The Journal of Immunology

Sequence and Phylogenetic Analysis of the Untranslated Promoter Regions for *HLA* Class I Genes

Veron Ramsuran, Pedro G. Hernández-Sanchez, Colm O'hUigin, Gaurav Sharma, Niamh Spence, Danillo G. Augusto, Xiaojiang Gao, Christian A. García-Sepúlveda, Gurvinder Kaur, Narinder K. Mehra and Mary Carrington

J Immunol 2017; 198:2320-2329; Prepublished online 1 February 2017;
doi: 10.4049/jimmunol.1601679
http://www.jimmunol.org/content/198/6/2320

The Journal of Immunology is published twice each month by
The American Association of Immunologists, Inc.,
1451 Rockville Pike, Suite 650, Rockville, MD 20852

Print ISSN: 0022-1767 Online ISSN: 1550-6606.

Sequence and Phylogenetic Analysis of the Untranslated Promoter Regions for *HLA* Class I Genes

Veron Ramsuran,*[†,1] Pedro G. Hernández-Sanchez,*[,‡] Colm O'hUigin,* Gaurav Sharma,[§,¶] Niamh Spence,*[,¶] Danillo G. Augusto,[‖] Xiaojiang Gao,* Christian A. García-Sepúlveda,[‡] Gurvinder Kaur,[§] Narinder K. Mehra,[§] and Mary Carrington*[,†]

Polymorphisms located within the *MHC* have been linked to many disease outcomes by mechanisms not yet fully understood in most cases. Variants located within untranslated regions of *HLA* genes are involved in allele-specific expression and may therefore underlie some of these disease associations. We determined sequences extending nearly 2 kb upstream of the transcription start site for 68 alleles from 57 major lineages of classical *HLA* class I genes. The nucleotide diversity within this promoter segment roughly follows that seen within the coding regions, with *HLA-B* showing the highest (~1.9%), followed by *HLA-A* (~1.8%), and *HLA-C* showing the lowest diversity (~0.9%). Despite its greater diversity, *HLA-B* mRNA expression levels determined in 178 European Americans do not vary in an allele- or lineage-specific manner, unlike the differential expression levels of *HLA-A* or *HLA-C* reported previously. Close proximity of promoter sequences in phylogenetic trees is roughly reflected by similarity of expression pattern for most *HLA-A* and *-C* loci. Although promoter sequence divergence might impact promoter activity, we observed no clear link between the phylogenetic structures as represented by pairwise nucleotide differences in the promoter regions with estimated differences in mRNA expression levels for the classical class I loci. Further, no pair of class I loci showed coordinated expression levels, suggesting that distinct mechanisms across loci determine their expression level under nonstimulated conditions. These data serve as a foundation for more in-depth analysis of the functional consequences of promoter region variation within the classical *HLA* class I loci. ***The Journal of Immunology*****, 2017, 198: 2320–2329.**

The *MHC* is the most polymorphic locus in the human genome (1, 2), and it is also the richest segment genome-wide in terms of disease associations, including bacterial, viral, inflammatory, and autoimmune diseases, cancer, neurologic disorders, and drug hypersensitivity (3). Assigning the mechanism for these disease associations is confounded by the extensive polymorphism, dense gene content with strong linkage disequilibrium between pairs of genes, and functional similarity of the molecules encoded by genes within the *MHC*.

Several studies have indicated that expression levels of certain *HLA* loci vary in an allotype dependent manner (4–16) and this variation has been shown to associate with certain diseases, such as outcome after HIV infection (4, 17), Crohn's disease (4), HBV clearance (8), graft-versus-host disease after unrelated hematopoietic cell transplantation (18, 19), Parkinson's disease (20), systemic lupus erythematosus (21), and certain cancers (22–24). The regulation of *HLA* class I expression is mediated by a series of core promoter motifs, which are located within 500 bp upstream of the start site and are relatively conserved (25). Binding sites for regulatory proteins further upstream from the core promoter have also been identified (26, 27) and it is likely that others exist, some of which may be polymorphic. Indeed, genetic variants and epigenetic modifications located within the 3′ and 5′ untranslated regions (UTRs) (6, 13, 25, 28, 29), as well as the coding region (14), have been shown to alter *HLA* allele expression levels, but so far, these mechanisms account for only a small percentage of the expression variation found within the respective genes.

In this study we provide extensive allelic sequence information for the 5′ UTR and upstream promoter regions of the classical *HLA* class I genes and the homologous region for one nonclassical

*Cancer and Inflammation Program, Leidos Biomedical Research, Inc., Frederick National Laboratory for Cancer Research, Frederick, MD 21702; [†]Ragon Institute of MGH, MIT and Harvard, Cambridge, MA 02139; [‡]Laboratorio de Genómica Viral y Humana, Facultad de Medicina de la Universidad Autónoma de San Luis Potosi, 78210 San Luis Potosi, Mexico; [§]Department of Transplant Immunology and Immunogenetics, All India Institute of Medical Sciences, New Delhi 110029, India; [¶]Laboratory Oncology, Dr. B.R. Ambedkar Institute Rotary Cancer Hospital, All India Institute of Medical Sciences, New Delhi 110029, India; and [‖]Laboratório de Genética Molecular Humana, Departamento de Genética, Universidade Federal do Paraná, Curitiba, CEP 81531-980, Brazil

[1]Current address: Center for the AIDS Program of Research in South Africa, School of Laboratory Medicine and Medical Sciences, College of Health Sciences, University of KwaZulu-Natal, Durban, South Africa.

ORCIDs: 0000-0001-6094-2575 (P.G.H.-S.); 0000-0003-2237-0902 (G.S.); 0000-0002-5191-3614 (N.S.); 0000-0002-1169-7857 (C.A.G.-S.); 0000-0002-1319-7012 (G.K.); 0000-0003-2682-7471 (N.K.M.); 0000-0002-2692-2180 (M.C.).

Received for publication October 5, 2016. Accepted for publication January 2, 2017.

This work was supported by federal funds from the National Cancer Institute, National Institutes of Health (Grants HHSN261200800001E, N02-CP-55504, R01-DA04334, and R01-DA12568). The content of this publication does not necessarily reflect the views or policies of the U.S. Department of Health and Human Services, nor does mention of trade names, commercial products, or organizations imply endorsement by the U.S. Government. This research was supported in part by the Intramural Research Program of the National Institute of Health, Frederick National Laboratory, Center for Cancer Research, a research Flagship grant from the South African Medical Research Council (Grant MRC-RFA-UFSP-01-2013/UKZN HIVEPI), and the Indian Council of Medical Research, Government of India.

The sequences presented in this article have been submitted to GenBank (https://www.ncbi.nlm.nih.gov/genbank/).under accession numbers KY427751–KY427809.

Address correspondence and reprint requests to Dr. Mary Carrington, National Cancer Institute Frederick, Frederick, MD 21702. E-mail address: carringm@mail.nih.gov

The online version of this article contains supplemental material.

Abbreviations used in this article: CV, coefficient of variation; EA, European American; β2M, β-2 microglobulin; qPCR, quantitative PCR; TFBS, transcription factor binding site; TSS, translation start site; UTR, untranslated region.

gene, *HLA-E*. Based on these sequences, the intralocus nucleotide diversity and the phylogenetic congruence between the 5′ UTR promoter and the gene region sequences for each of the three classical class I loci were determined. Expression variation observed among the 57 *HLA* class I lineages was examined in relation to promoter sequences. These data serve as a basis for functional studies aimed at identifying the variants that are involved in differential *HLA* allelic expression levels.

Materials and Methods

Sequencing

Cell lines, homozygous at the *HLA* class I loci, from the Tenth International Histocompatibility Workshop were selected for *HLA* class I promoter sequencing (30). These cell lines represent common alleles that predominantly originate from European American (EA) ancestry. We also selected common African alleles for promoter sequencing from homozygous donors recruited from the Frederick National Laboratory for Cancer Research Donor Program. At least two donors homozygous for a particular allele were selected for study.

Primers specific for each *HLA* gene interrogated (*HLA-A*, *-B*, *-C*, and *-E*) were designed using online Primer 3 software (31). The targeted region covering ∼2 kb upstream of the transcription start site was amplified with locus-specific primers (Supplemental Table I) using the following PCR conditions; 95°C for 5 min, 40 cycles of 95°C for 15 s, 60/63°C for 15 s and 72°C for 1 min, finally 72°C for 1 min. An annealing temperature of either 60 or 63°C was used depending on the locus. The purified PCR products were sequenced in both directions with at least two overlapping primer sets using the ABI-3130XL DNA analyzer (Applied Biosystems).

Sequences were analyzed using Sequencher (Gene Codes) and Mutation Surveyor (Soft Genetics) software and aligned with BioEdit (http://www.mbio.ncsu.edu/bioedit/bioedit.html) and Clustal. All donors homozygous for a specific *HLA* coding region were also homozygous in the 5′ UTR sequence. Previously identified transcription factor binding sites (TFBSs) were highlighted using the BioEdit program.

Putative TFBSs were obtained using the online software Alibaba 2.0 (32) with default settings.

Sequences were deposited to the GenBank database with the accession numbers KY427751–KY427809 (https://www.ncbi.nlm.nih.gov/genbank/).

