Published in final edited form as:

J Med Virol. 2012 November ; 84(11): 1703–1709. doi:10.1002/jmv.23398.

Impact of Seminal Cytomegalovirus Replication on HIV-1 Dynamics between Blood and Semen

Sara Gianella¹, Sanjay R. Mehta¹, Matthew C. Strain¹, Jason A. Young¹, Milenka V. Vargas¹, Susan J. Little¹, Douglas D. Richman^{1,2}, Sergei L. Kosakovsky Pond¹, and Davey M. Smith^{1,2}

¹University of California San Diego

²Veterans Affairs San Diego Healthcare System, San Diego

Abstract

The genital tract of individuals infected with HIV-1 is an anatomic compartment that supports local HIV-1 and CMV replication. This study investigated the association of seminal CMV replication with changes in HIV-1 clonal expansion, evolution and phylogenetic compartmentalization between blood and semen. Fourteen paired blood and semen samples were analyzed from four untreated subjects. Clonal sequences (n=607) were generated from extracted HIV-1 RNA (*env* C2-V3 region), and HIV-1 and CMV levels were measured in the seminal plasma by real-time PCR. Sequence alignments were evaluated for: (i) viral compartmentalization between semen and blood samples using Slatkin-Maddison and F_{ST} methods, (ii) different nucleotide substitution rates in semen and blood, and (iii) association between proportions of clonal HIV-1 sequences in each compartment and seminal CMV levels. Half of the semen samples had detectable CMV DNA, with at least one CMV positive sample for each patient. Seminal CMV DNA levels correlated positively with seminal HIV-1 RNA levels (Spearman $p=0.05$). A trend towards an association between compartmentalization of HIV-1 sequences sampled from blood and semen and presence of seminal CMV was observed (Cochran Q test $p=0.12$). Evolutionary rates between semen and blood HIV-1 populations did not differ significantly, and there was no significant association between seminal CMV DNA levels and the frequency of non-unique clonal HIV-1 sequences in the semen. In conclusion, the effects of CMV replication on HIV-1 viral and immunologic dynamics within the male genital tract are not significant enough to perturb evolution or disrupt compartmentalization in the genital tract.

*Corresponding author: Sara Gianella, University of California San Diego, 9500 Gilman Drive, MC 0679, La Jolla, CA 92093-0679, USA, Tel: (858) 552-8585 #2624; Fax: (858) 552-7445; gianella@ucsd.edu.

Author Contributions

SG participated in the study design, performed laboratory experiments and data analyses, and wrote the first version of the manuscript; SRM, MAS and JAY participated in data analysis, and revised the manuscript; MVV performed laboratory experiments, SJL and DMS enrolled patients and revised the manuscript; SKP, DDR, SJL and DMS designed the study, participated in data analysis and revised the manuscript. All authors read and approved the final manuscript.

Financial Disclosure

This work was supported by grants from the US National Institutes of Health AI69432, AI043638, MH62512, MH083552, AI077304, AI36214, AI047745, AI74621, GM093939 and AI080353, the James Pendleton Trust, and the Swiss National Science Foundation grant PBZHP3-125533. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests

SG does not have any commercial or other associations that might pose a conflict of interest. DDR has served as a consultant for Biota, Bristol-Myers Squibb, Chimerix, Gen-Probe, Gilead Sciences, J & J, Merck & Co, Monogram Biosciences, Tobira Therapeutics, and Vertex. DMS has received research support from ViiV Pharmaceuticals and has served as a consultant to Gen-Probe. SKP has served as a consultant to Gen-Probe. The remaining authors do not report any conflicts of interest.

Keywords

HIV-1; Cytomegalovirus; compartmentalization; evolution; semen

Introduction

Following the initial dissemination of human immunodeficiency virus type 1 (HIV-1) during primary infection, the virus begins to evolve rapidly. Differences in immunological surveillance, target cell characteristics, and efficiencies of drug penetration within different anatomic compartments, e.g. genital tract, central nervous system (CNS), may select for unique characteristics in response to local selective pressures [Wong et al., 1997; Delwart et al., 1998; Gupta et al., 2000; Günthard et al., 2001; Paranjpe et al., 2002; Pillai et al., 2005; Zarate et al., 2007]. For example, although levels of HIV-1 RNA in peripheral blood can correlate with those in seminal plasma [Vernazza et al., 1997; Speck et al., 1999; Pinto-Neto et al., 2002; Butler et al., 2008; Kalichman et al., 2008; Baeten et al., 2011], viral sequences from these two anatomical compartments frequently form segregated populations [Smith et al., 2004; Blackard, 2012]. The lack of genetic flow between viral populations in different tissues or locations is known as compartmentalization [Nickle et al., 2003]. Most research in HIV-1 pathogenesis has focused on the peripheral blood compartment, despite the fact that the majority of new infections occur by mucosal exposure to HIV-1 contained in genital secretions [Piot et al., 2001; Marks et al., 2006; UNAIDS, 2009]. Therefore, understanding the dynamics of HIV-1 in semen and characterizing the contribution of co-factors that increase viral load and sequence heterogeneity in the genital tract is critical for the rational design of successful prevention strategies.

