

RESEARCH ARTICLE

Frequency of occurrence of HIV-1 dual infection in a Belgian MSM population

Laura Hebberecht, Leen Vancoillie, Marlies Schauvliege, Delfien Staelens, Kenny Dauwe, Virginie Mortier, Chris Verhofstede*

Aids Reference Laboratory, Department of Clinical Chemistry, Microbiology and Immunology, Ghent University, Ghent, Belgium

* chris.verhofstede@ugent.be



Abstract

Introduction

HIV-1 dual infection is a condition that results from infection with at least two HIV-1 variants from different sources. The scarceness of information on this condition is partly due to the fact that its detection is technically challenging. Using next-generation sequencing we defined the extent of HIV-1 dual infection in a cohort of men who have sex with men (MSM).

Material & methods

Eighty-six MSM, diagnosed with HIV-1 subtype B infection between 2008 and 2013 were selected for next-generation sequencing of the HIV-1 *envelope* V3. Sequencing was performed on 2 plasma samples collected with an interval of > 6 months before the initiation of antiretroviral therapy. Maximum likelihood phylogenetic trees were inspected for dual infection, defined as the presence of two or more monophyletic clusters with $\geq 90\%$ bootstrap support and a mean between-cluster genetic distance of $\geq 10\%$. To confirm dual infection, deep V3 sequencing of intermediate samples was performed as well as clonal sequencing of the HIV-1 *protease-reverse transcriptase* gene.

Results

Five of the 74 patients (6.8%) for whom deep sequencing was successful, showed clear evidence of dual infection. In 4 of them, the second strain was absent in the first sample but occurred in subsequent samples. This was highly suggestive for superinfection. In 3 patients both virus variants were of subtype B, in 2 patients at least one of the variants was a subtype B/non-B recombinant virus.

Conclusions

Dual infection was confirmed in 6.8% of MSM diagnosed with HIV-1 in Belgium. This prevalence is probably an underestimation, because stringent criteria were used to classify viral variants as originating from a new infection event.

OPEN ACCESS

Citation: Hebberecht L, Vancoillie L, Schauvliege M, Staelens D, Dauwe K, Mortier V, et al. (2018) Frequency of occurrence of HIV-1 dual infection in a Belgian MSM population. PLoS ONE 13(4): e0195679. <https://doi.org/10.1371/journal.pone.0195679>

Editor: Luis Menéndez-Arias, Consejo Superior de Investigaciones Científicas, SPAIN

Received: October 20, 2017

Accepted: March 27, 2018

Published: April 6, 2018

Copyright: © 2018 Hebberecht et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: All relevant data are within the paper and its Supporting Information files.

Funding: The AIDS Reference Laboratory of Ghent is supported by the Belgian Ministry of Social Affairs through a fund within the Health Insurance System. The funder had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing interests: The authors have declared that no competing interests exist.

Introduction

Human immunodeficiency virus type 1 (HIV-1) dual infection occurs when an HIV-1 infected patient is infected by a new viral variant from another host. This results in the presence of two variants, each originating from an independent infection event. When the second infection occurs nearly simultaneously with the first, or at least before seroconversion, the condition is generally called co-infection. When the second infection occurs after seroconversion it is called superinfection [1]. One of the firmest indications that co- or superinfection is possible is the existence of recombinant viruses. Recombination between genetically distinct viruses can only occur after infection of a single cell with both variants. Recombinant viruses can remain restricted to a single patient (unique recombinant forms; URF) or be further transmitted. Recombinant forms detected in at least three unrelated individuals are called circulating recombinant forms (CRF). CRF contribute extensively to the genetic variability of HIV-1 [2, 3] and the constant increase in number of reported CRF (<https://www.hiv.lanl.gov/content/sequence/HIV/CRFs/CRFs.html>) suggests high frequencies of dual infection.

The first patient with HIV-1 dual infection was documented in 2002 [4]. Since then, several other, mostly isolated cases of dual infection, have been reported [1]. Estimates of the overall frequency of this condition vary between 0 and 20% [5–8]. The number of well documented superinfections, however, remains low. Difficulties to demonstrate this condition may be the most important reason for that. Heteroduplex mobility assays [9, 10], multi-region hybridization assays [11, 12] and bulk viral sequencing [13, 14] have been used to identify the presence of different viral variants in the same individual. But all these methods lack sensitivity and specificity and require additional clonal sequencing as confirmation [9, 14]. With the introduction of next-generation sequencing, enabling the detection of variants representing 1% or less of a virus population, new possibilities for sensitive detection of dual infection arose [14].

The majority of studies concentrated on intersubtype dual infections [8, 12, 15, 16]. Reports of dual infection with viruses of the same subtype (intrasubtype dual infection) are much more limited [17, 18], probably caused by the challenge to differentiate these closely resembling strains. Alternatively, it is possible that the susceptibility for re-infection is higher when the second variant is genetically more distinct from the one already present.

An additional difficulty when trying to demonstrate dual infection is that co-presence of the initial and the superinfecting strain may be limited in time. In an extensive study on the dynamics of the virus population after superinfection, Chaillon et al. [19] found that only in 2 of the 7 patients both strains co-circulated during the entire follow-up period. In 4 patients the original strain was rapidly overgrown by the superinfecting strain and in 1 patient a recombinant virus took over [19]. To what extent the level of genetic differences between the initial and the superinfecting strain influences the population dynamics remains unclear.

Further research on dual infection is needed to elucidate this. Research on co- and superinfection may be particularly useful for a better understanding of the host-virus interaction and provide valuable information for vaccine development.

In Belgium, the overall burden of HIV-1 is limited, but the prevalence in men having sex with men (MSM) is high, reaching up to 6% in some cities [20]. Because of the high HIV prevalence in this specific population, a high risk of dual infection can be assumed. As in other Western-European countries, the HIV-1 epidemic in MSM is dominated by subtype B [21]. In this study we therefore specifically focused on the detection of dual infection in HIV-1 subtype B infected MSM. Eighty-six MSM, diagnosed with HIV-1 infection between 2008 and 2013, were included. Roche 454 next-generation sequencing of the variable V3 region of the HIV-1 *envelope* gene was performed to identify dual infection. Clonal sequencing of the HIV-1 *protease-reverse transcriptase* region was done to confirm the selected cases of dual infection.

Materials and methods

Study population

Patients were selected retrospectively from the cohort of HIV-1 infected individuals followed at Ghent University Hospital. About half of the patients in this cohort are MSM diagnosed with HIV-1 subtype B infection. We focused specifically on this population and selected 86 patients diagnosed with a subtype B infection between 2008 and 2013 from whom leftover plasma, collected at two time points with an interval of at least 6 months before initiation of antiretroviral therapy (ART), was available.

Subtyping

Subtyping of the HIV-1 infection at diagnosis was done using the subtyping tool of Smartgene IDNS (Smartgene, Zug, Switzerland) and the *protease* and *reverse transcriptase* sequences generated for the purpose of baseline resistance analysis.

Subtyping of the different virus variants present in the patients with dual infection was done using COMET (Context-based Modeling for Expeditious Typing) HIV-1 (Version 1.0) [22] and all clonal *protease-reverse transcriptase* sequences.

