Distribution Agreement

In presenting this thesis or dissertation as a partial fulfillment of the requirements for an advanced degree from Emory University, I hereby grant to Emory University and its agents the non-exclusive license to archive, make accessible, and display my thesis or dissertation in whole or in part in all forms of media, now or hereafter known, including display on the world wide web. I understand that I may select some access restrictions as part of the online submission of this thesis or dissertation. I retain all ownership rights to the copyright of the thesis or dissertation. I also retain the right to use in future works (such as articles or books) all or part of this thesis or dissertation.

Signature:

Martin J. Deymier	Date

Natural Selection During Subtype C HIV-1 Transmission

By

Martin J. Deymier
Doctor of Philosophy

Graduate Division of Biological and Biomedical Science
Microbiology & Molecular Genetics

Eric Hunter, Ph.D.
Advisor

Rama Amara, Ph.D
Committee Member

Cynthia A. Derdeyn, Ph.D.
Committee Member

Arash Grakoui, Ph.D.
Committee Member

Bali Pulendran, Ph.D.
Committee Member

Accepted:

Lisa A. Tedesco, Ph.D.
Dean of the James T. Laney School of Graduate Studies

Date

Natural Selection During Subtype C HIV-1 Transmission

By

Martin J. Deymier
B.S., University of Arizona, 2010

Advisor: Eric Hunter, Ph.D.

An abstract of
A dissertation submitted to the Faculty of the
James T. Laney School of Graduate Studies of Emory University
in partial fulfillment of the requirements from the degree of
Doctor of Philosophy
in Microbiology & Molecular Genetics
2015

Abstract

Natural Selection During Subtype C HIV-1 Transmission
By Martin J. Deymier

HIV-1, the causative agent of AIDS, kills approximately 1.6 million people per year globally. The majority of infected persons live in sub-Saharan Africa, where disease burden is greatest. A major roadblock to the development of an effective vaccine is the extraordinary genetic diversity of HIV, both around the world as well as within a single infected individual. However, during transmission of HIV between individuals, a strong genetic bottleneck occurs, wherein from a large diverse population of viruses present in a chronically infected person, only a single genetic variant establishes infection in the naïve host. Identifying the viral characteristics of these transmitted/founder (TF) variants may provide a useful common target for any future vaccine.

We and others have strived to define common genetic and phenotypic characteristics of these TF virus variants, although few traits have so far been confirmed. Work from our lab established that a genetic bottleneck occurs during transmission, and confirmed that there is a transmission selection bias for consensus amino acid residues. This selection bias suggests that consensus amino acid residues lead to virus variants with increased transmission fitness.

In order to further investigate the phenotypic characteristics of HIV TF variants, we have examined the populations of full-length genome HIV from six linked heterosexual transmission pairs near the time of transmission. We confirmed that TF variants had more consensus-like genomes across all viral genes. However, following the generation of full-length infectious molecular clones, we did not observe a selection for variants with increased infectivity, *in vitro* replicative capacity, or inherent resistance to IFNα. Our findings indicate that although selection for more consensus-like variants suggests a selection of virus with increased transmission fitness, this fitness is not associated with *in vitro* measurements of virus replication. Continued efforts to describe these uniquely derived full-length genome TF viruses in comparison to their non-transmitted counterparts will offer additional insight into viral requirements of transmission and may provide potential common targets for an effective vaccine.

Natural Selection During Subtype C HIV-1 Transmission

By

Martin J. Deymier
B.S., University of Arizona, 2010

Advisor: Eric Hunter, Ph.D.

A dissertation submitted to the Faculty of the
James T. Laney School of Graduate Studies of Emory University
in partial fulfillment of the requirements from the degree of
Doctor of Philosophy
in Microbiology & Molecular Genetics
2015

Table of Contents

Chapter I : Introduction

Acquired immunodeficiency syndrome (AIDS) is a human chronic disease characterized by depressed cellular immunity due to a depletion of CD4+ T cells. In 1981, AIDS was first reported in previously healthy homosexual men living in Los Angeles who displayed symptoms of *Pneumocystis* pneumonia, a condition observed typically only in severely immunosuppressed individuals (1). Two years later the causative agent of AIDS was discovered by Luc Montagnier at the Pasteur Institute in Paris, wherein by culturing lymphoid biopsies from an individual at risk of AIDS, his group isolated and described a novel human lentivirus now known as human immunodeficiency virus (HIV) (2). Since the beginning of the HIV pandemic approximately 78 million people have been infected with HIV and about 39 million people have died of AIDS (3). As of 2012, about 35.3 million people live with HIV and in that same year approximately 1.6 million people died of AIDS-related deaths (3). This pandemic has been throughout the later years of the 20th century a medical challenge, which, far exceeded the expectations of the time and continues to require both a preventative vaccine and a cure.

