Human Molecular Genetics, 2015, Vol. 24, No. 15 4268–4275

doi: 10.1093/hmg/ddv158
Advance Access Publication Date: 1 May 2015
Original Article

ORIGINAL ARTICLE

Epigenetic regulation of differential *HLA-A* allelic expression levels

Veron Ramsuran[1,2], Smita Kulkarni[2], Colm O'huigin[1], Yuko Yuki[1], Danillo G. Augusto[1,2], Xiaojiang Gao[1] and Mary Carrington[1,2,*]

[1]Cancer Inflammation Program, Laboratory of Experimental Immunology, Leidos Biomedical Research Incorporated, Frederick National Laboratory for Cancer Research, Frederick, MD 21702, USA and [2]Ragon Institute of Massachusetts General Hospital, Massachusetts Institute of Technology, and Harvard University, Cambridge, MA 02139, USA

*To whom correspondence should be addressed. Tel: +1 3018461390; Fax: +1 3018466771; Email: carringm@mail.nih.gov

Abstract

MHC class I expression levels influence the strength of immune responses and represent another variable in determining outcome to disease beyond peptide binding alone. Identification of the *HLA* loci that vary in allelic expression levels and delineating the mechanism responsible for expression variation may provide the opportunity to modify their expression therapeutically. We have examined the expression levels of allelic lineages at the *HLA-A* locus in a sample of 216 European Americans using a real-time polymerase chain reaction assay, which amplifies all *HLA-A* lineages specifically with equal efficiency, and observed a gradient of expression that associates with *HLA-A* allelic lineage ($R = 0.6$, $P = 5 \times 10^{-25}$). DNA methylation of the *HLA-A* gene appears to contribute to the variation in *HLA-A* mRNA expression levels, as a significant inverse correlation was observed between *HLA-A* mRNA expression levels in untreated cells and the degree to which expression is increased after treatment of the cells with a DNA methyltransferase inhibitor ($R = 0.6$, $P = 2.8 \times 10^{-6}$). Further, deep-sequencing and immunoprecipitation assays revealed allelic lineage-specific methylation patterns within the *HLA-A* promoter region where increased DNA methylation levels correlated significantly with reduced *HLA-A* expression levels ($R = 0.89$, $P = 3.7 \times 10^{-9}$). These data demonstrate *HLA-A* allelic lineage-specific variation in expression levels, and DNA methylation as a likely factor in contributing to this variation.

Introduction

The highly polymorphic *HLA* class I and class II genes map to the human *MHC* and encode molecules that contribute to both the adaptive and innate immune responses. Genome-wide association studies have identified the *MHC* as the most rich 4 Mb region of the genome in terms of association with virtually all types of complex human disease, and for many of these diseases, they highlight this region as the most important in determining disease risk genome wide (1,2). Increasing expression of *MHC* class I molecules have been shown to improve the efficacy of antigen presentation and immune responses (3,4). In line with these human and murine models, higher HLA-C expression levels have been associated with enhanced human immunodeficiency virus (HIV) control, greater odds of generating HLA-C-restricted cytotoxic T lymphocyte (CTL) responses to the virus, and stronger immune pressure on HIV as measured by viral escape mutations (5,6). HLA-C expression levels also associate with risk of Crohn's disease, but in this case, low expression appears to confer protection (5). Opposing effects of expression levels on various diseases may indicate selection pressure to maintain differential expression across HLA allelic lineages, a model supported by genetic evolution analyses of the *HLA-C* locus (7).

Received: February 6, 2015. **Revised:** April 16, 2015. **Accepted:** April 27, 2015

While a minimal number of peptide–MHC complexes are sufficient for CTL killing *in vitro* (8), increasing density of these interactions improves the efficiency of cytotoxicity and alters the response to induce secretion of inflammatory cytokines (3,4). IFN-γ-mediated upregulation of HLA expression on host cells is a hallmark of the host response to acute infection (9), and exemplifies the impact that enhanced *HLA* expression levels have on the immune response. Indeed, certain human cancers and viruses have devised mechanisms to down-regulate HLA-class I proteins (10–13), apparently as a means of avoiding host immune recognition. There are two classical *MHC* class I genes within the chicken genome, which are expressed at distinct levels on the cell surface, and the locus encoding the low-expression type confers resistance to certain infectious diseases in this species (14–16). In this case, low expression correlates with greater promiscuity in peptide binding (17), a situation that may be essential in order to prevent overt self-reactivity. Overall, these data point to expression levels of MHC class I molecules as an important determinant in disease outcome, and when expression levels vary by allelic lineage, as shown for the *HLA-C* locus, this constitutes another specific characteristic of the lineage that may impact its function in addition to peptide binding specificity.

A polymorphic microRNA binding site in the *HLA-C* 3′UTR accounts for part of the differential allele-specific expression level of *HLA-C* allelic lineages (18). The *HLA-A* locus also has a miR-148a binding site in its 3′UTR, but unlike *HLA-C*, the seed region in the *HLA-A* locus is fixed for the form that is inhibited by miR-148a (7,18), so differential miR-148a binding cannot account for variation in *HLA-A* expression levels. Several studies have identified DNA methylation as an epigenetic mechanism responsible for varied expression levels of genes located within the *MHC* region (11–13,19–22). Variation in *HLA-A* methylation levels may play a role in certain cancers (10–13), but these studies have not addressed whether differential, allelic lineage-specific DNA methylation patterns occur at the locus and whether they affect *HLA-A* expression levels in a lineage-specific manner.

Here, we show that *HLA-A* mRNA expression levels vary significantly by allelic lineage due at least in part to differential DNA methylation patterns where low *HLA-A* expressing lineages have higher DNA methylation levels, and vice versa, in healthy European American (EA) donors. This mode of transcriptional regulation is distinct from *HLA-B* and *HLA-C*, as all *HLA-B* and *-C* lineages are completely unmethylated in spite of having the similar CpG target sites for methylation as *HLA-A*, further distinguishing the evolutionary history of the classical *HLA* class I loci.