Ethical approval

The study was approved by the Ethics Committee of Ghent University Hospital (reference number 2014/0173). All patients provided informed consent for the use of their leftover plasma for scientific research. All samples were anonymized before processing.

Viral RNA extraction and preparation of the amplicon library

Viral RNA was extracted from 140 μ L of EDTA plasma using the QIAamp Viral RNA Mini kit (Qiagen, Hilden, Germany). Ten μ L of the extracted RNA was reverse transcribed and subsequently amplified in one step with the Titan one tube RT-PCR kit (Roche, Basel, Switzerland) using two sense primers 5' -GAGGATATAATCAGTTTATGG-3' and 5' -GGATATAATCAGYYTATGGGA-3' and two antisense primers 5' -GGTGGGTGCTATTCCCTAATGG-3' and 5' -GGTGGGTGCTAYTCCYAAITG-3'. The resulting amplicon was used as template for the second, nested amplification with the FastStart High Fidelity PCR SystemTM (Roche, Basel, Switzerland). The nested primers 5' -TCAACHCAAYTRCTGTTAAATGG-3' and 5' -ATTTCTGGRTCYCKCCTG-3' were extended with Roche multiplex identifiers (MID 1 to MID 34). One unique MID was used per sample. Generated amplicons, containing 345 bp of the V3 region of *env* (HXB2 position 6,990 to 7,336) were visualized by electrophoresis on a 2% agarose gel. Positive reactions were purified with the Agencourt AMPure XP DNA purification kitTM (Analabs, Ghent, Belgium) and quantified using the Qubit 2.0 fluorometer and Qubit dsDNA HS Assay kitTM (Life technologies, Carlsbad CA, USA). Thereafter, amplicons were diluted to a concentration of 10^9 molecules / μ L, pooled and further diluted to a concentration of 10^6 molecules/ μ L.

Next-generation sequencing

Equimolar amplicon mixtures were added to DNA capture beads (Roche, Basel, Switzerland) at a ratio of 1.4 copies per bead. The resulting solution was used for emulsion PCR executed with the Lib-A emPCR kit (Roche, Basel, Switzerland). Next-generation sequencing of the PCR products was performed using the Roche 454 GS Junior next-generation sequencer according to the manufacturer's protocol.

Clonal sequencing of the HIV-1 *protease-reverse transcriptase*

Viral RNA was extracted from 140 μ L of EDTA plasma using the QIAamp Viral RNA Mini kit (Qiagen, Hilden, Germany) and reverse transcribed using SuperScript™ III reverse transcriptase (Life Technologies, Carlsbad CA, USA) and the antisense primer 5′-GGGATGTGTA CTTCTGAACTTAYTYTTGG-3′. The resulting cDNA was diluted to a concentration that resulted in a PCR success rate of maximum 30%. A nested PCR was run with Platinum™ Taq DNA Polymerase High Fidelity (Life Technologies, Carlsbad CA, USA), outer primers 5′-CTCAATAAAGCTTGCCCTTGAGTGC-3′ and 5′-GGGATGTGTA CTTCTGAACTTAYTYTTGG-3′ and inner primers 5′-AAGTAGTGTGTGCCCGTCTGT-3′ and 5′-CACCTGCCATCTGTTTTCCATA-3′. Amplicons were purified with Agencourt AMPure beads (Analisis, Ghent, Belgium) and subsequently subjected to sequencing of the *protease* and *reverse transcriptase* gene using the Big-Dye™ Terminator v3.1 Cycle Sequencing Kit (Life Technologies, Carlsbad CA, USA) following the procedure described before [23]. Sequences were analyzed using the 3500 Genetic Analyzer (Applied Biosystems, Lennik, Belgium). Proofreading was done using Smartgene IDNS (Smartgene, Zug, Switzerland). Sequences with mixed nucleotides were eliminated. Population HIV-1 *protease-reverse transcriptase* sequences were obtained through routine laboratory activities for the purpose of drug resistance analysis using an in-house Sanger sequencing protocol.

Data analysis

Sequences generated by next-generation sequencing were trimmed to 228 nucleotides. Identical reads were clustered before and after manual correction of homopolymer regions using in-house software and aligned in BioEdit [24]. Reads with a coverage of less than 5 or reads shorter than 228 nucleotides were removed.

For the sequences generated by clonal or population sequencing, proofreading was done using Smartgene IDNS (Smartgene, Zug, Switzerland). Concatenated sequences of 831 nucleotides long were then aligned in BioEdit [24].

Phylogenetic analysis

Phylogenetic trees were constructed with PhyML 3.0 using automatic model selection (<http://www.atgc-montpellier.fr/phyml/>). The resulting trees were visualized with iTol [25]. Genetic distances were calculated with MEGA 7 based on the Tamura-Nei model [26]. Highlighter plots were generated using the highlighter tool created by the Los Alamos sequence database (https://www.hiv.lanl.gov/content/sequence/HIGHLIGHT/highlighter_top.html).

Recombination analysis was performed with Simplot [27].

Time of infection

Dating of the infection at diagnosis was done using the BED HIV-1 Incidence EIA and the HIV-1 LAg-Avidity EIA (both from Sedia Biosciences Corporation, Portland, Oregon, USA). The analyses were performed on the first sample collected after diagnosis. Only a concordant recent infection result for both assays was accepted as a valid indication of recent infection, all other results were interpreted as long-term infections. This approach has been extensively validated for accuracy [28] and allows to reliably differentiate infections ≤ 130 days from infections > 130 days.

Results

Characteristics of the study population

Next-generation sequencing was attempted on both samples of 86 selected patients. The sequencing was considered successful if a read coverage of more than 500 was obtained. For

samples failing this criterion, RNA extraction and next-generation sequencing were repeated. Sufficient read coverage for both samples was finally obtained for 74 of the 86 patients (86%), with a median number of sequencing reads per sample of 1588 (IQR: 851–1915). The median time interval between collection of the two samples analyzed was 25 months (IQR: 15–39). Forty two patients of the 74 (57%) were classified as recently infected at the time of collection of the first sample, 32 were diagnosed with a long-term infection. The median age at diagnosis was 36 years (IQR: 28–46), the median viral load (VL) of the first sample collected was 4.45 log copies per milliliter (cp/mL) (IQR: 4.05–4.87 log cp/mL) and the median CD4⁺ T-cell count was 554 cells/ μ L (IQR: 422–702 cells/ μ L).

Identification of dual infection

All next-generation sequences were aligned to construct per patient alignments. These patient alignments were then randomly combined to assemble 6 large alignments, each including 24 or 25 patients. Each patient was included twice in this exercise but, although shuffling was done at random, care was taken to avoid that duplicates were included in the same alignment. A phylogenetic tree was constructed for each of the 6 alignments and visually examined for clustering. A patient was considered HIV-1 dual infected when his sequences were distributed over 2 or more clusters with a bootstrap value of $\geq 90\%$ in one of the phylogenetic trees. Ten patients (patients 08, 16, 17, 20, 29, 35, 50, 54, 60 and 66) fulfilled this criterion. For patients 08, 16, 35 and 66, sequences of at least one other patient separated the two phylogenetic clusters, indicating high genetic difference between the variants in these patients. An example phylogenetic tree is shown in [Fig 1](#).

