTATAAGGGTCCAGAAA	282
SP140_e02_a001_0-100	SP140	CCACCACAAACCTCTTGGAA	TGGCATAATTGAGTATGGACCTG	290
SP140_e09_a001_0-100	SP140	ACACTCTCAGAGATGCCTTTTT	AGAAGCTCTTGTGCAATCCA	308
C11orf30_e06_a001_0-100	C11orf30	GCCTGGACAGTTTTGGTGAA	TCCAGGAAGGCAGGTCTAATC	287
C1orf106_e07_a001_0-100	C1orf106	TGGCCTTACCAGGTTTCTTT	ATAGCTCCCACCGACTCCCT	300
CARD9_e09_a001_0-100	CARD9	AGGACCCCAGGTGAAGGAAG	CTAGCAGGTTGTTCTGGGCAT	245
GCKR_e02_a001_0-100	GCKR	CTTCCTGCTCCATCCTTGTC	CGATATTAGGAGCCATGGTGA	280
GCKR_e16_a001_0-100	GCKR	GCCTCTCCTGCTCCTCTTTC	CTGGTAACCCATGACCTTGC	263
KIF21B_e24_a002_0-100	KIF21B	ACCTCAGAAGGAATGGGCTG	AGGATCCTGAACTCTCTCTGC	290
LRRK2_e25_a001_0-100	LRRK2	GCTGTTCTTTGAAAGCAAATTGT	TAGAGGGCCTGGTAAGGAGG	282
LRRK2_e28_a001_0-100	LRRK2	CCAACAGGTTTTGCCCTTTT	TCCATCAAAGTCACAGAGAGTAGA	287
PTPN22_e05_a001_0-100	PTPN22	CTCTGAGAATCACGTGGCAT	TTTCTGAGTCTGGGATCCAT	273
PTPN22_e07_a001_0-100	PTPN22	CAAAGACCCTACCACTTCCCA	TGGCCCTTTCTCTGTATCCTGT	301
SP140_e14_a001_0-100	SP140	GCAGCACTGGAAACTCTTCTG	TCTCCTTTCACTGCTCCTCC	274
C11orf30_e04_a001_0-100	C11orf30	GCAATCTTGTCCCCCTTCTC	TCAGGAATCTGCTACACAATAGG	274
C11orf30_e16_a001_0-100	C11orf30	GTGGACAACTCTGGTTTTGGT	TCCCATGAAATCCAGGAAGG	286
C1orf106_e09_a001_52-100	C1orf106	TACGTGGTGGTGGCTGAGAG	CCAACAAAACAGCCCTCCTG	389
CARD9_e07_a001_0-100	CARD9	aaagattaaaaaTCCTCTGACAACTGC	GTATACCCCAAGGGCCAGAAT	289
GCKR_e18_a001_0-100	GCKR	TTCTTAGTTCCTCTGGCCTTTAG	TGGAGGTCCTTGTGAGCAGA	308
IRF8_e04_a001_0-100	IRF8	CCCAAACCAATGAAGACACTC	ACATGAGACTCCTCCTGGCA	273
KIF21B_e05_a001_0-100	KIF21B	TCAATCTGGTTTGCTATTTCTGA	CCAAGCCCTCACTTCTGCTT	266

KIF21B_e10_a001_0-100	KIF21B	TCTCTAGAAGCTTAGTCCTGCAC	gcgcTGGCCTATTTATCTTCT	244
KIF21B_e21_a002_0-100	KIF21B	CATCTGTGCCCATAGAGCCC	GAGACCAAGGTATCCGCAGG	251
MUC19_e14_a001_0-100	MUC19	GGAACAAGCTATCCCTGAAAA	GTACCATTGCTGCAGTTAGCC	258
P4HA2_e04_a001_0-100	P4HA2	GATGGCTTCCTTGTTCCTTG	CTTATCCCAGGGGACTTGCT	275
PRDM1_e05_a002_0-11	PRDM1	TCCCGAACATGAAAAGACGA	CCGTCAATGAAGTGGTGAAG	282
PRDM1_e06_a003_16-45	PRDM1	AAAAGGACCTCGATGACTTTAGAAGA	GGTAGGAGCCCAAACCTTCC	339
PTGER4_e03_cds_a002_29-80	PTGER4	ATCAAATGCCTCTTCTGCCG	CCAGTAAGACACTCTCTGAGTCC	331
PTPN22_e09_a001_0-100	PTPN22	AATGGGAAGTGCCTGCTGAG	TTGCCTAGGACCTCTTGCAT	273
SBNO2_e08_a001_0-100	SBNO2	AAAAGGGTTTGAACCTGTCCC	CTGGCCTAGAAGCTGGGAGT	360
SBNO2_e20_a001_0-100	SBNO2	AGACGCAGAGTGACCAGCAG	GAAGGCAGCCAACATGTCCT	325
SP140_e05_a001_0-100	SP140	CTGAGGTCTCTCCTCTGGTCA	GCAGGTACATGGCTGTGTGA	283
SP140_e27_a001_0-100	SP140	TGGGAATCCACAGAAGGTAGA	TTTCTTGGCACCAGAAAATG	279
SP140_e25_a001_0-100	SP140	AGACCCCATGTGTGAATCTTG	GAAAGGAGAAGGCTTGGGGTT	301
SP140_e22_a001_0-100	SP140	GACAGGGAGATGGTTTCTCATT	ACCCAATTTCAGACCCCATA	284
SP140_e18_a001_0-100	SP140	AGTCGACACTCCCAAAAGCA	ATTGACCTGTAGGGCAAAGG	276
SP140_e13_a001_0-100	SP140	GGACAGCCCTACCTGAGAGAT	GCTCCAAATTCAACCTCCTG	260
SP140_e11_a001_0-100	SP140	TGAACCCACCTACAGAACCC	CCAAATTCAACCTCCTGACA	279
SP140_e06_a001_0-100	SP140	CCTCAGGTCAGGTTGTTTGG	GAATTCCACCTCCTGACACC	275
SP110_e18_a001_0-100	SP110	GCATTTTATGCAAGAGAACCG	AAACAAAATCCCACACACCC	261
SP110_e04_a001_0-100	SP110	ATATACAGGTGTTGGATGGGAT	GGTACTGATGTGTGTCCCTTCA	332
SP110_e02_cds_a001_0-100	SP110	AATCAGGCATATTGGTGGGG	cccagccTTTGCTTCTATTTTT	245
SBNO2_e23_a002_0-100	SBNO2	AAAGCTTTGGAGAGCCTTCCT	CATCCAGCAGTTCGGTGAGT	289
SBNO2_e19_a001_0-100	SBNO2	TGGAATCCTGACCCACAGGT	GTCTCCTCCACCTGCTCCTC	305
SBNO2_e12_a001_0-100	SBNO2	GTAGTGAGCTCCAGTGGCCT	CCATTCCACACCTGTTCTGT	350
SBNO2_e05_cds_a001_0-100	SBNO2	ACTGCCAGAAGAGACGGACG	GACTAAACCGACATGCAAATG	368
PTPN22_e13_a002_76-100	PTPN22	TCTGCCATTTCATGGTTGTG	ATTGTGATGCTAGGAAGCGG	269
PTPN22_e12_a001_0-100	PTPN22	CATTTTTGCTGCTTTTGTGGT	GCTTCACTGAAAACCAGCCT	275
PRDM1_e08_cds_a002_35-85	PRDM1	CATCTCTGTAGCCTCAAGGTTC	CATGAGGGGTAGATCTGATGAC	309
TNFSF4_e03_cds_a001_0-100	TNFSF4	GGTGCCTGGTTTTAGATATTGC	AGGAATTAATTAAACAAATTTTCCCCT	420
TNFSF4_e01_cds_a001_0-100	TNFSF4	ATGAGCTGGTGGGAAAACAG	CTTCTTTCTGGGGAAAACTTCA	274
SP110_e14_a002_0-100	SP110	CCTGTCACCTATCTGTCCCG	AGTCTCTCTTGACCTGGAAAGG	281
GCKR_e14_a002_0-100	GCKR	tttagtcctcaagtctgtgcttTA	CCCAGAAAAACAACAGCTCC	242
C11orf30_e09_a002_0-100	C11orf30	TTTGCTTTCCTATCAAATTCTCTT	accgtgttcatggactggaag	381

C11orf30_e08_a003_0-100	C11orf30	GTTGTATATTCTCTGTTGAATTGCTT	ACCTTCATATCCAAAGCAGAGAAA	505
C11orf30_e13_a003_0-100	C11orf30	TTTTCATCATCCTGGTAAAACC	TTTCCTTCAAATTTCTACCCG	355
C11orf30_e15_a003_0-100	C11orf30	GGGGAGATAAAGCATGAAGTTTACAAT	CTCTAGAGAAAACATGCTTCATTCA	340
C11orf30_e20_a006_0-55	C11orf30	CTGGGGTGTTTGATTAGATTTT	AATGATCCTGCTTCACCAGT	356
CREM_e04_a002_0-100	CREM	GTCAGAAATGATTATTAGAAAGCTAGG	TTGACTTTCAACGGTGCCTG	589
CREM_e09_a002_0-100	CREM	GCAATGGTAATAAATAGACCCAGAGTA	AAAAACTATAAAATCACCACTCAATG	368
CREM_e13_a002_0-100	CREM	CAAAAGTATATCATTTCCAGATCAGT	AAAAAGTGTCCTAGAAGTTACCGA	290
CUL2_e05_a002_0-100	CUL2	GCCAAATCAGTGTCCTTAAAAA	TTTGCTTTCCTACTGACTTTAACTT	257
CUL2_e11_a002_0-100	CUL2	CCATGTCTACATCAGAAACCTCATAC	TCCTTGAACCTCTTAAAATGTGG	261
CUL2_e12_a002_0-100	CUL2	GCATTCCTTACATAAATGCCAA	GGCTGGTGGTTAATTATGGC	255
CUL2_e13_a003_0-100	CUL2	GCCAGCATTCTTGCGTAGAAC	AGAAGCACCCTTTGTAAGATTTATT	295
CUL2_e17_a002_0-100	CUL2	GGTTGATGGAGGAGAAAAGGTT	TGAATGGTTTTCATTGCTTTTG	277
CUL2_e19_a002_0-100	CUL2	CCAATGCTTAAATGTTTCCC	AGTTATGAGAGTGGAGGCTTTTT	272
FASLG_e04_cds_a002_0-100	FASLG	GGCCCACAGTTTTGCCTTAG	TGCCTGTAACAAAGAATCATAATGG	524
ICOSLG_e06_a003_0-100	ICOSLG	TTTTTGAAACTCTGAAAGACAGC	TCCAAGTAGGGTGTTTAGGGG	220
IL18R1_e04_a003_0-100	IL18R1	GGAATTGCTACTTTCCATTCTTTTT	TGGACTCTCAATTCCTACACAGTTT	373
LRRK2_e05_a003_0-100	LRRK2	GGGTCTACAAACCATTCACAGTCT	TTATAACTGTGTTCTACTTTTCCCAGT	325
LRRK2_e09_a004_0-100	LRRK2	GGCTTCTCCTAAAGCACACCTC	AGGAGATTATTTAGTGCCCAGCAT	387
LRRK2_e14_a003_0-100	LRRK2	AATTTCTTCACCATCGTAATTTTT	TGACTTCCTCCTATCATAGCAAA	498
LRRK2_e21_a003_0-100	LRRK2	AGATTTTACAAAGGGAATGGACT	ACACAACATAATCACAATTGCAC	298
LRRK2_e24_a003_0-100	LRRK2	GGTGTGTAAGGCAGAAATATTAGC	TCATTGCTATAAAAATGTCAGCATA	414
LRRK2_e29_a003_0-100	LRRK2	TCTCTAGCAAAATATATGCATTAACAG	AGCTCAGTTTTTGAATCCCTG	481
LRRK2_e42_a003_0-100	LRRK2	GCCTAAGTGTATGCCTCCTTGG	TTTAAGGAATTAAGCATACAACTACAA	339
LRRK2_e44_a003_0-100	LRRK2	GGGTTCCAGTTTAACAGTTTTAGGTTT	TGAATAATAATCCTACATTGGCTCTTA	424
LRRK2_e45_a003_0-100	LRRK2	AAGTAATGCAAGAAAGCAAAAAG	TGCACACAATTCAGTCGTTTA	374
LRRK2_e46_a003_0-100	LRRK2	ggaggagaacattaaggcca	tgattcttgctctcagggagtt	253
LRRK2_e47_a003_0-100	LRRK2	TGGAGTTTTGTTCTGTGGATACT	ACCCTTCTGCTTTCTGGAATTTT	426
MUC19_e33_a003_0-100	MUC19	CCTTGACTCCTTGTTGAAAAAT	AGCTTACATTTCTCTCTATGCAGAAC	300
PRDM1_e03_a003_0-100	PRDM1	AGCTTGTTAATTATGTAGGCATCA	AGCCACCCAGCTCTTTTAGC	349
PRDM1_e06_a007_0-20	PRDM1	GGCAGTTTTGCTTCAGTTCTCTC	TAAACGACCCGAGGGTAGAA	293