Although HIV and AIDS were only reported in the early 1980s, the history of the virus predates those discoveries approximately 80 years to the early 20th century where HIV-1 established itself in the human population following zoonotic transmission from chimpanzees in the Democratic Republic of Congo (4). Twelve different strains of HIV (four strains of HIV-1 and eight strains of HIV-2) circulate among humans today following twelve independent cross-

species transmissions from primates (reviewed in (5)). The two major lineages of HIV, HIV-1 and HIV-2, originated from chimpanzees (6) and sooty mangabeys respectively (7). HIV-1 is further divided into four main groups (M, N, O, & P) representative of the four independent zoonotic transmission events from primates to humans (8). The most common strain of HIV worldwide is HIV-1 group M, which as previously mentioned originated in central Africa, and has now further spread across the globe (4,9). Following multiple introductions into naïve populations around the world, HIV-1 group M can now be further separated in genetically distinct virus subtypes (A-K) and inter-subtype recombinants (8,10). A majority of the global health burden is centered in sub-Saharan Africa where 70% of the total number of people living with HIV are located. Subtype C HIV-1 is the most common strain in this disproportionately affected area, and represents approximately 50% of the global virus subtype prevalence (9). The genetic diversity of HIV is one of the major hurdles towards developing an effective HIV vaccine or cure. Although current highly active antiretroviral therapies (HAART) are effective at suppressing virus replication and controlling a patient's viral load, these drugs must be taken for perpetuity, and are cost-prohibitive in low-income countries (reviewed in (11)). If HAART therapy is not adhered to properly it can select for drug resistance mutations in the viral genome leading to the potential transmission of strains across naïve populations resistant to common antiretrovirals. The difficulty in treating and preventing HIV is in part due to its mutation driven diversity, as well as the ability of the virus to establish long-lived latent infection inside the host cell.

HIV was first discovered by introducing interferon blocking agents to cell cultures collected from individuals affected by AIDS, which allowed production of virions that could be observed under transmission electron microscopy (2). HIV was discovered to be a single-stranded RNA (ssRNA) virus of the genus *lentivirus* from the *retroviridae* family, most similarly related to previously known primate lentiviruses and viruses associated with chronic illnesses in ungulates and cats (6,7,12). This small, plasma-membrane enveloped *lentivirus* with a ssRNA genome of approximately ~9 kilobases (kbs), encodes 9 genes that allow it to effectively infect, replicate in, and disseminate through the host's target cells. As a typical retrovirus, HIV, encodes three main genes, the *gag*, *pol*, & *env* genes, which when translated generate the virus capsid structural proteins (Gag), the reverse transcriptase machinery (Pol), and the viral surface glycoprotein (Env), respectively. Additional accessory genes *vif*, *vpr*, *vpu*, *tat*, *rev* & *nef* all encode proteins which enhance virus replication by inhibiting host antiviral responses (Vif, Vpu), promote processes of virus replication such as integration, transcription, and transport of viral mRNAs (Vpr, Tat, Rev), or downregulate cellular host surface proteins in order to prevent superinfection of single cells as well as decrease the response of the human adaptive immune system to infected cells (Nef). These virally encoded genes along with host cell machinery are essential for the virus life cycle and for transmission to new naïve hosts.

The major host target cells for HIV-1 are CD4+ T lymphocytes, which play a major role in the human adaptive immune response to both bacterial and viral infections. Human CD4+ T cells perform many functions in priming the immune response following a new viral or bacterial infection. These lymphocytes, survey

the lymphoid tissues for antigen presenting cells displaying foreign peptides and can receive signals from these cells to further enhance the stimulation of macrophages to kill intracellular bacteria, help antigen-specific B cells generate antibodies, and provide stimulatory cytokines to CD8+ T cells (13). Since these lymphocytes are essential players for proper immune function, long-term depletion caused by HIV infection is the underlying cause of HIV disease progression and AIDS, allowing opportunistic pathogens to thrive within the immunocompromised host.

On the single cell level, the HIV replication cycle begins as the virion binds to its primary receptor, CD4 on the surface of the cell via the Env glycoprotein, triggering a conformational change and allowing portions of Env to become exposed, which bind to the coreceptor CCR5 or CXCR4 (14-16). Once the Env surface protein has attached to both receptors, a fusion triggering conformational event takes place allowing the virus to fuse with the cell membrane and deliver the structural core into the cytoplasm of the cell. The virus capsid delivers the viral ssRNA attached to virion incorporated reverse transcriptase (RT) and integrase proteins. Unique to the family of retroviruses this complex undergoes reverse transcription from ssRNA and generates a double-stranded DNA (dsDNA) copy of the viral genome prior to transport and integration into the cellular genomic DNA (17-19). In activated CD4+ T cells, following integration, the viral genome is transcribed by the cell's RNA transcriptional machinery along with viral accessory proteins to generate the required proteins for viral assembly of new genomic ssRNA molecules along with the viral capsid and surface glycoproteins needed for assembly and budding. This complete viral replication

