Supplementary Table 1: *IRGM/ZNF300* region tagging SNPs: association results from 931 cases and 976 controls

Variant	MA (+)	Phenotype	-/-	-/+	+/+	MAF (%)	p_trend	OR (95% CI)
rs10059011	A	Controls	279	436	164	43.5	0.164	1.10 (0.96,1.26)
		CD	238	448	166	45.8		
rs1150153	G	Controls	791	31	1	2.0	0.806	0.93 (0.58,1.53)
		CD	842	33	0	1.9		
rs1150154	T	Controls	817	20	1	1.3	0.396	1.28 (0.73,2.24)
		CD	809	28	0	1.7		
rs12659118	C	Controls	760	143	5	8.4	0.0017	1.43 (1.14,1.79)
		CD	650	169	12	11.6		
rs1277648	G	Controls	694	106	7	7.4	0.502	0.91 (0.70,1.19)
		CD	757	120	0	6.8		
rs13177730	C	Controls	678	97	0	6.3	0.335	0.87 (0.65,1.16)
		CD	767	88	3	5.5		
rs4958430	C	Controls	765	127	9	8.0	0.365	0.89 (0.69,1.14)
		CD	746	104	10	7.2		
rs9324665	G	Controls	674	120	6	8.3	0.063	0.78 (0.60,1.02)
		CD	698	103	1	6.5		

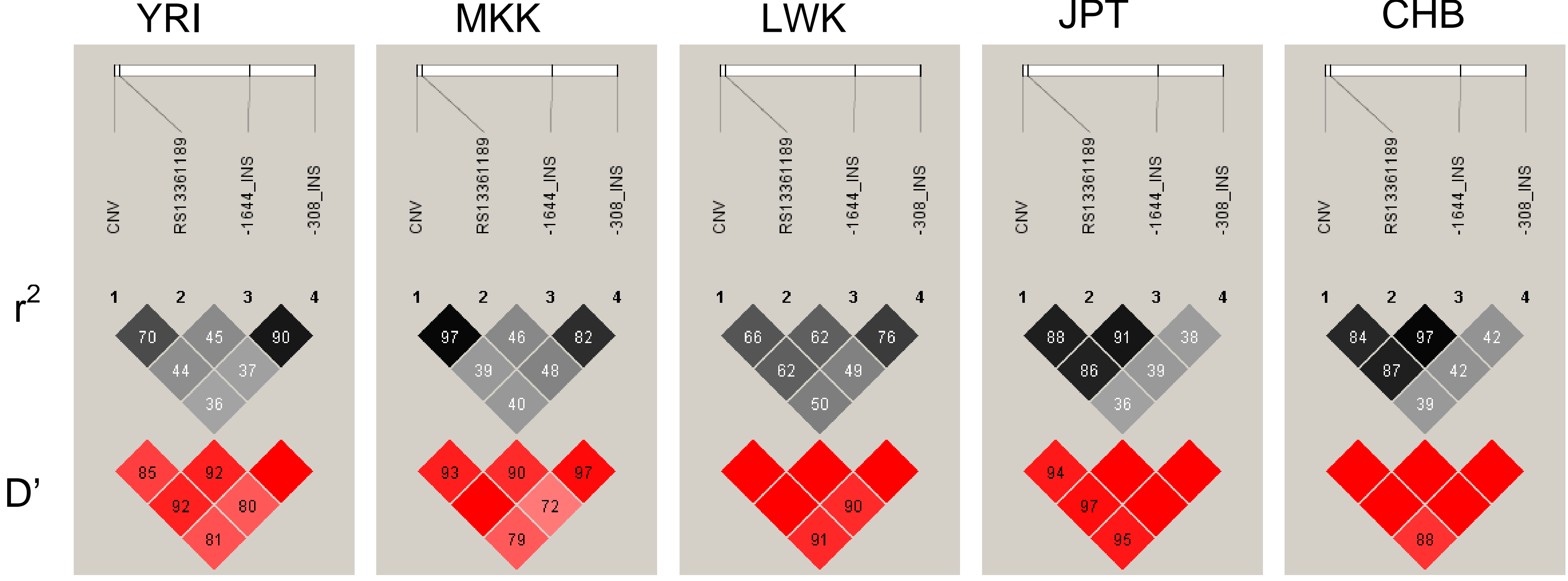
MA = minor allele, MAF = minor allele frequency (%), p_trend = p-value from Cochran-Armitage's trend test; OR = allelic odds ratio (95% confidence interval) for minor allele.