The mean genetic distances of the *env* sequences in the last sample collected from each of the 74 patients fluctuated between 0.92% and 19.94% (mean 3.28%; IQR 1.78% - 4.02%). For the 10 patients with a tree topology suggesting dual infection, the mean distances were 19.94%, 13.10%, 11.02%, 8.28%, 4.12%, 6.35%, 4.52%, 4.40%, 4.12% and 2.59% for patients 66, 08, 35, 16, 50, 54, 17, 29, 20 and 60 respectively. As the overall mean genetic distance may be biased in case of uneven distribution of the variants and is sensitive to differences in the number of sequences generated per patient, we also calculated the mean pairwise genetic distance between the sequences of the two clusters for those patients with a dichotomized tree topology. The obtained between-cluster mean distances were 56.18%, 30.65%, 26.63%, 24.78%, 12.40%, 9.80%, 8.80%, 8.76%, 8.10% and 7.80% for patients 66, 08, 16, 35, 50, 54, 60, 29, 17 and 20 respectively. Using a cut off of 10%, 5 patients were defined as dually infected. For the 5 other patients, dual infection was likely but the genetic distance was considered too low for definite conclusion. Highlighter plots constructed for the 5 patients with confirmed dual infection are shown in [Fig 2](#). The highlighter plots for the 5 patients with presumed dual infection are presented in [S1 Fig](#). Further analysis was limited to the patients with confirmed dual infection. The characteristics of these patients and information on the tested longitudinal samples is summarized in [Table 1](#).

Analysis of intermediate samples

Intermediate samples from the 5 patients with confirmed dual infection were also subjected to next-generation V3 sequencing. Three intermediate samples were available for patient 08 and patient 66 and 2 samples for patient 16 and patient 35. No intermediate samples were available for patient 50. The per patient phylogenetic trees ([Fig 3A–3D](#)) clearly revealed the presence of 2 genetically different viral variants in at least one of the intermediate samples in all 4 patients ([Table 1](#)).

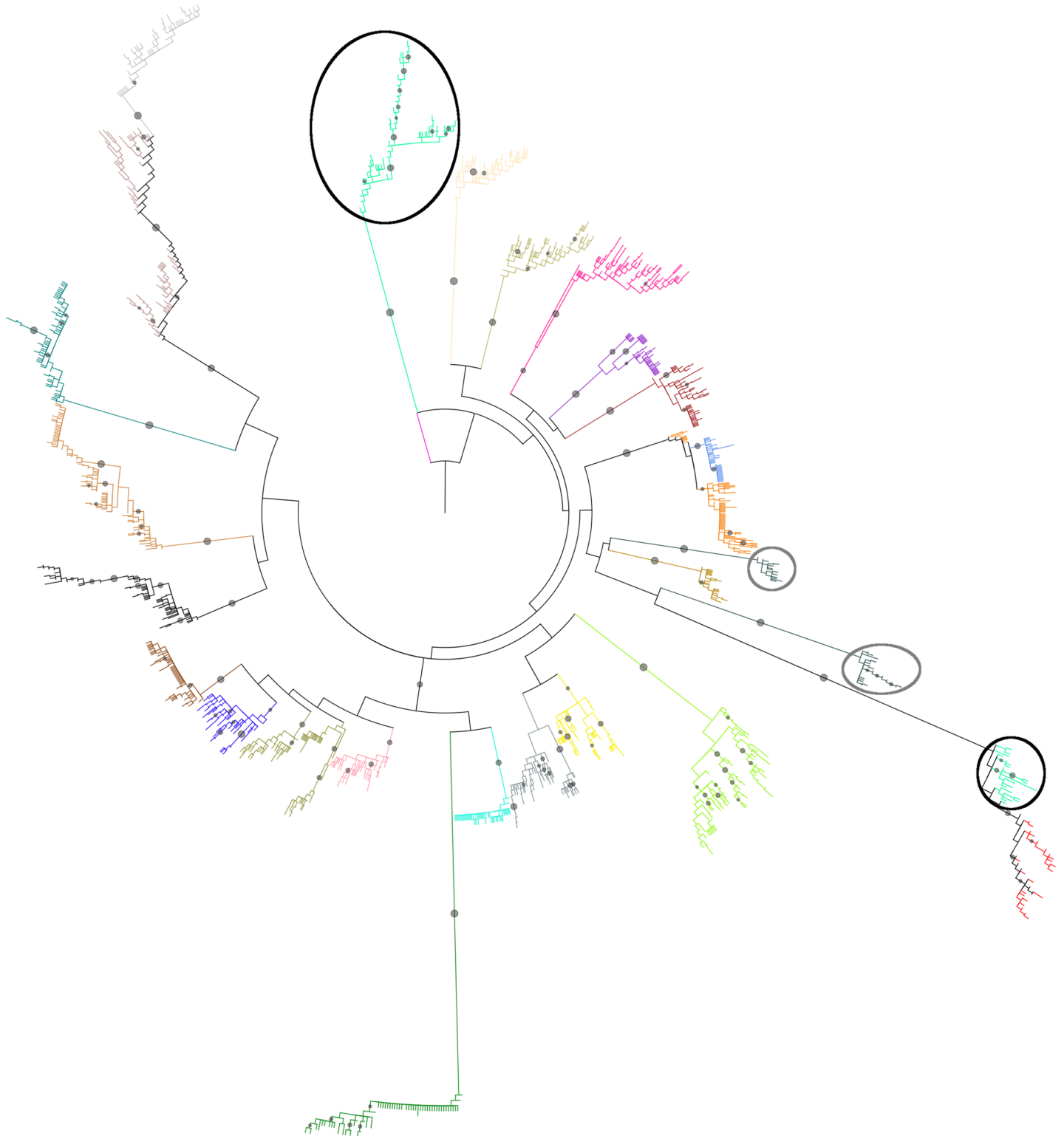


Fig 1. Example phylogenetic tree visualized with iTol. The tree, rooted on the HXB2 reference, presents the sequences of 25 patients. The sequences of the patients considered dual infected (patients 35 and 66) are marked with respectively a grey and black circle. Mid-branch filled grey circles indicate a bootstrap value $\geq 90\%$.