HLA-B *and* HLA-E *mRNA expression levels using real-time PCR*

Healthy EA donors from the Research Donor Program at the Frederick National Laboratory for Cancer Research were used to measure the *HLA-B* ($n = 178$) and *HLA-E* ($n = 27$) mRNA expression levels using quantitative PCR (qPCR). Primers are located in exon 4 and exons 2–3 for *HLA-B* and *-E*, respectively. *HLA-B*, *-E*, and the housekeeping gene *β-2 microglobulin* (*β2M*) were amplified with the PCR conditions previously described (5, 6). Primer specificity was verified by both melt curve analysis and sequencing of *HLA-B* and *-E* PCR amplicons. PCR amplification efficiencies were tested for all common *HLA-B* ($n = 18$) and *-E* ($n = 2$) allelic lineages. *HLA-B* and *-E* expression levels were normalized to *β2M* and calculated using the $2^{-\Delta\Delta Ct}$ method. Normalizing *HLA-B* and *-E* expression levels by the housekeeping genes *GAPDH*, *18s*, and *β-Actin* show very consistent results to those using *β2M* (data not shown).

Primer-BLAST checks and sequencing of PCR amplicons were performed to verify that only the specific targeted lineage corresponding to each given pair of *HLA-B* and *-E* primers was amplified. To further confirm accuracy of the *HLA-B* primers, an additional set of primers within the exon 5–7 region of the *HLA-B* locus was designed. The PCR amplification efficiencies and *HLA-B* specificity for the exon 5–7 primer set were shown to be equal across *HLA-B* allelic lineages.

Phylogenetic analysis

A comprehensive alignment, including the sequences generated in this study, full length genomic sequences obtained from the IMGT/HLA database (1), and sequences identified by Primer-BLAST searches of nucleotide databases, was constructed for phylogenetic analyses. The completed alignment consisted of up to 5.3 kb, which encompassed promoter (∼2 kb) and coding regions (∼3.3 kb) for 68 alleles from 57 lineages of the three classical *HLA* class I loci. Phylogenetic trees were constructed using the neighbor-joining method (33) based on nucleotide distances calculated by Kimura's 2-parameter method (34) across relevant segments of the alignment. Gaps were excluded on a pairwise basis.

Nucleotide diversity for homologous positions within the promoters of *HLA-A*, *-B*, and *-C* was estimated using Kimura's 2-parameter method (34) and examined by looking at intralocus nucleotide diversity within small (150 bp) windows with sliding window analysis across the alignment as previously described (35). Single representatives from the 57 distinct lineages of *HLA* class I genes were included in the sliding window analysis.

Statistical analyses

The estimated mean expression levels of individual *HLA-B* (178 donors) and *HLA-C* (256 donors) lineages were calculated using the linear model function from the linear regression R package (36) as previously described for *HLA-A* mRNA (6). All donors were of European descent. The magnitude and range of expression differences across the three loci are likely related to consequent phenotypic effects at each *HLA* locus. Although standard units are used to measure intralocus expression differences, comparisons between loci require a normalization procedure to correct for the different experimental measurement procedures developed for each locus. To standardize the measurements across loci, the coefficient of variation (CV) was determined by estimating the ratio of the SD to the mean of expression in a cohort of individuals typed at each of *HLA-A*, *-B*, and *-C* ($n = 172$). The CV can be either directly estimated from each of the 172 individual measurements of expression at each locus or alternatively, it can be estimated by utilizing mean effects of allelic lineages. In the first approach, expression was experimentally determined by qPCR for every individual at each locus allowing direct calculation of means, variances, and CV. For the second approach, expression was inferred by combining the mean effects of allelic lineages that make up the known genotypes of each individual. The individual's expression level is thus determined from the standard expression levels of the alleles making up his/her genotype. The latter approach should be less sensitive to the increases in CV due to experimental errors.

Spearman correlations and Mann–Whitney *U* test were performed using GraphPad Prism software version 6 (GraphPad Software, San Diego, CA).

Results

A region of 1.8–2 kb extending upstream of the translation start site (TSS) was sequenced for a total of 68 alleles from the three *HLA* classical class I genes: 19 *HLA-A*, 33 *HLA-B*, and 16 *HLA-C* alleles. The 68 alleles group into 57 distinct allelic lineages, comprising 17 of 21 known lineages at *HLA-A*, 27 of 36 known lineages at *HLA-B*, and 13 of 16 known lineages at *HLA-C*. Two or more donors homozygous for each allele were sequenced and covered the common allotypes present within individuals of European and African descent. A previous report provided sequences 859 bp upstream of the TSS for the 16 *HLA-C* lineages (37), which overlaps with about half of that provided herein, and these 859 bp sequences are identical to those determined in this study for the same segment of DNA.

Variation in TFBSs of the *HLA* gene core promoter (38–42) was identified between allelic lineages of *HLA-C* (Fig. 1A), where *HLA-C*03*, *-C*07*, and *-C*17*, which were previously identified as having low expression levels (4), possess variation in at least one of the eight *HLA-C* core promoter motifs. Relative to consensus, the *HLA-C*03* lineage contained a single nucleotide change in the TATA binding site, whereas the *HLA-C*07* lineage contained a single nucleotide change within the Enhancer A (κB2) binding site and a deletion polymorphism in the ISRE binding site. The *HLA-C*17* lineage contained single base changes within both the Enhancer A (κB1) and TATA binding sites, and two base changes within the Enhancer A (κB2) binding site. Furthermore, the *HLA-C*17* lineage possesses a 7 bp deletion located at position −37 to −43 bp upstream of the TSS, and an insertion of 6 bp is observed at position −181 to −187 (Fig. 1A). The potential for these variants to explain the low cell surface expression of *HLA-C*03*, *-C*07*, and *-C*17* will require experimental examination.

Among *HLA-A* lineages, the core promoter TFBSs are highly conserved. A single variant located at position −208 bp upstream of the TSS is observed in the Enhancer A (κB1) binding site for the *HLA-A*30* lineage (Fig. 2A). Variation in the *HLA-B* core

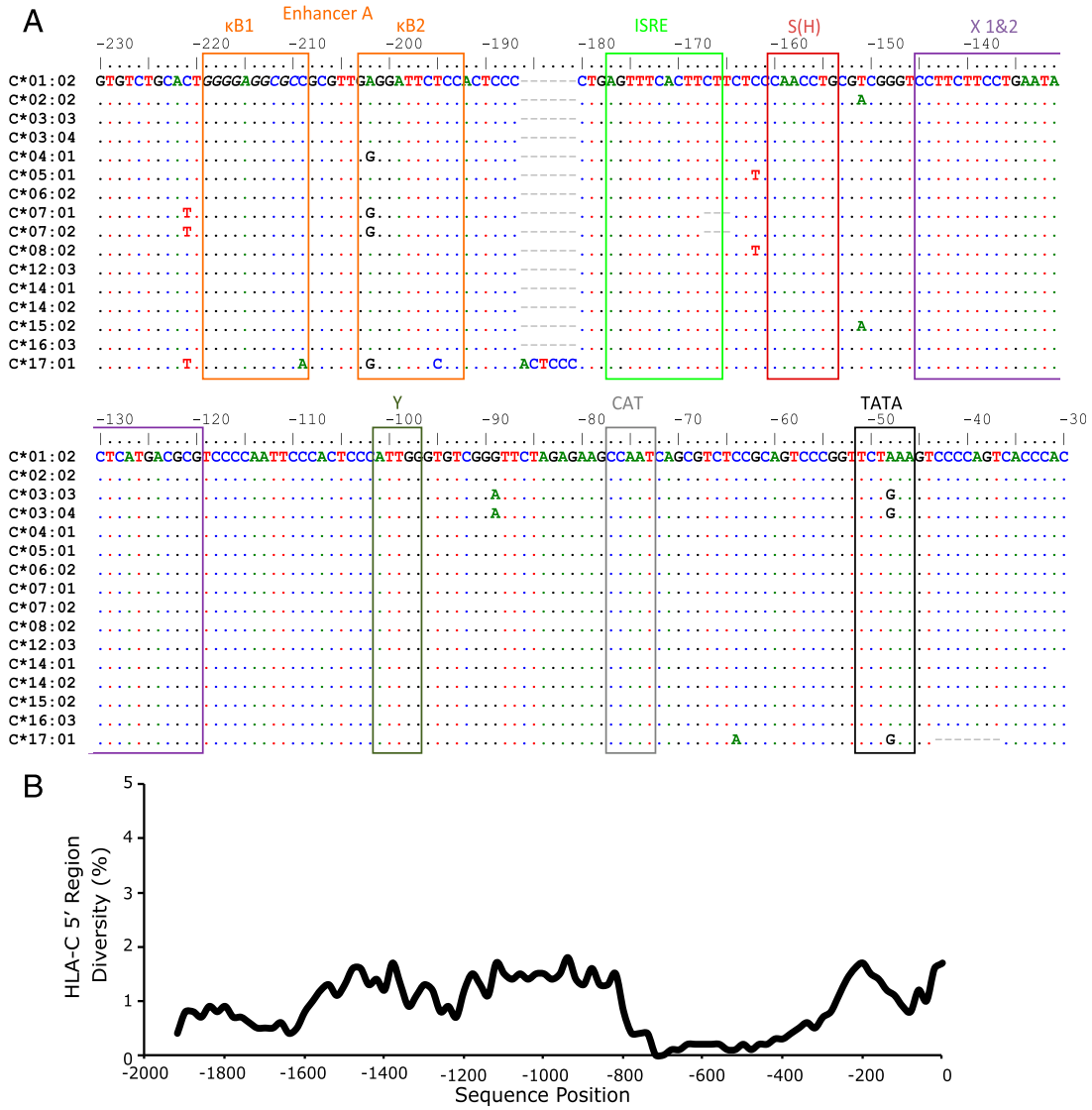


FIGURE 1. Transcription factor binding motifs in the *HLA-C* proximal promoter region and nucleotide diversity across ∼2 kb of the *HLA-C* promoter regions. (**A**) Highlighted regions represent confirmed TFBS motifs based on previously published data (38–42). Sequences for the *HLA-C* loci show a fragment of the total 2 kb upstream region sequenced. Dots represent consensus with the reference sequence, and dashes indicate a deletion. Sites are numbered according to distance from the TSS. (**B**) Intralocus *HLA-C* nucleotide diversity is examined within 150 bp windows using a sliding window analysis of aligned sequences consisting of a single representative of each allelic lineage (n = 13). The x-axis represents the nucleotide position in the alignment relative to the TSS (represented as 0).

promoter TFBSs is also limited with a few nucleotide changes in the ISRE, S (H)-box, and X1/X2 binding sites (Fig. 3A). In general, however, the *HLA-B* promoter region as a whole is more polymorphic than that of *HLA-A* or *-C*, in keeping with the relative polymorphism across their coding regions.

