The HIV-1 pandemic: does the selective sweep in chimpanzees mirror humankind’s future?

Natasja G de Groot1* and Ronald E Bontrop1,2

Abstract

An HIV-1 infection progresses in most human individuals sooner or later into AIDS, a devastating disease that kills more than a million people worldwide on an annual basis. Nonetheless, certain HIV-1-infected persons appear to act as long-term non-progressors, and elite control is associated with the presence of particular MHC class I allotypes such as HLA-B*27 or -B*57. The HIV-1 pandemic in humans arose from the cross-species transmission of SIVcpz originating from chimpanzees. Chimpanzees, however, appear to be relatively resistant to developing AIDS after HIV-1/SIVcpz infection. Mounting evidence illustrates that, in the distant past, chimpanzees experienced a selective sweep resulting in a severe reduction of their MHC class I repertoire. This was most likely caused by an HIV-1/SIV-like retrovirus, suggesting that chimpanzees may have experienced long-lasting host-virus relationships with SIV-like viruses. Hence, if natural selection is allowed to follow its course, prospects for the human population may look grim, thus underscoring the desperate need for an effective vaccine.

Keyword: AIDS, Chimpanzee, HIV-1, HLA, Human, MHC, Repertoire reduction, SIVcpz, Zoonosis

Review

AIDS: facts and figures

Acquired immunodeficiency syndrome (AIDS) is caused by the human immunodeficiency virus type 1 (HIV-1) and poses a major threat to human health. At present, over thirty million individuals are infected worldwide, and more than twenty five million people have died from AIDS since the beginning of the pandemic (UNAIDS 2011). Based on genetic characteristics, the virus strains can be clustered into different groups designated M, N, O and P. HIV-1 group M is responsible for the vast majority of infections recorded in human individuals globally [1,2]. Thanks to the currently available anti-retroviral (AR) drugs, the number of people dying yearly from AIDS reached a plateau around 2004, and the number has even begun to decline (UNAIDS 2011). However, it is important to mention that highly active anti-retroviral therapy (HAART) does not eradicate HIV-1 [3], and several studies have demonstrated that discontinuation of HAART may result in a rapid viral rebound [4-7]. Moreover, AR therapy carries with it high costs, resulting in only a limited number of HIV-1-infected individuals having access to it. Especially in developing countries, where most of the infected individuals live, AIDS is still a prominent health problem. Another worry is that sooner or later the virus may develop resistance to these drugs [8], and the spread of drug-resistant HIV-1 strains would be devastating. Therefore, a series of low-cost drugs and, more optimistically, an efficacious vaccine able to protect against infection are urgently needed.

HIV-1 in a nutshell: its genome and infection routes

The genome of HIV-1 is relatively small, approximately 10 kilobases, and encodes a limited number of gene products (Figure 1). This retrovirus belongs to the group of lentiviruses. Like most retroviruses, the HIV-1 genome contains the three functional proteins Gag, Pol, and Env, which are essential for the construction of the virus particle [9]. In addition, lentiviral genomes encode a number of regulatory (Tat and Rev) and accessory genes (Vif, Vpr, Vpu, and Nef). The Tat and Rev proteins are essential for replication of the virus. The accessory proteins are attached to viral messenger RNA, but they are not required for replication of the virus in vitro. However, in vivo these proteins are necessary for the replication and virulence of the virus [10].
HIV-1 possesses specific characteristics that hamper or have hampered the development of broadly efficacious and efficient anti-retroviral drugs and vaccines. On the one hand, the virus can cause a persistent infection and hide in an inactive form in cellular reservoirs. On the other hand, it can replicate quickly, and – depending on the copy number present in one individual – may produce $10^9$ to $10^{10}$ particles per day [11,12]. Moreover, the genome of HIV-1 encodes an error-prone reverse transcriptase characterized by a high mutation rate (approximately $3 \times 10^{-5}$ per nucleotide/replication cycle) [13]. This, in combination with a relatively small genome size and the high viral production rate within a given individual, may ultimately result in the establishment of a virus swarm. There is also the risk of emergence of new recombinant strains as multiple HIV-1 viruses, which are prone to exchange genetic material, infect a single cell. Hence, a high mutation rate in concert with recombination provides the biological ability for HIV-1 to escape immune recognition/control of its host.

HIV-1 enters the human body via mucosal sites, blood-contact, or breast milk [14]. After entry, the virus either directly infects CD4+ T cells, is recruited by dendritic cells (DC), or is taken up by macrophages. The latter two cell types act as a kind of Trojan horses, and transport HIV-1 through different layers of epithelial cells to the lymphoid tissues, where they can transfer the virus to CD4+ T cells [15,16]. DCs can either internalize HIV-1 into endosomes, and the cell-surface molecule that is used by the virus for entry into the cell is the C-type lectin receptor DC-SIGN [17], or they can recruit HIV-1 virions to their endosomal compartment, where the virions accumulate at regions that come into contact with neighboring T cells (referred to as the infectious synapse) [18]. Macrophages may phagocytize HIV-1, and can either function as a reservoir for long-term persistence of the virus or can be responsible for the transport and dissemination of the virus. The “primary” non-syncytium-inducing (NSI) HIV-1 viruses mainly infect CD4+ memory T cells that express CC-chemokine receptor-5 (CCR5). Later in the course of the infection, the NSI strains can switch to a syncytium-inducing (SI) phenotype, which can infect T cells by using the C-X-C-chemokine receptor-4 (CXCR4) as co-receptor. This switch is associated with a loss of sensitivity to chemokines (like RANTES, and MIP-1α and β), a rapid decrease in CD4+ T cells, and progression towards AIDS [19]. Moreover, an infection of memory CD4+ T cells that revert back to a resting state ensues as well, resulting in a dormant stage of the virus that is undetectable by the host’s immune system [20]. In essence, HIV-1 infection and its subsequent manipulations take place at the heart of the immune system.

