JOURNAL OF GENERAL VIROLOGY

RESEARCH ARTICLE
Capone *et al.*, *Journal of General Virology* 2018;99:1717–1728
DOI 10.1099/jgv.0.001159

MICROBIOLOGY SOCIETY

OPEN MICROBIOLOGY

Genetic diversity in the env V1-V2 region of proviral quasispecies from long-term controller MHC-typed cynomolgus macaques infected with $SHIV_{SF162P4cy}$

Alessia Capone,[1,2,3] Alessandra Lo Presti,[4] Leonardo Sernicola,[1] Stefania Farcomeni,[1] Flavia Ferrantelli,[5] Maria T. Maggiorella,[1] Edward T. Mee,[6] Nicola J. Rose,[6] Eleonora Cella,[7] Massimo Ciccozzi,[7] Barbara Ensoli[1] and Alessandra Borsetti[1,*]

Abstract

Intra-host evolution of human immunodeficiency virus (HIV) and simian immunodeficiency virus (SIV) has been shown by viral RNA analysis in subjects who naturally suppress plasma viremia to low levels, known as controllers. However, little is known about the variability of proviral DNA and the inter-relationships among contained systemic viremia, rate of reservoir reseeding and specific major histocompatibility complex (MHC) genotypes, in controllers. Here, we analysed the proviral DNA quasispecies of the *env* V1-V2 region, in PBMCs and in anatomical compartments of 13 long-term controller monkeys after 3.2 years of infection with simian/human immunodeficiency virus $(SHIV)_{SF162P4cy}$. A considerable variation in the genetic diversity of proviral quasispecies was present among animals. Seven monkeys exhibited *env* V1-V2 proviral populations composed of both clusters of identical ancestral sequences and new variants, whereas the other six monkeys displayed relatively high *env* V1-V2 genetic diversity with a large proportion of diverse novel sequences. Our results demonstrate that in $SHIV_{SF162P4cy}$-infected monkeys there exists a disparate pattern of intra-host viral diversity and that reseeding of the proviral reservoir occurs in some animals. Moreover, even though no particular association has been observed between MHC haplotypes and the long-term control of infection, a remarkably similar pattern of intra-host viral diversity and divergence was found within animals carrying the M3 haplotype. This suggests that in animals bearing the same MHC haplotype and infected with the same virus, viral diversity follows a similar pattern with similar outcomes and control of infection.

INTRODUCTION

Simian/human immunodeficiency virus (SHIV) infection in nonhuman primates (NHPs) has proven invaluable in providing insights into human immunodeficiency virus-1 (HIV-1) pathogenesis and for intervention studies [1]. Preclinical studies on different monkey species have indicated that a differential susceptibility to primate lentivirus-induced diseases could be related to the genetic diversity of the animals [2–6]. $SHIV_{SF162}$ is a C-C chemokine receptor type 5 (CCR5)-tropic virus capable of establishing persistent infection and causing simian acquired immune deficiency syndrome (AIDS) with a varying disease progression that is characteristic of HIV disease in humans [2, 7]. Rhesus macaques can be readily infected with $SHIV_{SF162}$ viral isolates, while cynomolgus macaques are more resistant to infection, consistent with the evidence of host factors that affect species-level differences in infection susceptibility and disease progression [8]. In cynomolgus monkeys, the acute phase of infection is followed by a prolonged asymptomatic phase, in which monkeys can efficiently control viral replication, maintaining it at very low levels during chronic infection in the absence of treatment, similar to HIV-infected humans

Received 8 June 2018; Accepted 20 September 2018
Author affiliations: [1]National HIV/AIDS Research Center, Istituto Superiore di Sanità, V.le Regina Elena 299, 00161 Rome, Italy; [2]Neuroimmunology Laboratory, Fondazione Santa Lucia, Rome, Italy; [3]Department of Biology and Biotechnology Charles Darwin, Sapienza University, Rome, Italy; [4]Department of Infectious Diseases, Istituto Superiore di Sanità, V.le Regina Elena 299, 00161 Rome, Italy; [5]National Center for Global Health, Istituto Superiore di Sanità, V.le Regina Elena 299, 00161 Rome, Italy; [6]Division of Virology, National Institute for Biological Standards and Control, Medicines and Healthcare product Regulatory Agency, South Mimms, Hertfordshire, EN6 3QG, UK; [7]Medical statistic and molecular epidemiology unit, University campus bio medico, Roma, Italy.
***Correspondence:** Alessandra Borsetti, alessandra.borsetti@iss.it
Keywords: $SHIV_{SF162P4cy}$; MCM; MHC; Viremic controllers; diversity; evolution.
Abbreviations: AIDS, acquired immune deficiency syndrome; AUC, area under the curve; bAb, binding antibody; CCR5, C-C chemokine receptor type 5; GI, gastrointestinal; HIV, human immunodeficiency virus; HLA, human leukocyte antigen; LTNP, long-term nonprogressor; Mamu, Macaca Mulatta; MCM, Mauritian cynomolgus macaques; MHC, major histocompatibility complex; nAb, neutralizing Ab; NJ, neighbour-joining; NHPs, nonhuman primates; PNLG, N-linked glycosylation; SIV, simian immunodeficiency virus; SHIV, simian/human immunodeficiency virus.
Four supplementary figures and two supplementary tables are available with the online version of this article.

who control infection and become long-term nonprogressors (LTNPs).

Despite control of viremia, human HIV controllers present ongoing evolution and divergence of viral RNA sequences, although at a significantly lower rate than that observed for typical progressors [9]. In contrast, the DNA proviral population is highly homogenous, mostly composed of ancestral sequences, suggesting that in these individuals ongoing replication does not permit the significant reseeding of the latent reservoir [10]. However, given the ability of HIV and SIV to establish a stable latent viral reservoir in anatomic sites early in infection, viral replication might persist in tissues; this is the case where immune responses developed by controllers fail to eradicate HIV/SIV infections [11]. Recent data suggest that latent reservoirs may be established before virus can be detected in the peripheral blood, making it difficult to treat HIV-1 early enough to avoid reservoir seeding. Latently infected cells remain undetectable by the immune system and can persist for years without losing their ability to produce infectious virus [10, 12]. There are conflicting data on whether new virus variants emerge in distinct anatomic sites by restriction of viral gene flow, with consequent viral evolution and divergence from virus present in the blood, or they disseminate to other anatomic sites through ongoing replication [13]. Thus, virus genetic variations across the body might provide insights to understand the potential correlation between suppression of viral replication and rate of reservoir reseeding in controllers.

Control of systemic viral suppression levels has been correlated with particular MHC class I alleles, such as human leukocyte antigen (HLA)-B27, B57 in humans, or Macaca mulatta (Mamu)-B17 and Mamu-B08 in macaques, although specific genotypes are only associated with, and not predictive of viral control [14–16]. The role of monkey models has been crucial for the elucidation of the mechanisms that underlie control of virus replication because the viral strain, host genotype as well as timing and route of infection can be controlled [17–22]. Furthermore, animal models enable extensive tissue collection at necropsy, facilitating studies of viral reservoirs and evaluation of virus variability [23–26].

In our previous work, we showed the effects of MHC haplotypes on early and late $SHIV_{SF162P4cy}$ infection in Mauritian cynomolgus macaques (MCM), highlighting the importance of considering host-related genetic background and immunological factors in the evaluation of vaccine efficacy in the different monkey species [20, 21].

Here, we describe the proviral DNA quasispecies diversity and phylogenetic relationships among proviral variants at the *env* V1-V2 region, in the peripheral blood and in lymphoid, gastrointestinal (GI) and genital anatomical compartments, of 13 long-term controller $SHIV_{SF162P4cy}$-infected MCM with defined MHC haplotypes. The pattern of changes of $SHIV_{SF162P4cy}$ DNA populations in the setting of long-term undetectable viremia over 170 weeks and the association with putative restrictive MHC alleles were also investigated.

RESULTS

A cohort of 13 monkeys experimentally infected with $SHIV_{SF162P4cy}$ was followed for 3.2 years.

LNTPs were defined as monkeys surviving $SHIV_{SF162P4cy}$ infection more than 2 years post infection (p.i.) and having plasma viral RNA copies below the detection limit (<50 copies ml^{-1}). In this cohort, all animals had stable CD4+ T and CD8+ T cell numbers; from week 46 they remained plasma RNA-negative throughout the period of observation of 170 weeks (data not shown), and no animal had signs of disease development.

