



OPEN ACCESS

ORIGINAL RESEARCH

Genetic diversity of HIV in seminal plasma remains higher than in blood after short-term antiretroviral therapy

Miguel Ángel López Zúñiga,¹ Natalia Chueca,^{2,3} Adolfo de Salazar ,² José Angel Fernández Caballero,² Alicia Gutierrez Valencia,⁴ David Vinuesa García,⁵ Mohamed Omar Mohamed Balgahata,⁶ Carmen Hidalgo Tenorio,⁷ Miguel Angel Lopez-Ruz,^{3,7} Federico Garcia  ^{2,3}

¹Internal Medicine Service, Jaen University Hospital, Jaen, Andalucía, Spain

²Microbiology Department, San Cecilio University Hospital, Granada, Spain

³Instituto de Investigación Biosanitaria Ibs.Granada, Granada, Spain

⁴Infection Disease Service, Virgen del Rocio University Hospital, Sevilla, Andalucía, Spain

⁵Infectious Disease Service, San Cecilio University Hospital, Granada, Spain

⁶Infectious Disease Department, Jaen University Hospital, Jaen, Spain

⁷Infectious Disease Service, University Hospital Virgen de las Nieves, Granada, Spain

Correspondence to

Dr Federico Garcia, Hospital Universitario San Cecilio, 18016 Granada, Spain; fegarcia@ugr.es

MAL-R and FG contributed equally.

Received 13 January 2020

Revised 10 March 2020

Accepted 17 March 2020

Published Online First

3 April 2020



© Author(s) (or their employer(s)) 2020. Re-use permitted under CC BY-NC. No commercial re-use. See rights and permissions. Published by BMJ.

To cite: López Zúñiga MÁ, Chueca N, de Salazar A, et al. *Sex Transm Infect* 2020;**96**:337–341.

ABSTRACT

Objective To provide insight on viral kinetics and genetic diversity of HIV in seminal plasma at baseline and 1 month after initiating antiretroviral therapy (ART).

Patients and methods Blood and seminal samples from patients with newly diagnosed HIV were obtained before ART initiation (T0) and 1 month after ART initiation (T1). HIV *env* genetic diversity was studied using deep sequencing Nextera and V3 chemistry in a MiSeq Illumina platform. The number of viral quasispecies (5% cut-off) and Shannon Index were used to analyse diversity.

Results Forty-seven ART-naive patients were recruited between September 2016 and November 2018. At enrolment, the number of quasispecies in blood (median 4 (IQR 2–5)) was lower than in the seminal compartment (median 6, (IQR 4–8)) ($p<0.01$); the Shannon Index was also higher ($p<0.001$) in the seminal compartment than in blood (1.77 vs 0.64). At T1, for the 13 patients with detectable HIV in both blood/seminal plasma, viral diversity remained higher ($p=0.139$) in seminal plasma (median 2 (IQR 1–4.5)) than in blood (median 1 (IQR 1–1.5)) Integrase inhibitors (INI)-based regimens achieved higher levels of undetectability and led more frequently to lower variability ($p<0.001$) than protease inhibitors (PI) or non-nucleoside reverse transcriptase inhibitors (NNRTI).

Conclusion We provide here further evidence of a larger genetic diversity in seminal plasma, both at diagnosis and short term after ART initiation. Our results strengthen previous findings on HIV diversity in seminal plasma. In addition, INIs decrease variability more rapidly than PI and NNRTI in both blood and seminal plasma.

INTRODUCTION

Infection with the HIV is one of the epidemics affecting the global population. Sexual route, followed by parenteral route, are the main transmission routes of HIV, and semen is the vector responsible for majority of sexual transmissions globally. Currently, quantification of viral load in blood is used in clinical practice to assess therapy effectiveness and disease control. However, it is known that the kinetics of the HIV viral replication in the blood compartment differs from its kinetics in the seminal

compartment. Several authors have reported the presence of viral DNA in seminal fluid with undetectable viral load in blood.^{1,2} The HIV DNA can be found in seminal leucocytes as non-integrated with linear or circular form or as integrated provirus.³ Nevertheless, cell-free HIV in seminal plasma seems to be actually infective¹ and infectivity is based on viral load. The presence of virus in this compartment can be intermittent, regardless of the time lapse from a negative result in the blood sample.⁴ Similarly, seminal positive result for HIV RNA was found despite a negative result for HIV RNA was found in blood.⁵

