Supplemental Data to: Follicular lymphoma protective HLA Class II variants correlate with increased HLA-DQB1 protein expression

Fenna C. M. Sillé¹, Lucia Conde², Jianqing Zhang², Nicholas K. Akers¹, Sylvia Sanchez¹, Joanna Maltbaek¹, Jacques E. Riby², Martyn T. Smith¹, Christine F. Skibola²

¹Division of Environmental Health Sciences, School of Public Health, University of California, Berkeley, CA 94720, USA ²Department of Epidemiology, Comprehensive Cancer Center, University of Alabama, Birmingham, AL 35294, USA

Correspondence: Christine F. Skibola, Department of Epidemiology, Comprehensive Cancer Center, University of Alabama, 1665 University Boulevard, Birmingham, AL 35294-0022, USA; tel 205-934-2950, fax 205-934-8665, email: <u>cskibola@uab.edu</u>

Supplemental Materials and Methods

Cell lines and culture conditions

Lymphoblastoid cells were maintained in T25 flasks in RPMI 1640 (ATCC, Manassas, VA) containing 20% heat-inactivated fetal bovine serum (FBS) at 37°C and 5% CO₂ in a humidified incubator.

Monocyte-derived dendritic cells

Peripheral monocytes were obtained from PBMCs from healthy controls (Astarte Biologics, Redmond, WA and Cellular Technology Limited, Shaker Heights, OH) (aged between 24-67 years, average 43 +/- 12 years), using a Monocyte Isolation Kit II according to manufacturer's recommendations (Miltenyi Biotec, Auburn, CA). Briefly, all non-monocytes were targeted with antibodies conjugated to biotin. Iron microbeads conjugated with streptavidin were used to capture the targeted cells on a magnetized column. Unstained monocytes were recovered in the column flow-through. Enrichment was checked by flow cytometry using CD14-FITC for monocyte detection, whereas nonmonocytes were stained with an anti-biotin antibody conjugated with RPE. The monocyte-enriched fraction was cultured in RPMI 1640 (ATCC, Manassas, VA), supplemented with 10% heat-inactivated FBS, 1% penicillin/streptomycin, 1mM sodium pyruvate, human interleukin-4 (50 ng/mL; BD biosciences, San Jose, CA) and granulocyte-macrophage colony-stimulating factor (50 ng/mL; BD biosciences, San Jose, CA) for 6 days at 37°C with 5% CO₂ in a humidified incubator. The medium was refreshed every two days. The CD11c surface marker was used to identify cultured myeloid dendritic cells by flow cytometry. On day 6, lipopolysaccharide (LPS, 500 ng/mL; Sigma, St. Louis, MO) was added to stimulate monocyte-derived dendritic cells. After 2 days of LPS stimulation, HLA surface expression was determined on activated dendritic cells (CD11c+ CD80+) by flow cytometry. Cell culture supernatants were collected for soluble HLA-DQB1 determination by ELISA and stored at -80°C prior to analysis by ELISA.

TaqMan genotyping

DNA from PBMCs and B-LCLs was extracted using the DNeasy Blood and Tissue Kit (Qiagen, Valencia, CA) and quantified on the NanoDrop ND-1000 spectrophotometer (Thermo Scientific, NanoDrop products, Wilmington, DE). Genotyping was performed using TaqMan SNP Genotyping Assays for rs2647012, rs10484561, rs6457327, rs3135006, rs4947344, rs17496307, rs311722 (Applied Biosystems - Life Technologies, Grand Island, NY) on the Bio-Rad CFX96 system (Bio-Rad, Hercules, CA). Replicate, blinded quality control samples were included to assess reproducibility of the above genotyping procedures. Ambiguous genotypes were re-genotyped as necessary. All plates contained positive (all 3 genotypes) and negative (no template) controls.