Antibody response kinetics showed that all animals mounted robust anti-Env binding antibody (bAb) responses with peak titres ranging between 1 : 800 and 1 : 25600, which declined by week 170 p.i. in 9 of 13 animals. The remaining four monkeys maintained detectable anti-Env bAb throughout this period of time. Similarly, at 170 weeks p.i., neutralizing Ab (nAb) titres diminished or became undetectable in all but three animals, which had persistent homologous nAbs.

For detection and quantification of viral DNA in the lymphoid, GI and genital anatomic locations of viral-suppressed animals at 170 weeks p.i., quantitative PCR assays targeting the *gag* gene were performed.

As shown in Table 1, low levels of proviral DNA were detected in PBMC (<1 to 11 copies μg^{-1}) whereas in tissues, the virus was most prominent in the GI jejunum/ileum tract (2 to 152 copies μg^{-1}). Lower levels of viral DNA were detected in lymphoid tissues such as axillary lymph nodes, inguinal lymph nodes, spleen (2 to 112 copies μg^{-1}) and genital tissues including epididymis, testis, prostate, penis, rectum (2 to 48 copies μg^{-1}), with epididymis (P=0.0191) and testis (P=0.0302) displaying higher levels of viral DNA compared to prostate (Fig. 1a). The low level of viral DNA measured in the tissues may be due, at least in part, to a relatively low frequency of target cells (CD4+ T; macrophages) in the tissue specimens [23, 25]. Interestingly, statistical analysis demonstrated a positive (relatively weak) correlation between genital viral DNA levels at necropsy and both acute phase plasma viremia (P=0.0485), (Fig. 1b), and the area under the curve (AUC) of the RNA levels over 170 weeks p.i., a parameter reflecting cumulative plasma virus production (P=0.0492) (Fig. 1c).

Analyses of *env* V1-V2 sequence diversity in $SHIV_{SF162P4cy}$ controllers

For each animal, the mean Env diversity from the parental clone was calculated by pairwise comparisons and represents the average viral changes emerging after infection.

According to the degree of diversity of *env* V1-V2 variants identified, we clustered animals into four different groups, revealing that both diversity and complexity of virus

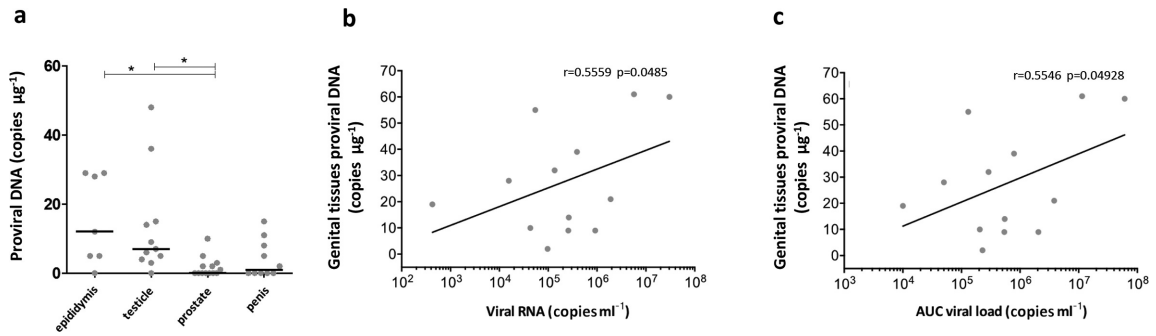

Fig. 1. Proviral load in tissue compartments and correlation between virological parameters in infected monkeys. $SHIV_{SF162P4cy}$ DNA viral load in epididymis, testicle, prostate and penis of infected macaques, by quantitative real-time PCR. Mean in each group is represented by black bars. Stars indicate statistical difference between organs with epididymis (*P*=0.0191) and testis (*P*=0.0302) displaying the higher levels of viral DNA as compared to prostate (a). The amount of proviral DNA detected in each of the genital tissues for each animal was plotted against the acute phase plasma viremia level for all animals (*P*=0.0485) (b). The amount of proviral DNA detected in each of the genital tissues for each animal was plotted against the area under the curve (AUC) of plasma viremia between weeks 1–170 p.i. (*P*=0.0492) (c).

quasispecies varied extensively among animals (Table 2). In group 1, the pattern observed in monkeys AU676, AQ271, AQ882, AG172, AS377, AK484 and AU427 was characterized by low genetic diversity as compared to the challenge virus. The mean genetic distances between $SHIV_{SF162P4cy}$ and PBMC at week 2 or week 170 p.i. ranged from 0.004 to 0.017 and from 0.001 to 0.061, respectively, whereas the overall mean pairwise distances from the inoculum, among tissue compartments, ranged from 0 to 0.075, with the exception of animal AS377, which exhibited higher genetic distances (0–0.295). This result indicated that in periphery and in the tissue compartments, the V1-V2 variants were composed mainly of ancestral sequences and some novel variants, which were widely disseminated during acute phase.

Group 2 monkeys (AS167, AH694 and AG981) displayed high levels of V1-V2 viral diversity (range, 0.006–0.310) with PBMC variants isolated at week 2 p.i. exhibiting a genetic distance ranging from 0.298 to 0.310. In contrast at 170 weeks p.i., quasispecies were more similar to those of the challenge virus, with a genetic distance ranging from 0.006 to 0.054. Despite the long-term follow up, quasispecies in the tissue compartments were similar to those of PBMC sampled at week 2 p.i. (genetic distance ranging from 0.006 to 0.041) but divergent from those of the challenge virus (range, 0.292–0.323). This finding indicated that V1-V2 variants present across tissues at necropsy were probably established during acute phase by spreading of viral variants arising early in infection.

Group 3 comprised two animals, AH960 and AK952. Here, V1-V2 quasispecies in the tissue compartments showed low diversity compared to those of week 2 PBMC and $SHIV_{SF162P4cy}$, (range 0.002–0.011), suggesting that they were derived from infected cell populations that had spread across tissues early in infection. In contrast, the high genetic diversity of week 170 PBMC variants (range, 0.187–0.414) indicated, most likely, a blood re-seeding from anatomic compartments, which were not investigated in this study.

Finally, group 4 comprised a single monkey, AP511, and was characterized by high diversity of week 2 PBMC variants as compared to $SHIV_{SF162P4cy}$ (0.29), whereas at 170 weeks, the V1-V2 proviral populations in both PBMC (0.077) and tissue (0.01) were quite homogeneous.

Overall these data show that a high viral diversity existed among infected animals. Moreover, virus challenge dose had no impact on the viral diversity at early and late time points of infection, since monkeys infected with low ($1.79MID_{50}$) or high ($179MID_{50}$) infectious doses showed comparable *env* V1-V2 diversity (0.03), which is supportive of our previous results [20, 21].

Phylogenetic analysis of *env* V1-V2 proviral quasispecies in $SHIV_{SF162P4cy}$ controllers

The phylogenetic relationship between *env* V1-V2 proviral quasispecies in the tissue compartments and those in the PBMC sampled at week 2 or 170 p.i. was assessed. A neighbour-joining (NJ) phylogenetic tree was generated initially to investigate the clustering of sequences within and between anatomic sites and PBMC. The potential genetic segregation of V1-V2 quasispecies between PBMC and anatomical compartments was confirmed by applying the Slatkin–Maddison test; this analysis can estimate and statistically measure the viral in/out gene flow between compartments.