Gene sequencing of proviral protease and DNA envelope genes has shown different variants in seminal plasma and blood.¹

The presence of HIV in seminal plasma seems to depend on various factors that may influence its presence in this compartment, even when it is undetectable in blood. Thus, the presence of a symptomatic sexually transmitted disease, such as acute urethritis,⁶ cytomegalovirus infection⁷ or herpes simplex virus type 2,⁸ increases the risk of HIV transmission without showing association with previous or asymptomatic urethritis. Similarly, seropositivity to herpes simplex virus type 2, the increase in cytokines in seminal plasma, the increase in activation of mucosal T cells and the number of compartmentalised viral quasispecies¹ seem to influence the presence of the virus in seminal plasma.

Time to virological suppression varies based on the ART regimen and compartment; being integrase inhibitors (INIs) and rilpivirine-based ART regimens the best ones in seminal plasma^{9,10} followed by protease inhibitors (PI).

Despite the above described, the seminal compartment is not well understood to date. Therefore, in the present study, we investigated viral kinetics and quasispecies in patients with HIV-infection initiating antiretroviral therapy (ART).

PATIENTS AND METHODS

Inclusion criteria and ethics statement

The inclusion criteria of our study were: male sex, HIV infected, ART-naive and over 18 years of age

patients. All approached patients agreed to participate in the study. Before entering the cohort, ethics approval was obtained from the Ethics Committee of San Cecilio University Hospital in Granada (Spain) (1184N-16). Patients signed a written informed consent to participate in the study. No patient had other sexually transmitted diseases at the time of this study.

Methods

Blood and seminal samples were obtained before ART initiation (T0) and 1 month (T1) after ART initiation. At diagnosis, age, sex, concomitant sexually transmitted infections, HIV viral load (RNA copies/mL), baseline resistance and CD4 cell count (cells/ μ L) were recorded (retrieved from their medical records). Treatment regimen was recorded at the study entry (T0). HIV viral load, CD4 count and HIV env diversity were recorded at T0 and at T1.

All blood and seminal plasma were obtained by venipuncture and masturbation, respectively, and stored at -80°C until transfer to the laboratory. Quantitative nucleic acid test for use on the Cobas 6800/8800 Systems (Roche Diagnostics) with a limit of detection 20 copies/mL was used for HIV RNA viral load test in both blood and seminal plasma. The diversity of HIV was studied in all viraemic samples using an in house, deep sequencing protocol, using the Nextera and V3 chemistry in a MiSeq Illumina platform. We amplified a 500 bp of the HIV type 1 (HIV-1) *env* C2V3C3 region for these studies.¹¹ We selected *env* because this is the HIV gene with the largest variability.^{12–15} Briefly, HIV RNA was extracted from blood and seminal plasma samples using the Magna Pure Compact System (Roche) following the manufacturer's instructions. This system is based on a magnetic-bead technology. RNA was eluted into 50 μ L of the kit buffer from sample volumes of 400 or 1000 μ L according to availability. Seminal plasma was obtained from semen samples by centrifugation at 2500 rpm for 10 min. After extraction, RNA was purified using Ampure Agencourt RNAClean XP (Beckman Coulter) before cDNA synthesis. Quantification was performed with a NanoDrop ND-1000 spectrophotometer (Thermo Fisher Scientific, Delaware, USA). The nucleic acid yield was determined by measuring absorbance ratios spectrophotometrically. The measurement included A260/280 nm for protein contamination and A260/230 nm for salt and phenol contamination.