Gene expression analysis by RT-qPCR

Total RNA was extracted from B-LCLs using the RNeasy kit (Qiagen, Valencia, CA). cDNA was generated by reverse transcriptase using the High-Capacity cDNA Reverse Transcription kit (Applied Biosystems - Life Technologies, Grand Island, NY). Dendritic cells were washed with 1x PBS and freeze-thawed at -80°C in 10 µL "catch-buffer" (RNase-free H₂0, 0.025% RNasin Plus RNase inhibitor [Promega, Madison, WI] and 0.25% Triton X). Lysed dendritic cells were then subjected to one-step RT-qPCR using the OneStep RT-PCR kit (Qiagen, Valencia, CA). Negative controls were performed without the reverse transcriptase. Duplex TaqMan real-time quantitative PCR was conducted in duplicate or triplicate wells on the BioRad CFX96 system (Bio-Rad, Hercules, CA). TaqMan Gene Expression Master Mix, and TaqMan Gene Expression Assays for human HLA-DOB1 (Hs03054971 m1), HLA-DOA1 (Hs03007426 mH), HLA-DRB1 (Hs04192464 mH), HLA-DPB1 (Hs03045105 m1), HCG22 (Hs00417248 m1), CIITA (Hs00172094 m1), ACTB (Hs99999903 m1) and GAPDH (Hs03929097 g1) were purchased from Applied Biosystems (Life Technologies, Grand Island, NY). Cycling parameters were as follows: 95°C for 10 min, 45 cycles of 95°C for 15 sec, and 60°C for 1 min. Results are presented as relative expression normalized to the ubiquitously expressed *GAPDH* or β -*Actin* to correct for differences in reverse transcription efficiency and were calculated using the $\Delta\Delta$ Ct method.

HLA-DQB1 splicing analysis by RT-PCR

RNA splicing transcripts for *HLA-DQB1* were evaluated by reverse transcription (RT)-PCR. Intron spanning primers were designed to capture exons 2-6 of *HLA-DQB1*, with the ability to detect transcripts without exon 4, without exon 5, or with neither exon (see Supplemental M&M Table I below). cDNA was generated as described above. Cycling parameters on the Bio-Rad CFX96 system were as follows: 98°C for 2 min, 39 cycles of 98°C for 1 sec, and 60°C for 12 sec. PCR products were run with 6x loading buffer (NEB, Ipswich, MA) on 2% Agarose TAE gels together with a 100 bp DNA ladder (NEB, Ipswich, MA), to determine the PCR product size.

Supplemental M&M Table I. Primers used in HLA-DQB1 splicing analysis

Name	Sequence (5'-3')	Location	Exons Amplified (nucleotides)
F1	GAG TGG AGC CCA CAG TGA	Exons 2-3	
R1	TCA GTG CAG AAG CC CT	Exons 6-5/4*	2/3/4/5/6 (434); 2/3/4/6 (410); 2/3/6 (299)
R2	CAG AAG CC CTG CTG GTG GA	Exons 6-5	2/3/4/5/6 (428); 2/3/5/6 (293)
R3	TCA GTG CAG AAG CC CTT TC	Exons 6-4	2/3/4/6 (410)

Cell-based Enzyme-linked immunosorbent assays (ELISA)

Total HLA protein expression was measured by cell-based ELISAs according to manufacturer's instructions (Base Kit 6, R&D Systems, Inc, Minneapolis, MN). In short, cells in their respective media were seeded in provided high-binding 96 well plates. Lymphoblastoid cells were spun down gently at 1200 rpm and then incubated at 37°C for 1 hr. Monocyte-derived DCs were cultured similarly as described in the section above. After incubation, cells were fixed and permeabilized with 8% formaldehyde at 4°C for at least 2 hours and up to 2 weeks. Cells were washed with wash buffer and background fluorescence was quenched with 0.6% H₂O₂. After 3 washes, cells were blocked for a minimum of 1 hour at room temperature. Cells were washed again and incubated overnight at 4°C with a mouse and rabbit primary antibody mixture in blocking buffer (see Supplemental M&M Table II below). After washes, cells were incubated with secondary antibodies in blocking buffer (HRP-conjugated anti-mouse IgG and AP-conjugated anti-rabbit IgG) for 2 hrs at room temperature. After final washes with 1xPBS, fluorogenic substrate F1 (Horse radish peroxidase substrate) was incubated for 20-60 min. after which substrate F2 (Alkaline phosphatase substrate) was added to the

mix. After additional 20-40 min. fluorescence was measured using a fluorescence microplate reader (BioTek FLx-800, BioTek,Winooski, VT) with excitation at 540 nm and emission at 600 nm (for substrate F1) and subsequently with excitation at 360 nm and emission at 450 nm (for substrate F2). HLA expression was measured relative to *GAPDH* [or cell counts measured by crystal violet staining – data not shown].