A recent paper reported that the level of CMV DNA detected in semen is associated with the absolute number and activation state of CD4+ T-lymphocytes found in the same compartment [Gianella et al., 2012b]. Previous work has also demonstrated that levels of CMV DNA detected in semen correlate positively with the HIV load within that compartment [Speck et al., 1999; Sheth et al., 2006; Gianella et al., 2012a]. In the CNS, the presence of inflammation (indicated by pleocytosis) is associated with disruption of the compartmentalization between HIV-1 populations from the CNS and the peripheral blood [Smith et al., 2009]. Similarly, seminal CMV replication could cause local inflammatory reaction with consequent trafficking of immune cells carrying virus from the blood into the genital tract. These incoming cells could eventually release virus locally in the genital tract, which would disrupt of the compartmentalization of HIV-1. Alternatively, inflammation in the genital tract may induce replication of virus within cells infected with HIV-1 already present within this compartment, which would result in divergence between genital and blood HIV-subpopulations.

The present study evaluated whether seminal CMV replication and the described associated genital tract inflammation influenced HIV-1 compartmentalization, the rate of viral evolution and clonal expansion in the genital compartments of men infected with HIV-1.

Materials and Methods

Ethics Statement

These studies were conducted with appropriate subject consent and were approved by the Human Research Protections Program at the University of California, San Diego, CA, USA. All study participants were provided written informed consent.

Study Participants and Clinical Data

Four antiretroviral naïve participants infected recently with HIV-1 were selected based on longitudinal measurements of CMV DNA levels in seminal plasma, i.e. having at least one seminal sample positive and one negative for detectable CMV DNA [Gianella et al., 2012b]. Fourteen paired blood and seminal samples were collected at 3 or 4 different time points from each participant. Blood plasma and peripheral blood mononuclear cells (PBMC) samples were separated as previously described [Butler et al., 2010], aliquoted, frozen, and stored at -80°C and -150°C , respectively. Semen was collected by masturbation without lubricant after 48 hours of abstinence. Viral transport medium (2 ml of RPMI 1640 with 2mMol L-glutamine and 10% fetal bovine serum, with addition of 100U/ml of penicillin, 100 $\mu\text{U}/\text{ml}$ of streptomycin, and 200U/ml of nystatin) was added to seminal samples at collection. Seminal plasma was separated from seminal cells by centrifugation at 700g for 12 min within 4 hours of collection and stored at -80°C and -150°C , as described previously [Smith et al., 2004; Butler et al., 2010]. Screening for *Neisseria gonorrhoeae* and *Chlamydia trachomatis* genital tract infection was performed with PCR of urine samples, and syphilis infection was evaluated by rapid plasma reagin titers in blood (RPR).

HIV-1 RNA in peripheral blood was quantified using the Amplicor HIV Monitor assay (Roche Molecular Systems Inc.). Clinical data, including baseline demographics, symptoms and resolution of sexually transmitted infections (STI), and standard laboratory values, were collected. HIV-1 subtype was determined using HIV-1 *pol* sequence data generated by Viroseq 2.0 (Applied Biosystems) using SCUEAL [<http://www.datamoney.org/>], [Kosakovsky Pond et al., 2009]].

Nucleic Acid Extraction, and HIV-1 RNA and CMV DNA quantification

RNA and DNA for viral quantification were extracted from 500 μL (for RNA) of blood and seminal plasma and 200 μL (for DNA) of seminal plasma using High Pure Viral RNA Kit (Roche, Switzerland) and QIAamp DNA Mini Kit (Qiagen, CA) respectively, according to manufacturers' protocol, as described previously [Gianella et al., 2012b]. Using the extracted RNA, HIV-1 cDNA was generated using the SuperScript III First-Strand Synthesis Kit (Invitrogen, CA) according to manufacturer's protocol, as described previously [Gianella et al., 2012b]. HIV-1 RNA as well as CMV DNA in seminal plasma were quantified by real-time PCR in an ABI 7900HT thermocycler (Applied Biosystems, CA) with 0.005 μM ROX as passive reference, as described previously [Gianella et al., 2012b].