PTPN22_e14_a003_0-100	PTPN22	ctcaaggctcacacatcagc	TGAATGAACAAGTGTCAACTTTACTG	336
PTPN22_e18_a003_0-100	PTPN22	CGGCATGTTTCCCAAAACTC	AGAATACATTTTAAAAGCCCCTAAT	285
SBNO2_e17_a003_0-100	SBNO2	CATGTCCCAAACCTTCCAGG	AGCTCACCATCCCGAGACAG	526
SP110_e03_a003_0-100	SP110	GTGCATATTCCACAGGGCTA	GCTTTCTTAAAACTTAGAAAACCATAA	344
SP140_e15_a003_0-100	SP140	AAAATAGAGTCACTGGGTCAAACTG	GTCAAGCAAATGCGGAGAAT	290
SP140_e20_a003_0-100	SP140	TGAAACTCTAGAGGGTTTTGGAA	CCTTCAGATAGGAGCTGCAAA	277
TNFSF4_e02_a002_0-100	TNFSF4	CCTTCCCTTAGGAAAAGAAGAG	AGGCTACTGCATTTCTGCTTCA	255
TRIB1_e02_a002_0-100	TRIB1	AACTTCTGTTCAGCTCCCTCAG	CCAACACACAGCACCTAACTAC	459
CDKAL1_e12_a001_43-100	CDKAL1	CAAGAAACAGTGAAACTTGTTGA	TGAGCAAGTTGTGTGAAGGTAG	200
CrohnsManual_MAST3_e21_a002_45-100	MAST3	ACAATGCCCAAGTTTGCCTT	CTCTCTTAAAGCCTTTGTTGGC	266
CrohnsManual_MAST3_e13_a002_53-100	MAST3	TTTGTGGAGCGTGACATTCT	GAACAGACCTCAACTTGGTCTAAAAAT	201
CrohnsManual_MAST3_e20_a001_0-100	MAST3	cggcTTACCATTCTTTTGTC	GAAGCTGGATCTGCCTACACTTC	202
CrohnsManual_MAST3_e12_a001_0-100	MAST3	ACAGGGGAGGAGGTAGCATC	GAAATGGGTCTATGCAGGACT	230
CrohnsManual_MAST3_e07_a001_0-100	MAST3	TCAGGTCCCATATCACAAGCC	GACATAGCCAGGAAGACAGAGTTC	201
CrohnsManual_MAST3_e05_a001_0-100	MAST3	TTCCTTCCATCTGTTCTTGTTC	cacaaCTATCCAACACCTTGAA	207
CrohnsManual_MAST3_e02_a001_0-100	MAST3	TAGTCTCTCTGAGAATGATCCCCAT	GGCAAGAATGAACATGAGATG	201
Manual_UBE2L3_e05_a003_3-100	UBE2L3	TGGCAGTAATCCAGTCCCTC	ACGGGGTCTCTGCTCACACT	205
Manual_UBE2L3_e05_a002_0-99	UBE2L3	GCCTGTGCCATTTCTGAGACT	TGGCAGATTTTAGTCCACAGGTC	221
Manual_UBE2L3_e04_a002_0-100	UBE2L3	CCAGCGCTCTTCCTTTGTCT	GCATCTTTTCATATTACTATTTCTTCC	212
Manual_UBE2L3_e03_a003_51-100	UBE2L3	TTTAAAACAAAGATCTATCACCCAA	GTACCTGAAAGATTCTCTGGATTC	211
Manual_UBE2L3_e03_a002_0-82	UBE2L3	CCTGTCTTGGCCATAGGTTTAT	TTCCAGTTTTCGGCACTAAT	226
CrohnsManual_C10orf22_e01_a001_0-19	C10orf22	TATTTTCCAATCAACAAACTGAACTAT	CTTCAGCTTGCTCAGGTTCTC	589
CrohnsManual_ZGPAT_e04_a001_0-100	ZGPAT	CACCATGCACAATCCTCTGG	GACCCTGGGAATCTGAGGAG	438
CrohnsManual_ZGPAT_e05_a00	ZGPAT	GGAATGTGTTTGACTTCCTCAATG	CTCCTGGATGCTCCTGATGTC	200

3_48-92				
CrohnsManual_ZGPAT_e05_a001_0-46	ZGPAT	CTCCTCAGATTCCCAGGGTC	AAACACATTCCGAGGAGCTG	347
CDKAL1_e05_a001_0-100	CDKAL1	GCCACTGGATATTTTTGAATGATTAAG	TCTTTTGGTATCAATTCTTACTTTGAT	213
CDKAL1_e13_a001_0-100	CDKAL1	CTGCTGTCAACTCAATTGTTTGTATGT	ACCATATGTTGGTGAAAATTGTACT	200
CDKAL1_e04_a001_0-100	CDKAL1	aaaaTCACTCAATGAAACTTACTTTTT	ACCCTTTAACTGCTAGGCTATTTAT	246
CDKAL1_e10_a001_0-100	CDKAL1	GCCACAGATTTGTTAATTATCTCTT	TTAGCATGCAAATAGGGTGC	233
CDKAL1_e16_cds_a002_0-64	CDKAL1	GATTGATATCACAACAGGGCTT	ACATCCTCAGCGCACAGTCTT	201
CDKAL1_e15_a001_0-100	CDKAL1	tgcaataattctataaatctgaaactg	AGCTGCTGTTTTACTAGCAGGGTAGA	297
CDKAL1_e07_a001_0-100	CDKAL1	AATGTGTAACATTGACTCAAGCAT	AAAGGTGCCTGGCAAGAGTT	229
CDKAL1_e11_a001_0-100	CDKAL1	AAATGATTAGTGTTTTCTCCTTCC	TGCCACAAGCTTTTGAAAGAAA	241
CDKAL1_e16_cds_a001_43-100	CDKAL1	GCGAGTTCCAGAATGGTGCT	TGGAGATGTCGCTTTCCTGTT	200
JAK2_e23_a001_0-100	JAK2	GATGGCCCTTAGTGTTCATTTAATTT	TGTTCTGTAAATCTACTTTGGTCTCAG	200
JAK2_e19_a001_0-100	JAK2	GGCAGAGTAAAACATTATTTCCACCTT	ACGGTTGCTTCATCTACAGC	205
JAK2_e12_a001_0-100	JAK2	GGCAGAGAGAATTTTCTGAACTATTTA	TGCTCTGAGAAAGGCATTAGAAAG	203
JAK2_e08_a001_0-100	JAK2	TGTTTTAGATGACACTTGGTCATAAT	TGGAGAGGAAATTAAATGAACTAAATA	246
JAK2_e06_a001_0-100	JAK2	GGCTCTGTAAATTCTACCCGTTTTTA	TTTCTGAGAACTATAAAATGAAACCAT	220
JAK2_e03_a001_0-100	JAK2	GGAAGCTGACCAAATGTTTTTAT	TGTAAAGTTTTTCCTGATTAAGTGAC	238
JAK2_e02_a001_0-100	JAK2	ACCTTTAATAATTCCTTTCTCTGCT	TGCCTTTTAGCATTAAGTGAGTTAC	201
LOC441108_e03_a001_0-100	LOC441108	TGTCTGAATCTAGAATAACAGGAAGA	GTGGCAGCACACCTCACAAC	570
CrohnsManual_ZNF365_e02_a001_0-86	ZNF365	gtgaatTCCTTGCCTTATATTTCC	ATCACGGACAAAGCCAGAGG	220
CrohnsManual_IL18RAP_e03_a002_29-100	IL18RAP	GTTAAGCCCCAGACAAATGC	AGCCCATGAAGAAGATACCATAAAAAC	205
CrohnsManual_IL18RAP_e03_a001_0-82	IL18RAP	CTAGACAAAATATCTATGTTCACAGGA	TTTGTGCATCACTTTGGCAG	205
CrohnsManual_IL18RAP_e07_a001_0-100	IL18RAP	TGTTACTGAATGGAGTGGTATTTTT	TCTGGGCTTACCTCCTCTCTTT	201
CrohnsManual_IL18RAP_e04_a001_0-100	IL18RAP	GTGGCAAATTTATCTGCTTAAAATATC	AAATCTCACTTACCAATGGTTCTC	202
CrohnsManual_ZNF365_e08_a00	ZNF365	TGGAAGAAAGAGCATTCCATC	ACCCCCTCAACCCTACTCTCAG	200

1_0-100				
CrohnsManual_ZNF365_e06_a001_0-100	ZNF365	GGCAGCACATAAGTGAAGAGTT	ATCCAAGTCCTCATCCAGAATCA	200
CrohnsManual_NCF4_e04_a001_0-100	NCF4	AGGCTCCTAGGACAGCTCTTTGT	CCCAGTTCCCAAAGCCATAC	200
CrohnsManual_ZNF365_e12_a001_0-100	ZNF365	CACAGGCTCTGTGTTCATCTC	tgacaGTAGGGTAGAGGAAGTGGAC	201
CrohnsManual_IL18RAP_e01_a001_0-100	IL18RAP	CTCAGGGCAGAGTTCTGAAT	TGACCACAGACTCACAGTGCTAAC	201
CrohnsManual_NCF4_e06_a001_0-100	NCF4	ATTTCAGCATCTTCTGTCTCCT	CAGATGCCCACAGTGTGAGAG	231
CrohnsManual_IL18RAP_e06_a001_0-100	IL18RAP	CACCAGCTTCCTTCTTGTTTTCT	TGCAGAGAACAGAATCCGTGAG	200
CrohnsManual_MAST3_e26_a001_0-100	MAST3	GTTCTGAAGCCTCCGGTTTT	CTACATCGTGTCATTCCATCCC	316
CrohnsManual_MAST3_e22_a001_0-100	MAST3	ATGGTGAGTTTGAGGGGCAC	CAGACAAAGGTGAGAGGGGC	305
CrohnsManual_MAST3_e14_a001_0-100	MAST3	gccTCATGACTACATCCGCC	CTCTGCCCTGATCCTCTTTG	254
CrohnsManual_MAST3_e10_a001_0-100	MAST3	GTGCTATTGGGCTAGCGTGG	CCTTGGTTTCCTCTTCCCAT	318
Manual_IL12B_e03_a001_0-82	IL12B	ttCCTTATGGAGCACATATAATCATC	GGAGCTGCTACACTCTCTGC	223
Manual_IL12B_e02_a001_0-100	IL12B	CCAAGGAATATACTGCACCTGAAT	AGCCTTATTAGGAAAATAAAGCATGT	234
Manual_IL12B_e06_a001_0-100	IL12B	GGGGCCTCCACAGAGAATAAT	CAAAGGTAACAAGAATTATTGACATTC	223
Manual_IL12B_e04_a001_0-100	IL12B	ATGCTGAGAAACCAGAGCAG	ACAATCAAAACCTTTCTGCAA	210
Manual_ATG16L1_e05_a002_38-100	ATG16L1	CCAGACCCTGAAGGATGAATATG	AGGATGGAAACCACTAACTGTG	201
Manual_ATG16L1_e05_a001_0-62	ATG16L1	AGGTACTATTCGTCCTCTGATGTC	TTCCGTAGTTTTCCTCAGTTTTCC	200
Manual_ATG16L1_e17_a001_0-100	ATG16L1	AATTGACAACCTTTCTTTTCTTACTG	CCTGAGATGAAGTCAAAGGC	202
Manual_ATG16L1_e12_a001_0-100	ATG16L1	GGCTAAAAATTGGTTTTCCTCTTC	AAAGTCCCCTAATTAAGCACCTG	200
Manual_ATG16L1_e06_a001_0-100	ATG16L1	GCTTCTAGAGAGCCCAAACCC	ACTTTTAAGAGAGCCAAATTTTCAG	200
Manual_ATG16L1_e15_a001_0-	ATG16L1	GGCCTACGTTACATTTCTCAGAT	TCTTCATGCAACCAGCACAT	219

100				
Manual_ATG16L1_e16_a001_0-100	ATG16L1	CTGAGCTCCCTCAGTGACAGT	GAAACATAGGTTGGGCACTGAAC	200
Manual_ATG16L1_e14_a001_0-100	ATG16L1	GGGTGGCAGCATTATGTAAACAG	GTGCCTCCAGTCAGGAGCTT	207
Manual_ATG16L1_e09_a001_0-100	ATG16L1	TGAGCAGTAAACCTCTGCAATCC	CCATATCAAGCGTGGTAGGGTT	200
Manual_ATG16L1_e13_a001_0-100	ATG16L1	TGTGGATACTTTGCCAGCAT	CATCTAAGCATATACACAGACAGAGAG	234
Manual_ATG16L1_e08_a001_0-100	ATG16L1	AGTGGGAAACATATTCCCAGTTAC	CTTATACACATTTAAAAAGGAAAGGTC	200
Manual_ATG16L1_e04_a001_0-100	ATG16L1	CGCCACCCACTGTGTAGTTTC	AGCAGAACAAAGGACATGGGA	209
Manual_ATG16L1_e02_a001_0-100	ATG16L1	GGTACCTAGCAAGTGTACATACGTT	TGAGATAAATTTCCTTAGAGGAGACA	213
Manual_NOD2_e04_a014_91-100	NOD2	CACCTCAAGTTGACATTTTGC	CATACCTGAACAGCAATGCCC	200
Manual_NOD2_e04_a011_69-84	NOD2	ACACCTCTTCAATTGTGGCA	GGATGGAGTGGAAGTGCTTG	287
Manual_NOD2_e04_a005_27-37	NOD2	TTAACCTTTGATGGCTTTGACGA	GAACTCGGTGCGGATGTACT	201
Manual_NOD2_e04_a004_18-31	NOD2	AATTTCTCTTTGTCTTCCCATTCAG	GTCTGGACAGAGGTGGGGTC	244
Manual_NOD2_e04_a002_6-18	NOD2	TGGAGGACATATACACAGAGAATG	CAAAGAGAAATTCCTGGAAGTCTTG	240
Manual_NOD2_e11_a001_0-100	NOD2	CCTTGAAGCTCACCATTGTATC	TCAGATCCTTCACATGCAGA	268
Manual_NOD2_e09_a001_0-100	NOD2	AGAGCACCACGAATTTTGCC	CTTCCCCAGAGCAGAGAATCC	206
Manual_NOD2_e08_a001_0-100	NOD2	AGGGAGGAGGACTGTTAGTTCATGT	CAAGAAAACTGCAGGATAGACTC	200
Manual_NOD2_e06_a001_0-100	NOD2	GCTTTGCTTCTGTGTCTCCTCTCT	AGGCCAAGACCAGATCAGACT	200
Manual_NOD2_e05_a001_0-100	NOD2	CTTGTCTTACTAGCTCCATTTTCA	AGCCCATCACTCACAGCTTC	221
Manual_NOD2_e03_a001_0-100	NOD2	CCTTCCCACATTGCTCCATC	TGCCCTTCCCTTTCTGATTAACT	205
Manual_IL23R_e10_a004_71-100	IL23R	AACTCAGTAGAGGAGGAAACCAC	TGCAAGGCAGCTTTCTCATATT	238
Manual_IL23R_e10_a003_46-78	IL23R	CCACCAGTTGATTCCTTAGACT	TCTGGAATAGTTTCACTGGGTG	227
Manual_IL23R_e10_a002_21-54	IL23R	GGACCCCTGGAGACAAGAGAC	ACAGAAAAAGCAAAATTAGGATGC	236
Manual_IL23R_e10_a001_0-24	IL23R	CCAGTTGGTTCTTTTAAATGTCTTTTG	AGCGAGTTTTGCGGGTAG	220
Manual_IL23R_e06_a001_0-100	IL23R	CACCACATTTTATTATTGTTACCCA	agcAGAAAAGATATATAAAGAAAAGCC	249
Manual_IL23R_e04_a001_0-100	IL23R	cctggccAATTAATTCAAACTAAATAC	AGGTCAGCACAAACTGGTTATTAT	274
Manual_IL23R_e05_a001_0-100	IL23R	GGCAAGTTTTAAACAGCCAGGT	GGCTTTTAAATCATTCATTCACAT	246
Manual_IL23R_e02_a002_37-100	IL23R	GAACTGCCAACCAAGGAAAC	TGCCCCAACACTTACATCCA	212