The Transcriptor One-Step RT-PCR (Roche) kit was used for cDNA synthesis. All PCRs were performed in 25 μ L reaction volumes containing 12.5 μ L 2X KAPA HiFi Hotstart ready mix (KAPA Biosystems, Woburn, Massachusetts, USA), 5 μ L of each forward and reverse primers (1 μ M) and 2.5 μ L of extracted RNA (10 ng). The PCR conditions were the following: initial denaturation at 95°C for 3 min, followed by cycles of denaturation at 95°C for 30s, annealing at 55°C for 30s and elongation at 72°C for 30s, with a final extension at 72°C for 5 min. PCR clean-up was performed using AMPure XP beads (Beckman Coulter, Indianapolis, Indiana, USA) to purify the *env* amplicon away from free primers and primer dimer species. Then, the next step was the index PCR, which attaches dual indices and Illumina sequencing adapters using the Nextera XT Index Kit (Illumina, San Diego, California, USA). The conditions of the second PCR were the following: 95°C for 3 min; 8 cycles of 95°C for 30s, 55°C for 30s, 72°C for 30s; 72°C for 5 min and hold at 4°C. The pooled PCR products were purified using AMPure XP beads (Beckman Coulter) before quantification. The resulting amplicons were sequenced at MiSeq Platform (Illumina), using a paired-end (2 \times 300 nt) sequencing system. All experiments were run on an equimolar concentration of amplicons.

The FASTA files obtained were demultiplexed to obtain a FASTA file for each sample and for the strand. Ugene tool¹⁶ was used for quality assessment of raw reads, which were assembled, aligned and screened based on the minimum length of 200 bp and a maximum length of 500 bp. Sequences were filtered and chimaera were removed using Usearch,¹⁷ accepting only those with $Q>30$ and a seed cluster of 30. Finally, the number of quasispecies in each sample was investigated using BLAST. Only the quasispecies representing more than 5% of the population were recorded; this limit was used to skew any method error, since this is the standard limit for resistance analysis.¹⁸

The diversity of the HIV quasispecies was calculated using the Shannon Index as previously reported¹⁹ using the entropy tool from Los Alamos HIV (https://www.hiv.lanl.gov/content/sequence/ENTROPY/entropy_one.html). This numerical score is a representative number and relative frequency of the HIV quasispecies found in each sample.

We were able to investigate HIV viral tropism using the HIV genotopheno tool²⁰ because our *env* fragment encompassed the V3 loop. The number and location of potential N-linked glycosylation sites for each sequence were estimated using the N-GLYCOSITE web tool from the Los Alamos HIV database (<http://www.hiv.lanl.gov/content/sequence/GLYCOSITE/glycosite.html>).

Normality of the variables has been analysed using the Shapiro-Wilks test for the statistical analysis. Baseline characteristics are expressed as proportions for categorical variables and as medians and IQR for continuous variables. Mann-Whitney U non-parametric test has been used to compare differences in the number of quasispecies between blood and seminal plasma. Wilcoxon test was used for the differences in the number of quasispecies before and after therapy. Wilcoxon test was applied, correcting the p value obtained by the number of comparisons, to analyse whether the decrease in the number of quasispecies was significant based on the type of regimen. Pearson χ^2 test was used for qualitative variables. To correlate the number of quasispecies with initial viral blood load (T0) and the nadir CD4 T-cell count, the Spearman's correlation coefficient was used. A $p<0.05$ was considered significant. Data have been analysed using the IBM SPSS software Statistics 19.

RESULTS

Forty-seven ART-naive patients with HIV-infection from the different participating locations (Virgen del Rocio University Hospital, Virgen de las Nieves University Hospital, San Cecilio University Hospital and Jaén University Hospital) were recruited for the study between September 2016 and November 2018. Patients had a median age of 30 years, a mean viral load of 29 2513 copies (log 4.56) and a mean CD4 count of 494 cells/ μ L. At enrolment (T0), 4 patients (8.51%) had clinical signs of acute HIV infection, 6 (12.8%) patients initiated a non-nucleoside reverse transcriptase inhibitors (NNRTI)-based regimen, 10 (21.3%) patients initiated a PI-based regimen and 31 (65.6%) patients initiated an INI-based regimen. Table 1 shows the baseline characteristics of the patients.