cycle takes approximately 2 days from entry to egress of new virions, and is followed by the death of the infected cell (20,21). As opposed to other retroviruses, HIV along with other *lentiviruses*, due to their accessory genes, can infect non-dividing cells and integrate their genomes into the host genome independent of cell cycle stage (22). HIV is often found as a latent provirus in infected cells contributing to the latent reservoir (reviewed in (23). Latently infected cells are thought to arrive by infection of partly activated CD4+ T cells followed by cellular relaxation and transcriptional silencing (24). The ability of HIV to establish a latent infection prevents antiretroviral treatment from effectively clearing the infection, since the virus rebounds to normal levels once treatment is interrupted (14-16,25). Although the impact of latent infection during HIV transmission is not known, in order to establish long-term infection within a newly infected individual the virus must undergo active replication cycles in the genital mucosal tissues of a naïve host.

The spread of HIV through the human population occurs primarily from sexual exposure of genital mucosa to virally infected genital fluids in the case of hetero- and homosexual contact, but the virus can also transmit from mother-to-child, and through direct blood-to-blood contact by contaminated transfusions or intravenous injections (reviewed in (26)). HIV risk of infection is highly dependent on the route of transmission. Virus exposure in men who have sex with men (MSM) through the rectum has an infection probability of 1 in 20 to 1 in 300, whereas infection by the male genital tract has a risk of approximately 1 in 700 to 1 in 3000 (26). Infection in the female genital tract carries a risk of 1 in 200 to 1 in 2000 (3,26). In the United States and Europe a majority of new

infections are caused by subtype B infections in MSM populations. On the other hand in Sub-Saharan Africa the majority of new infections are subtype C and transmitted via heterosexual contact in serodiscordant cohabiting partnerships, either male-to-female (MTF) or female-to-male (FTM) (3,27). It is known that the risk of infection can be affected by certain factors, such as donor viral load, or recipient genital ulcers and inflammation (4,28,29). Understanding the mechanisms required for viral transmission in these genital tissues, as well as the viral genetics surrounding the transmission process is crucial in determining preventative strategies for transmission.

The effectively low infection rate per exposure during human transmission indicates that this process is quite inefficient and that multiple barriers exist at the site of transmission to prevent infection (5,30). Mucosal surfaces of the human body are naturally coated with mucus, antibacterial peptides and a layer of epithelial cells, which protects the endothelial layers of skin (31). HIV must in this harsh environment find and infect a tissue resident CD4+ cell in order to establish itself in that host. In the female genital tract the epithelial cells of the vagina and ectocervix are a stratified squamous lining of cells, as opposed to the single columnar layer present in the endocervix, and has shown to allow virus particles to disseminate through it and come into contact with tissue resident Langerhans cells, and CD4+ T cells (6,32,33). Studies in primates and human cervicovaginal explant models have shown that susceptible cells are present in the basal layer of the epithelium, although pseudotyped virus particles have been seen to enter cells along all vaginal membranes as well as in the uterus (7,33-35). The surface of the male genitalia in both the outer and inner foreskin contain

keratinized layers of skin which substantially inhibit virus infection, yet still contain Langerhans cells and resident CD4+ T cells in the epithelial that become targets for initial infection (8,36). Circumcision has proved an effective preventative intervention for HIV infection and has been shown in clinical trials to reduce infection by 60% (8,10,37). The presence of a keratinized layer of cells on the male genital epithelium may explain why the risk per exposure event is lower in FTM transmission. Langerhans cells and dendritic cells have been implicated as potential sources of early infection or viral binding followed by their trafficking to the lymphatics to further transfer the virus to susceptible CD4+ T cells (9,35,38-42). Although the barriers the virus must cross to establish infection are great, infection still occurs, and the virus is able to penetrate through healthy host tissue. The ability of the virus to perform all of these functions from exposure to establishment of infection in a new host is defined as its transmission potential. A crucial process for HIV spread, the process of transmission, which is naturally inefficient may provide insight into mechanisms to enhance via vaccination strategies to help make the process impossible to new infections.