Manual_IL23R_e09_a001_0-100	IL23R	CCTAATCTCCTATATGATTGCCTGCT	AGACTGCACTCTTGGTCTCAAT	203
Manual_IL23R_e07_a001_0-100	IL23R	GGAAGAAACTCCGTTGGGAA	AAGACTACTCAAAAAGCCTAAGTTGT	203
Manual_IL23R_e03_a001_0-100	IL23R	ACAGCACCTCCTAAGTGTTATGT	ACTGGGGTGGAACTGCTTAT	202
Manual_IL23R_e01_a001_0-100	IL23R	GGGAAAAATGTTATGCTTTTTATTAT	TGTTGAATGAAAGATAGCAATAGATAC	213
Manual_PTPN2_e02_a001_0-100	PTPN2	CACTGTCAGTTACTAGTGCAGAAGC	TGCACTTCCTTTCACTTGTTTTTAT	289
STAT3_e08_a001_0-100	STAT3	TGAGGGAAAGGGACAAGGAT	TCACTTTGGTAATTAGCATCTTTCTTA	228
STAT3_e07_a001_0-100	STAT3	TTTCATCCTTTACCAGTTTTCTAGC	AAATGTTTTCTGACTTTGTTTTGG	200
CrohnsManual_ZGPAT_e01_a005_86-100	ZGPAT	GTGTCCGTGTGCTTTACCTG	GAAGCTGAGACAAAACTGCAAAG	201
CrohnsManual_ZGPAT_e01_a004_67-98	ZGPAT	aagagCTGAGTGGGACAAAGGT	GTTCTCCTTAAAGCGGCACT	200
CrohnsManual_ZGPAT_e01_a003_40-72	ZGPAT	AAGATGCTGAGTACCAGGCTTTC	CCAGGAGCTGTAGTAGGGCG	203
CrohnsManual_EGR2_e01_a006_49-65	EGR2	AGGGTAAAGTTACGGATTGTAGAGAGT	CTACCCACCACCTCCTTCCTATC	216
CDKAL1_e12_a002_0-86	CDKAL1	GTCAAGTTTTTACATTTTTGTCTTGT	TGTTCCATTTTTGCAGCAGG	201
CDKAL1_e08_a001_0-100	CDKAL1	GTGAAGTGTGTAACTCTTGCTAATGT	TTGAATATTTTATGAGCAAGTTTGAG	206
CDKAL1_e06_a001_0-100	CDKAL1	GCACTAACAACAGTTTCCAAGAGTAT	ATTTTCTCTTTTTGCATCAGGT	201
CDKAL1_e14_a001_0-100	CDKAL1	TCCTTTCTTTCCCTCCCCTT	ACTCTCCCCTTTAAATGCCCTT	201
CDKAL1_e09_a001_0-100	CDKAL1	GCACTTTTGTGTATGTTTTGAGGTATC	TCTTCAGAAAACTAAAGAAGTTGGT	300
Manual_SLC22A4_e05_a001_0-100	SLC22A4	GTATCCAGCCCTGCTGTTGTG	AGCTGATCTCAAGAGGAAACATCTG	216
Manual_SLC22A4_e03_a001_0-100	SLC22A4	CACATCTCATGTTTTGTGTTATACTG	TAGAGGGAAGCACCCCATGT	229
Manual_PTPN2_e06_a001_0-100	PTPN2	AGTCCACAATTTCGGGCAAG	TGGTTTCATGATGTCTATAAACTAAAA	222
Manual_PTPN2_e09_a001_0-100	PTPN2	AAATACACATGCACACAAACCCA	TACGCTGGCTGGGAAGATAAG	200
Manual_PTPN2_e01_a001_0-100	PTPN2	CTGTGAGGCAATCTAGAGGGTT	TCGTCTGTGTGTTTGGTAAATATG	201
Manual_IFI30_e01_a001_0-100	IFI30	TTAAAGGCGCTTATTTCCCAG	GAAGTCTGGGGAGAAAGAAGGAG	274
Manual_UBE2L3_e04_a001_0-100	UBE2L3	GCCACCCTCCTGCTGTAATCT	TCCATTATGCTTACTTTTCTGTACTT	210
Manual_UBE2L3_e02_a001_0-100	UBE2L3	GGGCATCAGGAGGGATCTTA	TCTGGTAGGAAGTGGAAGTGTT	201
Manual_C13orf31_e04_a002_31-100	C13orf31	TATGGCTGCAGTTTGGAAGACA	ATCAAACTTGCAGTTGAATGAA	200
Manual_C13orf31_e04_a001_0-	C13orf31	GGCTCATTTAAAAACCACATGATAA	TGCTGATTCCCTTGGAAGAGT	297

59				
Manual_C13orf31_e01_a003_47-82	C13orf31	GGCTGCCACTTTGTATACCATTAAA	TTCTTGAGCTGTGATTACGTTTAT	217
Manual_C13orf31_e01_a002_21-57	C13orf31	GCCAAGGCCAAGTTTCTCTGTAT	TGCCTGGGTACAATTACCTTA	221
Manual_C13orf31_e01_a001_0-30	C13orf31	GCAGGTGATTTATTTGGCATAAAAGTA	TTGTTCTCCATCCCTTTCAT	212
Manual_TNFSF15_e03_a001_0-100	TNFSF15	CTCATCTCTGAACTTCCTGCATAC	TGCTATGGGAAGCTGTAAGTGTT	201
Manual_TNFSF15_e02_a001_0-100	TNFSF15	CCCATGGTCTCCCGTAAAAC	AGCTGGAGGAGATCAAGGTTTCT	201
Manual_ZNF300_e01_a012_89-100	ZNF300	CCTGCCTGAAGAGTTTGTCTTG	TGACTCTTGTGATGACTCTTAAATTGT	221
Manual_ZNF300_e01_a011_81-94	ZNF300	GCATCATATTTATGGAATCTCTGTATT	AGGGATGGTTCATTGTGCTC	218
Manual_ZNF300_e01_a010_74-86	ZNF300	GGCTCACTCTGGCTAGATGATTTT	AGGCATCAGGGCATAAATATAA	208
Manual_ZNF300_e01_a009_66-78	ZNF300	ACACAGCTTTTCTCTTTAGTTTCC	ACCGAGTTGTTATAAGAGCAATTCAAG	210
Manual_ZNF300_e01_a008_57-69	ZNF300	CAACAAGATGAAACTTCTCACTGAA	TGGAAACGTTTTTAGAAATACACAATC	203
Manual_ZNF300_e01_a007_48-61	ZNF300	CCCGCATTCACTACATTCATAGG	AGAATTCATACTGGAAAGAAACCATA	213
Manual_ZNF300_e01_a006_40-54	ZNF300	GTGGTGTATAATCAGCTGTGACTT	TCTGAATGTGGAAAAGCCTTC	219
Manual_ZNF300_e01_a005_32-44	ZNF300	CCTCACATTGAGCACATTTGTA	GGGAAAAACCCTATGAATGTAGAGAG	201
Manual_ZNF300_e01_a004_24-36	ZNF300	TGTGTTCTCTGATGTATGATGAGC	TGTGAGAAGTCCCACCTCATT	206
Manual_ZNF300_e01_a003_14-27	ZNF300	CGGAAGGTGGGACTTCTGAG	ACCTTATGAATGTACTGAATGTGGAAA	223
Manual_ZNF300_e01_a002_7-20	ZNF300	CCCCAGTATGAATCCTCTGATGT	TGCCAGAAGTCACATCTCATTG	214
Manual_ZNF300_e01_a001_0-10	ZNF300	GCTTGAGCTAATACTAAGGCTTTTCTG	TGTGCTGAATGTGGAAAGGC	203
Manual_PDLIM4_e07_a001_0-92	PDLIM4	GACGCTGTCCTGTCTCGTTC	GAGTTCCACCTTGGCATTGG	229
Manual_PDLIM4_e03_a001_0-100	PDLIM4	TAGGAGACATATCTGACCATCACAAT	GAGGACATTCAGTGCAGGTG	210

Manual_NKX2-3_e02_a004_85-100	NKX2-3	TTTGTGAACGTGAGCAACCTA	CTGAGGAGCTAGACGTACATATTTT	248
Manual_NKX2-3_e01_a001_0-56	NKX2-3	GGATTTATTATTTGGACTGGACAAT	CAATTTCTCGCCCTCGTCTT	247
Manual_IFI30_e05_a001_0-100	IFI30	ATAACGGGGAGGAAGCTGAG	CTATGAGTGTCAAGCTGAGGGCTA	237
Manual_PTPN2_e04_a001_0-100	PTPN2	GCAGTTATTTGATGCTATGTGTATTT	TCCTGGGTTCCAATAACAAGA	266
Manual_UBE2L3_e01_a001_0-100	UBE2L3	GATGCATTCTGGGGAAGGAG	CTCACATTAACGACGCCCTC	387
Manual_ZNF300_e03_a001_0-100	ZNF300	GGCACTTGATTGGACAACTCTG	TCATTTCAGATTGAACAGGAATTT	208
Manual_ZNF300_e02_a001_0-100	ZNF300	CCTCAGGCACCAATCAATTAAA	TGCAGTCAATCTCAATTCTCTGTG	210
Manual_SLC22A5_e09_a002_0-100	SLC22A5	ATAAAGGGGTAGATGAGAGACCAAGTC	CAGACCCAGTGCATCAACACT	247
Manual_SLC22A5_e07_a001_0-100	SLC22A5	TTTTCCAGCTTTCTTCTGCAC	CCCCAAACCATAGATGCACA	276
Manual_SLC22A5_e08_a001_0-100	SLC22A5	CCTACTCCTACCCTCTTTCCTTT	GCTCATGGGACTTACCAAGG	231
Manual_SLC22A5_e03_a001_0-100	SLC22A5	GCTGGTTATCTGTCACTCTCCT	ATCAAGTACTCAACTCCAACCTG	221
Manual_SLC22A5_e10_a001_0-100	SLC22A5	CCTCAGTTCTTGTTTGTTTGGAGA	TCCTCTTCAGTTTCTCCCTTACT	200
Manual_SLC22A5_e06_a001_0-100	SLC22A5	GCTCTGAGTCTCTGACCACCTCTT	AGAAGCTTTGTCTGGAAGCCT	200
Manual_SLC22A5_e02_a001_0-100	SLC22A5	TTTTAAAAAGAAGTGAATGATACACCC	CACGCTTCTTCCTCAGTGCT	205
Manual_SLC22A4_e09_a001_0-100	SLC22A4	GATTGATGTTCTTATGTCCCG	AGGCATTTTGGTATCATTTTGTTAAA	205
Manual_ITLN2_e08_a001_0-100	ITLN2	AAATCTCTAAAACTGAGACCTGAGC	CTCTCCACAGCTCTCCACATAAA	201
Manual_ITLN2_e07_a001_0-100	ITLN2	GGAAACTGCAGCTCCTACTCTCA	TCCATTTGAGGATTCCTGACT	200
Manual_ITLN2_e06_a001_0-100	ITLN2	GGCTTCAAAGAGAGAATGCTGAGT	TGGCCTCAGTGGTCAGACTT	202
Manual_ITLN2_e03_a001_0-100	ITLN2	TTTATGACTGGACTTAGGGAGC	TGGCAGCATCTCTACTGTCCTC	203
Manual_ITLN2_e01_a001_0-100	ITLN2	GTTGGAAGATGGGTTCTCGC	CAAATCAAATTCCTCACAGTGTTTCTT	237
Manual_ITLN1_e06_a001_0-100	ITLN1	CTCCCAGAAGACACAATCCTGATA	TGTCAAACCGGGAATGAGAAT	201
Manual_ITLN1_e03_a001_0-100	ITLN1	TCCCTTTCCTACCAGGAGTT	TGTCTTTACTCCATGACTTTTTCTGTG	200
Manual_IRF1_e02_a001_0-100	IRF1	TAAGAAGCCATAAGGATCCAGG	GAGTTGATAGCCTCTTGCTCTT	200
Manual_IRF1_e01_a001_0-100	IRF1	TTGCCTAGAGGAATAAGAGGG	GTAGGGGGAAGCACACCTGAT	225

