Transcriptionally Abundant Major Histocompatibility Complex Class I Alleles Are Fundamental to Nonhuman Primate Simian Immunodeficiency Virus-Specific CD8$^+$ T Cell Responses

Melisa L. Budde,† Jennifer J. Lhost,† Benjamin J. Burwitz, Ericka A. Becker, Charles M. Burns, Shelby L. O’Connor, Julie A. Karl, Roger W. Wiseman, Benjamin N. Bimber, Guang Lan Zhang, William Hildebrand, Vladimir Brusic, and David H. O’Connor

Department of Pathology and Laboratory Medicine, University of Wisconsin—Madison, Madison, Wisconsin 53706; Wisconsin National Primate Research Center, Madison, Wisconsin 53706; Department of Microbiology and Immunology, University of Oklahoma Health Sciences Center, Oklahoma City, Oklahoma 73104; and Cancer Vaccine Center, Dana-Farber Cancer Institute, Boston, Massachusetts 02115

Received 10 November 2010/Accepted 13 January 2011

Simian immunodeficiency virus (SIV)-infected macaques are the preferred animal model for human immunodeficiency virus (HIV) vaccines that elicit CD8$^+$ T cell responses. Unlike humans, whose CD8$^+$ T cell responses are restricted by a maximum of six HLA class I alleles, macaques express up to 20 distinct major histocompatibility complex class I (MHC-I) sequences. Interestingly, only a subset of macaque MHC-I sequences are transcriptionally abundant in peripheral blood lymphocytes. We hypothesized that highly transcribed MHC-I sequences are principally responsible for restricting SIV-specific CD8$^+$ T cell responses. To examine this hypothesis, we measured SIV-specific CD8$^+$ T cell responses in MHC-I homozygous Mauritian cynomolgus macaques. Each of eight CD8$^+$ T cell responses defined by full-proteome gamma interferon (IFN-γ) enzyme-linked immunospot (ELISPOT) assay were restricted by four of the five transcripts that are transcriptionally abundant (>1% of total MHC-I transcripts in peripheral blood lymphocytes). The five transcriptionally rare transcripts shared by these animals did not restrict any detectable CD8$^+$ T cell responses. Further, seven CD8$^+$ T cell responses were defined by identifying peptide binding motifs of the three most frequent MHC-I transcripts on the M3 haplotype. Combined, these results suggest that transcriptionally abundant MHC-I transcripts are principally responsible for restricting SIV-specific CD8$^+$ T cell responses. Thus, only a subset of the thousands of known MHC-I alleles in macaques should be prioritized for CD8$^+$ T cell epitope characterization.

The best animal model to evaluate human immunodeficiency virus (HIV) vaccine proof-of-concept studies is simian immunodeficiency virus (SIV)-infected nonhuman primates (NHPs). Indeed, the focus on cellular immune responses against HIV and SIV has increased efforts to characterize major histocompatibility complex (MHC) class I alleles. MHC class I proteins are present on all nucleated cells, interact with multiple lymphocyte subsets, and are thus involved in both innate and adaptive immunity. A key function of MHC class I molecules is to present peptide fragments on the cell surface for examination by CD8$^+$ T cells (39). In humans, MHC class I loci are well defined, with only three classical polymorphic genes per haplotype (HLA-A, -B, and -C). In contrast, macaque MHC class I genes have undergone complex duplications and deletions, resulting in a heterogeneous assortment of up to 20 MHC class I loci per haplotype (4, 38).

Rhesus macaques of Indian origin are the best-characterized and most commonly studied NHP model for AIDS. Indeed, MHC class I polymorphism has been well characterized, with more than 1,000 MHC class I sequences from rhesus macaque species listed in GenBank. Due to this high level of diversity, most CD8$^+$ T cell response studies focus on a single well-defined MHC class I molecule, such as Mamu-A1*001, rather than an entire MHC class I haplotype. Fortunately, individual susceptibility to SIV/HIV can be influenced, both positively and negatively, by individual MHC class I transcripts (22, 24, 30, 31, 33, 55). Indeed, SIV controllers often contain a single protective MHC class I allele (e.g., Mamu-B*008 or Mamu-B*017) and are studied to gain an understanding of desirable CD8$^+$ T cell responses (11, 20–22, 24, 36). To date, 20 MHC class I NHP alleles that restrict CD8$^+$ T cell responses against SIV have been published (27) (Table 1). Unfortunately, it is still unclear which of the other 1,000 MHC class I alleles restrict CD8$^+$ T cell responses.

The introduction of high-resolution pyrosequencing has dramatically altered our ability to predict which MHC class I alleles are influential. Recent work by our group has employed pyrosequencing in the full-length discovery of MHC class I alleles from multiple NHP populations (53). In addition to detecting the full-length MHC class I alleles, we can define the abundance of each allele on a haplotype. Importantly, the transcript levels are consistent and reproducible both between MHC-matched animals and across experiments with the same animals. Intriguingly, around half of the identified MHC class I transcripts identified are expressed at less than 1% of the total MHC class I transcripts on total peripheral blood mono-
nuclear cells (PBMCs) (data not shown). Thus, to examine the importance of transcriptional abundance and to limit the confounding effect of unknown alleles, we examined CD8\textsuperscript{+} T cell responses in SIV-infected Mauritian cynomolgus macaques (MCMs).

MCMs are a geographically isolated population descended from a small founder population over the last 500 years (19). Consequently, MCMs have extremely simple MHC genetics consisting of seven major haplotypes, termed M1 to M7 (2, 29). The seven haplotypes vary in their gene contents and population frequencies; however, more than 13% of MCMs are homozygous. Additionally, we have thoroughly characterized the MCM MHC class I region (Mafa). Thus, all of the studied alleles and their relative transcript abundances are known. Here, we collectively show that transcriptionally abundant alleles are largely responsible for restricting SIV-specific CD8\textsuperscript{+} T cell responses. Thus, researchers can prioritize a subset of transcriptionally abundant MHC class I alleles for CD8\textsuperscript{+} T cell characterization, facilitating rapid correlation of NHP MHC genetics with successful immune responses.