Zoonotic infections: what may happen if we open Pandora’s box

More than 40 African non-human primate species are infected naturally with various simian immunodeficiency virus (SIV) strains, but progression towards AIDS is rarely observed [21,22]. From among these natural primate hosts, the SIV infection has been most thoroughly studied in three species, sooty mangabeys, African green monkeys, and mandrills, and has led to the insight that non-pathogenicity need not be linked to effective immune control. Apparently, many of the species have co-evolved with SIV infections, and have found ways to manage/control it. These non-human primates, however, are potential reservoirs for viruses that could be spread among related species. One example of such a zoonotic transmission is provided by sooty mangabeys (Cercocebus atys). This species, infected naturally with a SIV strain (SIVsm), is considered to represent the origin of the HIV-2 infection in humans, causing an AIDS-like disease [23]. The infection in humans is mostly confined to the West African territory, and indeed overlaps with the natural habitat of sooty mangabeys.

Chimpanzees (Pan troglodytes), humankind’s closest living relative, can also be infected with SIV [24]. Based on the geographic distribution as well as on morphological and genetic data, chimpanzees have been divided into four different subspecies/populations sharing a common ancestor approximately 1.5 million years ago [25-27], and incidences of natural SIVcpz infection have been recorded in contemporary animals from at least two of the four subspecies (Figure 2). These SIVcpz-strains appear to have a mosaic genome, consisting of gene segments from different types of SIV strains obtained from distinct Old World monkey species [28]. This suggests that chimpanzees have acquired their types of SIV infections by predation on Old World monkeys sharing the same habitat. In turn, the different HIV-1 groups M, N, O and P have arisen from at least four separate introductions of SIV into the human population. The initiator of the human pandemic (HIV-1 group M) is the SIVcpz strain derived from chimpanzees of the Central-African subspecies (Pan troglodytes troglodytes) [29] (Figure 2), and the HIV-1 group N also arose from an SIVcpz strain from this species. The HIV-1 groups
Figure 2 Map of the African continent, highlighting habitats of the four different chimpanzee subspecies. Populations that show evidence of contemporary natural infections with SIVcpz strains are Pan troglodytes troglodytes (light orange) and P.t. schweinfurthii (dark orange). Superimposed is a diagram illustrating the cross-species transmission events that led to the emergence of HIV-1 group M, the initiator of the human pandemic [28]. The lower panel in the figure illustrates the speciation events in the homo-pan lineage. The arrow indicates the putative time span of the MHC repertoire reduction in the pan lineage.
O and P most likely originated from gorillas (SIVgor) [30,31], in which there is an indication that P.t.t. animals are the source of SIVgor [32].

**Chimpanzees and AIDS**

Apart from humans, chimpanzees are one of the non-human primate species that is susceptible to infection with HIV-1. Recently, it has been shown that pig-tailed macaques can also be infected through a modified HIV-1 strain that only differs in the vif gene [33,34]. In the past, a substantial number of West-African chimpanzees (P. t. verus) were experimentally infected with different HIV-1 strains, but most of these animals seem to be relatively resistant to developing AIDS [35,36]. These experiments were conducted before it was known that HIV-1 was derived from a zoonotic transmission of a virus that has its roots in a chimpanzee reservoir. In the late 1980s and early 1990s, the first chimpanzees to be infected naturally with an HIV-1-related lentivirus were discovered [37-39]. In the past, substantial number of West-African chimpanzees (P. t. verus) were experimentally infected with different HIV-1 strains, but most of these animals seem to be relatively resistant to developing AIDS [35,36]. These experiments were conducted before it was known that HIV-1 was derived from a zoonotic transmission of a virus that has its roots in a chimpanzee reservoir. In the late 1980s and early 1990s, the first chimpanzees to be infected naturally with an HIV-1-related lentivirus were discovered [37-39]. Additional natural SIVcpz infections in captive and wild-living chimpanzees were described later [40,41]. Some of these chimpanzees were kept under close surveillance for long periods, but signs of an AIDS-like disease were not seen. Moreover, transmission of one of the viruses to West-African chimpanzees (animals in which natural infection with SIVcpz is not observed) also did not result in the progression to AIDS [42]. Recently, population analyses have shown that wild-ranging East-African chimpanzees (P. t. schweinfurthii) experience natural infections with SIVcpz. For one of the populations, evidence of AIDS-like symptoms has been documented [43], and AIDS-related symptoms have also been described for a naturally infected Central-African chimpanzee [44]. In addition to these observations of AIDS in SIVcpz-related infections, there is one case documenting that chimpanzees can develop AIDS after experimental infection with HIV-1 [45]. This chimpanzee of West-African origin was infected with three different HIV-1 isolates. The animal eventually developed AIDS, and subsequent analysis showed that a recombinant HIV-1 strain had emerged. Three other chimpanzees that were afterwards infected with this recombinant strain all developed signs of AIDS-like disease [46]. This set of experiments shows that viruses that have been edited in the human population, and that are re-introduced into their original host species, can cause AIDS. Thus, escape from immune control in chimpanzees is possible in exceptional cases, even within a species considered to be relatively resistant.

**Control of AIDS development in humans: an introduction to the MHC**

In HIV-1-infected human cohorts, certain individuals do not progress towards AIDS. Some of these long-term non-progressors (LTNP)/elite controllers (ECs) have been infected for more than two decades [47], and their situation appears to resemble that of HIV-1-infected chimpanzees in terms of their relative resistance to developing AIDS. The LTNP/EC status in humans is strongly associated with the presence of particular major histocompatibility complex (MHC) class I molecules, such as HLA-B*27:05 and -B*57:01 [48-54].

The MHC plays a central role in the induction of adaptive immune responses, and encodes two clusters of cell-surface proteins. In humans, these molecules are designated HLA-A, -B, and -C (class I cluster), and -DP, -DQ, and -DR (class II cluster), and almost all these genes display abundant levels of polymorphism [55]. MHC class I antigens are expressed on virtually all nucleated cells, and generally present peptides of 8-10 amino acids in length. One of the biological functions of MHC class I molecules is to sample degraded peptide fragments from pathogens that are able to establish intracellular infections. In most cases, such peptides are derived from viruses trying to manipulate the intracellular machinery of the host cell in favor of their own replication purposes. MHC class II molecules, on the other hand, display differential tissue distribution and usually bind substantially longer peptides. They sample degraded peptide fragments from pathogens that are establishing extra-cellular infections, or that have life stages that are taking place outside of the host cell. In the case of an intracellular infection, the immune system will try to kill the target cell by inducing apoptosis-mediated lysis by CD8+ cytotoxic T cells (CTL). Extra-cellular infections are usually eliminated by antibodies that are generated by B cells under the control of CD4+ T helper cells.