Results

HLA-A mRNA expression levels vary as a function of allelic lineage

The *HLA-A* mRNA expression levels, determined in 216 healthy EA donors using quantitative polymerase chain reaction (qPCR) primers that specifically amplify only the *HLA-A* gene, showed significant differential expression levels (Fig. 1). The expression levels determined were an intrinsic characteristic of each allelic lineage, as qPCR primers used to measure expression levels amplified all *HLA-A* lineages with equal efficiency and no significant correlation was observed between average mRNA expression levels and the PCR efficiency for the respective lineage (Supplementary Material, Fig. S1; Spearman correlation coefficient $R = 0.05$; $P = 0.87$). The three pairs of qPCR primers that were designed to amplify distinct coding regions in *HLA-A* mRNA were strongly correlated with one another (exon 2 versus exon 1–2, $R = 0.9$, $P < 0.0001$; exon 2 versus exon 2–3, $R = 0.87$, $P < 0.0001$; exon 1–2 versus exon 2–3, $R = 0.87$, $P < 0.0001$; Supplementary Material, Fig. S2A–C), providing further confirmation of accurate measurement of intrinsic expression levels.

Our previous study estimated the effects of individual HLA-C allelic lineages by simply calculating the raw average expression levels among all genotypes in which alleles of each particular lineage were found (5). Such expression estimates are heavily dependent on the observed allelic pairings and are influenced by the set of 'other' lineages present. Linear regression analysis is preferred for determining the average expression level of each lineage since it allows unbiased determinations of allelic lineage effects from a pool of genotypic data. Supplementary Material, Table S1 shows a comparison of estimates of lineage effects made by each method (raw average expression or linear regression) for the 14 lineages for which at least five individuals were present (three donors with rare lineages were omitted). As expected, linear regression gives a fit ($R^2 = 0.708$, $P < 10^{-54}$) that slightly improves on that found for averaged expression ($R^2 = 0.678$, $P < 10^{-50}$) (see Fig. 1), with most of the difference attributable to extreme high or low-expression lineages.

To determine the stability of measured expression levels over time and consistency in estimation of expression values, fresh

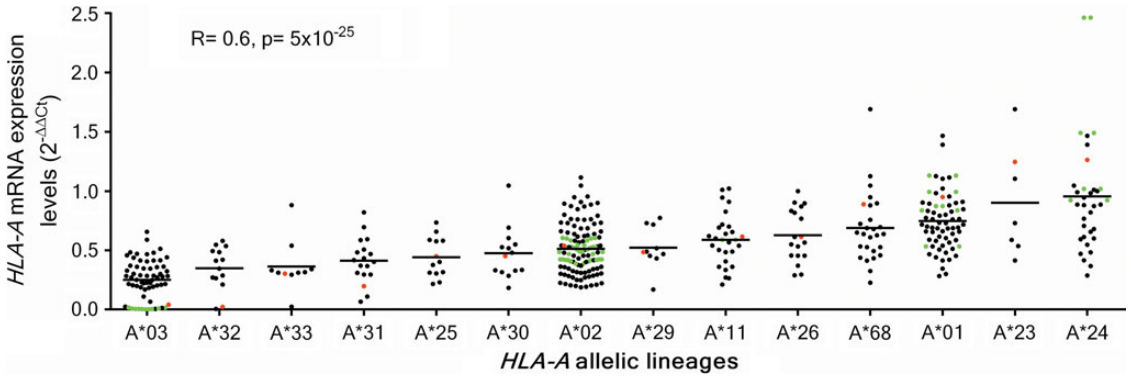

Figure 1. *HLA-A* gradient of expression across allelic lineages present in EAs. *HLA-A* mRNA expression level in 216 healthy donors was measured using qPCR. The average expression level of *HLA-A* was normalized to *β2M* mRNA using the $2^{-\Delta\Delta Ct}$ method. The relative expression value is plotted twice for each donor (i.e. once for each *HLA-A* lineage present). Expression levels correlate significantly with *HLA-A* lineage and are continuously distributed ($R = 0.6$, $P = 5 \times 10^{-25}$, ANOVA). Allelic lineages with ≥5 individuals are shown. The horizontal line indicates the raw average expression for the specific lineage. *HLA-A* homozygous individuals are marked in green dots. Red dots represent linear regression estimates of average expression levels for each *HLA-A* allelic lineage (estimated as a diploid homozygote).

blood samples were collected a second time, one year later, from 43 of the initial 216 donors and *HLA-A* mRNA expression levels were reassessed. Three distinct correlations were generated. Firstly, the expression level determined directly for each individual from the initial draw (year 1) was compared with that from the second draw (year 2) for each of the 43 individuals and a strong, highly significant correlation was observed (Supplementary Material, Fig. S2D; $R = 0.86$, $P < 0.0001$). Secondly, individual allelic expression values at year 1 and at year 2 were estimated by linear regression as described for all 215 donors, but restricting the analysis to 39 of the 43 repeat donors, which involved 11 distinct allelic lineages represented more than once among the 43 donors (i.e. 4 of the 43 donors carried lineages that were observed only once and these were not included in the analyses). When these 11 estimated mean lineage expression values were compared between year 1 and year 2 for each of the 39 individuals, a correlation of $R = 0.8628$ was observed ($P = 0.0003$). Finally, when the year 2 lineage expression values (as estimated by regression from the 39 donors, 11 allelic lineages) were compared with those previously obtained from the whole cohort (i.e. 216 donors with year 1 expression values estimated by regression), a correlation of 0.8006 was observed ($P = 0.0031$). These analyses indicate two important features of *HLA-A* expression levels: stability of expression levels over time in any given individual and consistency in assigning expression values, even when only a subset of 39 donors are employed for determining expression levels for each allelic lineage.