At baseline, we were able to sequence 47 blood and seminal samples. The mean number of sequences studied in blood was 68 458 (range 58 307–79 352). The mean number of sequences studied in seminal plasma was 65 435, (range 54 675–73 669). Using the prespecified 5% cut-off, the number of quasispecies in blood was lower than in the seminal compartment ($p<0.01$). In fact, a median of 4 (IQR 2–5) different quasispecies was found in the blood compartment, while a median of 6 (IQR 4–8), different

Table 1 Baseline characteristics of patients

	N=47
Age* (years)	30 (26–41)
Acute HIV Infection, n (%)	4 (8.51%)
Viral load (Log copies/mL)*	4.48 (3.98–5.06)
CD4 count (cells/ μ L)*	494.5 (301.75–641.5)
HIV Treatment, n (%)	
NNRTI-based	6 (12.8%)
PI-based	10 (21.3%)
INI-based	31 (65.6%)
Previous infections	
Chlamydia	4 (8.7%)
Syphilis	9 (19.1%)
Gonorrhoea	0 (0%)
HCV	2 (4.3%)
HPV	10 (21.3%)

*Median (IQR, 25th–75th percentile).

HCV, hepatitis C virus; HPV, human papillomavirus; INI, integrase inhibitors; NNRTI, non-nucleoside reverse transcriptase inhibitors; PI, protease inhibitors.

quasispecies were found in the seminal tract. All patients except one showed a higher number of quasispecies in seminal plasma than in the blood compartment. No statistically significant differences in viral diversity were found between patients with and without acute HIV infection. Similarly, the Shannon Index was higher ($p < 0.001$) in the seminal compartment than in blood (1.77 vs 0.64). Patients were infected predominantly with CCR5 viruses (80.9% in blood and 83% in seminal plasma). Interestingly, we found a different tropism from blood to seminal samples in two patients. These data are shown in [table 2](#).

Secondary analysis showed no relationship between viral diversity in blood and initial viral load ($p = 0.196$), nadir CD4 T-cell count ($p = 0.119$) or acute HIV infection ($p = 0.594$). Similarly, no relationship was found between acute HIV infection and viral diversity in seminal plasma ($p = 0.652$). Likewise, in 11 patients for whom we could analyse the seminal viral load, no relationship was found between viral diversity and the initial viral load ($p = 0.382$).

To further characterise the evolution of viral diversity in both the seminal and the systemic compartment, we studied *env* quasispecies in 33 patients for whom paired T0 and T1 samples from blood and seminal plasma were available. [Table 3](#) shows the number of viral quasispecies for each patient and the median of these values. At T1 (1 month after ART treatment), a significant decrease ($p < 0.001$) in both blood (T0: 3 (2–5) vs T1: 0 (0–1)) and seminal plasma (T0: 6 (4–7.5) vs T1: 0 (0–2)) was found. HIV-RNA was undetectable in both blood and seminal plasma in 16 patients (48.5%). Discordant results were found in 4/33 patients (12.1%) characterised by detectable HIV-1 RNA

Table 2 Virological characteristics in blood and seminal plasma

	Blood, n=47	Seminal plasma, n=47	P value
Quasispecies*	4 (2–5)	6 (4–8)	<0.001
Shannon Index*	0.64 (0.41–1.01)	1.77 (1.21–2.32)	<0.001
Tropism			0.789
CXCR4	9 (19.1%)	8 (17%)	
CCR5	38 (80.9%)	39 (83%)	
Sequences studied*	68 963 (57 699–79 352)	62 577 (54 335–73 669)	0.155

*Median (IQR, 25th–75th percentile).

CCR5, chemokine coreceptor type 5; CXCR4, chemokine receptor type 4.

Table 3 Number of viral quasispecies in each of the 33 patients from whom paired T0-baseline and T1-1 month after baseline-blood and seminal samples were available

