SUPPLEMENTARY TABLES

A. India	Discovery[ϕ]	Replication	Total
Cases (no.)	989 (889)	941 (889)	1930 (1778)
Male	584 (532)	561 (525)	1145 (1057)
Female	405 (357)	380 (364)	785 (721)
Mean age at study encounter ±SD (yr)	16 ±9.4 (15.9 ±9.3)	31.3 ±16.8 (31.4 ±16.8)	23.7 ±15.6 (23.9 ±15.7)
Range	1-62	3-73	1-73
Mean age at onset of VL ±SD (yr)	17.2 ±13.1	26.8 ±15.3	23.3 ±15.3
Religious Group (no.)			
Hindu	813 (727)	832 (786)	1645 (1513)
Muslim	176 (162)	109 (103)	285 (265)
Controls (no.)	1089 (977)	990 (948)	2079 (1925)
Male	623 (553)	559 (533)	1182 (1086)
Female	466 (424)	431 (415)	897 (839)
Mean age at study encounter ±SD (yr)	26.6 ±16.9 (26.2 ±16.7)	31.6 ±15.9 (31.5 ±15.8)	29.1 ±16.6 (28.9 ±16.5)
Religious Group (no.)			
Hindu	939 (849)	864 (830)	1803 (1679)
Muslim	150 (128)	126 (118)	276 (246)
B. Brazil	**Natal**	**Belém**	**Total**
Families (no.)	244	64	308
Affected individuals (no.)	231	126	357
Male	138	78	216
Female	93	48	141
Mean age at study encounter ±SD (yr)	13 ±14.2	10.7 ±8.9	12.2 ±12.6
Range (yr)	0-72	1-49	0-72
Mean age at onset of VL ±SD (yr)	10.4 ±13.9	8.2 ±8.4	9.6 ±12.2
Total Individuals (no.)	1714	258	1972
Male	776	141	917
Female	938	117	1055
Mean age at study encounter ±SD (yr)	22.5 ±18.7	18.9 ±15.3	22.1 ±18.3
Range (yr)	0-90	1-82	0-90

Supplementary Table 1. Baseline characteristics of samples for (A) the Indian case-control cohorts, and (B) the Brazilian family-based study. [ϕ]Numbers in brackets for India are after removing related individuals with less than 95% IBD0 as determined using a Hidden Markov Model.

			Allele			Indian Discovery			Brazilian discovery			Discovery Meta	Indian Replication			Total Meta
chr	rsID	position	0	1	Freq	P value	Beta	SE	P value	Beta	SE	P value	P value	Beta	SE	P value
2	rs2312548	69809264	A	C	0.55	2.24E-06	0.30	0.06	0.407	-0.08	0.09	5.64E-04	0.811	0.02	0.06	4.74E-03

Supplementary Table 2. Loci outside of the MHC showing evidence of association in Indian discovery but no evidence of association was found for these loci in replication. The Betas and allele frequency are shown for Allele 1.

