The HLA A03 supertype and several pan species major histocompatibility complex class I A allotypes share a preference for binding positively charged residues in the F pocket: implications for controlling retroviral infections

Groot, N.G. de; Heijmans, C.M.C.; Ru, A.H. de; Otting, N.; Koning, F.; Veelen, P.A. van; Bontrop, R.E.

AMERICAN SOCIETY FOR MICROBIOLOGY | Journal of Virology®

GENETIC DIVERSITY AND EVOLUTION

The HLA A03 Supertype and Several *Pan* Species Major Histocompatibility Complex Class I A Allotypes Share a Preference for Binding Positively Charged Residues in the F Pocket: Implications for Controlling Retroviral Infections

Natasja G. de Groot,[a] Corrine M. C. Heijmans,[a] Arnoud H. de Ru,[b] Nel Otting,[a] Frits Koning,[c] Peter A. van Veelen,[b] Ronald E. Bontrop[a,d]

[a]Comparative Genetics and Refinement, Biomedical Primate Research Centre, Rijswijk, The Netherlands
[b]Centre for Proteomics and Metabolomics, Leiden University Medical Center, Leiden, The Netherlands
[c]Immunohaematology and Blood Transfusion, Leiden University Medical Center, Leiden, The Netherlands
[d]Theoretical Biology and Bioinformatics, Utrecht University, Utrecht, The Netherlands

ABSTRACT The major histocompatibility complex (MHC) class I region of humans, chimpanzees (*Pan troglodytes*), and bonobos (*Pan paniscus*) is highly similar, and orthologues of *HLA-A*, *-B*, and *-C* are present in both *Pan* species. Based on functional characteristics, the different HLA-A allotypes are classified into different supertypes. One of them, the HLA A03 supertype, is widely distributed among different human populations. All contemporary known chimpanzee and bonobo MHC class I A allotypes cluster genetically into one of the six *HLA-A* families, the *HLA-A1/A3/A11/A30* family. We report here that the peptide-binding motif of the Patr-A*05:01 allotype, which is commonly present in a cohort of western African chimpanzees, has a strong preference for binding peptides with basic amino acids at the carboxyl terminus. This phenomenon is shared with the family members of the HLA A03 supertype. Based on the chemical similarities in the peptide-binding pocket, we inferred that the preference for binding peptides with basic amino acids at the carboxyl terminus is widely present among the human, chimpanzee, and bonobo MHC-A allotypes. Subsequent *in silico* peptide-binding predictions illustrated that these allotypes have the capacity to target conserved parts of the proteome of human immunodeficiency virus type 1 (HIV-1) and the simian immunodeficiency virus SIVcpz.

IMPORTANCE Most experimentally infected chimpanzees seem to control an HIV-1 infection and are therefore considered to be relatively resistant to developing AIDS. Contemporary free-ranging chimpanzees may carry SIVcpz, and there is evidence for AIDS-like symptoms in these free-ranging animals, whereas SIV infections in bonobos appear to be absent. In humans, the natural control of an HIV-1 infection is strongly associated with the presence of particular HLA class I allotypes. The ancestor of the contemporary living chimpanzees and bonobos survived a selective sweep targeting the MHC class I repertoire. We have put forward a hypothesis that this may have been caused by an ancestral retroviral infection similar to SIVcpz. Characterization of the relevant MHC allotypes may contribute to understanding the shaping of their immune repertoire. The abundant presence of MHC-A allotypes that prefer peptides with basic amino acids at the C termini suggests that these molecules may contribute to the control of retroviral infections in humans, chimpanzees, and bonobos.

KEYWORDS chimpanzee, HLA, bonobo, human immunodeficiency virus, major histocompatibility complex, peptide-binding motif, retroviruses, simian immunodeficiency virus

Citation de Groot NG, Heijmans CMC, de Ru AH, Otting N, Koning F, van Veelen PA, Bontrop RE. 2020. The HLA A03 supertype and several *Pan* species major histocompatibility complex class I A allotypes share a preference for binding positively charged residues in the F pocket: implications for controlling retroviral infections. J Virol 94:e01960-19. https://doi.org/10.1128/JVI.01960-19.

Editor Frank Kirchhoff, Ulm University Medical Center

Address correspondence to Natasja G. de Groot, groot@bprc.nl.

Received 21 November 2019
Accepted 11 February 2020
Accepted manuscript posted online 19 February 2020
Published 16 April 2020

The *Pan* species, comprising chimpanzees (*Pan troglodytes*) and bonobos (*Pan paniscus*), are the closest living relatives to humans, and they shared a common ancestor that lived approximately 5 to 6 million years ago (1). At the nonrepetitive DNA level, humans and chimpanzees display high sequence similarity, approximating 98.7% (2). Moreover, the common ancestry of these species is reflected by highly comparable immune systems. A notable difference between humans and chimpanzees, however, might be found in the way both species handle certain infections and whether they subsequently develop pathology. For example, without medical treatment, the majority of humans may develop AIDS upon infection with human immunodeficiency virus type 1 (HIV-1) within 5 to 10 years. Most chimpanzees that were experimentally infected with HIV-1, however, appear to be relatively resistant to developing the disease (3) although there is one report of a chimpanzee infected with a cocktail of HIV-1 strains that developed signs of AIDS-like symptoms (4). In this animal a recombinant HIV-1 strain emerged that after subsequent transmission caused AIDS-like symptoms in three other chimpanzees as well (5). Large-cohort studies have revealed that specific human individuals, named elite controllers, can naturally control an HIV-1 infection for more than 20 years. Particular major histocompatibility complex (MHC) class I molecules, designated human leukocyte antigen (HLA), appear to be strongly associated with this control (6–8). In the distant past, chimpanzees have experienced a selective sweep in their MHC class I repertoire, and although there may be several candidate pathogens that might have caused this, we have put forward the hypothesis that this may be due to a simian immunodeficiency virus (SIV)-like virus infection (9, 10). As a result of this selective sweep, contemporary living chimpanzees have a reduced MHC class I repertoire that encodes molecules that are able to bind conserved regions of HIV-1/SIVcpz and that are functionally similar to the HLA molecules strongly associated in the control of HIV-1 replication (11).

Chimpanzees, who have their natural habitat in Africa, are considered to be the zoonotic source of the virus that caused the AIDS pandemic in the human population (12). Based on their geographic distribution and mitochondrial DNA (mtDNA) variation, four different subspecies of chimpanzees are recognized: *P. troglodytes troglodytes*, which has its habitat in central Africa; *P. troglodytes schweinfurthii*, which lives in eastern Africa; *P. troglodytes verus*, which inhabits western Africa; and *P. troglodytes ellioti*, which populates Nigeria and parts of Cameroon (13, 14). The bonobo has its natural habitat in different sections of forest lying south of the Congo River (15, 16). Some contemporary animals of the *P. t. troglodytes* and *P. t. schweinfurthii* subspecies harbor natural infections with various strains of SIVcpz (12, 17–20), which is considered to be the ancestral progenitor of HIV-1 (12). Certain *P. t. troglodytes* and *P. t. schweinfurthii* populations display a high prevalence of SIVcpz infection, which may impact the fitness of infected animals (17, 20). There are no records of such contemporary SIVcpz infections in the other two subspecies and the bonobo (21–24). Although records of observations involving free-ranging chimpanzees with AIDS-like symptoms are rare (25, 26), the high prevalence of SIVcpz infections in some chimpanzee populations may influence birth and mortality rates (27). Nevertheless, some zoos and primate centers have housed naturally SIVcpz-infected chimpanzees for decades, without documenting AIDS-like illness or pathology (3, 28). This suggests that the immune system of these animals can control the infection rather effectively. There is one report, however, documenting AIDS-like disease in an elderly captive western African chimpanzee that was experimentally infected with an SIVcpz strain, isolated from another subspecies, 20 years ago (29).

MHC class I molecules play a central role in activating an adaptive immune response in all jawed vertebrate species and act as cell surface molecules that present intracellularly derived peptides to $CD8^+$ T cells. In humans, the classical MHC class I molecules are designated HLA-A, -B, and -C, and, in particular, the HLA-A and -B molecules are predominantly involved in antigen presentation to T cells. Generally, these molecules bind nonamer peptides, and in the event that such peptides are derived from a foreign antigen, this may result in the elimination of the infected cell by the induction of

apoptosis. The peptide-binding groove of the MHC class I molecule accommodates six different pockets, which may interact with its ligand. For most allotypes, the B and F pockets bind the anchor residues of a peptide in which the p2 position of a peptide anchors into the B pocket, and the p9/carboxyl (C) terminus of a peptide anchors into the F pocket. However, considering the polymorphic character of the MHC, the peptide-binding site of each allotype has its own chemical composition, and therefore a different repertoire and length of peptides may be bound. In humans, HLA-A and -B allotypes are grouped into supertypes based on similarity in their peptide-binding profiles, and six HLA-A (A01, A01A03, A01A24, A02, A03, and A24) and six HLA-B (B07, B08, B27, B44, B58, and B62) supertypes can be distinguished (30).

The orthologues of HLA-A and -B are present in chimpanzees and bonobos and designated Patr-A and -B (where Patr indicates *Pan troglodytes*) and Papa-A and -B (where Papa indicates *Pan paniscus*), respectively. In three chimpanzee subspecies, *P. t. troglodytes*, *P. t. schweinfurthii*, and *P. t. verus*, and in the bonobo, the MHC class I repertoire has been studied thoroughly at the genetic level (31–36). At this stage, 41 different *Patr-A* and 84 *Patr-B* alleles and 21 *Papa-A* and 32 *Papa-B* alleles have been described (37–40). The alleles within a given locus are highly related to each other, which is in line with an ancient selective sweep in the MHC region that occurred before the speciation of chimpanzees and bonobos (10, 32). For the chimpanzee, in the case of only six *Patr-A* and six *-B* alleles, the peptide-binding motif of the corresponding allotypes has been reported in the scientific literature (11, 41, 42). Here, we describe the definition of the peptide-binding profile of Patr-A*05:01, an allotype observed in the *P. t. verus* cohort, formerly housed at the Biomedical Primate Research Centre (BPRC), with an estimated 7% gene frequency (33). We compared the B- and F-pocket compositions of known HLA-A, Patr-A, and Papa-A allotypes with those of Patr-A*05:01 and demonstrated that humans, different chimpanzee subspecies, and bonobos share several MHC-A allotypes with pocket characteristics that are similar or identical to those of Patr-A*05:01. Subsequently, simulations of peptide binding using the NetMHCpan servers has illustrated that parts of the HIV-1/SIVcpz Gag and Nef proteome can be targeted by Patr-A*05:01 and other functionally related Patr-A and Papa-A allotypes, and corresponding cytotoxic T lymphocyte (CTL) epitopes have been defined in humans for the majority of these regions (www.hiv.lanl.gov).

RESULTS

Definition of the peptide-binding profile of Patr-A*05:01. The cohort of western African chimpanzees, formerly housed at the BPRC, is made up of descendants of a founder population of 35 animals that were caught in the wild. Most of these animals and their offspring were characterized thoroughly for their MHC repertoire by various complementary techniques, and mtDNA d-loop analyses showed that the initially imported animals were unrelated (33, 43). Within this founder population, 12 distinct *Patr-A* alleles were defined. Four of these alleles, *Patr-A*01:01*, *-A*03:01*, *-A*07:01*, and *-A*09:01*, displayed frequencies above 10% in the founder population (33), and for the corresponding allotypes the peptide-binding motif was defined (11, 41, 42). The alleles *Patr-A*04:01*, *-A*05:01*, *-A*06:01*, and *-A*14:01* showed frequencies of 8.8, 7, 7, and 7%, respectively, whereas the *Patr-A*02:01*, *-A*03:02*, *-A*06:02*, and *-A*11:01* alleles were relatively rare and showed frequencies of 1.8%.