Sequencing of HIV-1 *env*

Reverse transcription and nested PCR amplification of the coding region of HIV-1 *env* C2-V3 from HIV-1 RNA extracted from each blood and seminal sample were performed in triplicate using RETROscript (Applied Biosystems, CA) and Taq DNA polymerase (Roche, Switzerland) according to manufacturer's protocol in a 50 μl reaction volume and using primers V3Fout and V3Bout, as described previously [Gunthard et al., 1999]; subsequently 5 μl of first step PCR product was used in the second, nested PCR with primers V3Fin and V3Bin in a total volume of 50 μl [Gunthard et al., 1999]. Final nested PCR products were visualized by agarose gel electrophoresis and ethidium bromide staining. Replicates of final nested PCR products were pooled (proportionally to the band intensity) and cloned using the TOPO TATM cloning system (Invitrogen, CA). Twenty to thirty clones were selected by blue-white screening and expanded in 1.5 ml broth cultures [Gunthard et al., 1999]. Plasmid clones were purified using Qiagen Mini-prep kits (Qiagen, CA). Purified plasmids were sequenced bi-directionally with -20M13 primer (5'-sequence-3') or TOPO Forward primer (5'-sequence-3') using Prism Dye terminator kits (ABI, Foster City, CA) on an ABI 3100 Genetic Analyzer. Sequences were initially compiled, aligned, and manually edited using SequencherTM 4.0 (Genecodes, Ann Arbor, MI). Preliminary alignments were then

performed using the MUSCLE package [Edgar, 2004a; Edgar, 2004b]. Sequences were additionally codon-aligned [Kosakovsky Pond et al., 2009] to the HXB2 reference to ensure that each clone encoded a productive protein homologous to HXB2 C2-V3 *env*. Sequences from individual patients have been deposited in Genbank (P9: accession numbers JN393317- JN393413 and JN393423-JN393504; L3: JN886090-JN886221, L7: JN886222-JN886300 and JN886323-JN886407, X3: JN886408-JN886476 and JN886499-JN886563).

Quality control

To avoid cross-contamination, extraction and PCR procedures were performed in separate rooms and on separate days with different sets of micropipettes in laminar flow cabinets. For every experiment, negative controls were included to monitor cross-contamination. Sequences were checked for inter-sample and laboratory strain contamination by performing MEGABLAST homology searches against each other and against the online public Los Alamos HIV sequence database (<http://www.hiv.lanl.gov>; accessed August 2011).

Phylogenetic analysis

The impact of CMV replication on HIV-1 dynamics and evolution within the genital tract was evaluated using (i) compartmentalization analysis, (ii) evolutionary rate analysis, and (iii) clonal expansion analysis. The results were interpreted in relation to the presence or absence of CMV in the seminal compartment. All analyses were performed in the HyPhy package [Pond et al., 2005].

Compartmentalization analysis—Assessment for compartmentalization was performed using two different statistical approaches: the Slatkin-Maddison (SM) test [Slatkin and Maddison, 1989] and the Wright measure of population subdivision (F_{st}) [Zarate et al., 2007]. When present, multiple copies of the same viral genotype isolated from the same compartment at the same time point, may bias intra-compartment viral populations towards homogeneity and create artificial signal for compartmentalization. All such copies were represented by a single sequence in order to make compartmentalization analyses more conservative. The SM test is a tree-based method for evaluating compartmentalization. Briefly, the minimum number of inter-compartment migration events required to form the observed tree were inferred (by parsimony) using the phylogenetic tree inferred (by neighbor joining, using the Tamura Nei, TN93, distance) from each individual's *env* sequences and characterized by compartment of origin. This result was compared to the number of migration events for 1000 randomizations of compartment labels on the same tree. Evidence of restricted gene flow (compartmentalization) was documented (at $p = 0.05$) when <5% of the replicates required as many or fewer migration events as in the observed tree. F_{st} testing is a distance-based method for evaluating compartmentalization [Hudson et al., 1992a; Hudson et al., 1992b], and the procedure compares the mean pairwise genetic distance (TN93) between sequences sampled within a compartment to the mean pairwise genetic distance calculated from sequences from both of the compartments. Statistical significance is derived by performing a population structure randomization test. The time point was defined as compartmentalized if both tests showed significant results, i.e. p -value 0.05.

Evolutionary rate analysis—Mean pairwise genetic distances of viral populations sampled from blood and seminal plasma samples over time for each subject were calculated using HyPhy [Pond et al., 2005], by considering path lengths between all pairs of sequences in the maximum likelihood tree. Confidence intervals were obtained by profile likelihood as described previously [Noviello et al., 2007]. Because the chronological time of sampling was identical between compartments, it was sufficient to compare genetic distances (the

products of mean evolutionary rates and chronological times) directly, without the need to fit more complex models, such as relaxed molecular clocks.