			Allele			Brazilian discovery			Indian discovery			Discovery Meta	Indian Replication			Total Meta
chr	rsID	position	0	1	Freq	P value	Beta	SE	P value	Beta	SE	P value	P value	Beta	SE	P value
2	rs3748934	100376340	A	G	0.99	4.88E-06	-19.56	4.32	0.014	2.71	1.10	0.205	-	-	-	-
4	rs1355967	10836785	A	C	0.55	1.40E-07	-0.46	0.09	0.169	0.09	0.06	0.057	-	-	-	-
4	rs902174	185892165	A	G	0.73	4.60E-06	-0.42	0.09	0.902	0.01	0.06	0.014	-	-	-	-
6	rs9268878	32539270	A	T	0.31	2.41E-06	0.42	0.09	0.715	-0.03	0.07	9.35E-03	0.578	0.04	0.07	0.016
7	rs2527214	158447342	A	G	0.91	5.98E-06	-0.60	0.13	0.743	-0.03	0.09	5.56E-03	0.525	-0.06	0.09	0.010
10	rs2674355	126052643	A	G	0.04	4.01E-06	0.85	0.19	0.060	-0.46	0.24	0.013	-	-	-	-
11	rs2957710	10334795	A	G	0.17	3.18E-06	0.49	0.11	0.918	0.01	0.06	0.014	-	-	-	-
11	rs1484433	15756036	A	C	0.71	7.62E-06	-0.41	0.09	0.262	-0.08	0.07	4.05E-04	-	-	-	-
11	rs11031947	32691220	A	G	0.33	3.83E-06	0.41	0.09	0.684	-0.03	0.07	0.014	-	-	-	-
11	rs4078355	124436544	A	G	0.75	3.39E-06	0.52	0.11	0.948	0.00	0.07	0.020	0.825	-0.02	0.07	0.097
15	rs1549520	48256580	A	C	0.77	3.06E-06	-0.45	0.10	0.123	0.10	0.06	0.192	0.348	0.07	0.07	0.645
15	rs8036138	83693160	A	G	0.28	8.26E-06	0.41	0.09	-	-	-	-	-	-	-	-
18	rs8098585	36762038	A	G	0.60	6.37E-07	0.46	0.09	0.575	0.04	0.07	7.26E-04	-	-	-	-
18	rs12454166	53780638	A	G	0.09	3.96E-06	0.60	0.13	0.775	-0.03	0.10	9.61E-03	-	-	-	-

Supplementary Table 3. Loci showing evidence of association in Brazilian discovery but no evidence of association was found for these loci in replication. The Betas and allele frequency are shown for Allele 1.

					India discovery		Brazil discovery		Discovery Meta	India replication		Total Meta	
Chr	rsID	Position	Risk	RAF	P value	OR(95%CI)	P value	OR(95%CI)	P value	P value	OR(95%CI)	P value	OR(95%CI)
6	rs9271252	32688266	G	0.67	$1.26x10^{-9}$	1.50 (1.31-1.70)	$2.33x10^{-5}$	1.67 (1.32-2.13)	$1.82x10^{-13}$	-	-	-	-
6	rs9271255	32688335	A	0.67	$1.71x10^{-9}$	1.49 (1.31-1.70)	$2.36x10^{-5}$	1.67 (1.32-2.13)	$2.53x10^{-13}$	$1.48x10^{-5}$	1.35 (1.18-1.55)	$4.48x10^{-17}$	1.46 (1.33-1.59)
6	rs9271842	32702931	A	0.39	$1.49x10^{-10}$	1.50 (1.33-1.70)	$1.94x10^{-4}$	1.38 (1.16-1.64)	$1.47x10^{-13}$	-	-	-	-
6	rs9271858	32703201	G	0.39	$1.57x10^{-10}$	1.50 (1.33-1.70)	$2.01x10^{-4}$	1.38 (1.16-1.64)	$1.60x10^{-13}$	$1.94x10^{-5}$	1.33 (1.17-1.51)	$2.76x10^{-17}$	1.41 (1.30-1.52)
6	rs9272070	32707463	C	0.43	$1.66x10^{-10}$	1.49 (1.32-1.68)	$3.68x10^{-5}$	1.45 (1.21-1.73)	$2.60x10^{-14}$	-	-	-	-
11	rs9300005	29783790	G	0.60	$1.80x10^{-4}$	1.27 (1.12-1.44)	$3.16x10^{-4}$	1.38 (1.16-1.65)	$2.94x10^{-7}$	0.878	1.01 (0.89-1.15)	$9.46x10^{-5}$	1.17 (1.08-1.27)