**MATERIALS AND METHODS**

**High-throughput pyrosequencing.** cDNA was synthesized from total cellular RNA using a SuperScript III first-strand synthesis system (Invitrogen, Carlsbad, CA). Primary cDNA PCR products for each amplicon (either 190 bp, 368 bp, or 576 bp in length) were then generated and purified as previously described (53). Each PCR primer also contained a 10-bp multiplex identifier incorporated during the primary PCR with adapter sequences for Roche/454 Life Sciences (Roche/454) pyrosequencing. After being purified, samples were normalized and pooled in equimolar ratios for pyrosequencing. Pyrosequencing steps were performed with Roche/454 Genome Sequencer (GS) FLX and GS Junior instruments with titanium amplicon chemistry, using the manufacturer’s (454 Life Sciences, Branford, CT) protocols. Data analysis was performed using Genome Sequencer FLX software (454 Life Sciences, Branford, CT), CodonCode Aligner (CodonCode Corporation, Dedham, MA), and an in-house database of macaque MHC class I sequences. Transcript abundance was determined by dividing the total number of sequence reads detected for each distinct class I sequence by the total number of sequence reads for that animal.

**Nomenclature for MHC class I alleles.** Recently, the official nomenclature of macaque MHC class I transcripts has been updated. Briefly, colons are now included in the names such that Mafa-A4*0101 is now Mafa-A4*01:01. Additionally, MCM nomenclature has been updated so that the lineage is similar across macaque species, facilitating broad comparisons. More information on the nomenclature and sequences used in this paper can be found at http://www.ebi.ac.uk/ipd/mhc/nhp/index.html.

**Animals and infections.** In the current study, MCMs were used (data not shown). They were infected intrarectally with 50,000 50% tissue culture infective doses (TCID\textsubscript{50}) of molecularly cloned SIVmac239 Nef open virus. All animals were cared for at the Wisconsin National Primate Research Center (University of Wisconsin, Madison, WI) according to experimental protocols approved by the University of Wisconsin Research Animal Resources Center. A panel of microsatellite markers spanning the MHC region was used to determine the haplotype of the MCMs as previously described by Wiseman et al. (54). Genotypes were confirmed using high-throughput pyrosequencing as previously described (2).

**Peptide binding motifs.** Approximately 25 mg of Mafa-A1*063*02, *A*075*01, and *B*011*01 was harvested from the supernatants of transfected 721.221 cells and purified over an affinity column. Peptides were eluted from the Mafa class I region in 0.2 N acetic acid, which was increased to 10% acetic acid and heated to 78°C for 10 min. Peptides were separated from MHC heavy and light chains by ultrafiltration in a stirred cell with a 3-kDa-molecular-mass-cut-off cellulose membrane (Millipore, Bedford, MA). The peptide pools were flash frozen and lyophilized. The peptides were then reconstituted in 100 μl of PBS and concentrated by ultrafiltration in a stirred cell with a 3-kDa-molecular-mass-cut-off cellulose membrane (Millipore, Bedford, MA). The peptide pools were then reconstituted in 100 μl of PBS and concentrated by ultrafiltration in a stirred cell with a 3-kDa-molecular-mass-cut-off cellulose membrane (Millipore, Bedford, MA). The peptide pools were then reconstituted in 100 μl of PBS and concentrated by ultrafiltration in a stirred cell with a 3-kDa-molecular-mass-cut-off cellulose membrane (Millipore, Bedford, MA). The peptide pools were then reconstituted in 100 μl of PBS and concentrated by ultrafiltration in a stirred cell with a 3-kDa-molecular-mass-cut-off cellulose membrane (Millipore, Bedford, MA). The peptide pools were then reconstituted in 100 μl of PBS and concentrated by ultrafiltration in a stirred cell with a 3-kDa-molecular-mass-cut-off cellulose membrane (Millipore, Bedford, MA). The peptide pools were then reconstituted in 100 μl of PBS and concentrated by ultrafiltration in a stirred cell with a 3-kDa-molecular-mass-cut-off cellulose membrane (Millipore, Bedford, MA). The peptide pools were then reconstituted in 100 μl of PBS and concentrated by ultrafiltration in a stirred cell with a 3-kDa-molecular-mass-cut-off cellulose membrane (Millipore, Bedford, MA). The peptide pools were then reconstituted in 100 μl of PBS and concentrated by ultrafiltration in a stirred cell with a 3-kDa-molecular-mass-cut-off cellulose membrane (Millipore, Bedford, MA). The peptide pools were then reconstituted in 100 μl of PBS and concentrated by ultrafiltration in a stirred cell with a 3-kDa-molecular-mass-cut-off cellulose membrane (Millipore, Bedford, MA).