Soluble HLA-DQB1 detection by ELISA

Secreted HLA-DQB1 was measured in cell culture supernatants from B-LCLs and monocyte-derived DC cultures according to manufacturer's instructions (USCN, Life Science, Inc, Houston, TX). Briefly, stored (-80°C) cell culture supernatants from B-LCL or LPS-activated DCs were thawed and transferred to pre-coated 96-wells assay plates in parallel with serial dilutions of HLA-DQB1 protein standard. Plates were incubated for 2 hrs at 37°C after which, liquid was removed and detection reagent A was added to the plates for 1 hr at 37°C. Wells were washed 3 times with wash solution and incubated for 30 min at 37°C with detection reagent B. Wells were washed again for 5 times and TMB substrate solution was added for 15-25 min followed by stop solution. Absorbance was read immediately at 450nm using a microplate reader (Biotek PowerWave XS, BioTek, Winooski, VT).

Flow cytometry

Single-cell suspensions were diluted in flow cytometry buffer (2% fetal calf serum and 2 mM sodium azide in PBS) and incubated for 20 min with the relevant conjugated antibodies (see Supplemental M&M Table II below) on ice in the presence of anti-FcγRII/III antibody (2.4G2). Cells were fixed with 4% paraformaldehyde for 15 min before analysis. Fluorescence was measured for at least 10,000 cells per sample on a LSR Fortessa apparatus (BD Biosciences). The mean fluorescent intensity (MFI) was analyzed using FlowJo software (Tree Star, Inc., Ashland, OR). Values were normalized against a common sample and replicated experiments were subsequently combined for statistical analysis.

Western blotting

Sub-physiological temperatures can enhance exon 4 exclusion in HLA-DQB1 transcripts and consequently cells were maintained at 37°C for at least 16h prior to cell lysis [Kralovicova J. Human Molecular Genetics. 2004;13(24):3189-3202]. Cells were washed with PBS and lysed at 4°C for 30 min in RIPA buffer [50 mM Tris-HCl [pH 8.0], 150 mM NaCl, 0.2-2% SDS, 0.5% sodium deoxycholate, 1% NP-40, +/- 1% βmercaptoethanol (β -ME)] in the presence of 5 mM phenylmethylsulfonyl fluoride, and protease inhibitors (Complete, EDTA-free; Roche). Insoluble material was removed by centrifugation at 4°C for 15 min at 14,000xg, and protein concentrations were determined using the DC protein assay (Bio-Rad, Hercules, CA). Cellular proteins (50 µg/ sample) were incubated at RT for 30 minutes [loading buffer containing 0.2% SDS]. Protein samples (10 μ g) were resolved by 8-12% SDS polyacrylamide gel electrophoresis (PAGE) and transferred to PVDF membranes (both Bio-Rad, Hercules, CA). Blots were successively blocked with 5% skim milk and 0.05% Tween 20 in PBS for at least 1 h at RT and then probed overnight with mouse monoclonal anti-HLA-DQB1 (1:1000), mouse anti-HLA-DRB (1:1000), rabbit polyclonal anti-HLA-DQA1 (1:100), mouse anti-β-actin antibodies (1:10,000), or stripped and re-probed with mouse anti-human HLA-DPB1 antibody (1:1000) (see Supplemental M&M Table II below). Blots where then incubated with horseradish peroxidase-conjugated anti-mouse IgG or anti-rabbit IgG secondary antibodies (Santa Cruz Biotechnology) at a dilution of 1:5000 and revealed by enhanced chemiluminescence (Amersham, Buckinghamshire, UK).