The nucleotide diversity observed within the *MHC* region, ranging upwards of several percent for classical *MHC* genes, has previously been shown to be far higher than the genome average of ∼0.1% (43). This holds true for the regulatory region ∼2 kb upstream of the initiation codon of the classical class I genes, as well, where the intralocus nucleotide diversity across the region for the pool of *HLA-A* (Fig. 2B), *-B* (Fig. 3B), and *-C* (Fig. 1B) lineages is on average ∼1.8, ∼1.9, and ∼0.9% respectively, roughly following the diversity within the coding regions. Divergence across the three loci is also evident, in part due to the presence of insertion elements, which were identified in the University of California Santa Cruz browser (https://genome.ucsc.edu) as long interspersed nuclear elements MLT1E3 and L2 within *HLA-A*, and L1M4a1 within *HLA-B* and *-C*. Surprisingly, the most conserved regions for *HLA-B* and *-C* are not those of the core promoter region where multiple well-characterized TFBSs are positioned, but rather positions −700 to −500 bp, which are more conserved in *HLA-C* than any other 200 bp window examined across the remainder of the promoter region (generally above 1% diversity) (Figs. 1B, 3B). The lowest *HLA-B* diversity of ∼0.2% is located from positions −1720 to −1620 bp (Fig. 3B) and that for *HLA-A* of ∼0.5% is located near −500 bp (Fig. 2B).

The phylogenetic tree of allelic lineages for the *HLA-A* gene constructed from promoter sequences is generally congruent with that found for the remainder of the gene region, indicating that recombination has not scrambled the sequences extensively. Seven promoter clades are identified (Fig. 4A), consisting of *A*02/A*34/A*68*; *A*25/A*26/A*66*; *A*23/A*24*; *A*31/A*33*; *A*29/A*32/A*74; A*01/A*11;* and *A*03/A*30* lineages. These clades are also recovered for the remainder of the gene (Fig. 4B), apart from the *A*34* lineage being located in the *A*25/A*26/A*66* clade. The

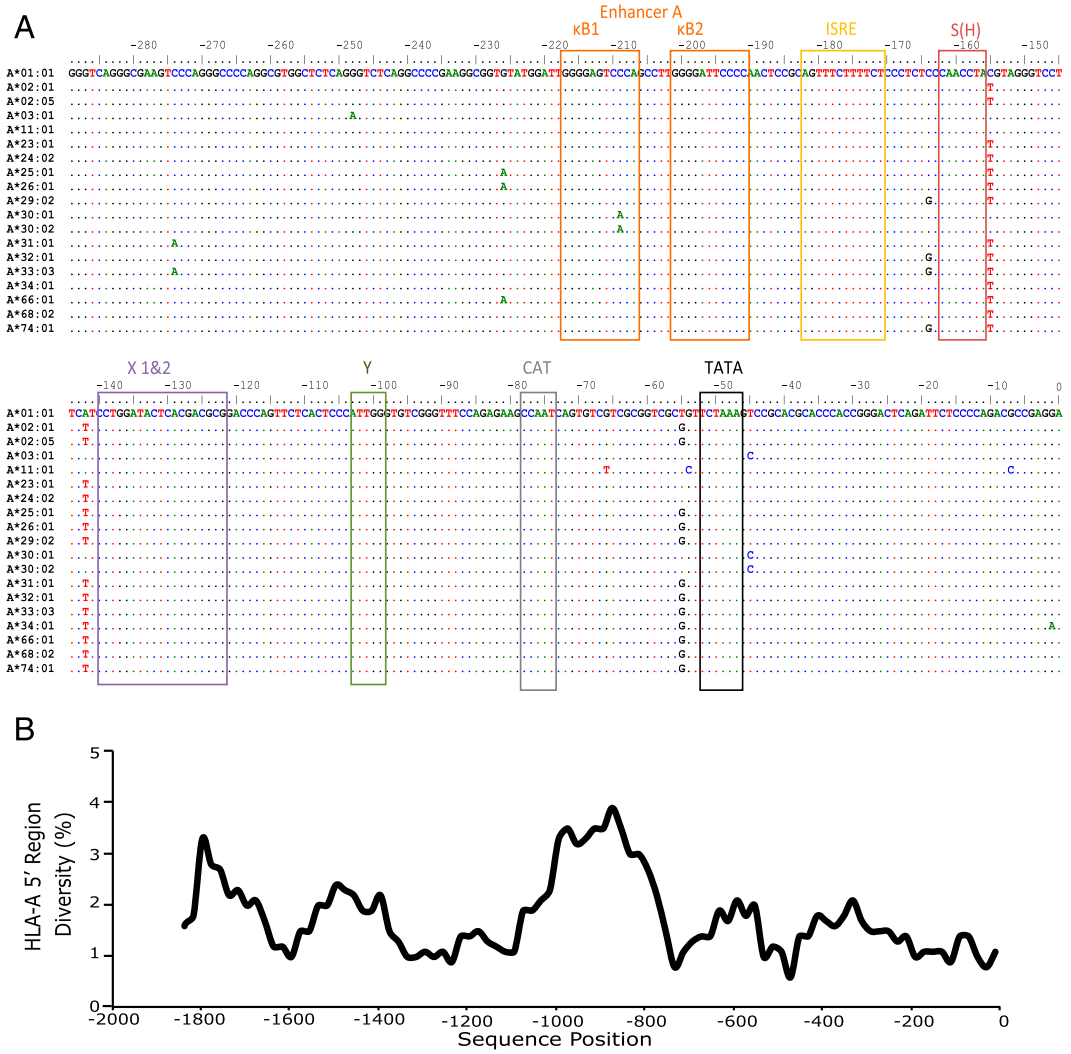


FIGURE 2. Transcription factor binding motifs in the *HLA-A* proximal promoter region and nucleotide diversity across ~2 kb of the *HLA-A* promoter regions. (**A**) Highlighted regions represent confirmed TFBS motifs based on previously published data (38–42). Sequences for the *HLA-A* loci show a fragment of the total 2 kb upstream region sequenced. Dots represent consensus with the reference sequence, and dashes indicate a deletion. Sites are numbered according to distance from the TSS. (**B**) Intralocus *HLA-A* nucleotide diversity is examined within 150 bp windows using a sliding window analysis of aligned sequences consisting of a single representative of each allelic lineage (n = 17). The x-axis represents the nucleotide position in the alignment relative to the TSS (represented as 0).

relative position of the *A*23/A*24* clade also differs between promoter and gene coding segments. Linking previously published *HLA-A* mRNA expression levels (6) with the phylogenetic trees indicated that certain lineages with similar or identical promoter regions possess similar mRNA expression levels, such as *A*23/A*24* and *A*31/A*33* (Fig. 4C), whereas others, such as *A*03* and *A*30*, do not. *A*03* and *A*30* show a 10-fold difference in average lineage expression levels, which may result in part from the eight variable sites observed between the two lineages within the ~2 kbp promoter region (Supplemental Table II).

Nine clades were identified within the HLA-B promoter tree (Fig. 5A): B*35/B*51/B*52/B*53/B*58; B*18/B*27/B*37; B*40/B*41/B*45/B*49/B*50; B*15/B*46; B*55/B*56; B*13/B*44/B*47; B*14/B*38/B*39; B*07/B*08/B*42; and B*57 (the latter being the only lineage-specific clade). The tree derived from coding sequences recovers the B*35/B*51/B*52/B*53/B*58 clade with incorporation of the B*57 lineage into the clade (Fig. 5B). The B*15/B*46 and B*55/B*56 pairings also reflect the promoter clades. However, the promoter-based B*40/B*41/B*45/B*49/B*50 clade splits, where B*40/B*41 groups with B*07/B*08/B*42 in the coding region. Likewise, where the promoter forms a B*13/B*44/B*47 clade, B*27/B*44/B*47 form a clade in coding regions and B*13 forms an isolated tree branch. These and other inconsistencies between the promoter and the coding region trees can be explained by infrequent recombination events that have displaced some, but not all, relationships. A parsimonious reconciliation of the trees requiring the fewest recombination events suggests that lineages sharing inconsistency arose after their common ancestral sequence underwent recombination, giving rise to novel associations across promoter and coding regions.