From the alleles characterized by intermediate frequencies, only the peptide-binding profile of Patr-A*04:01 has been reported (42). To extend the knowledge relating to the functional capacity of the alleles with intermediate frequencies, we wanted to define the peptide-binding motif of Patr-A*05:01. Therefore, a stable single-molecule-expressing cell line for Patr-A*05:01 was constructed by transfection. Subsequently, the MHC molecules were isolated by immunoprecipitation, and the pool of naturally bound peptides was extracted. The amino acid sequences of the isolated peptides were identified by tandem mass spectrometry. This one-round approach resulted in the identification of 737 natural peptides that included 8-, 9-, 10-, and 11-mers (see Table S1A in the supplemental material), with a Mascot ion score above

35.0 (44). Of the naturally eluted peptides, 2% were 8-mer, 30% were 9-mer, 43% were 10-mer, and 25% were 11-mer, which suggests that the binding of 10-mers is favored slightly by Patr-A*05:01.

The pool of eluted 9-, 10-, and 11-mer peptides illustrated that Patr-A*05:01 favors primarily the aliphatic and aromatic amino acids leucine (L), phenylalanine (F), tyrosine (Y), isoleucine (I), valine (V), and tryptophan (W) but also methionine (M), which contains sulfur in its side chain, at the second residue (p2) of the peptide. At the C terminus (pΩ) of the peptide, a very strong preference for the basic amino acid arginine (R) was observed, as on average approximately 91% of the eluted 9-, 10-, and 11-mer peptides contained that amino acid at this position (Fig. 1A and B). In addition, the basic amino acid lysine (K) is tolerated at the C terminus, and on average 5% of the 9-, 10-, and 11-mer peptides harbored this amino acid. The number of 8-mer peptides was too low to allow definition of a clear peptide-binding motif, but a preference for arginine at the C terminus is evident for the majority of the extracted peptides (Table S1A).

Several human, chimpanzee, and bonobo MHC-A allotypes share a preference for binding basic amino acids in the F pocket with Patr-A*05:01. The alpha 1 and 2 domains of an MHC class I molecule contain the amino acid residues constituting the B and F pockets. An overview of the chemical constitution of the B and F pockets of the chimpanzee and bonobo MHC-A allotypes as well as six defined HLA-A supertypes has been provided (Fig. 2). A comparison of Patr-A*05:01 with the other known Patr-A allotypes reveals that it shares a highly similar B pocket and identical F pocket composition with the Patr-A*04 lineage allotypes and Patr-A*27:01 (Fig. 2). This suggests that these allotypes can present analogous repertoires of peptides. This is substantiated by the experimentally defined peptide-binding motif of a member of the Patr-A*04 lineage, Patr-A*04:01, which displays similarity to the peptide-binding motif of Patr-A*05:01 (42). The Patr-A*27:01 allotype is observed in eastern (*P. t. schweinfurthii*) and central (*P. t. troglodytes*) African chimpanzees (34), and this indicates that the functional capacity to present peptides with arginine (R) and lysine (K) at the C-terminal end is distributed among the different subspecies.

The preference for arginine (R) and lysine (K) at the C-terminal end in the peptide-binding motif of Patr-A*05:01 is also observed for Patr-A*01:01 and -A*03:01 (Fig. 2) (11, 41, 42). The F pocket of Patr-A*01:01 and -A*03:01 differs slightly from that of Patr-A*05:01, and at position 97 a leucine (L) is replaced by an isoleucine (I) (L97I) (Fig. 2). Both chemical structures of the F-pocket, DDTLYILRDYWTKW and DDTLYIIRDYWTKW (underlining indicates sequences discussed above), are also found among other Patr-A allotypes (Fig. 2), suggesting that the preference for binding arginine (R) or lysine (K) at the C terminus is widely covered by various chimpanzee MHC-A allotypes. More than one-third of the currently known Patr-A allotypes encode one of the above-described F pockets, and these allotypes are observed in western, central, and eastern African chimpanzees (31, 33, 34).

Recently, along with other investigators, we have reported on the MHC class I variation in bonobos and concluded that the class I repertoire in this species, in particular *Papa-B*, is more reduced than that in chimpanzees (32, 34, 35). For the 21 *Papa-A* alleles, intron 2 analyses revealed the discovery of only two different sequences, which appear to be identical to the equivalents detected in chimpanzees, suggesting that the entire *Papa-A* repertoire stems from two ancestral structures (32). A comparison of the B- and F-pocket compositions of the Papa-A allotypes showed that in total approximately 62% have an F-pocket structure identical to that of Patr-A*05:01 or to Patr-A*01:01 and Patr-A*03:01 (Fig. 2). This illustrates that a large proportion of the Papa-A allotypes may also have the functional capacity to bind peptides with an arginine (R) or lysine (K) at the C-terminal end of the peptide. Predictions of the peptide-binding specificities of these Papa-A allotypes using the MHCcluster program (45) substantiated the preference for arginine or lysine at their F pockets (Fig. 3; Table S2). The high number of Papa-A allotypes that share these specific F-pocket compositions suggests that they are advantageous for the species.

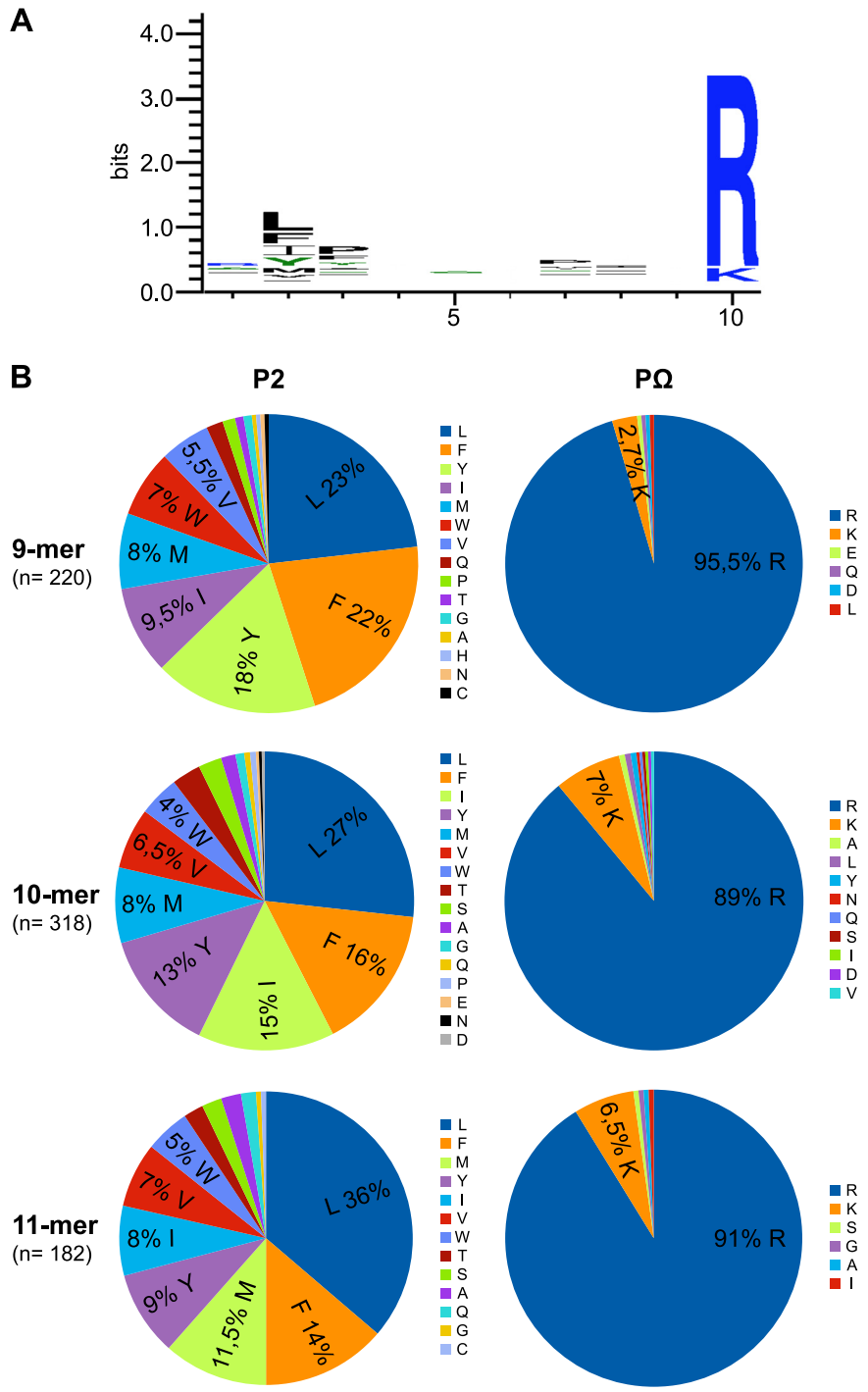

FIG 1 Peptide-binding motif of Patr-A*05:01. (A) The motif was extracted by comparing 318 eluted natural 10-mer peptides. (B) Pie charts illustrating the amino acid preferences (in percentages) at the second residue (p2) and carboxyl terminus (pΩ) of the natural, eluted 9-, 10-, and 11-mer peptides of Patr-A*05:01. The legend on the right of each pie chart depicts the amino acids in accordance with the conventional one-letter code. Table S1B in the supplemental material provides a complete overview of the different amino acids present at a particular position in the eluted 9-, 10-, and 11-mer peptides complemented with the number of times the amino acid has been detected at a particular position.

In humans, it is the A03 supertype that likes to bind the basic amino acids arginine (R), lysine (K), and histidine (H) at the C-terminal end of the peptide. To some extent, the A01A03 supertype also prefers these amino acids, next to aromatic amino acids (30). Reference allotypes within the A03 supertype cluster are HLA-A*03:01, -A*11:01, -A*31:

MHC allotype	B-Pocket											F-pocket																				Peptide-binding motif
	7	9	24	25	34	45	63	66	67	70	99	74	77	80	81	84	95	97	114	116	123	133	143	146	147							
HLA-B*57:01	Y	Y	A	V	V	M	E	N	M	S	Y	Y	N	I	A	Y	I	V	D	S	Y	W	T	K	W							
Patr-A*01:01	-	F	-	-	-	-	-	S	A	H	-	D	D	T	L	-	-	I	R	D	-	-	-	-	-							x-[FMLI(VAST)]-x-x-x-x-x-x-[RK]
Patr-A*02:01	-	F	-	-	-	-	-	S	A	H	-	D	D	T	L	-	-	I	R	D	-	-	-	-	-							
Patr-A*03:01	-	-	-	-	-	-	-	-	-	-	-	D	D	T	L	-	-	I	R	D	-	-	-	-	-							x-[ST]-x-x-x-x-x-x-[RK]
Patr-A*03:02	-	-	-	-	-	-	-	-	-	-	-	D	D	T	L	-	-	I	R	D	-	-	-	-	-							
Patr-A*03:03	-	-	-	-	-	-	-	S	V	-	-	-	-	T	L	-	-	I	R	D	-	-	-	-	-							
Patr-A*03:04	-	F	-	-	-	-	-	-	-	-	-	D	-	T	L	-	-	I	R	D	-	-	-	-	-							
Patr-A*03:05	-	F	-	-	-	-	-	-	-	-	-	D	-	T	L	-	-	I	R	D	-	-	-	-	-							
Patr-A*03:06	-	-	-	-	-	-	-	S	V	-	-	-	-	T	L	-	-	I	R	D	-	-	-	-	-							
Patr-A*03:07	-	-	-	-	-	-	-	S	V	-	-	D	D	T	L	-	-	I	R	D	-	-	-	-	-							
Patr-A*04:01	-	S	-	-	-	-	-	S	V	-	F	D	D	T	L	-	-	L	R	D	-	-	-	-	-							#-x-x-x-#-x-x-x-x-[RK]
Patr-A*04:02	-	S	-	-	-	-	-	S	V	-	F	D	D	T	L	-	-	L	R	D	-	-	-	-	-							
Patr-A*04:04	-	S	-	-	-	-	-	S	V	Q	F	D	D	T	L	-	-	L	R	D	-	-	-	-	-							
Patr-A*05:01	-	S	-	-	-	-	-	S	V	F	F	D	D	T	L	-	-	L	R	D	-	-	-	-	-							x-[LFIYMVW]-x-x-x-x-x-x-x-[R(K)]
Patr-A*06:01	-	S	-	-	-	-	-	S	V	-	-	D	-	T	L	-	-	I	R	D	-	-	-	-	-							
Patr-A*06:02	-	S	-	-	-	-	-	S	V	F	-	D	-	T	L	-	-	I	R	D	-	-	-	-	-							x-x-x-x-x-x-x-x-[Y]
Patr-A*07:01	-	S	-	-	-	-	-	S	V	G	F	-	-	T	L	-	-	L	E	L	-	-	-	-	-							x-[YFMLI(P)]-x-x-x-x-x-x-x-[VLMAFI]
Patr-A*08:01	-	S	-	-	-	-	-	S	V	G	F	-	-	T	L	-	-	I	E	L	-	-	-	-	-							
Patr-A*08:02	-	S	-	-	-	-	-	S	V	G	F	-	-	T	L	-	-	L	E	L	-	-	-	-	-							
Patr-A*08:03	-	S	-	-	-	-	-	S	V	G	F	-	-	T	L	-	-	I	E	L	-	-	-	-	-							
Patr-A*09:01	-	S	-	-	-	-	-	S	V	-	-	D	D	T	L	-	-	I	R	Y	-	-	-	-	-							x-[WYF(AMVISNQ)]-x-x-x-x-x-x-[MLIVA(RFTS)]
Patr-A*09:02	-	S	-	-	-	-	-	S	V	-	-	D	D	T	L	-	-	I	R	D	-	-	-	-	-							
Patr-A*10:01	-	S	-	-	-	-	-	-	-	F	F	-	-	T	L	-	-	L	E	H	-	-	-	-	-							
Patr-A*11:01	-	-	-	-	-	-	-	-	-	-	-	D	-	T	L	-	-	I	R	D	-	-	-	-	-							
Patr-A*12:01	-	S	-	-	-	-	-	S	V	-	F	D	-	T	L	-	-	L	E	H	-	-	-	-	-							
Patr-A*13:01	-	S	-	-	-	-	-	S	V	F	F	D	-	T	L	-	-	L	R	Y	-	-	-	-	-							
Patr-A*14:01	-	S	-	-	-	-	-	S	V	F	F	D	-	T	L	-	-	L	E	H	-	-	-	-	C							
Patr-A*15:01	-	S	-	-	-	-	-	S	V	F	F	D	-	T	L	-	-	L	E	H	-	-	-	-	-							
Patr-A*15:02	-	S	-	-	-	-	-	S	V	F	F	D	-	T	L	-	-	L	E	H	-	-	-	-	-							
Patr-A*15:03	-	S	-	-	-	-	-	S	V	F	F	D	-	T	L	-	-	L	E	H	-	-	-	-	-							
Patr-A*16:01	-	-	-	-	-	-	-	S	A	H	-	D	D	T	L	-	-	L	R	D	-	-	-	-	-							
Patr-A*17:01	-	-	-	-	-	-	-	S	T	-	-	N	D	T	L	-	-	I	R	D	-	-	-	-	-							
Patr-A*17:02	-	-	-	-	-	-	-	S	A	-	-	N	D	T	L	-	-	I	R	D	-	-	-	-	-							
Patr-A*17:03	-	-	-	-	-	-	-	-	-	-	-	D	D	T	L	-	-	I	R	D	-	-	-	-	-							
Patr-A*17:04	-	-	-	-	-	-	-	-	-	-	-	D	D	T	L	-	-	I	R	D	-	-	-	-	-							
Patr-A*18:01	-	S	-	-	-	-	-	S	V	F	F	D	D	T	L	-	-	L	E	L	-	-	-	-	-							
Patr-A*21:01	-	-	-	-	-	-	-	-	V	Q	-	D	D	T	L	-	-	I	R	D	-	-	-	-	-							
Patr-A*22:01	-	S	-	-	-	-	-	S	V	-	F	D	-	T	L	-	-	L	R	D	-	-	-	-	-							
Patr-A*24:01	-	S	-	-	-	-	-	S	V	G	-	D	-	T	L	-	-	I	R	D	-	-	-	-	-							
Patr-A*25:01	-	S	-	-	-	-	-	S	V	G	F	-	-	T	L	-	-	L	E	H	-	-	-	-	-							
Patr-A*26:01	-	S	-	-	-	-	-	S	V	-	F	-	-	T	L	-	-	L	E	H	-	-	-	-	-							
Patr-A*27:01	-	F	-	-	-	-	-	S	V	F	F	D	D	T	L	-	-	L	R	D	-	-	-	-	-							

MHC allotype	B-Pocket											F-pocket																				
	7	9	24	25	34	45	63	66	67	70	99	74	77	80	81	84	95	97	114	116	123	133	143	146	147							
HLA-B*57:01	Y	Y	A	V	V	M	E	N	M	S	Y	Y	N	I	A	Y	I	V	D	S	Y	W	T	K	W							
Papa-A*01:01	-	F	-	-	-	-	-	-	V	H	-	D	-	T	L	-	-	I	R	D	-	-	-	-	-							
Papa-A*02:01	-	S	-	-	-	-	-	S	V	G	F	N	-	T	L	-	-	I	E	Y	-	-	-	-	-							
Papa-A*03:01	-	F	-	-	-	-	-	S	V	Q	-	D	D	T	L	-	-	I	R	D	-	-	-	-	-							
Papa-A*03:02	-	-	-	-	-	-	-	S	V	H	-	D	D	T	L	-	-	I	R	D	-	-	-	-	-							
Papa-A*03:03	-	-	-	-	-	-	-	S	V	-	-	D	D	T	L	-	-	I	R	D	-	-	-	-	-							
Papa-A*03:04	-	F	-	-	-	-	-	S	V	Q	-	D	D	T	L	-	-	I	R	D	-	-	-	-	-							
Papa-A*03:05	-	-	-	-	-	-	-	-	V	Q	-	D	D	T	L	-	-	I	R	D	-	-	-	-	-							
Papa-A*03:06	-	F	-	-	-	-	-	S	V	Q	-	D	D	T	L	-	-	I	R	D	-	-	-	-	-							
Papa-A*04:01	-	F	-	-	-	-	-	S	V	G	F	-	-	T	L	-	-	L	R	D	-	-	-	-	-							
Papa-A*04:02	-	F	-	-	-	-	-	S	V	G	F	-	-	T	L	-	-	L	R	D	-	-	-	-	-							
Papa-A*05:01	-	S	-	-	-	-	-	S	V	H	-	D	D	T	L	-	-	L	R	D	-	-	-	-	-							
Papa-A*05:02	-	S	-	-	-	-	-	S	V	H	-	D	D	T	L	-	-	L	R	D	-	-	-	-	-							
Papa-A*05:03	-	S	-	-	-	-	-	S	V	H	-	D	D	T	L	-	-	L	R	D	-	-	-	-	-							
Papa-A*06:01	-	S	-	-	-	-	-	-	V	-	F	D	D	T	L	-	-	L	R	D	-	-	-	-	-							
Papa-A*06:02	-	S	-	-	-	-	-	-	V	-	F	D	D	T	L	-	-	L	R	D	-	-	-	-	-							
Papa-A*06:03	-	S	-	-	-	-	-	-	V	-	F	D	D	T	L	-	-	L	R	D	-	-	-	-	-							
Papa-A*08:01	-	F	-	-	-	-	-	-	V	Q	-	D	D	T	L	-	-	I	E	H	-	-	-	-	-							
Papa-A*08:02	-	-	-	-	-	-	-	-	V	Q	-	D	D	T	L	-	-	I	E	H	-	-	-	-	-							
Papa-A*08:03	-	F	-	-	-	-	-	-	V	Q	-	D	D	T	L	-	-	I	E	H	-	-	-	-	-							
Papa-A*09:01	-	-	-	-	-	-	-	-	V	-	-	D	D	T	L	-	-	I	R	D	-	-	-	-	-							
Papa-A*09:02	-	-	-	-	-	-	-	-	V	-	-	D	-	T	L	-	-	I	R	D	-	-	-	-	-							

HLA supertype (ref. allotype)																																
A01 (HLA-A*01:01)	-	F	-	-	-	-	-	-	-	H	-	D	-	T	L	-	-	I	R	D	-	-	-	-	-							
A01 A03 (HLA-A*30:01)	-	S	-	-	-	-	-	-	V	Q	-	D	D	T	L	-	-	I	E	H	-	-	-	-	-							
A01 A24 (HLA-A*29:02)	-	T	-	-	-	-	Q	-	V	Q	-	D	-	T	L	-	-	M	R	D	-	-	-	-	-							
A02 (HLA-A*02:01)	-	F	-	-	-	-	-	K	V	H	-	H	D	T	L	-	V	R	H	Y	-	-	-	-	-							
A03 (HLA-A*03:01)	-	F	-	-	-	-	-	-	V	Q	-	D	D	T	L	-	-	I	R	D	-	-	-	-	-							
A24 (HLA-A*23:01)	-	S	-	-	-	-	-	K	V	H	F	D	-	-	-	-	L	M	H	Y	-	-	-	-	-							

FIG 2 Chimpanzee (Patr) and bonobo (Papa) MHC class I A allotypes with their B- and F-pocket composition. The amino acid residues determining the B and F pocket are presented at the top. Those of HLA-B*57:01 are taken as a consensus. Identity to the consensus is indicated by dashes, whereas an amino acid replacement is indicated by the conventional one-letter code. The B- and F-pocket composition of the six HLA-A supertypes is displayed, and this is based on one of the reference allotypes of the supertype, which is indicated in parentheses. The peptide-binding motifs of six different Patr-A allotypes previously experimentally defined are provided. The number sign (#) in the motif of Patr-A*04:01 marks additional primary anchors. The experimentally defined peptide-binding motif of Patr-A*05:01, its corresponding B- and F-pocket composition, and the Patr-A*27:01 allotype with a highly similar B- and identical F-pocket composition are highlighted in yellow. Allotypes with an F-pocket composition identical to that of Patr-A*05:01 are highlighted in light blue; allotypes with an F-pocket composition highly similar to that of Patr-A*05:01 are highlighted in salmon (L97I).

01, -A*33:01 and A*33:03 (A*33:01/03), -A*66:01, -A*68:01, and -A*74:01 (30). The F-pocket composition of these reference allotypes has revealed four different chemical structures: DDTLYIIRDYWTKW, DDTLYIMQDYWTKW, DDTLYIRQDYWTKW, and DDTLYIMRDYWTKW. The first structure is present in HLA-A*03:01 and -A*11:01 and is identical (encoded by the same codon usage) to the F-pocket structure found for particular Patr-A and Papa-A allotypes (Fig. 2, highlighted in salmon). That suggests that this motif is old and has been preserved over a long evolutionary time span and was present in an ancestor of the human *Pan* lineage. The other three structures are detected in the group consisting of HLA-A*31:01, -A*33:01/03, and -A*74:01 in HLA-A*66:01, and in HLA-A*68:01, respectively, and for these structures no equivalents have been observed in the Patr-A and Papa-A allotypes, implying that they are human specific or were lost during the *Pan* species radiation. These structures differ from the first one at position 97 where isoleucine (I) is replaced by methionine (M) or arginine (R) (I97M or I97R) and at position 114 where arginine is replaced by a glutamine (R114Q) in the structures found in the group consisting of HLA-A*31:01, -A*33:01/03, -A*74:01 and in HLA-A*66:01. However, compared to the effects of the structures of HLA-A*03:01 and -A*11:01, the amino acid replacements at the indicated positions do not have an effect on the functional properties of these allotypes since they all preferentially bind peptides with a basic amino acid at the C terminus.