Methylation affects *HLA-A* expression levels

Numerous CpG sites are located in the promoter, exon 1–3, and intron 1–2 regions of the *HLA-A* gene (Fig. 2), including one CpG island located in the promoter and another spanning the intron 1–exon 3 region. The methylation levels at these sites were estimated using the methylation DNA immunoprecipitation (MeDIp) assay on DNA from healthy EA donors. Varying levels of methylated DNA as measured by qPCR after MeDIp for each individual tested were present in the promoter region (Supplementary Material, Fig. S3A), whereas the exon 2–3 and intron 2 regions showed no significant methylation levels in any of the *HLA-A* lineages tested (Supplementary Material, Fig. S3A). This result was confirmed by sequencing the bisulfite converted DNA (Supplementary Material, Fig. S3B and C). The *HLA-A* promoter region contained several cytosines at CpG sites in the sequence, indicating the presence of DNA methylation within the region (Supplementary Material, Fig. S3C), but cytosines at CpG sites were essentially absent from the gene body region that was analyzed (exon 2, intron 2 and exon 3; Supplementary Material, Fig. S3B).

A deep-sequencing approach was employed to determine methylated positions and the frequency of methylation at a given site. A bisulfite converted DNA fragment ~300 bp upstream of the transcription start site was subjected to deep sequencing. Individuals homozygous for either a high expressing lineage (*HLA-A*24*) or a low expressing lineage (*HLA-A*03*) were selected

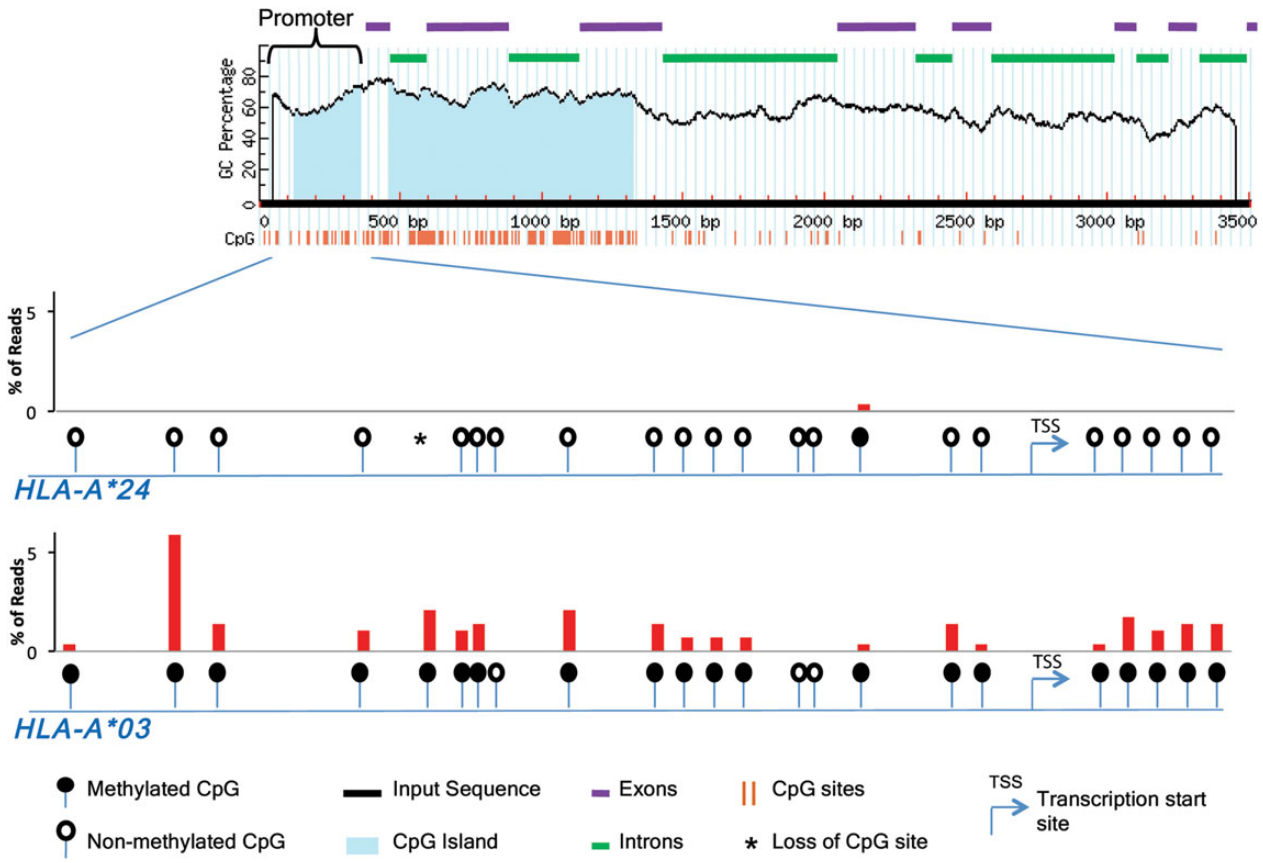

Figure 2. DNA methylation and CpG island locations throughout the *HLA-A* gene region. Using the 454 sequencing technique, the *HLA-A* promoter region showed variable number of methylated sites between individuals homozygous for the highest and lowest *HLA-A* expression lineages, *HLA-A*24* and *-A*03*, respectively. The percentage of reads containing a methylated CpG site is shown as red vertical bars for each position. The asterisk illustrates the site in *HLA-A*24* lacking a CpG as a result of polymorphism.

for sequence analysis (Fig. 2). The high-expression *HLA-A*24* lineage showed only one methylated position in the promoter region that was present in 70 reads. On the other hand, one or more individual CpG sites were found to be methylated in 74 of 289 reads (25.6%) across the promoter region of the low-expression *HLA-A*03* lineage. The frequency of sequences determined as methylated is found to be significantly different ($P < 10^{-4}$, Fisher''s exact test) for the two lineages tested. The percentage of reads depicted above each CpG site (Fig. 2) illustrates that the *HLA-A*03* lineage possesses a higher number of reads in which methylation was present compared with the *HLA-A*24* lineage. One CpG site is actually disrupted by a polymorphism present in the *HLA-A*24* lineage (indicated by asterisk, Fig. 2), slightly decreasing the number of potential methylation sites. Further, two methylation sites located immediately within well characterized transcription factor binding sites in the HLA class I regulatory complex (CRC) (23,24) were identified for the *HLA-A*03* lineage, but no methylated site was observed in any of these motifs for *HLA-A*24* lineage (Supplementary Material, Fig. S4), raising the possibility that allelic expression may be regulated differentially depending on methylation status within the CRC.