In order to initiate an immune response, each individual will present a range of antigenic peptides as dictated by their unique HLA repertoire. Thus, MHC polymorphism warrants that many individuals within a given population have the capacity to generate unique adaptive immune responses to eliminate infections. In other words, MHC polymorphism minimizes the chance that a given pathogen will eliminate an entire population. Several pathogens have evolved different ways to interfere with the MHC class I and II antigen presentation pathways in order to avoid immune recognition [56-59]. Indeed, it has been suggested that in human populations exposed to HIV-1 there is enrichment for particular HLA specificities [60,61].

**The MHC and resistance to the development of AIDS in chimpanzees**

Grosso modo, humans and chimpanzees share the same MHC class I and II loci. However, chimpanzees possess an additional oligomorphic locus designated Patr-AL (A-like), which is characterized by a differential haplotype distribution and low expression levels [62,63]. An overview of the reported number of alleles for each locus is provided (Table 1). With regard to lineages, in comparison to humans, chimpanzees have a reduced MHC class I
Table 1 Number of alleles detected for Mhc class I and II loci in humans (HLA) and chimpanzees (Patr)

| Locus | HLA | Patr |
|-------|-----|------|
| Mhc class I |     |      |
| A    | 2188 | 33   |
| B    | 2862 | 57   |
| C    | 1746 | 31   |
| E    | 11   | 2    |
| F    | 22   | 1    |
| G    | 50   | 1    |
| AL   | 0    | 5    |
| DRA1 | 36   | 3    |
| DPB1 | 159  | 29   |
| DQA1 | 49   | 7    |
| DQB1 | 193  | 10   |
| DRA  | 7    | 2    |
| DRB  | 1386 | 79   |

#Individuals analyzed: **100**

Data were retrieved from the databases www.ebi.ac.uk/imgt/hla [70] and www.ebi.ac.uk/ipd/mhc/nhp [71].

repertoire. The very first indications that chimpanzees may lack the evolutionary equivalents of particular HLA lineages became evident in serological studies [64]. In humans, the HLA-A locus alleles are divided into two lineages comprising six families (Table 2) [65,66], and the experiments of van Rood et al. [64] showed that chimpanzees lack the HLA-A2 related serotypes. Subsequent molecular studies suggested that chimpanzees might only possess the equivalents of the HLA-A1/A3/A11/A30 family [67]. Although the number of animals tested in this initial study was relatively small, the outcome was unexpected. Humans and chimpanzees are each other’s closest living relatives, and they shared a common ancestor about 5 million years ago [68]. Thus, the most logical explanation would be that humans and chimpanzees share for the A and B loci the same lineages, since they were inherited in a trans-species mode of evolution [69].

As a follow-up to McAdam’s study [67], the MHC class I repertoire of a pedigreed West-African chimpanzee colony comprising more than 30 wild-caught founder animals was analyzed. Again, only orthologs of the HLA-A1/A3/A11/A30 family were encountered [72]. Moreover, MHC analyses of West-African chimpanzees by other research teams, and of small numbers of Central- and East-African chimpanzees, also provided data indicating only the presence of HLA-A1/A3/A11/A30 family orthologs [72-76]. Although the results suggest that there was a repertoire reduction at the MHC class I region in chimpanzees, conclusions in this direction should be arrived at cautiously. The presence of one particular omnipresent HLA-A-like family in chimpanzees can, in theory, also be explained by convergent evolution. Additionally, the imbalance in sample size between the number of chimpanzees and humans analyzed for their MHC repertoire can hamper an accurate interpretation of the data. For instance, one can argue that particular alleles and/or lineages were missed in chimpanzees due to a limited sample size. In addition, the MHC class I and II gene products of different species/populations may experience different sorts of natural selection, depending on their habitat and on the pathogens that are present. Another viable explanation could be that there was an expansion of the HLA class I lineages. To answer such questions, we designed a study that investigated the influence of selection operating on the classical MHC class I loci. This involved the comparison of intron variation of chimpanzee and human MHC class I alleles. The different intron 2 variations observed in both species are clustering into various lineages, of which some are shared (Figures 3 and 4). On the basis of two well-defined populations, the MHC class I intron 2 data regarding twenty-five randomly chosen human Caucasian individuals [77] were compared with the data regarding our twenty-five P.t.v. chimpanzees. The χ² statistics showed that the intron 2 variation found in humans is 2.56 times higher (confidence interval (CI) 95% is 0.87-7.55, P = 0.07) for the MHC-A locus, and 2.64 times higher (CI 95% is 1.20-5.82, P = 0.01) for the B locus [78]. The human cohort used, however, was not typed for the HLA-C locus, and therefore we have reanalyzed the data using a different human cohort that is typed at high-resolution level for HLA-A, B, and C (random panel IHB). To determine for all three loci whether the number of intron 2 alleles detected in humans is significantly higher as compared to that in chimpanzees, the difference in unique number of alleles (Δne) was calculated by using bootstrap analyses. Additionally, for each locus, the ratio of the number of unique alleles in humans divided by the number of unique alleles in chimpanzees was calculated. The Δne and the ratios of unique alleles were considered statistically significant if their confidence interval did not enclose 0 or 1, respectively. The analyses revealed that all three loci show a statistical significant difference for the Δne and ratio in humans versus chimpanzees (Table 3). Thus, the approach confirms that chimpanzees indeed experienced a selective sweep targeting the MHC class I repertoire. The repertoire condensation was claimed to predate the (sub)speciation of chimpanzees, as the same
Intron lineages are shared between the different populations of chimpanzees, and was calibrated to have happened approximately 2-3 million years ago [78] (Figure 2). This suggests that in the distant past, ancestors of the contemporary chimpanzee populations must have been in close contact with each other, and that a plague may have affected the entire species. The fact that chimpanzees experienced a repertoire reduction in the MHC class I region, and the knowledge that most chimpanzees infected experimentally with HIV-1 are relatively resistant to developing AIDS, resulted in the hypothesis that the ancient repertoire reduction may have been caused by an HIV-1/SIVcpz-like retrovirus [78].