Following initial infection of a susceptible cell the virus must disseminate from the genital tract to local lymphatics and continue to efficiently replicate in lymphocytes before migrating to the gut-associated lymphoid tissues (GALT) where early after infection the majority of virus replication and CD4+ T cell depletion takes place (9,43). During the initial phase of replication in the tissue the virus must replicate in the presence of innate host antiviral factors such as NK cells and the antiviral effets of Interferon- α (IFNα) which is induced early

upon viral infection of cells in the tissue (44-46). Once the virus has evaded early innate immune responses the virus can spread to the lymph nodes of the host. Initial viral trafficking to the GALT has been associated with the virus Env gp120 of acute viruses binding to the α4β7 molecule on the surface of CD4+ T cells (11,47-49). α4β7 acts as a gut homing receptor on the surface of CD4+ T cells. It binds to the mucosal vascular addressin cell adhesion molecule, a ligand expressed on blood vessels of the GALT (50). In studies of rhesus macaques, in which HIV infected cells were tracked with a copper-tagged antibody that was visualized via PET scanning, virus was observed throughout the body, including lungs, and nasal cavities following initial peak viremia (14-16,51). Although the human immune system mounts a robust immune response against HIV, it cannot clear infection, and despite early innate as well as adaptive cellular and humoral responses, through selection of de-novo mutations the virus rapidly escapes from these immune pressures (17-21,44,52,53). The viral diversity established by the continued battle between the virus and cytotoxic T cell and antibody-mediated responses leads to a diverse viral quasispecies within the host. This genetic diversity is not only due to the intrinsic error rate of the HIV reverse transcriptase but also due to the high viral loads and viral turnover during infection, along with viral recombination (20,54-57). As mentioned previously, the between host and within host HIV diversity makes effective universal vaccine strategies difficult. Although the global diversity is great, a unique opportunity occurs during the HIV transmission process wherein during this inefficient process it has been shown that only single viral variants establish infection in a naïve host. This is commonly known as the transmission bottleneck or founder

effect, where from a diverse population, following transmission, a reduction in diversity occurs wherein a single viral variant, the Transmitted/Founder virus, establishes infection (58-60). This process has been studied extensively as since understanding the viral requirements for transmission may provide a common target against HIV for preventative vaccines.

Landmark studies published in 1992 and 1993, for the first time analyzed viral sequences from early time points in primary HIV infection, as well as, in some cases, viral sequences from the donors of linked heterosexual, homosexual and mother-to-child transmission events (61-64). Wolinsky *et al.* analyzed sequences of the V3-V5 region of HIV *env* gene from three separate mother to child transmission events and determined that sequences within the infant after transmission were less diverse than those found in their mothers. They also observed that the sequence in the mother that most resembles the infant's HIV sequence is a minor variant in the mother's plasma, and that a potential n-linked glycosylation (PNG) site at the start of V3 which was conserved in all three mothers, was never transmitted to the infant. The following studies by Wolf *et al.* (1992), Zhu *et al. (1993)*, and Zhang *et al. (1993),* presented similar findings in early primary infections of heterosexual, homosexual, and parenteral transmission events where almost all the *env* sequences in recipients were genetically homogeneous (62-64). Wolf *et al.* also mentioned that in the two heterosexual transmissions described, the transmitted virus appears as a minor variant in the blood of the transmitters. In parallel with the genetic analyses, phenotypic studies of early isolates compared to late isolates were carried out and discovered that isolates from early stages of infection generally grew more slowly,

did not induce syncytia between infected cells, and retained a capability to replicate in primary macrophages as well as in activated CD4+ lymphocytes. In contrast, later in infection, variants appear that are highly cytopathic upon virus culture and lost their ability to grow in macrophages (62,63,65,66). These studies all laid down a fundamental framework for more extensive studies of the events occurring during the time of HIV transmission. A few possibilities explaining the genetic homogeneity of virus in recipients were put forth based on these preliminary observations: 1) selective transmission of a variant capable of establishing infection in mucosal tissues, 2) transmission of multiple variants followed by viral outgrowth of a most fit variant or 3) low virus inoculum or variant sequestration in the donor genital tract. Each provides a basis for which a genetic bottleneck might occur at transmission, and each may provide information that may lead to specific targets for microbicide development or rational vaccine design.

The discovery of HIV obligate coreceptor usage in the mid 1990s lead to the observation that HIV can use different coreceptors, and that previously defined cell tropism phenotypes were dependent on coreceptor expression. CCR5 and CXCR4 were discovered to be the main coreceptors for HIV, and provided the basis for what at the time were called macrophage-tropic/non-syncitium and T-cell tropic/syncytium inducing phenotypes, respectively (15,16,67,68). This clarified the observation that subtype B viruses purified from acute time points were typically CCR5 tropic, and that later time points during disease saw the rise of CXCR4 tropic strains (61-64,69). Although the switch to CXCR4 occurs in subtype B chronic infections, subtype C viruses typically retain CCR5 tropism

(70). A critical clinical discovery that people homozygous for a deletion of a 32 amino acids in their CCR5 gene were protected from acquiring HIV by mucosal exposure, provided more evidence that of the two major coreceptors used by HIV, only one, CCR5, was required for HIV transmission (62,63,65,66,71). This presented evidence for selection against CXCR4 tropic viruses occurring during mucosal transmission of HIV.