Patient	T0		T1		Antiretroviral regimen
	Blood	Seminal plasma	Blood	Seminal plasma	
1	3	6	1	2	TDF+FTC+RPV
2	5	7	0	0	TDF+FTC+RPV
3	4	6	1	1	TDF+FTC+RPV
4	1	3	1	1	TDF+FTC+RPV
5	2	4	1	2	TDF+FTC+DRV/rtv
6	4	7	2	4	TDF+FTC+DRV/rtv
7	7	10	3	6	TDF+FTC+DRV/rtv
8	3	4	1	1	TDF+FTC+DRV/rtv
9	4	5	1	1	TDF+FTC+DRV/rtv
10	2	3	0	0	TDF+FTC+DRV/rtv
11	5	8	0	0	TDF+FTC+EVG/c
12*	2	4	0	0	TAF+FTC+EVG/c
13	2	3	0	0	TAF+FTC+EVG/c
14	3	6	0	0	TAF+FTC+EVG/c
15	2	3	1	2	TAF+FTC+EVG/c
16	1	2	0	0	TAF+FTC+EVG/c
17	3	6	1	2	TAF+FTC+EVG/c
18	2	7	0	0	TAF+FTC+EVG/c
19	8	11	0	6	TAF+FTC+EVG/c
20	4	2	0	1	TAF+FTC+EVG/c
21	3	5	0	0	TAF+FTC+EVG/c
22*	7	11	3	5	TAF+FTC+EVG/c
23*	3	6	0	5	TAF+FTC+EVG/c
24	2	4	1	6	TAF+FTC+EVG/c
25	5	7	0	3	TAF+FTC+DRV/c
26	6	11	0	0	3TC+ABC+DTG
27	3	6	0	0	3TC+ABC+DTG
28	3	5	0	0	3TC+ABC+DTG
29	5	8	0	0	3TC+ABC+DTG
30*	6	9	0	0	3TC+ABC+DTG
31	1	2	0	0	3TC+ABC+DTG
32	4	8	0	0	3TC+ABC+DTG
33	2	4	1	1	3TC+ABC+DTG

*Patients with acute infection. 0: undetectable viral load, no quasispecies were detected.

ABC, abacavir; DRV/rtv, darunavir/ritonavir; DTG, dolutegravir; DTG, dolutegravir; EVG/c, elvitegravir/cobicistat; FTC, emtricitabine; RPV, rilpivirine; TAF, tenofovir alafenamide; 3TC, lamivudine; TDF, tenofovir disoproxil fumarate.

in the seminal plasma but undetectable HIV-1 RNA in the blood plasma.

Patients treated with INI were more likely to reach undetectability in both blood and seminal plasma ($p < 0.001$) than patients treated with a PI or with an NNRTI-based regimen, even from viral blood loads (T0) higher than other regimens (INI 47 423 (4502–339 750), IP 8220 (4730–42 700), NNRTI 46 300 (32 050–83 650)). [Table 4](#) shows the quasispecies according to the prescribed treatment regimen regarding the 33 patients, we found that INI led to a decrease in the number of quasispecies in both samples ($p < 0.001$); this decrease was not found with PI or with NNRTI. Regarding the number of quasispecies at T1 in the 13 patients with positive viral load in seminal plasma and blood, viral diversity remained higher in seminal plasma (median 2 (IQR 1–4.5)) than in blood (median 1 (IQR 1–1.5)), $p = 0.139$.

Table 4 Number of quasispecies according to the prescribed treatment regimen

	N	T0*	T1*	
Blood				
INI	22	3 (2–5)	0 (0–0.25)	
PI	7	4 (2–5)	1 (0–2)	
NNRTI	4	3.5 (1.5–4.75)	1 (0.25–1)	
Seminal plasma				
INI	22	6 (3.75)	0 (0–2)	
PI	7	5 (4–7)	2 (0–4)	
NNRTI	4	6 (3.75–6.75)	1 (0.25–1.75)	
	N	T0*	T1*	P value
Blood				
INI	22	3 (2–5)	0 (0–0.25)	<0.001
PI	7	4 (2–5)	1 (0–2)	0.051
NNRTI	4	3.5 (1.5–4.75)	1 (0.25–1)	0.327
Seminal plasma				
INI	22	6 (3.75–8)	0 (0–2)	<0.001
PI	7	5 (4–7)	2 (0–4)	0.051
NNRTI	4	6 (3.75–6.75)	1 (0.25–1.75)	0.204

*Median (IQR, 25th–75th percentile).

INI, integrase inhibitors; NNRTI, non-nucleoside reverse transcriptase inhibitors; PI, protease inhibitors.

DISCUSSION

Understanding HIV viral genetic diversity in seminal plasma of patients with HIV infection is of utmost importance to elucidate HIV sexual transmission. To our knowledge, our study is one of the largest studies analysing paired blood and seminal plasma samples from ART-naïve patients with HIV-1 infection. We found a consistent larger number and variability of HIV-1 *env* viral quasispecies in seminal plasma than in blood by using deep sequencing. After starting first-line treatment, we also found that HIV-1 *env* viral diversity was higher in seminal plasma than in blood, and HIV remained detectable in seminal plasma in more patients than in blood.