Clonal expansion analysis—Proportions of unique variants among all sampled clones were estimated by dividing the number of unique sequences found in each compartment for each time point by the total number of sequences obtained. This additional analysis was performed to investigate a possible relationship between CMV replication and monoclonal viral expansion within the genital tract.

Statistics

Effect size calculations were performed using an online sample size estimation tool [Eng, 2003]. Comparisons between groups (CMV DNA positive and negative) were performed using Cochran Q test or Mann-Whitney test, while correlation analyses were performed using a univariate non-parametric rank correlation test (Spearman). Level of significance was $p = 0.05$ for each performed analysis.

Results

Participant demographics, clinical factors and viral loads

All study participants were antiretroviral-naïve and infected with HIV-1 subtype B virus, and reported sex with other men as their HIV risk factor. All were Caucasian men, with one reporting Hispanic ethnicity. Median age at baseline was 33 years (range: 30–36). From these participants, longitudinal paired semen and blood samples (three to four time points) were collected over a median follow-up period of 133 days (range: 70–1155 days). The median estimated duration of infection (EDI) at baseline was 156 days (range: 97 to 194 days). Median CD4 count was 804 cells/ml (range: 331–1380 cells/ml) and the median blood plasma HIV-1 level was 4.7 HIV-1 RNA $\log_{10}$ copies/ml (range: 3.9–5.1 HIV RNA $\log_{10}$ copies/ml). The median seminal plasma HIV-1 level was 3.45 HIV RNA $\log_{10}$ copies/ml (range: 1.21–5.78 HIV-1 RNA $\log_{10}$ copies/ml). All participants infected with HIV-1 had positive CMV serology and undetectable levels of CMV in blood plasma. In a univariate non-parametric analysis of the samples, the level of CMV DNA present in the semen positively correlated with the level of HIV-1 RNA ($p=0.05$, $r=0.53$) found in the seminal fluid. One participant (patient X3) tested positive for syphilis at the time of his third sample collection and was treated. Detailed viral characteristics in semen and blood compartments (HIV-1 RNA and CMV DNA levels) are summarized in supplementary table 1.

Sequence Analysis

A median of 22 *env* C2V3 sequences (range: 18–30 sequences) for each timepoint were generated for both semen and blood samples with a median length of 400 nucleotides (range: 248–414 nucleotides) for a total of 607 sequences. Inferred phylogenetic trees are provided in supplementary figure 1.

Compartmentalization analysis—Seven out of 14 paired blood and semen samples were compartmentalized (both SM and F_{st} tests $p<0.05$) demonstrating genetic segregation of viral populations within the blood and male genital tract. These were patient L7 at time points 1, 2 and 3; patient P9 at time points 1, 3 and 4; and patient L3 at time point 3 (see table 1). Compartmentalization was not inferred for patient X3 at any of the sampled time points. The results of the two tests (SM and F_{st}) were mostly consistent except for time point 1 for patient L3. Here, the SM formally indicated compartmentalization ($p=0.03$) while F_{st} did not ($p=0.47$). Per *a priori* definition the time point was treated as not compartmentalized.

Five out of seven compartmentalized samples had detectable CMV DNA, while only two of the seven non-compartmentalized samples presented positive seminal CMV DNA. Although not significant (Cochran Q test, $p=0.12$), in this small sample, there was a trend for CMV DNA positive samples to demonstrate more HIV-1 compartmentalization than CMV DNA negative samples.

Evolutionary rate analysis—Differences in viral evolutionary rate were evaluated within the two compartments (i.e. blood and genital tract) using mean pairwise sequence divergence analysis. No difference in evolutionary rates between semen and blood was observed, with an average of 0.013 for blood and 0.014 expected nucleotide substitutions per site for semen and overlapping confidence intervals for each individual analyzed (table 2).

Clonal expansion analysis—To investigate whether presence of CMV replication was associated with increased clonal viral expansion, the proportions of unique viral sequences for each time point were tabulated (see supplementary table 2). In blood plasma a median of 59.2% (range: 36.4–100%) unique sequences was found, while in semen there was a median of 46.6% (range: 15–89.5%) unique sequences. There was an association between presence of CMV DNA and the proportion of unique sequences ($p<0.05$ in blood and $p=0.02$ in semen by Mann-Whitney), but this analysis was confounded by the strong association between time since EDI and the proportion of unique sequences ($p<0.001$ and $r=0.79$ for semen; $p<0.05$ and $r=0.55$ for blood). The association between EDI and diversity has been previously reported in much larger cohorts in which the effect is presumably unrelated to CMV [Wei et al., 1995; Shankarappa et al., 1999; Rieder et al., 2011]. Thus the data here do not provide evidence for an association between CMV and unique sequences. In fact, in patient L7 and X3, where CMV DNA was intermittently detected only in earlier time-points, the proportion of unique sequences in semen still increased over time independently of CMV replication (with exception of time-point 4 for patient L7).