A successive comparative genomic approach, using microsatellite markers located in the MHC region and markers mapping at a variety of other chromosomes, revealed that both chimpanzees and humans show similar amounts of variation for most of the non-MHC markers. Nonetheless, a significant lower allelic variation for the majority of the markers mapping in the chimpanzee MHC region was documented [79]. These data support the claim that chimpanzees possess a reduced MHC repertoire. Subsequent multi-locus demographic analyses highlighted the fact that chimpanzees experienced a selective sweep that mainly targeted the chromosomal segment carrying the MHC class I region [79].

The evolution of chimpanzee MIC genes: further substantiation of selection

Activated CD4+ memory T cells are selectively and rapidly depleted after an HIV-1/SIV infection [80,81]. This cell type is present in abundance in the intestine, but other mucosal surfaces (lung, vagina) that are frequently exposed to environmental pathogens also contain a high percentage of activated lymphocytes [82]. In humans and macaques in particular, the intestine is the prominent site of CD4+ T-cell depletion during the first few weeks after exposure to HIV-1/SIV [81,83]. The major histocompatibility complex class I chain-related gene (MIC) molecules are highly expressed on the gastrointestinal epithelium [84], and the genes encoding these molecules map to chromosome 6 within the MHC region. In humans, seven genes (MIC A to MIC G) have been distinguished, though only the MIC A and -B genes produce functional transcripts and display high levels of polymorphism [70]. The molecules are thought to be involved in signaling stress to the immune system.

Chimpanzees appear to have only one functional MIC gene (localized near the Patr-B locus), which has an intermediate character resembling in part the human MICA and -B genes. A large ancestral deletion of 95 kb resulted in the birth of a hybrid Patr-MICA/B gene [85]. We have investigated MIC gene polymorphism in chimpanzees, and found that the Patr-MICA/B fusion gene controls only one lineage showing moderate allelic variation [86]. Thus, as compared to humans, chimpanzees have a reduced MIC gene repertoire, which is consistent with the selective sweep observed for the Patr-B locus. Moreover, the hybrid MICA/B gene appears to represent a fixed entity in all chimpanzee (sub)species. This is peculiar, as most other primate species possess haplotypes that carry both MICA and -B genes. Hence, somewhere during chimpanzee evolution a haplotype with a hybrid MICA/B gene was generated, and all other haplotypes carrying MICA and -B genes have gone missing. As the MICA/B gene maps next to the Patr-B gene, it is possible that selection on particular B alleles favored the preferential selection of this hybrid MIC gene by way of a piggyback effect. However, it cannot be excluded that the hybrid MICA/B gene itself, or the combination of this gene with particular Patr-B alleles, provided a selective advantage.

In humans, MIC is the ligand for the NKG2D receptor, which is expressed on natural killer (NK) cells, γδ T cells, and particular CD8αβ T cells [87]. A cellular “stress” signal, triggered for instance by a viral infection or malignant transformation, may upregulate the expression of MIC, which can ultimately lead to an immune response. The Patr-MICA/B fusion molecule is recognized by human Vδ1 γδ T cells specific for MICA and B, suggesting a conserved recognition site [88]. In addition, the NKG2D
receptor of humans and chimpanzees is also highly similar [89]. Whether MIC plays a role in NK effector responses against HIV-1/SIVcpz has yet to be proven, as does the functional role of the Patr-MICA/B gene in viral infections [86].

Signs of selection beyond the MHC in chimpanzees

HIV-1 targets CD4+ T cells by the use of the CD4 receptor and different co-receptors. Orthologs of the human CD4 receptor, and of the CCR5 and CXCR4 co-receptors, are present in chimpanzees. In humans, several genetic CCR5 modifications are associated with resistance/susceptibility to HIV-1 infection [90]. The CCR5 variant that possesses a 32-base pair deletion (CCR5-Δ32) conferring nearly complete resistance to HIV-1 infection in homozygous individuals is present in approximately 1% of the Caucasian population [91]. In chimpanzees, no 32-base
Table 3 Difference in the number of unique alleles ($\Delta n$ (95% CI)) and the ratio of unique alleles (95% CI) in humans as compared to chimpanzees for the different MHC class I loci

| Locus   | $\Delta n$ (95% CI) | Ratio (95% CI) |
|---------|---------------------|---------------|
| Mhc-A   | 4.00 (2.00 - 7.00)  | 1.80 (1.33 - 2.40) |
| Mhc-B   | 5.00 (2.00 - 8.00)  | 1.63 (1.22 - 2.17) |
| Mhc-C   | 6.00 (3.00 - 8.00)  | 2.25 (1.60 - 3.67) |

For the statistical analysis, 50 haplotypes were re-sampled 100,000 times and with replacement, separately for both species. The $\Delta n$ was calculated by subtracting the $n_c$ of the chimpanzee from the $n_h$ of the human, and the ratio of unique alleles in the two species was calculated by dividing the $n_h$ of the human by the $n_c$ of the chimpanzee. The median was defined as the 50,000th value and the lower and upper confidence limits were defined as the 25th and 99.75th values emanating from the bootstrapping. Calculations were performed using the R statistical package (version 2.15.3).

pair deletion in the CCR5 gene has been observed, and the relative resistance of chimpanzees to developing AIDS after an HIV-1 infection cannot be attributed to this genetic modification [92-94]. Variations in the human 5’ cis-regulatory region of CCR5 (5’CCR5) were found to be associated with different transcription levels that influence HIV-1 entry and may affect disease progression [95]. Hence, the human 5’CCR5 haplotype that shows the lowest promoter activity resulting in control of AIDS development is the most common haplotype in chimpanzees [96].