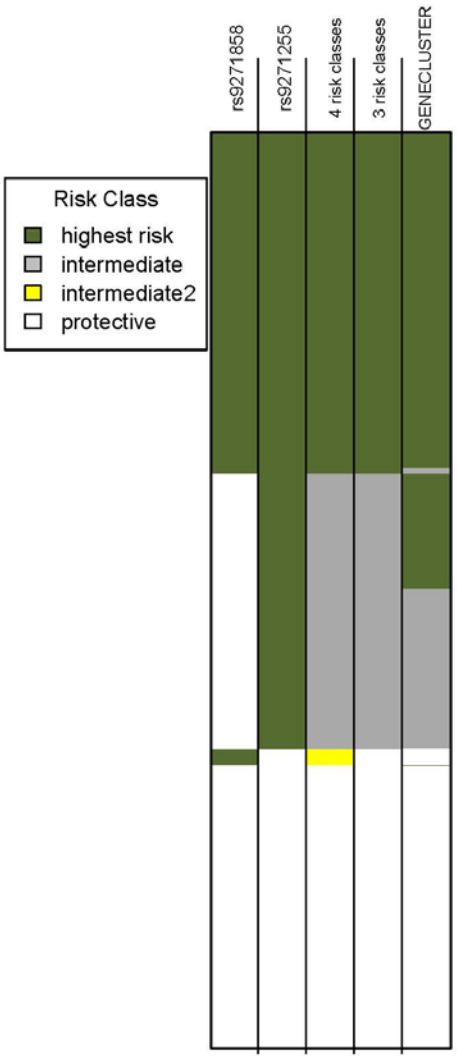
Supplementary Table 4. Association results for SNPs that showed the strongest evidence of association in the combined Indian and Brazilian analysis, and the results where available in the Indian Replication cohort. Analysis was carried out using a multiplicative (on the odds scale) allelic model, with the odds ratio (OR) shown relating to the effect of the risk allele. RAF risk allele frequency, '-' SNP data not available.

DRB1	Indian Discovery				Brazilian Discovery			
	Protective	Other	Total	Frequency	Protective	Other	Total	Frequency
*01	1	0	1	0.006	11	0	11	0.077
*03	0	8	8	0.052	0	10	10	0.07
*04	0	14	14	0.091	0	10	10	0.07
*07	0	30	30	0.195	0	12	12	0.085
*08	0	0	0	0	0	19	19	0.134
*09	0	3	3	0.019	0	0	0	0
*10	0	5	5	0.032	0	5	5	0.035
*11	0	10	10	0.065	0	22	22	0.155
*12	0	6	6	0.039	0	4	4	0.028
*13	0	12	12	0.078	0	28	28	0.197
*14	0	13	13	0.084	0	11	11	0.077
*15	50	0	50	0.325	6	0	6	0.042
*16	2	0	2	0.013	4	0	4	0.028
Total	53	101	154	1	21	121	142	1

Supplementary Table 5. The HLA DRB1 2-digit alleles grouped by whether or not they are carried on the SNP haplotype identified as protective in the Indian data. In both the Indian and Brazilian data the protective risk group is perfectly correlated with DRB1 *15, *16 and *01.