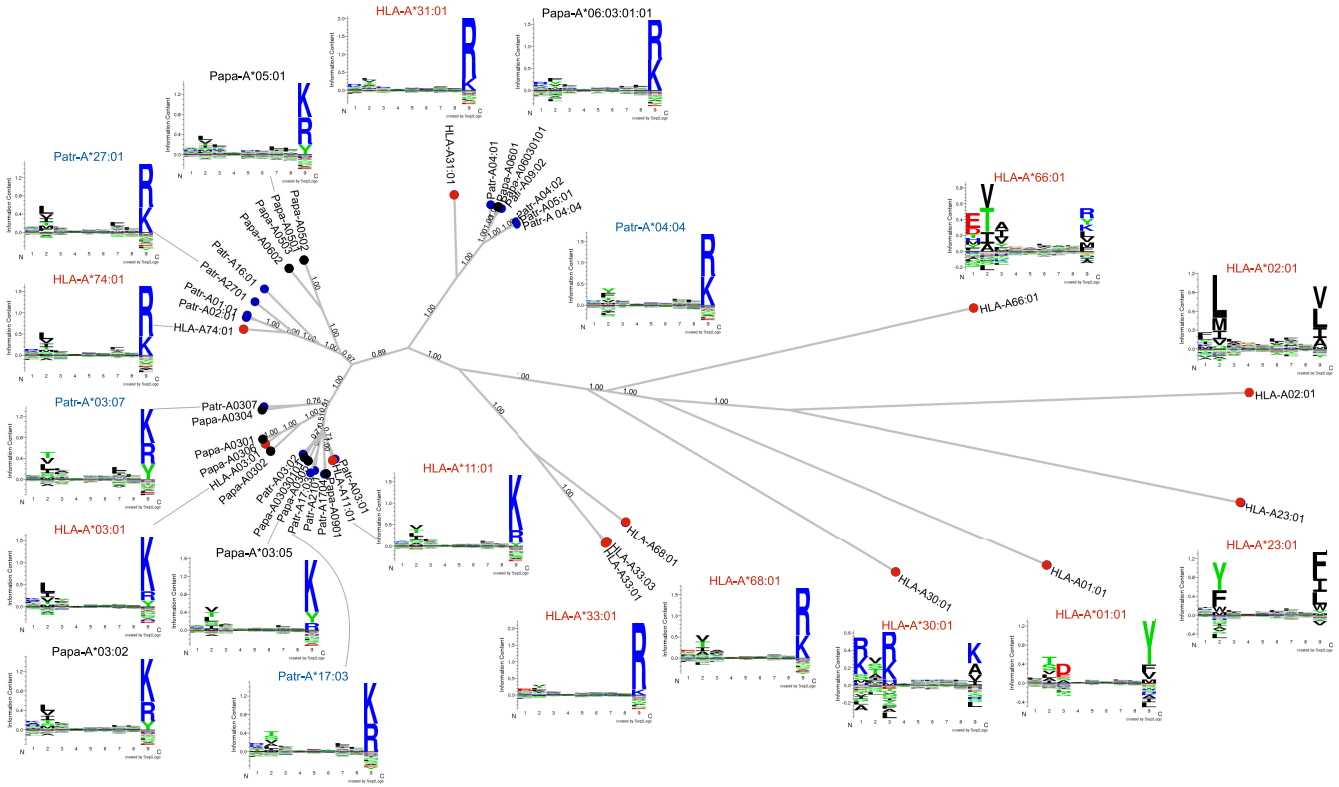

FIG 3 Representative tree illustrating the predicted functional similarity between human (HLA, in red), chimpanzee (Patr, in blue), and bonobo (Papa, in black) MHC class I A allotypes. The reference allotypes of the HLA A01, A02, and A24 supertypes, all reference allotypes with an HLA A03 supertype association, and the Patr- and Papa-A allotypes with an F-pocket structure similar (L97I) or identical to that of Patr-A*05:01 (Fig. 2) are included in the analyses. Illustrative sequence logo plots for several allotypes are provided.

Taken together, the data illustrate that several Patr-A, Papa-A, and HLA-A allotypes share the functional capacity to bind a similar repertoire or even perhaps an extensive set of identical peptides, with a preference for basic amino acids (R, K, or H) at the C-terminal end. The maintenance of this functional capacity at high frequencies in the contemporary living *Pan* species populations (>33% for Patr and ~62% for Papa allotypes), of which the ancestor was affected by an ancient selective sweep in the MHC class I repertoire, suggests that these types of molecules must have played and still play an important biological role in the immune response of these species.

Capacity to target the HIV-1/SIVcpz proteome: an *in silico* prediction. In humans, several HLA allotypes are associated with protection from developing disease after an HIV-1 infection, such as HLA-B*27 (7). This allotype favors binding peptides with an arginine (R) at the p2 position. Within the chimpanzee and bonobo, MHC class I allotypes with a functional capacity similar to that of HLA-B*27 are present (11, 46). In particular, in bonobos these types of allotypes are present at high frequencies (46). Moreover, in the rhesus macaque SIV infection model, control of infection is observed, and also in this species protection is associated with the presence of particular MHC class I allotypes such as Mamu-B*003, -B*008, and -B*017, which all share peptide-binding characteristics similar to those of HLA-B*27 (47, 48). Furthermore, Mamu-A2*05, a highly frequent specialist MHC class I molecule that has family members in other macaque species, has HLA-B*27-like peptide-binding features. This allotype is able to bind SIV epitopes that are also recognized in the context of Mamu-B*003 and -B*008, where they evoke strong $CD8^+$ T-cell responses in SIV-infected rhesus macaques (49). Moreover, other rhesus macaque MHC class I allotypes prefer basic amino acids at the C-terminal end of the peptide (50–52), similar to what is described above for the HLA A03 supertype and for particular Patr-A and Papa-A allotypes.

For humans, the website of the HIV databases (www.hiv.lanl.gov) provides an overview of the known CTL/CD8$^+$ epitopes with their respective HLA restriction elements. For the *in silico* prediction that was performed, we used HIV-1 HXB2 as a model species and focused on the Gag protein as it is generally accepted that adaptive immune responses to this gene product play a crucial role in controlling HIV-1 replication in infected human individuals (53, 54). For the reference allotypes of the HLA A03 supertype (HLA-A*03:01, -A*11:01, -A*31:01, -A*33:01/03, -A*66:01, -A*68:01, and -A*74:01), several HIV-1 (HXB2) Gag CTL epitopes have been defined. In addition, next to Gag, a selection of the reference allotypes of the HLA A03 supertype are seen to dominantly target CTL epitopes of the Nef protein (55, 56). To investigate whether the Patr-A and Papa-A allotypes that prefer basic amino acids at the C-terminal end (pΩ) of the peptide can target conserved regions of the HIV-1/SIVcpz proteome in a manner similar to that of the reference allotypes of the HLA-A03 supertype, we used the NetMHCpan algorithm to predict peptide-binding affinities (57, 58).

The predictions were performed with overlapping nonamer peptides of the Gag sequence of HIV-1 HXB2 and the Nef reference sequence (HIV-1 B-clade) (www.hiv.lanl.gov), as MHC class I molecules generally have a strong preference for binding this peptide length. The predictions generated by the NetMHCpan version 3.0 and 4.0 servers showed that Patr-A and Papa-A allotypes that prefer basic amino acids at pΩ can indeed target regions within HIV-1 Gag (HXB2) and Nef (HIV-1 B-clade), and a different number of nonamer peptides was predicted to bind with high affinity (indicated by the prediction tool as strong binder [SB]) (total data set is not shown).

For the Gag protein, seven nonamer peptides were identified that stand out (RLRPGGKKK, ATLYCVHQR, MVHQAISPR, HQAAMQMLK, NSATIMMQR, IMMQRGNFR, and ELYPLTSLR), and they were predicted to bind with either high (strong binder; SB) or intermediate (weak binder; WB) affinity to many of the Patr-A and Papa-A allotypes that prefer basic amino acids at pΩ (Fig. 4A). Five of the seven peptides (RLRPGGKKK, ATLYCVHQR, MVHQAISPR, HQAAMQMLK, and IMMQRGNFR) are identical or nearly identical to well-defined human CTL epitopes, and these have particular reference allotypes of the HLA A03 supertype as their restriction elements, whereas for two peptides (NSATIMMQR and ELYPLTSLR) the HLA-restriction elements are not known (Table 1). Moreover, the *in silico* prediction revealed that peptide RWIILGLNK, which is a thoroughly defined HLA-B*27-restricted CTL epitope, can also be targeted by several of the Patr-A and Papa-A allotypes that prefer basic amino acids at pΩ (Fig. 4A, positions 259 to 277; Table 1).

For the Nef protein, two nonamer peptides (AVDLSHFLK and RXALRHVAR) were found that are identical or nearly identical to CTL epitopes defined for reference allotypes of the HLA A03 supertype, and were predicted to bind with high (SB) or intermediate (WB) affinity to several of the indicated Patr-A and Papa-A allotypes (Fig. 4B; Table 1). Moreover, the prediction tools found two additional peptides in the Nef protein (SIVGWPTIR and KLVPVEPEK) that can be targeted by several of the Patr-A and Papa-A allotypes that prefer basic amino acids at pΩ. These two peptides were predicted to bind to particular reference allotypes of the HLA A03 supertype as well (Fig. 4B) although a corresponding CTL epitope has not yet been reported in the HIV databases (www.hiv.lanl.gov).

Overall, the *in silico* predictions revealed that the Patr-A and Papa-A allotypes that prefer basic amino acids at pΩ can target conserved regions of the HIV-1 Gag and Nef proteins, and for these regions HLA A03 supertype-restricted CTL epitopes are also defined. These observations suggest that in addition to Patr-B and Papa-B allotypes, which have functional properties analogous to those of the AIDS-protective HLA-B*57 and -B*27 allotypes (11, 46), the gene products of the MHC-A locus may play an additional role in controlling retroviral infections such as HIV-1/SIVcpz.

DISCUSSION

By studying the MHC in our closest living relatives, the chimpanzee and the bonobo, information may be gathered with regard to how this region and its polymorphisms