Antibody	Manufacturer	Host	Clone	Fluorochrome	Assay	Product ID
HLA-DQB1	SCB	Mouse	Genox 3.53	FITC	FC	SC53313
HLA-DQA1	Abcam	mouse	HI118	PerCP	FC	Ab91329
HLA-DR	SCB	mouse	M11C11	PE	FC	SC73366
HLA-DPB1	SCB	mouse	BRAFB6	PE	FC	SC33719
HLA-A/B/C	eBioscience	mouse	W6/32	APC	FC	17-9983-42
HLA-DQA1	LSB	Rabbit	Polyclonal IgG	Unconjugated	WB	LS-C164916
HLA-DQB1	LSB	Mouse	4B7-E2	Unconjugated	WB	LS-B6229
HLA-DQB1	MLS	Mouse	1a3	Unconjugated	WB	P01137F
HLA-DPB1	Sigma	Mouse	6C6	Unconjugated	WB	WH0003115m1
HLA-DPB1	SCB	Mouse	BRAFB6	Unconjugated	WB	SC33719
CIITA	SCB	Mouse	7-1H	Unconjugated	WB	SC13556
β-actin	Abgent	Mouse	8H10D10	Unconjugated	WB	AO1215a
GAPDH	SCB	Rabbit	Polyclonal IgG	Unconjugated	WB	SC25778

Supplemental M&M Table II. Antibodies used in immune-assays

APC, allophycocyanin; APC-H7, allophycocyanin-H7; FITC, fluorescein isothiocyanate; PE, phycoerithrin; PE-Cy5, phycoerithrin-cyanine-5; PE-Cy7, phycoerithrin-cyanine-7; PERCP-Cy5.5, peridinin chlorophyll protein complexcyanine-5.5. FC=flow cytometry; WB=Western blotting; cELISA=cell-based Enzyme-linked immunosorbent assay. Abcam, Cambridge, MA; Abgent, San Diego, CA; SCB=Santa Cruz Biotechnology. Santa Cruz, CA; LSB=LifeSpan Bioscience, Inc. Seattle, WA; MLS=Meridian Life Science, Inc, Memphis, TN.

Bioinformatics

To investigate potential underlying regulatory mechanisms, the SNPs were analyzed for overlap with known and predicted regulatory elements using RegulomeDB (http://regulome.stanford.edu), a database that, based on overlap with known and predicted regulatory elements, classifies SNPs according to their potential functional effect using a tiered category system (1 to 6) where lower scores represent higher regulatory potential [Boyle AP, *et al.* Genome Res. 2012;22(9):1790-1797]. Predicted regulatory elements were visualized using the human genome browser at UCSC [Kent WJ, *et al.* Genome Res. 2002;12(6):996-1006].

Supplemental Table Legends

Table S1. Selected variants associated with follicular lymphoma for further post-GWAS analysis

Selected independent follicular lymphoma associated variants rs2647012, rs10484561, rs6457327, variants in LD, rs3117222, and FL-associated HLA allelotypes. Additionally, 2 surrogate SNPs for rs10484561 were selected, as well as 32 SNPs from our previous eQTL study in LD with rs2647012 that showed the most significant changes for either HLA-DQB1, HLA-DQA1 or HLA-DRB1 expression in either RNAseq data set analyzed [Conde L et al., Am J Hum Genet 2013, 92 (1):126-30]. a/A=minor/Major allele. MAF ca/co = the minor allele frequency for cases / controls. * rs3117222 is in LD with *HLA-DPB1*0301* (r^2 =0.24, D'=0.96) and correlated with increased *HLA-DPB1* expression in our previous meta-analysis [Skibola C *et al.*, BMC Genomics 2012, 13:516]. ** SNPs in LD with rs6457327, but not rs6457327 itself, were associated with *C6orf26* in our previous eQTL study. Highlighted in bold italics are the top eQTLs from our previous study [Conde L *et al.*, Am J Hum Genet 2013, 92 (1):126-30].

Table S2. HLA mRNA transcript levels in B-lymphoblastoid cell lines

HLA mRNA transcript levels in B-lymphoblastoid cell lines were determined by RTqPCR using gene-specific Taqman gene expression assays and calculated using the $\Delta\Delta$ CT method. a/A=minor/Major allele, NA=not available, Rho Corr. Coeff.=Spearman's rank correlation coefficient. In bold are P-values that were deemed significant when P≤2.1x10-3 (after Bonferroni correction for the six independent loci and four main genes (HLA-DQB1, HLA-DQA1, HLA-DRB1, HLA-DQA2) at α =0.05). Highlighted in bold italics are the top eQTLs from our previous study [Conde L *et al.*, Am J Hum Genet 2013, 92 (1):126-30]. HLA-DQB1 amino acids are grouped by amino acids in high LD. HLA-DQB1 alleles within this data set carrying the respective amino acids are listed per group.