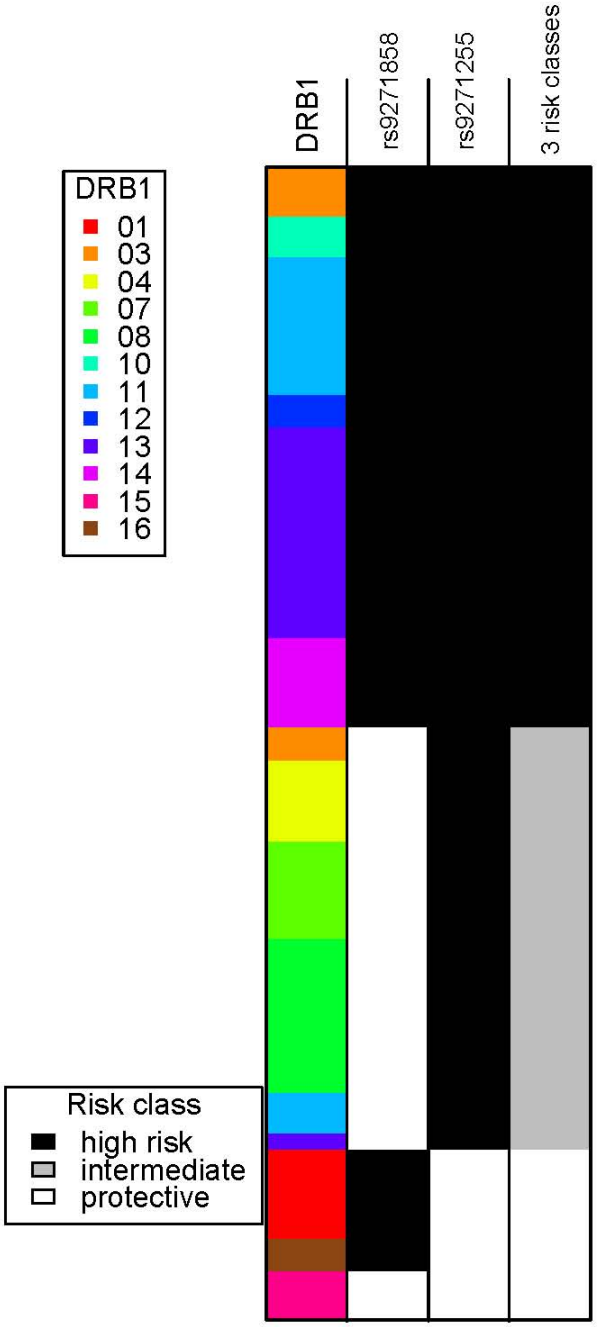
SUPPLEMENTARY FIGURES

Figure S1



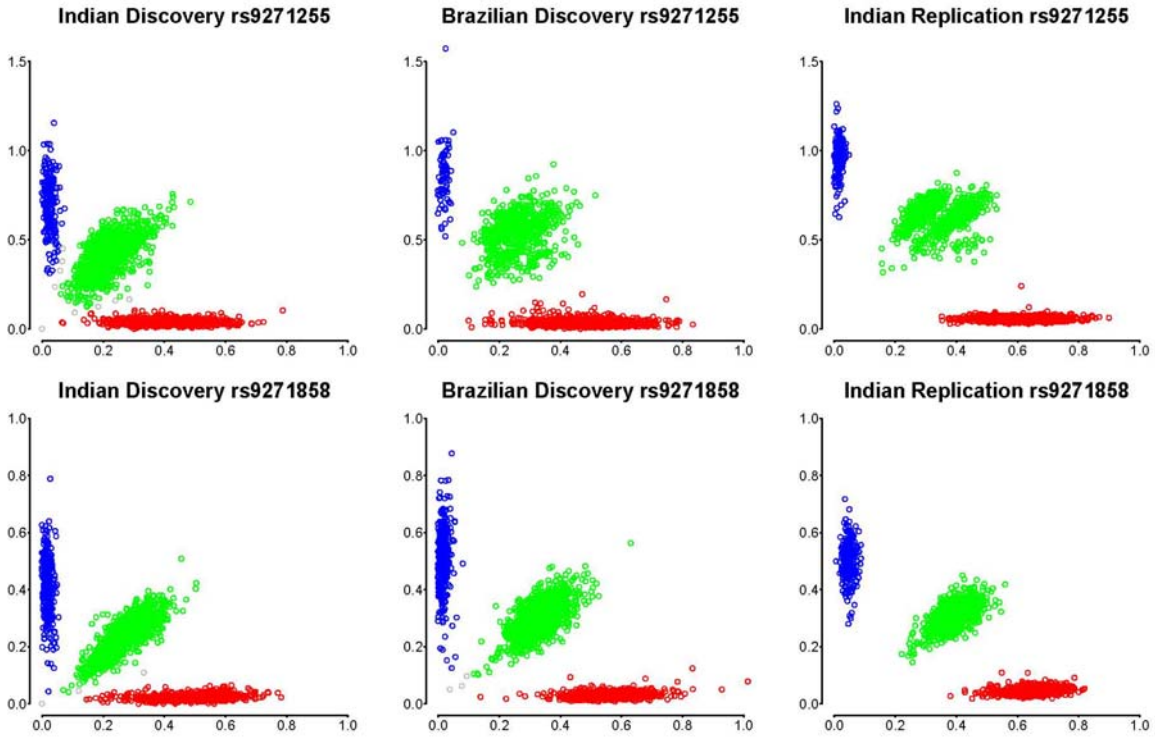
Supplementary Figure 1 Schematic representation of haplotype analyses at the HLA region in the Indian discovery data. The haplotypes were classified using two different models: a GENECLUSTER Bayesian approach and conditional analysis. Each row represents an individual haplotype and each column shows a different haplotype classification model, with the colours denoting which class in the model that individual's haplotype is in. The three and four haplotype models are based on the genotypes at the top SNP and the top conditional SNP, which are shown in the two left-most columns. The three- and the four- risk class model differ only in the haplotypes coloured yellow, which are the rarest class (GG). This rare class does not show a statistically different risk than the protective class (AG), and is thus grouped with the protective in the three-risk model. The GENECLUSTER analysis provided an independent method and also fit a three-risk model to the Indian discovery data. GENECLUSTER implements a Bayesian approach and it found more evidence for two mutations (three risk groups) than for a single mutation (two risk groups).

