UvA-DARE (Digital Academic Repository)

Within-host HIV-1 evolution in relation to viral coreceptor use and host environment
Edo Matas, D.

Link to publication

Citation for published version (APA):
Edo Matas, D. (2011). Within-host HIV-1 evolution in relation to viral coreceptor use and host environment

General rights
It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations
If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: http://uba.uva.nl/en/contact, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.

UvA-DARE is a service provided by the library of the University of Amsterdam (http://dare.uva.nl)

Download date: 18 Dec 2018
Emergence of CXCR4-using HIV-1 variants in an individual infected by a donor with exclusively CCR5-using variants at all stages of his HIV-1 infection
ABSTRACT

CCR5-using (R5) HIV-1 variants predominate early in infection. In 50% of HIV-1 subtype B infections, CXCR4-using viruses emerge prior to AIDS diagnosis, preceding an accelerated disease progression. Determinants for the emergence of CXCR4-using virus variants remain largely unresolved.

Here we analyzed replication-competent clonal HIV-1 variants and viral RNA from serum longitudinally sampled from a donor who had exclusively R5 variants during his entire disease course and from his recipient in whom CXCR4-using variants emerged.

Detailed phylogenetic analysis revealed evolution of CXCR4-using variants from the transmitted R5 variants in the recipient. Over time, R5 variants in the donor optimized coreceptor use, as reflected by increasing resistance to inhibition by anti-CCR5 antibodies, and increased the number of PNGS and V3 charge of their envelopes. Emergence of CXCR4-using variants in the recipient was preceded by a selective sweep and an increase in the envelope number of PNGS and V3 charge of the R5 viral population, which at the time of establishing infection had an efficiency of CCR5 use and envelope molecular properties similar to the late stage R5 variants from the donor. We conclude that phenotype and envelope molecular characteristics of late stage viruses can be preserved upon transmission. The high efficiency in CCR5 use and the more positively charged V3 envelope region of the recently transmitted R5 variants may have been instrumental to the emergence of CXCR4-using viruses in the specific immune environment of the recipient. However, the contribution of viral and host factors in this evolutionary process remains to be elucidated.

Diana Edo-Matas\textsuperscript{1,*}, Andrea Rachinger\textsuperscript{1,*}, Laurentia C. Setiawan\textsuperscript{1}, Brigitte Boeser-Nunnink\textsuperscript{1}, Angélique B. van 't Wout\textsuperscript{1,2}, Philippe Lemey\textsuperscript{3} and Hanneke Schuitemaker\textsuperscript{1,2}

\textsuperscript{1}Department of Experimental Immunology, Sanquin Research, Landsteiner Laboratory, part of the Center for Infectious Diseases and Immunity Amsterdam (CINIMA) at the Academic Medical Center of the University of Amsterdam, Amsterdam, The Netherlands; \textsuperscript{2}Present address: Crucell Holland BV, Leiden, The Netherlands; \textsuperscript{3}Rega Institute for Medical Research, K.U.Leuven, Belgium

* authors contributed equally to the study

Submitted for publication
INTRODUCTION

Human Immunodeficiency Virus type I (HIV-1) entry into its target cells is mediated by the interaction of the viral envelope protein with CD4 and a coreceptor on the target cell surface. The main coreceptors used by HIV-1 in vivo are the chemokine receptors CCR5 and CXCR4.

CCR5-using (R5) viruses predominate in the early stages of HIV-1 infection irrespective of the route of transmission and persist throughout the course of the disease. Viruses capable of using CXCR4 emerge in 50% of HIV-1 subtype B infected individuals prior to AIDS diagnosis, preceding a more rapid CD4+ T cell decline and accelerated disease progression. Progression to AIDS in the remaining infected individuals occurs in the presence of solely R5 HIV-1 variants. Throughout the course of infection, the biological properties of the R5 viruses commonly evolve.

Indeed, late stage R5 HIV-1 variants show more rapid replication, higher cytopathicity and a more efficient use of CCR5, the latter reflected by a decreased sensitivity to inhibition by CCR5 antagonists and the ability to use CCR5-CXCR4 chimeric coreceptors. It is currently unknown whether the phenotypic characteristics of late stage R5 variants are preserved upon transmission to a new individual and how they are affected by the immune system in a new host.

The predominance of R5 variants early in infection implies a major role for this viral phenotype in HIV-1 transmission or in initial dissemination and establishment of infection. This is underscored by the fact that individuals who are homozygous for a 32 base pair deletion in the CCR5 gene, a mutation that results in a truncated CCR5 protein that is not expressed on the cell surface, are relatively resistant to HIV-1 infection, supporting that establishment of HIV-1 infection by CXCR4-using variants is an unlikely event. From the few reported HIV-1 infected individuals with a CCR5 Δ32/Δ32 genotype, which has a prevalence of about 1% in the Caucasian population, only CXCR4-using variants could be isolated providing formal proof that CXCR4-using variants can actually be transmitted. Indeed, CXCR4-using variants have been detected in 3-17.2% of recently infected patients, with frequencies varying depending on whether a genotypic or a phenotypic assay was used to determine viral coreceptor use. In those studies, the earliest sampling ranged from 14-183 days after SC, implying that the detection of CXCR4-using variants early in infection in those HIV-1 infected individuals, is either the result of the transmission of both R5 and CXCR4-using variants or, in case that there had been selective transmission of R5 variants, evolution towards a CXCR4-using phenotype must have occurred at least within the first two weeks upon infection.

Several mechanisms, acting at mucosal and non-mucosal level, have been suggested to be responsible for the selective advantage of R5 viruses during HIV-1 transmission and/or initial dissemination, which may contribute to their predominance in the early phase of infection. However, those mechanisms neither provide an explanation for the exclusive presence of R5 variants early in infection in individuals who became infected via needle sharing during injection drug use and in whom CXCR4-using variants may eventually emerge, nor do they explain the absence of CXCR4-using variants in some HIV-1 infected individuals in their entire disease course.
Here, we studied HIV-1 evolution in a donor-recipient pair in which the donor harbored only R5 variants during his entire course of infection, whereas in the recipient CXCR4-using variants eventually emerged. We compared evolutionary rates and dN/dS ratios based on gp120 envelope sequences, efficiency of CCR5 use, and molecular characteristics of the gp120 envelope gene of longitudinally obtained R5 variants, and in particular from R5 variants obtained prior and after transmission, from donor and recipient. With the detailed study of the evolutionary history of this transmission chain in relation to the variation of viral phenotype characteristics and envelope molecular properties, we attempt to define the role of viral factors and host environment in the emergence of CXCR4-using variants.