By assuming a wide range of demographic histories, Wooding et al. [96] demonstrated that the human 5’CCR5 promoter region experienced balancing selection, and that, in contrast, the chimpanzee equivalent was affected by a selective sweep. This suggests that if the chimpanzee 5’CCR5 promoter region is linked to functional variants that influence progression to AIDS, it may contribute to the relative resistance of chimpanzees to developing AIDS after an HIV-1/SIVcpz infection. Corresponding to this finding, MacFie et al. [97] demonstrated differential patterns of diversity for the HIV-related loci CCR5, CXCR4, and CX3CR1 in three chimpanzee subspecies (West- and Central-African, and Nigeria-Cameroon) [97]. For the CCR5 locus, they showed that it has low levels of diversity for all three subspecies, and this appeared to be tightly centered, as flanking loci displayed normal variation in all subspecies. The results suggest that the CCR5 locus experienced a selective sweep and that this may have predated subspeciation. For the Central-African chimpanzees, natural infections with SIVcpz are documented [39], and a selective sweep at the CCR5 locus may be related to recent co-evolution with SIVcpz. For the other two subspecies, natural infections with SIVcpz are not documented in the contemporary living animals [44,98]. The fact that these subspecies do show evidence for selection at CCR5 could indicate that they were infected with SIVcpz-like virus strains in the past, and that infections are rare or absent at present. Alternatively, the entire ancestral chimpanzee population may have been infected with an SIVcpz-like virus, resulting in a selection event at that stage for which the signature still can be measured. Depending on novel infections with SIVcpz strains, particular selection forces are still operative.

Different cellular restriction factors, such as tripartite motif protein 5 alpha (TRIM5alpha) and apolipoprotein B mRNA-editing enzyme 3G (APOBEC3G), can target the intracellular replication of HIV-1 [99]. The APOBEC3G gene underwent strong positive selection, and during primate evolution this gene experienced several episodes of adaptations [100]. Polymorphisms in the human APOBEC3G gene have been documented, but its relationship with control of viral replication needs further investigation [101]. The chimpanzee APOBEC3DE, one of the other seven members of the APOBEC family, has potent antiretroviral activity against HIV-1. This is thought to be driven by an ancient lentiviral selective pressure dating back approximately 2 to 6 million years [102], which is consistent with our hypothesis that chimpanzees experienced a selective sweep caused by HIV-1/SIVcpz-like or a closely related retrovirus prior to chimpanzee (sub)speciation [78].

In addition, chimpanzees experienced selection on the fourth component (C4) gene of the complement system that maps in the MHC class III region [103]. Most humans possess two copies of the gene, designated C4A and C4B, and both genes are known to display size polymorphism due to the insertion of a complete endogenous retrovirus of 6.3 kb [104]. The long version of the C4A gene is found in humans, orangutans, and a variety of Old World monkeys [103,105]. However, chimpanzees only possess the equivalents of the short versions of the C4A and C4B genes. Since the presence of the long C4A gene predates human and Old World monkey speciation, the equivalent of this gene has most likely been lost during chimpanzee evolution.

Therefore, different research angles provide evidence that chimpanzees, as compared to humans, have experienced selection operating on the MHC and other genomic regions during their evolution. Chimpanzees shared a common ancestor with the bonobo (Pan paniscus) approximately 2 million years ago. A recent comparison of the bonobo and chimpanzee genome revealed that in chimpanzees, the MHC region in particular has experienced positive selection [106].

Ancient selective sweep in chimpanzees most likely caused by an SIV-like retrovirus: supporting data from functional studies

Control of AIDS development in HIV-1-infected human cohorts is strongly associated with the presence of HLA-B*27:05 and -B*57:01 [48,52-54,60], and adaptive immune responses to the HIV-1 Gag protein are thought to play an important role in the control of viral replication [107,108].
Hence, the Gag protein was taken as a model system to measure immune responses to this protein in chimpanzees. More specifically, we investigated whether the selective sweep had resulted in the preferential selection of Patr class I allotypes, and, if so, whether these allotypes could target similar HIV-1/SIV Gag regions as do HLA molecules associated with control of AIDS development. Therefore, the peptide-binding motifs of four Patr class I molecules, occurring at high frequency in a thoroughly characterized West-African chimpanzee population, were determined [109]. The obtained motifs were used to scan the HIV-1/SIV<sub>cpz</sub> Gag proteins for potential CTL epitopes, and the relevant peptides were subsequently tested in binding studies. Two of the studied Patr molecules have peptide-binding motifs that resemble those of HLA-B*27 or -B*57, and can target similar areas of the HIV-1/SIV<sub>cpz</sub> Gag protein (Figure 5). In addition, the two other studied allotypes, although divergent in their peptide-binding anchors, also appear to target the conserved areas of the HIV/SIV<sub>cpz</sub> proteome similar to the AIDS-controlling HLA-B*27 and -B*57 molecules. Thus, particular human and chimpanzee allotypes share similar qualitative functional characteristics.

Recently, an in silico approach comparing human and chimpanzee MHC class I alleles illustrated that, with regard to the peptide-binding repertoire, the Patr-A molecules show signs of a selective sweep [110]. The Patr-A allotypes, in general, appear to have promiscuous binding profiles, and as a consequence they can target conserved areas of the Gag protein. The chimpanzee-specific Patr-AL molecule has a peptide-binding motif similar to HLA-A*02 [111]. However, as compared to the classical MHC
class I loci, Patr-AL has distinct characteristics (low expression level and present only on approximately 50% of the chimpanzee MHC haplotypes), and further investigation is needed as to whether this molecule contributes significantly to adaptive immune responses.