CrohnsManual_IL18RAP_e02_a002_38-100	IL18RAP	CTCTCACCAAAACAAGTCCCTGA	TGGAAATGTTTTCAGATAGAAATGTA	250
CrohnsManual_IL18RAP_e10_a003_57-100	IL18RAP	ACTTGGAGAGGCTTAAAATCAGT	TCCAGGGCTCATTTCACCAT	202
CrohnsManual_ZGPAT_e06_a001_0-100	ZGPAT	GGGTCTGCATGTTGGAGGAC	CACGCTTCGTCCATAGTGCT	278
CrohnsManual_NCF4_e09_a001_0-83	NCF4	CCCTTTGATTATCCCTGACTT	GTAGTTGTCCTTCTGCGTGATGT	204
JAK2_e16_a001_37-100	JAK2	GGGCAGAATTAGCAAACCTTAT	AGGGCCCAAATGACATCAAG	219
JAK2_e01_a001_0-100	JAK2	CTCTTACAGGCAAATGTTCTGA	TGCATAAAAAGGAAAATGCTGT	296
CrohnsManual_IRGM_e01_a002_20-55	IRGM	AGTTAACATCACTATGGCAGGG	GTTCTCCAGGGTTGTGGTGG	211
CrohnsManual_IRGM_e01_a001_0-24	IRGM	TTTTTGCCACACCATAAGCATT	ACATCCCATTGCCAGAGTCC	479
CrohnsManual_IL18RAP_e02_a001_0-57	IL18RAP	GCTAGAATTCCATTTTCTAATGTGTTT	AGGTTGTTGGTACCATTGGACA	250
CrohnsManual_NCF4_e08_a002_29-78	NCF4	CATCAGCACCATCAAGTCTGT	GTAAGGAGGCATCAGGCTGG	239
CrohnsManual_IL18RAP_e10_a002_25-69	IL18RAP	CTACAAGCAGCAGTGAATCTTGC	AGGCATGTGGTAGCGCATTT	201
CrohnsManual_IL18RAP_e10_a001_0-40	IL18RAP	TCCATTGCAAATAATCAAATGTT	TGGCTCTTGGAAGTAACAGAACTTAAT	209
CrohnsManual_ZNF365_e10_a002_50-100	ZNF365	TCAGATCTGGAAACCACCTCA	GTTCTGACACTTGAATAGGTCTGG	201
CrohnsManual_ZNF365_e10_a001_0-64	ZNF365	TGTTCCTGTTAACTGGGGTT	GTCACGGTTGCTTCCAGTTG	211
CrohnsManual_EGR2_e02_a001_0-100	EGR2	AATCCTCTCCTGCGATTCCC	GAGTTGGGTCTCCAGGTTGTG	291
CrohnsManual_ZNF365_e07_a001_0-100	ZNF365	GTGGTTGTATTACCTAGTTATTTTTGC	GCATGACCATGCACATCACTC	227
CrohnsManual_IL18RAP_e04_a002_0-100	IL18RAP	AAAGTTTTTCTGTGGCAAATTTAT	TTTGCTGCTCAGATCTCAAAT	244
CrohnsManual_IL18RAP_e09_a002_28-100	IL18RAP	GGAGCTCTTTTCCAAGTGAGG	AGACTTGGGGATTGCATTGAAC	200
CrohnsManual_IL18RAP_e09_a001_0-90	IL18RAP	TGATTCAAGCATATGTATGTGTACC	TGGAGCCACATCTCTTTCAA	250
JAK2_e21_a001_0-100	JAK2	GAGTCCACATATCAAGTAACTGTCTT	aaaatcaaaCGAACAAACAAAAA	220

JAK2_e20_a001_0-100	JAK2	GGGAATATATAGGGTTAAGACCATT	AAGTTTGAAGTCTGTGCTCTAAAAT	303
JAK2_e09_a001_0-100	JAK2	GTGACTATCCCTCCCTTTCTTTAT	ATAAACAATCCATGACCAGTAATTT	277
JAK2_e05_a002_32-88	JAK2	ATTCAGCAATTCAGCCAATGC	TGTTTCCCTCTTGACCACTGAAT	200
JAK2_e05_a001_0-53	JAK2	GTTTTGTTTTTCTGTATGTGCTTTT	AGGCAGACTGCAGAGTTTCCA	214
STAT3_e23_a001_0-100	STAT3	AAAAATCTGGAACCACAAAGTTAGTA	CAACTGAACTAGTTTTCCCTGTCT	210
STAT3_e22_a001_0-100	STAT3	TTATAGGGACAAAGTCTGTCAACC	CATTGTGTCTTGTCAACCCCC	202
STAT3_e21_a001_0-100	STAT3	CCCAGGGATAACTGAGGATATTAGA	TGAGTTTCAAATCAGTCCTGCC	203
STAT3_e17_a001_0-100	STAT3	CTTCAGGCAGGTCCTACTGG	ACCCCAAGCTGAAAATGTACTACT	219
STAT3_e15_a001_0-100	STAT3	GGGCACCAACTAAAAGGAGG	GCGGGTGAAGAGATTTCCAA	200
STAT3_e14_a001_0-100	STAT3	AATCATTCCACCTTCTCTTGTAAA	tggccTAAGAGTGTCCTTCTC	211
STAT3_e13_a001_0-100	STAT3	ACAGGTTGATGTTTCTAATTCTGGG	GAGAGGGCTAGGGAGAGCCT	200
STAT3_e11_a001_0-100	STAT3	GGGATCTGAATCACAGGGGA	AGGTGGTCAAAGTAGGCTTTT	204
STAT3_e10_a001_0-100	STAT3	AGAGTCTCTAGTTCAAATGATGTCTGT	AGGTTATGGGAGAGTTACTGACTTTTT	200
STAT3_e05_a001_0-100	STAT3	TCAACTCAACAACACAAACTCACTTTC	TTCTGTTCCCAAGGAAATCTT	200
STAT3_e20_a002_33-100	STAT3	CCTCCTTGGGAATGTCAGGATA	ACCTAGCTGTAGGTTCCATGATCT	200
STAT3_e19_a001_0-100	STAT3	CAACTAGAAGCAGTGATGAGGC	CAATAACAACATTGTTCCTCCTCCTT	240
STAT3_e16_a001_0-100	STAT3	GGAGGGAGAAGGGGTGAAAT	TCCCTCATCTAAACAAGCAAATGT	200
JAK2_e14_a001_0-100	JAK2	GGACTGATATTTGAATATATGTGCGTT	ACCACTGCCCAAGTAAAGCTTAGTA	232
JAK2_e11_a001_0-100	JAK2	CGTTCTCCATCTTTACTCATTCTTTT	TTAAACAGCATAAACTACATGAACAAT	227
Manual_UBE2L3_e01_a002_0-100	UBE2L3	tgGttgtaactgccatctgatt	GATCCGGAGCCTAGGACG	315
Manual_ITLN1_e05_a002_21-100	ITLN1	TGTTGTAGTTGGCCCAGTTG	GAACCCTCAGCTCTCAGACA	256
Manual_ITLN1_e05_a001_0-78	ITLN1	TCACCCGAGTGGGTAAGAAG	CTTCTGTGACATGACCTCTGGG	232
Manual_ITLN1_e02_a001_0-100	ITLN1	CAGCAATGAGCAAGAGAAGGT	TGAATGCTGTTTTCTGCCTC	278
Manual_ATG16L1_e10_a001_0-100	ATG16L1	ATGAGAATGACTGGGTTTGACA	CAAGGGTGCATCATCTCAAGTTT	200
Manual_ATG16L1_e07_a001_0-100	ATG16L1	CACAGAACTGTTCAGGAGAATCAA	TCCACAGATGTGCCTTCATATTG	201
Manual_UBE2L3_e05_a001_0-100	UBE2L3	CTGAGACTGTGTTAACCCCC	TCACACTTGCTGGAACCAATC	230
Manual_UBE2L3_e03_a001_0-100	UBE2L3	TGGTGTGTTCATTTTGATCTCT	TCTGGATTCTGTCTGATCTGG	314

Manual_C13orf31_e02_a002_49-100	C13orf31	GTAGTTCCAAACGGAGAGATCC	AAAAAGCAGATACAAAAGTTAAATAGC	201
Manual_C13orf31_e02_a001_0-87	C13orf31	GGTCTAATTGAGAGACTGGTATTTTGA	ATTTCTCCACATTAAATCCTGC	230
Manual_C13orf31_e03_a001_0-100	C13orf31	CCAAAACATTTTTGTTGCACATT	TGCACAACAAAATGAGTAAAATTC	249
Manual_TNFSF15_e01_a002_28-71	TNFSF15	TAGATGGGCTGGAACCAGTTG	CCAGAGTCGGGAGACTACTTC	215
Manual_SLC22A4_e01_a002_30-100	SLC22A4	TATGTCAGTCGTGTTCCTGG	GTATGAACAGCAGGTTTGAGG	412
Manual_SLC22A4_e01_a001_0-92	SLC22A4	TAGTTGCAAGTTTCGGAGCG	GTAGACGTCCTGGCTGAACTCC	404
Manual_SLC22A4_e08_a001_0-100	SLC22A4	GTATCACTTCTAAAAGCACCATTG	AGGAAGAACAAATGCACAAGA	261
Manual_SLC22A4_e07_a002_30-100	SLC22A4	GGAGATGCCTACCTGAACTGT	TGGGTCTTTTCTGGTAAGCTGTT	203
Manual_SLC22A4_e07_a001_0-71	SLC22A4	CACACCATCCCTTTGTCATTTTTAC	TGCAGCTATGATATAACGCCT	201
Manual_SLC22A4_e04_a001_0-100	SLC22A4	GTTAATATGCTAATACTCCTCCCTTG	TGGGGACGAGTCACACTCAC	228
Manual_SLC22A4_e10_a001_0-100	SLC22A4	GGATGGAGCATTTTGAGGAG	TCTCCACAGGGTCTTATTTTTCTGT	200
Manual_SLC22A4_e06_a001_0-100	SLC22A4	CATAGCCAAAGATACTTCCTTACTACC	TGTGCTGGATATCTGCATTTC	200
CrohnsManual_ZNF365_e01_a006_84-100	ZNF365	GAAAAAGCAGGAAGTTCAGAGAC	TTGCTTAGCATTTGCTTGTAAGG	205
CrohnsManual_ZNF365_e01_a004_49-74	ZNF365	AGAGAGGCCTGTGTCCTATGT	CTTTGGTGAGTTTATCAATTCTCTT	200
CrohnsManual_ZNF365_e01_a002_14-38	ZNF365	GTGGAGACCATACCAGATTTAGAAGC	AGCTCGGTTTCTGCTTTACCA	201
CrohnsManual_ZNF365_e01_a001_0-21	ZNF365	CTTGTAAACTACCTATTTCCCCC	TCGTAGCTGTGACTGAACTCC	204
CrohnsManual_IRGM_e01_a004_66-100	IRGM	GGTTGCATCTGCACAATTCAG	AATGAAGACAGGAATTAGTATTCACAT	210
CrohnsManual_IRGM_e01_a003_43-76	IRGM	GTGCCTCCTATTTCTCTTCCCAC	AGAACTTCTTTCCCATGTCCTCA	201

Supplementary Table 3. List of 115 nonsynonymous, nonsense, splice-site variants selected for follow-up typing in large case control samples.