**Predictive scoring matrix.** Predictive scoring matrices have been built for the prediction of peptides binding to Mafa-A1*063*02, -B*011*01, and -B*075*01 molecules. The matrices have been integrated into an online prediction server, PREDmafa (http://cvc.dfci.harvard.edu/mafa/). Given an input protein sequence, PREDmafa breaks the sequence into overlapping peptides of the length appropriate to the prediction matrix selected and calculates a binding score for

**TABLE 1. Transcriptionally abundant alleles known to restrict CD8\textsuperscript{+} T cell responses**

<table>
<thead>
<tr>
<th>Macaque species</th>
<th>New name of restricting transcript</th>
<th>Old name of restricting transcript</th>
<th>No. of animals examined</th>
<th>Average transcript abundance (%)</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rhesus</td>
<td>Mamu-A1*002:01</td>
<td>Mamu-A*02</td>
<td>10</td>
<td>18.8</td>
<td>26, 41, 50, 52</td>
</tr>
<tr>
<td></td>
<td>Mamu-B*004:01</td>
<td>Mamu-B*04</td>
<td>2</td>
<td>17.6</td>
<td>5, 7, 8</td>
</tr>
<tr>
<td></td>
<td>Mamu-A*007:01</td>
<td>Mamu-A*07</td>
<td>4</td>
<td>16.5</td>
<td>43</td>
</tr>
<tr>
<td></td>
<td>Mamu-B*008:01</td>
<td>Mamu-B*08</td>
<td>10</td>
<td>15.2</td>
<td>21</td>
</tr>
<tr>
<td></td>
<td>Mamu-A*001:01</td>
<td>Mamu-A*01</td>
<td>10</td>
<td>15.1</td>
<td>1, 6, 12, 35</td>
</tr>
<tr>
<td></td>
<td>Mamu-B*048:01</td>
<td>Mamu-B*48</td>
<td>5</td>
<td>14.8</td>
<td>44</td>
</tr>
<tr>
<td></td>
<td>Mamu-A*008:01</td>
<td>Mamu-A*08</td>
<td>8</td>
<td>12.7</td>
<td>51</td>
</tr>
<tr>
<td></td>
<td>Mamu-A*011:01</td>
<td>Mamu-A*11</td>
<td>5</td>
<td>12.4</td>
<td>5, 7, 8, 45</td>
</tr>
<tr>
<td></td>
<td>Mamu-B*003:01</td>
<td>Mamu-B*03</td>
<td>2</td>
<td>8.7</td>
<td>5, 7, 8</td>
</tr>
<tr>
<td></td>
<td>Mamu-A*001:01</td>
<td>Mamu-A*01</td>
<td>8</td>
<td>7.6</td>
<td>25, 56</td>
</tr>
<tr>
<td></td>
<td>Mamu-B*012:01</td>
<td>Mamu-B*12</td>
<td>9</td>
<td>7.3</td>
<td>51</td>
</tr>
<tr>
<td></td>
<td>Mamu-B*017:01</td>
<td>Mamu-B*17</td>
<td>5</td>
<td>3.6</td>
<td>5, 7, 8, 15, 32</td>
</tr>
<tr>
<td>Pigtail</td>
<td>Mane-A*084</td>
<td>Mane-A*10/16</td>
<td>10</td>
<td>20.6</td>
<td>10, 48, 49</td>
</tr>
<tr>
<td></td>
<td>Mane-A*019</td>
<td>Mane-A*11/12</td>
<td>1</td>
<td>17.4</td>
<td>10, 48</td>
</tr>
<tr>
<td></td>
<td>Mane-B*123:01</td>
<td>Mane-B*10</td>
<td>5</td>
<td>8.5</td>
<td>10, 28, 48</td>
</tr>
<tr>
<td></td>
<td>Mane-A*014:01</td>
<td>Mane-A*17</td>
<td>6</td>
<td>0.8</td>
<td>10, 28, 40</td>
</tr>
<tr>
<td>Cynomolgus</td>
<td>Mafa-B*104:01\textsuperscript{b}</td>
<td>Mafa-B*44</td>
<td>4</td>
<td>64.6</td>
<td>34</td>
</tr>
<tr>
<td></td>
<td>Mafa-A*004:01w:01</td>
<td>Mafa-A*02</td>
<td>3</td>
<td>12.7</td>
<td>13</td>
</tr>
<tr>
<td></td>
<td>Mafa-A*0163</td>
<td>Mafa-A*25</td>
<td>5</td>
<td>9.3</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>Mafa-A*0101:01</td>
<td>Mafa-A*29</td>
<td>5</td>
<td>1.8</td>
<td>3</td>
</tr>
</tbody>
</table>

\textsuperscript{a} Mamu-B*001:01-restricted epitopes may not bind the allele with biologically relevant affinities.

\textsuperscript{b} Mafa-B*104:01 was confirmed in our lab to restrict Gag NA9 and has not been published.
technique (3). Briefly, PBMCs were isolated using Ficoll-Paque Plus (GE Healthcare Bioscience, Uppsala, Sweden) and density centrifugation from whole blood containing EDTA. Subsequently, 10^6 cells in 100 μl of R-10 medium (RPMI 1640 medium; HyClone, Logan, UT) supplemented with 10% fetal calf serum, 1% antibiotic-antimycotic (HyClone), and 1% l-glutamine (HyClone) were added to precoated monkey gamma interferon (IFN-γ) ELISPOTPLUS plates (Mabtech, Inc., Mariemont, OH) with the peptide of interest (10 μM). Full-proteome peptides contained pools of 10 peptides totaling 10 μM, or 1 μM each peptide. For a positive control, 10 μM concanavalin A was added in duplicate to a set of wells. A minimum of three wells in duplicate did not receive any stimulant and served as negative controls. All samples were repeated in duplicate or triplicate. Wells were imaged using an AID ELISPOT reader, and spots were counted using an automated program with fixed parameters. For the assay to be considered successful, a minimum of 100 spots was required in each positive-control well. Experimental responses exceeding the arithmetic mean of the negative-control wells plus two standard deviations were considered positive. The limit of detection was set at 70 spot-forming cells per million PBMCs.

**Generation of M3/M3 BLCLs.** MHC-matched B-lymphoblastoid cell lines (BLCLs) were generated as previously described (47). Briefly, PBMCs were isolated by density centrifugation from whole blood containing EDTA. B cells were then immortalized with medium from an S549 cell line containing herpesvirus papio. Cells were maintained in R-10 medium.