Sexual contact is currently the main route of transmission of HIV-1. The male genital tract has been characterised as a reservoir,^{21 22} a compartment²³ and a drug sanctuary.²⁴ HIV compartmentalisation and diversity in seminal plasma has been previously studied using various methodologies. Ghosn *et al*²⁵ analysed the genotypic resistance pattern in patients with HIV established on ART with a median of seven previous antiretroviral treatments (range 3–15), showing different genotypic resistance patterns in the blood and genital compartments of six patients. Interestingly, phylogenetic analysis of clones of HIV protease gene showed that viral strains could originate from local production in seminal plasma. Pillai *et al*²⁶ analysed HIV-1 *env* sequences derived from paired blood and seminal plasma samples from Los Alamos HIV Sequence Database, suggesting the male genital tract as a selective environment contributing to the genetic bottlenecks associated with the sexual transmission of HIV-1.²⁶ In consistency with our study, Jiao *et al*²⁷ reported a higher genetic diversity of HIV-1 in the seminal compartments and blood of seven men who have sex with men (MSM) with early HIV-1 infection.

It is well known that HIV can remain positive in the seminal plasma after starting ART for longer periods than in blood.^{10 28} Du *et al*³ described how amounts of HIV RNA and HIV DNA remained high in seminal plasma of Chinese patients with HIV infection after 6 months undergoing an ART, even when HIV

RNA was undetectable in blood. In addition, we evaluated longitudinally the genetic diversity in the seminal plasma after 1 month of initiating first-line ART. To our knowledge, this is the first study reporting a higher genetic diversity in the seminal plasma even after initiation of first-line ART.

Similarly, several studies have addressed HIV-1 shedding in the rectal and vaginal tracts. Although using different methods and measures to address diversity, similarly to our study, HIV diversity was consistently higher in these compartments than in blood. García-Payá *et al*²⁹ reported a higher HIV-1 RNA rectal shedding in virologically suppressed MSM, suggesting that blood HIV-1 RNA measurements may not necessarily reflect the viral concentrations at the rectal reservoir. No certain explanation to the greater viral diversity in seminal plasma than in the blood is available to date; a possible cause could be a lower immune selection in the genitourinary tract compared with blood and, therefore, a lower purifying selection in the seminal plasma.

One of the major advantages of preventing HIV transmission are presented in the final report of the PARTNER study. In this study, Rodger *et al*³⁰ suggested that the risk of HIV transmission in gay couples through condomless sex when HIV viral load is suppressed is effectively zero. Our findings in seminal plasma, along with the above-mentioned studies conducted in seminal plasma and in rectal mucosa must be interpreted with caution.

Our study has several limitations, the short-term (1 month) follow-up being its main limitation. Although our study was longitudinal, our follow-up was only extended 1 month after ART initiation. Finding a larger variability in seminal plasma than in blood at this follow-up, despite being relevant, must not be interpreted as a caveat for HIV spread, especially after the results of the PARTNER study, where patients had been on ART for a median (IQR) of 4.3 (1.8–9.3) years. Together with the short-term follow-up, our sample size was also limited. Finally, although our study would have been strengthened if full-genome sequencing had been available, we chose the *env* gene, which in the absence of full-genome is the gold standard to study HIV variability.

Viral load in seminal plasma has been shown to become negative earlier after treatment with INI and rilpivirine than with PI.⁹ The influence of the various therapeutic regimens on variation in the number of quasispecies in blood or seminal plasma is still unknown. In our study, we found a significant decrease in the number of quasispecies in T1 compared with T0 in both blood and seminal plasma with INI, but no significant decrease was found with PI or with NNRTI.

In summary, we provide here strong evidence of a larger genetic variability in seminal plasma than in blood at diagnosis and short term after ART initiation. Our results strengthen previous findings on HIV diversity in seminal plasma. Furthermore, INIs decrease the number of quasispecies more efficiently than PI and NNRTI in both blood and seminal plasma.

Key messages

- ▶ There is a greater viral HIV diversity in seminal plasma than in blood before treatment.
- ▶ There is a greater viral HIV diversity in seminal plasma than in blood 1 month after treatment.
- ▶ Integrase inhibitors decrease the viral load in seminal plasma more rapidly than protease inhibitors or non-nucleosides, regardless of the viral load found before starting treatment.

