Supplementary Methods

Genotyping. Cases were genotyped for a panel of 2,360 SNPs in the MHC region using the combined MHC Exon-Centric and Mapping panels from Illumina. Full genome profiling of MCIS controls and all children was conducted using the Illumina 660K SNP array or the Illumina ImmunoChip. QA/QC criteria were applied to all genotype data resulting in the exclusion of SNPs that were genotyped in <80% of samples and/or had a minor allele frequency (MAF) <1%. The 80% threshold was set to maximize the number of SNPs available for HLA imputation. Whole genome sequencing (WGS) data was available for ITMI control families. We excluded SNPs that were genotyped in <90% of samples and/or had an MAF <1%. ITMI WGS methods have been previously described [1]. We verified familial relationships for all mother-child duos.

Classical HLA allele imputation. In order to investigate mother-child relationships at classical HLA loci, post-QA/QC genotype data was used to impute markers across the extended MHC using SNP2HLA [2]. Imputation was conducted separately for each chip. SNP2HLA uses the Type 1 Diabetes Genetics Consortium (T1DGC) as the reference. After imputation, we excluded variants with a low measure of imputation accuracy (r^2 <0.3). As an additional QC measure, we compared imputation results at the allelic level to typed two-field *DRB1* data available for a subset of MCIS participants (n=2,136) using the method described by Raychaudhuri et al. [3]. The imputation procedure correctly called 93.1% of alleles. Using a chi-square test and a significance threshold of α=0.05, we compared allele frequencies by platform among controls. We did not find statistically significant differences by platform for the *DRB1* alleles investigated (data not shown).

Allelic frequencies were within the range reported for individuals of European ancestry as published in the online database www.allelefrequencies.net (Download 10/23/2015).

Population substructure. In order to minimize confounding by genetic ancestry, the MCIS recruited mothers who self-identified as non-Hispanic white. ITMI mothers were selected for having >90% European ancestry using markers across the genome and multidimensional scaling analysis conducted in PLINK (version 1.90p) [4] using the Human Genome Diversity Project as the reference panel [5]. Due to differences in coverage depending on genotyping or sequencing platform, we used the software package STRUCTURE (version 2.3.4) [6] to adjust for potential stratification by ancestry in the combined MCIS-ITMI dataset. We used 271 ancestry informative markers (AIMs) available for the majority of participants in the combined dataset. The markers were a subset of markers that differentiate between northern and southern European populations [7]. AIMs were available for 95% of study participants (n=1,239). In our analyses, we did not find evidence of stratification by ancestry. There were no statistically significant differences ($p <0.05$) in *DRB1* allele frequencies between control participants with and without AIM data.

REFERENCES

[1] Bodian DL, Klein E, Iyer RK, Wong WS, Kothiyal P, Stauffer D *et al.* Utility of whole-genome sequencing for detection of newborn screening disorders in a population cohort of 1,696 neonates. Genetics in medicine : official journal of the American College of Medical Genetics, 2015.

[2] Jia X, Han B, Onengut-Gumuscu S, Chen WM, Concannon PJ, Rich SS *et al.* Imputing amino acid polymorphisms in human leukocyte antigens. PLoS One, 2013;8**:**e64683.

[3] Raychaudhuri S, Sandor C, Stahl EA, Freudenberg J, Lee HS, Jia X *et al.* Five amino acids in three HLA proteins explain most of the association between MHC and seropositive rheumatoid arthritis. Nat Genet, 2012;44**:**291-6.

[4] Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, Bender D *et al.* PLINK: a tool set for whole-genome association and population-based linkage analyses. Am J Hum Genet, 2007;81**:**559-75.

[5] Cann HM, de Toma C, Cazes L, Legrand MF, Morel V, Piouffre L *et al.* A human genome diversity cell line panel. Science, 2002;296**:**261-2.

[6] Falush D, Stephens M, Pritchard JK. Inference of population structure using multilocus genotype data: linked loci and correlated allele frequencies. Genetics, 2003;164**:**1567-87.

[7] Barcellos LF, May SL, Ramsay PP, Quach HL, Lane JA, Nititham J *et al.* High-density SNP screening of the major histocompatibility complex in systemic lupus erythematosus demonstrates strong evidence for independent susceptibility regions. PLoS genetics, 2009;5**:**e1000696.