**Generation of MHC class I transfectants.** Transfectants expressing Mafa-A1*063:02, -A4*01:01, -B*011:01, or -B*075:01 were generated using HLA-deficient human B cell line 721.221 as previously described (46). Briefly, full-length amplicons were ligated into pcDNA 3.1(+) (Invitrogen, Carlsbad, CA) between the NotI and KpnI restriction sites. Using a Nucleofector C kit (Amaxa, Gaithersburg, MD), the constructs were then transfected into the 721.221 cells. After 3 weeks, transfectants were screened for MHC class I expression using W6/32- phycoerythrin (PE) antibody courtesy of David Watkins. Subsequently, transfectants expressing MHC class I transcripts were magnetically isolated, using a kit from Miltenyi Biotec (Auburn, CA). After magnetic selection, transfectants were maintained under drug selection conditions in R-10 medium containing G418 (Mediatech, Manassas, VA).

**Generation of peptide-specific CD8^+ T cell lines.** CD8^+ T cell lines were generated by incubating 5 × 10^6 freshly isolated PBMCs with 2 μM peptide in complete medium (RPMI 1640 medium supplemented with 15% fetal calf serum, 1% antibiotic-antimycotic, and 1% l-glutamine) with 100 IU of interleukin-2 (NIH AIDS Research and Reference Reagent Program). After 1 week, and every 2 weeks thereafter, the cells were restimulated with peptide-pulsed, irradiated (9,000 rads) BLCLs, as previously described (3).

**ICS.** Intracellular expression of IFN-γ and tumor necrosis factor-α (TNF-α) was used to measure the activation of CD8^+ T cell lines as previously described.
Briefly, 10^5 peptide-pulsed BLCLs or MHC class I transfectants were incubated with 2 × 10^5 cells from corresponding CD8+ T cell lines for 5 h in the presence of brefeldin A (Sigma-Aldrich, St. Louis, MO). Cells were surface stained with CD8 Pacific Blue (PB) and CD3 Alexa Fluor 700 (BD Biosciences, San Jose, CA), washed, and fixed with 2% paraformaldehyde at 4°C for at least 30 min. Cells were then washed, permeabilized, and stained intracellularly with IFN-γ fluorescein isothiocyanate (FITC) and TNF-α PE (BD Biosciences) for 30 min before being washed and fixed in 2% paraformaldehyde. Following staining, samples were run on a BD LSRII instrument, and data were analyzed using FlowJo software, version 8.8.2 (TreeStar, Ashland, OR). All results of intracellular cytokine staining (ICS) were confirmed in at least three individual experiments.

### Tetramers and tetramer staining

Tetramers were generated by David Price and Emma Gostick as previously described (16). Tetramers were conjugated to either allophycocyanin (APC) or PE courtesy of Nancy Wilson. Fresh and frozen PBMCs were stained with tetramers as previously described by Allen et al. (1). Briefly, 5 × 10^6 cells were resuspended in 100 μl of RPMI containing 5 μg/ml of the tetramer of interest for 90 min. Subsequently, cells were surface stained with CD3 AF-700, CD4 PE-Cy5.5, and CD8 PB (BD Biosciences), washed, and fixed with 2% paraformaldehyde. Samples were run on a BD LSRII instrument, and data were analyzed with FlowJo software, version 8.8.2.

### IVSAs

Previously described techniques were modified to examine the ability of CD8+ T cell lines to suppress viral replication in vitro (14). Briefly, CD8+ T cells were removed from PBMCs using a kit from Miltenyi Biotech, and the remaining target cells were stimulated with concanavalin A. After 4 days, activated target cells were magnetoinfected with SIVmac239 as previously described (42). Next, 2.5 × 10^6 infected target cells were plated in 96-well round-bottom plates with SIV-specific CD8+ T effector cells added at effector-to-target ratios of 1:1, 1:10, and 1:20. Targets and effectors were cocultured for 4 days before being stained with CD8 PB, CD4 APC, and CD3 AF-700 (BD Biosciences). Plates were washed twice and fixed with 2% paraformaldehyde at 4°C for 30 min. Plates were then washed and stained with intracellular Gag p27 conjugated to FITC (NIH AIDS Research and Reference Reagent Program) in the presence of brefeldin A (Sigma-Aldrich, St. Louis, CA), washed, and stained with CD8 Pacific Blue (PB) and CD3 Alexa Fluor 700 (BD Biosciences) for 30 min before being washed and fixed in 2% paraformaldehyde at 4°C. After 30 min, plates were washed twice and fixed with 2% paraformaldehyde. Samples were run on a BD LSRII instrument with a high-throughput system (HTS), and data were analyzed with FlowJo software, version 8.8.2. In vitro viral suppression assays (IVSAs) were performed in triplicate in at least two separate experiments.

### Viral sequencing

Cell-free plasma was obtained using Ficoll-Paque Plus density centrifugation from whole blood containing EDTA. Following the manufacturer’s instructions, viral RNA was isolated from the plasma using a Qiagen MinElute viral RNA isolation kit (Qiagen, Valencia, CA). Sixteen overlapping PCR amplicons spanning the entire SIVmac239 genome were then amplified using a Qiagen OneStep reverse transcription (RT)-PCR kit (Qiagen). The conditions for RT-PCR were as described previously (37). Products were then gel purified and bidirectionally sequenced with a DYEnamic energy transfer terminator cycle sequencing kit (GE Healthcare, Piscataway, NJ) on an ABI 3730xl instrument (Applied Biosystems, Foster City, CA). CodonCode Aligner (CodonCode Corporation, Dedham, MA) was then used to compare sequences to the molecularly cloned SIVmac239 reference sequence. To ensure the consistency of sample analysis, a custom assembly script was used. Mutations in sequences at 48 weeks postinfection were confirmed using ultradeep pyrosequencing as previously described (37).