The difficulty of obtaining samples at early time points after infection, and from linked donors of HIV transmission, hindered successful investigation of viral characteristics. Twelve years after the initial discoveries of HIV genetic homogeneity after transmission to a new host, a more in depth analysis was performed using samples from a previously established cohort of serodiscordant couples in Lusaka Zambia, founded by Dr. Susan Allen. Derdeyn *et al*. sequenced cloned HIV-1 subtype C & G *env* genes from 8 heterosexually linked transmission pairs, and performed genetic and phenotypic studies (58). A strong genetic bottleneck was observed, with a single genetic variant transmitted in each pair, and *env* genes found in the linked recipient soon after transmission encoded significantly shorter and less glycosylated gp120 V1-V4 regions than the average of the donor Env at the time of transmission (58). This first demonstration of statistically confirmed genetic signatures associated with HIV transmission garnered a lot of attention toward the HIV selection hypothesis. These findings were also confirmed for 18 more subtype C acutely infected individuals (72), as well as in subtype A acutely infected individuals compared to a database of matched chronic sequences, and in 13 subtype D and A transmission pairs from the Rakai district of Uganda (73,74). These genetic signatures were not found in

subtype B transmissions, suggesting that different virus subtypes may have unique transmission characteristics (73,75). Many caveats from bulk amplification and cloning of *env* genes existed, including possible genomic recombination by template switching during amplification, the introduction of *Taq* errors in the early steps of amplification, and cloning bias of certain Env sequences, therefore a new method of single genome amplification followed by direct amplicon sequencing was developed in order to more precisely describe the sequence of recipient envelopes (76). By directly sequencing single genome amplified *env* from 102 subjects with acute subtype B infection, and inferring a consensus by assuming a star like- phylogenetic random evolution of virus for the first few weeks of infection, Keele et al. to determined the precise *env* sequence that initiated infection in these patients, the Transmitted/Founder (TF) virus (59). This for the first time deduced exactly which virus Envelope initiated infection in the naïve host, and observed that of the 102 subjects, 78 had evidence of being infected by a single virus variant (59). The search for genetic signatures continued, and by sequencing *env* SGA amplicons, from acute and chronically subtype B infected individuals, two significant transmission associated signatures were found (77). The most significant of these signatures was the presence of a Histidine at amino acid position 12 of the signal peptide, which may effect Env expression, and incorporation of Env into virions (78). A virion containing higher levels of Env may have a distinct advantage at establishing in infection where there are few available target cells that may have low levels of CD4.

In certain cases multiple variants establish infection following transmission. An in depth analysis of 171 subtype B and C transmission events

found that infection with more than one variant did not follow a Poisson distribution and therefore these cannot be seen as independent transmission of multiple variants but are caused by some mechanism, such as damage to the mucosal barrier (79). In a cohort of subtype A and C heterosexual transmission pairs, the presence of inflammatory genital infections abolished the strict genetic bottleneck and lead to transmission of multiple variants (60). So far in sexually transmitted subtype C and B infections, about ~85% are initiated by a single viral variant, while others are initiated by only a few. These transmissions of multiple variants occurred most often in case of men who have sex with men, and intravenous drug users where as few as 60% and 40% of transmissions are caused by a single viral variant, depending on the cohorts (80-82). Transmission of multiple variants occurs when sexually transmitted infections and associated inflammation abrogate the mucosal barrier by inducing breaks in the epithelial lining. Inflammation induced by these infections may also increase the numbers of activated CD4+ T cell targets for the virus in the genital mucosa.

The role in transmission of the viral quasispecies in the genital tract fluids remained to be determined (reviewed in (83)). Were all these genetic signatures due to limited population diversity in the donor genital tract or was this really due to mucosal selection at the site of entry in the recipient? As previously mentioned, the first reports on the transmission bottleneck suggested that transmitted viruses were minor variants of the donor's plasma (61). A follow up study comparing viral genetic diversity within and between transmission pairs using homoduplex assays from virus in their plasma, seminal fluid and seminal cells indicated that the transmitted virus was present both as cell-free and cell-