HLA-B mRNA expression levels were determined in 178 EA donors, and unlike allele-specific differential expression at *HLA-A* and *-C*, we observed very minor differential expression across alleles (Fig. 5C). A fold change of only 1.2 was observed between the CV expression levels of the highest (*B*13*) and the lowest (*B*52*) lineages at *HLA-B* (Fig. 5C, 5D), whereas the fold change within the *HLA-A* and *-C* loci is 3.8 and 2.2, respectively, when comparing the highest and lowest CV for expression levels at each of these loci.

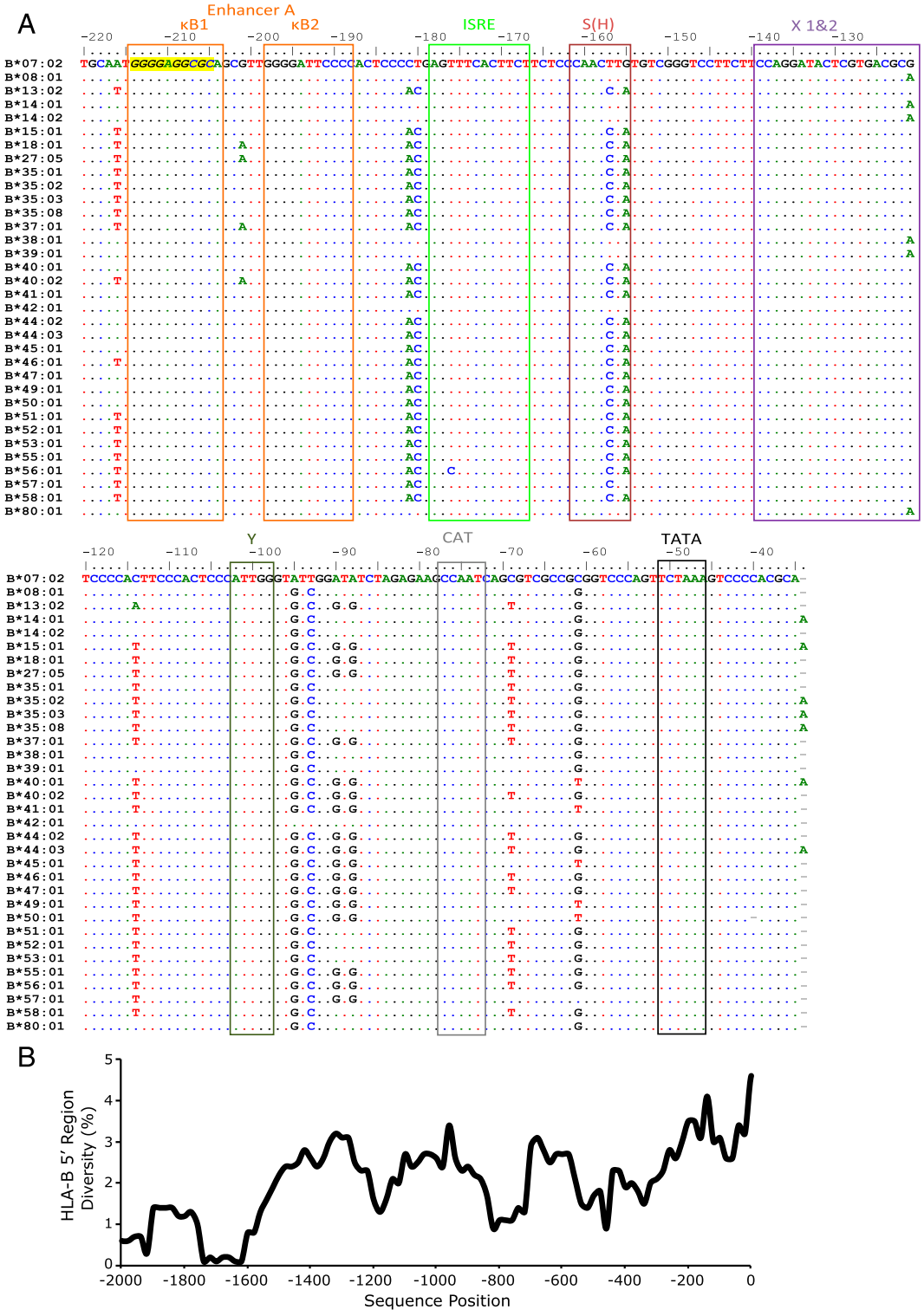
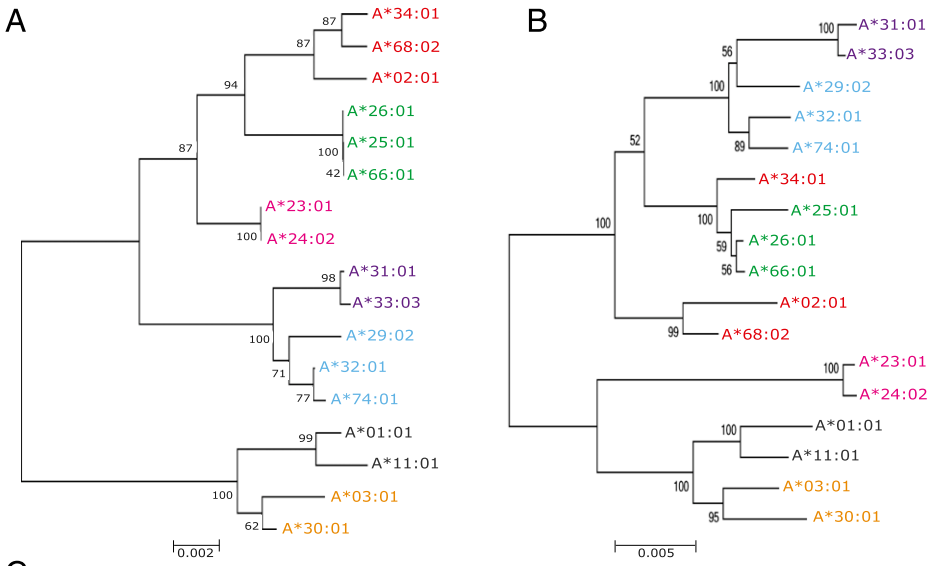


FIGURE 3. Transcription factor binding motifs in the *HLA-B* proximal promoter region and nucleotide diversity across ~2 kb of the *HLA-B* promoter regions. (**A**) Highlighted regions represent confirmed TFBS motifs based on previously published data (38–42). Sequences for the *HLA-B* loci show a fragment of the total 2 kb upstream region sequenced. Dots represent consensus with the reference sequence, dashes indicate a deletion, and yellow highlights represent Sp1 binding site. Sites are numbered according to distance from the TSS. (**B**) Intralocus *HLA-B* nucleotide diversity is examined within 150 bp windows using a sliding window analysis of aligned sequences consisting of a single representative of each allelic lineage (n = 27). The x-axis represents the nucleotide position in the alignment relative to the TSS (represented as 0).

HLA-B mRNA expression levels were measured using a set of qPCR primers that are specific only to *HLA-B* lineages, as determined by sequencing and online Primer-BLAST software (data not shown). The primer set, which targets exon 4, was shown to amplify all common *HLA-B* lineages with equal efficiency (Supplemental Fig. 1A), and no correlation was observed between PCR efficiency of the primers for each *HLA-B* lineage and average expression level of each lineage (Spearman correlation coefficient; $R^2 = 0.22$, $p = 0.36$, Supplemental Fig. 1A). A second set of qPCR primers spanning *HLA-B* exons 5–7 was



C

HLA-A Allelic Lineages	Estimated *HLA-A* expression levels	Standard Error
*A*32*	0.02	0.10
*A*03*	0.04	0.04
*A*31*	0.20	0.08
*A*33*	0.30	0.12
*A*25*	0.44	0.10
*A*30*	0.46	0.10
*A*29*	0.48	0.12
*A*02*	0.54	0.03
*A*26*	0.62	0.08
*A*11*	0.62	0.07
*A*68*	0.88	0.07
*A*01*	0.94	0.04
*A*23*	1.24	0.14
*A*24*	1.26	0.06