### RESULTS

**Alleles known to restrict CD8+ T cell responses are transcriptionally abundant.** MHC class I expression in macaques is significantly more complex than in humans. As characterization of the entire macaque MHC class I region continues to evolve, it will be important to define which alleles are effective in restricting CD8+ T cell responses. Indeed, macaques can have a heterogeneous assortment of up to 20 MHC class I loci per haplotype (4, 38). Due to this high level of polymorphism, multiple shared MHC class I alleles are rarely observed in unrelated feral NHP populations. Thus, most published CD8+ T cell response studies in macaques focus on a single well-defined MHC class I allele rather than an entire MHC class I haplotype.

To date, only 20 published MHC class I NHP alleles have been shown to restrict CD8+ T cell responses (Table 1). To examine any trends in the transcriptional abundances of these alleles, we retrospectively examined high-throughput pyrosequencing data from more than 200 animals. Interestingly, 19 of these 20 alleles are transcriptionally abundant (expressed on more than 1% of total MHC class I transcripts in PBMCs). One allele in a pigtail macaque, Mane-A4*014:01, was expressed on less than 1% of total MHC class I transcripts in PBMCs (Table 1). Collectively, these results suggest that initially focusing on transcriptionally abundant MHC class I alleles on a haplotype will maximize our understanding of effective SIV-specific CD8+ T cell responses.
Only transcriptionally abundant M3 MHC class I transcripts restrict epitopes identified by IFN-γ ELISPOT assay. Variation in the numbers of MHC class I alleles expressed per haplotype combined with an incomplete knowledge of all possible macaque MHC class I alleles hinders research of cellular immunity in macaque species. In contrast, MCMs have extremely simple MHC genetics consisting of seven major haplotypes, termed M1 to M7. These seven haplotypes vary in their gene contents and population frequencies; however, more than 16% of MCMs carry at least one copy of the MHC class I M3 haplotype, and nearly 4% are M3 homozygous. Additionally, the MHC genetics of MCMs have been solved; therefore, all of the transcripts on the M3 haplotype that could potentially restrict SIV-specific CD8+ T cell responses are known.

Each MCM haplotype contains an average of 10 distinct MHC class I transcripts (range, 9 to 12), double the number of HLA class I transcripts expressed on a human haplotype. Additionally, while humans express relatively equal distributions of MHC class I transcripts on total PBMCs, nearly half of the MCM MHC class I transcripts, termed minor transcripts, are expressed at a frequency of less than 1% (see supplemental Fig. 1 at https://xnight.primate.wisc.edu:8443/labkey/filecontent/WNPRC/WNPRC_Laboratories/oconnor/public/publications

FIG. 3. Detection of 12 potential SIV-specific peptide responses by IFN-γ ELISPOT assay. Freshly isolated PBMCs from eight SIVmac239-infected M3 MCMs were pulsed with 10 μM peptide in triplicate and examined for their abilities to stimulate a response. Positive responses exceeded the average value for the negative control plus two standard deviations, with a limit of detection set at 70 spot-forming cells per million PBMCs. Positive responses were also required to be present in at least two animals or at two time points. (A) Responses observed during acute infection, denoted as less than 20 weeks postinfection. (B) Responses observed during chronic infection, denoted as more than 20 weeks postinfection. SFC, spot-forming cells.
The M3 haplotype encodes five major classical MHC class I transcripts, Mafa-A1*063:02 (22.3%), Mafa-A4*01:01 (1.7%), Mafa-B*011:01 (22.9%), Mafa-B*075:01 (37.5%), and Mafa-B*079:01 (4.8%), as well as four minor transcripts, Mafa-A2*05:11 (0.7%), Mafa-A2*05:11 (0.3%), Mafa-B*165:01 (0.1%), and Mafa-B*070:02 (0.1%), and several nonclassical transcripts, including Mafa-A1*063:01 (0.4%), a B-like transcript (Fig. 1). Interestingly, Mafa-A1*063 and Mafa-A4*01:01 are expressed on the three most common haplotypes, M1, M2, and M3.

To examine the haplotype as a functional unit during SIV infection, we infected M3 MCMs and scanned the entire SIV proteome with overlapping peptide pools. Using this conventional technique, we identified eight responses present in at least two animals; three of these responses were previously described (Fig. 2). Of the three previously described responses, two were restricted by Mafa-A1*063 and one was restricted by Mafa-A4*01:01 (3). To determine the MHC restriction levels of the five novel CD8+ T cell responses identified by full-proteome ELISPOT assay, we created CD8+ T cell lines via stimulation with the 15-mer peptides that produced positive responses and performed ICS (see Fig. 4A). For a positive control, CD8+ T cell lines were stimulated with peptide-pulsed MHC-matched BLCLs. We identified that Env 333-347 was restricted by Mafa-A1*063:02; Rev 59-67, Tat 41-55, and Gag 145-159 were restricted by Mafa-B*075:01; and Gag 25-39 was restricted by Mafa-B*011:01 (see Fig. 4B). Importantly, all of the CD8+ T cell responses in M3 MCMs identified by full-
Characterization of three M3 MHC class I peptide binding motifs. Peptide binding motifs can facilitate identification of novel epitopes without prior knowledge of immune responses. Based on information from the full-proteome ELISPOT assay, we generated peptide binding motifs for the three most frequent M3 Mafa-A and -B transcripts. Combined, these three transcripts account for 87% of the MHC class I sequences expressed on total PBMCs from M3 MCMs. The peptide binding motif for Mafa-A1*063:02 was previously published (3). The binding motifs for Mafa-B*011:01 and -B*075:01 were defined by sequencing endogenous peptides bound to these MHC proteins and identifying the amino acids that bound most frequently at each position (Table 2).