associated virus in the donor genital tract, and was also generally a minor variant of the genital tract (84). A more recent analysis with modern techniques was required, therefore once SGA techniques were described, a more in depth study was performed to define the role of genital tract virus compartmentalization during HIV transmission bottleneck. Single genome amplification of the V1-V4 region of Env from genital tract samples of eight subtype C infected heterosexual transmission pairs, whose plasma sequences were already determined, showed significant genital compartmentalization of viral variants (85). However, the virus in the donor quasispecies that most resembled the TF virus was a minor variant of the genital tract in some cases and of the plasma in others, and was not derived from the most abundant subpopulation in the genital tract, suggesting a non-stochastic process for transmission (85). These findings suggested that viral selection was occurring during transmission. A larger investigation on the genetic level of viral selection recently confirmed a selection bias for consensus amino acids was occurring during transmission (86). It was shown by defining the individual amino acid sites in the *gag*, *pol,* and *nef* genes from 137 linked transmission pairs, that amino acids that were consensus for the cohort had a specific advantage during the transmission process, independent of their frequency within the chronically infected donor (86). All of these studies together have brought the field to the understanding that upon transmission a process of selection occurs for viruses that may be less adapted to that individual but more closely resemble transmitted/founder viruses of the past, and hence may retain a particular phenotype which includes a transmission fitness advantage (74,87).

The founder effect that occurs during HIV transmission has been validated in multiple cohorts, in different modes of transmission, with all major subtypes, yet the method by which natural selection consensus-like genomes and shorter, less glycosylated, and CCR5 tropic Envelopes occurs has yet to be explained (58-60,74). Following the ability to accurately determine the TF sequence of acutely infected individuals many phenotypic studies were performed by cloning these Envs and comparing them to chronic Envs in assays that probe characteristics possibly required for transmission. The initial study from the Zambia cohort also used Env pseudotyped viruses from the heterosexually linked transmission pairs in neutralization assays, and found that the shorter, less glycosylated envelopes were also less resistant to donor plasma antibodies (58). This suggests that antibody escape in a host may compromise the transmission potential of a certain viral variant, resulting in significant effort to characterize the neutralization profile of TF virus Envs (58,59). Although natural antibody responses during acute infection neutralize circulating variants, these responses do not target cross-reactive neutralizing epitopes between clades (88). It has been suggested that the shorter, less glycosylated envelope glycoproteins may allow for easier access to CD4 and therefore might produce more infectious virus, although further investigation revealed that subtype C TF variants were resistant to soluble CD4 (72). It has since been found that a subset of individuals generate broadly neutralizing antibodies that are potent against a range of HIV-1 subtypes that could be used to neutralize transmitting viruses (reviewed in ((89,90)).

The more neutralization sensitive, shorter, less glycosylated envelopes also lead to the question of whether a virus Env that has evolved and adapted to its

host by escaping nAbs may have lost the ability to transmit. Genetic analysis of HIV sequences from heterosexual transmission pairs were conducted to determine whether the early/acute viruses that infect the new individual more closely resemble viruses from earlier time points in their previous host. This was at first performed by analyzing HIV sequences from linked transmission pairs near the time of transmission and calculating the distances of recipient and donor viruses to their most recent common ancestor (MRCA). Sagar *et al.* for the first time provided evidence that recipient viruses were closer to the MRCA than donor viruses by calculating distances of recipient and donor Envelope sequences to their MRCA from 10 subtype D and 10 subtype A linked transmission pairs (74). This work is supported by the fact that in two separate studies of cohorts containing subtypes D, A, and B, intra-host diversity was found to be greater than inter-host diversity (87,91). A longitudinal analysis of viruses from donors also found that the TF virus more closely resembled earlier viruses in the donor than the viruses circulating at the time of transmission (87). These findings are consistent with the observation that consensus-like amino acids are selected for during transmission (86). They all suggest that viruses that acquire polymorphisms during adaptation to the chronically infected individual evolve away from a virus that has the highest potential transmission fitness (86). All of the genetic findings point to a fact that a more consensus-like/ancestral viral genome has the enhanced transmission associated properties, although the underlying mechanisms are not known.

Phenotypic studies of Envs cloned from TF/acute and chronic viruses were carried out to determine differences in virus characteristics associated with

transmission, unfortunately few differences were found. Subtype C acute and chronic Envs from heterosexual transmissions pairs showed no greater efficiency of receptor utilization when compared on HeLa Cells expressing varying amounts of CD4 and CCR5 (92). Additional phenotypic analyses of these acute/early Envs and chronic Envs found that there was no overall difference in these Envs from transmission pairs in their infection of CD4+ T cells, or macrophages, which they infected much less efficiently than activated T cells. The only difference was that early Envs are more restrained in their use of CCR5 (93). Subtype B TF and chronic Env pseudotyped viruses were tested in a variety of phenotypic assays only to show that they were phenotypically equivalent in coreceptor tropism, CCR5 utilization efficiencies, primary CD4 T cell subset tropism, dendritic cell *trans*-infections, fusion kinetics, and neutralization sensitivities. However, TF Envs were modestly more sensitive to CD4 binding site antibodies (94). A further analysis, of Envs derived from 40 TF and 47 chronic viruses from subtypes B and C, and their ability to mediate infection of cells with high levels of CCR5 in the presence of maravoric showed that TF viruses were less resistant to maraviroc inhibition (95). Overall there is no evidence that TF Envs preferentially bind low levels of CD4 and CCR5, or preferential usage of alternate coreceptors although they may be confined to use a specific conformation of CCR5 based on sensitivity to maraviroc.