Table S3. HLA mRNA transcript levels in LPS-activated monocyte-derived dendritic cells

HLA mRNA transcript levels in LPS-activated monocyte-derived dendritic cells were determined by RT-qPCR using gene-specific Taqman gene expression assays and calculated using the $\Delta\Delta$ CT method. a/A=minor/Major allele, NA=not available, Rho Corr. Coeff.=Spearman's rank correlation coefficient. In bold are P-values that were deemed significant when P≤2.1x10-3 (after Bonferroni correction for the six independent loci and four main genes (HLA-DQB1, HLA-DQA1, HLA-DRB1, HLA-DQA2) at α =0.05). In italics are P-values that are approaching significance (2.1x10⁻³<P≤5.0x10⁻²).

Table S4. Total HLA protein expression in B-lymphoblastoid cell lines measured by cell-based ELISA

Cell-based ELISA using protein-specific antibodies was used to determine HLA-DQB1, HLA-DRB1, HLA-DQA1, HLA-DQA2, HLA-DPB1 and CIITA protein expression in B-lymphoblastoid cell lines. P-values were calculated based on ratios of [target protein]/[GAPDH] that have been normalized to one randomly selected sample. a/A=minor/Major allele, NA=not available, Rho Corr. Coeff.=Spearman's rank correlation coefficient. In bold are P-values that were deemed significant when $P \le 2.1 \times 10^{-3}$ (after Bonferroni correction for the six independent loci and four main genes (HLA-DQB1, HLA-DQA1, HLA-DRB1, HLA-DQA2) at α =0.05). In italics are P-values that are approaching significance ($2.1 \times 10^{-3} < P \le 5.0 \times 10^{-2}$). Highlighted in bold italics are the top eQTLs from our previous study [Conde L *et al.*, Am J Hum Genet 2013, 92 (1):126-30]. HLA-DQB1 amino acids are grouped by amino acids in high LD. HLA-DQB1 alleles within this data set carrying the respective amino acids are listed per group.

Table S5. Surface HLA protein expression on B-lymphoblastoid cell lines measured by flow cytometry

Flow cytometry using protein-specific antibodies was used to determine total HLA (using pan-HLA antibodies), HLA-DQB1, HLA-DRB1, HLA-DQA1, HLA-DPB1 and Class I (using antibodies against HLA-A/B/C) protein expression in B-lymphoblastoid cell lines. P-values were calculated based on MFI data that have been normalized to one randomly selected sample. a/A=minor/Major allele, NA=not available, Rho Corr. Coeff.=Spearman's rank correlation coefficient. In bold are P-values that were deemed

significant when P \leq 2.1x10-3 (after Bonferroni correction for the six independent loci and four main genes (HLA-DQB1, HLA-DQA1, HLA-DRB1, HLA-DQA2) at α =0.05). In italics are P-values that are approaching significance (2.1x10⁻³<P \leq 5.0x10⁻²). Highlighted in bold italics are the top eQTLs from our previous study [Conde L *et al.*, Am J Hum Genet 2013, 92 (1):126-30]. HLA-DQB1 amino acids are grouped by amino acids in high LD. HLA-DQB1 alleles within this data set carrying the respective amino acids are listed per group.

Table S6. Total HLA protein expression in primary LPS-activated monocyte derived dendritic cells measured by cell-based ELISA

Cell-based ELISA using protein-specific antibodies was used to determine HLA-DQB1, HLA-DRB1, HLA-DQA1, HLA-DPB1 and CIITA protein expression in primary LPS-activated monocyte-derived dendritic cells. P-values were calculated based on ratios of [target protein]/[GAPDH] that have been normalized to one randomly selected sample. a/A=minor/Major allele, NA=not available, Rho Corr. Coeff.=Spearman's rank correlation coefficient. In italics are P-values that are approaching significance (2.1x10⁻³ $< P \le 5.0x10^{-2}$).

Table S7. Surface HLA protein expression in primary LPS-activated monocyte derived dendritic cells measured by flow cytometry

Flow cytometry using protein-specific antibodies was used to determine HLA-DQB1, HLA-DRB1, HLA-DQA1, HLA-DQA2, HLA-DPB1 and Class I (using antibodies against HLA-A/B/C) protein expression in primary LPS-activated monocyte-derived dendritic cells. P-values were calculated based on MFI data that have been normalized to one randomly selected sample. a/A=minor/Major allele, NA=not available, Rho Corr. Coeff.=Spearman's rank correlation coefficient. In bold are P-values that were deemed significant ($P \le 2.1 \times 10^{-3}$). In italics are P-values that are approaching significance (2.1x10⁻³).