In the founder population that we analyzed thoroughly, 94% of the chimpanzees possess at least one of the four Patr class I allotypes studied (Figure 5). Moreover, many chimpanzees were observed to express multiple allotypes, which are able to bind peptides derived from various conserved Gag regions. This has led us to suggest that chimpanzees may have developed a “double-lock” strategy to respond to an HIV-1/SIV<sub>cpz</sub> infection, meaning that the chance of the virus escaping by mutations has been severely diminished. As only four Patr allotypes have been studied at this point, this quantitative aspect of the immune response might become more prominent if more peptide-binding motifs become available. Evaluation of the immune response data of three different HIV-1-infected chimpanzees revealed that the animals displayed broadly reactive CTL responses to conserved epitopes of the Gag protein [109,112,113]. Thus, the functional characteristics of the chimpanzee MHC class I repertoire suggest that the ancient selective sweep was caused by a lentiviral pandemic. The effect of this sweep can still be measured, as the chimpanzee MHC class I repertoire resembles HLA-B*27, and only three CD8<sup>+</sup> T-cell epitopes appear to be responsible for the T-cell responses controlling replication of the pathogenic SIV<sub>mac251</sub> in these ECs [118]. Two of the epitopes, Vif RL9 and Nef RL10, are conserved, and similar motifs are present in SIV<sub>cpz</sub> and HIV-1 (Figure 6). Modeling the peptide binding of the proteins Vif and Nef to HLA-B<sup>+</sup>27:05 and the four studied Patr class I allotypes showed that they can target these conserved regions of Vif RL9 and Nef RL10 as well (Figure 6). This may suggest that, in different species, evolutionarily unrelated MHC class I molecules possessing similar peptide-binding motifs are important for control of the lentiviral infection.

Conclusions

The selective sweep in chimpanzees: a mirror of humankind’s future?

A remaining question concerns how many chimpanzees died during the pandemic that may have been caused by an HIV-1/SIV<sub>cpz</sub>-like retrovirus. This is a difficult question to answer, but it is likely that the number must have been huge. Natural history has provided a few examples in which the introduction of a “novel” virus into a naïve population resulted in mass mortality. For instance, contact between European individuals and the native population of America led to the death of millions of Amerindians as a result of infection with measles, small pox, and other viruses [120]. The deliberate infection of rabbits with the myxoma virus exterminated approximately 99% of the rabbit population in Australia [121]. Natural infections in animals can also have strong effects on population size, as is illustrated by the rinderpest epidemic in African buffalo [122], and by the decimation of seals by morbilli and influenza viruses [123,124].

Modern humans, who have their cradle in Africa, have existed for about 150,000-200,000 years [125], and display an impressive amount of MHC class I polymorphism [126]. Recent cohort studies have demonstrated that particularly those individuals equipped with HLA-B*27 and/or -B*57 are able to control an HIV-1 infection [49,50]. Nevertheless, in these individuals one also sees that an escape by the virus may occur, although mostly with a cost to viral fitness [127-129]. Furthermore, there is an additional effect on viral control caused by a phenomenon such as heterozygous advantage [130,131]. Chimpanzees may reflect the latter stages of selection. First, the Patr molecules have the ability to respond to an HIV-1/SIV<sub>cpz</sub> infection. Second, a quantitative aspect seems to be operative in chimpanzees, as each individual may possess several Patr molecules that can respond to different conserved parts of the HIV-1/SIV<sub>cpz</sub> proteome, termed “double lock” strategy.
The selective sweep in chimpanzees resulted in the preferential selection of Patr class I allotypes that can target conserved areas of Gag similar to those targeted by HLA-B*27/B*57. In humans, these allotypes also seem to predispose for the immune-mediated pathologies ankylosing spondylitis and psoriasis, respectively. Evidence for the presence of these immune-mediated pathologies in chimpanzees is very low [132,133]. In humans, the exact role of HLA-B*27 in ankylosing spondylitis is also not known. Moreover, chimpanzee MHC class I molecules can target similar conserved motifs such as HLA-B*27/B*57, but this does not tell us whether the structure of the chimpanzee MHC class I molecules are comparable to HLA-B*27/B*57, as these molecules belong to different ancestral lineages [72,112].

In general, a virus survives best if it is able to replicate and disseminate itself within a population without killing all of its hosts. Such a host/virus state of equilibrium is, for instance, reached in different non-human primate species that are infected naturally with SIV strains. For HIV-1 and its human hosts, the battle seems to be in full swing, and without proper treatment the human population may be hit hard by this pandemic. For humans up until now, control of HIV-1 replication, without treatment, is significantly associated with the MHC region, in particular with the MHC class I molecules HLA-B*27:05 and -B*57:01, which can target conserved regions of the HIV-1 Gag protein. The observations adduced in this review have important consequences for vaccine design, as different HLA specificities may target different sections of the HIV-1 proteome. One major implication is that many human individuals may not possess HLA class I molecules that have the capacity to bind conserved HIV-1 epitopes that are associated with control of AIDS development. A truly protective vaccine has not yet been reported. Therefore, until that moment arrives, prevention of the infection itself will be, for the individual, one of the most important goals as regards surviving the current HIV-1 pandemic.

### Competing interests
The authors declare that they do not have competing interests.

### Authors’ contributions
NG and RB wrote the manuscript. Both authors have given their final approval of the manuscript.

---

| Consensus | Vif RL9 | Nef RL10 |
|-----------|---------|----------|
| SIVmac239 | VRRAIRGEQL | ARRHRILDILEK |
| SIVcpz   | iRKAllGerv | rrReILD1wyh |
| HBX2     | IRKALLGHIV | QRRQDLIDLWIYH |

**Figure 6** Conserved regions around the Mamu-B*008 epitopes Vif RL9 and Nef RL10 that can be targeted by HLA and Patr class I molecules. P2 and PΩ (carboxyl-terminus) represent the anchor-binding positions of the MHC class I molecules; amino acids that are preferred on these position are indicated by the conventional one-letter code, whereas amino acids indicated between brackets are tolerated. The epitopes Vif RL9 and Nef RL10 in SIVmac239 are indicated by a red line [118]. For HIV-1 (HBX2) and SIVcpz, the consensus sequences surrounding these Vif and Nef epitopes are given (www.hiv.lanl.gov). A lower-case letter in the SIVcpz consensus indicates a variable position. A red arrow indicates that the respective MHC class I molecule is predicted to target the SIVmac239 consensus sequence. A grey arrow indicates prediction of the MHC molecule to target the SIVcpz/HBX2 consensus sequence. SB stands for solid binding (affinity < 100 nM); WB stands for weak binding (affinity between 100 to 700 nM). Binding affinity of the peptides was predicted using the NetMHCpan algorithm [119], except for Patr-B*03:01 (white dotted arrows), for which binding was assumed based on agreement in the anchor binding positions.
immunodeficiency virus infection in wild-living eastern chimpanzees (Pan troglodytes schweinfurthii). J Virol 2003, 77:7545–7562.