In addition to the key receptors for virus fusion, CD4 and CCR5, it was recently found that α4β7 integrin on T cells forms complexes with CD4 molecules and increases susceptibility of these T cells to HIV infection (96). The precise mechanism by which HIV Env binds and signals through α4β7 derives from a

three amino acid sequence on the V2 loop (97). HIV Envs with transmission associated reduced n-linked glycosylation were found to bind to α4β7 more efficiently, a possible phenotypic preference for early transmitting viruses (49,98). It has also been shown that targeting α4β7 via antibodies in macaques can reduce the efficiency and delay time to viral infection through intra-vaginal infection (48). Finally, in subtype C viruses from South Africa TF variants were observed to have increased dependence on α4β7 binding for efficient *in vitro* replication, and that this dependence waned with time (98). More studies are necessary to determine the role of α4β7 T cells in the genital mucosa and whether TF and non-transmitted donor variants may differentially bind and infect these cells.

The inability to find consistent, statistically strong phenotypic differences between Env clones derived from TF and chronic viruses argued that other regions of the viral genome may be important in accurately defining the phenotypic properties of TF and chronic viruses. The analysis of full-length genomes from acutely infected individuals was first conducted in 2009 (99). Using SGA techniques and a model of evolution as previously defined for envelope, 12 full length genome TF sequences were inferred for acutely infected individuals of both subtype B and C viruses (99). Three subtype C full length infectious molecular clones (IMC) were generated and were shown to replicate efficiently in activated CD4+ T cells but only modestly in macrophages (99). The generation of 10 more clade B TF IMCs and 6 subtype C TF IMCs confirmed these results and determined that in comparison to chronic viruses, TF viruses infected CD4+ T cells with equal efficiency, had low replication in monocyte derived

macrophages, and had no T cell subset specificities between T central memory, or T effector memory targets (100,101). Subtype A, D, and A/D recombinant full length TF IMCs all were also replication competent and it was found that subtype D TF had slightly greater replication capacities than Subtype A TFs (102). The most thorough analysis to date of TF IMCs compared 27 TF with 14 chronic control viruses of both subtypes B and C, finding that TF viruses were more infectious than chronic viruses, had more Env per particle, and were also better captured and transferred to T cells by Dendritic Cells (103). Subtype differences were found when TF and chronic viruses were used to infect CD4+ T cells in the presence of IFNα. Subtype B TFs were significantly more resistant to IFNα treatment than were the subtype B chronic viruses, whereas no significant differences were seen with subtype C viruses (103). Further analysis of the resistance to the antiviral effects of IFNα found that TF viruses were more resistant compared to 6-month consensus derived viruses from the same patients for both subtypes B and C (104). Although these analyses were the first to measure the virus phenotype over the entire genome of the TF and chronic circulating viruses, none of these studies with full length IMCs used linked transmission pairs since all of the chronic control viruses were derived from separate chronically infected patients or longitudinal samples, therefore comparisons cannot be linked to transmission.

In order to determine the replication and genetic requirements for HIV transmission, we have collaborated with the Zambia Emory HIV Research Project to generate full-length TF and non-transmitted (NT) IMCs from linked heterosexual transmission pairs. In the work presented here, we describe the

methods by which we have accurately and efficiently generated these reagents and applied this panel of viruses to assays that may have relevant importance for transmission (105). In the case of six heterosexual transmission pairs we have confirmed previous findings that TF viruses are more neutralization sensitive to contemporaneous autologous antibodies than the majority of donor NT viruses, and show across the entire viral genome that the transmission bottleneck selects for viruses with more consensus-like genes (106). We set out to establish the role of viral replicative capacity and resistance to IFNα in selection of TF viruses. We observed that TF viruses did not have increased infectivity, replicative capacity or IFNα resistance than the median of the diversity of variants present in the donor quasispecies. Together this work suggests that although selection clearly occurs, transmission fitness is not defined by *in vitro* viral replication. More subtle mechanisms of virus behavior *in vivo* during the transmission process will need to be assessed to define the transmission potential of HIV variants.

Future studies using full-length genome IMCs from linked transmission pairs assayed in relevant genital mucosal cells and tissues may be needed to clarify the genetic requirements of HIV at transmission. The human mucosal barrier provides a strong block towards the establishment of infection and naturally selects for viral variants that have a certain transmission phenotypes. It is crucial that we discover the genetic and phenotypic viral requirements for transmission in order to take advantage of the mechanisms required via a preventative intervention. Only by understanding the viral and host aspects of HIV transmission will an effective vaccine be established, which is the key to reducing the spread of this virus and stopping one of the great global pandemics.