<https://doi.org/10.1371/journal.pone.0195679.g001>

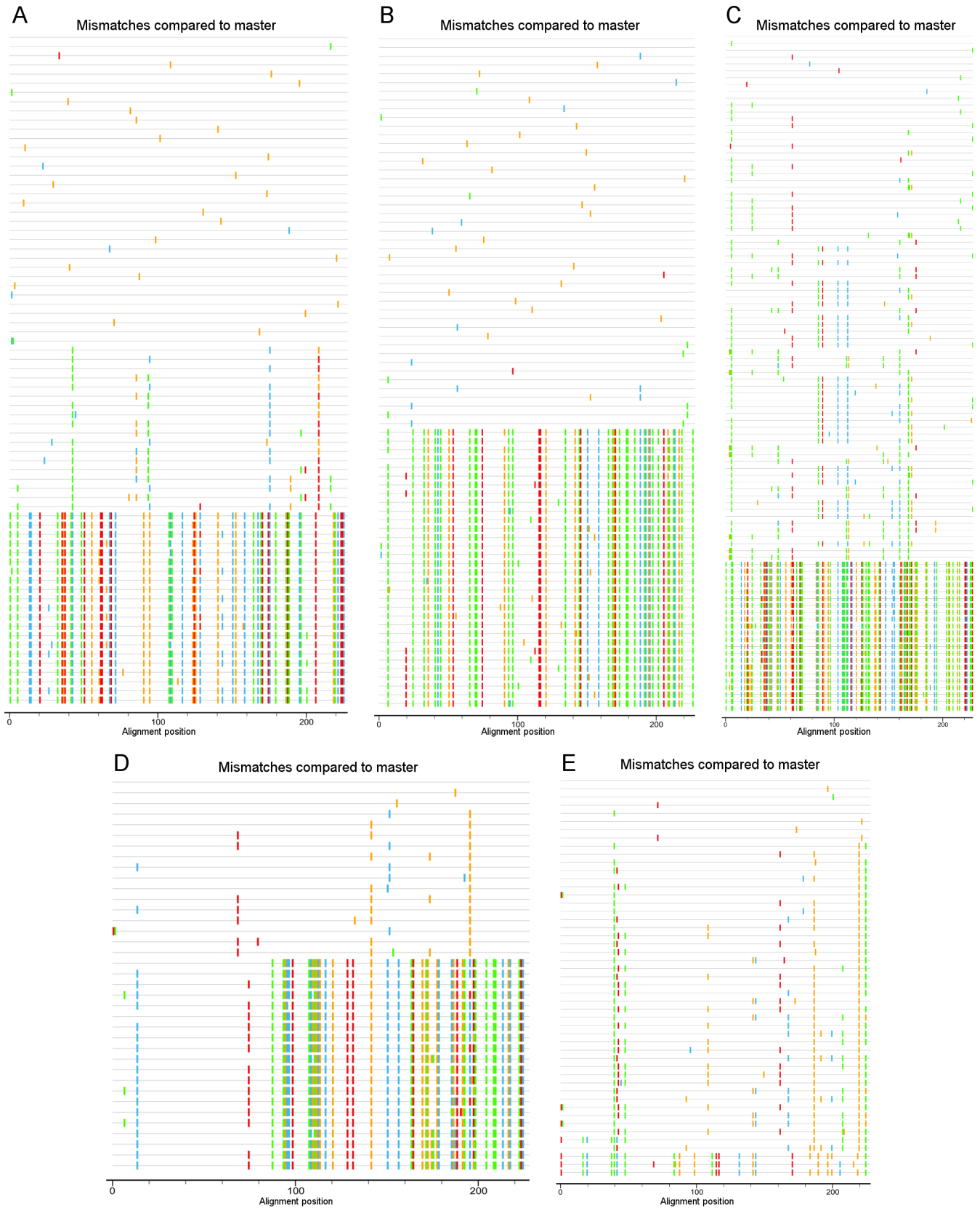


Fig 2. Highlighter plots constructed with the *env* V3 sequences of the 5 patients with dual infection. (A) Patient 08, (B) Patient 16, (C) Patient 66, (D) Patient 35, (E) Patient 50. Sequences are aligned to the most abundant sequence in the initial sample of each patient. Mutations compared to the first sequence are color coded.

<https://doi.org/10.1371/journal.pone.0195679.g002>

A significant increase in viral load coinciding with the presumed time point of superinfection was observed in 2 of the superinfected patients. The viral load increased with 2.19 log cp/mL in patient 08 and with 0.94 log cp/mL in patient 16 (Table 1).

Analysis of HIV-1 *protease-reverse transcriptase* sequences

Limiting dilution clonal sequencing of the *protease-reverse transcriptase* (*pol*) gene was performed for the viruses isolated from the last sample of the 5 patients with dual infection. Respectively 11, 14, 21, 11 and 20 clonal sequences were obtained for patients 08, 16, 35, 50 and 66. They were aligned and analyzed together with sequences obtained as part of routine drug resistance analysis. Sequence alignment, construction of the phylogenetic tree and highlighter plots was done as described above for the *env* sequences. Before phylogenetic analysis, the alignment was supplemented with a selection of 162 reference sequences from MSM reflecting the overall sequence variability in the cohort from where the study patients were selected. Inspection of the resulting phylogenetic tree confirmed the dichotomized topology

Table 1. Characteristics of the 5 patients with HIV-1 dual infection.

Patient ID	Country of origin	Infection stage at diagnosis	Sample date	Time since first sample (days)	CD4 count (cells/ μ L)	VL (log c/mL)	Viral variants	Subtype variant 1*	Subtype variant 2*
08	Belgium	Long-term	6/06/2008	0	1060	4.54	1	B	Unassigned
			20/02/2009	259	757	4.19	1	(B with fragments	
			22/02/2010	626	660	5.13	1 + 2	of A1/D)	
			24/01/2011	962	627	4.76	1 + 2		
			19/12/2011	1291	525	5.12	1 + 2		
16	Belgium	Recent	26/01/2009	0	403	2.71	1	B	B
			9/06/2009	134	419	2.98	1 + 2		
			30/10/2009	277	318	5.17	1 + 2		
			19/01/2010	358	298	4.83	1 + 2		
35	Belgium	Recent	24/02/2010	0	1020	4.53	1	Majority B	B
			10/03/2010	14	624	4.58	1	with fragments	
			16/06/2010	112	968	4.97	1 + 2	of G	
			9/01/2013	1050	353	4.32	1 + 2		
50	France	Long-term	13/01/2011	0	482	3.13	1 + 2	B	B
			26/03/2012	438	498	3.47	1 + 2		
66	Belgium	Long-term	19/03/2012	0	555	4.45	1	B	B
			20/02/2013	338	587	4.69	1 + 2		
			3/07/2013	471	414	4.28	1 + 2		
			27/11/2013	618	452	4.83	1 + 2		
			26/03/2014	737	338	4.76	1 + 2		