**MATERIAL AND METHODS**

**Study subjects and samples**

Longitudinal cryopreserved peripheral blood mononuclear cells (PBMC) samples from two homosexual men (subject 1 (cohort ID:19858) and subject 2 (cohort ID:19308)) who participated in the Amsterdam Cohort Studies on HIV-1 infection and AIDS (ACS, http://www.amsterdamcohortstudies.org), with an imputed or documented seroconversion (SC) date, respectively, were used for this study. Subject 1 remained therapy-naïve while subject 2 received changing antiretroviral therapy (AZT+3TC; ddi+d4T+Indinavir) starting at the third-last analyzed time point. Time points studied are shown in Table 1. Longitudinal data on CD4 counts and plasma viral load are shown in Figure 1.

The Amsterdam Cohort Studies are conducted in accordance with the ethical principles set out in the declaration of Helsinki and were approved by the Academic Medical Center institutional medical ethics committee. Written informed consent was obtained from all cohort participants.

**Isolation of replication-competent clonal HIV-1 variants from PBMC**

Replication-competent clonal HIV-1 variants were isolated by co-cultivation of serial dilutions of cryopreserved patient PBMC with PHA-stimulated PBMC from HIV-1 seronegative donors as described previously. From each time point, a median of 6 clonal HIV-1 variants (range, 4-15) was subjected to DNA isolation, PCR and sequencing (Table 1).

**DNA isolation, PCR and env gp120 C1-V5 sequencing from replication-competent clonal HIV-1 variants from PBMC**

Total DNA was isolated from 0.5-1 x 10^6 HIV-1 infected cells using a modification of the L6 isolation method. Gp120 env PCR amplification was performed with one outer PCR with primers TB3 forward (5’-GGCCTTATAGACATAGTTAGCC-3’) and OFM19 reverse (5’-GCACCTAAGGGCAACGCTTTATTAGGCTTA-3’) and a nested PCR with primers env1aTOPO forward (5’-ACCACATTACCACGAGGCTTTATTAGGCTTA-3’) and envN reverse (5’-CTGCTAATCAGGGAAGTAGCCTTGTGT-3’) using the Expand High Fidelity Taq Polymerase kit (Roche) and the following amplification cycles: 2
HIV-1 evolution in a donor-recipient pair

**Table 1. Study subjects, time points, number of sequences analyzed and number of clonal HIV-1 variants tested for sensitivity to anti-CCR5 monoclonal antibodies.**

<table>
<thead>
<tr>
<th>Cohort ID</th>
<th>Sample ID</th>
<th>Subject</th>
<th>Sampling time after SC or imputed SC date* (months)</th>
<th>Time to X4 emergence a (months)</th>
<th>Time since SC of the donor (months)</th>
<th>Serum Nr. of sequences</th>
<th>Clon.a.l HIV-1 variants Nr. of sequences</th>
<th>Sensitivity to anti-CCR5 antibodies b</th>
</tr>
</thead>
<tbody>
<tr>
<td>19858</td>
<td>1 (donor)</td>
<td>40*</td>
<td>101</td>
<td>n.a.</td>
<td>n.a.</td>
<td>6</td>
<td>0</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td></td>
<td>66*</td>
<td></td>
<td>13</td>
<td>0</td>
<td>n.a.</td>
<td>n.a.</td>
<td>n.a.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>90*</td>
<td></td>
<td>13</td>
<td>0</td>
<td>n.a.</td>
<td>n.a.</td>
<td>n.a.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>111*</td>
<td></td>
<td>n.a.</td>
<td>n.a.</td>
<td>4</td>
<td>6</td>
<td>4</td>
</tr>
<tr>
<td>19308</td>
<td>2 (recipient)</td>
<td>3</td>
<td>-36</td>
<td>105</td>
<td>13</td>
<td>0</td>
<td>n.a.</td>
<td>n.a.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>8</td>
<td>-31</td>
<td>105</td>
<td>13</td>
<td>0</td>
<td>n.a.</td>
<td>n.a.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>26</td>
<td>-13</td>
<td>124</td>
<td>13</td>
<td>0</td>
<td>n.a.</td>
<td>n.a.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>28.8</td>
<td>-10.2</td>
<td>126.6</td>
<td>n.a.</td>
<td>4</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td></td>
<td>29.4</td>
<td>-9.6</td>
<td>127.2</td>
<td>10</td>
<td>2</td>
<td>n.a.</td>
<td>n.a.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>37</td>
<td>-2</td>
<td>134</td>
<td>12</td>
<td>2</td>
<td>n.a.</td>
<td>n.a.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>41</td>
<td>2</td>
<td>139</td>
<td>9</td>
<td>5</td>
<td>8</td>
<td>7</td>
</tr>
<tr>
<td></td>
<td></td>
<td>50</td>
<td>11</td>
<td>148</td>
<td>5</td>
<td>7</td>
<td>n.a.</td>
<td>n.a.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>53</td>
<td>14</td>
<td>151</td>
<td>4</td>
<td>12</td>
<td>4</td>
<td>11</td>
</tr>
<tr>
<td></td>
<td></td>
<td>59</td>
<td>20</td>
<td>156</td>
<td>11</td>
<td>4</td>
<td>n.a.</td>
<td>n.a.</td>
</tr>
</tbody>
</table>

* Time to emergence of CXCR4-using viruses (X4) (estimated date between last negative and first positive MT-2 test).

b Number of clonal HIV-1 variants tested for sensitivity to anti-CCR5 monoclonal antibodies.

R5: Predicted and/or in vitro tested CCR5 use.