Figure S2



Supplementary Figure 2 Schematic representation of the HLA and SNP phased 142 Brazilian discovery haplotypes in the HLA region. Each row represents an individual haplotype. The leftmost column shows the HLA alleles at DRB1, colour coded as shown in the top legend. The other three columns show different risk classification models, as described in the main text, with black/grey/white shading denoting which class in the model that haplotype is in. The protective haplotype in the three risk class model correlates perfectly with chromosomes carrying DRB1 *15, *16 and *01.

Figure S3



Supplementary Figure 3 Cluster plots of rs9271858 and rs9271255.
The A and B channel intensities are plotted and coloured by genotype, where missing (no call) genotypes are grey, red and blue are AA and BB homozygotes and green are heterozygotes. There are three plots for each SNP showing from left to right, the genotypes of the Indian discovery individuals, Brazil discovery individuals and the Indian replication. The top row is for rs9271255 and the bottom row for rs9271858.

Supplementary Note

Common variants in the HLA-DRB1-HLA-DQA1 Class II region are associated with susceptibility to visceral leishmaniasis

CONSORTIUM MEMBERSHIPS

Membership of LeishGEN Consortium

Principal Investigators

Jenefer M. Blackwell (Chair)[1,2], Muntaser E. Ibrahim[3], Selma M.B. Jeronimo[4], Shyam Sundar[5], Mary E. Wilson[6]

Key Personnel

Michaela Fakiola[1], Sanjana Mehrotra[5], E. Nancy Miller[1], Anshuman Mishra[5], Hiba Mohamed[3], Gloria R. Monteiro[4], Núbia N. Pontes[4], Madhukar Rai[5], S.P. Singh[5]

Key Collaborators

Frank Christiansen[7], Heather J. Cordell[8], Henio G. Lacerda[9], Linda Smith[7], Olivia Sousa[4], Campbell Witt[7]

[1] Cambridge Institute for Medical Research, University of Cambridge School of Clinical Medicine, Addenbrooke's Hospital, Cambridge CB2 0XY, UK; [2] Telethon Institute for Child Health Research, Centre for Child Health Research, The University of Western Australia, Subiaco, Western Australia 6008, Australia; [3] Endemic Diseases Institute, University of Khartoum, Khartoum, Sudan; [4] Department of Biochemistry, Center for Biosciences, Universidade Federal do Rio Grande do Norte, Natal, RN 59078-970, Brazil; [5] Institute of Medical Sciences, Banaras Hindu University, Varanasi – 221 005, India; [6] Departments of Internal Medicine and Microbiology, University of Iowa and the VA Medical Center, Iowa City, IA, USA 52242, USA; [7] School Pathology and Laboratory Medicine, University of Western Australia, Department of Clinical Immunology, PathWest Laboratory Medicine WA, Royal Perth Hospital & Fremantle Hospital, Wellington Street , Perth, Western Australia 6000, Australia; [8] Institute of Genetic Medicine, Newcastle University, Newcastle upon Tyne NE1 3BZ, UK; [9] Department of Infectious Diseases, Universidade Federal do Rio Grande do Norte, Natal, RN 59078-970, Brazil.