* Subtyping based on the *protease-reverse transcriptase* sequence

VL: Viral Load; MSM: Men who have Sex with Men

<https://doi.org/10.1371/journal.pone.0195679.t001>

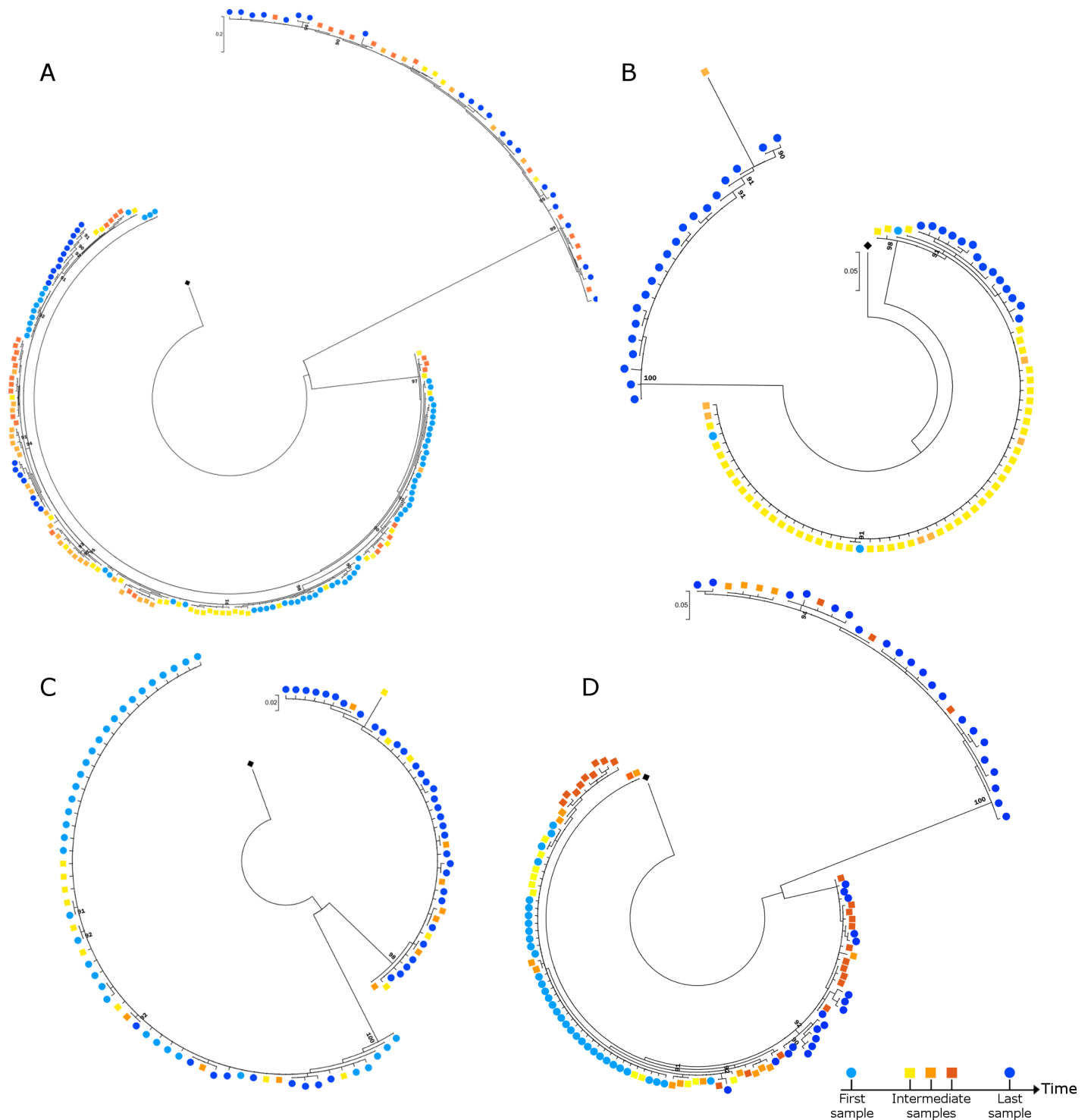


Fig 3. Individual phylogenetic trees of the 4 patients with dual infection from whom more than 2 longitudinal samples were analyzed. (A) Patient 66, (B) Patient 35, (C) Patient 16, (D) Patient 08. Bootstrap values $\geq 90\%$ are displayed. Black diamond = HXB2.

<https://doi.org/10.1371/journal.pone.0195679.g003>

for patients 08, 50 and 66 and revealed a division over at least 3 clusters for patients 16 and 35 (S2 Fig). Highlighter plots (Fig 4) showed patterns highly suggestive for recombination between the initial and superinfecting strains in patients 16, 66 and 35, an assumption that was

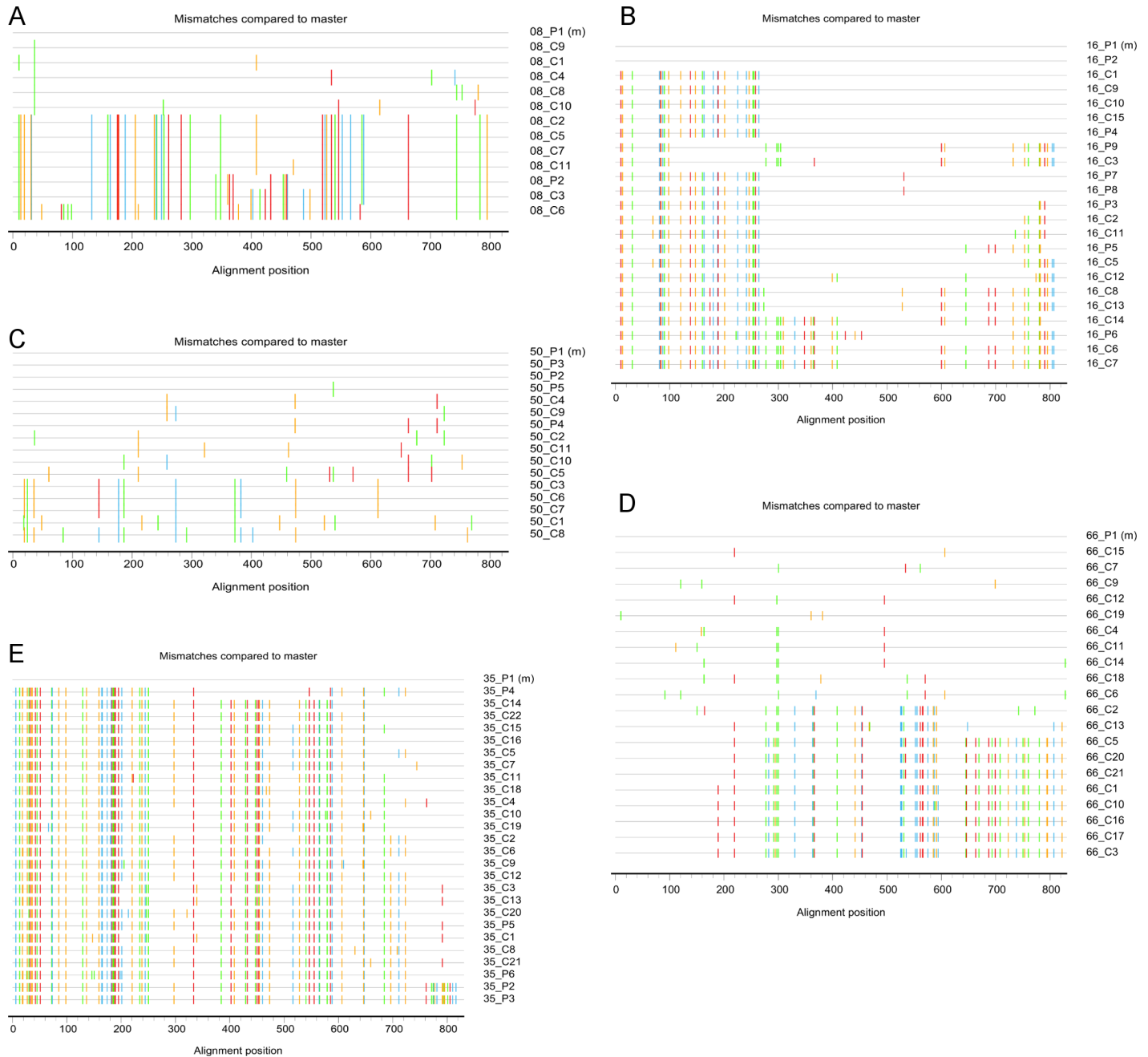


Fig 4. Highlighter plots constructed with the *pol* sequences of the 5 patients with dual infection. (A) Patient 08, (B) Patient 16, (C) Patient 50, (D) Patient 66, (E) Patient 35. Sequences are aligned to the sequence of the first sample collected from each patient. Mutations compared to this sequence are color coded. P = population sequencing (numbered by sampling date from old to recent); C = clonal sequencing (numbered by the amount of clones sequenced).