Phylogenetic analysis identified three clades (A, B and C). Most of the V1-V2 sequences clustered into clade B and were closely related both to the $SHIV_{SF162P4cy}$ inoculum and to the reference sequence (accession number JN205735) (Figs 2 and S1–S3, available in the online version of this article). Specifically, 10 out of 13 macaque quasispecies

Table 1. Load of proviral DNA (copie μg^{-1}) in PBMC and tissues at necropsy (week 170)

	AS167	AH960	AS377	AU427	AP511	AG172	AH694	AU676	AQ882	AK484	AK952	AQ271	AG981
PBMC	<1	3	<1	5	<1	7	3	<1	<1	<1	4	11	<1
Axillary LN	17	13	8	8	NA	<1	NA	59	8	10	1	1	10
Inguinal LN	42	0	2	4	NA	NA	11	<1	9	116	NA	0	28
Spleen	<1	<1	3	15	<1	8	16	17	<1	11	16	18	<1
Jejunum	<1	2	36	5	<1	6	1	13	<1	28	102	<1	7
Ileum	3	NA	2	7	20	NA	72	61	152	52	20	<1	5
Epididymis	NA	5	NA	29	29	<1	28	NA	NA	12	5	NA	NA
Testis	48	4	5	<1	3	9	6	36	14	7	15	NA	NA
Prostate	5	<1	<1	10	<1	<1	3	2	<1	<1	<1	<1	2
Penis	5	11	<1	<1	<1	NA	15	<1	<1	NA	8	2	NA
Rectum	2	NA	5	<1	<1	NA	3	23	<1	<1	NA	7	NA

All values are the means for five different tissue samples of each tissue. NA, not available due to lack of tissue or due to low detectable viral DNA

located on clade B were intermixed, suggesting that in each monkey there has been an exchange of variants between periphery and tissue compartments, with the exception of week 170 PBMC variants of some animals that clustered on distinct branches, namely on clade C or externally to the clades.

This result was supported by the Slatkin–Maddison test that determined in each monkey statistically significant gene flows from week 170 PBMC variants to genital (range 20 %), gastrointestinal (range 50 %) and lymphoid tissues (range 50 %), and indicated a lack of anatomical compartmentalization. Interestingly, within the B clade five, statistically significant partially segregated sub-clusters, from 4 to 8, were identified indicating weak compartmentalization of quasispecies sustained by gene flow analysis.

In contrast, clade A included most of the V1-V2 sequences from macaques AS167, AH694, AG981 and week 2 PBMC variants of monkey AP511 (clade A, 100 % bootstrap value). In animals AS167, AH694 and AG981, tissue quasispecies genetically distant from the virus inoculum, clustered together with week 2 PBMC variants. This suggested that the population of quasispecies present in the acute phase persisted in the tissues over 170 weeks of infection, as clearly shown by sub-cluster 3 present in week 2 PBMC and tissue variants of monkey AG981. In contrast, in the same animals, PBMC variants present at weeks 170 p.i. were located on clade B and mostly related to the virus inoculum. This result pointed at a compartmentalization between blood and tissue at 170 weeks p.i., and a possible reseeding of the PBMC variants by reservoir from other anatomical sources. The Slatkin–Maddison analysis confirmed the absence of gene flow from/to PBMC and anatomical compartments at 170 weeks p.i.

Finally, clade C included V1-V2 variants derived from monkey AK484 epididymis and from week 170 PBMC of monkey AH960 and AK952. In monkey AH960, we observed statistically supported phylogenetic evidence of compartmentalization of PBMC and tissue quasispecies at week 170 p.i.

Analysis of single-site mutations

Analysis of the sequence alignment of all monkey quasispecies of $SHIV_{SF162P4cy}$, performed by Highlighter analysis, identified mutations shared across the animal variants (identical sequences were not included in the analysis) (Fig. S4). The majority of virus variants carried substitutions at eight codons: four at the amino acids K134, N135, A136, D148; two at the potential PNG sites K140, K158 in V1 region and two at amino acids R164 and K190 in the V2 region (Table S1). These variations appeared late in infection and simultaneously in blood cells and tissues, except for R164K/G changes that were found in both anatomical compartments either at weeks 2 and 170 p.i., suggesting a substantial amount of viral variant migrations, as already described.

Six out of eight substitutions detected in the *env* V1-V2 region of the viral variants have been previously reported by Balfe *et al.*, [27] in the $SHIV_{SF162P3}$ isolate that they termed the signature of the P3 variant. This signature made the V1-V2 regions of the long-term infected monkeys more like V1-V2 regions of $SHIV_{SF162P3}$ than those of the inoculum virus $SHIV_{SF162P4cy}$ or $SHIV_{SF162P4}$. This tendency was common to all macaques, suggesting that the same selection pressure drove the changes. As both $SHIV_{SF162P4cy}$ and $SHIV_{SF162P4}$ viruses were obtained from *in vivo* passage of $SHIV_{SF162P3}$, it is also possible but speculative that, in the absence of selective pressure upon transmission of $SHIV_{SF162P4cy}$ to a new host, the mutations were lost and reverted to $SHIV_{SF162P3}$ thereby conferring a fitness advantage to the virus during long-term infection. Overall, even if viral populations diverged to different levels in different animals, reversion by *de novo* mutation after passage to MHC-disparate hosts can occur, likely due to the absence of CTL response against specific epitopes.

Table 2. Mean genetic distances among $SHIV_{SF162P4cy}$ DNA variants in PBMC and tissue compartments of long-term infected monkeys

Group	Monkeys	Proviral diversity[a] from $SHIV_{SF162P4cy}$		Tissue compartment proviral diversity[a] from virus inoculum and PBMC			
		PBMC week 2	PBMC week 170	SHIV SF162P4cy	PBMC week 2	PBMC week 170	S.E.
1	**AU676**	**0.004** S.E. 0.004	**0.005** S.E.0.002	**0–0.007**	**0.002–0.006**	**0.005–0.012**	**<0.005**
	AQ271	**0.006** S.E. 0.003	**0.022** S.E. 0.011	**0–0.009**	**0.006–0.016**	**0.022–0.032**	**<0.007**
	AQ882	**0.006** S.E. 000.2	**0.009** S.E. 0.007	**0.001–0.007**	**0.006–0.013**	**0.010–0.017**	**<0.005**
	AG172	**0.010** S.E. 000.4	**0.047** S.E. 0.009	**0–0.007**	**0.010–0.015**	**0.047–0.056**	**<0.010**
	AS377	**0.012** S.E. 0.006	**0.061** S.E. 0.016	**0–0.295**	**0.009–0.296**	**0.061–0.075**	**<0.007**
	AK484	**0.017** S.E. 0.007	**0.001** S.E. 0.001	**0–0.008**	**0.014–0.023**	**0.001–0.009**	**<0.005**
	AU427	**0.017** S.E. 000.7	**0.001** S.E. 0.001	**0–0.007**	**0.015–0.023**	**0.001–0.008**	**<0.007**
2	**AS167**	**0.298** S.E. 0.056	**0.054** S.E.0.010	**0.3–0.318**	**0.009–0.061**	**0.355–0.374**	**<0.014**
	AH694	**0.303** S.E. 0.052	**0.013** S.E. 0.005	**0.295–0.303**	**0.006–0.012**	**0.312–0.320**	**<0.003**
	AG981	**0.310** S.E.0.053	**0.006** S.E. 0.005	**0.292–0.323**	**0.014–0.041**	**0.297–0.328**	**<0.013**
3	**AH960**	**0.007** S.E. 0.003	**0.189** S.E. 0.003	**0–0.005**	**0.007–0.012**	**0.187–0.19**	**<0.003**
	AK952	**0.006** S.E. 0.003	**0.410** S.E. 0.074	**0.002**	**0.008**	**0.414**	**<0.003**
4	**AP511**	**0.291** S.E. 0.052	**0.077** S.E. 0.017	**0–0.011**	**0.291–0.299**	**0.077–0.088**	**<0.005**

a, mean intrasample genetic distance of viral sequences analysed at week 2 and week 170 p.i.

Increase of predicted glycans in monkeys with divergent *env* V1-V2 regions

To evaluate the glycosylation pattern of the viral variants in animal tissues and PBMC of controller monkeys, we sought differences in the number of potential N-linked glycosylation (PNLG) sites of *env* V1-V2 regions using the N-GlycoSite tool (reference strain, accession number JN205735) [28]. There was no difference in the lengths of V1-V2 sequences (P=0.30) and only a moderate variation in the total number of PNLG among compartments as compared to the virus inoculum; however there were no significant trends overall. By contrast, there was a difference in total PNLG changes, when the sequences were grouped according to clade A (7 macaques/134 sequences), clade B (13 macaques/351 sequences) or clade C (3 macaques/15 sequences), with 57 PNLG site changes that occurred in clade A as compared to 21 in clade B and 55 in clade C. More specifically, statistical analysis showed that in clade B a low number of PNLG changes (19 lost/4 acquired) occurred, as compared to clade A (13 lost/44 acquired) and clade C (20 lost/35 acquired), respectively. According to the genetic divergence levels, while the number of PNLG in low divergent sequences remained relatively conserved (clade B), more divergent sequences included in clade A had a gain of PNLG sites as compared to clade B (P=0.0472) (Fig. 3b), whereas in clade C sequences showed both loss and acquisition of new PNLG sites as compared to both clade A (P=0.0012; P=0.0001) and B (P=0.0004; P=0.0001), respectively (Fig. 3a, b). In particular a Lys-to-Asn change at position 158 (K158N) in V2 region (clade A and C), encoding a novel glycosylation site involved in the neutralizing antibody responses, was absent in blood cells early in infection but appeared at week 170 p.i. in both blood and tissue compartments (P=0.0475) (Fig. 3c).