The peptide binding motifs were then used to create a peptide-scoring matrix for the prediction of SIVmac239 peptides that could bind Mafa-A1*063:02, -B*011:01, and -B*075:01. Briefly, the SIV genome was input into PREDmafa, an online prediction server, broken down into overlapping peptides of the lengths set in the program, and scored for similarity to each M3 MHC class I peptide binding motif. The overlapping SIV peptides were then displayed in a table with their prediction scores; the higher the score, the greater the similarity to the peptide binding motif. For cost effectiveness, we assessed the top 1 to 2% of SIV-derived peptides for CD8+ T cell recognition: 24 SIV peptides predicted to bind Mafa-A1*063:02 and -B*075:01 and 48 SIV peptides predicted to bind Mafa-B*011:01 (see supplemental Table 2 at https://xnight.primate.wisc.edu:8443/labkey/filecontent/WNPRC/WNPRC_Laboratories/oconnor/public/publications/begin.view).

*Ex vivo* detection of 12 potential peptide responses in M3 MCMs. To determine whether the 96 peptides predicted to bind either Mafa-A1*063:02, -B*011:01, or -B*075:01 identified CD8+ T cell responses in *vivo* more efficiently than traditional whole-proteome techniques, we performed IFN-γ ELISPOT assays using freshly isolated PBMCs from eight SIVmac239-infected M3 MCMs. Twelve of the 120 peptides examined elicited IFN-γ production in PBMCs: Mafa-A1*063:02 (3/24), -B*075:01 (6/24), and -B*011:01 (3/48), (Fig. 3; see also supplemental Table 2 at https://xnight.primate.wisc.edu:8443/labkey/filecontent/WNPRC/WNPRC_Laboratories/oconnor/public/publications/begin.view). Additionally, three responses were observed by both scanning the SIV proteome with the peptide binding motif and the traditional full-proteome ELISPOT assay.

Considerable variability existed in the magnitude and timing of the peptides with positive responses. Only one peptide, Env 338-346, was positive in all of the M3 MCMs tested. Also, the positive responses were observed for peptides within proteins expressed early and late during SIV replication. As expected, peptides predicted to bind Mafa-B*011:01 or -B*075:01 did not produce positive responses in animals lacking the M3 haplotype (data not shown). Due to the exceptionally high frequency of Mafa-A1*063 (88% of MCMs), we were unable to examine the peptides in animals lacking this transcript. No reactivity to these peptides was observed in uninfected animals (data not shown), confirming SIV specificity.

Seven polyclonal CD8+ T cell lines were successfully created via stimulation with the peptides that produced positive responses by peptide scanning (Fig. 3). To confirm the MHC restriction of the CD8+ T cell lines, 721.221 transfectants expressing either Mafa-A1*063:02, -A4*01:01, -B*075:01, or -B*011:01 were constructed and used for ICS (Fig. 4A). All CD8+ T cell lines were activated by peptide-pulsed MHC-matched BLCLs. Two cell lines, Tat 59-67 and Env 338-346, were activated by transfers expressing Mafa-A1*063:02. Five cell lines, Rev 59-67, Gag 459-467, Env 620-628, Gag 146-154, and Gag 221-229, were activated by transfers expressing Mafa-B*075:01, as predicted (Fig. 4B). We were unable to establish CD8+ T cell lines against any of the three peptides predicted to be restricted by Mafa-B*011:01; however, one response restricted by Mafa-B*011:01 was identified by full-proteome ELISPOT assay of the M3 animals.

Characterization of nine novel SIV-specific CD8+ T cell responses. To determine the optimal epitopes bound by Mafa-A1*063:02, -B*011:01, or -B*075:01 identified by either full-
proteome ELISPOT assay or peptide scanning, we pulsed BLCLs with panels of overlapping 8-, 9-, 10-, or 11-mer peptides. CD8+ T cells were then added, and activation was measured by intracellular staining of IFN-γ and TNF-α (data not shown). Peptides producing positive responses were further examined by pulsing BLCLs with serial dilutions of these peptides to determine optimal responses (Fig. 5). Five of the CD8+ T cell lines started from peptides predicted by PREDMafa were specific to only that peptide, and no serial dilutions were performed. As a result of this mapping, four functional MHC class I tetramers, Env RF9, Rev SP10, Gag HL9, and Tat QA8, were generated (Table 3).

To examine the antiviral efficacy of the nine novel SIV-specific CD8+ T cell responses, we modified an ex vivo viral suppression assay (VSA) (23). As controls, target cells with no effectors were included, and samples were normalized to the percentages of infection present in those wells (Fig. 6A). Targets and effectors were MHC matched; thus, viral suppression occurred as a result of SIV-specific CD8+ T cell responses and not because of the presence of allogeneic antigens (14).

SIV-specific CD8+ T cell lines suppressed viral replication at various levels. Five cell lines, Gag TV9, Env TL9, Gag PR9, Rev SP10, and Gag HL9, were able to suppress more than 90% of viral replication at all effector-to-target ratios examined (Table 3 and Fig. 6B). Tat QA8, Gag KA10, Tat CF9, and Env RF9 suppressed at least 40% of viral replication at a 1:1 effector-to-target ratio and less than 40% at the 1:10 and 1:20 effector-to-target ratios (Table 3 and Fig. 6B). The SIV-specific CD8+ T cell lines that suppressed more than 90% of viral replication were directed against epitopes from both late (Gag and Env) and early (Rev) SIV proteins; however, all responses were restricted by Mafa-B*075:01, the most frequent transcript restricting epitope. The frequency of the restricting transcript.