Variants that were successfully designed and typed in follow-up case control samples are above the dashed line. Variants above the dashed line are successfully genotyped across studies.

chr:position	alleles	gene	type	Aasub		dbSNP (v132)
chr11:75931037	CG	C11orf30	ns	Ala896Gly		
chr13:43353233	AG	C13orf31	ns	Lys38Glu		rs34414396
chr1:199144649	AT	C1orf106	ns	Tyr333Phe		rs41313912
chr1:199145124	GC	C1orf106	ns	Arg397Thr		
chr9:138379413	CG	CARD9	splice		1	
chr9:138382026	CG	CARD9	ns	Val385Leu		rs3124993
chr9:138384709	TA	CARD9	ns	Glu270Val		rs114895119
chr6:167469973	GA	CCR6	ns	Ala89Thr		
chr6:167470814	CT	CCR6	ns	Ala369Val		rs17860852
chr10:35354137	AG	CUL2	splice		5	
chr10:35362198	GC	CUL2	ns	Leu338Val		
chr2:27579941	AC	GCKR	ns	Gln234Pro		
chr2:27599688	GT	GCKR	ns	Glu586Amb		
chr19:18146896	CT	IFI30	ns	Pro60Leu		
chr2:102354958	CT	IL18R1	ns	Thr139Ile		rs34216045
chr2:102407134	AG	IL18RAP	ns	Asp136Gly		rs11695455
chr2:102434852	GT	IL18RAP	ns	Val527Leu		
chr1:67407799	GA	IL23R	ns	Arg86Gln		rs76575803
chr1:67421184	GA	IL23R	ns	Gly149Arg		rs76418789
chr1:67478488	GA	IL23R	ns	Val362Ile		rs41313262
chr16:84500209	CT	IRF8	ns	Thr96Met		
chr16:84509524	CT	IRF8	ns	Ala201Val		
chr5:150207929	GC	IRGM	ns	Glu17Asp		
chr5:150208159	CA	IRGM	ns	Thr94Lys		rs72553867
chr1:159181643	CT	ITLN2	ns	Ala297Thr		
chr1:159189045	TG	ITLN2	ns	His61Pro		
chr9:5062561	GA	JAK2	ns	Gly571Ser		
chr9:5113003	GT	JAK2	splice		1	
chr1:199244650	GA	KIF21B	ns	Ser106Leu		
chr11:76049066	CT	LRRC32	ns	Ala407Thr		rs79525962
chr11:76049353	CG	LRRC32	ns	Gly311Ala		rs35130967

chr12:38929912	CG	LRRK2	ns	Leu286Val		
chr12:38994128	CT	LRRK2	ns	Pro1542Ser		rs33958906
chr12:39020469	GA	LRRK2	ns	Gly2019Ser		rs34637584
chr12:39026953	AG	LRRK2	ns	Asn2081Asp		rs33995883
chr19:18102335	CT	MAST3	ns	Arg390Trp		rs55994961
chr12:39225467	GA	MUC19	ns	Gly36Arg		
chr12:39226476	GA	MUC19	ns	Val56Met		
chr12:39229997	GA	MUC19	ns	Ala167Thr		rs80014625
chr12:39230776	GC	MUC19	ns	Val219Leu		rs73115384
chr16:49299292	CT	NOD2	ns	Thr189Met		rs61755182
chr16:49302189	AG	NOD2	ns	Asn289Ser		rs5743271
chr16:49302254	CT	NOD2	ns	Arg311Trp		rs104895427
chr16:49302393	AC	NOD2	ns	Asp357Ala		rs104895469
chr16:49302615	CT	NOD2	ns	Ser431Leu		rs104895431
chr16:49303373	CT	NOD2	ns	Arg684Trp		rs5743276
chr16:49303430	CT	NOD2	ns	Arg703Cys		rs5743277
chr16:49303700	GA	NOD2	ns	Val793Met		rs104895444
chr16:49308311	AG	NOD2	ns	Asn852Ser		rs104895467
chr16:49308343	AG	NOD2	ns	Met863Val		rs104895447
chr16:49314041	GC	NOD2	ns	Gly908Arg		rs2066845
chr16:49314041	GC	NOD2	ns	Gly908Arg		rs2066845
chr17:35333045	GA	ORMDL3	splice		3	
chr5:131573931	GC	P4HA2	ns	Phe218Leu		
chr5:131635620	CT	PDLIM4	ns	Ala225Val		rs10479001
chr6:106642946	GA	PRDM1	ns	Gly74Ser		rs2185379
chr6:106650242	AG	PRDM1	ns	Ile117Met		
chr6:106659789	GA	PRDM1	ns	Ser220Asn		
chr6:106659950	GC	PRDM1	ns	Ala274Pro		
chr6:106660076	CT	PRDM1	ns	Leu316Phe		
chr6:106661707	GA	PRDM1	ns	Ala577Thr		
chr5:40727650	GA	PTGER4	ns	Val294Ile		rs111866313
chr18:12804310	GT	PTPN2	ns	Asn250Lys		
chr18:12804315	GT	PTPN2	ns	Leu249Met		
chr1:114182437	GT	PTPN22	ns	His370Asn		rs72650671
chr1:114196212	CT	PTPN22	ns	Arg263Gln		rs33996649
chr1:114203227	TC	PTPN22	splice		4	rs72650669
chr5:131685803	GA	SLC22A4	ns	Arg227His		
chr5:131757267	GT	SLC22A5	ns	Gly484Val		rs28383480
chr5:131757279	GA	SLC22A5	ns	Arg488His		rs28383481
chr2:230751117	CT	SP110	ns	Gly483Arg		

chr2:230785927	CT	SP110	ns	Gly126Ser		rs41309088
chr2:230882885	GA	SP140	ns	Met687Ile		
chr1:171422483	CT	TNFSF4	ns	Ser116Asn		
chr9:116706178	AC	TNFSF8	ns	Tyr187Asp		
chr20:61837700	GT	ZGPAT	ns	Lys507Asn		
chr10:63818322	GA	ZNF365	ns	Arg302His		
chr2:233829561	CA	ATG16L1	ns	Pro64Thr		
chr2:233838333	GT	ATG16L1	ns	Arg149Leu		
chr1:199145060	CT	C1orf106	ns	Arg376Cys		
chr1:199147649	CT	C1orf106	ns	Arg554Cys		rs61745433
chr6:21339191	CG	CDKAL1	ns	Ser561Cys		
chr10:64244198	CT	EGR2	ns	Arg69Lys		
chr19:18147146	AG	IFI30	ns	Glu114Gly		rs76227216
chr2:102367893	AG	IL18R1	splice		3	
chr1:67445177	CT	IL23R	splice		4	rs115546002
chr1:67457857	GT	IL23R	ns	Asp271Tyr		
chr5:131850401	GC	IRF1	ns	Ser133Arg		
chr16:84509646	TC	IRF8	ns	Tyr242His		
chr5:150208020	AG	IRGM	ns	Thr48Ala		
chr1:159118484	CT	ITLN1	ns	Asp98Asn		
chr9:5060023	CA	JAK2	ns	His538Asn		
chr9:5071730	GT	JAK2	ns	Glu814Ocr		
chr11:76048907	CA	LRRC32	ns	Ala460Ser		
chr12:39031792	AG	LRRK2	ns	Tyr2189Cys		rs35658131
chr22:35603714	GA	NCF4	ns	Arg308Gln		
chr16:49303695	GA	NOD2	ns	Arg791Gln		rs104895464
chr18:12804331	GT	PTPN2	ns	Asn243Lys		
chr18:12820943	GT	PTPN2	ns	Ser120Amb		
chr5:131757779	GT	SLC22A5	ns	Met530Ile		
chr2:230744672	GA	SP110	ns	Thr624Met		
chr17:37744273	GT	STAT3	splice		2	
chr17:37744300	GT	STAT3	ns	Asn175Lys		
chr17:37744353	GT	STAT3	ns	Leu158Ile		
chr8:126517797	GA	TRIB1	ns	Val341Ile		
chr22:20305819	TC	UBE2L3	ns	Ile109Thr		
chr20:61810863	GC	ZGPAT	ns	Ala163Pro		
chr5:150255792	CG	ZNF300	ns	Arg401Thr		
chr5:150255916	TC	ZNF300	ns	Lys360Glu		
chr5:150256176	GT	ZNF300	ns	Thr273Lys		
chr5:150256395	AG	ZNF300	ns	Leu200Pro		

chr5:150257885	GT	ZNF300	ns	Pro66Gln	
chr10:63806049	TC	ZNF365	ns	Cys31Arg	
chr10:63806482	CA	ZNF365	ns	Thr175Asn	
chr10:63828504	GT	ZNF365	splice		1

Supplementary Table 4. Identification of independent rare risk and protective variants associated with IBD.

We identify IVS11+1G>C to be protective against IBD with an estimated odds ratio of 0.29 strong (4-fold) protective effect. 5 independent rare mutations in *NOD2* are identified to be associated with Crohn's Disease including R311W, R703C, S431L + V793M, N852S, M863V + fs1007insC. Additional variants conferring protection to IBD are identified in *IL23R*, and *CUL2* and risk missense variants in *IL18RAP*, *C1orf106*, *MUC19*, and *PTPN22.*

CD versus UC + HC (CD loci)				**Targeted Replication**			**International Immunochip**			**Combined**	
Gene,Mutation	**chr:position**	**Sequencing Based Counts**		(CD,UC + HC)			(CD, UC + HC)			(CD, UC + HC)	
		Case Counts	Control Counts	Observed	Expected	(CD AF, UC + HC AF)	Observed	Expected	(CD AF, UC + HC AF)	OR	(L95, U95)
NOD2,p.M863V+ fs1007insC	chr16:49308343	6	1		(7969,10179)			(6544,16126)		(14523,26305)	
				107	69.34	(.0067,.00157)	47	25.53	(.0036,.0011)	4.02	(2.8,5.07)
						6.73E-11			*2.15E-07*	***<1e-16***	
NOD2,p.N852S	chr16:49308311	5	0		(7962,9590)			(6542,16121)		(14504,25711)	
				73	54.12	(.0046,.0021)	14	9.28	(.001,.000465)	2.47	(1.55,3.93)
						0.00017			*0.0338*	***2.90E-05***	
NOD2,p.R703C	chr16:49303430	7	6		(3090,4100)			(8416,17183)		(11506,21283)	
				66	47.88	(.011,.0054)	133	103.12	(.0079,.0052)	1.51	(1.12,2.03)
						0.00025			*1.59E-04*	***2.33E-07***	
NOD2,p. S431L	chr16:49302615	3	5		(7949,9569)			(6545,16124)		(14494,25693)	
				62	87	(.0039,.0019)	50	39.52	(.0038,.0026)	1.45	(1.07,1.95)
						0.0014			*0.023*	***0.00025***	
NOD2,p.V793M	chr16:49303700	4	4		(2227,3252)			(6949,16156)		(9176,19408)	
				15	10.11	(.00337,.0015)	56	43.86	(.004,.0026)	1.45	(1.07,1.95)
						0.0217			*0.0127*	***0.002***	
NOD2,p.R311W	chr16:49302254	3	1		(3010,5506)			(6950,16149)		(9960,21654)	
				10	7.49	(.0017,.00099)	20	14.14	(.0014,.00073)	2.28	(1.37,3.79)
						0.118			*0.029*	***0.00143***	
IL18RAP,p.V527L	chr2:102434852	4	2		(7920,9561)			(4131,10336)		(12051,19897)	
				57	41.94	(.0036,.0015)	2	0.9	(.00025,0)	3.03	(1.95,4.73)
						0.0006			*0.0456*	***2.90E-04***	
MUC19,p.V56M	chr12:39226476	3	2		(2227,3253)			(4963,11324)		(7190,14577)	
				13	8.79	(.0029,.00138)	3	0.95	(.0003,.00004)	4.32	(1.93,9.67)
						0.033			*0.11*	***0.00546***	

IBD versus HC (CD + UC loci)				**Targeted Replication**			**International Immunochip**			**Combined**	
Gene,Mutation	**chr:position**	**Sequencing Based Counts**		(IBD, HC)			(IBD, HC)			(IBD, HC)	
		Case Counts	Control Counts	Observed	Expected	(IBD AF, HC AF)	Observed	Expected	(IBD AF, HC AF)	OR	(L95, U95)
CARD9,p.IVS11+1C>G	chr9:138379413	0	6		(10439,5933)			(16420,10707)		(26859,16640)	
				42	69.3	(.002,.0058)	78	137.11	(.0024,.0071)	0.29	(0.22,0.37)
						1.90E-08			*3.33E-16*	***<1e-16***	
IL23R,p.V362I	chr1:67478488	9	12		(5321,6112)			(12241,10426)		(17562,16538)	
				140	143.56	(.0131,.0127)	271	319.05	(.011,.0152)	0.72	(0.63,0.83)
						0.27			*2.70E-05*	***1.18E-05***	
IL23R,p.G149R	chr1:67421184	2	1		(4629,5305)			(13789,10707)		(18418,16012)	
				24	29.28	(.0026,.0045)	69	87.89	(.0025,.0043)	0.6	(0.45,0.79)
						0.064			0.0013	***3.20E-04***	
CUL2,p.IVS17+ 5A>G	chr10:35354137	4	4		(5582,1684)			(16387,10707)		(21969,12391)	
				63	65.8	(.0056,.0063)	213	245.88	(.0065,.0092)	0.72	(0.60,0.86)
						0.2			0.0004	***3.45E-04***	
PTPN22,p.H370N	chr1:114182437	12	6		(5583,1682)			(21997,12393)		(21997,12393)	
				33	31.6	(.003,.002)	101	85.9	(.0031,.002)	1.6	(1.16,2.24)
						0.3			*0.0046*	***6.20E-03***	
C1orf106,p.Y333F	chr1:19914464	12	6		(13991,8486)			NA		(13991,8486)	
				333	150.72	(.013,.01)	-	-	-	1.44	(1.02,2.06)
						0.009			*NA*	***0.009***	

Supplementary Table 5. Conditional analysis of NOD2 variants.

Conditional analysis and haplotype based test demonstrates independent contributions from R703C and R311W (A subset of all genotyped samples were used for the conditional analysis – as not all samples had both variants typed).

Conditional Analysis NOD2 (*p.*R311W,*p.*R703C)								
SNP	**A1**	**TEST**	**NMISS**	**OR**	**L95**	**U95**	**STAT**	**P**
*p.*R311W	A	ADD	33563	2.115	1.35	3.314	3.269	0.001078
*p.*R703C	A	*p.*R703C	33563	1.433	1.177	1.744	3.585	3.37E-04

Supplementary Table 6. Conditional analysis for rare variants in IL18RAP and MUC19.

Conditional analysis for rare variants in IL18RAP and MUC19 demonstrate independent contributions from GWAS associated common variant (A subset of all genotyped samples were used for the conditional analysis – as not all samples had both variants typed).