Effects of MHC haplotype on viral diversity

In humans, HLA genotypes HLA-B57 and HLA-B27 are associated with slow HIV disease progression [29]. Likewise alleles Mamu-A01, Mamu-B08, *Mamu-B17* in rhesus macaques, or haplotypes M3 and M6 in cynomolgus monkeys, are associated with superior control of SIV or SHIV infection [15, 17–22]. To evaluate the influence of MHC haplotype on the outcome of long-term infection, animals were arranged into four groups, each containing three animals heterozygous for M1 (AH694, AQ271, AU427), M3 (AS167, AK484, AG981), M4 (AH960, AK952, AU427) or M7 (AU676, AQ882, AP511) haplotype (Table S2). For statistical analysis, monkey AU427 with M1/M4 haplotype was included in both M1 and M4 groups, whereas recombinant haplotypes (monkey AG172) and the only M5 haplotype (monkey AS377) were excluded.

No statistically significant haplotype advantage was observed in terms of provirus copy numbers, CD4 T-cell counts and anti-Env Ab response on long-term control of viral infection, likely due to the small number of animals (data not shown).

In contrast, sequence diversity differed significantly between either MHC haplotype or within the same group. As shown in Table 3, the mean genetic distances among variants with $SHIV_{SF162P4cy}$ inoculum ranged from 0.023 in M7 to 0.183 in M3 haplotype, and within variants of the same group

from 0.040 in M7 to 0.183 in M3 haplotype. This difference was substantial for M3, with about 8, 5.5 and 2.6 times the distance from $SHIV_{SF162P4cy}$ and nearly 4.5, 3 and 1.6 times the distance within the same haplotype from M7, M4 and M1. After all, according to the diversity level of *env* V1-V2 variants, two out of three M3 animals were included in the second group of genetic diversity displaying higher levels of viral diversity (see Table 2)

Most of the sequences of M3 macaques segregated to clade A. With the exception of the epididymis sequences of monkey AK484 and week 170 PBMC of animal AH960 and AK952 located on clade C, all of the other sequences were included on clade B. Although the observation regarding changes that emerged after infection was not statistically significant due to the small number of animals, each group of animals harboured specific amino acid changes.

Mutation K158N was observed almost exclusively in tissue variants of M3 animals. However only a trend toward statistical significance was noted for the presence of this mutation in M3 as compared to M1 and M4 haplotypes (P=0.0518) (data not shown). Conversely, a change at residue R164 in V1 was recurring in 90 % of monkey tissues, with the exception of the week 170 PBMC of monkeys with M3 and M4 haplotypes. Tissues of M1, M3 and M4 animals, as well as week 170 PBMC of M4- and M7-positive monkeys, carried a mutation at amino acid 190 in V2 with the number of changes per sequence significantly greater in tissues of M3- as compared to M7-positive monkeys (P=0.0482) (data not shown). Thus, at least for K158 and K190, the M3 haplotype was associated with a signature in tissues when compared to the other haplotypes. Considering that M1 and M3 haplotypes share the same class IA alleles, it is possible that in animals bearing M3 haplotype, class I B alleles can influence virus diversity.

It was of interest to analyse whether the total number of PNLG sites in variants of different haplotypes increased over time and/or was related to *env* V1-V2 diversity. Based on diversity levels of V1-V2 regions, analysis of M3- and M4-animal sequences revealed a gain of PNLG sites. In particular, statistical analysis showed that M4 haplotype variants from week 170 PBMC had an increase of PNLG sites as compared to M1 (P=0.0018), M3 (P=0.0002), M7 (P=0.0004) (Fig. 3d). These data suggest that haplotype-dependent mechanisms may be involved in the generation of major V1-V2 viral variants during chronic infection.

DISCUSSION

In HIV controllers the proviral population is extremely homogeneous with no divergence over time and the majority of PBMC-associated proviral sequences represent ancestral variants without sequence replenishment from viruses in plasma [13, 30, 31]. Recently it has been demonstrated that PBMC proviral reservoir reseeding in HIV controllers can be possible from sanctuary tissue sites [13]. However, studies are needed to determine whether HIV-1 is broadly distributed or compartmentalized across tissues. Obtaining human samples for testing is difficult, however the macaque model of infection facilitates extensive tissue collection and detailed analyses of viral populations in blood and tissues.

In this study of long-term controller monkeys, we generated phylogenetic evidence for high proviral genetic diversity at the *env* V1-V2 regions in PBMC and lymphoid, gastrointestinal and genital compartments over a period of approximately 3 years. Two distinct patterns of intra-host viral diversity were observed. Seven animals (AU676, AQ271, AQ882, AG172, AS377, AK484 and AU427) displayed homogeneous proviral populations (<1 % diversity) mainly composed of large clusters of identical sequences. In contrast, the other six animals (AS167, AH694 AG981, AH960, AK952, AP511) showed greater diversity in their quasispecies (>3 %), comprising both clusters of identical and unique sequences, similar to long-term controllers with detectable plasma viremia [13, 30]; these were difficult to directly correlate in animals with viral suppression. However, a examination of the quasispecies diversity revealed a more complex picture. As described in Results, four different groups, according to the diversity levels of *env* V1-V2 variants, were identified. PBMC and tissue-associated V1-V2 sequences in group 1 animals showed no or minimal divergence over time, with the possibility that ancestral variants persisted for more than 3 years of infection, potentially resulting from clonal expansion of CD4+ T lymphocytes as the substantial reservoir. In contrast, a different scenario appeared within group 2 macaques. In these animals, the levels of *env* V1-V2 diversity detected in PBMC at week 2 p.i. and in tissues compartments, were much higher than those estimated for proviral PBMC at week 170 p.i. This result is consistent with an infection of blood cells during acute phase with variants highly divergent from challenge virus, followed by wide dissemination and compartmentalization in anatomical sites. Conversely, the low genetic divergence of week 170 PBMC variants, can be explained by a re-seeding of the PBMC proviral reservoir from compartments, for example B cell follicles or brain, that were not investigated in this study. In sharp contrast, in group 3 monkeys, *env* V1-V2 proviral populations in tissues at necropsy were genetically distant from those in PBMC at the same time point, suggesting the existence of ongoing viral replication in other anatomical compartments as previously described [32].

Examination of sequence polymorphisms highlighted that within individual hosts over time and in relation to the virus inoculum, several common changes led to an ancestral sequence, in this case $SHIV_{SF162P3}$, suggesting that the virus recovers ancestral features upon transmission to the new host. Some changes arose faster than others, with the most rapid mutations arising within structurally conserved residues, such as that at codon 164 that lies within the V2 loop, and which is known to harbour neutralizing antibody epitopes. Another change leading to a glycan insertion was found at codon 158 of virus in several animals, suggesting that probably the same selection pressure was driving this