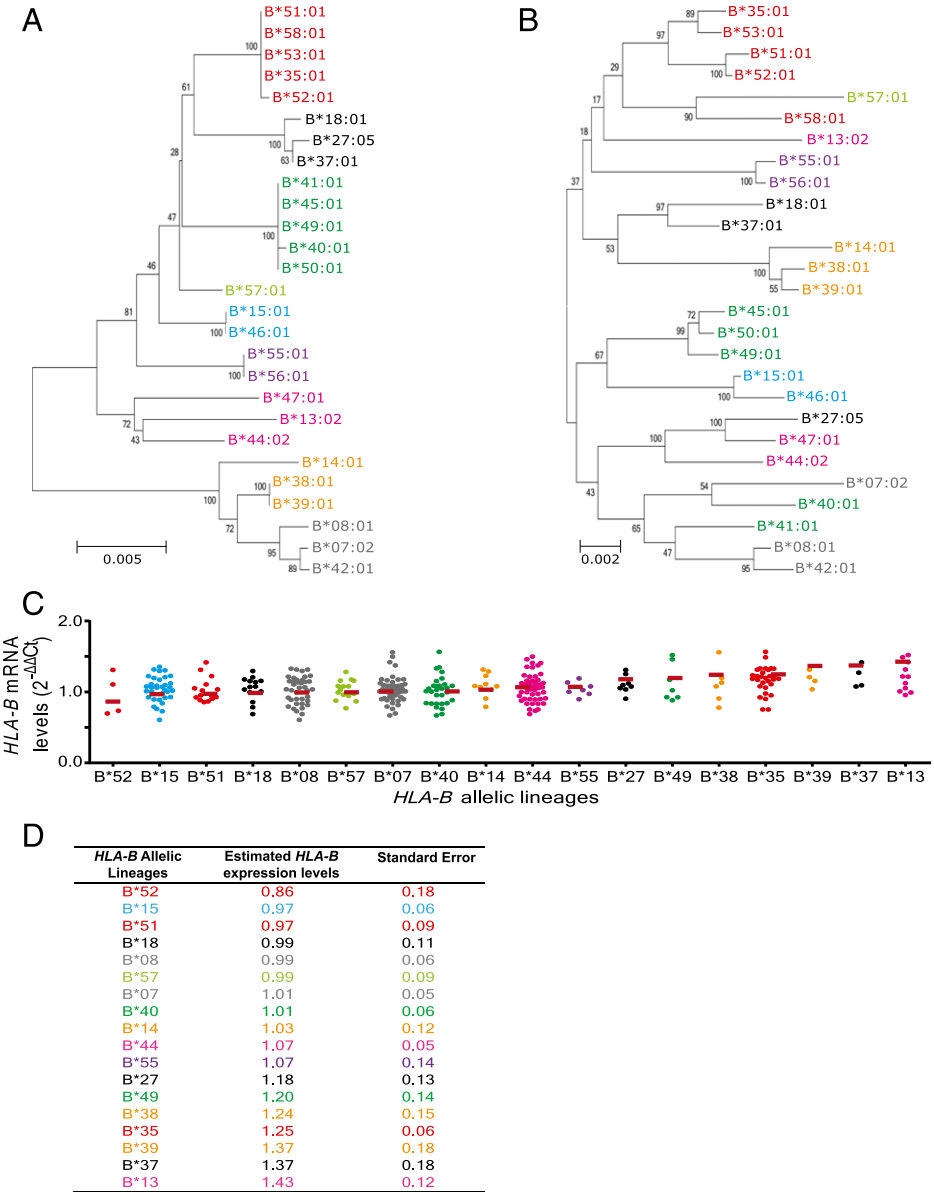
FIGURE 4. Phylogenetic promoter and gene region analysis, and estimated mRNA expression levels for the *HLA-A* gene. Neighbor-joining phylogenetic trees derived from ∼2 kb of promoter region (**A**) and ∼3 kb of gene region (**B**) sequences for 17 lineages within the *HLA-A* gene. Numbers on the nodes represent the bootstrap recovery for that node over 500 replications, and the scale bar represents the number of substitutions along a given branch length. Linear regression estimates of *HLA-A* expression levels for each lineage are calculated from a cohort of 215 EA healthy donors (**C**). Estimated mRNA expression levels and the SE represent each allelic lineage as a diploid homozygote. Clades within the promoter tree are color coded and maintained in the gene region tree and in the estimated expression levels.

generated, which amplified all *HLA-B* lineages with equal efficiency and validated the data obtained using exon 4 primers (data not shown). A significant positive correlation between *HLA-B* mRNA expression levels using exon 4 versus exon 5–7 primer sets was observed for 61 EA healthy donors ($R^2 = 0.81$, $p < 0.0001$, Supplemental Fig. 1B). *HLA-B* mRNA expression levels were also shown to be quite stable in 30 EA donors for whom measurements were repeated 3 y after the initial measurement ($R^2 = 0.72$, $p < 0.0001$, Supplemental Fig. 1C).

The *HLA-C* phylogenetic trees for the promoter region and the coding region show very similar clade structure (Fig. 6A versus 6B), the two main differences involving the *C*03* lineage, which groups with *C*01/C*04/C*14* in the coding region, but forms a distinct (though weakly supported) clade in the promoter region, and *C*02*, which groups weakly with *C*15*in the coding region, but again forms a distinct weakly supported clade in the promoter region. Similar to *HLA-A* (6) and *HLA-B* (Supplemental Fig. 1C), the mRNA expression levels of *HLA-C* are stable across time, as observed in 68 EA donors whose expression was measured 3 y apart (2012 and again in 2015) ($R^2 = 0.62$, $p < 0.0001$, Supplemental Fig. 1D). *C*01* and *C*14*, which are distinguished by only four nucleotides in the promoter region, are the two lineages with the greatest difference in mRNA expression levels (2 fold) within a promoter region clade (Supplemental Table II). These two lineages differ at position −824, where an Sp1 is predicted for the *C*01* lineage (Alibaba 2.0 software (32)), but not the *C*14* lineage, which could explain the difference in their expression levels.

It is notable that despite the allele-specific variation in expression level for *HLA-A* and *HLA-C* mRNA (5, 6), the extent of promoter region variation between pairs of alleles overall does not correlate with differential mRNA expression levels of those alleles. There is no clear link between the phylogenetic structure as represented by pairwise nucleotide differences in the promoter region with estimated differences in mRNA expression levels observed for *HLA-A* (Spearman correlation coefficient; $R^2 = 0.003$, Fig. 7A), *HLA-B* (Spearman correlation coefficient; $R^2 = 0.0007$, Fig. 7B), and *HLA-C* (Spearman correlation coefficient; $R^2 = 0.025$, Fig. 7C). Note that the nonindependence of pairwise comparisons in Fig. 6 (due to shared phylogeny) indicates that the significance values are not valid. Nevertheless, these data suggest that only very few specific variants in the promoter region are likely to account for differential expression across lineages.

Promoter sequences for the *HLA-A*, *-B* and *-C* genes form distinct locus-specific clades (data not shown) in a manner that largely reflects locus specificity of the coding region. The common ancestral locus that gave rise to *MHC-A*, *-B* and *-C* underwent duplication to give rise to two loci, one of which was the ancestral *MHC-A*, whereas the other locus underwent duplication at a later date, giving rise to the ancestral *MHC-B* and *-C* loci (44). These relationships are reflected in the promoter tree where *HLA-A* branches outside of an *HLA-B* and *-C* clade and indicate that interlocus conversion events have not noticeably contributed to the diversity of the promoter. This absence contrasts to findings in the 3′ UTR where interlocus conversion between *HLA-B* and *HLA-C* is strongly implicated in microRNA-induced control of expression (35).



D

HLA-B Allelic Lineages	Estimated *HLA-B* expression levels	Standard Error
B*52	0.86	0.18
B*15	0.97	0.06
B*51	0.97	0.09
B*18	0.99	0.11
B*08	0.99	0.06
B*57	0.99	0.09
B*07	1.01	0.05
B*40	1.01	0.06
B*14	1.03	0.12
B*44	1.07	0.05
B*55	1.07	0.14
B*27	1.18	0.13
B*49	1.20	0.14
B*38	1.24	0.15
B*35	1.25	0.06
B*39	1.37	0.18
B*37	1.37	0.18
B*13	1.43	0.12

FIGURE 5. Phylogenetic promoter and gene region analysis, and measured and estimated mRNA expression levels for the *HLA-B* gene. Neighbor-joining phylogenetic trees derived from ∼2 kb of promoter region (**A**) and ∼3 kb of gene region (**B**) sequences for 27 lineages within *HLA-B* gene. Numbers on the nodes represent the bootstrap recovery for that node over 500 replications, and the scale bar represents the number of substitutions along a given branch length. *HLA-B* mRNA expression levels from 178 EA healthy donors is shown as measured using qPCR (**C**). *HLA-B* mRNA levels were calculated using the $2^{-\Delta\Delta Ct}$ method and *B2M* as the housekeeping control. Relative expression levels are plotted twice for each donor, i.e., once for each lineage present. Lineages found in ≥4 individuals are shown, and the horizontal line represents the estimated *HLA-B* expression levels for each lineage as calculated using linear regression. Estimated mRNA expression levels and the SE represent each allelic lineage as a diploid homozygote (**D**). Clades within the promoter tree are color coded and maintained within the gene region tree, and in the measured and estimated expression levels.