5′-Aza-CdR is an inhibitor of DNA methyltransferase and treatment with the drug results in greater expression of genes/lineages that are more heavily methylated intrinsically. *HLA-A*, *-B* and *-C* mRNA expression levels were measured on both 5′-Aza-CdR treated and untreated cells after 24 h of culture (Fig. 3A–C). The fold change in expression levels on the treated versus untreated cells was compared with the expression levels at zero hour untreated cells (i.e. intrinsic expression levels) in 50 donors. Peripheral blood mononuclear cells (PBMCs) from subjects with intrinsically lower *HLA-A* expression (i.e. more heavily methylated) showed a significantly higher fold change in *HLA-A*

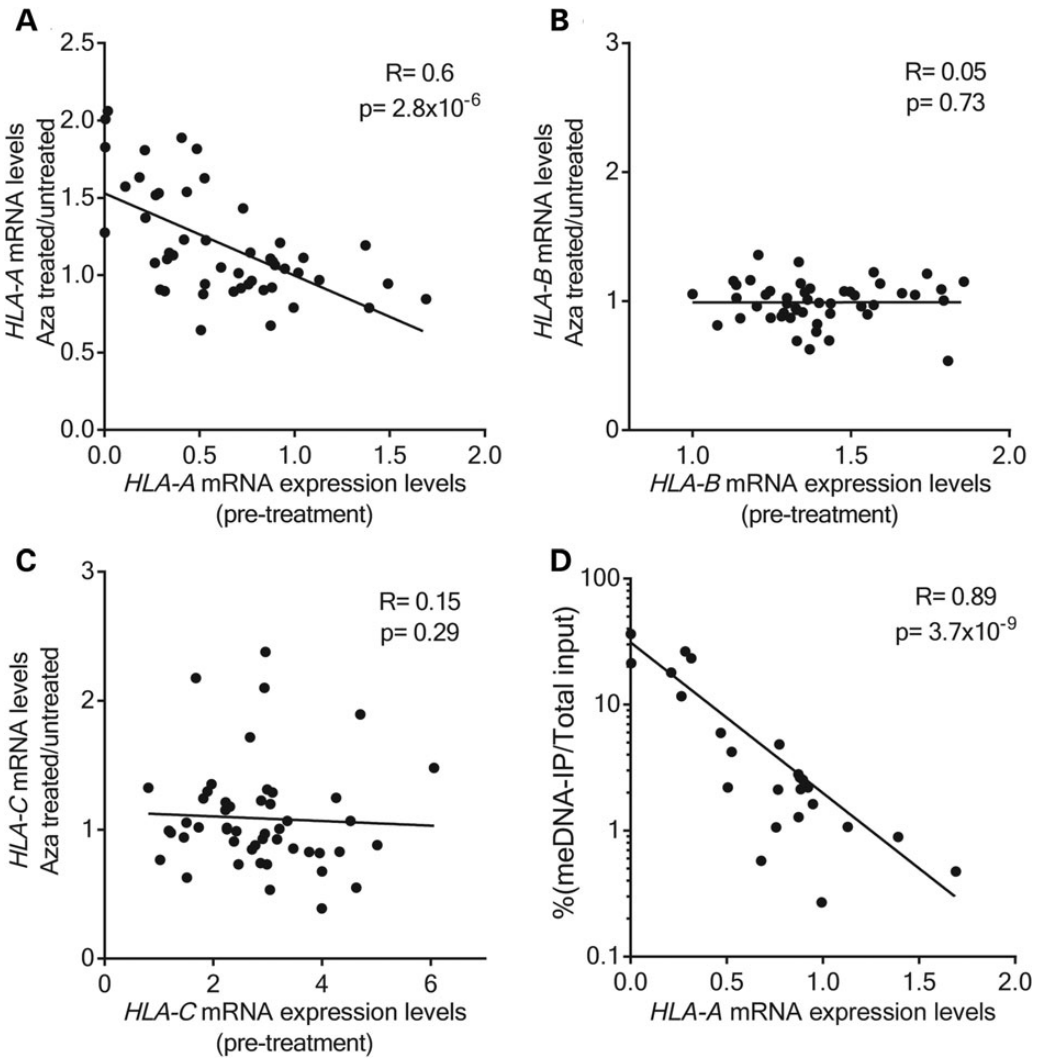

Figure 3. *HLA-A* expression level negatively correlates with DNA methylation level. (**A**) *HLA-A* mRNA levels were measured in PBMCs from each of 50 healthy EA donors as follows: immediately after collection of cells (i.e. pre-treatment); after culture for 24 h with 5-aza-2′-deoxycytidine (5′-Aza-CdR), a DNA methyltransferase inhibitor that causes hypomethylation of DNA; after culture for 24 h in the absence of 5′-Aza-CdR. The fold change in expression level of 5′-Aza-CdR treated/untreated after 24 h in culture correlated negatively with *HLA-A* expression level at pre-treatment. (**B** and **C**) *HLA-B* (B) and *-C* (C) mRNA levels were measured using the 50 healthy donors treated with 5′-Aza-CdR as described in A. (**D**) Sonicated DNA fragments from 25 healthy EA donors were immunoprecipitated using a methylation specific antibody (MeDIp assay). The MeDIp fragments were then quantified using qPCR and compared with levels of the input samples. A significant inverse correlation between the amount of methylated DNA and the level of *HLA-A* mRNA expression across donors was observed, where individuals that express *HLA-A* at lower levels have a higher proportion of methylated DNA. In all plots, each dot represents a single donor.

expression levels between treated and untreated cells after 24 h in culture as compared with that from subjects with intrinsically higher expression levels ($R = 0.6$, $P = 2.8 \times 10^{-6}$; Fig. 3A). However, unlike *HLA-A* there was no effect of 5′-Aza-CdR treatment on *HLA-B* and *HLA-C* expression, despite these loci having a similar number of CpG sites ($R = 0.05$, $P = 0.73$, Fig. 3B and $R = 0.15$, $P = 0.29$, Fig. 3C).