Membership of Wellcome Trust Case Control Consortium 2

Management Committee

Peter Donnelly (Chair)[1,2], Ines Barroso (Deputy Chair)[3], Jenefer M Blackwell[4, 5], Elvira Bramon[6] , Matthew A Brown[7] , Juan P Casas[8] , Aiden Corvin[9], Panos Deloukas[3], Audrey Duncanson[10], Janusz Jankowski[11], Hugh S Markus[12], Christopher G Mathew[13], Colin NA Palmer[14], Robert Plomin[15], Anna Rautanen[1], Stephen J Sawcer[16], Richard C Trembath[13], Ananth C Viswanathan[17], Nicholas W Wood[18]

Data and Analysis Group
Chris C A Spencer[1], Gavin Band[1], Céline Bellenguez[1], Colin Freeman[1], Garrett Hellenthal[1], Eleni Giannoulatou[1], Matti Pirinen[1], Richard Pearson[1], Amy Strange[1], Zhan Su[1], Damjan Vukcevic[1], Peter Donnelly[1,2]

DNA, Genotyping, Data QC and Informatics Group
Cordelia Langford[3], Sarah E Hunt[3], Sarah Edkins[3], Rhian Gwilliam[3], Hannah Blackburn[3], Suzannah J Bumpstead[3], Serge Dronov[3], Matthew Gillman[3], Emma Gray[3], Naomi Hammond[3], Alagurevathi Jayakumar[3], Owen T McCann[3], Jennifer Liddle[3], Simon C Potter[3], Radhi Ravindrarajah[3], Michelle Ricketts[3], Matthew Waller[3], Paul Weston[3], Sara Widaa[3], Pamela Whittaker[3], Ines Barroso[3], Panos Deloukas[3].

Publications Committee
Christopher G Mathew (Chair)[13], Jenefer M Blackwell[4,5], Matthew A Brown[7], Aiden Corvin[9], Chris C A Spencer[1]

[1] Wellcome Trust Centre for Human Genetics, University of Oxford, Roosevelt Drive, Oxford OX3 7BN, UK; [2] Dept Statistics, University of Oxford, Oxford OX1 3TG, UK; [3] Wellcome Trust Sanger Institute, Wellcome Trust Genome Campus, Hinxton, Cambridge CB10 1SA, UK; [4] Telethon Institute for Child Health Research, Centre for Child Health Research, University of Western Australia, 100 Roberts Road, Subiaco, Western Australia 6008; [5] Cambridge Institute for Medical Research, University of Cambridge School of Clinical Medicine, Cambridge CB2 0XY, UK; [6] Department of Psychosis Studies, NIHR Biomedical Research Centre for Mental Health at the Institute of Psychiatry, King's College London and The South London and Maudsley NHS Foundation Trust, Denmark Hill, London SE5 8AF, UK; [7] University of Queensland Diamantina Institute, Brisbane, Queensland, Australia; [8] Dept Epidemiology and Population Health, London School of Hygiene and Tropical Medicine, London WC1E 7HT and Dept Epidemiology and Public Health, University College London WC1E 6BT, UK; [9] Neuropsychiatric Genetics Research Group, Institute of Molecular Medicine, Trinity College Dublin, Dublin 2, Eire; [10] Molecular and Physiological Sciences, The Wellcome Trust, London NW1 2BE; [11] Department of Oncology, Old Road Campus, University of Oxford, Oxford OX3 7DQ, UK , Digestive Diseases Centre, Leicester Royal Infirmary, Leicester LE7 7HH, UK and Centre for Digestive Diseases, Queen Mary University of London, London E1 2AD, UK; [12] Clinical Neurosciences, St George's University of London, London SW17 0RE; [13] King's College London Dept Medical and Molecular Genetics, King's Health Partners, Guy's Hospital, London SE1 9RT, UK; [14] Biomedical Research Centre, Ninewells Hospital and Medical School, Dundee DD1 9SY, UK; [15] King's College London Social, Genetic and Developmental Psychiatry Centre, Institute of Psychiatry, Denmark Hill, London SE5 8AF, UK; [16] University of Cambridge Dept Clinical Neurosciences, Addenbrooke's Hospital, Cambridge CB2 0QQ, UK; [17] NIHR Biomedical Research Centre for Ophthalmology, Moorfields Eye Hospital NHS Foundation Trust and UCL Institute of Ophthalmology, London EC1V 2PD, UK; [18] Dept Molecular Neuroscience, Institute of Neurology, Queen Square, London WC1N 3BG, UK.