Conditional Analysis MUC19 (rs11564258,*p.*V56M)								
SNP	**A1**	**TEST**	**NMISS**	**OR**	**L95**	**U95**	**STAT**	**P**
*p.*V56M	A	ADD	23528	3.008	1.337	6.769	2.662	0.007769
rs11564258	A	rs11564258	23528	1.486	1.325	1.666	6.759	1.39E-11

Conditional Analysis IL18RAP (rs2058660,*p.*V527L)								
SNP	**A1**	**TEST**	**NMISS**	**OR**	**L95**	**U95**	**STAT**	**P**
*p.*V527L	A	ADD	23521	11.51	1.374	96.44	2.253	0.02426
rs2058660	A	rs2058660	23521	1.486	1.066	1.166	4.741	2.13E-06

Supplementary Table 7. Additive genetic variance explained.

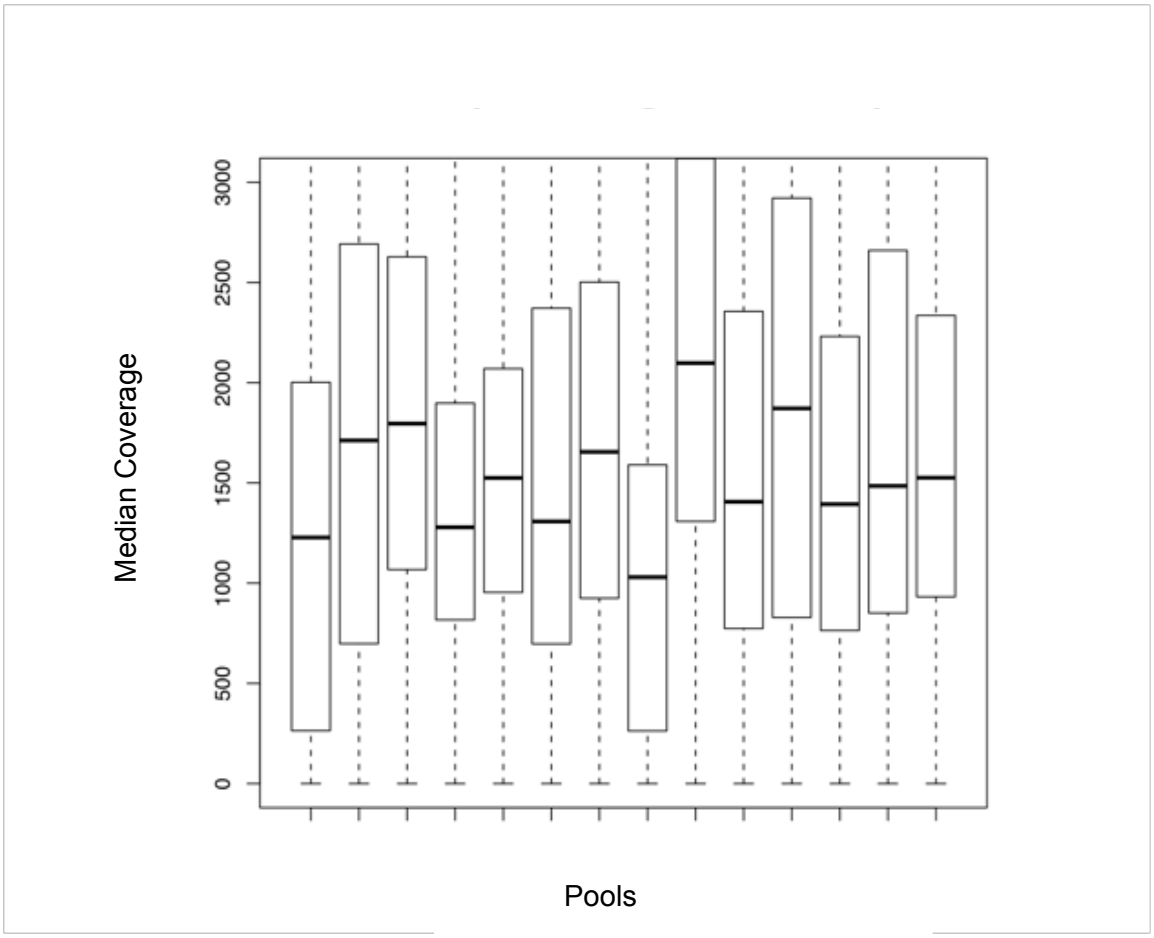
(Estimate) denotes without M863V included in the calculation.

Variant	% Additive Genetic Variance Explained		RR
	All Samples	AJ	
NOD2, M863V	0.370	0.759	4.02
NOD2, N852S	0.160	0.553	2.47
NOD2, S431L + V793M	0.016	0.009	1.5
NOD2, R703C Only	0.040	0.051	1.61
NOD2, R311W Only	0.004	0.000	1.82
NOD2, R703C + R311W	0.028	0.000	2.7
CARD9, IVS11+1C>G	0.320	0.158	3.48
IL18RAP,V527L	0.147	0.130	2.79
PTPN22,H370N	0.016	0.000	1.6
MUC19,V56M	0.167	1.413	4.32
C1orf106,Y333F	0.043	0.043	1.44
CUL2, IVS17+5A>G	0.018	0.000	1.39
IL23R,V362I	0.028	0.000	1.34
IL23R,G149R	0.036	0.000	1.68
TOTAL	1.39 (1.02)	3.12 (2.36)	

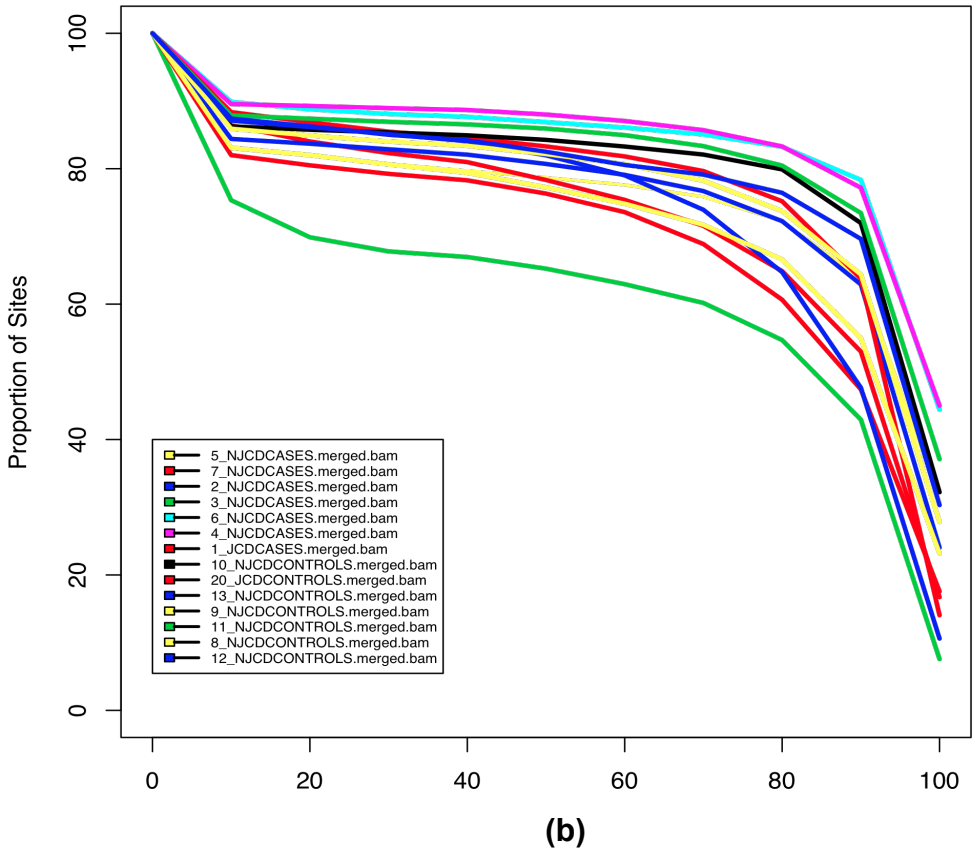
Supplementary Figures

Supplementary Figure 1. Coverage summary and power to detect variants.

(a) 1500X median coverage per pool, corresponding to an average of 30X per samples (15X per individual chromosome) was achieved for the CD pooled sequencing experiments. Power to detect a singleton is approximately greater than or equal to 60% in greater than 80% of all targeted genomic positions for the **(b)** CD experiment.



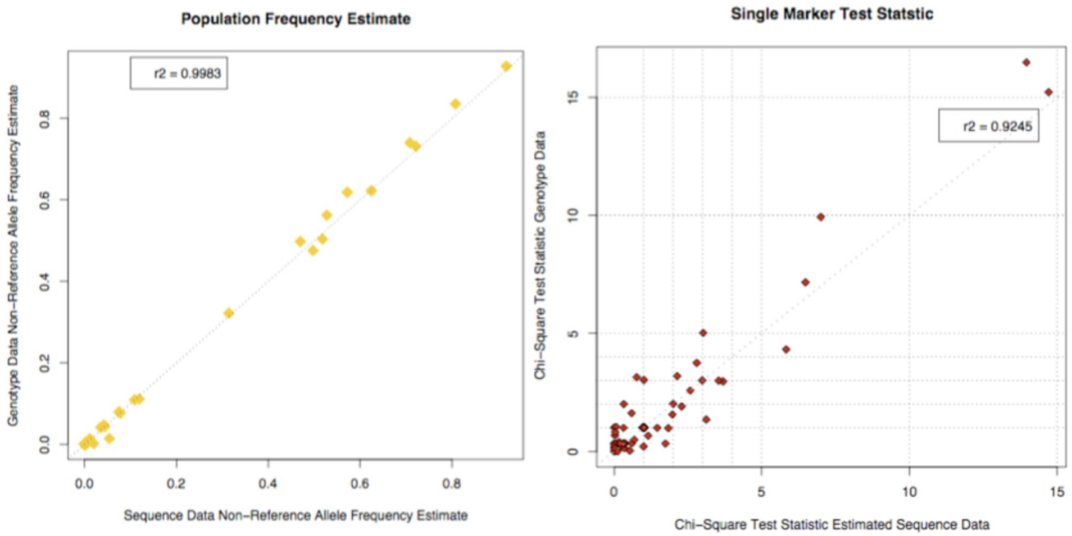
(a)


Proportion of Sites
0
20
40
60
80
100
5_NJCDCASES.merged.bam
7_NJCDCASES.merged.bam
2_NJCDCASES.merged.bam
3_NJCDCASES.merged.bam
6_NJCDCASES.merged.bam
4_NJCDCASES.merged.bam
1_JCDCASES.merged.bam
10_NJCDCONTROLS.merged.bam
20_JCDCONTROLS.merged.bam
13_NJCDCONTROLS.merged.bam
9_NJCDCONTROLS.merged.bam
11_NJCDCONTROLS.merged.bam
8_NJCDCONTROLS.merged.bam
12_NJCDCONTROLS.merged.bam

(b)

Supplementary Figure 2. Pooled sequence allele frequency and genotyped allele frequency correlation.

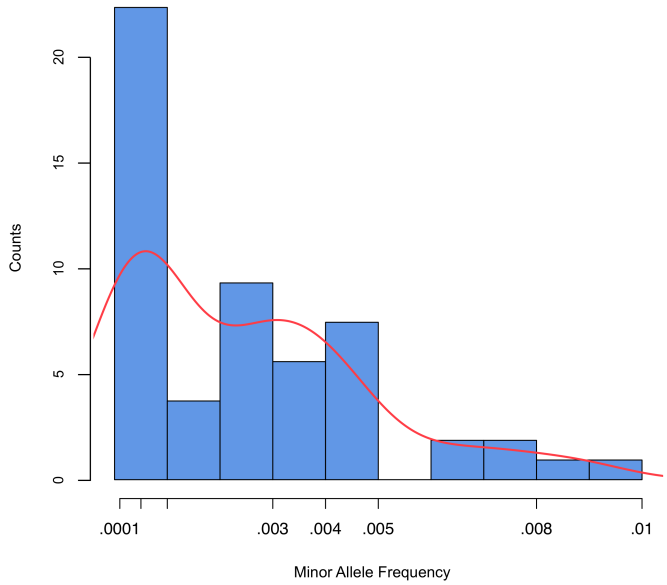
We observe high correlation between pooled sequencing estimated frequencies and genotype estimated frequencies.



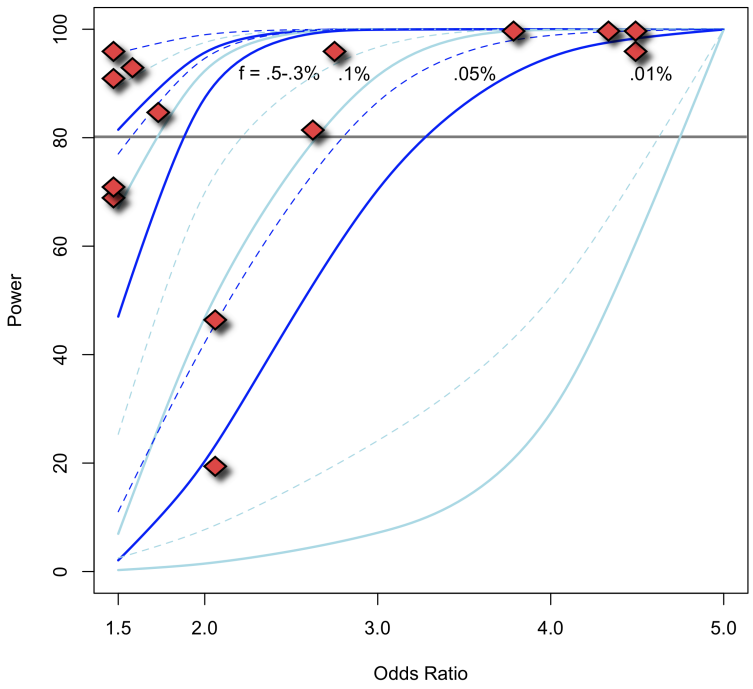
(a) **(b)**

Supplementary Figure 3. Allele frequency distribution and study-wide power.