5′-Aza-CdR treatment results in genome-wide demethylation and it is possible that demethylation of genes other than *HLA-A* itself may account for the correlation shown in Figure 3A. For example, demethylation of a transcription factor that upregulates *HLA-A* may enhance *HLA-A* expression levels, accounting for higher *HLA-A* mRNA levels upon 5′-Aza-CdR treatment, rather than demethylation of the *HLA-A* gene itself. In order to test directly for a correlation between methylation levels of the *HLA-A* gene itself and levels of *HLA-A* mRNA expression, we measured the methylation levels in 25 donors using the methylation specific antibody in a MeDIp assay and compared these values with intrinsic *HLA-A* expression levels in the same donors. We observed a strong negative correlation between the MeDIp methylation levels and *HLA-A* mRNA expression levels ($R = 0.89$, $P = 3.7 \times 10^{-9}$; Fig. 3D), confirming results from deep sequencing and the 5′Aza-CdR assays.

Discussion

Growing evidence over the past several years points to the importance of MHC class I and II lineage expression levels in an effective immune response against both pathogens and tumors (3–6,14–16,25), illustrating the need to determine consistent patterns of differential expression levels across allelic lineages of each class I and II gene and to identify the mechanisms by which differential expression occurs. This endeavor is hampered by the extensive polymorphism at each *HLA* locus, which restricts the ability to obtain monoclonal antibodies (Mabs) that recognize all lineages of a given locus with equal affinity. On the other hand, the degree of similarity across the genes poses a problem in generating Mabs that specifically recognize a single locus and do not cross-react with lineages of other loci. As such, no HLA-A specific Mab is presently available, so we measured allelic lineage-specific expression levels for the *HLA-A* locus by targeting mRNA using *HLA-A*-specific primers that recognize all *HLA-A* lineages with equal efficiency, which can be measured definitively.

We identified DNA methylation in the *HLA-A* promoter region as a factor regulating the expression of *HLA-A*, where higher methylation levels were observed in low-expression *HLA-A* lineages, and vice versa. This correlation is consistent with vast literature linking higher methylation levels to inhibition of transcription, resulting in lower gene expression. Differential allelic lineage-specific methylation patterns at the *HLA-A* locus result in corresponding lineage-specific expression levels, endowing each lineage with a distinguishing characteristic beyond peptide binding specificity.

Allele-specific methylation patterns have previously been linked to corresponding allele-specific expression levels at several other loci (26), as well. We identified locus-specific methylation within the MHC region, where differential lineage-specific methylation patterns are observed within the *HLA-A* locus, but not within the *HLA-B* and *-C* loci. While the classical class I genes may be regulated by common means to some extent, distinct mechanisms of regulation between the genes are being identified, including the differential miR-148A inhibition across *HLA-C* lineages, but not across *HLA-B* or *-A* lineages (18), and the mode by which interferon-γ regulates *HLA-A* expression as compared with *HLA-B* and *-C* (27). Such distinctions in regulatory mechanisms may reflect somewhat specialized functions across the loci. For example, every *HLA-C* lineage is capable of regulating natural killer (NK) cell activity by serving as a ligand for killer cell immunoglobulin-like receptors (KIR) expressed on NK cells, whereas most *HLA-A* and *HLA-B* lineages are not recognized by KIR (28). Distinctions are also seen in the estimated intensity of natural selection, with the selective coefficient for *HLA-B* exceeding that of *HLA-A* and *HLA-C* by approximately 3-fold and 16-fold, respectively (29). Thus, distinct regulatory mechanisms may have evolved for the three class I loci in order to accommodate specialized locus-specific functions.

Variation in *cis*-acting regulatory elements, including transcription factor binding sites, CpG dinucleotides, insulators (which block the interaction between enhancers and promoters), and variants that affect long-range chromosome structure (30,31) have been identified as possible mechanisms affecting allele-specific methylation (30,31). The high-expression *HLA-A*24* lineage, along with several other lineages, possess a polymorphism in the promoter region that disrupts a methylation site (CpG→TpG), but this CpG site is intact and can be methylated in the low-expression *HLA-A*03* lineage. Such variation can potentially result in differential methylation levels between lineages, contributing to lineage-specific expression levels. A CCCTC-binding factor (CTCF) binding site located about 1 kb upstream of the *HLA-A* transcription start site was identified using the CTCF binding site database (CTCFBSDB) (32) (data not shown). The function of CTCF bound DNA complexes can lead to gene repression, activation, silencing or chromatin insulation (33), so the role of this variant in the CTCF binding site (minor allele frequency of 0.46) upstream of *HLA-A* may result in *HLA-A* expression variation. The presence of DNA methylation within a CTCF binding site has been shown to cause a decrease in CTCF bound (34). Lineage-specific methylation patterns within the CTCF binding site upstream of *HLA-A* could potentially contribute further to differential *HLA-A* lineage-specific expression levels.

Several mechanisms are likely to contribute to the variation in expression levels of *HLA-A* lineages, and characterizing these mechanisms may present the potential for *HLA-A*-specific regulation of expression levels as drug targets. The *HLA-A* expression and methylation levels presented herein may be useful in genetic association studies of diverse diseases.