Baseline characteristics of the samples. Supplementary Table 1 provides baseline characteristics of the samples used in the Indian discovery and replication studies, and for the Brazilian families.

Further analyses of the SNP signal at the HLA in the Indian data. We tested in the Indian discovery data for departures from the simple genetic model in which risks combine multiplicatively with each copy of the risk allele at the top SNP, rs9271858, and saw no evidence for such departures (p = 0.642 for comparing the two-parameter model with the simple multiplicative model using the mixed-model).

Following conditional analysis two SNPs, rs9271858 and rs9271255, showed evidence of independent association in the HLA. The genotypes at these two SNPs split the data into four risk groups, where the highest risk haplotype is defined by the alleles GA, and the protective haplotype by AG at rs9271858 and rs9271255 respectively (Supplementary Figure 1). Where the protective haplotype (AG) is set as the baseline with OR=1, the highest risk haplotype (GA) has OR(95%CI)=1.71(1.47-2.00) and the two other haplotypes have OR(95%CI)=0.92(0.56-1.51) (GG) and 1.27(1.08-1.49) (AA). In the Indian Immunochip replication data the haplotypes have OR(95%CI)= 1.09(0.69-1.71) (GG) and 1.19(1.01-1.41) (AA) and the highest risk haplotype (GA) has OR(95%CI)=1.50(1.27-1.78).

One of the haplotypes (GG) is very rare, with a frequency of 0.017. It also does not have a significantly different OR from the protective baseline group. We grouped this intermediate haplotype together with the protective haplotype, and tested a model with three risk groups (Supplementary Figure 1). The four risk group model did not fit the data significantly better than the three risk group model in either the discovery (P=0.73) or Indian replication (P=0.71). The model with three risk groups also has a

higher maximised log-likelihood than the two-SNP model (the difference in log-likelihoods is 1.01) and both models have the same number of parameters, thus separating risk into three classes gives the best fit to the data.

HLA allele haplotype analyses: Linkage disequilibrium across this HLA class II region makes it difficult to pinpoint etiological genes, but a natural first step was to focus on the adjacent highly polymorphic classical HLA class II genes, DRB1 and DQB1, that form strong immunological candidates. We selected (see Online Methods) a sample of 112 Indian discovery and 142 Brazilian SNP phased haplotypes along with classical HLA typing at DRB1 and DQB1 to investigate whether the SNP associations correlated with specific HLA alleles. The protective haplotype identified by both GeneCluster (Supplementary Figure 1) and the two SNP haplotype correlated perfectly with DRB1 *15, *16 and *01 in the Indian data (Supplementary Figure 2), suggesting a protective role for these alleles. This was also observed in HLA and SNP phased haplotypes available in a Caucasian dataset (a subset of the 1958 birth cohort), providing a degree of robustness to the finding. In the HLA-typed Indian discovery sample, DRB1*15 is at a high frequency (32.5%), with *16 and *01 at much lower frequencies (1.3% and 0.6% respectively). In the 142 HLA and SNP phased Brazilian haplotypes, the protective haplotype identified in the Indian data also perfectly tagged the HLA DRB1 *15, *16 and *01 2-digit allele groups (Supplementary Table 5, Supplementary Figure 2).