Retrospective Power Calculation

As a consequence of the limited number of samples/participants included in this study, the analysis did not reach statistical significance, despite a difference in the compartmentalization pattern between samples with and without detectable CMV replication. A retrospective power analysis was performed to estimate the theoretical effect size needed in order to achieve significant results (with 5% significance level and 80% power). Using SM and F_{st} tests to analyze our data, a theoretical effect size of 38 different time points was estimated (assuming 72% versus 28% compartmentalized time points for CMV positive and negative, respectively). This power calculation suggested that a significant difference in compartmentalization would potentially be observed by including a minimum of 24 additional time points possibly from different subjects.

Discussion

A recent study from our group demonstrated that detectable CMV DNA in the male genital tract was associated with increased numbers of T-lymphocytes and increased immune activation of these cells in semen [Gianella et al., 2012b]. In the present dataset, levels of CMV DNA in the semen also correlated with levels of seminal HIV-1 RNA, as described in similar works [Speck et al., 1999; Sheth et al., 2006], but no positive correlation could be found between levels of HIV-1 RNA detected in semen and in blood.

The inferred lack of association between blood and seminal HIV-1 RNA levels is consistent with the hypothesis that the male genital tract has a unique immunologic milieu and often

harbors a distinct HIV-1 subpopulation. This study aimed to determine if the presence of CMV replication in the male genital tract were associated with the disruption of this compartmentalization, causing the intermixing of cells infected with HIV-1 and/or circulating cell-free virus between the peripheral blood and the genital tract, as described previously by our group for the CNS [Smith et al., 2009]. To this end, cell-free HIV-1 RNA extracted from 14 paired semen and blood samples with (n=7) and without (n=7) detectable seminal CMV DNA was sequenced, and multiple analyses to assess the level of compartmentalization between the two viral populations were performed. Whether the presence of CMV replication in seminal fluid had an impact on HIV-1 evolution or clonal viral expansion in the genital tract was investigated.

Our analysis suggested that the presence of CMV replication in semen was not associated with the disruption of HIV-1 compartmentalization between semen and blood. In fact, in those samples with detectable CMV DNA, HIV-1 compartmentalization was more likely to be observed: 72% versus 28% compartmentalized time points for CMV DNA positive and negative samples, respectively (Cochran Q $p=0.12$). Increased compartmentalization between HIV-1 sub-populations suggests that compartment-derived virus was more likely to be produced locally; therefore the inference that seminal samples with positive CMV DNA were more likely associated with compartmentalization means that the presence of CMV in the male genital tract enhances local HIV-1 replication. This observation is supported by a positive association between CMV DNA and HIV RNA levels.

Subsequently, the impact of seminal CMV replication on viral evolution or clonal expansion within the genital compartment was investigated. No significant difference in the mean evolutionary rate between seminal and blood viral population was observed. Also, no association between CMV replication and the extent of clonal viral expansion within the genital tract was found. These results suggested that, although the presence of CMV in the genital tract was associated with increased numbers of lymphocytes and increased activation rates of circulating CD4+ T-cells [Gianella et al., 2012b], the immune response to CMV in the genital tract was neither significant enough to disrupt its status as a separate compartment nor to perturb HIV-1 evolution or enhance clonal expansion.

The primary limitation of this study was its small sample size. Although 14 paired blood and semen samples were analyzed for HIV-1 RNA and CMV DNA levels, and a total of 607 *env* C2V3 sequences were generated, the sample size was still too small to achieve statistical significance assuming the observed effect size between detectable CMV in seminal plasma and HIV-1 compartmentalization, viral evolution and clonal expansion. Moreover, since this study was conducted on a small number of early HIV-1 infected subjects with relatively preserved immunity, our observation can not be generalized to other populations like chronically infected or subjects with lower CD4+ T-cell count

These results suggest that the presence of detectable CMV replication in the male genital tract is likely associated with local immune activation and enhanced HIV-1 replication with a subsequent increased compartmentalization. As a consequence of the high prevalence of CMV shedding in individuals infected with HIV-1 [Lang and Kummer, 1972; Lange et al., 1984; Bresson et al., 2003], the connections between CMV and HIV-1 are likely important, but the impact of local CMV replication and immune activation are likely to be clinically subtle and may not have much effect on the evolution of HIV-1 within the genital tract.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We are grateful to all the participants in the San Diego Primary Infection Cohort, to Caroline Ignacio for excellent technical support, to Steffney Rought and Celsa Spina for helpful discussion. We also thank and commemorate our dear friend and outstanding research colleague, Marek Fischer, for all his contributions and support to our research over many years.