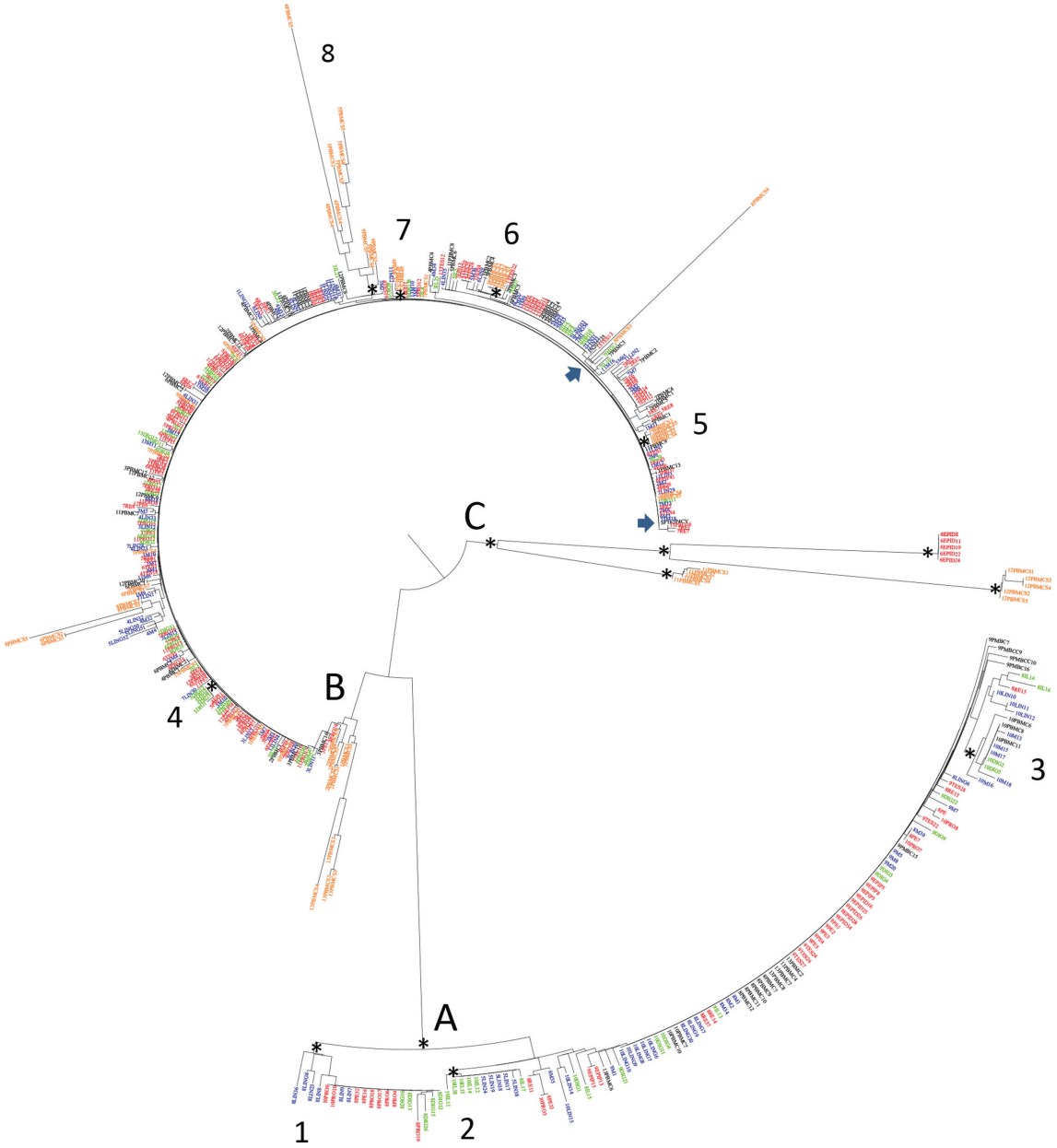

Fig. 2. Phylogenetic analysis of *env* V1-V2 sequences from 13 $SHIV_{SF162P4cy}$ controller monkeys. The neighbour-joining (NJ) tree was rooted by the midpoint rooting. The consensus of the inoculum and the reference sequence (JN205735) were indicated by a blue arrow. The scale bar at the bottom of the tree indicated 0.04 nucleotide substitutions per site. One asterisk (*) along the branches represents significant statistical support for the clusters subtending that branch (bootstrap support >75 %). Main clades were highlighted. The compartments have been indicated by different colours (genital: red; gastro enteric: green; lymphoid: blue; peripheral blood mononuclear cells (PBMC) week 2 p.i.: black; PBMC week 170 p.i.: yellow). Tissues are indicated as follow: LIN, Axillar lymph node; LING, inguinal lymph node; M, spleen; TES, testicle; RE, rectum; EPI, epididymis; PRO, prostate; DIG, jejunum; IL, ileum; PBMC, week 2 p.i., PBMCS, week 170 p.i. Monkey IDs correspond to progressive numbers from 1 to 13 (AU676=1; AQ271=2; AQ882=3; AG172=4; AS377=5; AK484=6; AU427=7; AS167=8; AH694=9; AG981=10; AH960=11; AK952=12; AP511=13).

mutation. Neutralization escape of HIV/SIV has been previously associated with a high number of glycosylation sites and deletions in the external regions of the Env protein [33]. Since N158 has been described in a rhesus macaque as an escape mutant from autologous immune recognition at week 6 p.i. [27], it is suggested that in the monkeys investigated the same driving force has induced the N158 mutation in the early stage of infection, maintaining the phenotype both in blood and tissues over 170 weeks, as proviral DNA has an extremely long half-life and can persist for years.

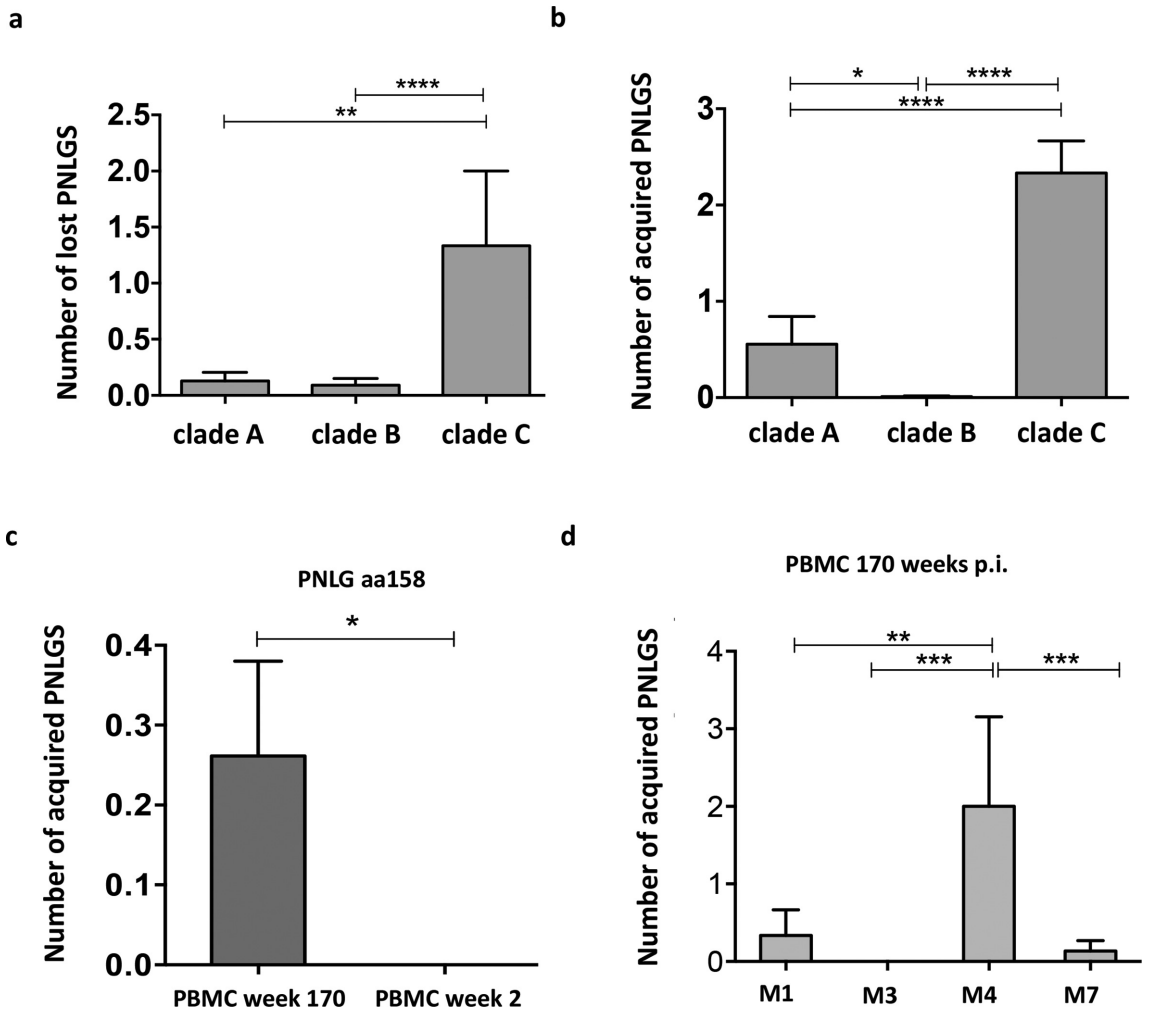

Fig. 3. Numbers of N-linked glycosylation (PNLG) sites within *env* V1-V2 regions by comparison of the average PNLG site numbers in clade A, B and C sequences. Stars indicate statistical difference between clade C with clade A (loss PNLG, *P*=0.0012; gain PNLG *P*=0.0001) and clade B (loss PNLG, *P*=0.0004; gain PNLG, *P*=0.0001) (a, b), and between clade A with clade B (gain PNLG, *P*=0.0472), respectively (b). K158N mutation was absent in proviral DNA from PBMC at week 2 as compared week 170 p.i. (*P*=0.0475) (c). M4 animals had PBMC viral sequences at week 170 p.i. with a statistical different gain of PNLG sites as compared to M1 (*P*=0.0018), M3 (*P*=0.0002), M7 (*P*=0.0004) animals (d).