Materials and Methods

HLA-A mRNA expression levels using real-time PCR

Healthy EA donors ($n = 216$) were recruited from the Research Donor Program at the Frederick National Laboratory for Cancer Research. The *HLA-A* mRNA expression levels were measured using qPCR as previously described in detail (35). Total RNA was extracted from PBMCs using RNeasy Universal kit (Qiagen), treated with RNAse-free DNAse to remove genomic DNA, and quantitated using HT RNA Lab Chip (Caliper, Life Sciences). The RNA samples with a quality score of >8 were used for quantitative analysis of *HLA-A* mRNA expression. Reverse transcription was performed using the high capacity RNA to cDNA kit (Applied Bioscience). Power SYBR green PCR mastermix (Applied Biosystems) was used to run the qPCR assay on the ABI7900HT machine (Applied Bioscience). All primers sets used in this study are available upon request. *HLA-A* and the housekeeping gene *beta-2 microglobulin* (*β2M*) were amplified with the PCR conditions previously described (35). Primer specificity was verified by melt curve

analysis and by sequencing the *HLA-A* PCR amplicons. PCR amplification efficiencies were tested for all common *HLA-A* allelic lineages (Supplementary Material, Fig. S1) using serial dilutions of cDNA to generate standard curves. The *HLA-A* expression level was normalized to *β2M* and calculated using the $2^{-\Delta\Delta Ct}$ method (36) (where Ct is the threshold cycle).

A primer-BLAST and sequencing of PCR amplicons ensured that the primers amplified *HLA-A* specifically and that no *HLA-B*, -C or any other non-classical class I lineages were amplified. The reliability of the primers was verified by designing two additional sets of primers within the *HLA-A* locus, one set in exon 1–2 region and the second set in exon 2–3. The PCR amplification efficiencies and *HLA-A* specificity for both sets of primers were shown to be equal across *HLA-A* allelic lineages.

Detection of DNA methylation by sequencing

The CpG island prediction software MethPrimer (37) was used to identify CpG islands within the *HLA-A* gene region, for which default settings were applied (island size >100 bp, GC percentage >50%, and the observed/expected CpG ratio >0.6 (37)). Based on these predictions, appropriate PCR primers that are specific for the promoter, exon 2, exon 3, and intron 2 of *HLA-A* were designed using the MethPrimer (37) software (primers available upon request). Sodium bisulfite conversion was performed on genomic DNA from PBMCs using the EZ DNA methylation™ Kit (Zymo Research). In the presence of the sodium bisulfite reagent, unmethylated cytosines are converted to uracil, whereas the methylated cytosines are unreactive under the same treatment. The bisulfite converted DNA was then amplified (95°C for 15 min, 45 cycles of 95°C for 30 s, 60°C for 45 s, 72°C for 30 s and one cycle of 72°C for 10 min), which converts uracils to thymidines. Sanger sequencing was performed and the sequences were analyzed using the Mutation Surveyor® software (Soft Genetics).

454 sequencing (Roche) was employed both to verify methylated positions and to determine the frequency at which each site was actually methylated in PBMC DNA from a given individual. Multiplex identifiers were used to barcode the PCR amplicons (Roche). PCR products using bisulfite-treated DNA as the template were purified using the Qiaquick PCR Purification Kit (Qiagen), followed by measurement of the products using Quant-iT™ PicoGreen® dsDNA Assay Kit (Invitrogen). After amplicon pooling, emulsion PCR was performed, and DNA-containing beads were recovered, enriched and sequenced using the 454 GS Junior Titanium Series as per manufacturer's instructions (Roche). Online Galaxy software was used to determine the number of methylation sites for each *HLA-A* lineage tested.

Detection of DNA methylation by immunoprecipitation assay

HLA-A methylation levels were measured in the promoter, exon2, exon3 and intron 2 regions using the MagMeDIP Kit (Diagenode) in 25 healthy EA donors. Genomic DNA was sonicated using the Bioruptor ® (Diagenode) to an average size of 300 bp. Methylated DNA was immunoprecipitated with a methylation specific antibody, anti-5meC. After several washes, the immunoprecipitated DNA was used in a qPCR assay to determine methylation levels. The enrichment of the *HLA-A* specific genomic regions targeted was assessed relative to the input DNA as the percentage of Ct_{MeDIp}/Ct_{input}, where Ct_{MeDIp} is the Cycle threshold for the MeDIp sample and Ct_{input} is the Cycle threshold for the input sample.

5′-Aza-2-deoxycytidine treatment

PBMCs from EA healthy donors ($n = 50$) were treated with 10 μM 5′-aza-2-deoxycytidine (5′-Aza-CdR; Sigma) or with dimethyl sulfoxide (DMSO; treatment control). A subset of cells was stored at time zero. The 5′-Aza-CdR or DMSO-treated cells were maintained at 37°C and 5% CO_2 in RPMI (Lonza) with 10% fetal calf serum (Lonza) for 24 h. RNA was then extracted from these and from untreated cells that had been frozen at the zero hour time point using the RNeasy Universal kit (Qiagen); *HLA-A*, *HLA-B* and *HLA-C* expression levels were measured by qPCR. Data are represented as fold change of 5′-Aza-CdR treated versus DMSO treated and compared with the time zero *HLA-A* expression levels.

Statistical analyses

The linear regression function (lm) of the R package (38) was used to determine the effects of individual lineages found among the 216 EA donors on *HLA-A* mRNA expression. For functional and evolutionary considerations as well as for increased sample size, we group alleles into their respective two-digit allelic lineages, without further subdivision. The lineages of the two alleles carried by each individual were used as explanatory variables in the linear regression. To enhance accuracy, only individuals having lineages present on at least five chromosomes in the whole sample were considered in the analysis. There were 14 such lineages present among 213 of the 216 donors, representing the *HLA-A*01*, *-A*02*, *-A*03*, *-A*11*, *-A*23*, *-A*24*, *-A*25*, *-A*26*, *-A*29*, *-A*30*, *-A*31*,*-A*32*,*-A*33* and *-A*68* lineages. Regression values for each lineage were determined with intercept set to zero. Coefficients of determination (multiple R^2) were based on deviations from the mean value and were used to test the goodness of fit between predicted and observed donor *HLA-A* mRNA expression. To avoid negative predictions for allelic lineage expression levels, we slightly adjusted all regression values for *HLA-A* lineages towards the mean expression level using an *ad hoc* shrinkage correction (Supplementary Material, Table S1).