References

- Baeten JM, Kahle E, Lingappa JR, Coombs RW, Delany-Moretlwe S, Nakku-Joloba E, Mugo NR, Wald A, Corey L, Donnell D, Campbell MS, Mullins JI, Celum C. Genital HIV-1 RNA predicts risk of heterosexual HIV-1 transmission. *Sci Transl Med*. 2011; 3:77ra29.
- Blackard, JT. Current HIV research. 2012. HIV compartmentalization: A Review On A Clinically Important Phenomenon.
- Bresson JL, Clavequin MC, Mazon MC, Mengelle C, Scieux C, Segondy M, Houhou N. Risk of cytomegalovirus transmission by cryopreserved semen: a study of 635 semen samples from 231 donors. *Hum Reprod*. 2003; 18:1881–1886. [PubMed: 12923143]
- Butler DM, Delport W, Kosakovsky Pond SL, Lakdawala MK, Cheng PM, Little SJ, Richman DD, Smith DM. The origins of sexually transmitted HIV among men who have sex with men. *Sci Transl Med*. 2010; 2:18re11.
- Butler DM, Smith DM, Cachay ER, Hightower GK, Nugent CT, Richman DD, Little SJ. Herpes simplex virus 2 serostatus and viral loads of HIV-1 in blood and semen as risk factors for HIV transmission among men who have sex with men. *AIDS*. 2008; 22:1667–1671. [PubMed: 18670228]
- Delwart EL, Mullins JI, Gupta P, Learn GH Jr, Holodniy M, Katzenstein D, Walker BD, Singh MK. Human immunodeficiency virus type 1 populations in blood and semen. *J Virol*. 1998; 72:617–623. [PubMed: 9420266]
- Edgar RC. MUSCLE: a multiple sequence alignment method with reduced time and space complexity. *BMC Bioinformatics*. 2004a; 5:113. [PubMed: 15318951]
- Edgar RC. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic acids research*. 2004b; 32:1792–1797. [PubMed: 15034147]
- Eng J. Sample size estimation: how many individuals should be studied? *Radiology*. 2003; 227:309–313. [PubMed: 12732691]
- Gianella S, Morris SR, Anderson C, Spina CA, Vargas MV, Young JA, Richman DD, Little SJ, Smith DM. Herpesviruses and HIV-1 Drug Resistance Mutations Influence the Virologic and Immunologic Milieu of the Male Genital Tract. *AIDS*. 2012a In Press.
- Gianella S, Strain MC, Rought SE, Vargas M, Little SJ, Richman DD, Spina CA, Smith DM. Associations between the Virologic and Immunologic Dynamics in Blood and in the Male Genital Tract. *J Virol*. 2012b; 86:1307–1315. [PubMed: 22114342]
- Gunthard HF, Frost SD, Leigh-Brown AJ, Ignacio CC, Kee K, Perelson AS, Spina CA, Havlir DV, Hezareh M, Looney DJ, Richman DD, Wong JK. Evolution of envelope sequences of human immunodeficiency virus type 1 in cellular reservoirs in the setting of potent antiviral therapy. *J Virol*. 1999; 73:9404–9412. [PubMed: 10516049]
- Günthard HF, Havlir DV, Fiscus S, Zhang ZQ, Eron J, Mellors J, Gulick R, Frost SDW, Brown AJL, Schleif W. Residual Human Immunodeficiency Virus (HIV) Type 1 RNA and DNA in Lymph Nodes and HIV RNA in Genital Secretions and in Cerebrospinal Fluid after Suppression of Viremia for 2 Years. *J Infect Dis*. 2001; 183:1318–1327. [PubMed: 11294662]
- Gupta P, Leroux C, Patterson BK, Kingsley L, Rinaldo C, Ding M, Chen Y, Kulka K, Buchanan W, McKeon B, Montelaro R. Human immunodeficiency virus type 1 shedding pattern in semen correlates with the compartmentalization of viral Quasi species between blood and semen. *J Infect Dis*. 2000; 182:79–87. [PubMed: 10882584]
- Hudson RR, Boos DD, Kaplan NL. A statistical test for detecting geographic subdivision. *Molecular biology and evolution*. 1992a; 9:138–151. [PubMed: 1552836]
- Hudson RR, Slatkin M, Maddison WP. Estimation of levels of gene flow from DNA sequence data. *Genetics*. 1992b; 132:583–589. [PubMed: 1427045]

