HIV-1 superinfection in homosexual men
Rachinger, A.

Citation for published version (APA):
Chapter 5

High-resolution phylogenetic reconstruction methods support the absence of HIV-1 superinfection in a documented donor-recipient pair during long-term follow-up

Andrea Rachinger¹, Esther F. Gijsbers¹, Judith A. Burger¹, Precious Manyenga¹, Tom Derks van de Ven¹, Ineke G. Stolte²,³, Angélique B. van ’t Wout¹, Philippe Lemey⁴, Hanneke Schuitemaker¹

¹Department of Experimental Immunology, Sanquin Research, Landsteiner Laboratory, and Center for Infection and Immunity Amsterdam (CINIMA) at the Academic Medical Center of the University of Amsterdam, Amsterdam, The Netherlands.
²Cluster of Infectious Diseases, Department of Research, Public Health Service of Amsterdam, Amsterdam, The Netherlands.
³Department of Internal Medicine, CINIMA, Academic Medical Center of the University of Amsterdam, Amsterdam, The Netherlands.
⁴Department of Microbiology and Immunology, Rega Institute, K.U. Leuven, Belgium.

Manuscript submitted for publication
Abstract

Background Close to fifty cases of superinfection with heterologous HIV strains have been reported, suggesting that infection with HIV-1 does not necessarily elicit protection against incoming virus variants. Here, in an HIV-1 donor-recipient pair that continued to have unprotected sexual intercourse after transmission, we studied whether superinfection with homologous HIV-1 could be demonstrated in any of the two individuals.

Results Analysis was performed on 263 env and 151 gag sequences from twelve donor and ten recipient serum samples spanning 8.6 and 5.3 years of follow-up, respectively. Consistent Maximum Likelihood and Bayesian Markov Chain Monte Carlo phylogenetic analyses confirmed the initial transmission of HIV-1. However, all phylogenetic analyses convincingly supported the absence of subsequent superinfection in this donor-recipient pair over the course of HIV-1 infection.

Conclusion In this homosexual donor-recipient pair that after the initial HIV-1 transmission event continued unprotected sexual intercourse, also with other partners, no evidence for superinfection could be demonstrated. These results may imply that protective immunity against incoming homologous HIV-1 may exist. However, the influence of external factors contributing to the specific risk setting needs to be addressed.

Background

To date, almost fifty cases of HIV-1 superinfection have been documented worldwide [1-23]. The occurrence of HIV-1 superinfection may imply that infection with HIV-1 does not elicit protective immunity against newly incoming HIV-1 variants. Some reports suggested a role for adaptive immunity in the prevention of superinfection [24,25] while others did not find a protective effect [2,26]. Moreover, similar incidence rates for initial HIV-1 infection and HIV-1 superinfection [9,13,18,20], and reports on higher numbers of superinfections as compared to earlier studies due to more frequent sampling [18,23] argue against efficient protection by host immune responses elicited by the initial HIV-1 strain towards incoming heterologous HIV-1. As to date only superinfections with heterologous viral strains have been reported, protection against homologous HIV-1 might still be a possibility.

Here we performed a longitudinal in-depth analysis to identify possible HIV-1 superinfection in an HIV-1 transmission couple of which the therapy-naïve HIV-1 positive partners continued to engage in unprotected sexual intercourse after the transmission event, also with other partners. Although extensive phylogenetic analyses of HIV-1 envelope (C2-C4) and gag sequences from serial serum samples isolated over 8.6 years and 5.3 years from the donor and the recipient respectively confirmed the initial transmission event, all phylogenetic analyses convincingly supported the absence of superinfection.
Results

Confirmation of the transmission event between donor and recipient and of absent HIV-1 superinfection with heterologous HIV-1

Env and gag sequences derived from the donor and the recipient were aligned with a reference panel of published and unpublished sequences generated from participants of the Amsterdam Cohort Studies. The derived env and gag ML trees confirmed that viruses of the donor and the recipient were very closely related to each other. In particular, the sequences of the recipient were nested in a large cluster of exclusively donor sequences that was maximally supported by bootstrap analysis (data not shown). This viral evolutionary pattern, combined with the epidemiological history given by the two men, confirmed transmission of HIV-1 from patient 18814 to his partner, patient 18766, and was consistent with absence of HIV-1 superinfection with heterologous HIV-1.