X4: Predicted and/or in vitro tested CXCR4 use.

n.a.: Non-applicable.

min 30s 94°C, 9 cycles of 15s 94°C, 45s 50°C, 6min 68°C, 30 cycles of 15s 94°C, 45s 53°C, 6min 68°C, followed by 10min at 68°C and cooling to 4°C. PCR products were purified using ExoSAP-IT (USB) according to the manufacturer’s protocol. Sequencing conditions were 5’ at 94°C, 30 cycles of 15'' at 94°C, 10'' at 50°C, 2’ at 60°C and 10' at 60°C. Sequencing of gp120 C1-V5 region corresponding to HXB2 nucleotide positions 6465 to 7636 was performed using BigDye Terminator v1.1 Cycle Sequencing kit (ABI Prism, Applied Biosystems) according to the manufacturer’s protocol using the primers Seq1 forward (5’-TACATAATGTTTGGGACCACATGCC-3’), Seq4 reverse (5’-CTTGTATTGTTGTTGGGTCTTGTTGAC-3’), Seq5 forward (5’-GTCAACTCAACTGCTGTTAATGGC-3’) and Seq2 reverse (5’-TCCTTCATATCTCCTCCTCCAGGTC-3’). Sequences were analyzed on the 3130 xl Genetic Analyzer (Applied Biosystems).
RNA isolation from serum, RT-PCR, PCR amplification, molecular cloning of multiple PCR products and sequencing of env gp120 C2-C4

Env gp120 C2-C4 sequences from viral RNA in serum from subject 1 and 2 were available from an earlier study (Rachinger et al, manuscript submitted for publication). HIV-1 RNA isolation from serum samples, cDNA synthesis, molecular cloning of multiple PCR products and sequencing of env gp120 C2-C4 PCR products (549 nucleotides, corresponding to HxB2 envelope nucleotide positions 811-1290) was performed as described previously. From each time point a median of 3 C2-C4 env (range, 1-5) PCR products was generated and cloned. A median of 13 env gp120 C2-C4 env (range, 2-16) sequences was generated per time point (Table 1).

Prediction and determination of coreceptor use

The two study participants were routinely tested at approximately 3-monthly intervals for the presence of replication-competitive CXCR4-using HIV-1 variants in PBMC using the MT-2 assay. CCR5 and CXCR4 coreceptor use of replication-competitive clonal HIV-1 variants at the time points under study was determined by testing the ability of the virus to replicate in MT-2 cells, U87.CD4.CCR5 cells and U87.CD4.CXCR4 cells, as described previously. V3 sequence based prediction tools (Position Specific Scoring Matrix (PSSM) and the geno2pheno [coreceptor] method (FPR=5%) were used to confirm the coreceptor use of the replication-competitive clonal HIV-1 variants and to predict coreceptor use of sequences derived from viral RNA in serum.

Phylogenetic and molecular clock analysis

Nucleotide sequences were aligned using ClustalW (BioEdit v.7.0.9) and edited manually.

A Maximum Likelihood (ML) tree was reconstructed from published and unpublished env gp120 C2-C4 sequences from ACS participants, merged with an international panel of highly related, but not epidemiologically linked sequences (downloaded from the Los Alamos database). The best-fit nucleotide substitution model (TVM+I+G), selected by hierarchical likelihood tests (hLRTs, Modeltest v3.7) was implemented in
HIV-1 evolution in a donor-recipient pair

the heuristic search for the best ML tree applying NNI (nearest-neighbor-interchange) branch-swapping algorithm using PAUP*4.0.70, starting with a Neighbor-Joining (NJ) tree constructed under the Hasegawa-Kishino-Yano (HKY85) model of evolution. The resulting ML tree was rooted with the earliest sequence available and displayed with Dendroscope. Statistical support for nodes was generated using NJ bootstrapping (1000 repeats).

Subsequently, ML and Bayesian MCMC analysis were performed on the alignment of \textit{env} gp120 C1-V5 sequences derived from clonal HIV-1 variants from subjects 1 and 2 (alignment A), and on a set of \textit{env} gp120 C2-C4 sequences derived from viral RNA in serum aligned to the clonal HIV-1 variants but trimmed to the length of the C2-C4 region (alignment B). ML phylogenies were reconstructed using PAUP*, applying the best-fit nucleotide substitution model and the TBR (tree-bisection-reconnection) branch-swapping algorithm. The ML trees were rooted using one sequence of the earliest time point from subject 1. Bayesian phylogenetic reconstruction was performed using Markov Chain Monte Carlo (MCMC) analysis implemented in BEAST v1.5.4.73. BEAST focuses on rooted, time-measured phylogenetic trees with a coalescent prior. We applied a general time-reversible (GTR) model of nucleotide substitution with gamma-distributed rate variation among sites, a lognormal relaxed clock model and a flexible Bayesian skyride tree prior. MCMC chains were run sufficiently long (100 million generations) to ensure stationarity and adequate effective sample sizes as diagnosed using Tracer (http://tree.bio.ed.ac.uk/software/tracer/). Maximum clade credibility (MCC) trees were annotated using TreeAnnotator and visualized with FigTree (http://tree.bio.ed.ac.uk/software/figtree/).

To compare evolutionary rates of viruses between subject 1 and subject 2, Bayesian MCMC analysis was performed on the \textit{env} gp120 C1-V5 sequences from clonal HIV-1 variants for each patient separately. For subject 2, an additional analysis was performed solely on the clonal R5 variants sequences. BEAST MCMC analyses were run applying both a strict and a lognormal relaxed clock model and using the same substitutions models as mentioned above (chain length 200 million); models were compared using Bayes factor testing.

Selective pressure analysis

To quantify the differences in selective pressure between HIV-1 variants from subject 1 and 2, and between the R5 and CXCR4-using variants within subject 2, we estimated synonymous/non-synonymous substitution rate ratios (dN/dS) using codon substitution models implemented in the codeml program from the PAML package. Specifically, we fitted different models that allow dN/dS to vary among lineages. Our comparison included a model that assumes a single dN/dS for lineages from subject 1 and 2 (Model A), a model that specifies a different dN/dS for subject 1 and 2 (Model B) and a model that allows a different dN/dS for subject 1, the subject 2 R5 variants clade and the subject 2 CXCR4-using variants clade (Model C). Analyses were performed on alignment A and nested models were compared using the likelihood ratio test.