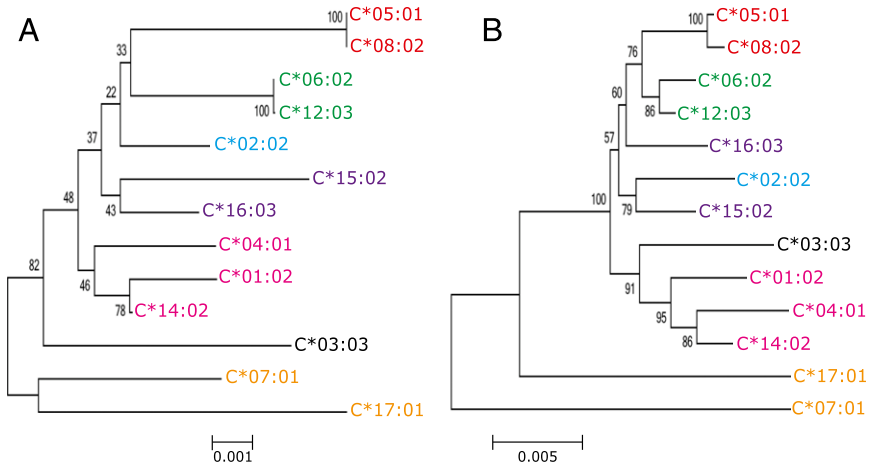
Expression levels of *HLA-A*, *-B*, and *-C* mRNA were measured in 172 EA individuals and tested for correlations between pairs of loci. Very poor correlations in expression levels were detected between the pairs *HLA-A* and *-B* (Spearman correlation coefficient; $R^2 = 0.0002$, $p = 0.87$, Fig. 8A), *HLA-A* and *-C* (Spearman correlation coefficient; $R^2 = 0.02$, $p = 0.05$, Fig. 8B), and *HLA-B* and *-C* (Spearman correlation coefficient; $R^2 = 0.02$, $p = 0.07$, Fig. 8C). These data indicate that unlike the HLA-class II genes, which are regulated by common factors and display strong correlations in expression levels across loci (29, 45), under healthy conditions, the classical class I genes are regulated largely independently, presumably by at least some distinct mechanisms.

The CV for expression of *HLA-A*, *-B*, and *-C* was estimated in these 172 individuals for whom complete three-locus data were available, and used to determine the relative variation in expression for the three loci. Both observed and expected CV, where expected values are based on regression estimates of allelic effects for each genotype, were determined. The CVs increased in order of *HLA-B* (lowest) to *HLA-A* (highest) and the ratio relative to *HLA-B* is 1–2.84 (*HLA-C*) to 3.5 (*HLA-A*) for observed, and 1–4 (*HLA-C*) to 5.6 (*HLA-A*) for expected expression. Observed expression fully incorporates experimentally determined variation, whereas experimental noise can be minimized using expected variation.

No mRNA expression variation was observed between the only two common *HLA-E* alleles, *HLA-E*01:01* and *-E*01:03*, as measured by qPCR in 27 individuals (Supplemental Fig. 1E, $p = 0.95$). The qPCR primers amplified both alleles with virtually the same PCR efficiency of 1.95 and 1.96, respectively. *HLA-E*01:01* and *-E*01:03* differ in the coding region by only a single nonsynonymous nucleotide change, and in keeping with their high sequence similarity, the ∼1.2 kb sequence upstream of the TSS for these alleles was identical (data not shown).

Discussion

Differential *HLA* expression levels have recently been shown to associate with a number of diseases (4, 5, 8, 18, 19, 46), and the sequences provided herein serve as an initial step in identifying novel variants responsible for allele-specific expression variation. Despite substantial variation in the promoter regions of the *HLA* class I genes, which reflects that observed in the coding region of these loci, a Spearman correlation test revealed no strong correlation between



C

HLA-C Allelic Lineages	Estimated *HLA-C* expression levels	Standard Error
C*03	0.70	0.22
C*07	1.63	0.14
C*17	2.15	1.84
C*12	2.18	0.37
C*14	2.27	0.96
C*16	2.31	0.56
C*05	3.20	0.30
C*02	3.22	0.47
C*04	3.41	0.23
C*15	3.76	0.57
C*08	3.87	0.52
C*06	3.95	0.26
C*01	4.47	0.56

FIGURE 6. Phylogenetic promoter and gene region analysis, and estimated mRNA expression levels for the *HLA-C* gene. Neighbor-joining phylogenetic trees derived from ~2 kb of promoter region (**A**) and gene region (**B**) sequences for 13 lineages of the *HLA-C* gene. Numbers on the nodes represent the bootstrap recovery for that node using 500 replications, and the scale bar represents the number of substitutions along a given branch length. Linear regression estimates of *HLA-C* mRNA expression levels for each lineage were calculated from a cohort of 256 EA healthy donors (**C**). Estimated mRNA expression levels and the SE represent each allelic lineage as a diploid homozygote. Clades within the promoter tree are color coded and maintained in the gene region tree and in the estimated expression levels.

similarity in promoter sequence and expression levels between pairs of alleles, suggesting that only a limited number of promoter region variants are likely to impact allele-specific expression variation.

Three *HLA-C* lineages, *HLA-C*03*, *-C*07* and *-C*17*, possess the lowest expression levels and also contain variants within confirmed TFBSs of the core promoter. These three lineages all carry the intact

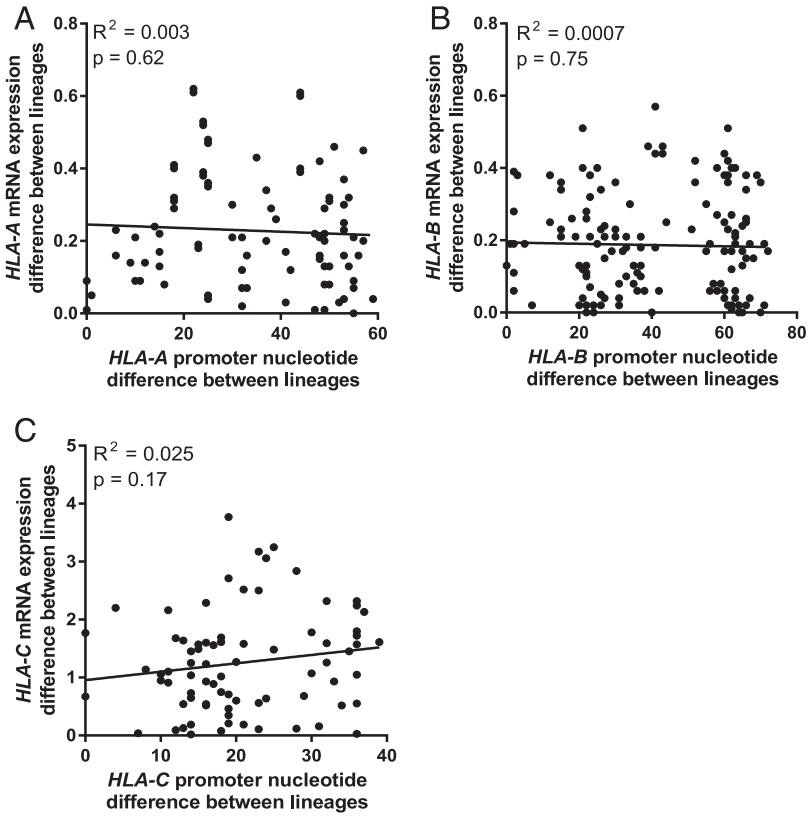


FIGURE 7. The number of nucleotide differences in the promoter region does not correlate with change in mRNA expression level between pairs of lineages at the classical class I loci. The number of promoter region nucleotide differences was tested for a correlation with the difference in expression levels across all pairs of lineages at the (**A**) *HLA-A* (Spearman correlation coefficient; $R^2 = 0.003$, $p = 0.62$), (**B**) *HLA-B* (Spearman correlation coefficient; $R^2 = 0.0007$, $p = 0.75$), and (**C**) *HLA-C* (Spearman correlation coefficient; $R^2 = 0.025$, $p = 0.17$) loci.

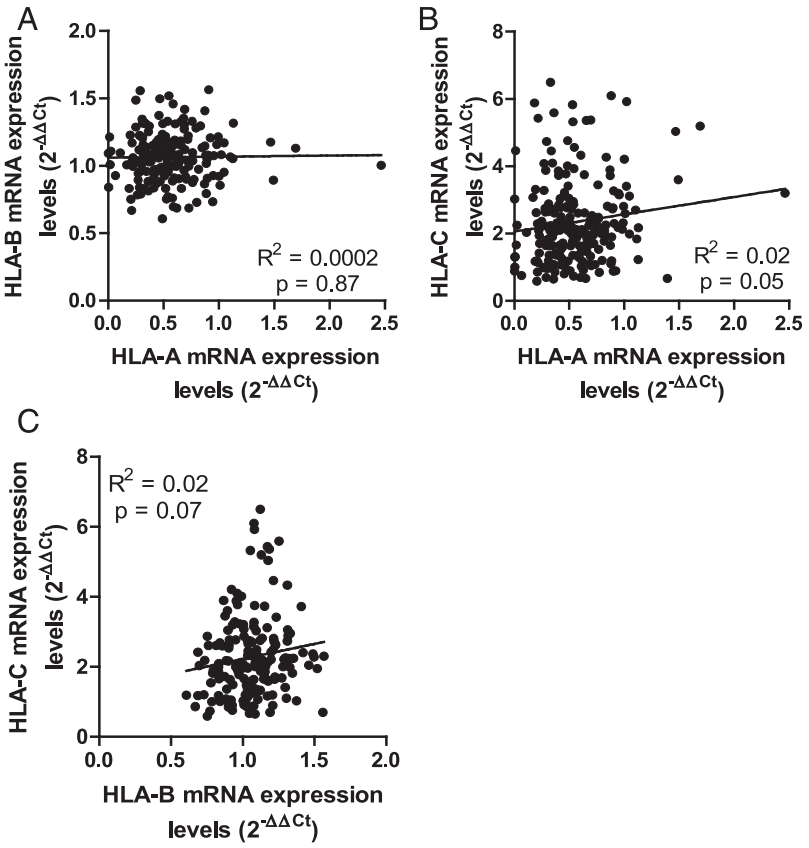


FIGURE 8. Lack of a correlation in mRNA expression levels between pairs of classical class I loci. *HLA-A*, *-B*, and *-C* mRNA expression levels measured in 172 EA donors show no correlation when comparing (**A**) *HLA-A* versus *HLA-B* (Spearman correlation coefficient; $R^2 = 0.0002$, $p = 0.87$), (**B**) *HLA-A* versus *HLA-C* (Spearman correlation coefficient; $R^2 = 0.02$, $p = 0.05$), or (**C**) *HLA-B* versus *HLA-C* (Spearman correlation coefficient; $R^2 = 0.02$, $p = 0.07$).

microRNA-148a binding site in their 3′ UTR, which has been shown to reduce *HLA-C* expression levels (28). It is possible that the disrupted core promoter, along with microRNA regulation of these three alleles, results in their low expression levels relative to other *HLA-C* alleles, highlighting multiple distinct mechanisms determining differential expression of the *HLA-C* alleles.