No evidence for HIV-1 superinfection with homologous HIV-1

To explore in more detail the viral evolutionary history in this couple and possible HIV-1 transmissions between the two partners over time, we performed Maximum Likelihood (ML) analysis of the patients’ env and gag sequences without reference panels. The best-fit nucleotide substitution models (env: K81uf+I+G, gag: TVM+I+G) were applied to estimate three different ML topologies for each dataset, applying the three different heuristic search branch-swapping algorithms NNI, SPR and TBR. Each ML tree was rooted using the consensus sequence of the first time point of the donor. In env as well as in gag, tree topologies supported a single transmission event from patient 18814 to patient 18766 at about nine months after seroconversion of the donor, when the highest plasma viral load was detected with 5.4*10^5 copies/mL. Tree topologies generated by all three different search algorithms displayed a temporal structure of sequences. For env as well as for gag, sequences derived from the recipient formed a monophyletic cluster independent of the search algorithm used, supporting the absence of additional transmission events between both patients, hence inferring absence of superinfection with homologous virus (for the env and gag ML trees generated by TBR search algorithm, see additional material).

To gain additional support for the absence of additional HIV-1 transmission events between the two HIV-1 positive partners we complemented the ML analyses with a Bayesian Markov Chain Monte Carlo (MCMC) analysis approach that infers a posterior distribution of time-measured phylogenies from temporally-spaced sequence data [27]. For both env and gag, we represent the phylogenetic relationships using a maximum clade credibility tree in Figure 1 and Figure 2, respectively. Both MCMC trees corroborate a single transmission event from the donor to the recipient; a single recipient cluster branches off from the donor tree at
5.4 (2.8-7.6) months and 9.2 (6.39-12.19) months respectively after seroconversion of the donor. Almost all trees (99% and 100% respectively) in the posterior distribution support a monophyletic cluster of the recipient and estimates for the most recent common ancestor of this cluster are 9.8 (6.4-12.7) months and 12.5 (9.5-14.5) months respectively.

Figure 1: Maximum clade credibility tree summarizing the Bayesian MCMC analysis of the heterochronous env sequences. The tree includes 241 env sequences obtained at 23 time points. The recipient cluster is indicated with a gray box. Monophyly of sequences isolated from the recipient is supported by a posterior probability value of 1.0. A time scale in months is indicated below. The tree topology supports absence of HIV-1 superinfection.

The mean evolutionary rates were estimated at 0.0015 (0.0013-0.0017) substitutions per site per month and 0.00069 (0.00052-0.00075) substitutions per site per months for env and gag respectively, and the coefficients of variation for the rates (0.77 (0.59-0.96) and 0.44 (0.25-0.62) for env and gag respectively) indicated significant non-clock like behavior in both gene regions justifying the use of relaxed molecular clocks. In conclusion, we found evidence for only a single transmission event that has occurred according to the viral evolutionary history, which led to the infection of the recipient by the donor. No further transmission events between the two partners, thus superinfection by a homologous virus of any of the partners, were supported (Fig.1, 2).
Figure 2: Maximum clade credibility tree summarizing the Bayesian MCMC analysis of the heterochronous gag sequences. The tree includes 135 gag sequences obtained at 23 time points. The recipient cluster is indicated with a gray box. Monophyly of sequences isolated from the recipient is supported by a posterior probability value of 0.99. A time scale in months is indicated below. The tree topology supports absence of HIV-1 superinfection.