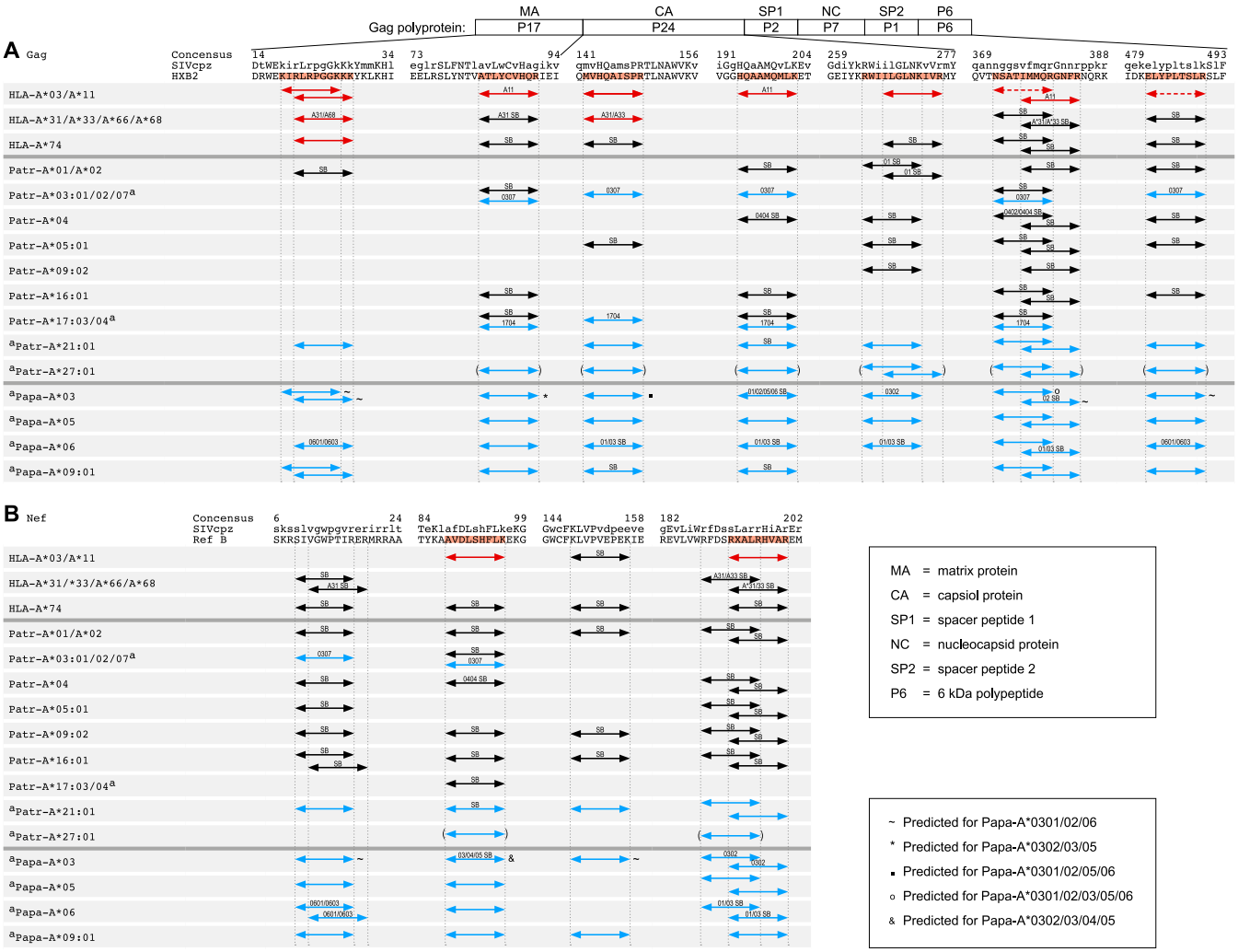

FIG 4 Overview of the HIV-1 Gag (A) and Nef (B) protein regions that are potentially targeted by chimpanzee (Patr) and bonobo (Papa) MHC-A allotypes that prefer basic amino acids at the C-terminal end of the peptide. Parts of the HIV-1 Gag-HXB2 and Nef-B clade and SIVcpz consensus sequences (www.hiv.lanl.gov) are provided at the top. The Gag polyprotein is schematically drawn to indicate from which proteins the amino acid sequences are derived (abbreviations explained in the legend). A lowercase letter in the SIVcpz consensus indicates a variable position. The HLA-A, Patr-A, and Papa-A allotypes investigated (see Table S3 in the supplemental material) are indicated or shown as groups based on a sharing of the lineage (e.g., Patr-A*04 is comprised of the allotypes Patr-A*04:01, -A*04:02, and -A*04:04), identical F-pocket structure (HLA-A*03 and -A*11 and Patr-A*01 and -A*02), or high similarity in predicted peptide-binding profiles (HLA-A*31, -A*33, -A*66, and -A*68). The HIV-1 Gag and Nef CTL epitopes (www.hiv.lanl.gov) that are defined for reference allotypes of the HLA-A03 supertype association (HLA-A*03:01, -A*11:01, -A*31:01, -A*33:01/03, -A*66:01, -A*68:01, and -A*74:01) are indicated by red arrows, and if applicable above the red arrow the CTL-restricted HLA allotype is indicated (details are given in Table 1). A dotted red arrow indicates that the HLA restriction is not specified for the CTL epitope. Black arrows indicate the peptides predicted to bind to the corresponding allotypes using the NetMHCpan 4.0 server. For Patr-A*03:07, -A*17:04, -A*21:01, and -A*27:01 and all Papa-A allotypes, peptide-binding predictions were performed using the NetMHCpan 3.0 server and are indicated by the blue arrows. SB refers to predicted strong binding; e.g., 01/03 SB refers to strong binding of specific alleles in a group, and arrows without any indication indicate peptides that show intermediate binding affinity. When peptide binding is predicted only for a subset of a group of HLA-A, Patr-A, or Papa-A allotypes, this is also indicated by printing the corresponding allotype(s) or four-digit allele number above the arrow or by referring to the legend in the figure. The arrows for Patr-A*27:01 are shown in brackets because most of the predicted binding affinities for this allotype are low (range, 421.3 to 5143.9 nM). A detailed specification of the predicted binding affinity data is provided in Table S4.

evolved during hominoid evolution. Moreover, the characterization of the functional properties of different MHC allotypes in these species may facilitate our understanding of how the repertoire was shaped as a result of selective pressures that impact immunological responses (59). In the present study, we have characterized the peptide-binding motif of Patr-A*05:01, an MHC class I molecule present in moderate frequencies in a population of western African chimpanzees (33). The Patr-A*05:01 allotype primarily binds aliphatic and aromatic amino acids at the p2 position of the peptide side and has a very strong preference for arginine (R) at the C-terminal end of the peptide, which docks into the F pocket. A comparison of the F-pocket structures of the MHC-A locus

TABLE 1 List of defined HIV-1 epitopes in humans that are also predicted to bind to Patr-A and Papa-A allotypes that prefer basic amino acids at pΩ

Protein and epitope[a]	HXB2 start[c]	HXB2 end[c]	Subprotein[d]	HXB2 DNA contig	Subtype(s)	HLA allotype(s)
Gag						
KIRLRPGGK	18	26	p17 (18–26)	841..867	A, B, CRF01_AE	A*0301, A*30, A11, A3, A69, B*82, B27, B7
RLRPGGKKK	20	28	p17 (20–28)	847..873	A, A1, B, C, CRF02_AG, multiple	A*0201, A*0301, A*30, A*74, A11, A3, A30, A31, A68, B42, B62, B7, C*01
ATLYCVHQR	83	91	p17 (83–91)	1036..1062	B	A*11, A*1101
MVHQAISPR	142	150	p24 (10–18)	1213..1239	B, C	A*11, A*31, A*3303, A3, A3 supertype, A33, B*68
HQAAMQMLK	194	202	p24 (62–70)	1369..1395	A1, B	A11, B52
RWIILGLNK	264	272	p24 (132–140)	1579..1605	B	B27
ILGLNKIVR	267	275	p24 (135–143)	1588..1614	B	A11, A3
NSATIMMQR	372	380	p2p7p1p6 (9–17)	1903..1929	B	
IMIQKGNFR[b]	376	384	p2p7p1p6 (13–21)	1915..1941	B	A*1101
ELYPLASLR[b]	482	490	p2p7p1p6 (119–127)	2233..2259	B	
Nef						
AVDLSHFLK	84	92		9046..9072	B	A*0301, A*1101
RLAFHHVAR[b]	188	196		9358..9384	B	A11, A3, B52

[a]All HIV-1 epitopes are from human (www.hiv.lanl.gov). All epitopes except those highlighted in gray are identical or highly similar to the peptides predicted to bind to most of the indicated allotypes in Fig. 4.
[b]Epitopes are highly similar to IMMQRGNFR and ELYPLTSLR indicated in the HXB2 (Gag) consensus sequence and to RXALRHVAR indicated in the Ref B (Nef) consensus sequence in Fig. 4.
[c]Amino acid position.
[d]Amino acid positions are given in parentheses.

gene products revealed that HLA A03 supertype allotypes as well as over 33% of the chimpanzee and 62% of the bonobo MHC-A allotypes share an F-pocket composition that is either identical or highly similar to that of Patr-A*05:01. This suggests that these three species share many MHC-A molecules with the capacity to present peptides with an arginine (R) at the C-terminal end.

The allotype Patr-A*27:01, present in the eastern and central African chimpanzee subspecies (34), shares an identical F pocket and a comparable B pocket with Patr-A*05:01, which is present in western African chimpanzees (Fig. 2, highlighted in yellow). Only at amino acid position 9 of the B pocket is a serine replaced by a phenylalanine (S9F). Compared to serine (S), phenylalanine (F) is a large amino acid with an aromatic ring as a side chain. The presence of phenylalanine (F) at position 9 is also evident for Patr-A*01:01, -A*02:01, -A*03:04, and -A*03:05 (Fig. 2). For Patr-A*01:01, the peptide-binding motif has been defined, and it prefers to bind peptides with phenylalanine (F), methionine (M), leucine (L), or isoleucine (I) at the p2 position (41). These preferences match those found for Patr-A*05:01 and suggest that Patr-A*27:01 is able to bind an identical or highly related array of peptides.

Next to *Patr-A*04:01* and -A**05:01*, *Patr-A*06:01* and *-A*14:01* were also found to display moderate allele frequencies (7%) in the BPRC chimpanzee founder population (33). The Patr-A*06:01 and -A*06:02 allotypes share comparable B-pocket structures and identical F-pocket structures (Fig. 2). For Patr-A*06:02, the peptide-binding motif has been defined and shows that it is promiscuous for the residues docking into the B pocket and prefers to bind peptides with a tyrosine (Y) residue at the C-terminal end. The Patr-A*14:01 allotype has a B-pocket structure identical to that of Patr-A*05:01; however, its F-pocket structure is unique (Fig. 2). Therefore, to characterize the binding specificities of Patr-A*14:01, it will be necessary to resolve its peptide-binding motif in the future.

In humans, the *HLA-A* gene is divided into two lineages comprising six families (60, 61), and this classification is based on the substitution pattern of 33 diagnostic nucleotide positions. From a genetic perspective, the HLA-A2 lineage is comprised of the *HLA-A2*, *-A10*, and *-A19* families, whereas the HLA-A3 lineage consists of the *HLA-A9*, *-A1/A3/A11/A30*, and *-A80* families. Members of the HLA A03 supertype association cluster genetically into different *HLA-A* families. The HLA-A*03:01 and -A*11:01 allotypes

cluster into the *HLA-A1/A3/A11/A30* family. HLA-A*31:01, -A33:01/03, and -A*74:01 cluster into *HLA-A19*; HLA-A*66:01 clusters into *HLA-A10*; and HLA-A*68:01 clusters into the *HLA-A2* family. Thus, at the genetic level, the alleles encoding the reference allotypes of the A03 supertype appear to be distributed among different *HLA-A* families; however, at the functional level, these alleles encode allotypes sharing a preference for binding basic amino acids in their F pockets. Chimpanzees and bonobos possess only MHC class I *A* alleles that are orthologous to the *HLA-A1/A3/A11/A30* family, and alleles that cluster into the other five human *HLA-A* families are absent (32, 33, 62). In the first instance, the homology at the genetic level may suggest that also at the functional level the Patr-A and Papa-A repertoires have been narrowed owing to the selective sweep. Nevertheless, the seven defined chimpanzee peptide-binding motifs, as well as the chemical composition of the F pockets of the allotypes encoded by the *Patr-A* and *Papa-A* alleles, illustrate that chimpanzees and bonobos harbor allotypes with HLA A01 and A03 supertype characteristics and allotypes with an F-pocket structure that has a promiscuous binding characteristic (i.e., Patr-A*09:01) (Fig. 2). Therefore, although the known *Patr-A* and *Papa-A* alleles cluster genetically into only one of the six *HLA-A* families, at the functional level they seem to cover a broader repertoire and cluster into different HLA-A supertype specificities.

In the human population, members of the HLA A02, A03, and B07 supertypes are present in the greatest abundances (30, 51). Approximately 33% of the chimpanzee and 62% of the bonobo MHC-A allotypes prefer to bind basic amino acids in the F pocket, which is a characteristic shared with the A03 supertype. Extrapolation of these data to the chimpanzee and bonobo founder populations that were previously analyzed for their MHC class I repertoire (11, 32) has revealed that 46% of the chimpanzee founder haplotypes and 70% of the bonobo founder haplotypes contain a Patr-A or Papa-A allotype with the F pocket characteristic of the HLA A03 supertype.

A survey of the literature has illustrated that the presence of members of the HLA A03 supertype in HIV-1-infected human individuals is associated with either a lower viral load or control of the development of AIDS (7, 63–65). In particular, CTL responses against the conserved parts of the HIV-1 Nef and Gag proteins are often related to allotypes that fit into the HLA A03 supertype (54–56). The *in silico* modeling of nonamer peptide binding showed that several of these conserved parts of HIV-1 Nef and Gag can also be targeted by Patr-A and Papa-A allotypes that share F-pocket characteristics with the A03 supertype (Fig. 4). This may suggest that the selective sweep in the progenitor species of chimpanzees and bonobos resulted in a skewing of the repertoire that favored the positive selection of A-locus allotypes that can target conserved structures of the retroviral proteome. The modeling was performed for nonamers as this is the preferential peptide length that binds to most HLA allotypes. Caution is advised regarding this assumption, however, because HLA-A*11:01 ligand sequencing, for instance, has revealed that the range of HIV-1 Gag and Nef peptides varies from 7- to 20-mers in length, and these peptides showed binding and are recognized by T cells (66).