Handling editor Anna Maria Geretti

Twitter Carmen Hidalgo Tenorio @carmenhtenorio

Acknowledgements This study is part of the Ph.D. thesis of M.A. López Zúñiga, University of Granada, in Clinical Medicine and Public Health. We acknowledge Nutraceutical Translations for English language editing of this manuscript.

Contributors MALZ: investigation, methodology, supervision, writing—original draft, writing—review and editing. NC: microbiological analysis, investigation. AdS: microbiological analysis, investigation. JAFc: microbiological analysis, investigation. AGV: microbiological analysis, investigation. DVG: investigation. MOMB: investigation. CHT: investigation. MALR: investigation, methodology, supervision, writing—original draft, writing—review and editing. FG: funding acquisition, investigation, methodology, supervision, writing—original draft, writing—review and editing.

Funding This paper has been funded in part by grants from Fondo de Investigación Sanitaria (www.isciii.es) (PI18/ 00819), Plan Nacional de I+D+I and Fondo Europeo de Desarrollo Regional-FEDER (www.redes/redes/inicio) (RD16/0025/0040).

Competing interests None declared.

Patient consent for publication Not required.

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement All data relevant to the study are included in the article or uploaded as supplementary information.

Open access This is an open access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited, appropriate credit is given, any changes made indicated, and the use is non-commercial. See: <http://creativecommons.org/licenses/by-nc/4.0/>.