Quasispecies represent a compromise between evasion from the host immune response and lower ability of replication, thus recovery of an ancestral state may reflect restoration of virus fitness that was lost as a result of immunological escape in the previous host and important for viral replication. In this setting, in long-term controller monkeys infected with the same virus stock, the divergent patterns of genetic diversity of *env* V1-V2 proviral populations may be due to different control mechanisms of viral replication or to more effective $SHIV_{SF162P4cy}$-specific immune responses. Usually escape mutations lead to lower replicative viral fitness and, in the absence of immune pressure, an escape mutant virus 'reverts' to the wild-type phenotype, for example after transmission to MHC I mismatched new hosts [34–36]. Depending on the timing of wild-type virus emergence, it has a higher or lower probability of survival. If it arises early and expands sufficiently during the period of high target cell availability, then it will not be lost during the contraction phase of virus infection [34]. However, the

Table 3. Mean pairwise distances between DNA $SHIV_{SF162P4cy}$ variants in monkeys with different haplotypes

Haplotype	Mean distance (SE)	
	With $SHIV_{SF162P4cy}$	Within the same haplotype
M1	0.070 (S.E., 0.011)	0.108 (S.E., 0.016)
M3	0.183 (S.E., 0.024)	0.183 (S.E., 0.023)
M4	0.033 (S.E., 0.011)	0.059 (S.E., 0.007)
M7	0.023 (S.E., 0.012)	0.040 (S.E., 0.005)

presence of these mutations was not associated with neither an increase in viral load nor disease progression. The selective advantage of these mutations may represent a virus better adapted to the host and, in turn, less pathogenic to cynomolgus monkeys. This is the case for the nonpathogenic molecular clone $SHIV_{SF162Pc}$, which contains mutations associated with the $SHIV_{SF162P3}$*env* 120gp that have been described in earlier work [37]. In fact, it appears that in cynomolgus monkeys the long-term persistent infection with SIV or SHIV relies on host responses, with immune pressure-driven sequence changes leading to the emergence of less pathogenic viruses [37]. In this, the effect of protective MHC genotypes is widely considered a major determinant of viral control and has been shown to significantly influence the outcome of HIV/SIV infection in their respective hosts [38–40].

Data presented from MHC heterozygous M3 macaques indicated that *env* V1-V2 proviral DNA during the early and the long-term phase of infection harboured similar sequences, showing the same high levels of viral diversity and a signature in tissues. Since the M1 and M3 haplotypes share MHC class I A alleles, it is possible that in M1/M3 animals the different and more numerous MHC class I B alleles encoded on each haplotype drive virus diversity. In these macaques as the virus diversified, convergent changes occurred in specific residues within V1-V2 region from independent animals, such as that at the predicted glycosylation site at position 158. We did not observe any particular MHC haplotypes associated with the long-term control of infection, however, we determined a remarkably similar pattern of intra-host viral diversity within heterozygous M3 haplotype. This has also been described by others in M3 haplotype animals [40] that identified M3-restricted CD8+ T cell epitopes selecting for high-frequency mutation in chronic infection [18, 39, 40]. This suggested that in heterozygous animals bearing the same MHC haplotype and infected with the same virus, viral diversity could follow a similar pattern in animals with similar outcome. Our results reveal that in a setting of natural suppression of viral replication, controllers do not form a homogeneous group, because of high genetic variation in the proviral compartment.

METHODS

Animals and infections

Male adult cynomolgus macaques imported from Mauritius were housed according to the European guidelines for non-human primate care (EEC, Directive No. 86–609, 24 November 1986). The study protocol was approved by the ethics committee of the Istituto Superiore di Sanità. Animals were clinically examined under ketamine hydrochloride anaesthesia (10 mg kg^{-1}). Macaques used in this study were part of different experimental protocols as naıve or control animals [20, 21]. Animals were inoculated intrarectally with the same $SHIV_{SF162P4cy}$ virus stock as previously described [20].

Sample collection, plasma viral load and proviral DNA measurement

Blood was collected throughout the infection and at the time of euthanasia. Tissue was collected, cut into fragments and stored at −80 °C. DNA was extracted from five distinct fragments of each tissue using DNeasy Blood and Tissue Kit (QIAGEN, Italy). To avoid sample cross-contamination, nucleic acid extractions were performed using cleaning precautions and separate storage of templates and reagents. Plasma viremia was detected as previously described by using a 'one step' quantitative reverse transcriptase (RT)-PCR (TaqMan) assay with a threshold limit for detection of 50 RNA eq ml^{-1} [20]. To determine proviral load, DNA was extracted from 400 µl of whole citrated blood by using the QIamp DNA Blood Mini Kit (QIAGEN, Italy). Tissues from axillary and inguinal lymph nodes, spleen, jejunum, ileum, epididymis, testis, prostate, penis and rectum were processed and total DNA was extracted from five different fragments of tissue using the DNeasy Blood and Tissue Kit according to the manufacturer's instructions. Proviral DNA copies were quantitated using the TaqMan real-time PCR with a threshold limit of detection of 1 copy μg^{-1} DNA. SIVmac251 plasmids were used as standards to calculate SIV DNA copy numbers. Probe and primers specifically amplifying a region of 71 bp within the *gag* sequence of SIVmac251 (GI:334657) were designed and thermal cycling conditions were used as previously described [20, 21]. Samples were analysed in duplicate and positive and negative controls were used to rule out sample contamination.

PCR AND SEQUENCING

Viral DNA levels in tissues of some animals were extremely low. To ensure that the V1-V2 *env* regions were representative of the virus populations present in each compartment, all specimens were subjected to DNA limiting dilution (<1 template copy per reaction) and endpoint concentrations of template DNA measured such that nested PCR resulted in <50 % of positive reactions [41].

The first round of PCR was performed in 25 µl of (High Fidelity DNA polymerase PCR master mix Invitrogen). First round primers: (HXB2 position 6203–6222)-RP27 (5′GAA AGA GCA GAA GAC AGT GG) 7523–7542 LD2R (5′AAT GTA TGC CCC TCC CAT CA), second-round primers: (HXB2 position 6592–6606) LD4 (5′CCC CAC TCT GTG TTA CTC) 7175–7158 LD9 (5′AAA TGC TTT CCC CGG TCC) amplifying a 583 bp fragment encompassing the V1-V2 *env* region of $SHIV_{SF162P4cy}$.

First round PCR amplification conditions: 94 °C for 12 m and 35 cycles of 94 °C for 90 s, 60 °C for 90 s and 72 °C for 120 s, ending with an extension step at 72 °C for 10 m. A 5 µl aliquot of first-round PCR product was used in the second round in 25 µl of PCR master mix for 40 cycles. PCR products were analysed by gel electrophoresis, purified and directly sequenced.