Discussion

The observation that HIV-1 superinfection can occur has lead to the assumption that antiviral immunity elicited against the initial HIV-1 variant does not protect against newly incoming HIV-1 variants. However, all superinfection reports dealt with superinfection with heterologous HIV-1 variants, which does not preclude that elicited HIV-specific immunity could protect against superinfection with homologous viral variants. As superinfection may accelerate the clinical course of HIV-1 infection, knowledge on the possibility and impact of superinfection with homologous virus is relevant for couples between whom HIV-1 transmission has occurred and who may wish to continue to have unprotected sexual intercourse. Here, we performed a longitudinal study on the occurrence of HIV-1 superinfection in an HIV-1 transmission couple of which the partners continued to have unprotected sexual intercourse after the HIV-1 transmission event, also with other partners. We applied multiple molecular phylogenetic techniques to analyze 241 C2-C4 env and 135 gag sequences generated from a total of 23 time points covering
time periods of 8.6 and 5.3 years of HIV-1 positive follow-up of the donor and the recipient respectively. This covered a risk period for HIV-1 superinfection with homologous HIV-1 of 5.3 years, being the time period during which both partners were HIV-1 positive and engaging in sexual intercourse. Despite this in-depth analysis, we could not find any evidence for superinfection with homologous HIV-1 in the partners over time. HIV-1 superinfection with heterologous virus was discarded based on ML analysis.

However, intermittent superinfection may have been missed due to insufficient sampling. Moreover, superinfecting strains present at too low abundance may also have remained undetected. Therefore, we can only conclude that superinfection with a persisting HIV-1 variant that contributes significantly to the viral quasispecies can be excluded.

In a previous study, two HIV-1 transmission couples similar to the couple in our study were studied for the occurrence of superinfection with homologous virus [28]. However, in that study only two time points for each of these two couples were investigated. Moreover, both partners of each couple were on antiretroviral therapy reducing the risk for superinfection; first, by suppressing viral replication resulting in lowered or absent viral load in the potential donor and second, by the prophylactic activity of cART on a newly incoming virus before establishment of systemic superinfection in the potential recipient. Therefore, no firm conclusions on the probability of superinfection with homologous HIV-1 could be drawn from this study [28]. Both patients in our study remained therapy-naïve throughout the entire study period with detectable viral loads at levels that are likely to support HIV-1 transmission [29].

As mentioned above, all published cases to date involved HIV-1 superinfection by heterologous HIV-1 strains indicating that immune responses generated after initial HIV-1 infection may be insufficient in protecting against newly incoming heterologous viruses. This is supported by studies describing similar incidence rates for initial HIV-1 infection and HIV-1 superinfection [9,13,18,20] and more recent publications with higher numbers of detected superinfections in more in-depth longitudinal cohort studies [18,23]. Although a certain degree of protection against heterologous HIV-1 strains may be generated by the HIV-1 specific adaptive immune responses [24,25], the opposite has been described as well [2,26]. This and the constantly growing number of documented HIV-1 superinfections underscores the dilemma that research is currently facing with regard to understanding mechanisms involved in immune responses against HIV-1 and their possible role in resistance against superinfection.
Conclusions

Our longitudinal case study allowed detailed analyses of the homologous virus variants within two HIV-1 positive partners, with a clear separation of their HIV-1 quasispecies in ML and MCMC trees (additional MCMC analyses see additional material). Both phylogenetic methods rejected HIV-1 superinfection.

Assuming that superinfection has indeed not occurred within the couple studied here, two mechanisms, perhaps in combination, might have played a role. Firstly, HIV-1 specific immune responses might have eliminated incoming virus during ongoing unprotected sex. In our case, since both partners were infected with a highly related virus, immune responses might have been specific enough to eliminate the partner's virus variants and thus have protected against homologous HIV-1 superinfection. However, both partners also engaged in sex with additional partners (data not shown) and we did not find any evidence of superinfection with a heterologous viral variant either. Therefore, external factors modulating the risk of HIV-1 transmission may also be relevant for HIV-1 superinfection. Previous studies have demonstrated that viral inoculum size in semen, often represented by the plasma viral load, frequency of unprotected (receptive) anal intercourse, number of partners, number of sex acts, presence of other sexually transmitted infections and mucosal trauma are all associated with initial HIV-1 infection [29-32]; these factors may be relevant for HIV-1 superinfection as well.