To investigate the accumulation of synonymous and non-synonymous substitutions over time, we also estimated branch lengths for the MCC phylogeny of subject 1 and 2 based on alignment A in expected synonymous (E[S]) and nonsynonymous (E[N]) substitutions using a local codon model implemented in HyPhy. We subsequently
plotted E[S] and E[N] root-to-tip divergences using Path-o-gen (http://tree.bio.ed.ac.uk/software/pathogen/).

**Sensitivity of HIV-1 to anti-CCR5 monoclonal antibodies mediated inhibition in a PBMC-based assay**

PBMC from 3 to 4 different healthy seronegative blood donors were isolated by Ficoll-Isopaque density gradient centrifugation and stimulated for 3 days in Iscove’s Modified Dulbecco’s Medium (IMDM, Lonza) supplemented with 10% fetal calf serum (FCS; Hyclone), 1µg/ml phytohemagglutinin (PHA;Welcome), 100U/ml Penicillin and 100 µg/ml Streptomycin (Pen/Strep;Gibco Brl), 5µg/ml Ciproxin (Bayer) in a culture flask at a cell density of 5x10⁶/ml. In a 96-well tissue culture plate, a final inoculum of 35 TCID₅₀ (50% tissue culture infectious dose) was used to infect 1x10⁵ PHA-stimulated PBMC previously incubated for 1h at 37°C with threefold serial dilutions of monoclonal antibody (MAb) RoAb13 or RoAb952 (highest concentration = 30µg/ml)⁸², kindly provided by Dr. Andreas Jeckle (Roche). Cultures were incubated for one week at 37°C and 10% CO₂. A median of 5 clonal HIV-1 variants (range, 2-6) were tested per time point (Table 1); each clonal HIV-1 variant was tested in triplicate. On day 7, a third of the culture was transferred to a new plate containing 1x10⁵ new PHA-stimulated PBMC per well. Virus production in culture supernatants at days 7 and 11 was analyzed using an in-house p24 antigen capture enzyme-linked immunosorbent assay (ELISA)⁸³. The percent inhibition was calculated by determining the reduction in p24 production in the presence of the antibody as compared to the cultures that lacked antibody. 50% inhibitory concentrations (IC₅₀) were determined by linear regression.

**Analysis of envelope molecular properties**

Potential N-linked Glycosylation Sites (PNGS) were identified using N-Glycosite⁸⁴ at the HIV database website (http://www.hiv.lanl.gov/content/sequence/GLYCOSITE/glycosite.html). Charge was calculated by counting all charged amino acid residues per sequence, where R and K were counted as +1, H as +0.293, and D and E as -1.

**Statistical analysis**

Statistical analyses were performed in SPSS 16 software package. Longitudinal changes in sensitivity to anti-CCR5 MAbs Rob13 and RoAb952, in charge and PSSM score of the gp120 V3 region and in number of PNGS in gp120 C1-V5, were assessed using a Kruskal-Wallis test. Differences in sensitivity to anti-CCR5 MAbs Rob13 and RoAb952, in charge and PSSM score of the gp120 V3 region and in number of PNGS in gp120 C1-V5, between two different time points were evaluated using the Mann-Whitney U test.
RESULTS

Phylogenetic analysis of env sequences demonstrates HIV-1 transmission from an HIV-1 infected individual with R5 variants only, to an individual in whom eventually CXCR4-using HIV-1 variants developed

A Maximum likelihood (ML) phylogenetic tree based on env gp120 C2-C4 sequences from 103 participants of the ACS merged with a highly-related but not epidemiologically linked international reference panel of 148 sequences, revealed well-supported clustering of subject 1 and subject 2 derived sequences (data not shown). Moreover, sequences from subject 2 were nested in a larger cluster of sequences that were exclusively derived from subject 1 (both clusters supported by bootstrap analysis), suggesting that direct HIV-1 transmission had occurred from subject 1 to subject 2. Interestingly, routine testing using the MT-2 assay for the presence of CXCR4-using variants in subject 1, the potential donor, never yielded a positive test, suggestive of the absence of CXCR4-using variants, while subject 2, the potential recipient, tested positive in the MT-2 test at 3.4 years after seroconversion.

To explore in more detail the HIV-1 evolutionary history in these two subjects and the origin of CXCR4-using variants in the recipient, we performed ML as well as Bayesian MCMC analysis of the two subject’s viral env sequences. The ML tree (Supplementary Figure A1A) and the Bayesian maximum clade credibility (MCC) tree (Figure 2A) topologies for the alignment of env gp120 C1-V5 sequences reconstructed from clonal HIV-1 variants (alignment A) supported a single transmission event from donor to recipient, showing a temporal structure of sequences and monophyletic clustering of the recipient sequences, which branched off from late donor sequences. Moreover, the phylogenies also demonstrated monophyletic clustering of the recipient’s CXCR4-using variants sequences, which descended from his R5 variants. The ML tree (Supplementary Figure A1B) and Bayesian MCC tree (Figure 2B) topologies for the shorter alignment of env gp120 C2-C4 sequences derived from clonal HIV-1 variants and viral RNA in serum (alignment B) also showed a temporal structure of sequences. However in these trees, recipient sequences did not cluster monophyletically, as single donor sequences (three in ML tree, five in MCMC tree) from later time points fell within the recipient’s cluster. Additionally, the majority of sequences with predicted and/or in vitro tested CXCR4 use of the recipient were part of a main CXCR4-using sequence cluster but three of them constituted a separate cluster. Interestingly, these three sequences were derived from serum at 10 or 2 months before the estimated date of emergence of CXCR4-using variants and provided discordant results by the V3 coreceptor prediction methods. This may imply that these viral variants from serum were intermediates between CCR5- and CXCR4-using viruses that failed to produce replication-competent progeny. This was further supported by the absence of viral descendants from these three virus variants. In conclusion, the most informative alignment (alignment A) clearly supports a single transmission event from donor to recipient. Due to lower evolutionary information, the shorter alignment B did not yield accurate reconstructions of the phylogenetic relationships. However, the clustering of donor sequences from late time points with recipient sequences from early time points
further underscores the high sequence similarity between viral populations isolated from subject 1 and 2, and hence their close genetic relationship.