- Kalichman SC, Di Berto G, Eaton L. Human immunodeficiency virus viral load in blood plasma and semen: review and implications of empirical findings. *Sex Transm Dis*. 2008; 35:55–60. [PubMed: 18217225]
- Kosakovsky Pond SL, Posada D, Stawiski E, Chappey C, Poon AF, Hughes G, Fearnhill E, Gravenor MB, Leigh Brown AJ, Frost SD. An evolutionary model-based algorithm for accurate phylogenetic breakpoint mapping and subtype prediction in HIV-1. *PLoS Comput Biol*. 2009; 5:e1000581. [PubMed: 19956739]
- Lang DJ, Kummer JF. Demonstration of cytomegalovirus in semen. *N Engl J Med*. 1972; 287:756–758. [PubMed: 4341443]
- Lange M, Klein EB, Kornfield H, Cooper LZ, Grieco MH. Cytomegalovirus isolation from healthy homosexual men. *JAMA*. 1984; 252:1908–1910. [PubMed: 6088818]
- Marks G, Crepaz N, Janssen RS. Estimating sexual transmission of HIV from persons aware and unaware that they are infected with the virus in the USA. *AIDS*. 2006; 20:1447–1450. [PubMed: 16791020]
- Nickle DC, Shriner D, Mittler JE, Frenkel LM, Mullins JI. Importance and detection of virus reservoirs and compartments of HIV infection. *Curr Opin Microbiol*. 2003; 6:410–416. [PubMed: 12941414]
- Noviello CM, Pond SL, Lewis MJ, Richman DD, Pillai SK, Yang OO, Little SJ, Smith DM, Guatelli JC. Maintenance of Nef-mediated modulation of major histocompatibility complex class I and CD4 after sexual transmission of human immunodeficiency virus type 1. *Journal of virology*. 2007; 81:4776–4786. [PubMed: 17329339]
- Paranjpe S, Craig J, Patterson B, Ding M, Barroso P, Harrison L, Montelaro R, Gupta P. Subcompartmentalization of HIV-1 Quasispecies between Seminal Cells and Seminal Plasma Indicates Their Origin in Distinct Genital Tissues. *AIDS Res Hum Retroviruses*. 2002; 18:1271–1280. [PubMed: 12487815]
- Pillai SK, Good B, Pond SK, Wong JK, Strain MC, Richman DD, Smith DM. Semen-specific genetic characteristics of human immunodeficiency virus type 1 env. *J Virol*. 2005; 79:1734–1742. [PubMed: 15650198]
- Pinto-Neto LF, Vieira NF, Soprani M, Cunha CB, Dietze R, Ribeiro-Rodrigues R. Lack of correlation between seminal and plasma HIV-1 viral loads is associated with CD4 T cell depletion in therapy-naïve HIV-1+ patients. *Mem Inst Oswaldo Cruz*. 2002; 97:563–567. [PubMed: 12118292]
- Piot P, Bartos M, Ghys PD, Walker N, Schwartlaender B. The global impact of HIV/AIDS. *Nature*. 2001; 410:968–973. [PubMed: 11309626]
- Pond SL, Frost SD, Muse SV. HyPhy: hypothesis testing using phylogenies. *Bioinformatics*. 2005; 21:676–679. [PubMed: 15509596]
- Rieder P, Joos B, Scherrer A, Kuster H, Braun D, Grube C, Niederöst B, Leemann C, Gianella S, Metzner K, Böni J, Weber R, Gönthard H. Characterization of Human Immunodeficiency Virus Type 1 (HIV-1) Diversity and Tropism in 145 Patients With Primary HIV-1 Infection. *Clin Infect Dis*. 2011
- Shankarappa R, Margolick JB, Gange SJ, Rodrigo AG, Upchurch D, Farzadegan H, Gupta P, Rinaldo CR, Learn GH, He X, Huang XL, Mullins JI. Consistent viral evolutionary changes associated with the progression of human immunodeficiency virus type 1 infection. *Journal of virology*. 1999; 73:10489–10502. [PubMed: 10559367]
- Sheth PM, Danesh A, Sheung A, Rebbapragada A, Shahabi K, Kovacs C, Halpenny R, Tilley D, Mazzulli T, MacDonald K, Kelvin D, Kaul R. Disproportionately high semen shedding of HIV is associated with compartmentalized cytomegalovirus reactivation. *J Infect Dis*. 2006; 193:45–48. [PubMed: 16323130]
- Slatkin M, Maddison WP. A cladistic measure of gene flow inferred from the phylogenies of alleles. *Genetics*. 1989; 123:603–613. [PubMed: 2599370]
- Smith DM, Kingery JD, Wong JK, Ignacio CC, Richman DD, Little SJ. The prostate as a reservoir for HIV-1. *AIDS*. 2004; 18:1600. [PubMed: 15238781]
- Smith DM, Zarate S, Shao H, Pillai SK, Letendre SL, Wong JK, Richman DD, Frost SD, Ellis RJ. Pleocytosis is associated with disruption of HIV compartmentalization between blood and cerebral spinal fluid viral populations. *Virology*. 2009; 385:204–208. [PubMed: 19100592]