Despite high levels of nucleotide diversity found in the *HLA-B* promoter region, we observe a low level of *HLA-B* expression variation across donors and find no evidence for allele-specific mRNA expression levels for *HLA-B*, further distinguishing *HLA-B* from *HLA-A* and *-C*. The stability of *HLA-B* mRNA expression levels measured 3 y apart indicates that expression differences across subjects are genuine and are not simply due to measurement fluctuation. The lack of allele-specific expression along with the stability of *HLA-B* mRNA expression levels suggests that *trans*, rather than *cis*-polymorphism may modify *HLA-B* expression. Although we find no evidence for *HLA-B* lineage-specific mRNA expression differences, recent reports suggest that cell surface expression levels of HLA-B do vary in a lineage-specific manner to some extent (47, 48), which may result from lineage-specific translational events. Using three Abs (Tu149, B1.23.2, and 22E1), Chappell et al. (48) showed that the surface expression levels of HLA-B*57, -B*27, -B*07, and -B*35 inversely correlated with promiscuity of peptide binding, where B*57 was expressed at the highest level, followed by B*27, B*07, and B*35 in descending order. A second study used the anti-Bw4 Ab clone 0007 to measure the surface expression of Bw4 alleles, including HLA-B*57 and -B*27 (47). The anti-Bw4 Ab, however, indicated that HLA-B*27 has higher expression than -B*57 (47). These studies were performed with different cell types, which may result in a distinct hierarchy of HLA-B allotype expression levels. As differential cell surface expression levels are likely to impact HLA-B restricted T cell responses, it will be important to pursue this line of study more fully.

Differential allele-specific expression levels have been observed at loci within the MHC, including the *TNF*, *LTA*, *HSP70*, *ZFP57*, and *HLA-DRB1* genes (48–50). This regulation is thought to involve a complex interplay between genetic and epigenetic mechanisms (50), which is also likely to be true for *HLA-A*, where methylation levels are known to associate with differential expression levels of *HLA-A* alleles (6). Interestingly, methylation does not appear to regulate expression levels of *HLA-B* or *-C*.

A highly efficacious vaccine against SIV in rhesus monkeys has been shown to elicit MHC-E–restricted $CD8^+$ T cell responses against a broad array of SIV peptides (49, 50). These responses are likely to explain in part the remarkable protection conferred by this vaccine, sparking interest in considering HLA-E in vaccine design as a mediator of protection against HIV-1. There are only two common alleles of *HLA-E*, *-E*01:01* and *-E*01:03*, which differ by a single amino acid (Gly to Arg at position 107, respectively). *HLA-E*01:03* is expressed at a higher level on the cell surface relative to *HLA-E*01:01* (15), which may affect their functional activity. We found no evidence for allele-specific expression variation in *HLA-E* transcript levels and no difference in the sequence ~1.2 kb upstream of the TSS between these two common lineages, suggesting that any variation in cell surface expression of the allotypes are due to posttranslational modification.

Regulation of *HLA* class I expression levels appears to occur at both the transcriptional and translational level, depending on the locus. The work described herein provides a foundation for greater understanding of the mechanisms responsible for differential regulation of class I allelic expression, which may lead to the ability to manipulate *HLA* expression levels in diseases where class I expression levels are known to associate with outcome.

Acknowledgments

We thank the participants from the Research Donor Program at the Frederick National Laboratory for Cancer Research.

Disclosures

The authors have no financial conflicts of interest.