Bayesian MCMC analysis based on env gp120 C1-V5 sequences from clonal HIV-1 variants (alignment A) estimated time to the most recent common ancestor (TMRCA) of the recipient at 98 (95% credible intervals (CI) = 93.9-99.8) months after SC of the donor. This estimate dates the TMRCA for the recipient, which puts a lower bound on date of HIV-1 transmission, back to around 12 July 1991, which corresponds to the SC date of the recipient (14 July 1991).

**Evolutionary rate of HIV-1 variants from donor and recipient**

The mean evolutionary rate in the HIV-1 transmission chain was estimated to be 0.00012 (95% CI = 0.0009-0.0015) substitutions per site per month for the env gp120 C1-V5 sequences (alignment A), but the coefficient of variation (0.44 (95% CI = 0.22-0.71)) indicated considerable rate variation among branches, which justifies the
HIV-1 evolution in a donor-recipient pair

Donor:
1: 40 months since SC
2: 66 months since SC
3: 96 months since SC
4: 111 months since SC

Recipient:
1: 101 months since SC of the donor
2: 115 months since SC of the donor
3: 124 months since SC of the donor
4: 126.6 months since SC of the donor
5: 127.2 months since SC of the donor
6: 134 months since SC of the donor
7: 136 months since SC of the donor
8: 148 months since SC of the donor
9: 151 months since SC of the donor
10: 156 months since SC of the donor
**HIV-1 evolution in a donor-recipient pair**

Use of a relaxed clock for the transmission history. To assess whether evolutionary rates were different between HIV-1 variants in the donor and in the recipient, we performed Bayesian MCMC analysis of the clonal HIV-1 variants sequences for each patient separately. For subject 2, we also performed an additional analysis solely on the clonal R5 variants sequences. Bayes factor testing indicated that a strict clock provided a good fit for subject 1, but a relaxed clock appeared to be more appropriate for subject 2 (for all clonal HIV-1 variants sequences as well as clonal R5 variants sequences separately, data not shown). For both the strict and relaxed clock estimates (Figure 3), the mean evolutionary rate of the donor’s R5 variants sequences was 1.6-fold lower than the estimated evolutionary rate of the recipient’s R5 and CXCR4-using variants sequences. However, the differences were less pronounced when the mean evolutionary rate of the donor’s R5 variants sequences was compared to the mean evolutionary rate of the recipient’s R5 variants sequences (1.3 and 1.4-fold difference for strict and relaxed clock estimates, respectively) and credible intervals were even more broadly overlapping.

**Estimation of dN/dS ratios of HIV-1 variants from donor and recipient**

To estimate dN/dS ratios, three different models were fitted to the alignment of donor and recipient env gp120 C1-V5 sequences derived from clonal HIV-1 variants (alignment A): model A (a single dN/dS ratio for all branches), model B (separate dN/dS ratios for branches in the donor and recipient clade) and model C (separate dN/dS ratios for branches in the donor clade, in the recipient R5 variants clade and in the recipient CXCR4-using variants clade) (Table 2). Likelihood ratio testing revealed that Model B provided a significantly better fit to the data than the simpler single ratio model (Model B, p=0.012, and that Model C did not provide any improvement in fit respect to Model 2 (p=0.391). Although dN/dS ratios for both subjects were < 1 under model B, indicating predominantly negatively selection, the higher dN/dS ratio in the recipient compared to the donor may be attributed to more neutral evolution in
general or a bout of positive selection in the viral evolutionary history in this subject. Given that Model 3 did not lead to a better fit, we cannot conclude any difference in dN/dS between R5 and CXCR4-using sequence clades and that selection pressure on the CXCR4-using viral population, although contributing to the overall dN/dS ratio for the recipient, it is not solely responsible for the differences with the dN/dS ratio of the donor.

To examine in more detail the differences observed in dN/dS ratio estimates between donor and recipient, we estimated the branch lengths of the MCC tree for alignment A in expected synonymous (E[S]) and nonsynonymous (E[N]) changes using a codon model that allows for a different synonymous and nonsynonymous substitution rate along each branch (a `local’ codon model[8]) (Supplementary Figures A2 A and B). We subsequently plotted E[S] and E[N] divergence over time in Figure 4. A marked increase in E[N] could be observed between the first and second sampling time point of the recipient (101 and 126.6 months since SC of the donor, respectively), whereas the E[S] increase remains fairly monotonous in both the donor and recipient. In fact, between this points the E[N] divergence exceeds the E[S] divergence whereas

Table 2. dN/dS estimates

<table>
<thead>
<tr>
<th>Model</th>
<th>Branch set</th>
<th>dN/dS</th>
<th>Likelihood</th>
<th>LRT comparison</th>
<th>Δ and p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model A</td>
<td>D &amp; R(R5+X4)</td>
<td>0.67</td>
<td>-5017.9</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Model B</td>
<td>D</td>
<td>0.54</td>
<td>-5014.7</td>
<td>ModelA vs ModelB</td>
<td>6.37 (p=0.012)</td>
</tr>
<tr>
<td></td>
<td>R (R5+X4)</td>
<td>0.88</td>
<td>-5014.7</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Model C</td>
<td>D</td>
<td>0.54</td>
<td>-5014.3</td>
<td>ModelA vs ModelC</td>
<td>0.74 (p=0.391)</td>
</tr>
<tr>
<td></td>
<td>R (R5)</td>
<td>0.82</td>
<td>-5014.3</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>R (X4)</td>
<td>1.12</td>
<td>-5017.9</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

D: donor; R: recipient; R5: branches for clonal R5 variants sequences; X4: branches for clonal CXCR4-using variants sequences; dN/dS: non-synonymus/synonymus rates ratio; LRT: Likelihood Ratio Test; D represents twice the log likelihood difference between the simple and more complex model and p represents the probability that D would be obtained under the simple model (for one degree of freedom, the chi-square cut-off is 3.84 at the 0.05 significance level).