Supplementary Table 1. 1987 American College of Rheumatology Classification Criteria

	RA Cases
1987 ACR criteria, N (%)	**(n=170)**
Morning stiffness	107 (62.9)
Arthritis of 3 or more joint areas	149 (87.7)
Arthritis of hands and joints	168 (98.8)
Symmetric arthritis	159 (93.5)
Rheumatoid nodules	58 (34.1)
Serum rheumatoid factor	117 (68.8)
Radiographic changes	93 (54.7)

Supplementary Table 2. Frequency of the 16 HLA-DRB1 haplotypes defined by amino acids at positions 11, 71, and 74 among rheumatoid arthritis cases and controls in the Mother-Child Immunogenetic Study (MCIS)

Haplotype	DRB1 amino acid position			RA Cases (n=170)		Controls (n=995)		Haplotype frequency (%)		Classical *DRB1* alleles
	11	71	74	N of heterozygotes	N of homozygotes	N of heterozygotes	N of homozygotes	RA	Controls	
1	Val	Lys	Ala	72	13	146	7	28.9	8.7	*04:01
2	Val	Arg	Ala	39	4	114	7	13.9	7.0	*04:08, *04:05, *04:04, *10:01
3	Leu	Arg	Ala	44	3	126	13	14.7	8.3	*01:02, *01:01
4	Pro	Arg	Ala	2	0	18	0	0.6	1.0	*16:01
5	Val	Arg	Glu	0	0	11	0	0.0	0.6	*04:03, *04:07
6	Asp	Arg	Glu	3	0	18	0	0.9	1.0	*09:01
7	Val	Glu	Ala	2	0	24	0	0.6	1.3	*04:02
8	Ser	Lys	Ala	3	0	53	1	0.9	3.0	*13:03
9	Pro	Ala	Ala	32	2	226	28	10.6	15.3	*15:01, *15;02
10	Gly	Arg	Gln	19	1	209	18	6.2	13.3	*07:01
11	Ser	Arg	Ala	14	0	218	11	4.1	13.0	*11:01, *11:04, *12:01
12	Ser	Arg	Glu	5	0	55	1	1.5	3.1	*14:01
13	Leu	Glu	Ala	3	0	14	0	0.9	0.8	*01:03
14	Ser	Arg	Leu	2	0	50	3	0.6	3.0	*08:01, *08:04
15	Ser	Lys	Arg	35	1	221	14	10.9	13.5	*03:01
16	Ser	Glu	Ala	12	2	119	6	4.7	7.1	*11:02, *11:03, *13:01, *13:02

HLA-DRB1 haplotypes according to amino acids at positions 11, 71, and 74 as defined in studies of rheumatoid arthritis [1] [2].

[1] Raychaudhuri S, Sandor C, Stahl EA, Freudenberg J, Lee HS, Jia X *et al.* Five amino acids in three HLA proteins explain most of the association between MHC and seropositive rheumatoid arthritis. Nat Genet, 2012;44:291-6.

[2] Viatte S, Plant D, Han B, Fu B, Yarwood A, Thomson W *et al.* Association of HLA-DRB1 haplotypes with rheumatoid arthritis severity, mortality, and treatment response. Jama, 2015;313:1645-56.

Supplementary Table 3. Case-control association between children carrying 1 or 2 risk alleles and mothers' rheumatoid arthritis status adjusted for mother's carrier status of risk alleles

Risk allele group	**Proportion with children carrying alleles**		**RA (Mother) OR (95% CI)** [1]	***p*-value**
	Cases N=170	Controls N=995		
HLA-DRB1	n (%)	n (%)		
Shared Epitope +	133 (78)	391 (39)	2.8 (1.8-4.4)	<0.001
DERAA +	47 (28)	220 (22)	1.7 (1.1-2.6)	0.03
Valine + (AA position 11)	101 (59)	270 (27)	2.1 (1.4-3.1)	<0.001
Lysine + (AA position 71)	115 (68)	389 (39)	2.4 (1.6-3.5)	<0.001
Alanine + (AA position 74)	163 (96)	863 (87)	3.2 (1.3-7.6)	0.01
HLA-B				
Aspartic acid + (AA position 9)	52 (31)	225 (23)	1.3 (0.8-2.0)	0.34
HLA-DPB1				
Phenylalanine + (AA position 9)	164 (98)	932 (94)	4.5 (1.3-15.7)	0.02

[1] OR = odds ratio; 95% CI = 95% confidence interval for the association between having at least one child (born prior to diagnosis for cases) with one or two alleles (+) of each risk allele group compared to none and RA among mothers. Estimates adjusted for mother's carrier status of the *DRB1* "shared epitope", DERAA, AA position 11, 71, and 74 **haplotype** (16 total), *HLA-B* AA position 9, *HLA-DPB1* AA position 9 and number of live births. AA = amino acid.