In summary, this case study, in which we performed frequent sampling during a long follow-up period, did not reveal HIV-1 superinfection. Although it is tempting to speculate on a role for adaptive immunity in the protection against superinfection with a homologous virus, the influence of external factors contributing to the specific risk setting needs to be addressed.

Methods

Patients

We studied a homosexual male couple from which both partners participated in the Amsterdam Cohort Studies on HIV Infection and AIDS (ACS) and who both seroconverted for HIV-1 antibodies during follow-up. Patient 18814 (donor) seroconverted in July 1987 and his partner patient 18766 (recipient) seroconverted one year later, in July 1988. Both men reported that patient 18766 contracted HIV-1 from his partner, patient 18814. The donor started combination antiretroviral therapy (cART) in March 1996, while the recipient was diagnosed with AIDS in June 1994 and died from AIDS in September 1995. During routine visits, both participants
filled out questionnaires on their sexual behavior, which at that time included questions regarding number of partners, sexual techniques and condom use. The Amsterdam Cohort Studies are conducted in accordance with the ethical principles of the Helsinki declaration and written informed consent was obtained prior to data collection from every participant. The ACS were approved by the Medical Ethics Committee of the Academic Medical Center.

**Serum samples**

Cryopreserved serial serum samples with detectable viral load (VL>1,000 copies/mL plasma) were used for analysis (Fig.3). From the donor, we studied twelve therapy-naïve serum samples spanning a time period of eight years and seven months starting in August 1987, which was eleven days post-SC until March 1996 when he started cART. Median time between investigated time points was nine months (range, 3-13). In addition, we studied one serum sample from July 2000, which was the first sample with detectable viral load after initiation of cART (Fig.3).

**Figure 3: Plasma viral loads in the donor-recipient pair.** Samples from the donor and the recipient are indicated by squares and triangles, respectively; filled symbols indicate samples that were investigated. The recipient was diagnosed with AIDS in June 1994, and died of an AIDS-related death in September 1995. The donor started combination antiretroviral therapy (cART) in March 1996.

From the recipient, ten therapy-naïve serum samples were used for analysis covering a time period of five years and three months starting in February 1989, which was seven months after SC until May 1994, with a median time of six months (range, 3-12) between time points (Fig.3). Viral load (VL) in the serum samples from the donor that were used for analysis and obtained at the first twelve time points varied from $1.9 \times 10^4$ – $5.4 \times 10^5$ copies/mL (median VL of $8.1 \times 10^4$ copies/mL)
and was 1.1*10^3 copies/mL at the one time point during antiretroviral therapy in July 2000. Median VL in the serum samples of the recipient was 6.8*10^4 copies/mL, ranging from 6.8*10^3 – 3.6*10^5 copies/mL (Fig.3).