![Figure 4. Root-to-tip divergence based on MCC branch lengths estimated in expected Nonsynonymos (E[N]) and Synonymous Substitutions (E[s]) for clonal HIV-1 variants from donor and recipient. White-filled symbols and black lines represent donor divergences whereas grey-filled symbols and grey lines represent the recipient divergences. Lines represent moving averages for E[S] (dotted lines) and E[N] (full lines).](image)
they both keep a roughly similar pace during the evolutionary history in general. This is a clear signal for a selective sweep between these time points. The E[N] and E[S] trees also show that while there is a long branch between these time points in the E[N] tree, there is hardly any divergence in the E[S] tree. This suggests that the increase in E[N] between the first two time points of the recipient is responsible for the higher dN/dS in the donor as compared to the recipient. The fact that this selection is restricted to the time between these time points explains why the evolutionary rate is only moderately faster in general between donor and recipient and why dN/dS is not particularly higher for the CXCR4-using variants.

**CCR5 use of R5 variants from donor and recipient**

In order to determine the efficiency of CCR5 use of R5 variants, before and after transmission, we tested the sensitivity of longitudinally obtained R5 variants from donor and recipient to two anti-CCR5 monoclonal antibodies directed against the CCR5 N-terminal domain (MAb Rob13) and the second extracellular loop (ECL2) of CCR5 (MAb RoAb952). A minimum of two and a maximum of six R5 variants per time point were tested (Table 1).

R5 variants from the donor showed an increasing resistance to MAb Rob13-mediated inhibition overtime (Figure 5A), as shown by a significant increase in IC₅₀ values per time point (P = 0.012), but not to MAb RoAb952-mediated inhibition (Figure 5B). However, we did observe a trend towards higher IC₅₀ values of MAb RoAb952 for R5 variants from the later time point (111 months after SC; average IC₅₀ (ng/ml) = 850.8±724.8) as compared to R5 variants isolated at earlier time points ((40, 66 and 90 months after SC; average IC₅₀ (ng/ml) = 138±102.3); P = 0.0685)).

A decrease in IC₅₀ values between R5 variants isolated at the first time point and at later time points from the recipient was observed for the two anti CCR5 MAbs (Figure 5A, B), although these differences did not approach statistical significance.

R5 variants from donor and recipient isolated close to the transmission event showed similar sensitivity to inhibition by the two MAbs (Figure 5A, B), as there were no significant differences between IC₅₀s of the earliest R5 variants from the recipient and IC₅₀s of R5 variants isolated from the donor prior (90 months after SC) and after (111 months after SC) the transmission event (98 months after SC). This leads us to conclude that these variants used CCR5 with the same efficiency.

**Envelope molecular properties of HIV-1 variants from donor and recipient**

Variation in charge and PSSM score of the env gp120 V3 region and in number of Potential N-linked Glycosylation Sites (PNGS) of the env gp120 C1-V5 region was studied, over the course of infection, for R5 variants from the donor and for R5 and CXCR4-using variants from the recipient.

The charge of the V3 regions for R5 variants from both donor and recipient increased over time (P = 0.03 and P = 0.032 respectively; Figure 6A) but was always lower than for CXCR4-using variants of the recipient. Whereas no change in PSSM score was observed for R5 variants from the donor, PSSM score for the R5 variants from the recipient significantly increased over time (P < 0.0001; Figure 6B). R5 variants
isolated at the last three time points from the recipient had overall more positively charged V3 regions and higher PSSM score than R5 variants from the donor ($P < 0.0001$ in both cases).

An overall increase in number of PNGS over time was observed in R5 variants of both patients ($P = 0.031$ and $P = 0.002$ respectively; Figure 6C). For R5 variants of the recipient, this increase was more pronounced, reaching the maximum peak at the second time point, after which a decrease in number of PNGS was observed in both R5 and CXCR4-using variants.

The V3 charge for the earliest R5 variants from the recipient was similar to the V3 charge for the R5 variants isolated from the donor prior (90 months after SC) and after (111 months after SC) the transmission. PSSM score between the earliest R5 variants from the recipient and the genetically most related

![Figure 5](image)

**Figure 5.** Susceptibility to anti-CCR5 monoclonal antibodies mediated inhibition of longitudinally obtained clonal R5 variants from donor and recipient. Geomean IC$_{50}$ values and standard deviation per time point are shown. *P*-value for Kruskal-Wallis test for donor. A: susceptibility to Rob13 monoclonal antibody; B: susceptibility to Rob952 monoclonal antibody; IC$_{50}$: 50% inhibitory concentration; SC: seroconversion.

The V3 charge for the earliest R5 variants from the recipient was similar to the V3 charge for the R5 variants isolated from the donor prior (90 months after SC) and after (111 months after SC) the transmission. PSSM score between the earliest R5 variants from the recipient and the R5 variants from the donor isolated prior (90 months after SC) transmission was similar, but PSSM score of the R5 variants isolated from the donor isolated after (111 months after SC) transmission was significantly higher than for the earliest R5 variants of the recipient. Moreover, the number of PNGS was also similar for the earliest R5 variants from the recipient and the genetically most related

![Figure 6](image)

**Figure 6.** Envelope molecular properties from viral populations from donor and recipient during disease course. A: average charge and standard deviation of the gp120 V3 region; B: average PSSM score and standard deviation of the gp120 V3 region; C: average number of PNGS and standard deviation in gp120 C1-V5; *P*-values for Kruskal-Wallis test for donor; $^bP$-values for Kruskal-Wallis test for recipient; *Mann-Whitney U test; SC: seroconversion.
R5 variants from the donor, which according to the phylogenetic analysis (Figures 2A and Supplementary Figure A1A), were the variants isolated after the transmission (111 months after SC). However, the number of PNGS of the earliest R5 variants from the recipient was lower than in R5 variants isolated from the donor prior the transmission event (90 months after SC).

**DISCUSSION**

The predominance of R5 HIV-1 variants early in infection as a result of the selective transmission of R5 variants is still object of debate and the origin of CXCR4-using HIV-1 variants is not yet understood. Emergence of CXCR4-using viruses might either reflect (re-)appearance of these viral variants from a reservoir that was established at the time of infection or de novo evolution from R5 viruses. The first scenario assumes transmission of both R5 and CXCR4-using variants, with preferential expansion of R5 variants and preservation of replication-competent CXCR4-using variants in body compartments outside the peripheral blood during the initial stage of infection. The second scenario, on the contrary, implies the selective transmission and/or initial outgrowth of R5 variants, which may acquire the ability to use CXCR4 through evolution.