Supplementary Table 2: Sequences of overlapping nested primers used for sequencing a 2.9 kb region upstream of *IRGM*.

IRGM_PROM_PCR_fwd*	5'-ACAATGAGTGTGTGAAACAGACCT-3'
IRGM_PROM_PCR_rev*	5'-CATAGTGATGTTAACTGGTGTCCTG-3'
IRGM_PROM_SEQ1_fwd	5'-CAAGGGCTGAAGAGTGCAG-3'
IRGM_PROM_SEQ1_rev	5'-GTCCATTTTGACAGGGTGCT-3'
IRGM_PROM_SEQ2_fwd	5'-AGGTTTGCAAATGCACCA-3'
IRGM_PROM_SEQ2_rev	5'-AGGGGCCAGGTATTTGAGAC-3'
IRGM_PROM_SEQ3_fwd	5'-AGGGTCCGTGGCTTCATT-3'
IRGM_PROM_SEQ3_rev	5'-TGAGATGGCTTGAAAAGCAG-3'
IRGM_PROM_SEQ4_fwd	5'-GCAATGGGAATTGGATGTTT-3'
IRGM_PROM_SEQ4_rev	5'-CGCCTGTAGTCCCAGCTAAT-3'

*Primers also used to generate PCR product.

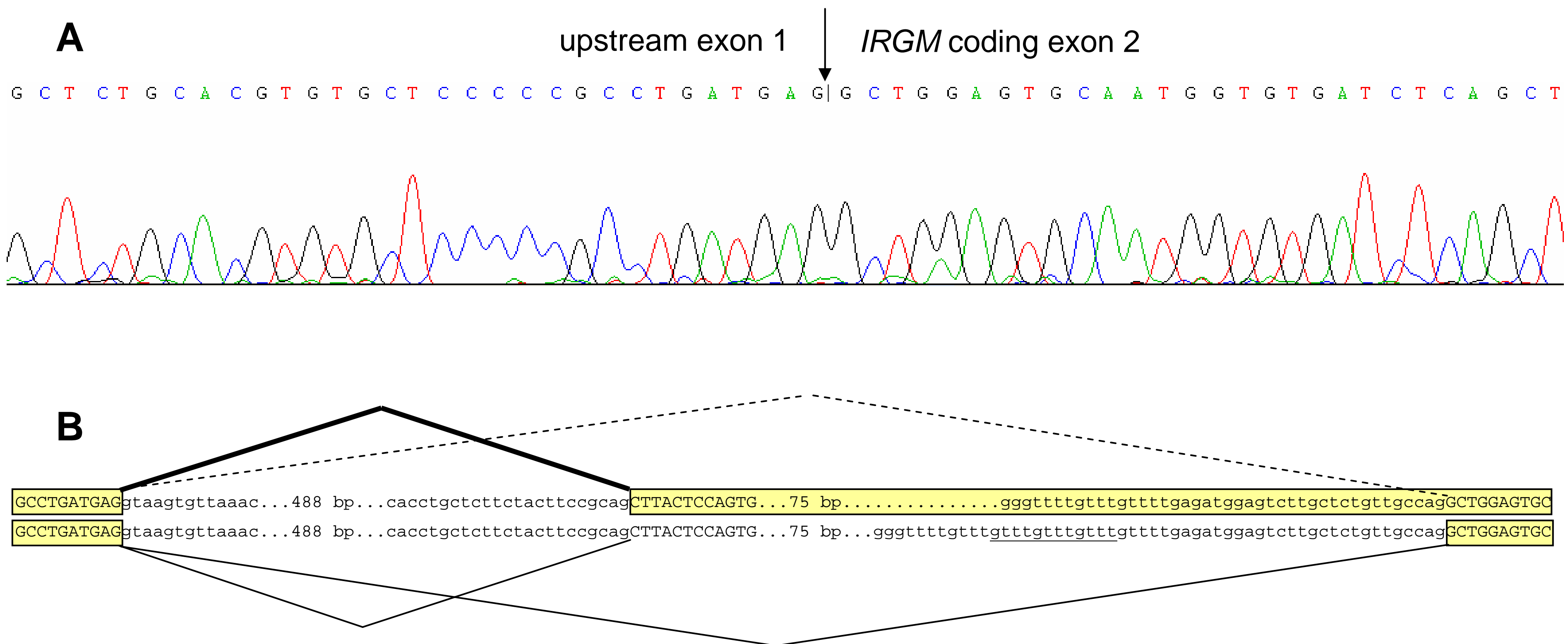
Supplementary Figure 1. Linkage disequilibrium of IRGM variants in Asian and African populations