The presence of basic amino acids in the different parts of the Gag polyprotein of retroviruses is crucial for genomic RNA binding, capsid assembly, and trafficking (67–70). Nevertheless, it is necessary to consider that in the literature there are reports showing no significant association between control of an HIV-1 infection and the presence of the HLA A03 supertype (71, 72). The abundant presence of MHC allotypes in humans and the *Pan* species that share the preference for basic amino acids at the C terminus suggests that these types of molecules must fulfill an important biological function in these species, and they may play a role in the control of an early HIV-1/SIV infection. Several nonhuman primate species are infected naturally with various SIV strains, and although progression toward AIDS is rarely observed, this certainly may occur (73). In free-ranging chimpanzees, natural infections with SIVcpz strains may result in lower birth rates and increased mortality (27). Studies in captive-living chimpanzees naturally infected with SIVcpz showed, however, that AIDS-like symptoms have not been the cause of death but that this is rather due to cardiac disease (28). In this

line, in the ancestor of the contemporary living chimpanzees and bonobos, an ancient retroviral infection may have selected for the survival of MHC haplotypes that carried a combination of A and B allotypes that can target a differential number of conserved epitopes of the retroviral Gag and Nef proteome. Recognition of these epitopes can occur by way of multiple motifs. In other words, if there is an escape from immune recognition for one allotype, there may be another allotype that prevents such an escape. We have named this phenomenon a double-lock strategy (11). The fact that some human studies fail to find a significant association between control of an HIV-1 infection and the presence of the HLA A03 supertype may be due to the abundant distribution of different HLA allotypes with a preference for binding arginine (R) or lysine (K) at the C-terminal end of a peptide. It will be difficult to tease out an allotype association in population studies, and the present report substantiates the premise that such studies may be best conducted at the supertype level.

MATERIALS AND METHODS

Patr-A*05:01-transfected cell line. In the past, the Biomedical Primate Research Centre housed a large population of chimpanzees that originated with 35 founder animals that had been caught in the wild in Sierra Leone and belonged to the western African subspecies (*Pan troglodytes verus*). Most of these animals have been characterized on the molecular level for their *Mhc* class I gene polymorphism (33), and Epstein-Barr virus transformed B-cell lines are available as a source for RNA/DNA. A B-cell line of a Patr-A*05:01-positive chimpanzee (Diana) was cultured, and cells were harvested to extract total RNA (RNeasy minikit; Qiagen), which served as a template to synthesize cDNA (RevertAid First Strand cDNA synthesis kit; Thermo Fisher Scientific). Patr class I amplification was performed using the primers 5′-GGGAAGCTTGGACTCAGAATCTCCCCAGACGCCGAG-3′ and 5′-GCGATGCATTCTCAGTCCCTCACAAGGC AGCTGTC-3′ as described previously (33). The underlined nucleotides in the primer sequences indicate the restriction element sites for HindIII and NsiI, respectively. Target DNA and vector pcDNAI/Neo were digested with HindIII and NsiI (Invitrogen) and subsequently ligated, cloned, and sequenced as previously described (11, 74). The transfection of the Patr-A*05:01-positive construct into K562 cells was performed as previously described (75), and MHC class I-positive cells were enriched by cell sorting using Dynabeads (Dynal) coated with W6/32 anti-HLA ABC antibody.

Determination of the peptide-binding motif of Patr-A*05:01. The K562 Patr-A*05:01-positive transfectant was cultured up to 10 billion cells in Iscove's modified Dulbecco's medium (IMDM), which was supplemented with 5% fetal calf serum (FCS) and 0.2 mg/ml G418. The culture was checked routinely for MHC expression and mycoplasma infection as described previously (11). The cells were harvested every second day and washed with phosphate-buffered saline (PBS), and the cell pellet was stored at −80°C. From the collection of cells, the MHC molecules were isolated by immunopurification. Subsequently, the pool of naturally bound peptides was eluted, and the amino acid sequences were identified by tandem mass spectrometry as described previously (11). Peptides with Mascot ion scores lower than 35 were discarded.

Data visualization, *in silico* prediction of peptide binding and peptide-binding profiles. WebLogo, version 3.6.0, was used for the visualization of the peptide-binding motif of Patr-A*05:01 (76). In a sequence logo, the height of the column of letters is equal to the information content at that position. The height of a letter in a column is proportional to the frequency of the corresponding amino acid at that position. The amino acid color code is according to the chemical property: green, polar; blue, basic; and black, hydrophobic amino acids.

The NetMHCpan server (77), freely available from the Internet, can predict binding of peptides to any MHC class I molecule of known sequence using artificial neural networks (http://www.cbs.dtu.dk/services/NetMHCpan/). To predict peptide binding to HLA-A and most Patr-A molecules, the NetMHCpan version 4.0 server was used (57). The chimpanzee allotypes Patr-A*03:07, -A*17:04, -A*21:01, -and A*27:01 and all Papa-A allotypes are not routinely included in the NetMHCpan 4.0 server; however, when the deduced protein sequence (FASTA format) was entered into the required field for these allotypes and peptide-binding predicting was performed, this server did not provide information on the BindLevel. Therefore, predictions for these allotypes were performed with the NetMHCpan 3.0 server (58). The NetMHCpan 3.0 server predicts the binding affinity (in nanomolars) also for the nonstandard allotypes included. This binding affinity (in nanomolars) is an indication of the strength or weakness of the binding of a peptide to the MHC molecule of interest (strong binding [SB], affinity on average of $<$100 nM; intermediate binding [weak, WB], affinity on average between 100 and 900 nM). Of note, regarding *in vivo* T cell recognition studies, 500 nm has been recognized as a peptide-binding threshold for immunogenicity (52). For both servers, 9-mers was chosen as the peptide length, a percent rank of 0.5 was used as the threshold for a strong binder, and a percent rank of 2 was used for a weak binder.

With the MHCcluster, version 2.0, server (45), which is freely available from the Internet, MHC molecules can be clustered functionally on the basis of their predicted binding specificities (http://www.cbs.dtu.dk/services/MHCcluster/). The server predicts for each selected MHC allotype the binding to a set of predefined natural peptides. The server provides a heat map, a graphical tree-based visualization that illustrates the functional relationship of the MHC molecules analyzed, a zip file containing the sequence logos of the predicted binding motifs of the MHC molecules, and a file with the estimated accuracies of the predicted sequence motifs (estimated from the distance to the nearest MHC molecules included in

the training of the peptide-binding prediction method) (45). The MHCcluster, version 2.0, server was used to cluster functionally the allotypes that are not routinely included in the NetMHCpan version 3.0 and 4.0 servers (Patr-A*03:07, -A*17:04, -A*21:01, and -A*27:01 and all Papa-A allotypes), together with particular reference allotypes of the six different HLA-A supertypes as references. The accuracy predictions for the Patr-A and Papa-A allotypes included in the analyses are provided (see Table S2 in the supplemental material). The generic tree, predicted by MHCcluster, including all Patr- and Papa-A allotypes as well as representatives of all HLA supertypes, gave a predicted topology and results for the logo plots similar to the representative tree depicted in Fig. 3.

SUPPLEMENTAL MATERIAL

Supplemental material is available online only.

SUPPLEMENTAL FILE 1, PDF file, 0.3 MB.

ACKNOWLEDGMENTS

We thank D. Devine for editing the manuscript and F. van Hassel for preparing the figures.

REFERENCES

1. Fujiyama A, Watanabe H, Toyoda A, Taylor TD, Itoh T, Tsai SF, Park HS, Yaspo ML, Lehrach H, Chen Z, Fu G, Saitou N, Osoegawa K, de Jong PJ, Suto Y, Hattori M, Sakaki Y. 2002. Construction and analysis of a human-chimpanzee comparative clone map. Science 295:131–134. https://doi.org/10.1126/science.1065199.
2. Prado-Martinez J, Sudmant PH, Kidd JM, Li H, Kelley JL, Lorente-Galdos B, Veeramah KR, Woerner AE, O'Connor TD, Santpere G, Cagan A, Theunert C, Casals F, Laayouni H, Munch K, Hobolth A, Halager AE, Malig M, Hernandez-Rodriguez J, Hernando-Herraez I, Prüfer K, Pybus M, Johnstone L, Lachmann M, Alkan C, Twigg D, Petit N, Baker C, Hormozdiari F, Fernandez-Callejo M, Dabad M, Wilson ML, Stevison L, Camprubí C, Carvalho T, Ruiz-Herrera A, Vives L, Mele M, Abello T, Kondova I, Bontrop RE, Pusey A, Lankester F, Kiyang JA, Bergl RA, Lonsdorf E, Myers S, Ventura M, Gagneux P, Comas D, Siegismund H, Blanc J, Agueda-Calpena L, Gut M, Fulton L, Tishkoff SA, Mullikin JC, Wilson RK, Gut IG, Gonder MK, Ryder OA, Hahn BH, Navarro A, Akey JM, Bertranpetit J, Reich D, Mailund T, Schierup MH, Hvilsom C, Andrés AM, Wall JD, Bustamante CD, Hammer MF, Eichler EE, Marques-Bonet T. 2013. Great ape genetic diversity and population history. Nature 499:471–475. https://doi.org/10.1038/nature12228.
3. Heeney JL, Dalgleish AG, Weiss RA. 2006. Origins of HIV and the evolution of resistance to AIDS. Science 313:462–466. https://doi.org/10.1126/science.1123016.
4. Novembre FJ, Saucier M, Anderson DC, Klumpp SA, O'Neil SP, Brown CR, Hart CE, Guenthner PC, Swenson RB, McClure HM. 1997. Development of AIDS in a chimpanzee infected with human immunodeficiency virus type 1. J Virol 71:4086–4091. https://doi.org/10.1128/JVI.71.5.4086-4091.1997.
5. Novembre FJ, de Rosayro J, Nidtha S, O'Neil SP, Gibson TR, Evans-Strickfaden T, Hart CE, McClure HM. 2001. Rapid $CD4^+$ T-cell loss induced by human immunodeficiency virus type 1_{NC} in uninfected and previously infected chimpanzees. J Virol 75:1533–1539. https://doi.org/10.1128/JVI.75.3.1533-1539.2001.
6. Fellay J, Shianna KV, Ge D, Colombo S, Ledergerber B, Weale M, Zhang K, Gumbs C, Castagna A, Cossarizza A, Cozzi-Lepri A, De Luca A, Easterbrook P, Francioli P, Mallal S, Martinez-Picado J, Miro JM, Obel N, Smith JP, Wyniger J, Descombes P, Antonarakis SE, Letvin NL, McMichael AJ, Haynes BF, Telenti A, Goldstein DB. 2007. A whole-genome association study of major determinants for host control of HIV-1. Science 317:944–947. https://doi.org/10.1126/science.1143767.
7. Goulder PJ, Walker BD. 2012. HIV and HLA class I: an evolving relationship. Immunity 37:426–440. https://doi.org/10.1016/j.immuni.2012.09.005.
8. Kiepiela P, Leslie AJ, Honeyborne I, Ramduth D, Thobakgale C, Chetty S, Rathnavalu P, Moore C, Pfafferott KJ, Hilton L, Zimbwa P, Moore S, Allen T, Brander C, Addo MM, Altfeld M, James I, Mallal S, Bunce M, Barber LD, Szinger J, Day C, Klenerman P, Mullins J, Korber B, Coovadia HM, Walker BD, Goulder PJ. 2004. Dominant influence of HLA-B in mediating the potential co-evolution of HIV and HLA. Nature 432:769–775. https://doi.org/10.1038/nature03113.
9. de Groot NG, Bontrop RE. 2013. The HIV-1 pandemic: does the selective sweep in chimpanzees mirror humankind's future? Retrovirology 10:53. https://doi.org/10.1186/1742-4690-10-53.
10. de Groot NG, Otting N, Doxiadis GG, Balla-Jhagjhoorsingh SS, Heeney JL, van Rood JJ, Gagneux P, Bontrop RE. 2002. Evidence for an ancient selective sweep in the MHC class I gene repertoire of chimpanzees. Proc Natl Acad Sci U S A 99:11748–11753. https://doi.org/10.1073/pnas.182420799.
11. de Groot NG, Heijmans CM, Zoet YM, de Ru AH, Verreck FA, van Veelen PA, Drijfhout JW, Doxiadis GG, Remarque EJ, Doxiadis II, van Rood JJ, Koning F, Bontrop RE. 2010. AIDS-protective HLA-B*27/B*57 and chimpanzee MHC class I molecules target analogous conserved areas of HIV-1/SIVcpz. Proc Natl Acad Sci U S A 107:15175–15180. https://doi.org/10.1073/pnas.1009136107.
12. Hahn BH, Shaw GM, De Cock KM, Sharp PM. 2000. AIDS as a zoonosis: scientific and public health implications. Science 287:607–614. https://doi.org/10.1126/science.287.5453.607.
13. Bowden R, MacFie TS, Myers S, Hellenthal G, Nerrienet E, Bontrop RE, Freeman C, Donnelly P, Mundy NI. 2012. Genomic tools for evolution and conservation in the chimpanzee: Pan troglodytes ellioti is a genetically distinct population. PLoS Genet 8:e1002504. https://doi.org/10.1371/journal.pgen.1002504.
14. Gagneux P, Wills C, Gerloff U, Tautz D, Morin PA, Boesch C, Fruth B, Hohmann G, Ryder OA, Woodruff DS. 1999. Mitochondrial sequences show diverse evolutionary histories of African hominoids. Proc Natl Acad Sci U S A 96:5077–5082. https://doi.org/10.1073/pnas.96.9.5077.
15. Kawamoto Y, Takemoto H, Higuchi S, Sakamaki T, Hart JA, Hart TB, Tokuyama N, Reinartz GE, Guislain P, Dupain J, Cobden AK, Mulavwa MN, Yangozene K, Darroze S, Devos C, Furuichi T. 2013. Genetic structure of wild bonobo populations: diversity of mitochondrial DNA and geographical distribution. PLoS One 8:e59660. https://doi.org/10.1371/journal.pone.0059660.
16. Zsurka G, Kudina T, Peeva V, Hallmann K, Elger CE, Khrapko K, Kunz WS. 2010. Distinct patterns of mitochondrial genome diversity in bonobos (Pan paniscus) and humans. BMC Evol Biol 10:270. https://doi.org/10.1186/1471-2148-10-270.
17. Boué V, Locatelli S, Boucher F, Ayouba A, Butel C, Esteban A, Okouga A-P, Ndoungouet A, Motsch P, Le Flohic G, Ngari P, Prugnolle F, Ollomo B, Rouet F, Liégeois F. 2015. High rate of simian immunodeficiency virus (SIV) infections in wild chimpanzees in northeastern Gabon. Viruses 7:4997–5015. https://doi.org/10.3390/v7092855.
18. Peeters M, Fransen K, Delaporte E, Van den Haesevelde M, Gershy-Damet GM, Kestens L, van der Groen G, Piot P. 1992. Isolation and characterization of a new chimpanzee lentivirus (simian immunodeficiency virus isolate cpz-ant) from a wild-captured chimpanzee. AIDS 6:447–451. https://doi.org/10.1097/00002030-199205000-00002.
19. Peeters M, Honore C, Huet T, Bedjabaga L, Ossari S, Bussi P, Cooper RW, Delaporte E. 1989. Isolation and partial characterization of an HIV-related virus occurring naturally in chimpanzees in Gabon. AIDS 3:625–630. https://doi.org/10.1097/00002030-198910000-00001.
20. Rudicell RS, Piel AK, Stewart F, Moore DL, Learn GH, Li Y, Takehisa J, Pintea L, Shaw GM, Moore J, Sharp PM, Hahn BH. 2011. High prevalence of simian immunodeficiency virus infection in a community of