For sequencing of the $SHIV_{SF162P4cy}$, VI-V2 *env* region was amplified from 500 µl of inoculum using a QIAmp viral RNA Kit (Qiagen) followed by RT with Superscript III RT (Invitrogen) according to the manufacturer's instructions with reverse primer 5′ TTG GCC TCA CTG ATA CCC CT. To ensure amplification of a single viral envelope sequence representative of the virus populations, synthesized cDNA was titrated to a single copy, where PCR-positive wells constitute about 30 % of the reactions. At this dilution, most positive wells would contain amplicons derived from a single cDNA molecule. For nested PCR: first-round primers: (HXB2 position 6203–6222-RP27) (5′GAA AGA GCA GAA GAC AGT GG) 7523–7542 LD2R (5′AAT GTA TGC CCC TCC CAT CA), second-round primers: (HXB2 position 6592–6606 LD4) (5′CCC CAC TCT GTG TTA CTC) 7175–7158 LD9 (5′AAA TGC TTT CCC CGG TCC). First-round PCR amplification conditions: 95 °C for 12 m and 35 cycles of 95 °C for 90 s, 60 °C for 90 s and 72 °C for 120 s and ending with an extension step at 72 °C for 10 m. A 5 µl aliquot of first-round PCR product was used in the second round in 25 µl of PCR master mix for 40 cycles. Thirty two *env* sequences were derived from the challenge virus stock and directly sequenced. Chromatograms with dual or ambiguous peaks were excluded from the analysis. The mean genetic distance within the V1-V2 *env* regions of the challenge virus was 0.002 %, which is consistent with the *env* diversity estimated for $SHIV_{SF162}$ [42, 43].

Microsatellite analysis and allele-specific PCR

MHC class I*A* and I*B* and class II haplotypes were determined by microsatellite PCR with resolution of recombinant class I*B* haplotypes by allele-specific PCR as previously described [17].

Neutralization assay, anti-Env binding antibodies and flow cytometry

Plasma nAbs were assessed using a viral infectivity assay based on TZM-bl cells infected with the $SHIV_{SF162P4cy}$ [20, 21]. Percent neutralization was calculated relative to the negative control infection, containing pre-challenge plasma of the same monkey. Neutralizing Ab titres were estimated as the reciprocal plasma dilution resulting in 50 % inhibition of infection (ID50).

Enzyme-linked immunosorbent assay (ELISA) plates were coated with an oligomeric, SF162 strain-derived Env protein and the assay performed as previously reported [20, 21]. Env-specific IgG bAb titres were calculated as the reciprocal plasma dilution giving optical density (OD) readings three standard deviations (SD) above negative control, normal MCM plasma samples [20, 21].

PBMC were stained with anti-CD4 and anti-CD8 mAb (Becton-Dickinson, USA), and analysed with a FACScan cytometer as described previously [20, 21].

Phylogenetic analysis

V1-V2 *env* sequences collected from 13 monkeys and the $SHIV_{SF162P4cy}$ inoculum were used for phylogenetic analysis. SHIV(sf162p4) GenBank: JN205735 was used as reference genome. Multiple sequence alignments were obtained with Bioedit v 7.[44] followed by manual editing. The NJ phylogenetic tree was generated with the LogDet model, calculations were performed with PAUP* software version 4.0, according to Swofford and Sullivan (Swofford and Sullivan)-[45]. Statistical support for specific clades was obtained by bootstrapping values (1000 replicates) and bootstrap values>75 % were considered statistically supported.

The software MEGA 6 [46] allowed the calculation of the genetic distances from the parental clone and between the different tissue compartments, by using the LogDet model. The evaluation of the mutational pattern of the sequences with respect to the sequence of the inoculum $SHIV_{SF162P4cy}$ was performed on the alignment obtained with Bioedit. The evaluation of the N-linked glycosylation sites' variation was performed by using the server N-GlycoSite under the HIV sequence database by using the obtained alignment [28].

Migration analysis was conducted with the Mac Clade v. 4 program to test the viral in/out gene flows among the distinct compartments (tissues) using a modified version of the Slatkin and Maddison test as already described [47]. Specifically, gene flow analysis was performed classifying the V1-V2 *env* sequences into different groups, based on the specific compartments from which the sequences were sampled. A one-character data matrix was obtained by assigning to each taxon in the tree a one-letter code indicating its group of origin. The putative origin of each ancestral sequence (i.e. internal node) in the tree was inferred by finding the most parsimonious reconstruction of the ancestral character. The final tree-length, i.e. the number of observed migrations in the genealogy, was compared to the tree-length distribution of 10 000 trees, after random joining-splitting.

The numbers of PNLGs in the V1-V2 sequence were determined by using the tool at https://www.hiv.lanl.gov/content/sequence/GLYCOSITE/glycosite.html.

Statistical methods

Statistical tests were performed using Graph Pad Prism software. The nonparametric Mann–Whitney test was used to determine *P*-values when comparing two groups that were not normally distributed. After controlling for normal distribution, groups of animals were segregated on the basis of MHC haplotypes and mean viral load, proviral DNA and CD4+ T cell count were compared between groups by two-way ANOVA using Tukey's multiple comparison test. To determine the relationship among virological, immunological and MHC haplotypes during infection, the analysis of variance, in which all haplotypes were adjusted for each other in a multivariate model, was applied. All statistical tests were carried out at a two-sided 5 % significance level.

Funding information
This work was supported by grants from Gilead Science (41504D1018) and in part by the NIHR Centre for Research in Health Protection at the Health Protection Agency (UK; now Public Health England).

Acknowledgements
We thank P. Pupino Carbonelli for veterinarian assistance; M. Baesso, F. Incitti, F. Costa, A. Baldi, F. Grasso, A. Gallinaro, for technical assistance; S. De Menna, S. Tobelli and F. Fedeli for administrative support.

Conflicts of interest
The authors declare that there are no conflicts of interest.

Ethical statement
Animals were housed according to the European guidelines for nonhuman primate care (EEC, Directive No. 86-609, 24 November 1986). The study protocol was approved by the ethics committee of the Istituto Superiore di Sanità.