<table>
<thead>
<tr>
<th>Protein</th>
<th>Sequence</th>
<th>Allele specificity</th>
<th>Optimal epitope</th>
<th>Optimal response</th>
<th>Full-proteome ELISPOT</th>
<th>Peptide scanning</th>
<th>% viral suppression</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gag</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TV9</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TL9</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PR9</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HL9</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rev</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SP10</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gag</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TV9</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TL9</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PR9</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HL9</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tat</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>QA8</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 3. Characterization of all M3-restricted SIV-derived peptides identified

---

The percentage of viral suppression was determined at a 1:20 effector-to-target ratio. Data was confirmed by only a single experiment in triplicate. The percentage of viral suppression was determined at a 1:20 effector-to-target ratio.
**DISCUSSION**

MHC class I expression in macaques is significantly more complex than that in humans, and as MHC genotyping technologies become more accurate and efficient, a more thorough understanding of which alleles are effective in restricting SIV-specific CD8\(^+\)/H11001 T cell responses is necessary. Indeed, macaques can have a heterogeneous assortment of up to 20 MHC class I loci per haplotype (4, 38). Due to this high level of polymorphism, most CD8\(^+\)/H11001 T cell response studies focus on a single well-defined MHC class I allele. Consequently, only 20 published MHC class I NHP alleles are known to restrict SIV-specific CD8\(^+\)/H11001 T cell responses against SIV. Importantly, using high-throughput pyrosequencing we found that 19 of the 20 published alleles are transcriptionally abundant (expressed on more than 1% of total MHC class I transcripts in PBMCs). Thus, we hypothesized that researchers should prioritize transcriptionally abundant MHC class I alleles for CD8\(^+\)/H11001 T cell characterization.

To determine if the repertoire of SIV-specific immune responses generated by an individual was dependent on the transcriptionally abundant alleles on a haplotype, we infected M3 MCMs and scanned the entire SIV proteome by IFN-\(\gamma\) ELISPOT assay with overlapping peptide pools (Table 3). In support of our observation, the four M3 MCM alleles that restricted SIV-specific CD8\(^+\)/H11001 T cell responses were transcriptionally abundant by high-throughput pyrosequencing. To expedite the identification of additional novel M3 SIV-specific CD8\(^+\)/H11001 T cell responses, we characterized the peptide binding motifs for the three major Mafa transcripts expressed on the M3 haplotype, Mafa-A1*06302, B*07501, and B*07501, and subsequently generated an online prediction algorithm to scan the entire SIV genome for potential epitopes. We then scanned the top 1 to 2% of predicted SIV epitopes (96 peptides in total) by ELISPOT assay and identified 12 (12%) positive potential CD8\(^+\)/H11001 T cell responses. It is unclear why a majority (84/96) of the predicted SIV epitopes examined did not produce positive responses by ELISPOT assay; however, we hypothesize that these epitopes either have a poor affinity for the allele or are not recognized by the T cell receptor.

**FIG. 6.** *In vitro* viral suppression by SIV-specific CD8\(^+\)/H11001 T cell lines varies. MHC-matched targets were infected with SIVmac239, and CD8\(^+\)/H11001 T cell effectors were added at effector-to-target ratios of 1:1, 1:10, and 1:20. (A) Representative gating strategy used throughout the analysis to determine the percentage of p27\(^+\) cells. Cells were first gated for lymphocytes based on the forward and side scatter, and then CD8\(^+\)/H11001 T cells were eliminated. The p27\(^+\) cells were then gated, and the number of gated cells was expressed as a percentage of the number of target cells and normalized to the values for representative target wells with no effectors. (B) *In vitro* viral suppression by CD8\(^+\)/H11001 T cell lines.
eral blood lymphocytes can vary greatly, ranging from 70% to
that alleles, which are transcriptionally rare on total peripheral
subsets (J. Greene, submitted for publication). This suggests
peripheral blood lymphocytes can be abundant on distinct cell
MHC class I transcripts that are expressed at low levels in total
the M3 haplotype. Additionally, we discovered that specific
motifs can be made for the five transcriptionally rare alleles on
presentation, we are actively examining whether peptide binding
ine if transcriptionally rare alleles play a role in epitope pre-
search.

Notably, this combination of bioinformatics and epitope map-
facilitates rapid identification of SIV-specific CD8+ T cell
responses. The success rate for identification of peptide targets
is 12%, surpassing the 1% (8/840) of epitopes that were iden-
tified using the conventional method of testing overlapping
pools spanning the whole SIV genome. Bioinformatics-driven
peptide scanning missed five responses identified by full-pro-
teome ELISPOT assay; however, we believe this is a limitation
of scanning only 1 to 2% of the top predicted peptides. We are
confident that if we scanned the top 10% of predicted peptides
we would find all of the responses identified by full-proteome
ELISPOT assay. Importantly, the online prediction server
PREDmafa can also be harnessed to identify M3-restricted CD8+ T cell responses against other pathogens (e.g., CMV
and influenza), allowing the MCM model to be better inte-
grated into multiple areas of human infectious disease re-
search.

The relatively simple, solved MHC class I genetics of MCMs
may allow dissection of the advantages and disadvantages of
expressing a complex number of MHC class I alleles. Five of
the M3 MCM MHC class I transcripts identified are expressed
at less than 1% of total MHC class I transcripts in peripheral
blood lymphocytes. We have not identified any MCM CD8+ T
cell responses restricted by MHC class I alleles expressed at
less than 1% of total peripheral blood lymphocytes. To exa-
mine if transcriptionally rare alleles play a role in epitope pre-
sentation, we are actively examining whether peptide binding
motifs can be made for the five transcriptionally rare alleles on
the M3 haplotype. Additionally, we discovered that specific
MHC class I transcripts that are expressed at low levels in total
peripheral blood lymphocytes can be abundant on distinct cell
subsets (J. Greene, submitted for publication). This suggests
that alleles, which are transcriptionally rare on total peripheral
blood lymphocytes, may influence a yet-unappreciated func-
tion in vivo.