RNA isolation, RT-PCR and PCR amplification
Viral RNA was isolated from 140μl serum (Qiamp Viral Mini Kit, Qiagen, Hilden, Germany) and eluted in 50μl from which 10μl RNA was reverse transcribed into cDNA (SuperScript™ First strand synthesis system, Invitrogen, Carlsbad, CA) using the env-specific primer Seq2 (5’-TCCTCCATATCTCCTCCAGGTC-3’) and the gag-specific primer gag outer rev (5’-GCCTGTCTCTCAGTAC-3’), respectively. For generating the env C2-C4 fragment primer couples Seq2 - Seq3 (5’-TATGGGATCAAAGCTAAAGCATG-3’) and Seq5 (5’-GTCAACTCAACTGCTTAAATGGC-3’) were used for the first- and second-round PCR, respectively (GoTaq Flexi DNA Polymerase, Promega, Madison, WI). For amplification of gag fragments, the primer couples gag fw (5’-CGAGCGAGGACTCGGCTTGCTG-3’) - gag outer rev (5’-GCCTGTCTCTCAGTAC-3’) and BssHII fw (5’-TGCTGAAGCGCCCGCAGGC-3’) - p17rev (5’-CATAACTCTTTGGTATGG-3’) were used in the first- and second-round PCR, respectively (GoTaq Flexi DNA Polymerase, Promega, Madison, WI). For both fragments, 5μl cDNA input was used for the first-round PCR, and 2μl first-round PCR product as input in the second-round PCR. Both, first- and second-round PCR reactions were carried out in a total volume of 25μl and the amplification for both PCR was as follows: one cycle at 94°C for 5 minutes, 35 cycles at 94, 50 and 70°C of respectively 45, 30 and 90 sec, followed by a final elongation step of 10 min at 70°C resulting in the final PCR products env C2-V3-C3-V4-C4 (549 nucleotides, HXB2 positions 7012-7560) and gag (1175 nucleotides, HXB2 positions 704-1878) respectively.

Molecular cloning and sequencing of C2-C4 env and gag
From each serum sample a median of two second-round PCR products was cloned (range, 1-4) into the pGEM T Easy Vector System (Promega, Madison, WI), transformed into competent DH5α E.coli (Invitrogen, Carlsbad, CA) and plated on LB agar using blue/white screening. White colonies were picked at random (4-32 per time point). The vector primers T7 (5’-TAATACGACTCACTATAGGG-3’) and SP6 (5’-GATCAGTACAGCTATAG-3’) were used to amplify the cloned env and gag PCR products with the above described PCR program. After purification (ExoSAP-IT, USB, Cleveland, OH), PCR products were sequenced with T7 and SP6 (env) and T7, SP6, p17fw (5’-GCTAAACAGTGGGGGAGCCGACATG-3’) and p24_5I (gag) (Big Dye terminator v1.1 cycle sequencing kit, Applied Biosystems, Foster City, CA). Four to twenty colonies per cloning reaction were sequenced. Sequences were determined with an automated DNA sequencer (Applied Biosystems, Foster City, CA). A total of
241 env sequences and 135 gag sequences were included in the phylogenetic analysis with a median of eleven env and six gag sequences per time point (range, 1-19).

**Phylogenetic analysis**

Clonal sequences were aligned using the ClustalW algorithm [33] and alignments were manually edited in BioEdit (BioEdit, version 7.0.5.3; Ibis Biosciences). For confirmation of the HIV-1 transmission event between the two partners, a local reference panel of published and unpublished sequences generated from participants of the Amsterdam Cohort Studies was added to the patients’ env and gag alignments. The best-fitting nucleotide substitution models for each region were chosen with Modeltest v3.7 [34] using the hierarchical likelihood tests (hLRTs). Subsequently, phylogenetic analyses were performed using PAUP*4.0 [35]. A Neighbour-Joining (NJ) tree was inferred, followed by a heuristic search for a Maximum Likelihood (ML) tree with the best-fit substitution model and tree-bisection-reconnection (TBR) as the branch-swapping algorithm. Statistical support for nodes was generated with bootstrapping on the NJ tree (1000 repeats). The ML tree was rooted with the most distant sequence. Further, Maximum Likelihood phylogenies were estimated using the couple’s env and gag alignments without reference sequences in PAUP applying three different heuristic search branch-swapping algorithms: nearest-neighbour-interchange (NNI), subtree-pruning-regrafting (SPR) and tree-bisection-reconnection (TBR). Again, the corresponding nucleotide substitution model was chosen as described above and applied in the ML analysis. ML trees were displayed with Dendroscope [36].