<https://doi.org/10.1371/journal.pone.0195679.g004>

confirmed by Simplot analysis. The presence of recombinants explains the further branching of the tree. In patient 16, 50 and 66 all sequences were identified as subtype B. In patient 35 and 08 respectively the initial and the superinfecting strain were identified as a subtype B recombinant with fragments attributed to subtype G (patient 35) and subtype A and D (patient 08).

Discussion

Phylogenetic analysis of sequences of the V3 region in the HIV-1 *envelope* gene showed clear indications for dual infection in 5 (6.8%) of 74 randomly selected MSM diagnosed with HIV-1 in Belgium between 2008 and 2013. This figure is most probably an underestimation, because stringent criteria were used to consider viral variants as originating from dual infection. For 5 other patients the topology of the phylogenetic tree was suggestive for dual infection but the genetic difference between the variants was not high enough to confidently exclude that they evolved from the same founder virus. For 64 (86.5%) of the 74 patients phylogenetic analysis of V3 sequences could not reveal any indication for dual infection.

The study further concentrated on the 5 patients with high evidence of dual infection. Additional support for the dual infection was sought by deep sequencing of HIV-1 V3 in intermediate samples and by clonal analysis of a larger fragment in the *protease-reverse transcriptase* gene. In 4 of the 5 patients, a single variant was present in the initial sample with the second variant detected only in samples collected at a later time point. This is a strong argument for superinfection. In one patient both variants were already present in the first sample, hampering discrimination between co- and superinfection. Clonal analysis of the *protease-reverse transcriptase* gene of the variants present in the last sample of the 5 patients confirmed the presence of at least two genetically different variants. Interestingly, in 3 of the 5 patients the highlighter plots revealed the presence of new recombinant forms containing fragments of both variants. Based on this information we are confident to state that in the tested MSM population, at least 4 (5.4%) patients are identified with superinfection and 1 (1.4%) with co- or superinfection. For an additional 5 patients, indications but no convincing evidence for dual infection were found. These numbers are within the estimated overall frequency of dual infection reported by others [5–8].

The fact that after diagnosis some MSM continue to have unprotected sex with known seropositive partners may be one explanation for the observed high frequency of superinfection. Some have also suggested that the threshold for superinfection is lower in the early stages of infection when the immune response has not yet come to full maturation [29–31]. Research on this is limited but it may be significant that of the 4 superinfected patients identified, 2 were diagnosed early after infection and superinfection was demonstrated respectively 4 and 5 months after diagnosis. Whether dual or superinfection has implications on the clinical course of the disease is not clear. In line with the observation of others [18, 32–34], we noticed an increase in viral load following superinfection in patient 08 and 16. It is therefore important to consider the possibility of superinfection in case of unexpected rises in viral load.

Extension of the research on dual infection is partly hampered by the lack of a good detection method. We used next-generation sequencing to screen the genetic variability of the virus population and defined dual infection based on analysis of the phylogenetic tree, calculation of the between-cluster mean genetic distance and visual inspection of the highlighter plots. The outcome of this approach was validated using clonal sequencing of the HIV-1 *protease-reverse transcriptase* gene. While establishing the method, the lack of well-defined measures of naturally occurring inpatient evolution in V3 posed a difficulty. Initially all patients for whom the V3 sequences were divided over at least two phylogenetic clusters with a bootstrap support of at least 90% were considered for further analysis. Others have used a less stringent bootstrap cut-off of 80% [35–37]. In most studies on dual infection the criterion of phylogenetic linkage is extended with genetic distance calculation and in general a cut-off for pairwise *env* genetic distance of 5% is used to discriminate dual infection from inpatient evolution [18, 34, 37]. The overall mean distance however may be influenced by the distribution of the variants and will be lower if one variant is much less represented than the other. We therefore opted for the

use of between-cluster mean distances. These between-cluster distances were high, especially considering the fact that most were intrasubtype dual infections, but the tree topology and the highlighter plots indeed confirmed the large differences between the variants.

It is obvious that the chance to detect dual infection increases when the variants are genetically more distinct. Difficulties to conclude on dual infection with genetically more closely linked variants may explain why intersubtype dual infection is more frequently reported than intrasubtype dual infection. Considering the fact that many geographic areas have a predominant subtype [38] one would expect a higher frequency of intrasubtype recombination. It is possible that the host immune defense is more successful in preventing re-infection with a highly similar virus than in preventing re-infections with a virus of a more aberrant genetic background, but there are no data today to support this hypothesis. Luan et al. detected 18 dual infections in a population of 64 participants from the Liaoning province in northeastern China and found equal amounts of intrasubtype and intersubtype dual infections [37]. Redd et al. examined 149 patients from Uganda and detected 7 dual infections, 4 intrasubtype and 3 intersubtype [39].

To our knowledge, only one study so far has attempted to analyze the dynamics of the viral population after dual infection. The results of this study revealed different patterns, with individuals in whom both the original and second variant remained detectable over longer periods of time, individuals in whom one of both variants rapidly overgrew the other variant, and individuals in whom a recombinant virus took over [19]. This implies that also the timing of the sample collection is important for reliable detection of dual infection, even if highly sensitive techniques are applied. Our analysis of the *protease-reverse transcriptase* sequences clearly revealed the presence of recombinants in 3 of the 5 patients with dual infection. This is remarkable, considering the relatively small genomic fragment analyzed, and deserves further investigation.

A potential hurdle when studying dual infection is the difficulty to discriminate between co-infection and superinfection. Presence of 2 variants in a single sample can indicate either of both [40] and only examination of longitudinal samples will allow to discriminate between co- and superinfection. A limitation of the deep sequencing method that we used for broad scale screening of dual infection is the short sequence read length. This is a limitation not only of the Roche 454 technology but of most of the currently available next-generation sequencing technologies. Ideally, more than one region of the genome must be sequenced to increase the likeliness of detecting dual infection [41].

In conclusion, using next-generation sequencing, we found a high frequency of dual infection in our HIV-1 infected MSM population. Methods to easily identify dually infected individuals will facilitate more in-depth study of this phenomenon and of its consequences for the patient and the virus. This may provide interesting new insights on the interplay between virus and host, information that will be of particular interest for future vaccine research.