Spearman correlations and Mann–Whitney tests were performed using GraphPad Prism software version 6 (GraphPad Software, San Diego California USA).

Supplementary Material

Supplementary Material is available at *HMG* online.

Acknowledgements

We would like to thank the participants from the Research Donor Program at the Frederick National Laboratory for Cancer Research. We would also like to thank; Dr George Nelson and Dr Richard Apps for assistance and helpful discussions and Ms Sara Bass for technical support.

Conflict of Interest statement. None declared.

Funding

This work was supported by federal funds from the National Cancer Institute, National Institutes of Health (NIH), [grant numbers HHSN261200800001E, N02-CP-55504, R01-DA04334 and R01-DA12568]. The content of this publication does not necessarily reflect the views or policies of the Department of Health and Human Services, nor does mention of trade names, commercial products or organizations imply endorsement by

the US Government. This Research was supported in part by the Intramural Research Program of the NIH, Frederick National Lab, Center for Cancer Research.

References

1. Carrington, M. and Walker, B.D. (2012) Immunogenetics of spontaneous control of HIV. *Annu. Rev. Med.*, **63**, 131–145.
2. Welter, D., MacArthur, J., Morales, J., Burdett, T., Hall, P., Junkins, H., Klemm, A., Flicek, P., Manolio, T., Hindorff, L. *et al.* (2014) The NHGRI GWAS Catalog, a curated resource of SNP-trait associations. *Nucleic Acids Res.*, **42**, D1001–D1006.
3. Faroudi, M., Utzny, C., Salio, M., Cerundolo, V., Guiraud, M., Muller, S. and Valitutti, S. (2003) Lytic versus stimulatory synapse in cytotoxic T lymphocyte/target cell interaction: manifestation of a dual activation threshold. *Proc. Natl. Acad. Sci. U. S. A.*, **100**, 14145–14150.
4. Reits, E.A., Hodge, J.W., Herberts, C.A., Groothuis, T.A., Chakraborty, M., Wansley, E.K., Camphausen, K., Luiten, R.M., de Ru, A.H., Neijssen, J. *et al.* (2006) Radiation modulates the peptide repertoire, enhances MHC class I expression, and induces successful antitumor immunotherapy. *J. Exp. Med.*, **203**, 1259–1271.
5. Apps, R., Qi, Y., Carlson, J.M., Chen, H., Gao, X., Thomas, R., Yuki, Y., Del Prete, G.Q., Goulder, P., Brumme, Z.L. *et al.* (2013) Influence of HLA-C expression level on HIV control. *Science*, **340**, 87–91.
6. Blais, M.E., Zhang, Y., Rostron, T., Griffin, H., Taylor, S., Xu, K., Yan, H., Wu, H., James, I., John, M. *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.
7. O'Huigin, C., Kulkarni, S., Xu, Y., Deng, Z., Kidd, J., Kidd, K., Gao, X. and Carrington, M. (2011) The molecular origin and consequences of escape from miRNA regulation by HLA-C alleles. *Am. J. Hum. Genet.*, **89**, 424–431.
8. Purbhoo, M.A., Irvine, D.J., Huppa, J.B. and Davis, M.M. (2004) T cell killing does not require the formation of a stable mature immunological synapse. *Nat. Immunol.*, **5**, 524–530.
9. Koeffler, H.P., Ranyard, J., Yelton, L., Billing, R. and Bohman, R. (1984) Gamma-interferon induces expression of the HLA-D antigens on normal and leukemic human myeloid cells. *Proc. Natl. Acad. Sci. U. S. A.*, **81**, 4080–4084.
10. Menon, A.G., Morreau, H., Tollenaar, R.A., Alphenaar, E., Van Puijenbroek, M., Putter, H., Janssen-Van Rhijn, C.M., Van De Velde, C.J., Fleuren, G.J. and Kuppen, P.J. (2002) Down-regulation of HLA-A expression correlates with a better prognosis in colorectal cancer patients. *Lab. Invest.*, **82**, 1725–1733.
11. Nie, Y., Yang, G., Song, Y., Zhao, X., So, C., Liao, J., Wang, L.D. and Yang, C.S. (2001) DNA hypermethylation is a mechanism for loss of expression of the HLA class I genes in human esophageal squamous cell carcinomas. *Carcinogenesis*, **22**, 1615–1623.
12. Serrano, A., Tanzarella, S., Lionello, I., Mendez, R., Traversari, C., Ruiz-Cabello, F. and Garrido, F. (2001) Rexpression of HLA class I antigens and restoration of antigen-specific CTL response in melanoma cells following 5-aza-2′-deoxycytidine treatment. *Int. J. Cancer*, **94**, 243–251.
13. Ye, Q., Shen, Y., Wang, X., Yang, J., Miao, F., Shen, C. and Zhang, J. (2010) Hypermethylation of HLA class I gene is associated with HLA class I down-regulation in human gastric cancer. *Tissue Antigens*, **75**, 30–39.
14. Kaufman, J., Milne, S., Gobel, T.W., Walker, B.A., Jacob, J.P., Auffray, C., Zoorob, R. and Beck, S. (1999) The chicken B locus is a minimal essential major histocompatibility complex. *Nature*, **401**, 923–925.
15. Kaufman, J., Volk, H. and Wallny, H.J. (1995) A "minimal essential Mhc" and an "unrecognized Mhc": two extremes in selection for polymorphism. *Immunol. Rev.*, **143**, 63–88.
16. Wallny, H.J., Avila, D., Hunt, L.G., Powell, T.J., Riegert, P., Salomonsen, J., Skjodt, K., Vainio, O., Vilbois, F., Wiles, M. V. *et al.* (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. U. S. A.*, **103**, 1434–1439.
17. Koch, M., Camp, S., Collen, T., Avila, D., Salomonsen, J., Wallny, H.J., van Hateren, A., Hunt, L., Jacob, J.P., Johnston, F. *et al.* (2007) Structures of an MHC class I molecule from B21 chickens illustrate promiscuous peptide binding. *Immunity*, **27**, 885–899.
18. Kulkarni, S., Savan, R., Qi, Y., Gao, X., Yuki, Y., Bass, S.E., Martin, M.P., Hunt, P., Deeks, S.G., Telenti, A. *et al.* (2011) Differential microRNA regulation of HLA-C expression and its association with HIV control. *Nature*, **472**, 495–498.
19. Satoh, A., Toyota, M., Ikeda, H., Morimoto, Y., Akino, K., Mita, H., Suzuki, H., Sasaki, Y., Kanaseki, T., Takamura, Y. *et al.* (2004) Epigenetic inactivation of class II transactivator (CIITA) is associated with the absence of interferon-gamma-induced HLA-DR expression in colorectal and gastric cancer cells. *Oncogene*, **23**, 8876–8886.
20. Rakyan, V.K., Hildmann, T., Novik, K.L., Lewin, J., Tost, J., Cox, A.V., Andrews, T.D., Howe, K.L., Otto, T., Olek, A. *et al.* (2004) DNA methylation profiling of the human major histocompatibility complex: a pilot study for the human epigenome project. *PLoS Biol.*, **2**, e405.
21. Guillaudeux, T., Rodriguez, A.M., Girr, M., Mallet, V., Ellis, S.A., Sargent, I.L., Fauchet, R., Alsat, E. and Le Bouteiller, P. (1995) Methylation status and transcriptional expression of the MHC class I loci in human trophoblast cells from term placenta. *J. Immunol.*, **154**, 3283–3299.
22. Bormann, F., Sers, C., Seliger, B., Handke, D., Bergmann, T., Seibt, S., Lehrach, H. and Dahl, A. (2011) Methylation-specific ligation detection reaction (msLDR): a new approach for multiplex evaluation of methylation patterns. *Mol. Genet. Genomics*, **286**, 279–291.
23. Cereb, N., Lee, S., Maye, P., Kong, Y. and Yang, S.Y. (1994) Non-random allelic variation in the regulatory complex of HLA class I genes. *Hum. Immunol.*, **41**, 46–51.
24. Cereb, N. and Yang, S.Y. (1994) The regulatory complex of HLA class I promoters exhibits locus-specific conservation with limited allelic variation. *J. Immunol.*, **152**, 3873–3883.
25. Thomas, R., Thio, C.L., Apps, R., Qi, Y., Gao, X., Marti, D., Stein, J.L., Soderberg, K.A., Moody, M.A., Goedert, J.J. *et al.* (2012) A novel variant marking HLA-DP expression levels predicts recovery from hepatitis B virus infection. *J. Virol.*, **86**, 6979–6985.
26. Schalkwyk, L.C., Meaburn, E.L., Smith, R., Dempster, E.L., Jeffries, A.R., Davies, M.N., Plomin, R. and Mill, J. (2010) Allelic skewing of DNA methylation is widespread across the genome. *Am. J. Hum. Genet.*, **86**, 196–212.
27. Browne, S.K., Roesser, J.R., Zhu, S.Z. and Ginder, G.D. (2006) Differential IFN-gamma stimulation of HLA-A gene expression through CRM-1-dependent nuclear RNA export. *J. Immunol.*, **177**, 8612–8619.
28. Kulkarni, S., Martin, M.P. and Carrington, M. (2008) The Yin and Yang of HLA and KIR in human disease. *Semin. Immunol.*, **20**, 343–352.
29. Satta, Y., O'HUigin, C., Takahata, N. and Klein, J. (1994) Intensity of natural selection at the major histocompatibility complex loci. *Proc. Natl. Acad. Sci. U. S. A.*, **91**, 7184–7188.