Supplementary Figure 2. Allelic imbalance in *IRGM* mRNA compared to genomic DNA. Examples of sequencing electropherograms for exonic SNP rs10065172 in genominc DNA and cDNA from 8 CD individuals. Position of SNP is indicated by lower case letters in sequence above. Mean ratios of peak heights of 'wild type C' versus 'risk T' allele are given for both gDNA and cDNA and a comparison of this distribution of values was statistically significant (p=0.0152). A $cDNA_{C/T}$:$gDNA_{C/T}$ >1 indicates higher expression of the C allele.

gDNA AACTACc/tTGATGG	cDNA AACTACc/tTGATGG	C/T ratio cDNA	C/T ratio gDNA	cDNA: gDNA
		30.46	0.84	36.27
		353.44	0.69	515.2
		2.52	0.97	2.6
		1.82	0.89	2.04
		3.17	0.79	4.01
		23.3	0.79	29.4
		4.98	0.71	7.01
		0.50	0.71	0.70

Supplementary Figure 3. A) Sequencing electropherogram of cDNA from primary lymphocytes of an individual homozygous for the risk allele at -308 microsatellite [$(GTTT)_5$/ $(GTTT)_5$] showing use of an alternative splice acceptor site 159 bp downstream of the canonical splice site. **B)** Schematic summary of semi-quantitative splicing behaviour of *IRGM* transcripts in individuals homozygous for low-risk (above) and high-risk (below) haplotypes. High risk: bold line is predominant transcript, dotted line is minority transcript. Low risk: usage of both sites. Yellow boxes – exons of transcripts. Underline – additional 3xGTTT repeats in -308 risk allele.



Supplementary Methods

Fine mapping

Custom competitive allele specific assays for the 8 tagging SNPs (Supplementary table 1) were designed via PrimerPicker software (http://www.kbioscience.co.uk/primer-picker/index.htm, sequences available on request) and genotyped using KASPar chemistry (KBioscience, Hertforshire, UK) in over 931 CD cases and 976 controls. Allelic discrimination of the two fluorescent probes (FAM and VIC) was carried out via endpoint read (allelic discrimination) on ABI7900HT Sequence detection system.

Sequencing of ZNF300

The 6 exons of *ZNF300* including splice sites and all coding sequence were amplified in 9 independent PCR reactions using 8pmol of each primer pair (available on request) 1x PCR Master mix (Promega) and 25 ng of template genomic DNA in a 10µl reaction. PCR conditions were as follows: 5 min at 95°C followed by 30 cycles of 30s at 95°C, 30s at 55°C and 30s at 72°C with a final extension step of 10 min at 72°C. Subsequent downstream forward and reverse cycle sequencing was performed after ExoSAP-IT clean up (USB Europe, Staufen, Germany) in 18 independent reactions using 8pmol of each of either the forward or reverse primer used to generate the PCR product (Supplementary Table 4) and 0.25µl of ABI BigDye v3.1 (Applied Biosystems) in a 5µl reaction volume and using standard conditions. Products were analysed on an ABI3730xl DNA sequencer (Sequence analysis, Applied Biosystems) and aligned to the published genomic sequence using the Sequencher 4.7 package (GeneCodes).