Supporting information

S1 Fig. Highlighter plots of the 5 patients with potential indications of dual infection. (A) Patient 54, (B) Patient 60, (C) Patient 29, (D) Patient 20, (E) Patient 17. Sequences are aligned to the most abundant sequence in the first sample collected. Mutations are color coded. (PDF)

S2 Fig. Phylogenetic tree containing the *pol* sequences from the 5 dual infected patients. This maximum likelihood phylogenetic tree was constructed using *pol* sequences from 162 MSM and the 5 dual infected patients (08, 16, 35, 50 and 66), rooted on the HXB2 reference. Visualization with iTol. Mid-branch filled grey circles indicate a bootstrap value $\geq 90\%$. (PDF)

Author Contributions

Conceptualization: Laura Hebberecht, Leen Vancoillie, Virginie Mortier, Chris Verhofstede.

Data curation: Laura Hebberecht, Kenny Dauwe.

Formal analysis: Laura Hebberecht.

Funding acquisition: Chris Verhofstede.

Investigation: Laura Hebberecht.

Methodology: Laura Hebberecht, Leen Vancoillie, Marlies Schauvliege, Delfien Staelens, Kenny Dauwe.

Project administration: Laura Hebberecht, Chris Verhofstede.

Resources: Marlies Schauvliege, Delfien Staelens, Kenny Dauwe.

Software: Leen Vancoillie.

Supervision: Chris Verhofstede.

Validation: Chris Verhofstede.

Visualization: Laura Hebberecht.

Writing – original draft: Laura Hebberecht.

Writing – review & editing: Leen Vancoillie, Kenny Dauwe, Virginie Mortier, Chris Verhofstede.

References

1. Waters L, Smit E. HIV-1 superinfection. *Curr Opin Infect Dis.* 2012; 25(1):42–50. <https://doi.org/10.1097/QCO.0b013e32834ef5af> PMID: 22156898
2. Soares de Oliveira AC, Pessoa de Farias R, da Costa AC, Sauer MM, Bassichetto KC, Oliveira SM, et al. Frequency of subtype B and F1 dual infection in HIV-1 positive, Brazilian men who have sex with men. *Virology.* 2012; 9:223. <https://doi.org/10.1186/1743-422X-9-223> PMID: 23021203
3. Smyth RP, Davenport MP, Mak J. The origin of genetic diversity in HIV-1. *Virus Res.* 2012; 169(2):415–29. <https://doi.org/10.1016/j.virusres.2012.06.015> PMID: 22728444
4. Jost S, Bernard MC, Kaiser L, Yerly S, Hirschel B, Samri A, et al. A patient with HIV-1 superinfection. *N Engl J Med.* 2002; 347(10):731–6. <https://doi.org/10.1056/NEJMoa020263> PMID: 12213944
5. Sheward DJ, Ntale R, Garrett NJ, Woodman ZL, Abdool Karim SS, Williamson C. HIV-1 Superinfection Resembles Primary Infection. *J Infect Dis.* 2015; 212(6):904–8. <https://doi.org/10.1093/infdis/jiv136> PMID: 25754982
6. Vidal N, Diop H, Montavon C, Butel C, Bosch S, Ngole EM, et al. A novel multiregion hybridization assay reveals high frequency of dual inter-subtype infections among HIV-positive individuals in Cameroon, West Central Africa. *Infect Genet Evol.* 2013; 14:73–82. <https://doi.org/10.1016/j.meegid.2012.11.017> PMID: 23232100
7. Redd AD, Ssemwanga D, Vandepitte J, Wendel SK, Ndembu N, Bukonya J, et al. Rates of HIV-1 superinfection and primary HIV-1 infection are similar in female sex workers in Uganda. *AIDS.* 2014; 28(14):2147–52. <https://doi.org/10.1097/QAD.0000000000000365> PMID: 25265078
8. Ssemwanga D, Lyagoba F, Ndembu N, Mayanja BN, Larke N, Wang S, et al. Multiple HIV-1 infections with evidence of recombination in heterosexual partnerships in a low risk Rural Clinical Cohort in Uganda. *Virology.* 2011; 411(1):113–31. <https://doi.org/10.1016/j.virol.2010.12.025> PMID: 21239033
9. Rachinger A, van de Ven TD, Burger JA, Schuitemaker H, van 't Wout AB. Evaluation of pre-screening methods for the identification of HIV-1 superinfection. *J Virol Methods.* 2010; 165(2):311–7. <https://doi.org/10.1016/j.jviromet.2010.02.016> PMID: 20178816
10. Manigart O, Courgnaud V, Sanou O, Valéa D, Nagot N, Meda N, et al. HIV-1 superinfections in a cohort of commercial sex workers in Burkina Faso as assessed by an autologous heteroduplex mobility procedure. *Aids.* 2004; 18(12):1645–51. PMID: 15280775