30. Shoemaker, R., Deng, J., Wang, W. and Zhang, K. (2010) Allele-specific methylation is prevalent and is contributed by CpG-SNPs in the human genome. *Genome Res.*, **20**, 883–889.
31. Tycko, B. (2010) Allele-specific DNA methylation: beyond imprinting. *Hum. Mol. Genet.*, **19**, R210–R220.
32. Ziebarth, J.D., Bhattacharya, A. and Cui, Y. (2013) CTCFBSDB 2.0: a database for CTCF-binding sites and genome organization. *Nucleic Acids Res.*, **41**, D188–D194.
33. Weth, O. and Renkawitz, R. (2011) CTCF function is modulated by neighboring DNA binding factors. *Biochem. Cell Biol.*, **89**, 459–468.
34. Bell, A.C. and Felsenfeld, G. (2000) Methylation of a CTCF-dependent boundary controls imprinted expression of the Igf2 gene. *Nature*, **405**, 482–485.
35. Kulkarni, S., Qi, Y., O'HUigin, C., Pereyra, F., Ramsuran, V., McLaren, P., Fellay, J., Nelson, G., Chen, H., Liao, W. *et al.* (2013) Genetic interplay between HLA-C and MIR148A in HIV control and Crohn disease. *Proc. Natl. Acad. Sci. U. S. A.*, **110**, 20705–20710.
36. Livak, K.J. and Schmittgen, T.D. (2001) Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. *Methods*, **25**, 402–408.
37. Li, L.C. and Dahiya, R. (2002) MethPrimer: designing primers for methylation PCRs. *Bioinformatics*, **18**, 1427–1431.
38. R Development Core Team (2008) *R: A Language and Environment for Statistical Computing*, R Foundation for Statistical Computing, Vienna, Austria.