References

1. Robinson, J., J. A. Halliwell, J. D. Hayhurst, P. Flicek, P. Parham, and S. G. Marsh. 2015. The IPD and IMGT/HLA database: allele variant databases. *Nucleic Acids Res.* 43: D423–D431.
2. Abecasis, G. R., D. Altshuler, A. Auton, L. D. Brooks, R. M. Durbin, R. A. Gibbs, M. E. Hurles, and G. A. McVean, 1000 Genomes Project Consortium. 2010. A map of human genome variation from population-scale sequencing. [Published erratum appears in 2011 *Nature* 473: 544.] *Nature* 467: 1061–1073.
3. Welter, D., J. MacArthur, J. Morales, T. Burdett, P. Hall, H. Junkins, A. Klemm, P. Flicek, T. Manolio, L. Hindorff, and H. Parkinson. 2014. The NHGRI GWAS catalog, a curated resource of SNP-trait associations. *Nucleic Acids Res.* 42: D1001–D1006.
4. Apps, R., Y. Qi, J. M. Carlson, H. Chen, X. Gao, R. Thomas, Y. Yuki, G. Q. Del Prete, P. Goulder, Z. L. Brumme, et al. 2013. Influence of HLA-C expression level on HIV control. *Science* 340: 87–91.
5. Kulkarni, S., Y. Qi, C. O'HUigin, F. Pereyra, V. Ramsuran, P. McLaren, J. Fellay, G. Nelson, H. Chen, W. Liao, et al. 2013. Genetic interplay between HLA-C and MIR148A in HIV control and Crohn disease. *Proc. Natl. Acad. Sci. USA* 110: 20705–20710.
6. Ramsuran, V., S. Kulkarni, C. O'huigin, Y. Yuki, D. G. Augusto, X. Gao, and M. Carrington. 2015. Epigenetic regulation of differential HLA-A allelic expression levels. *Hum. Mol. Genet.* 24: 4268–4275.
7. Thomas, R., R. Apps, Y. Qi, X. Gao, V. Male, C. O'hUigin, G. O'Connor, D. Ge, J. Fellay, J. N. Martin, et al. 2009. HLA-C cell surface expression and control of HIV/AIDS correlate with a variant upstream of HLA-C. *Nat. Genet.* 41: 1290–1294.
8. Thomas, R., C. L. Thio, R. Apps, Y. Qi, X. Gao, D. Marti, J. L. Stein, K. A. Soderberg, M. A. Moody, J. J. Goedert, et al. 2012. A novel variant marking HLA-DP expression levels predicts recovery from hepatitis B virus infection. *J. Virol.* 86: 6979–6985.
9. Isa, A., J. O. Nehlin, H. J. Sabir, T. E. Andersen, M. Gaster, M. Kassem, and T. Barington. 2010. Impaired cell surface expression of HLA-B antigens on mesenchymal stem cells and muscle cell progenitors. *PLoS One* 5: e10900.
10. Mulrooney, T. J., A. C. Zhang, Y. Goldgur, J. E. Boudreau, and K. C. Hsu. 2015. KIR3DS1-specific D0 domain polymorphisms disrupt KIR3DL1 surface expression and HLA binding. *J. Immunol.* 195: 1242–1250.
11. Ferstl, B., T. Zacher, B. Lauer, N. Blagitko-Dorfs, A. Carl, and R. Wassmuth. 2004. Allele-specific quantification of HLA-DQB1 gene expression by real-time reverse transcriptase-polymerase chain reaction. *Genes Immun.* 5: 405–416.
12. Nepom, G. T., J. Chung, and K. A. West. 1995. Differential expression of HLA-DQB1 alleles in a heterozygous cell line. *Immunogenetics* 42: 143–148.
13. Beaty, J. S., K. A. West, and G. T. Nepom. 1995. Functional effects of a natural polymorphism in the transcriptional regulatory sequence of HLA-DQB1. *Mol. Cell. Biol.* 15: 4771–4782.
14. Dellgren, C., J. O. Nehlin, and T. Barington. 2015. Cell surface expression level variation between two common human leukocyte antigen alleles, HLA-A2 and HLA-B8, is dependent on the structure of the C terminal part of the alpha 2 and the alpha 3 domains. *PLoS One* 10: e0135385.
15. Strong, R. K., M. A. Holmes, P. Li, L. Braun, N. Lee, and D. E. Geraghty. 2003. HLA-E allelic variants. Correlating differential expression, peptide affinities, crystal structures, and thermal stabilities. *J. Biol. Chem.* 278: 5082–5090.
16. Hviid, T. V., S. Hylenius, C. Rørbye, and L. G. Nielsen. 2003. HLA-G allelic variants are associated with differences in the HLA-G mRNA isoform profile and HLA-G mRNA levels. *Immunogenetics* 55: 63–79.
17. Blais, M. E., Y. Zhang, T. Rostron, H. Griffin, S. Taylor, K. Xu, H. Yan, H. Wu, I. James, M. John, et al. 2012. High frequency of HIV mutations associated with HLA-C suggests enhanced HLA-C-restricted CTL selective pressure associated with an AIDS-protective polymorphism. *J. Immunol.* 188: 4663–4670.
18. Petersdorf, E. W., T. A. Gooley, M. Malkki, A. P. Bacigalupo, A. Cesbron, E. Du Toit, G. Ehninger, T. Egeland, G. F. Fischer, T. Gervais, et al. 2014. HLA-C expression levels define permissible mismatches in hematopoietic cell transplantation. *Blood* 124: 3996–4003.
19. Petersdorf, E. W., M. Malkki, C. O'hUigin, M. Carrington, T. Gooley, M. D. Haagenson, M. M. Horowitz, S. R. Spellman, T. Wang, and P. Stevenson. 2015. High HLA-DP expression and graft-versus-host disease. *N. Engl. J. Med.* 373: 599–609.
20. Wissemann, W. T., E. M. Hill-Burns, C. P. Zabetian, S. A. Factor, N. Patsopoulos, B. Hoglund, C. Holcomb, R. J. Donahue, G. Thomson, H. Erlich, and H. Payami. 2013. Association of Parkinson disease with structural and regulatory variants in the HLA region. *Am. J. Hum. Genet.* 93: 984–993.
21. Raj, P., E. Rai, R. Song, S. Khan, B. E. Wakeland, K. Viswanathan, C. Arana, C. Liang, B. Zhang, I. Dozmorov, et al. 2016. Regulatory polymorphisms modulate the expression of HLA class II molecules and promote autoimmunity. *ELife* 5: e12089–e12141.
22. Hiraki, A., N. Fujii, T. Murakami, K. Kiura, K. Aoe, H. Yamane, K. Masuda, T. Maeda, K. Sugi, Z. Darzynkiewicz, et al. 2004. High frequency of allele-specific down-regulation of HLA class I expression in lung cancer cell lines. *Anticancer Res.* 24(3a): 1525–1528.
23. Maeurer, M. J., S. M. Gollin, W. J. Storkus, W. Swaney, J. Karbach, D. Martin, C. Castelli, R. Salter, A. Knuth, and M. T. Lotze. 1996. Tumor escape from immune recognition: loss of HLA-A2 melanoma cell surface expression is associated with a complex rearrangement of the short arm of chromosome 6. *Clin. Cancer Res.* 2: 641–652.
24. Ibrahim, E. C., Y. Allory, F. Commo, B. Gattegno, P. Callard, and P. Paul. 2003. Altered pattern of major histocompatibility complex expression in renal carcinoma: tumor-specific expression of the nonclassical human leukocyte antigen-G molecule is restricted to clear cell carcinoma while up-regulation of other major histocompatibility complex antigens is primarily distributed in all subtypes of renal carcinoma. *Am. J. Pathol.* 162: 501–508.
25. Cereb, N., and S. Y. Yang. 1994. The regulatory complex of HLA class I promoters exhibits locus-specific conservation with limited allelic variation. *J. Immunol.* 152: 3873–3883.
26. Howcroft, T. K., J. D. Weissman, A. Gegonne, and D. S. Singer. 2005. A T lymphocyte-specific transcription complex containing RUNX1 activates MHC class I expression. *J. Immunol.* 174: 2106–2115.
27. L'Haridon, M., P. Paul, J. G. Xerri, H. Dastot, C. Dolliger, M. Schmid, N. de Angelis, L. Grollet, F. Sigaux, L. Degos, and C. Gazin. 1996. Transcriptional regulation of the MHC class I HLA-A11 promoter by the zinc finger protein ZFX. *Nucleic Acids Res.* 24: 1928–1935.
28. Kulkarni, S., R. Savan, Y. Qi, X. Gao, Y. Yuki, S. E. Bass, M. P. Martin, P. Hunt, S. G. Deeks, A. Telenti, et al. 2011. Differential microRNA regulation of HLA-C expression and its association with HIV control. *Nature* 472: 495–498.
29. Boss, J. M. 1997. Regulation of transcription of MHC class II genes. *Curr. Opin. Immunol.* 9: 107–113.
30. Prasad, V. K., and S. Y. Yang. 1996. Allele assignment for HLA-A, -B, and -C genes to the tenth International Histocompatibility Workshop cell lines. *Tissue Antigens* 47: 538–546.
31. Untergasser, A., I. Cutcutache, T. Koressaar, J. Ye, B. C. Faircloth, M. Remm, and S. G. Rozen. 2012. Primer3–new capabilities and interfaces. *Nucleic Acids Res.* 40: e115.
32. Grabe, N. 2002. AliBaba2: context specific identification of transcription factor binding sites. *In Silico Biol.* 2: S1–S15.
33. Saitou, N., and M. Nei. 1987. The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol. Biol. Evol.* 4: 406–425.
34. Kimura, M. 1980. A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. *J. Mol. Evol.* 16: 111–120.
35. O'huigin, C., S. Kulkarni, Y. Xu, Z. Deng, J. Kidd, K. Kidd, X. Gao, and M. Carrington. 2011. The molecular origin and consequences of escape from miRNA regulation by HLA-C alleles. *Am. J. Hum. Genet.* 89: 424–431.
36. R Development Core Team. 2008. *R: A Language and Environment for Statistical Computing.* R Foundation for Statistical Computing, Vienna, Austria. Available at: https://www.r-project.org/. Accessed: February 11, 2015.
37. Xu, Y., Z. Deng, C. O'hUigin, D. Wang, S. Gao, J. Zeng, B. Yang, S. Jin, and H. Zou. 2011. Characterization and polymorphic analysis of 4.5 kb genomic full-length HLA-C in the Chinese Han population. *Tissue Antigens* 78: 102–114.
38. Johnson, D. R. 2003. Locus-specific constitutive and cytokine-induced HLA class I gene expression. *J. Immunol.* 170: 1894–1902.
39. van den Elsen, P. J., T. M. Holling, H. F. Kuipers, and N. van der Stoep. 2004. Transcriptional regulation of antigen presentation. *Curr. Opin. Immunol.* 16: 67–75.
40. Peltenburg, L. T., and P. I. Schrier. 1994. Transcriptional suppression of HLA-B expression by c-Myc is mediated through the core promoter elements. *Immunogenetics* 40: 54–61.
41. Gobin, S. J., V. Keijsers, M. van Zutphen, and P. J. van den Elsen. 1998. The role of enhancer A in the locus-specific transactivation of classical and nonclassical HLA class I genes by nuclear factor kappa B. *J. Immunol.* 161: 2276–2283.
42. Gobin, S. J., M. van Zutphen, A. M. Woltman, and P. J. van den Elsen. 1999. Transactivation of classical and nonclassical HLA class I genes through the IFN-stimulated response element. *J. Immunol.* 163: 1428–1434.
43. Horton, R., R. Gibson, P. Coggill, M. Miretti, R. J. Allcock, J. Almeida, S. Forbes, J. G. Gilbert, K. Halls, J. L. Harrow, et al. 2008. Variation analysis and gene annotation of eight MHC haplotypes: the MHC Haplotype Project. *Immunogenetics* 60: 1–18.
44. Boyson, J. E., C. Shufflebotham, L. F. Cadavid, J. A. Urvater, L. A. Knapp, A. L. Hughes, and D. I. Watkins. 1996. The MHC class I genes of the rhesus monkey. Different evolutionary histories of MHC class I and II genes in primates. *J. Immunol.* 156: 4656–4665.
45. Swanberg, M., O. Lidman, L. Padyukov, P. Eriksson, E. Akesson, M. Jagodic, A. Lobell, M. Khademi, O. Börjesson, C. M. Lindgren, et al. 2005. MHC2TA is associated with differential MHC molecule expression and susceptibility to rheumatoid arthritis, multiple sclerosis and myocardial infarction. *Nat. Genet.* 37: 486–494.
46. Wallny, H. J., D. Avila, L. G. Hunt, T. J. Powell, P. Riegert, J. Salomonsen, K. Skjødt, O. Vainio, F. Vilbois, M. V. Wiles, and J. Kaufman. 2006. Peptide motifs of the single dominantly expressed class I molecule explain the striking MHC-determined response to Rous sarcoma virus in chickens. *Proc. Natl. Acad. Sci. USA* 103: 1434–1439.
47. Boudreau, J. E., T. J. Mulrooney, J. B. Le Luduec, E. Barker, and K. C. Hsu. 2016. KIR3DL1 and HLA-B density and binding calibrate NK education and response to HIV. *J. Immunol.* 196: 3398–3410.
48. Chappell, P., K. Meziane, M. Harrison, Ł. Magiera, C. Hermann, L. Mears, A. G. Wrobel, C. Durant, L. L. Nielsen, S. Buus, et al. 2015. Expression levels of MHC class I molecules are inversely correlated with promiscuity of peptide binding. *ELife* 4: e05345–e05367.
49. Hansen, S. G., M. Piatak, Jr., A. B. Ventura, C. M. Hughes, R. M. Gilbride, J. C. Ford, K. Oswald, R. Shoemaker, Y. Li, M. S. Lewis, et al. 2013. Immune clearance of highly pathogenic SIV infection. *Nature* 502: 100–104.
50. Hansen, S. G., H. L. Wu, B. J. Burwitz, C. M. Hughes, K. B. Hammond, A. B. Ventura, J. S. Reed, R. M. Gilbride, E. Ainslie, D. W. Morrow, et al. 2016. Broadly targeted $CD8^+$ T cell responses restricted by major histocompatibility complex E. *Science* 351: 714–720.