In macaques, MHC class I transcript levels on total periph-
eral blood lymphocytes can vary greatly, ranging from 70% to
less than 0.1% (see supplemental Fig. 1 at https://xnight.primate
wisc.edu:8443/labkey/filecontent/WNPRC/WNPRC_Laboratories
/oconnor/public/publications/begin.view). To examine the rela-
tionship between transcriptional abundance and protein ex-
pression, we evaluated one-dimensional isoelectric focusing
(ID-IEF) assay performed on eight MCMs (17). The ID-IEF
strongly suggests the expression of at least five common MHC
class I alleles in two of the MCMs that share the M3 haplo-
type (A1M and A7M). Based on the sensitivity of this assay, the five
shared MHC class I alleles likely correspond to the five tran-
scriptionally abundant alleles on the M3 haplotype. In conclu-
sion, using our improved knowledge of MCM haplotypes we
are actively characterizing the protein expression of all seven
MCM haplotypes in order to better understand transcriptional
abundance.

Rhesus macaques are the best-characterized NHP model for
AIDS. Unfortunately, the evaluation of CD8+ T cell restric-
tion by all rhesus alleles is not feasible with current technolo-
Thus, to identify rhesus alleles that should be prioritized for
epitope discovery, we examined transcript abundances of 731
rhesus macaques. In Table 4, we summarized the top 25 allele
lineages that are expressed at a high frequency on total PBMCs
and that are present in more than 5% of rhesus macaques. We
chose a limit of 5% of rhesus macaques to ensure that the
lineages would be well represented in the animals used in
research. Eight of the 12 allele lineages known to restrict
SIV-specific responses (Table 1) were identified, supporting
the idea that allele triage on the basis of transcript abundance
could identify other alleles that restrict CD8+ T cell responses.
Importantly, the four allele lineages that were missed were
expressed at high levels on total PBMCs but were present in
less than 5% of rhesus macaques. Collectively, this table pro-
vides a guide for which allele lineages SIV researchers using
rhesus macaques should prioritize for future epitope discovery
studies.

In conclusion, MHC class I expression in macaques is sig-
nificantly more complex than that in humans, with more than
1,000 MHC class I sequences from rhesus macaque species
characterized to date. Here, for the first time, using a pre-

<table>
<thead>
<tr>
<th>Non-synonymous Mutation</th>
<th>Synonymous Mutation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tat 59-67</td>
<td>CCYHCQCF (C)</td>
</tr>
<tr>
<td>Env 338-346</td>
<td>RPKQAWCF (S)</td>
</tr>
<tr>
<td>Rev 59-68</td>
<td>SFPDPPTDTP (S)</td>
</tr>
<tr>
<td>Gag 459-467</td>
<td>TAPPEDPAV (S)</td>
</tr>
<tr>
<td>Env 620-628</td>
<td>TVWPNASL (S)</td>
</tr>
<tr>
<td>Gag 221-229</td>
<td>PapQQGQRK (S)</td>
</tr>
<tr>
<td>Gag 146-154</td>
<td>HLPSPRTL (S)</td>
</tr>
<tr>
<td>Tat 42-49</td>
<td>FYVRPLEA (S)</td>
</tr>
<tr>
<td>Gag 28-37</td>
<td>KHYKMHVWA (S)</td>
</tr>
</tbody>
</table>

FIG. 7. Viral sequencing of the nine novel epitopes reveals inconsistent variation in four M3 homozygous MCMs. Plasma was obtained from
each animal at 8, 20, and >48 weeks postinfection, and viral RNA was sequenced. Pink boxes represent nonsynonymous mutations. Blue boxes
represent synonymous mutations. Dashes denote amino acids that are unchanged from the wild-type reference sequence. X, amino acid position
in which the viral variation is not yet fixed to a single amino acid; N/A, sequences that were unavailable.
screening of T-cell epitopes followed by experimental testing we were able to deal with the combinatorial complexity of alleles and simultaneously characterize epitopes from transcriptionally abundant alleles on the same haplotype. Additionally, using high-resolution pyrosequencing and traditional full-proteome ELISPOT assays, we determined that transcriptionally abundant alleles are largely responsible for restricting SIV-specific CD8\(^+\) T cell responses. Thus, in order to improve the correlation of MHC genetics with successful immune responses in all macaque species, researchers should prioritize a subset of transcriptionally abundant MHC class I alleles for epitope discovery.

**ACKNOWLEDGMENTS**

This research was supported by National Institutes of Health grants R01 AI077376-01 and R21 AI082880-01 and grant R01 SR01AI084787-02 from the National Institute of Allergy and Infectious Diseases. Work for this manuscript was conducted in a facility constructed with support from Research Facilities Improvement Program grants RR15459-01 and RR020141-01.

We thank David Watkins and his lab members Shari Piaskowski, Matthew Reynolds, and Nancy Wilson. Additionally, we thank David Price and Emma Gostick for their aid in generating tetramers and Ann Detmer for her help with viral sequencing. We thank the staff at the Wisconsin National Primate Research Center, a facility supported by grant P51 RR000167 from the National Center for Research Resources, a component of the National Institutes of Health.

Melia L. Budde is a student in the Cellular and Molecular Biology Program at the University of Wisconsin—Madison.

**REFERENCES**