ORCID iDs

Adolfo de Salazar <http://orcid.org/0000-0002-4318-858X>

Federico Garcia <http://orcid.org/0000-0001-7611-781X>

REFERENCES

- Gupta P, Leroux C, Patterson BK, *et al.* Human immunodeficiency virus type 1 shedding pattern in semen correlates with the compartmentalization of viral quasispecies between blood and semen. *J Infect Dis* 2000;182:79–87.
- Chéret A, Durier C, Mélard A, *et al.* Impact of early cART on HIV blood and semen compartments at the time of primary infection. *PLoS One* 2017;12:e0180191.
- Du P, Liu A, Jiao Y, *et al.* Hiv RNA and proviral HIV DNA can be detected in semen after 6 months of antiretroviral therapy although HIV RNA is undetectable in blood. *Microbiol Immunol* 2016;60:187–95.
- Ferraretto X, Estellat C, Damond F, *et al.* Timing of intermittent seminal HIV-1 plasma viral load under combination antiretroviral therapy. *Plos One* 2014;9:88922.
- López-Ruz MA, Navas P, López-Zúñiga MA, *et al.* Effect of monotherapy with Darunavir/Ritonavir on viral load in seminal fluid, and quality parameters of semen in HIV-1-positive patients. *PLoS One* 2016;11:e0159305.
- Laga M, Manoka A, Kivuvu M, *et al.* Non-Ulcerative sexually transmitted diseases as risk factors for HIV-1 transmission in women: results from a cohort study. *AIDS* 1993;7:95–102.
- Gianella S, Smith DM, Vargas MV, *et al.* Shedding of HIV and human herpesviruses in the semen of effectively treated HIV-1-infected men who have sex with men. *Clin Infect Dis* 2013;57:441–7.
- Looker KJ, Elmes JAR, Gottlieb SL, *et al.* Effect of HSV-2 infection on subsequent HIV acquisition: an updated systematic review and meta-analysis. *Lancet Infect Dis* 2017;17:1303–16.
- Gutiérrez Valencia A, Benmarzouk-Hidalgo OJ, Rivas-Jeremías I, *et al.* Viral kinetics in semen with different antiretroviral families in treatment-naive human immunodeficiency virus-infected patients: a randomized trial. *CID* 2017;65:551–6.
- Fabrizio C, de Gennaro N, Volpe A, *et al.* HIV-RNA decay in paired blood and semen samples of subjects receiving their first dolutegravir-based art regimen. *Journal of Clinical Virology* 2018;109:45–9.
- Villarreal J-L, Gutiérrez J, Palacio L, *et al.* Characterization of HIV type 1 envelope sequence among viral isolates circulating in the Northern region of Colombia, South America. *AIDS Res Hum Retroviruses* 2012;28:1779–83.
- Rao M, Peachman KK, Kim J, *et al.* HIV-1 variable loop 2 and its importance in HIV-1 infection and vaccine development. *Curr HIV Res* 2013;11:427–38.
- deCamp A, Hraber P, Bailer RT, *et al.* Global panel of HIV-1 env reference strains for standardized assessments of vaccine-elicited neutralizing antibodies. *J Virol* 2014;88:2489–507.
- Kitawi RC, Hunja CW, Aman R, *et al.* Partial HIV C2V3 envelope sequence analysis reveals association of coreceptor tropism, envelope glycosylation and viral genotypic variability among Kenyan patients on HAART. *Viral J* 2017;14:29.
- Hamoudi M, Simon-Lorieri E, Gasser R, *et al.* Genetic diversity of the highly variable V1 region interferes with human immunodeficiency virus type 1 envelope functionality. *Retrovirology* 2013;10:114.
- Okonechnikov K, Golosova O, Fursov M, *et al.* Unipro UGENE: a unified bioinformatics toolkit. *Bioinformatics* 2012;28:1166–7.
- Edgar RC, Flyvbjerg H. Error filtering, pair assembly and error correction for next-generation sequencing reads. *Bioinformatics* 2015;31:3476–82.
- Herbert A, Mbunah HA, Bertagnolio S, *et al.* Low-Abundance drug-resistant HIV-1 variants in antiretroviral drug-naïve individuals: a systematic review of detection methods, prevalence, and clinical impact. *J Infect Dis* 2019;pii: jiz650.
- Hoffman NG, Seillier-Moiseiwitsch F, Ahn J, *et al.* Variability in the human immunodeficiency virus type 1 gp120 Env protein linked to phenotype-associated changes in the V3 loop. *J Virol* 2002;76:3852–64.
- Lengauer T, Sierra S, *et al.* Bioinformatics prediction of HIV coreceptor usage. *Nat Biotechnol* 2007;25:1407–10.
- Smith DM, Kingery JD, Wong JK, *et al.* The prostate as a reservoir for HIV-1. *AIDS* 2004;18:1600–2.
- Zhang H, Dornadula G, Beumont M, *et al.* Human immunodeficiency virus type 1 in the semen of men receiving highly active antiretroviral therapy. *N Engl J Med* 1998;339:1803–9.
- Delwart EL, Mullins JL, Gupta P, *et al.* Human immunodeficiency virus type 1 populations in blood and semen. *J Virol* 1998;72:617–23.
- Taylor S, van Heeswijk RP, Hoetelmans RM, *et al.* Concentrations of nevirapine, lamivudine and stavudine in semen of HIV-1-infected men. *AIDS* 2000;14:1979–84.
- Ghosn J, Viard J-P, Katlama C, *et al.* Evidence of genotypic resistance diversity of archived and circulating viral strains in blood and semen of pre-treated HIV-infected men. *AIDS* 2004;18:447–57.
- Pillai SK, Good B, Pond SK, *et al.* Semen-specific genetic characteristics of human immunodeficiency virus type 1 env. *J Virol* 2005;79:1734–42.
- Jiao Y-M, Chen G-L, Zhu W-J, *et al.* Higher viral load and genetic diversity of HIV-1 in seminal compartments than in blood of seven Chinese men who have sex with men and have early HIV-1 infection. *Microbiol Immunol* 2017;61:239–46.
- Fernández-González M, García JA, Padilla S, *et al.* Rectal and seminal HIV-1 RNA decay towards virological suppression in infected MSM initiating dolutegravir/abacavir/lamivudine. *J Antimicrob Chemother* 2020;75:668–74.
- García-Payá E, Fernández M, Padilla S, *et al.* High-Grade anal intraepithelial neoplasia is associated with HIV-1 RNA rectal shedding in virologically suppressed MSM. *AIDS* 2018;32:1017–24.
- Rodger AJ, Cambiano V, Bruun T, *et al.* Risk of HIV transmission through condomless sex in serodifferent gay couples with the HIV-positive partner taking suppressive antiretroviral therapy (partner): final results of a multicentre, prospective, observational study. *Lancet* 2019;393:2428–38.