References
1. **Policicchio BB, Pandrea I, Apetrei C.** Animal models for HIV cure research. *Front Immunol* 2016;7:12.
2. **Polacino P, Larsen K, Galmin L, Suschak J, Kraft Z *et al.*** Differential pathogenicity of SHIV infection in pig-tailed and rhesus macaques. *J Med Primatol* 2008;37 Suppl 2:13–23.
3. **Borsetti A, Baroncelli S, Maggiorella MT, Bellino S, Moretti S *et al.*** Viral outcome of simian-human immunodeficiency virus SHIV-89.6P adapted to cynomolgus monkeys. *Arch Virol* 2008;153: 463–472.
4. **Veazey RS, Ling B.** Short Communication: comparative susceptibility of rhesus macaques of indian and chinese origin to vaginal simian-human immunodeficiency virus transmission as models for HIV prevention research. *AIDS Res Hum Retroviruses* 2017;33: 1199–1201.
5. **Palesch D, Bosinger SE, Tharp GK, Vanderford TH, Paiardini M *et al.*** Sooty mangabey genome sequence provides insight into AIDS resistance in a natural SIV host. *Nature* 2018;553:77–81.
6. **Huot N, Bosinger SE, Paiardini M, Reeves RK, Müller-Trutwin M.** Lymph node cellular and viral dynamics in natural hosts and impact for HIV cure strategies. *Front Immunol* 2018;9:780.
7. **Tan RC, Harouse JM, Gettie A, Cheng-Mayer C.** In vivo adaptation of SHIV(SF162): chimeric virus expressing a NSI, CCR5-specific envelope protein. *J Med Primatol* 1999;28:164–168.
8. **del Prete GQ, Lifson JD, Keele BF.** Nonhuman primate models for the evaluation of HIV-1 preventive vaccine strategies: model parameter considerations and consequences. *Curr Opin HIV AIDS* 2016;11:546–554.
9. **O'Connell KA, Brennan TP, Bailey JR, Ray SC, Siliciano RF *et al.*** Control of HIV-1 in elite suppressors despite ongoing replication and evolution in plasma virus. *J Virol* 2010;84:7018–7028.
10. **Rouzioux C, Avettand-Fenoël V.** Total HIV DNA: a global marker of HIV persistence. *Retrovirology* 2018;15:30.
11. **Salemi M, Rife B.** Phylogenetics and phyloanatomy of HIV/SIV intra-host compartments and reservoirs: the key role of the central nervous system. *Curr HIV Res* 2016;14:110–120.
12. **Huang SH, Ren Y, Thomas AS, Chan D, Mueller S *et al.*** Latent HIV reservoirs exhibit inherent resistance to elimination by CD8+ T cells. *J Clin Invest* 2018;128:876–889.
13. **de Azevedo SSD, Caetano DG, Côrtes FH, Teixeira SLM, dos Santos Silva K *et al.*** Highly divergent patterns of genetic diversity and evolution in proviral quasispecies from HIV controllers. *Retrovirology* 2017;14:29.
14. **Goulder PJ, Watkins DI.** Impact of MHC class I diversity on immune control of immunodeficiency virus replication. *Nat Rev Immunol* 2008;8:619–630.
15. **Nomura T, Matano T.** Association of MHC-I genotypes with disease progression in HIV/SIV infections. *Front Microbiol* 2012;3:234.
16. **Silver ZA, Watkins DI.** The role of MHC class I gene products in SIV infection of macaques. *Immunogenetics* 2017;69:511–519.
17. **Mee ET, Berry N, Ham C, Aubertin A, Lines J *et al.*** Mhc haplotype M3 is associated with early control of SHIVsbg infection in Mauritian cynomolgus macaques. *Tissue Antigens* 2010;76:223–229.
18. **O'Connor SL, Lhost JJ, Becker EA, Detmer AM, Johnson RC *et al.*** MHC heterozygote advantage in simian immunodeficiency virus-infected Mauritian cynomolgus macaques. *Sci Transl Med* 2010;2: 22ra18.
19. **Aarnink A, Dereuddre-Bosquet N, Vaslin B, Le Grand R, Winterton P *et al.*** Influence of the MHC genotype on the progression of experimental SIV infection in the Mauritian cynomolgus macaque. *Immunogenetics* 2011;63:267–274.
20. **Borsetti A, Maggiorella MT, Sernicola L, Bellino S, Ferrantelli F *et al.*** Influence of MHC class I and II haplotypes on the experimental infection of Mauritian cynomolgus macaques with SHIVSF162P4cy. *Tissue Antigens* 2012;80:36–45.
21. **Borsetti A, Ferrantelli F, Maggiorella MT, Sernicola L, Bellino S *et al.*** Effect of MHC haplotype on immune response upon experimental $SHIV_{SF162P4cy}$ infection of Mauritian cynomolgus macaques. *PLoS One* 2014;9:e93235.
22. **Seki S, Nomura T, Nishizawa M, Yamamoto H, Ishii H *et al.*** In vivo virulence of MHC-adapted AIDS virus serially-passaged through MHC-mismatched hosts. *PLoS Pathog* 2017;13:e1006638.
23. **Kline C, Ndjomou J, Franks T, Kiser R, Coalter V *et al.*** Persistence of viral reservoirs in multiple tissues after antiretroviral therapy suppression in a macaque RT-SHIV model. *PLoS One* 2013;8: e84275.
24. **Kearney MF, Anderson EM, Coomer C, Smith L, Shao W *et al.*** Well-mixed plasma and tissue viral populations in RT-SHIV-infected macaques implies a lack of viral replication in the tissues during antiretroviral therapy. *Retrovirology* 2015;12:93.
25. **Matusali G, Dereuddre-Bosquet N, Le Tortorec A, Moreau M, Satie AP *et al.*** Detection of simian immunodeficiency Virus in Semen, Urethra, and Male Reproductive Organs during Efficient Highly Active Antiretroviral Therapy. *J Virol* 2015;89:5772–5787.
26. **Feder AF, Kline C, Polacino P, Cottrell M, Kashuba ADM *et al.*** A spatio-temporal assessment of simian/human immunodeficiency virus (SHIV) evolution reveals a highly dynamic process within the host. *PLoS Pathog* 2017;13:e1006358.
27. **Balfe P, Shapiro S, Hsu M, Buckner C, Harouse JM *et al.*** Expansion of quasispecies diversity but no evidence for adaptive evolution of SHIV during rapid serial transfers among seronegative macaques. *Virology* 2004;318:267–279.
28. **Zhang M, Gaschen B, Blay W, Foley B, Haigwood N *et al.*** Tracking global patterns of N-linked glycosylation site variation in highly variable viral glycoproteins: HIV, SIV, and HCV envelopes and influenza hemagglutinin. *Glycobiology* 2004;14:1229–1246.
29. **Mclaren PJ, Carrington M.** The impact of host genetic variation on infection with HIV-1. *Nat Immunol* 2015;16:577–583.
30. **Bello G, Casado C, Sandonis V, Alvaro-Cifuentes T, dos Santos CA *et al.*** Plasma viral load threshold for sustaining intrahost HIV type 1 evolution. *AIDS Res Hum Retroviruses* 2007;23:1242–1250.
31. **Boritz EA, Darko S, Swaszek L, Wolf G, Wells D *et al.*** Multiple Origins of Virus Persistence during Natural Control of HIV Infection. *Cell* 2016;166:1004–1015.
32. **Fukazawa Y, Lum R, Okoye AA, Park H, Matsuda K *et al.*** A B cell follicle sanctuary permits persistent productive simian immunodeficiency virus infection in elite controllers. *Nat Med* 2015;21: 132–139.
33. **Seki S, Matano T.** CTL Escape and Viral Fitness in HIV/SIV Infection. *Front Microbiol* 2011;2:267.
34. **Kawashima Y, Pfafferott K, Frater J, Matthews P, Payne R *et al.*** Adaptation of HIV-1 to human leukocyte antigen class I. *Nature* 2009;458:641–645.
35. **Loh L, Reece JC, Fernandez CS, Alcantara S, Center R *et al.*** Complexity of the inoculum determines the rate of reversion of SIV Gag CD8 T cell mutant virus and outcome of infection. *PLoS Pathog* 2009;5:e1000378.
36. **Huang KH, Goedhals D, Carlson JM, Brockman MA, Mishra S *et al.*** Progression to AIDS in South Africa is associated with both reverting and compensatory viral mutations. *PLoS One* 2011;6: e19018.

37. Hsu M, Harouse JM, Gettie A, Buckner C, Blanchard J *et al.* Increased mucosal transmission but not enhanced pathogenicity of the CCR5-tropic, simian AIDS-inducing simian/human immunodeficiency virus SHIV(SF162P3) maps to envelope gp120. *J Virol* 2003;77:989–998.
38. Deeks SG, Walker BD. Human immunodeficiency virus controllers: mechanisms of durable virus control in the absence of antiretroviral therapy. *Immunity* 2007;27:406–416.
39. Sutton MS, Ellis-Connell A, Moriarty RV, Balgeman AJ, Gellerup D *et al.* Acute-phase CD4+ T cell responses targeting invariant viral regions are associated with control of live-attenuated simian immunodeficiency virus. *J Virol* 2018;87:9353–9364.
40. Cain BT, Pham NH, Budde ML, Greene JM, Weinfurter JT *et al.* T cell response specificity and magnitude against SIVmac239 are not concordant in major histocompatibility complex-matched animals. *Retrovirology* 2013;10:116.
41. Rodrigo AG, Goracke PC, Rowhanian K, Mullins JI. Quantitation of target molecules from polymerase chain reaction-based limiting dilution assays. *AIDS Res Hum Retroviruses* 1997;13:737–742.
42. Tsai L, Trunova N, Gettie A, Mohri H, Bohm R *et al.* Efficient repeated low-dose intravaginal infection with X4 and R5 SHIVs in rhesus macaque: implications for HIV-1 transmission in humans. *Virology* 2007;362:207–216.
43. Varela M, Landskron L, Lai RP, McKinley TJ, Bogers WM *et al.* Molecular evolution analysis of the human immunodeficiency virus type 1 envelope in simian/human immunodeficiency virus-infected macaques: implications for challenge dose selection. *J Virol* 2011;85:10332–10345.
44. Hall TA. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucl Acid Symp Ser* 1999;41:95–98.
45. Swofford D, Sullivan J. Phylogeny inference based on parsimony and other methods with PAUP*. In: Salemi M and Vandamme AM (editors). *The phylogenetic handbook-A practical approach to DNA and protein phylogeny*. New York: Cambridge University Press; 2003. pp. 160–206.
46. Tamura K, Stecher G, Peterson D, Filipski A, Kumar S. MEGA6: molecular evolutionary genetics analysis version 6.0. *Mol Biol Evol* 2013;30:2725–2729.
47. Slatkin M, Maddison WP. A cladistic measure of gene flow inferred from the phylogenies of alleles. *Genetics* 1989;123:603–613.