42. Heeney JL, Rutjens E, Verschoor EJ, Niphuis H, ten Haaf P, Rouse S, McClure H, Baffa-Jaghoorsingh S, Bogers W, Salas M, et al: Transmission of simian immunodeficiency virus SIVcpz and the evolution of infection in the presence and absence of concurrent human immunodeficiency virus type 1 infection in chimpanzees. J Virol 2006, 80:2708–2718.

43. Keele BF, Jones JT, Terio KA, Estes JD, Rudicell RS, Wilson ML, Li Y, Leaman GH, Beasley TM, Schumacher-Stankey J, et al: Increased mortality and AIDS-like immunopathology in wild chimpanzees infected with SIVcpz. Nature 2009, 460:515–519.

44. Etienne L, Nemier M, LeBlon M, Bibila GT, Fouppououognigni Y, Roussel D, Nana A, Okjo CT, Tarnoule U, Aghokeng AF, et al: Characterization of a new simian immunodeficiency virus strain in a naturally infected Pan troglodytes troglodytes chimpanzee with AIDS related symptoms. Retrovirology 2011, 8:4.

45. Novembre FJ, Sauier M, Anderson DC, Klump SP, O'Neil SP, Brown CR 2nd, Hart CE, Guenther PK, Svenson RB, McClure HM: Development of AIDS in a chimpanzee with human immunodeficiency virus type 1. J Virol 1997, 71:4086–4097.

46. Deeks SG, Walker BD: Human immunodeficiency virus controllers: mechanisms of durable virus control in the absence of antiretroviral therapy. Immunity 2007, 27:406–416.

47. Carrington M, OBrien SJ: The influence of HLA genotype on AIDS. Annu Rev Med 2003, 54:535–551.

48. Dalmasso C, Carpentier W, Meyer L, Rouxoux C, Goujard C, Chalx ML, Lambotte O, Aventaud-Fenelle V, Le Clerc S, de Senneville LD, et al: Distinct genetic loci control plasma HIV-1 RNA and cellular HIV-DNA levels in HIV-1 infection: the ANRS genome wide association 01 study. PLoS One 2008, 3:e3097.

49. Fellay J, Sha金华 LV, De Ce, Dolombo S, Ledergerber B, Weale M, Zhang K, Gumbs C, Castagna A, Cossarizza A, et al: A whole-genome association study of major determinants for host control of HIV-1. Science 2007, 317:944–947.

50. Goulder PJ, Bunce M, Krausa P, McIntyre K, Crowley S, Morgan B, Edwards A, Carrington M, O'Brien SJ, Rinaldo C: An update of the limits of HLA restriction of HIV-1 infection: the ANRS genome wide association 01 study. Retrovirology 2011, 9:A18.

51. Klein L: Parental HLA type and susceptibility to chronic diseases: a case control study. Am J Hum Genet 1987, 41:155–162.

52. de Groot NG, Otting N, Arguello R, Watkins DI, Doshi GD, Madrigal JA, Bontrop RE: Major histocompatibility complex class I A locus alleles are related to only one of the six families of human A locus alleles. J Immunol 1995, 154:6421–6429.

53. Fujimura A, Watanabe H, Toyoda A, Taylor TD, Itoh T, Tsai SF, Park HS, Yasaki ML, Lebrach C, Chen Z, et al: Construction and analysis of a human-chimpanzee comparative clone map. Science 2002, 295:131–134.

54. Goulder PJ, Bunch M, McCollum H, Lopez R, Parham P, Marsh SG: The IMGT/HLA database. Nucleic Acids Res 2011, 39:D1171–D1176.

55. de Groot NG, Otting N, Robinson JJ, Blancher A, Lafont BA, Marsh SG, O'Connor DH, Ehnta T, Walter L, Watkins DJ, Bontrop RE: Nomenclature report on the major histocompatibility complex genes and alleles of Great Ape, Old and New World monkey species. Immunogenetics 2012, 64:615–631.

56. de Groot NG, Otting N, Arguello R, Watkins DI, Doshi GD, Madrigal JA, Bontrop RE: Major histocompatibility complex class I diversity in a West African chimpanzee population: implications for HIV research. Immunogenetics 2000, 51:398–409.

57. Adams EJ, Cooper S, Thomson G, Parham P: Common chimpanzees have greater diversity than humans at two of the three highly polymorphic MHC class I genes. Immunogenetics 2000, 51:410–424.

58. Caggiari L, Simula MP, Marzotto A, Shina M, Rehemann B, De Re V: Identification of novel chimpanzee class I and II alleles using an improved sequence-based typing strategy. Hum Immunol 2006, 67:833–72.

59. Chen ZW, Hughes AL, Ghim SH, Lehtin LN, Watkins DJ: Two more chimpanzee Patr-A& B locus alleles related to the HLA-A1/A3/A11 family. Immunogenetics 1993, 38:238–240.

60. Kowalski H, Erickson AL, Cooper S, Domena JD, Parham P, Walker CM: Patr- A & B, the orthologues of HLA-A and B, present hepatitis C virus epitopes to CD8+ cytotoxic T cells from two chronically infected chimpanzees. J Exp Med 1996, 183:1761–1775.

61. Zhang L, van Bree S, Gibbey Y, van Veen A, van Rood JJ, Class FH: Comparison of the humoral and cytotoxic T-lymphocyte response to individual HLA class I alleles in highly immunized chimpanzeas. Hum Immunol 1991, 30:156–161.