savanna chimpanzees. J Virol 85:9918–9928. https://doi.org/10.1128/JVI.05475-11.
21. Leendertz SA, Locatelli S, Boesch C, Kucherer C, Formenty P, Liegeois F, Ayouba A, Peeters M, Leendertz FH. 2011. No evidence for transmission of SIVwrc from western red colobus monkeys (Piliocolobus badius badius) to wild West African chimpanzees (Pan troglodytes verus) despite high exposure through hunting. BMC Microbiol 11:24. https://doi.org/10.1186/1471-2180-11-24.
22. Li Y, Ndjango JB, Learn GH, Ramirez MA, Keele BF, Bibollet-Ruche F, Liu W, Easlick JL, Decker JM, Rudicell RS, Inogwabini BI, Ahuka-Mundeke S, Leendertz FH, Reynolds V, Muller MN, Chancellor RL, Rundus AS, Simmons N, Worobey M, Shaw GM, Peeters M, Sharp PM, Hahn BH. 2012. Eastern chimpanzees, but not bonobos, represent a simian immunodeficiency virus reservoir. J Virol 86:10776–10791. https://doi.org/10.1128/JVI.01498-12.
23. Locatelli S, Harrigan RJ, Sesink Clee PR, Mitchell MW, McKean KA, Smith TB, Gonder MK. 2016. Why are Nigeria-Cameroon chimpanzees (Pan troglodytes ellioti) free of SIVcpz infection? PLoS One 11:e0160788. https://doi.org/10.1371/journal.pone.0160788.
24. Van Heuverswyn F, Li Y, Bailes E, Neel C, Lafay B, Keele BF, Shaw KS, Takehisa J, Kraus MH, Loul S, Butel C, Liegeois F, Yangda B, Sharp PM, Mpoudi-Ngole E, Delaporte E, Hahn BH, Peeters M. 2007. Genetic diversity and phylogeographic clustering of SIVcpzPtt in wild chimpanzees in Cameroon. Virology 368:155–171. https://doi.org/10.1016/j.virol.2007.06.018.
25. Etienne L, Nerrienet E, LeBreton M, Bibila GT, Foupouapouognigni Y, Rousset D, Nana A, Djoko CF, Tamoufe U, Aghokeng AF, Mpoudi-Ngole E, Delaporte E, Peeters M, Wolfe ND, Ayouba A. 2011. Characterization of a new simian immunodeficiency virus strain in a naturally infected Pan troglodytes chimpanzee with AIDS related symptoms. Retrovirology 8:4. https://doi.org/10.1186/1742-4690-8-4.
26. Keele BF, Jones JH, Terio KA, Estes JD, Rudicell RS, Wilson ML, Li Y, Learn GH, Beasley TM, Schumacher-Stankey J, Wroblewski E, Mosser A, Raphael J, Kamenya S, Lonsdorf EV, Travis DA, Mlengeya T, Kinsel MJ, Else JG, Silvestri G, Goodall J, Sharp PM, Shaw GM, Pusey AE, Hahn BH. 2009. Increased mortality and AIDS-like immunopathology in wild chimpanzees infected with SIVcpz. Nature 460:515–519. https://doi.org/10.1038/nature08200.
27. Rudicell RS, Holland Jones J, Wroblewski EE, Learn GH, Li Y, Robertson JD, Greengrass E, Grossmann F, Kamenya S, Pintea L, Mjungu DC, Lonsdorf EV, Mosser A, Lehman C, Collins DA, Keele BF, Goodall J, Hahn BH, Pusey AE, Wilson ML. 2010. Impact of simian immunodeficiency virus infection on chimpanzee population dynamics. PLoS Pathog 6:e1001116. https://doi.org/10.1371/journal.ppat.1001116.
28. Greenwood EJ, Schmidt F, Kondova I, Niphuis H, Hodara VL, Clissold L, McLay K, Guerra B, Redrobe S, Giavedoni LD, Lanford RE, Murthy KK, Rouet F, Heeney JL. 2015. Simian immunodeficiency virus infection of chimpanzees (Pan troglodytes) shares features of both pathogenic and non-pathogenic lentiviral infections. PLoS Pathog 11:e1005146. https://doi.org/10.1371/journal.ppat.1005146.
29. Barbian HJ, Jackson-Jewett R, Brown CS, Bibollet-Ruche F, Learn GH, Decker T, Kreider EF, Li Y, Denny TN, Sharp PM, Shaw GM, Lifson J, Acosta EP, Saag MS, Bar KJ, Hahn BH. 2017. Effective treatment of SIVcpz-induced immunodeficiency in a captive western chimpanzee. Retrovirology 14:35. https://doi.org/10.1186/s12977-017-0359-0.
30. Sidney J, Peters B, Frahm N, Brander C, Sette A. 2008. HLA class I supertypes: a revised and updated classification. BMC Immunol 9:1. https://doi.org/10.1186/1471-2172-9-1.
31. Adams EJ, Cooper S, Thomson G, Parham P. 2000. Common chimpanzees have greater diversity than humans at two of the three highly polymorphic MHC class I genes. Immunogenetics 51:410–424. https://doi.org/10.1007/s002510050639.
32. de Groot NG, Heijmans CMC, Helsen P, Otting N, Pereboom Z, Stevens JMG, Bontrop RE. 2017. Limited MHC class I intron 2 repertoire variation in bonobos. Immunogenetics 69:677–688. https://doi.org/10.1007/s00251-017-1010-x.
33. de Groot NG, Otting N, Arguello R, Watkins DI, Doxiadis GG, Madrigal JA, Bontrop RE. 2000. Major histocompatibility complex class I diversity in a West African chimpanzee population: implications for HIV research. Immunogenetics 51:398–409. https://doi.org/10.1007/s002510050638.
34. Maibach V, Hans JB, Hvilsom C, Marques-Bonet T, Vigilant L. 2017. MHC class I diversity in chimpanzees and bonobos. Immunogenetics 69:661–676. https://doi.org/10.1007/s00251-017-0990-x.
35. Wroblewski EE, Guethlein LA, Norman PJ, Li Y, Shaw CM, Han AS, Ndjango JN, Ahuka-Mundeke S, Georgiev AV, Peeters M, Hahn BH, Parham P. 2017. Bonobos maintain immune system diversity with three functional types of MHC-B. J Immunol 198:3480–3493. https://doi.org/10.4049/jimmunol.1601955.
36. Wroblewski EE, Norman PJ, Guethlein LA, Rudicell RS, Ramirez MA, Li Y, Hahn BH, Pusey AE, Parham P. 2015. Signature patterns of MHC diversity in three Gombe communities of wild chimpanzees reflect fitness in reproduction and immune defense against SIVcpz. PLoS Biol 13:e1002144. https://doi.org/10.1371/journal.pbio.1002144.
37. de Groot NG, Otting N, Maccari G, Robinson J, Hammond JA, Blancher A, Lafont BAP, Guethlein LA, Wroblewski EE, Marsh SGE, Shiina T, Walter L, Vigilant L, Parham P, O'Connor DH, Bontrop RE. 2019. Nomenclature report 2019: major histocompatibility complex genes and alleles of Great and Small Ape and Old and New World monkey species. Immunogenetics 72:25–36. https://doi.org/10.1007/s00251-019-01132-x.
38. Maccari G, Robinson J, Ballingall K, Guethlein LA, Grimholt U, Kaufman J, Ho CS, de Groot NG, Flicek P, Bontrop RE, Hammond JA, Marsh SG. 2017. IPD-MHC 2.0: an improved inter-species database for the study of the major histocompatibility complex. Nucleic Acids Res 45:D860–D864. https://doi.org/10.1093/nar/gkw1050.
39. Robinson J, Halliwell JA, Hayhurst JD, Flicek P, Parham P, Marsh SG. 2015. The IPD and IMGT/HLA database: allele variant databases. Nucleic Acids Res 43:D423–D431. https://doi.org/10.1093/nar/gku1161.
40. de Groot NG, Otting N, Robinson J, Blancher A, Lafont BAP, Marsh SGE, O'Connor DH, Shiina T, Walter L, Watkins DI, Bontrop RE. 2012. Nomenclature report on the major histocompatibility complex genes and alleles of Great Ape, Old and New World monkey species. Immunogenetics 64:615–631. https://doi.org/10.1007/s00251-012-0617-1.
41. Sidney J, Asabe S, Peters B, Purton KA, Chung J, Pencille TJ, Purcell R, Walker CM, Chisari FV, Sette A. 2006. Detailed characterization of the peptide binding specificity of five common Patr class I MHC molecules. Immunogenetics 58:559–570. https://doi.org/10.1007/s00251-006-0131-4.
42. Sidney J, Peters B, Moore C, Pencille TJ, Ngo S, Masterman KA, Asabe S, Pinilla C, Chisari FV, Sette A. 2007. Characterization of the peptide-binding specificity of the chimpanzee class I alleles A 0301 and A 0401 using a combinatorial peptide library. Immunogenetics 59:745–751. https://doi.org/10.1007/s00251-007-0243-5.
43. de Groot NG, Garcia CA, Verschoor EJ, Doxiadis GG, Marsh SG, Otting N, Bontrop RE. 2005. Reduced MIC gene repertoire variation in West African chimpanzees as compared to humans. Mol Biol Evol 22:1375–1385. https://doi.org/10.1093/molbev/msi127.
44. Hassan C, Kester MG, de Ru AH, Hombrink P, Drijfhout JW, Nijveen H, Leunissen JA, Heemskerk MH, Falkenburg JH, van Veelen PA. 2013. The human leukocyte antigen-presented ligandome of B lymphocytes. Mol Cell Proteomics 12:1829–1843. https://doi.org/10.1074/mcp.M112.024810.
45. Thomsen M, Lundegaard C, Buus S, Lund O, Nielsen M. 2013. MHCcluster, a method for functional clustering of MHC molecules. Immunogenetics 65:655–665. https://doi.org/10.1007/s00251-013-0714-9.
46. de Groot NG, Stevens JMG, Bontrop RE. 2018. Does the MHC confer protection against malaria in Bonobos? Trends Immunol 39:768–771. https://doi.org/10.1016/j.it.2018.07.004.
47. Loffredo JT, Sidney J, Bean AT, Beal DR, Bardet W, Wahl A, Hawkins OE, Piaskowski S, Wilson NA, Hildebrand WH, Watkins DI, Sette A. 2009. Two MHC class I molecules associated with elite control of immunodeficiency virus replication, Mamu-B*08 and HLA-B*2705, bind peptides with sequence similarity. J Immunol 182:7763–7775. https://doi.org/10.4049/jimmunol.0900111.
48. Mothe BR, Sidney J, Dzuris JL, Liebl ME, Fuenger S, Watkins DI, Sette A. 2002. Characterization of the peptide-binding specificity of Mamu-B*17 and identification of Mamu-B*17-restricted epitopes derived from simian immunodeficiency virus proteins. J Immunol 169:210–219. https://doi.org/10.4049/jimmunol.169.1.210.
49. de Groot NG, Heijmans CMC, de Ru AH, Janssen GMC, Drijfhout JW, Otting N, Vangenot C, Doxiadis GGM, Koning F, van Veelen PA, Bontrop RE. 2017. A specialist macaque MHC class I molecule with HLA-B*27-like peptide-binding characteristics. J Immunol 199:3679–3690. https://doi.org/10.4049/jimmunol.1700502.
50. de Groot NG, Heijmans CM, de Ru AH, Hassan C, Otting N, Doxiadis GG, Koning F, van Veelen PA, Bontrop RE. 2013. Unique peptide-binding motif for Mamu-B*037:01: an MHC class I allele common to Indian and Chinese rhesus macaques. Immunogenetics 65:897–900. https://doi.org/10.1007/s00251-013-0734-5.