- Speck CE, Coombs RW, Koutsky LA, Zeh J, Ross SO, Hooton TM, Collier AC, Corey L, Cent A, Dragavon J, Lee W, Johnson EJ, Sampoleo RR, Krieger JN. Risk factors for HIV-1 shedding in semen. *Am J Epidemiol.* 1999; 150:622–631. [PubMed: 10490002]
- UNAIDS. AIDS epidemic update: 2010. Geneva: UNAIDS and WHO; 2009.
- Vernazza PL, Gilliam BL, Dyer J, Fiscus SA, Eron JJ, Frank AC, Cohen MS. Quantification of HIV in semen: correlation with antiviral treatment and immune status. *AIDS.* 1997; 11:987. [PubMed: 9223732]
- Wei X, Ghosh SK, Taylor ME, Johnson VA, Emini EA, Deutsch P, Lifson JD, Bonhoeffer S, Nowak MA, Hahn BH, et al. Viral dynamics in human immunodeficiency virus type 1 infection. *Nature.* 1995; 373:117–122. [PubMed: 7529365]
- Wong JK, Ignacio CC, Torriani F, Havlir D, Fitch NJ, Richman DD. In vivo compartmentalization of human immunodeficiency virus: evidence from the examination of pol sequences from autopsy tissues. *J Virol.* 1997; 71:2059–2071. [PubMed: 9032338]
- Zarate S, Pond SL, Shapshak P, Frost SD. Comparative study of methods for detecting sequence compartmentalization in human immunodeficiency virus type 1. *J Virol.* 2007; 81:6643–6651. [PubMed: 17428864]

Table 1

Compartmentalization analysis between blood and genital tract

ID	TP	Test	P-value	Comp	CMV
L7	TP1	SM	0	Yes	nd
		Fst	0		
	TP2	SM	0	Yes	3.97
		Fst	0.02		
TP3	SM	0.01	Yes	3.02	
		Fst	0.01		
	TP4	SM	0.79	No	nd
		Fst	0.8		
<hr/>					
X3	TP1	SM	0.98	No	3.5
		Fst	0.71		
	TP2	SM	0.72	No	nd
		Fst	0.33		
TP3	SM	0.58	No	nd	
		Fst	0.72		
<hr/>					
P9	TP1	SM	0	Yes	nd
		Fst	0		
	TP2	SM	1	No	nd
		Fst	0.97		
TP3	SM	0	Yes	6.81	
		Fst	0		
	TP4	S-M	0	Yes	4.45
		F_st	0		
<hr/>					
L3	TP1	SM	0.03	No	nd
		Fst	0.47		
	TP2	SM	0.09	No	4.13
		Fst	0.5		
TP3	SM	0	Yes	4.24	

ID	TP	Test	P-value	Comp	CMV
<hr/>					
Est					
0					

ID: subject identification number, TP: time point;

Test performed to determine presence of compartmentalization between HIV-1 populations in blood and seminal plasma: SM: Slatkin-Maddison, Est: Wright-s measure of population subdivision with relatives p-values

Samples is defined as compartmentalized (Comp) if both tests show a p-value <0.05

CMV: Levels of cytomegalovirus in semen are reported as Log10 copies/ml, nd: not detectable

Table 2

Evolutionary rates in blood and semen samples

ID	Sample	Evolutionary distance	CI
L7	BP	0.011	0.008–0.015
	SP	0.010	0.007–0.014
X3	BP	0.007	0.005–0.009
	SP	0.006	0.004–0.009
P9	BP	0.024	0.020–0.030
	SP	0.027	0.020–0.035
L3	BP	0.010	0.007–0.013
	SP	0.012	0.008–0.017

ID: subject identification number, BP: blood plasma, SP: seminal plasma

Evolutionary distances estimated by computing mean pairwise path lengths in the maximum likelihood phylogenetic tree under the GTR + G + I model of sequence evolution, reported in expected nucleotide substitutions per site

CI: confidence intervals (profile likelihood)