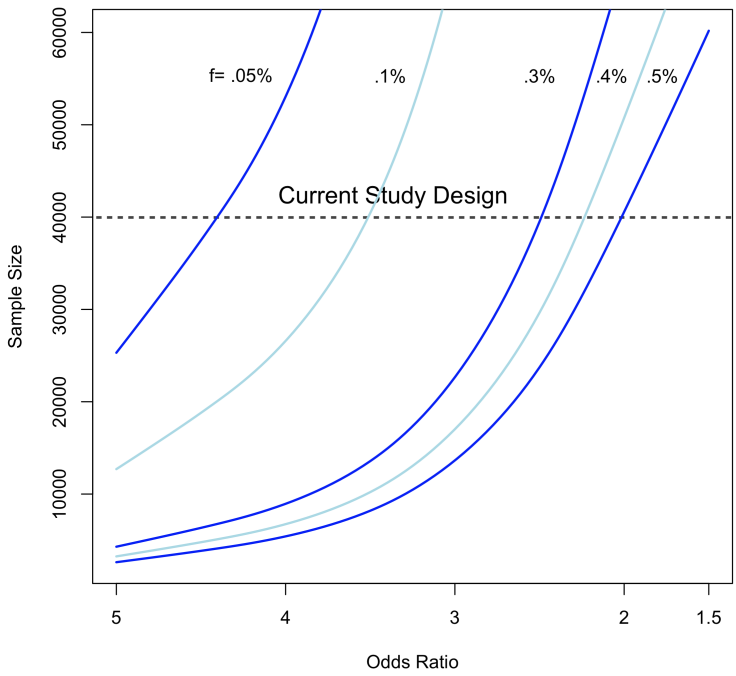
(a) The allele frequency distribution estimated in a combined dataset of 13167 CD disease patients, 12153 UC disease patients, and 15331 healthy controls for the 70 variants tested for association to IBD in follow up genotyping. (b) Experimental-wide power (P value <= .007, solid-line) and nominal power (P value <= .01, dashed line) to detect association at varying odds ratio and allele frequencies (percentage). In red diamonds we overlay the variants that are highlighted in the manuscript. (c) Required number of samples to reach experimental-wide significance for minor allele frequency range [.0005-.005] at varying odds ratio.



(a)



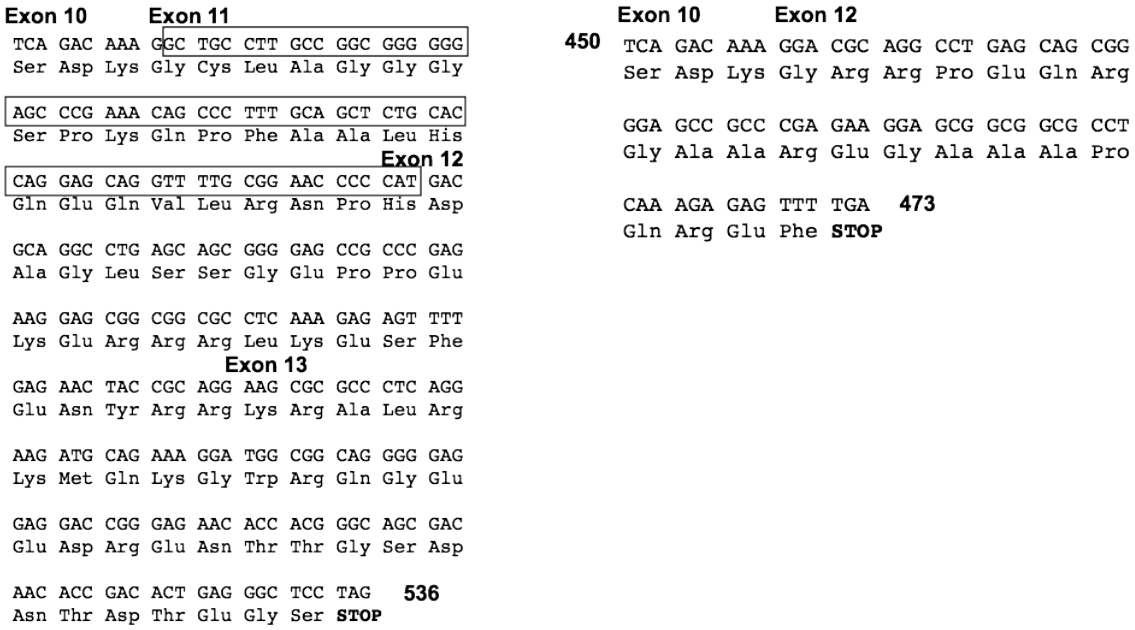
(b)


Sample Size
60000
50000
40000
30000
20000
10000
f= .05%
.1%
.3%
.4%
.5%
Current Study Design
5
4
3
2
1.5
Odds Ratio

(c)

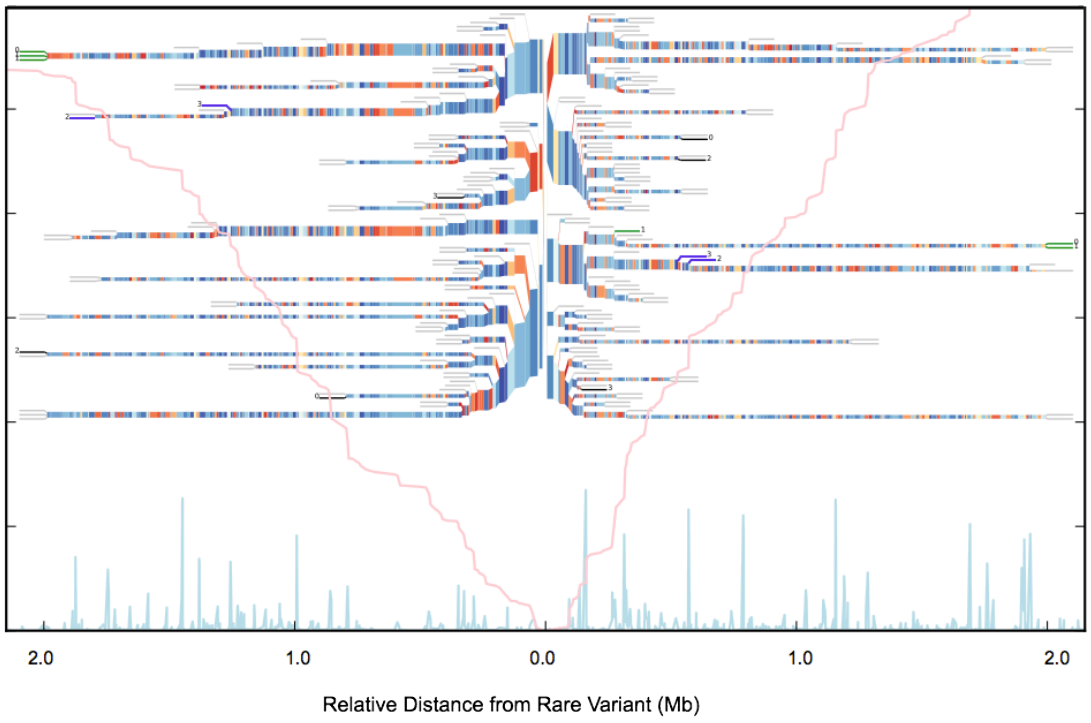
Supplementary Figure 4. CARD9 reference protein and predicted splicing.

CARD9 reference protein including exons 10, 11, 12, and 13 and predicted truncated CARD9 protein after exon-skipping due to IVS11+1G>C.



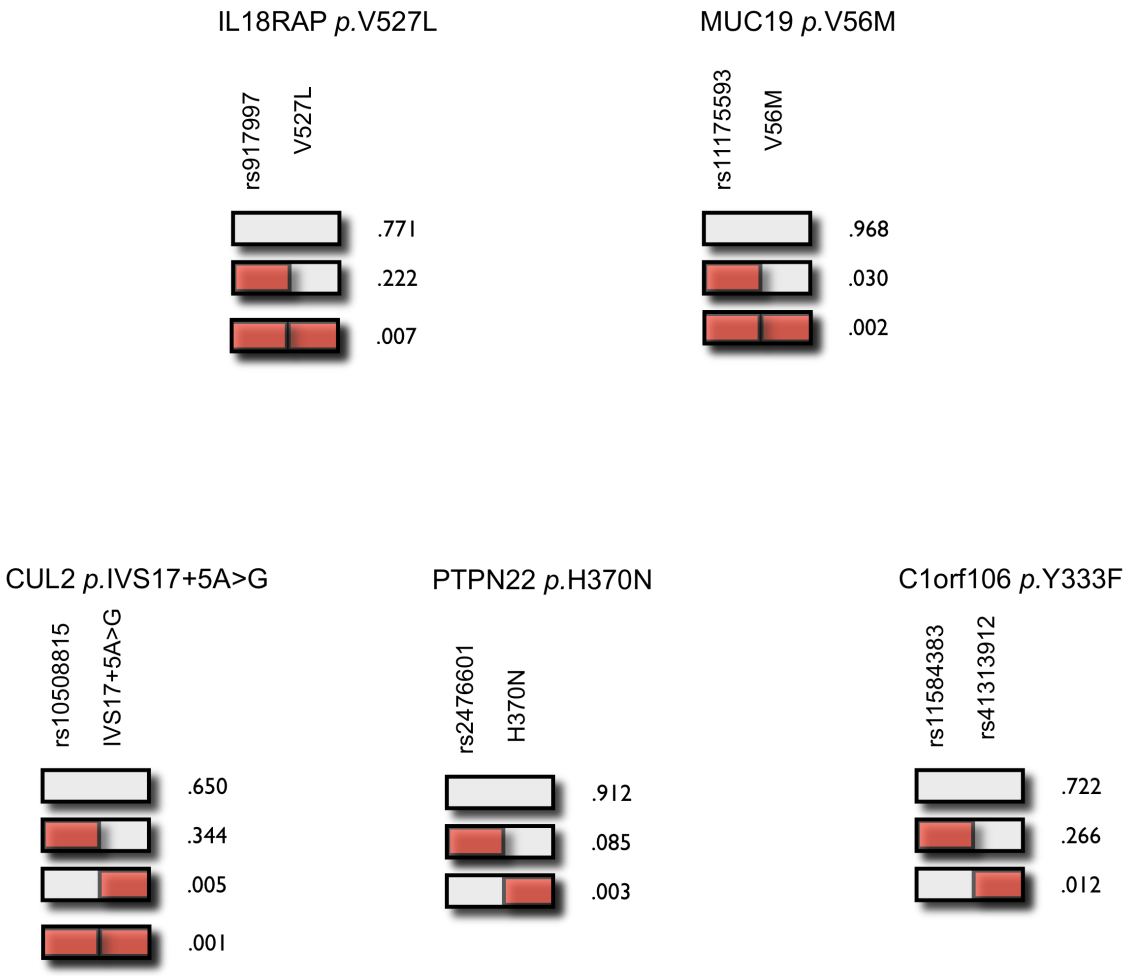
Supplementary Figure 5. Shared extended haplotypes of Ashkenazi Jews with *NOD2 p.N852S* and *p.M863V* mutation.

N852S and M863V are observed at higher frequencies in Ashkenazi Jews compared to non-Jewish disease patients with European ancestry (4%,2%,.5%,.%5) with a 4Mb extended shared segments. Homozygote individual for *p.N852S* (labeled 1) has at least a 2 Mb shared segment.



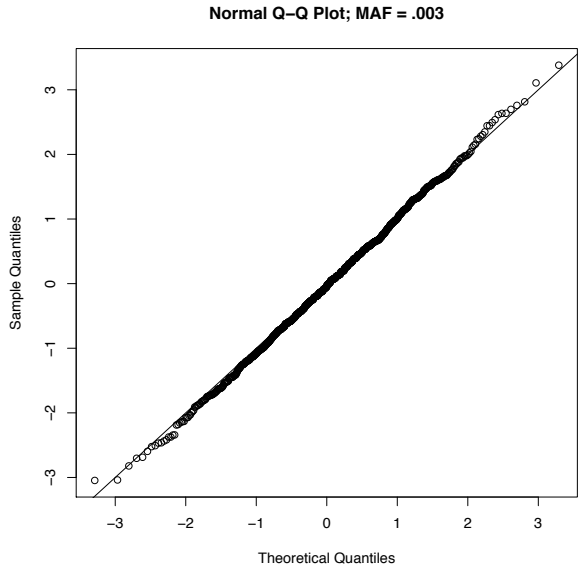
Supplementary Figure 6. LD structure of associated rare variants and common associated GWAS variant.

Haplotype Structure demonstrates that rare variant and common GWAS associated variant in *CUL2*, *PTPN22*, and *C1orf106* are in linkage equilibrium, whereas in *MUC19* and *IL18RAP* they are not. Conditional Analysis (Table S4) demonstrates an independent and an additional contributing effect from the rare variants in *IL18RAP* and *MUC19*.