62. de Groot NG, Otting N, Doshi GD, Baffa-Jaghoorsingh SS, Heeney JL, van Rood JJ, Gagneux P, Bontrop RE: Evidence for an ancient selective sweep in the MHC class I gene repertoire of chimpanzeas. Proc Natl Acad Sci USA 2002, 99:11748–11753.

63. de Groot NG, Heijmans CM, de Groot N, Otting N, de Vos-Rouweller AJ, Remarque EJ, Bonhommme M, Doshi GD, Coureau-Roy B, Bontrop RE: Pinpointing a selective sweep to the chimpanzee MHC class I region by comparative genomics. Mol Ecol 2008, 17:2074–2088.

64. Brenchley JM, Schacker TW, Price DA, Taylor JH, Beilman GJ, Nguyen P, Khoruts A, Larson M, Hawn TR, Douek DC: CD4+ T cell depletion during all stages of HIV disease occurs predominantly in the gastrointestinal tract. J Exp Med 2004, 200:749–759.

65. Mehndru S, Poiesz MA, Tenner-Racz K, Horowitz A, Hurley A, Hogan C, Boden D, Racz P, Markowitz M: Primary HIV-1 infection is associated with preferential depletion of CD4+ T lymphocytes from effector sites in the gastrointestinal tract. J Exp Med 2004, 200:761–770.
Stress management: MHC class I and class I-like HIV-1 replication. Mol Biol Evol 2005, 22:1375–1385.

Gleimer M, Parkham P. Stress management: MHC class I and class II-like molecules as reporters of cellular stress. Immunology 2003, 109:469–477.

Shum BF, Flodin UF, Muir DG, Rajaligan R, Khakoo SI, Cleland S, Guethlein LA, Uhrlberg M, Parkham P. Conservation and variation in human and common chimpanzee CD94 and NKG2 genes. J Immunol 2002, 168:240–252.

Bontrop RE, Broos LA, Otting N, Jonker MJ. Comparative sequencing of non-human primate CCR-5 chemokine receptor gene. J Mol Evol 1998, 45:1200–1209.

Pretet JL, Zerbib AC, Girard M, Guillet JG, Butor C. Apobec3DE in chimpanzees has been driven by positive selection. J Immunol 2002, 168:2015–2021.

Gonzalez EB, Banshod M, Sato N, Mumndi S, Dhanda R, Catano G, Cabrera LA, Uhrberg M, Parham P. Patterns of further structural polymorphism of the complement C4 genes in old world primates. Exp Clin Immunogenet 2001, 18:131–138.

Mudd PA, Martins MA, Ericsen AJ, Tully DC, Power KA, Bean AT, Piaskowski MD, Peters B, Sidney J, Pedersen LE, Sette A, Lund O, Buus S, Nielsen M. Characterization of one family member, HERV-K(C4), located in the C4q subregion of the C4 complement gene cluster. Nucleic Acids Res 2005, 33:7713–7720.

Sloot G, Deelder AM, Adema GJ, Zwaal RF, Haase AT. Characterization of one family member, HERV-K(C4), located in the C4q subregion of the C4 complement gene cluster. Nucleic Acids Res 2005, 33:7713–7720.

Van Heuswyn F, Li Y, Bailes E, Neel C, Lafay B, Keele BF, Shaw KS, Takehisa J, Kraus MH, Loul S, et al. Rapid detection of the ERV-K(C4) retropiral insertion reveals further structural polymorphism of the complement C4 genes in old world primates. Exp Clin Immunogenet 2001, 18:131–138.

Rittner C: Characterization of one family member, HERV-K(C4), located in the C4q subregion of the C4 complement gene cluster. Nucleic Acids Res 2005, 33:7713–7720.

McMichael AJ: Environmental and social influences on emerging infectious diseases: past, present and future. Philos Trans R Soc Lond B Biol Sci 2004, 359:949–958.

Silvers L, Inglis B, Labudovic A, Janssens PA, van Leeuwen BH, Kerr PJ: Virulence and pathogenesis of the MSW and MSD strains of Californian myxoma virus in European rabbits with genetic resistance to myxomatosis compared to rabbits with no genetic resistance. Virology 2006, 348:72–83.
African buffalo maintain high genetic diversity in the major histocompatibility complex in spite of historically known population bottlenecks. *Mol Ecol* 1998, 7:1315–1322.

Pathologic and immunocytochemical studies of morbillivirus infection in striped dolphins (*Stenella coeruleoalba*). *Vet Pathol* 1992, 29:1–10.

Mass mortality of harbor seals: pneumonia associated with influenza A virus. *Science* 1982, 215:1129–1131.

Why did modern human populations disperse from Africa ca. 60,000 years ago? A new model. *Proc Natl Acad Sci USA* 2006, 103:9381–9386.

Population biology of antigen presentation by MHC class I molecules. *Science* 1996, 272:67–74.

Escape from the dominant HLA-B27-restricted cytotoxic T-lymphocyte response in Gag is associated with a dramatic reduction in human immunodeficiency virus type 1 replication. *J Virol* 2007, 81:12382–12393.

Variable fitness impact of HIV-1 escape mutations to cytotoxic T lymphocyte (CTL) response. *PLoS Pathog* 2009, 5:e1000365.

Escape from the dominant HLA-B*57/B*5801 human immunodeficiency virus type 1 elite controllers select for rare gag variants associated with reduced viral replication capacity and strong cytotoxic T-lymphocyte *corrected* recognition. *J Virol* 2009, 83:2743–2755.

Escape from the dominant HLA-B^*27^-restricted cytotoxic T-lymphocyte response in Gag is associated with a dramatic reduction in human immunodeficiency virus type 1 replication. *J Virol* 2007, 81:12382–12393.

Submit your next manuscript to BioMed Central and take full advantage of:

- Convenient online submission
- Thorough peer review
- No space constraints or color figure charges
- Immediate publication on acceptance
- Inclusion in PubMed, CAS, Scopus and Google Scholar
- Research which is freely available for redistribution

Submit your manuscript at www.biomedcentral.com/submit