Here, we had the unique opportunity to study in detail HIV-1 transmission from an individual who only harbored R5 variants during his entire disease course, to an individual in whom CXCR4-using variants emerged after a 3.4 year-period during which only R5 variants were detected. Transmission of R5 variants from donor to recipient appeared to be the result of a single transmission event as supported by phylogenetic analysis. The possibility that both R5 and CXCR4-using variants were transmitted was unlikely as CXCR4-using variants remained undetected in PBMC and serum of the donor during his entire follow-up in the ACS, which lasted from 2.8 until 11.7 years after SC. Moreover, phylogenetic analysis showed that CXCR4-using variants that emerged in the recipient descended from the transmitted and initially expanded R5 viral population. In conclusion, our data suggest that emergence of CXCR4-using variants in the recipient was not caused by the re-activation of a viral reservoir established soon after infection, but by evolution from the transmitted and continuously evolving R5 variants in this individual.

The selective forces playing a role in the emergence of CXCR4-using variants are still not clearly defined and there is currently no explanation for the absence of those variants during the entire disease course of some HIV-1 infected individuals. Although our study describes only a single transmission pair with differential evolution of HIV-1 coreceptor usage, we highlight evolutionary, phenotypic and clinical parameters that may be associated with the emergence of CXCR4-using variants in the recipient. Detailed analysis of selection pressure revealed a selective sweep during the first two time points after transmission in the recipient, whereas selection pressure varied at a similar pace in donor and recipient at other time frames of the evolutionary history of this transmission chain. The higher ratio of nonsynonymous versus synonymous substitution rates (dN/dS) for the viruses of the recipient, which could not be explained by positive selection on the CXCR4-using variants population, can thus be attributed
to this specific period (approximately within the first 2.4 years after transmission) of stronger positive selection in the recipient. The fact that positive selection in the recipient is restricted to this time range also explains that viral evolutionary rate, which is the resultant of the viral mutation rate, viral generation time and the selective pressure acting on the virus, in the recipient is only moderately faster than in the donor. In addition, whilst optimization of CCR5 use was observed over time in the donor, already the earliest recipient’s R5 variants displayed an efficiency of CCR5 use that was similar to the late R5 variants of the donor, which were isolated close to the moment of transmission, and maintained this phenotype during the follow-up time of 4 years. Interestingly, donor and recipient were also different with respect to clinical parameters. While the donor had a median CD4 count of 760 cells/µl (range: 1490-420) with a median viral load (VL) of 10^{3.5} copies/ml (range: 10^{3}-10^{3.9}) during the follow-up period spanning 2.8-7 years after SC, the recipient had a median CD4 count of 405 cells/µl (range: 660-260) with median VL of 10^{4.5} copies/ml (range: 10^{4.3}-10^{4.8}) already during the first 3.3 years of infection, before the emergence of CXCR4-using variants (Figure 1). The lower CD4 counts and higher VL at early disease stage in the recipient could be secondary to the efficient CCR5 use of the R5 virus population that initiated the infection in the recipient. Alternatively or in addition, it could reflect intrinsic characteristics of the recipient’s immune system: lower target cell availability from the beginning of the infection could at least partially explain the stronger selection pressure exerted against the viruses of the recipient during the first 2.4 years of infection.

Given the high mutation rate and rapid replication dynamics of HIV-1, and considering that only few amino acid changes in the envelope V3 region are sufficient for a switch from CCR5 to CXCR4 coreceptor use^{89-95}, emergence of CXCR4-using variants – if beneficial for the virus - would be expected to occur rapidly and relatively early in the course of infection in every patient. However, absence of CXCR4-using variants during the entire disease course is observed in some HIV-1 infected individuals^{29,96} and in the individuals in whom those variants emerge, this generally does not occur at an early disease stage. The decreased replication rate, reduced coreceptor efficiency and high susceptibility to neutralizing antibodies^{97-99} of intermediate variants, suggest that the transition from CCR5- to CXCR4-usage involves a stage of reduced viral fitness. Therefore, it cannot be excluded that additional, compensatory mutations, most likely context dependent, are required to create the background in which the essential V3 region mutations indeed result in a successfully replicating CXCR4-using HIV-1 variant. In our study, the transmitted R5 variants, as a result of adaptation to the environment of the donor, may have had already a backbone with part of the mutations required for the transition to a CXCR4-using phenotype and evolution simply continued in the recipient. Indeed, the gp120 envelope V3 regions of late R5 variants from the recipient were overall more positively charged than of R5 variants from the donor and had higher PSSM scores. Given that the V3 region of a CXCR4-using envelope is generally more positively charged than in a CCR5-using envelope^{89,90,100,101} and it has a higher PSSM score^{66}, this may indicate that the envelope composition of the recipient’s R5 viruses may have been more similar to the one of a CXCR4-using virus. The fact that emergence of CXCR4-using variants was not observed in the last follow-up time point of the donor (2.5 years after the last time point studied), implies that the recipient’s
environment played a role on the acquisition of CXCR4 use of those R5 variants potentially more prone to switch. We previously demonstrated an increasing incidence of CXCR4-using HIV-1 variants when CD4 counts had dropped below 500 cells/µl\textsuperscript{96} which suggested that host immune surveillance may counteract the development of these viruses. In this perspective, the maintenance of a functioning immune system for a longer period in the donor, as reflected by his higher CD4 counts, may help to explain the absence of CXCR4-using variants in this patient.

We also observed an initial increase in PNGS in the viral envelope sequences of R5 variants from donor and recipient. This may reflect the continuous viral escape from the autologous humoral immune response in the host\textsuperscript{102-108}. Given that this increase in PNGS is much more pronounced between the first two time points of the recipient and this coincides with a time frame of strong positive selection, it seems most likely caused by humoral immune pressure. We hypothesize that the envelope changes associated with the increased efficiency of CCR5 use may impose the exposition of envelope regions targeted by the antibody response, making the virus more vulnerable to particular immune responses. As a result viruses may be selected for a denser glycan shield that can occlude these regions. This is further supported by the observation that this period of selective sweep in the recipient coincided with an increase, albeit not statistically significant, in sensitivity to anti-CCR5 MAbs, which suggests that immune escape occurred at the expense of the efficiency of CCR5 use. In the recipient, the decrease in PNGS in both R5 and CXCR4-using variants later in the course of infection could be a consequence of a declining humoral immune pressure, the latter being a sign of the impairment of the immune system, which, as discussed before, might be one of the factors favouring the emergence of CXCR4-using variants.