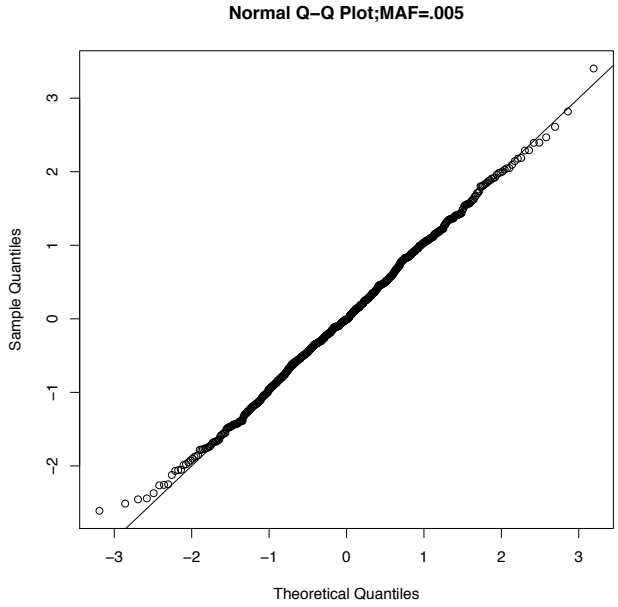


Supplementary Figure 7. Null simulations for M.A.R.V.

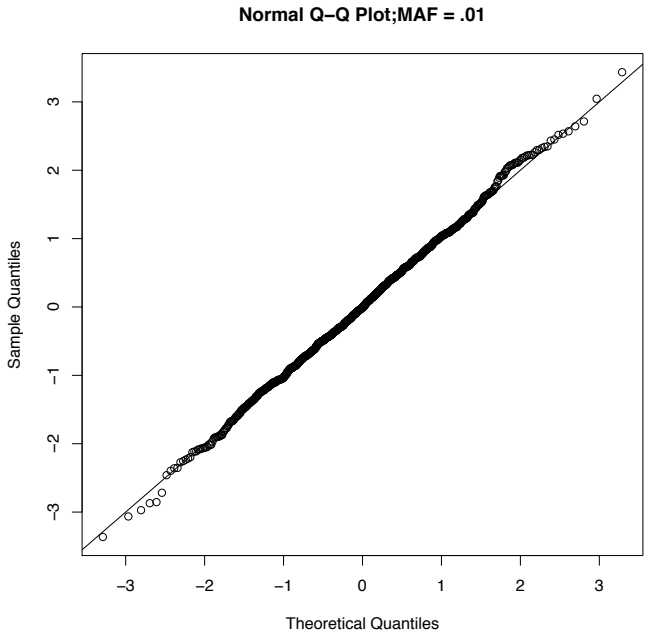
Distribution of quantiles under the null hypothesis of no disease association under three different scenarios: a) MAF = .003; b) MAF = .005; c) MAF = .01. 1,000 simulations under the null demonstrates properly calibrated test statistic values. Simulations done with 20,000 cases versus 20,000 controls with 10 stratas of 2,000 cases versus 2,000 matched controls.



(a)



(b)


Normal Q–Q Plot;MAF = .01
Sample Quantiles
Theoretical Quantiles
−3
−2
−1
0
1
2
3

(c)

Consortium Authors

The Principal Investigators of the NIDDK IBDGC are: Steven R Brant[1], Judy H Cho[2,3], Richard H Duerr[4], Dermot P B McGovern[5], John D Rioux[6], Mark S Silverberg[7]

[1]Johns Hopkins University School of Medicine, Department of Medicine, and Johns Hopkins University Bloomberg School of Public Health, Department of Epidemiology, Baltimore, Maryland, USA.
[2]Department of Medicine, Yale University, New Haven, Connecticut, USA.
[3]Department of Genetics, Yale University, New Haven, Connecticut, USA.
[4]Division of Gastroenterology, Hepatology and Nutrition, Department of Medicine, University of Pittsburgh School of Medicine, and Department of Human Genetics, Graduate School of Public Health, University of Pittsburgh, Pittsburgh, Pennsylvania, USA.
[5]Inflammatory Bowel and Immunobiology Research Institute, Cedars-Sinai Medical Center, Los Angeles, California, USA.
[6]Université de Montréal and the Montreal Heart Institute, Research Center, Montréal, Québec, Canada.
[7]Mount Sinai Hospital Inflammatory Bowel Disease Group, University of Toronto, Toronto, Ontario, Canada.

The International Inflammatory Bowel Disease Genetics Consortium consists of the following people: Tariq Ahmad[6], Carl A. Anderson[3], Vito Annese[15,64], Robert N. Baldassano[20], Tobias Balschun[8], Murray Barclay[10], Jeffrey C. Barrett[3], Theodore M. Bayless[21], Joshua C. Bis[11], Stephan Brand[22], Steven R. Brant[21], Suzanne Bumpstead[3], Carsten Buning[23], Judy H. Cho[19,63], Albert Cohen[24], Jean-Frederick Colombel[25], Mario Cottone[26], Mauro D'Amato[57], Renata D'Inca[29], Mark J. Daly[65], Ted Denson[27], Marla Dubinsky[30], Richard H. Duerr[47,61], Cathryn Edwards[31], David Ellinghaus[1], Tim Florin[32], Denis Franchimont[33], Andre Franke[1], Richard Gearry[10], Michel Georges[13], Jurgen Glas[22,34,35] , Gossum33 Andre Van Gossum[33], Anne M. Griffiths[55], Stephen L. Guthery[36], Hakon Hakonarson[4,20], Talin Haritunians[14], Jean-Pierre Hugot[39], Dirk J de Jong[66], Luke Jostins[3], Subra Kugathasan[59], Gerd Kullak-Ublick[54], Anna Latiano[15], Debby Laukens[28], Ian Lawrance[40], James Lee[9], Charlie W. Lees[7], Marc Lemann[41], Arie Levine[42], Cecile Libioulle[43], Edouard Louis[43], John C. Mansfield[60], Christopher G. Mathew[16], Dermot P.B. McGovern[2,14], Mitja Mitrovic[12,67], Grant W. Montgomery[17], Craig Mowat[44], William Newman[45], Orazio Palmieri[15], Julián Panés[46], Miles Parkes[9], Anne Phillips[44], C.Y. Ponsioen[52], Uros Potočnik[67], Natalie J. Prescott[16], Deborah D. Proctor[19], Graham L. Radford-Smith[5], Miguel Regueiro[47], John D. Rioux[56], Rebecca Roberts[10], Jerome I. Rotter[14], Paul Rutgeerts[48], Jeremy Sanderson[49], Miquel Sans[46], Jack Satsangi[7], Stefan Schreiber[1,62], Philip Schumm[18], Frank Seibold[50], Yashoda Sharma[19], Mark S. Silverberg[51], Lisa A. Simms[5], A. Hillary Steinhart[51], Stephan R. Targan[2], Kent

D. Taylor[14], Leif Torkvist[53], Severine Vermeire[48], Jonas Halfvarson[37], H.W. Verspaget[38], Martine De Vos[28], Thomas Walters[55], Kai Wang[4], Rinse K. Weersma[58], David Whiteman[17], Cisca Wijmenga[12]

[1] Institute of Clinical Molecular Biology, Christian-Albrechts-University Kiel, Schittenhelmstr. 12, D-24105 Kiel, Germany
[2] Inflammatory Bowel and Immunobiology Research Institute, Cedars-Sinai Medical Center, Los Angeles, California, USA.
[3] Wellcome Trust Sanger Institute, Wellcome Trust Genome Campus, Hinxton, Cambridge, UK
[4] Center for Applied Genomics, The Children's Hospital of Philadelphia, Philadelphia, PA 19104, USA
[5] Inflammatory Bowel Disease Research Group, Queensland Institute of Medical Research, Brisbane, Australia.
[6] Peninsula College of Medicine and Dentistry, Barrack Road, Exeter, UK
[7] Gastrointestinal Unit, Molecular Medicine Centre, University of Edinburgh, Western General Hospital, Crewe Road, Edinburgh, UK
[8] popgen Biobank, Christian-Albrechts University Kiel, D-24105 Kiel, Germany
[9] Gastroenterology Research Unit, Addenbrooke's Hospital, University of Cambridge, Cambridge, UK
[10] Department of Medicine, University of Otago, Christchurch 8140, New Zealand
[11] Cardiovascular Health Research Unit, Department of Medicine, University of Washington, Seattle, USA
[12] Department of Genetics, University Medical Center Groningen, Groningen, the Netherlands
[13] Department of Genetics, Faculty of Veterinary Medicine, University of Liège B43, 20 Bd de Colonster, 4000 Liège, Belgium
[14] Medical Genetics Institute, Cedars-Sinai Medical Center, Los Angeles, California, USA
[15] Unit of Gastroenterology, IRCCS-CSS Hospital, San Giovanni Rotondo, Italy
[16] Department of Medical and Molecular Genetics, King's College London School of Medicine, Floor 8 Tower Wing, Guy's Hospital, London, UK
[17] Molecular Epidemiology, Queensland Institute of Medical Research, Brisbane, Australia 4006
[18] Department of Health Studies, University of Chicago, Chicago, Illinois, USA
[19] Section of Digestive Diseases, Department of Medicine, Yale University, New Haven, Connecticut, USA
[20] Department of Pediatrics, Center for Pediatric Inflammatory Bowel Disease, The Children's Hospital of Philadelphia, Philadelphia, USA
[21] Inflammatory Bowel Disease Center, Dept. of Medicine, Johns Hopkins University School of Medicine, Baltimore, Maryland, U.S.A
[22] Department of Medicine II, University Hospital MunichGrosshadern, Ludwig-Maximilians-University, Munich, Germany
[23] Department of Gastroenterology, Charité, Campus Mitte, Universitätsmedizin Berlin, Berlin, Germany
[24] Montreal Jewish General Hospital, Montréal, Québec, Canada

[25] Registre EPIMAD, Université de Lille, Lille, France
[26] Unit of Gastroenterology, Cervello Hospital, Palermo, Italy
[27] Pediatric Gastroenterology, Cincinnati Children's Hospital. Medical Center. 3333 Burnet Ave, Cincinnati, USA
[28] Department of Hepatology and Gastroenterology, Ghent University Hospital, Ghent, Belgium
[29] Division of Gastroenterology, University Hospital Padua, Italy
[30] Department of Pediatrics, Cedars Sinai Medical Center, Los Angeles, CA, USA
[31] Torbay Hospital, Torbay, Devon, UK
[32] Department of Gastroenterology, Mater Health Services, Brisbane, Australia 4101
[33] Department of Gastroenterology, Erasmus Hospital, Free University of Brussels, Brussels, Belgium
[34] Department of Preventive Dentistry and Periodontology, Ludwig-Maximilians-University, Munich, Germany
[35] Department of Human Genetics, RWTH Aachen, Germany
[36] Department of Pediatrics, University of Utah School of Medicine, Salt Lake City, UT, USA
[37] Department of Medicine, Örebro University Hospital, Örebro, Sweden
[38] Dept of Gastroenterology, Leiden University Medical Center, Leiden, The Netherlands
[39] Université Paris Diderot, Paris, France
[40] School of Medicine and Pharmacology, The University of Western Australia, Fremantle, Australia 6160
[41] GETAID group, Université Paris Diderot, Paris, France
[42] Pediatric Gastroenterology Unit, Wolfson Medical Center and Sackler School of Medicine, Tel Aviv University, Israel
[43] Division of Gastroenterology, CHU, Université de Liège, Liège, Belgium
International IBD Genetics Consortium – Meta-analysis for Crohn's disease v2
[44] Dept of Medicine, Ninewells Hospital and Medical School, Dundee, UK
[45] Department of Medical Genetics, University of Manchester, Manchester, UK
[46] Department of Gastroenterology, Hospital Clínic/IDIBAPS. CIBER EHD. Barcelona, Spain
[47] Division of Gastroenterology, Hepatology and Nutrition, School of Medicine, University of Pittsburgh, Pittsburgh, Pennsylvania, USA
[48] Division of Gastroenterology, University Hospital Gasthuisberg, Leuven, Belgium
[49] Dept Gastroenterology, Guy's & St Thomas' NHS Foundation Trust, St Thomas' Hospital, London, UK
[50] Division of Gastroenterology, Inselspital, University of Bern, Bern, Switzerland
[51] Mount Sinai Hospital Inflammatory Bowel Disease Centre, University of Toronto, Canada
[52] Department of Gastroenterology, Academic Medical Center, Amsterdam, the Netherlands
[53] Department of Clinical Science Intervention and Technology, Karolinska Institutet, Stockholm, Sweden

[54] Division of Clinical Pharmacology and Toxicology University Hospital Zurich, CH-8091 Zurich, Switzerland
[55] The Hospital for Sick Children, University of Toronto, Ontario, Canada
[56] Université de Montréal and the Montreal Heart Institute, Research Center, Montréal, Québec, Canada
[57] Department of Biosciences and Nutrition, Karolinska Institute, Stockholm, Sweden
[58] Department of Gastroenterology, University Medical Center Groningen, Groningen, The Netherlands
[59] Department of Pediatrics; Emory University School of Medicine, Atlanta, GA, USA
[60] Institute of Human Genetics, Newcastle University, Newcastle upon Tyne, UK
[61] Department of Human Genetics, Graduate School of Public Health, University of Pittsburgh, Pittsburgh, Pennsylvania, USA
[62] Department for General Internal Medicine, Christian-Albrechts-University, Schittenhelmstr. 12, D-24105 Kiel, Germany
[63] Department of Genetics, Yale School of Medicine, New Haven CT, USA
[64] Unit of Gastroenterology, University Hospital Careggi Florence, Italy
[65] Center for Human Genetic Research, Massachusetts General Hospital, Harvard Medical School, Boston, Massachusetts, USA
[66] Department of Gastroenterology and Hepatology, Radboud University Medical Center Nijmegen, the Netherlands
[67] University of Maribor, Faculty of Medicine, Center for human molecular genetics and pharmacogenomics, Slomskov trg 15, 2000 Maribor