11. van der Kuyl AC, Cornelissen M. Identifying HIV-1 dual infections. *Retrovirology*. 2007; 4:67. <https://doi.org/10.1186/1742-4690-4-67> PMID: 17892568
12. Herbinger KH, Gerhardt M, Piyasirisilp S, Mloka D, Arroyo MA, Hoffmann O, et al. Frequency of HIV type 1 dual infection and HIV diversity: analysis of low- and high-risk populations in Mbeya Region, Tanzania. *AIDS Res Hum Retroviruses*. 2006; 22(7):599–606. <https://doi.org/10.1089/aid.2006.22.599> PMID: 16831083
13. Fang G, Weiser B, Kuiken C, Philpott SM, Rowland-Jones S, Plummer F, et al. Recombination following superinfection by HIV-1. *AIDS*. 2004; 18(2):153–9. PMID: 15075531
14. Redd AD, Quinn TC, Tobian AAR. Frequency and implications of HIV superinfection. *The Lancet Infectious Diseases*. 2013; 13(7):622–8. [https://doi.org/10.1016/S1473-3099\(13\)70066-5](https://doi.org/10.1016/S1473-3099(13)70066-5) PMID: 23726798
15. Andreani G, Espada C, Ceballos A, Ambrosioni J, Petroni A, Pugliese D, et al. Detection of HIV-1 dual infections in highly exposed treated patients. *Virology*. 2011; 8:392. <https://doi.org/10.1186/1743-422X-8-392> PMID: 21824422
16. Chohan B, Lavreys L, Rainwater SM, Overbaugh J. Evidence for frequent reinfection with human immunodeficiency virus type 1 of a different subtype. *J Virol*. 2005; 79(16):10701–8. <https://doi.org/10.1128/JVI.79.16.10701-10708.2005> PMID: 16051862
17. Kiwelu IE, Novitsky V, Margolin L, Baca J, Manongi R, Sam N, et al. Frequent intra-subtype recombination among HIV-1 circulating in Tanzania. *PLoS One*. 2013; 8(8):e71131. <https://doi.org/10.1371/journal.pone.0071131> PMID: 23940702
18. Wagner GA, Pacold ME, Kosakovsky Pond SL, Caballero G, Chaillon A, Rudolph AE, et al. Incidence and prevalence of intrasubtype HIV-1 dual infection in at-risk men in the United States. *J Infect Dis*. 2014; 209(7):1032–8. <https://doi.org/10.1093/infdis/jit633> PMID: 24273040
19. Chaillon A, Wagner GA, Hepler NL, Little SJ, Kosakovsky Pond SL, Caballero G, et al. Dynamics of viral evolution and neutralizing antibody response after HIV-1 superinfection. *J Virol*. 2013; 87(23):12737–44. <https://doi.org/10.1128/JVI.02260-13> PMID: 24049166
20. Vanden Berghe W NC, Buvé A, Beelaert G, Franssen K, Laga M. A venue-based HIV prevalence and behavioural study among men who have sex with men in Antwerp and Ghent, Flanders, Belgium, October 2009 to March 2010. *Euro Surveill*. 2011; 16(28).
21. Chalmet K, Staelens D, Blot S, Dinakis S, Pelgrom J, Plum J, et al. Epidemiological study of phylogenetic transmission clusters in a local HIV-1 epidemic reveals distinct differences between subtype B and non-B infections. *BMC Infect Dis*. 2010; 10:262. <https://doi.org/10.1186/1471-2334-10-262> PMID: 20822507
22. Struck D, Lawyer G, Ternes AM, Schmit JC, Bercoff DP. COMET: adaptive context-based modeling for ultrafast HIV-1 subtype identification. *Nucleic Acids Res*. 2014; 42(18):e144. <https://doi.org/10.1093/nar/gku739> PMID: 25120265
23. Vancoillie L, Mortier V, Demecheleer E, Schauvliege M, Vandekerckhove L, Vogelaers D, et al. Drug resistance is rarely the cause or consequence of long-term persistent low-level viraemia in HIV-1-infected patients on ART. *Antivir Ther*. 2015; 20(8):789–94. <https://doi.org/10.3851/IMP2966> PMID: 25945956
24. Hall T. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symposium Series*. 1999; 41:95–8.
25. Anisimova M, Gil M, Dufayard JF, Dessimoz C, Gascuel O. Survey of branch support methods demonstrates accuracy, power, and robustness of fast likelihood-based approximation schemes. *Syst Biol*. 2011; 60(5):685–99. <https://doi.org/10.1093/sysbio/syr041> PMID: 21540409
26. Kumar S, Stecher G, Tamura K. MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for Bigger Datasets. *Mol Biol Evol*. 2016; 33(7):1870–4. <https://doi.org/10.1093/molbev/msw054> PMID: 27004904
27. Lole KS, Bollinger RC, Paranjape RS, Gadkari D, Kulkarni SS, Novak NG, et al. Full-length human immunodeficiency virus type 1 genomes from subtype C-infected seroconverters in India, with evidence of intersubtype recombination. *J Virol*. 1999; 73(1):152–60. PMID: 9847317
28. Verhofstede C, Franssen K, Van Den Heuvel A, Van Laethem K, Ruelle J, Vancutsem E, et al. Decision tree for accurate infection timing in individuals newly diagnosed with HIV-1 infection. *BMC Infect Dis*. 2017; 17(1):738. <https://doi.org/10.1186/s12879-017-2850-6> PMID: 29187159
29. Koelsch KK, Smith DM, Little SJ, Ignacio CC, Macaranas TR, Brown AJ, et al. Clade B HIV-1 superinfection with wild-type virus after primary infection with drug-resistant clade B virus. *AIDS*. 2003; 17(7):F11–6. <https://doi.org/10.1097/01.aids.0000060361.78202.4c> PMID: 12700477
30. Cornelissen M, Pasternak AO, Grijsen ML, Zorgdrager F, Bakker M, Blom P, et al. HIV-1 dual infection is associated with faster CD4+ T-cell decline in a cohort of men with primary HIV infection. *Clin Infect Dis*. 2012; 54(4):539–47. <https://doi.org/10.1093/cid/cir849> PMID: 22157174

31. Wagner GA, Landais E, Caballero G, Phung P, Kosakovsky Pond SL, Poignard P, et al. Intrasubtype B HIV-1 Superinfection Correlates with Delayed Neutralizing Antibody Response. *J Virol*. 2017;91(17).
32. Redd AD, Mullis CE, Wendel SK, Sheward D, Martens C, Bruno D, et al. Limited HIV-1 superinfection in seroconverters from the CAPRISA 004 Microbicide Trial. *J Clin Microbiol*. 2014; 52(3):844–8. <https://doi.org/10.1128/JCM.03143-13> PMID: 24371237
33. Ronen K, Richardson BA, Graham SM, Jaoko W, Mandaliya K, McClelland RS, et al. HIV-1 superinfection is associated with an accelerated viral load increase but has a limited impact on disease progression. *AIDS*. 2014; 28(15):2281–6. <https://doi.org/10.1097/QAD.0000000000000422> PMID: 25102090
34. Pacold ME, Pond SL, Wagner GA, Delpont W, Bourque DL, Richman DD, et al. Clinical, virologic, and immunologic correlates of HIV-1 intraclade B dual infection among men who have sex with men. *AIDS*. 2012; 26(2):157–65. <https://doi.org/10.1097/QAD.0b013e32834dcd26> PMID: 22045341
35. Pernas M, Casado C, Fuentes R, Perez-Elias MJ, Lopez-Galindez C. A dual superinfection and recombination within HIV-1 subtype B 12 years after primoinfection. *J Acq Imm Def*. 2006; 42(1):12–8.
36. Kiwelu IE, Novitsky V, Margolin L, Baca J, Manongi R, Sam N, et al. HIV-1 subtypes and recombinants in Northern Tanzania: distribution of viral quasispecies. *PLoS One*. 2012; 7(10):e47605. <https://doi.org/10.1371/journal.pone.0047605> PMID: 23118882
37. Luan H, Han XX, Yu XO, An MH, Zhang H, Zhao B, et al. Dual Infection Contributes to Rapid Disease Progression in Men Who Have Sex With Men in China. *J Acq Imm Def*. 2017; 75(4):480–7.
38. Pacold M, Smith D, Little S, Cheng PM, Jordan P, Ignacio C, et al. Comparison of methods to detect HIV dual infection. *AIDS Res Hum Retroviruses*. 2010; 26(12):1291–8. <https://doi.org/10.1089/aid.2010.0042> PMID: 20954840
39. Redd AD, Mullis CE, Serwadda D, Kong X, Martens C, Ricklefs SM, et al. The rates of HIV superinfection and primary HIV incidence in a general population in Rakai, Uganda. *J Infect Dis*. 2012; 206(2):267–74. <https://doi.org/10.1093/infdis/jis325> PMID: 22675216
40. Novitsky V, Moyo S, Wang R, Gaseitsiwe S, Essex M. Deciphering Multiplicity of HIV-1C Infection: Transmission of Closely Related Multiple Viral Lineages. *PLoS One*. 2016; 11(11):e0166746. <https://doi.org/10.1371/journal.pone.0166746> PMID: 27893822
41. Piantadosi A, Ngayo MO, Chohan B, Overbaugh J. Examination of a second region of the HIV type 1 genome reveals additional cases of superinfection. *AIDS Res Hum Retroviruses*. 2008; 24(9):1221. <https://doi.org/10.1089/aid.2008.0100> PMID: 18729772