Bayesian phylogenetic reconstruction was performed using Markov Chain Monte Carlo (MCMC) analysis implemented in BEAST [27]. BEAST focuses on rooted, time-measured phylogenetic trees with a coalescent prior. We analyzed both the couple’s env and gag alignments using a general time-reversible nucleotide substitution model with gamma-distributed rate variation, a lognormal relaxed clock model [37] and a flexible Bayesian skyride tree prior [38]. MCMC analyses were run for 50 million generations and diagnosed using Tracer (http://tree.bio.ed.ac.uk/software/tracer). Maximum clade credibility trees were obtained using TreeAnnotator and visualized in FigTree (http://tree.bio.ed.ac.uk/software/figtree/).

**Acknowledgements**

The Amsterdam Cohort Studies on HIV infection and AIDS, a collaboration between the Public Health Service of Amsterdam, the Academic Medical Center of the University of Amsterdam, the Sanquin Blood Supply Foundation, the University Medical Center Utrecht, and the Jan van Goyen Medical Center, 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. Financial support for this study was obtained from the Academic Medical Center at the University of Amsterdam. We thank Peter van Swieten for excellent technical assistance. The authors are indebted to the participants of the ACS. The study could not be performed without their dedication.

Reference List


**Supplementary Methods**

*Phylogenetic analysis:* Bayesian Markov Chain Monte Carlo (MCMC) phylogeny estimates were performed with BEAST v1.4.8 [34]. Chain lengths and factors of the Bayesian MCMC run of the patient’s env and gag datasets were optimized during analysis. For the couple’s env and gag sequences, one MCMC chain each was run for 150 x 10^6 and 200 x 10^6 generations, respectively, trees were sampled every 1000 generations. Parameters were analyzed in Tracer v1.4 (Rambaut A, Drummond AJ (2007) Tracer v1.4, [http://beast.bio.ed.ac.uk/Tracer](http://beast.bio.ed.ac.uk/Tracer)) and the Effective Sampling Size (ESS) of parameters was calculated and being >200 ensured convergence of the MCMC analysis. The model parameters for both datasets, env and gag, consisted of the GTR model as the nucleotide substitution model, with estimated base frequencies, under a relaxed clock model (uncorrelated log-normal distribution). Here, a population genetic model assuming a constant population size over time was applied. The final Bayesian consensus trees were created in TreeAnnotator v.1.4.8 with a burn-in of 10% of all sampled trees; final MCMC trees were displayed in FigTree ([http://tree.bio.ed.ac.uk/software/figtree/](http://tree.bio.ed.ac.uk/software/figtree/)) (data not shown). Both trees supported absence of HIV-1 superinfection due to monophyly of sequences isolated from the recipient (posterior probability of 1.0).
Supplementary Figures

Fig.S1: Maximum likelihood (ML) tree using the tree-bisection-reconnection (TBR) search algorithm of heterochronous env sequences generated from serum samples of the donor and the recipient (R). In total, 241 env sequences of 23 time points were included in the ML analysis. Sequences of the donor are represented in black, sequences of the recipient in gray. The recipient cluster (R) is indicated with a box. Monophyly of sequences (bootstrap value of 1.0) isolated from the recipient supported absence of HIV-1 superinfection within this couple over time. Scale bar represents nucleotide substitutions per site.

Fig.S2: Maximum likelihood (ML) tree using the tree-bisection-reconnection (TBR) search algorithm of heterochronous gag sequences generated from serum samples of the donor and the recipient (R). In total, 135 gag sequences of 23 time points were included in the ML analysis. Sequences of the donor are represented in black, sequences of the recipient in gray. The recipient cluster (R) is indicated with a box. Monophyly of sequences (bootstrap value of 1.0) isolated from the recipient supported absence of HIV-1 superinfection within this couple over time. Scale bar represents nucleotide substitutions per site.