Our data also showed that R5 HIV-1 variants from a recently infected individual utilized CCR5 with similar efficiency and had similar envelope molecular properties as the late R5 variants from the donor, who at the time of transmission was chronically infected. This suggests that adaptations that support virus replication, efficient target cell infection and evasion from the host immune response can be preserved upon transmission, which may contribute to the evolution of HIV-1 phenotype at a population level\textsuperscript{109-111}. The notion that the acute phase of HIV-1 infection is associated with an increased risk of transmission relative to chronic infection and that transmission during this early phase plays a pivotal role in HIV-1 spread\textsuperscript{112-117} could however explain why the adaptation of HIV-1 at a population level seems to be a slow process.

Although our study is based on a single transmission case, we have shown here de novo evolution of CXCR4-using variants from transmitted R5 variants. Emergence of CXCR4-using variants in the recipient was associated with an R5 viral population subjected to a selective sweep after transmission, with an efficiency of CCR5 use similar to the late stage R5 variants from the donor and with an envelope with more positively V3 charged region. These viral characteristics, which are at least partly the result of the recipient’s host selective pressure on the transmitted HIV-1 variants adapted to the former’s host environment, may have been instrumental to the emergence of CXCR4-using viruses under the specific conditions of the new host environment. However, given both the characteristics of the viral lineage establishing infection and host environment differed in our study, the interplay of viral and host factors on this evolutionary process remains to be further elucidated.
ACKNOWLEDGEMENTS

The Amsterdam Cohort Studies on HIV infection and AIDS, a collaboration between the Amsterdam Health Service, the Academic Medical Center of the University of Amsterdam, Sanquin Blood Supply Foundation, the University Medical Center Utrecht, and the Jan van Goyen Clinic are part of the Netherlands HIV Monitoring Foundation and financially supported by the Center for Infectious Disease Control of the Netherlands National Institute for Public Health and the Environment. We are indebted to all participants for their continuous participation in the study.

This research was financially supported by the Landsteiner Foundation for Blood Transfusion Research (LSBR grant 0317), the European Community’s Seventh Framework Programme NGIN (FP7/2007-2013) under grant agreement n° 201433 and by the Academic Medical Center at the University of Amsterdam and the Netherlands AIDS fund (grant nr. 6006). PL is supported by a postdoctoral fellowship from the Fund for Scientific Research (FWO) Flanders.

REFERENCES

HIV-1 evolution in a donor-recipient pair


HIV-1 evolution in a donor-recipient pair


HIV-1 evolution in a donor-recipient pair

Peripheral blood mononuclear cells during primary HIV-1 infection and impact on disease progression. AIDS 24:2305-2312.


Phylogenetic Handbook. A practical approach to phylogenetic analysis and hypothesis testing.


85. Westby, M., M. Lewis, J. Whitcomb, M. Youle, A.L. Pozniak, I.T. James, T.M. Jenkins, M. Perros and R.E. van der. 2006. Emergence of CXCR4-using human immunodeficiency virus type 1 (HIV-1) variants in a minority of HIV-1-infected patients following treatment with the CCR5 antagonist maraviroc is from a pretreatment CXCR4-using virus reservoir. J. Virol. 80:4909-4920.


106. Sagar, M., X. Wu, S. Lee and J. Overbaugh. 2006. Human immunodeficiency virus type 1 V1-V2 envelope loop sequences expand and add glycosylation sites over the course of infection, and these modifications affect antibody neutralization sensitivity. J. Virol. 80:9586-9598.


HIV-1 evolution in a donor-recipient pair


SUPPLEMENTARY MATERIAL

Figure A1. Maximum Likelihood trees of heterochronous env sequences. A: The tree includes env gp120 C1-V5 sequences derived from clonal HIV-1 variants isolated from peripheral blood mononuclear cells (PBMC). B: The tree includes env gp120 C2-C4 sequences derived from clonal HIV-1 variants isolated from PBMC and env gp120 C2-C4 sequences generated from viral RNA isolated from serum. Black: clonal R5 variants sequences of the donor; dark grey: clonal R5 variants sequences of the recipient; light grey: clonal CXCR4-using variants sequences of the recipient. Numbers indicate bootstrap values. The scale bar (horizontal line) indicates branch length corresponding to nucleotide substitutions per site. SC: seroconversion.
HIV-1 evolution in a donor-recipient pair

B

Donor:
1: 45 months since SC
2: 66 months since SC
3: 90 months since SC
4: 111 months since SC

Recipient:
1: 121 months since SC of the donor
2: 115 months since SC of the donor
3: 124 months since SC of the donor
4: 126.6 months since SC of the donor
5: 127.2 months since SC of the donor
6: 134 months since SC of the donor
7: 139 months since SC of the donor
8: 140 months since SC of the donor
9: 151 months since SC of the donor
10: 156 months since SC of the donor
HIV-1 evolution in a donor-recipient pair

Donor:
1: 40 months since SC
2: 66 months since SC
3: 90 months since SC
4: 111 months since SC

Recipient:
1: 101 months since SC of the donor
4: 126.6 months since SC of the donor
7: 139 months since SC of the donor
9: 151 months since SC of the donor
HIV-1 evolution in a donor-recipient pair

Figure A2. Maximum credibility trees for clonal HIV-1 variants from donor and recipient with branch lengths in Expected Synonymous (E[S]) and Nonsynonymous (E[N]) Substitutions. A: E[S] MCC tree based on env gp120 C1-V5 sequences derived from clonal HIV-1 variants; B: E[N] MCC tree based on env gp120 C1-V5 sequences derived from clonal HIV-1 variants. Black: clonal R5 variants sequences of the donor; dark grey: clonal R5 variants sequences of the recipient; light grey: clonal CXCR4-using variants sequences of the recipient. The scale bar (horizontal line) indicates branch length corresponding to E[S] or E[N] substitutions per site respectively. SC: seroconversion.