51. Mothe BR, Southwood S, Sidney J, English AM, Wriston A, Hoof I, Shabanowitz J, Hunt DF, Sette A. 2013. Peptide-binding motifs associated with MHC molecules common in Chinese rhesus macaques are analogous to those of human HLA supertypes and include HLA-B27-like alleles. Immunogenetics 65:371–386. https://doi.org/10.1007/s00251-013-0686-9.
52. Southwood S, Solomon C, Hoof I, Rudersdorf R, Sidney J, Peters B, Wahl A, Hawkins O, Hildebrand W, Mothe BR, Sette A. 2011. Functional analysis of frequently expressed Chinese rhesus macaque MHC class I molecules Mamu-A1*02601 and Mamu-B*08301 reveals HLA-A2 and HLA-A3 supertypic specificities. Immunogenetics 63:275–290. https://doi.org/10.1007/s00251-010-0502-8.
53. Borghans JAM, Mølgaard A, de Boer RJ, Keşmir C. 2007. HLA alleles associated with slow progression to AIDS truly prefer to present HIV-1 p24. PLoS One 2:e920. https://doi.org/10.1371/journal.pone.0000920.
54. Goulder PJ, Watkins DI. 2008. Impact of MHC class I diversity on immune control of immunodeficiency virus replication. Nat Rev Immunol 8:619–630. https://doi.org/10.1038/nri2357.
55. Chakraborty S, Rahman T, Chakravorty R, Kuchta A, Rabby A, Sahiuzzaman M. 2013. HLA supertypes contribute in HIV type 1 cytotoxic T lymphocyte epitope clustering in Nef and Gag proteins. AIDS Res Hum Retroviruses 29:270–278. https://doi.org/10.1089/AID.2012.0160.
56. Stephens HA. 2005. HIV-1 diversity versus HLA class I polymorphism. Trends Immunol 26:41–47. https://doi.org/10.1016/j.it.2004.11.001.
57. Jurtz V, Paul S, Andreatta M, Marcatili P, Peters B, Nielsen M. 2017. NetMHCpan-4.0: improved peptide-MHC class I interaction predictions integrating eluted ligand and peptide binding affinity data. J Immunol 199:3360–3368. https://doi.org/10.4049/jimmunol.1700893.
58. Nielsen M, Andreatta M. 2016. NetMHCpan-3.0; improved prediction of binding to MHC class I molecules integrating information from multiple receptor and peptide length datasets. Genome Med 8:33. https://doi.org/10.1186/s13073-016-0288-x.
59. de Groot NG, Heijmans CMC, Bontrop RE. 2017. AIDS in chimpanzees: the role of MHC genes. Immunogenetics 69:499–509. https://doi.org/10.1007/s00251-017-1006-6.
60. Domena JD, Hildebrand WH, Bias WB, Parham P. 1993. A sixth family of HLA-A alleles defined by HLA-A*8001. Tissue Antigens 42:156–159. https://doi.org/10.1111/j.1399-0039.1993.tb02186.x.
61. Kato K, Trapani JA, Allopenna J, Dupont B, Sy Y. 1989. Molecular analysis of the serologically defined HLA-Aw19 antigens. A genetically distinct family of HLA-A antigens comprising A29, A31, A32, and Aw33, but probably not A30. J Immunol 143:3371–3378.
62. McAdam SN, Boyson JE, Liu X, Garber TL, Hughes AL, Bontrop RE, Watkins DI. 1995. Chimpanzee MHC class I A locus alleles are related to only one of the six families of human A locus alleles. J Immunol 154:6421–6429.
63. Koehler RN, Walsh AM, Saathoff E, Tovanabutra S, Arroyo MA, Currier JR, Maboko L, Hoelscher M, Hoelsher M, Robb ML, Michael NL, McCutchan FE, Kim JH, Kijak GH. 2010. Class I HLA-A*7401 is associated with protection from HIV-1 acquisition and disease progression in Mbeya, Tanzania. J Infect Dis 202:1562–1566. https://doi.org/10.1086/656913.
64. Selvaraj P, Swaminathan S, Alagarasu K, Raghavan S, Narendran G, Narayanan P. 2006. Association of human leukocyte antigen-A11 with resistance and B40 and DR2 with susceptibility to HIV-1 infection in south India. J Acquir Immune Defic Syndr 43:497–499. https://doi.org/10.1097/01.qai.0000233312.36226.76.
65. Valenzuela-Ponce H, Alva-Hernandez S, Garrido-Rodriguez D, Soto-Nava M, Garcia-Tellez T, Escamilla-Gomez T, Garcia-Morales C, Quiroz-Morales VS, Tapia-Trejo D, Del Arenal-Sanchez S, Prado-Galbarro FJ, Hernandez-Juan R, Rodriguez-Aguirre E, Murakami-Ogasawara A, Mejia-Villatoro C, Escobar-Urias IY, Pinzon-Meza R, Pascale JM, Zaldivar Y, Porras-Cortes G, Quant-Duran C, Lorenzana I, Meza RI, Palou EY, Manzanero M, Cedillos RA, Alaez C, Brockman MA, Harrigan PR, Brumme CJ, Brumme ZL, Avila-Rios S, Reyes-Teran G, Mesoamerican HIV Project Group. 2018. Novel HLA class I associations with HIV-1 control in a unique genetically admixed population. Sci Rep 8:6111. https://doi.org/10.1038/s41598-018-23849-7.
66. Yaciuk JC, Skaley M, Bardet W, Schafer F, Mojsilovic D, Cate S, Stewart CJ, McMurtrey C, Jackson KW, Buchli R, Olvera A, Cedeno S, Plana M, Mothe B, Brander C, West JT, Hildebrand WH. 2014. Direct interrogation of viral peptides presented by the class I HLA of HIV-infected T cells. J Virol 88:12992–13004. https://doi.org/10.1128/JVI.01914-14.
67. Dostalkova A, Kaufman F, Krizova I, Kultova A, Strohalmova K, Hadravova R, Ruml T, Rumlova M. 2018. Mutations in the basic region of the Mason-Pfizer monkey virus nucleocapsid protein affect reverse transcription, genomic RNA packaging, and the virus assembly site. J Virol 92:e00106-18. https://doi.org/10.1128/JVI.00106-18.
68. Hearps AC, Wagstaff KM, Piller SC, Jans DA. 2008. The N-terminal basic domain of the HIV-1 matrix protein does not contain a conventional nuclear localization sequence but is required for DNA binding and protein self-association. Biochemistry 47:2199–2210. https://doi.org/10.1021/bi701360j.
69. Lingappa JR, Dooher JE, Newman MA, Kiser PK, Klein KC. 2006. Basic residues in the nucleocapsid domain of Gag are required for interaction of HIV-1 gag with ABCE1 (HP68), a cellular protein important for HIV-1 capsid assembly. J Biol Chem 281:3773–3784. https://doi.org/10.1074/jbc.M507255200.
70. Yuan X, Yu X, Lee TH, Essex M. 1993. Mutations in the N-terminal region of human immunodeficiency virus type 1 matrix protein block intracellular transport of the Gag precursor. J Virol 67:6387–6394. https://doi.org/10.1128/JVI.67.11.6387-6394.1993.
71. Elahi S, Dinges WL, Lejarcegui N, Laing KJ, Collier AC, Koelle DM, McElrath MJ, Horton H. 2011. Protective HIV-specific $CD8^+$ T cells evade Treg cell suppression. Nat Med 17:989–995. https://doi.org/10.1038/nm.2422.
72. Lazaryan A, Song W, Lobashevsky E, Tang J, Shrestha S, Zhang K, Gardner LI, McNicholl JM, Wilson CM, Klein RS, Rompalo A, Mayer K, Sobel J, Kaslow RA, HIV Epidemiology Research Study, Reaching for Excellence in Adolescent Care and Health Study. 2010. Human leukocyte antigen class I supertypes and HIV-1 control in African Americans. J Virol 84:2610–2617. https://doi.org/10.1128/JVI.01962-09.
73. Sodora DL, Allan JS, Apetrei C, Brenchley JM, Douek DC, Else JG, Estes JD, Hahn BH, Hirsch VM, Kaur A, Kirchhoff F, Muller-Trutwin M, Pandrea I, Schmitz JE, Silvestri G. 2009. Toward an AIDS vaccine: lessons from natural simian immunodeficiency virus infections of African nonhuman primate hosts. Nat Med 15:861–865. https://doi.org/10.1038/nm.2013.
74. de Groot NG, Heijmans CM, de Groot N, Doxiadis GG, Otting N, Bontrop RE. 2009. The chimpanzee Mhc-DRB region revisited: gene content, polymorphism, pseudogenes, and transcripts. Mol Immunol 47:381–389. https://doi.org/10.1016/j.molimm.2009.09.003.
75. Zoet YM, Eijsink C, Kardol MJ, Franke-van Dijk ME, Wilson GL, de Paus R, Mickelson E, Heemskerk M, van den Elsen PJ, Claas FH, Mulder A, Doxiadis II. 2005. The single antigen expressing lines (SALs) concept: an excellent tool for screening for HLA-specific antibodies. Hum Immunol 66:519–525. https://doi.org/10.1016/j.humimm.2005.01.007.
76. Crooks GE, Hon G, Chandonia JM, Brenner SE. 2004. WebLogo: a sequence logo generator. Genome Res 14:1188–1190. https://doi.org/10.1101/gr.849004.
77. Hoof I, Peters B, Sidney J, Pedersen LE, Sette A, Lund O, Buus S, Nielsen M. 2009. NetMHCpan, a method for MHC class I binding prediction beyond humans. Immunogenetics 61:1–13. https://doi.org/10.1007/s00251-008-0341-z.