Table S8. Selected variants associated with follicular lymphoma analyzed with RegulomeDB

SNPs were analyzed for overlap with known and predicted regulatory elements using RegulomeDB (http://regulome.stanford.edu) [Boyle AP, et al. Genome Res. 2012;22(9):1790-1797]. RegulomeDB scores: 1= Likely to affect binding and linked to expression of a gene target, >4= Minimal binding evidence, N/A = Not Available. Highlighted in italics are the top SNPs from our previous eQTL study [Conde L et al., Am J Hum Genet 2013, 92 (1):126-30] and in bold are the SNPs that showed significant differences in protein expression in this study. "+" and "-" means the presence or absence, respectively of the type of regulatory element queried. Chromatin Structure DNase-seq; Chromatin Structure FAIRE = chromatin structure changes based on Formaldehyde-Assisted Isolation of Regulatory Elements; Motifs|Footprinting = transcription factor binding motifs based on DNaseI sensitivity; Motifs|PWM = transcription factor binding motifs based on position weight matrices; Single Nucleotides|X|eQTL = singlenucleotides affecting gene transcription based on expression quantitative trait loci (eQTL) analysis; Chromatin Structure DNase-seq = chromatin structure changes based on DNA-sequencing; Protein Binding|ChIP-seq = protein binding sites based on Chromatin immunoprecipitation (ChIP) followed by high-throughput DNA sequencing (ChIP-seq).

Supplemental Figure legends

Figure S1. HLA expression in B-lymphoblastoid cell lines measured by Western blotting

Western blot analysis of HLA-DQB1, HLA-DQA1, HLA-DRB1 and HLA-DPB1 expression in B-LCLs. Samples are clustered based on rs3135006. 0, 1 and 2 are the number of variant alleles per carrier. Separate blots were probed with (A) mouse antihuman HLA-DQB1 antibody and (B) rabbit polyclonal anti-human HLA-DQA1 antibody. Another blot was probed with (D) mouse anti-human HLA-DRB1 antibody and stripped and re-probed with mouse anti-human HLA-DPB1 antibody (E). Samples were quantitatively analyzed compared to β-actin (C and F).

Figure S2. HLA-DQB1 splicing in B-lymphoblastoid cell lines

HLA-DQB1 splicing was detected by reverse transcription PCR. Data are clustered for rs3135006, which is in high LD with rs2647012 and is one of the SNPs with the most significant effect on transcripts levels. 0,1 and 2 are the number of variant alleles per carrier. F1, R1, R2 and R3 are forward and reverse primers (Supplemental M&M). Expected PCR lengths: F1+R1: 434bp / 410bp / 299bp when full length transcript / no exon 5 / no exon 4, respectively; F1+R2: 428bp / 293bp with or without exon 4, respectively; F1+R3: 410bp without exon 5. Band patterns for F1+R1 indicate all cell lines express transcripts with exon 4 but omitting exon 5, with the exception of GM07055. F1+R2 indicate again that only GM07055 expresses transcripts including exon 5. F1+R3 demonstrate that all cell lines examined express exon 4.

Figure S3. Regulatory elements predicted to overlap with selected FL eQTLs

Displayed are the 40 selected SNPs within the region chr6:31,000,000-33,150,000, with overlapping regulatory elements, using the UCSC genome browser [Kent WJ, *et al*. Genome Res. 2002;12(6):996-1006].

Genes: The gene tracks shown are for protein-coding genes only, based on gene predictions from RefSeq, GenBank, CCDS, Rfam, and the tRNA Genes track. The tracks shown are only protein-coding genes, labeled with their gene symbol. H3K27Ac: The H3K27Ac histone mark is the acetylation of lysine 27 of the H3 histone protein, often found near active regulatory elements) on the B-LCL cell line GM12878 from ENCODE. TFBS: These tracks show regions where transcription factors, proteins responsible for modulating gene transcription, bind to DNA as assayed by ChIP-seq data on the B-LCL cell line GM12878 from ENCODE.



Figure S2



Figure S3