Chapter II : Particle infectivity of HIV-1 full-length genome infectious molecular clones in a subtype C heterosexual transmission pair following high fidelity amplification and unbiased cloning

Martin J. Deymier[a], Daniel T. Claiborne[a], Zachary Ende[a], Hannah K. Ratner[a], William Kilembe[b], Susan Allen[b,c], Eric Hunter[a#]

[a]Emory Vaccine Center, Yerkes National Primate Research Center, Emory University, Atlanta, GA, USA

[b]Zambia-Emory HIV Research Project (ZEHRP), Lusaka, Zambia

[c]Department of Pathology and Laboratory Medicine, Emory University, Atlanta, GA, USA

Published in *Virology, 2014, Vols. 468-470*

Abstract

The high genetic diversity of HIV-1 impedes high throughput, large-scale sequencing and full-length genome cloning by common restriction enzyme based methods. By applying newly developed methods that employ a high-fidelity polymerase for amplification and an unbiased fusion-based cloning strategy, we have generated several HIV-1 full-length genome infectious molecular clones from an epidemiologically linked transmission pair. These clones represent both the transmitted/founder virus and phylogenetically diverse non-transmitted variants selected from the chronically infected individual's diverse quasispecies near the time of transmission. We demonstrate extremely limited PCR-induced mutations in full-length single clones derived from single genome amplicons. Furthermore, all genomes produced functional virus with a range of infectivities, belying the previous assumption that a majority of circulating viruses in chronic HIV-1 infection are defective. Thus, these methods provide important tools to update current protocols in molecular biology that could be universally applied to the study of human viral pathogens.

Introduction

Human immunodeficiency virus type 1 (HIV-1) is a human lentivirus that causes Acquired Immunodeficiency Syndrome (AIDS) and currently infects approximately 30 million people worldwide, killing ~1.3 million people annually (1). HIV-1 is spread predominantly through heterosexual transmission in Sub-Saharan Africa, where the health burden is greatest (1,2). HIV-1 infection leads to chronic disease in humans that is able to persist despite a robust cellular immune response as well as constant antibody-mediated pressure (3,4). HIV-1 evades immune pressures by generating escape mutations via its high viral turnover and error-prone reverse transcriptase enzyme, as well as undergoing latent infection of memory CD4+ T cells (5). Both rapid genetic mutation and latency also allow HIV-1 to develop resistance and persist against modern regimens of highly active antiretroviral therapy (6). The genetic diversity of HIV-1, a result of the accumulation of mutations in the viral genome, means that it exists as a complex quasispecies of many genetically-distinct viral genomes in chronically-infected individuals.

Little is known about the biological activity of the many circulating variants in the plasma of chronically-infected individuals. Although it is suggested that a majority of proviral integrated genomes comprising the latent reservoir are defective, this has been poorly characterized within the RNA viral quasispecies (7,7,8). Research to date describing the biological and molecular characteristics of HIV-1 from chronically infected individuals have, for the most part, used isolates grown out of patient plasma *in vitro*, or laboratory-adapted clonal strains (9,10). However, these viruses likely do not accurately represent the

viruses that exist *in vivo*, as they have been grown and passaged in culture, which can change their phenotypic properties (11). Accurately measuring the viral characteristics of full-length HIV-1 genomes derived from the quasispecies of a chronically-infected individual is important for determining the fitness of the quasispecies, which may impact pathogenesis (12). It will also provide a much sought after panel of viruses with which to compare Transmitted/Founder (TF) viruses in transmission studies in order to determine viral requirements for transmission (12,13).

In ~80% of heterosexual transmission events of HIV-1, a single TF virus is transmitted and establishes infection in the naïve host from the diverse quasispecies in the chronically-infected donor (14-17). In order to determine biological correlates associated with this genetic bottleneck, it is imperative to compare the TF virus with full-length viruses present in the donor near the time of transmission. In the context of transmission studies to date, no full-length TF virus has been compared to the non-transmitted full-length genome viruses of the cognate donor (13). A recent study generated full-length clones from clonal clusters in chronically-infected individuals because they hypothesized that the inferred common ancestors of these clusters encoded persistently replicating viruses in contrast to those viruses not in clusters (18). Previous strategies attempting to clone full-length genomes from isolates derived from chronically-infected patients failed to produce any functional genomes prior to rescue by an adapted Envelope gene (19,20).

We suggest here that the lack of knowledge in this area results from the absence of appropriate methodology to accurately amplify and clone full-length

genomes. The main methodological limitations of studies to date are the generation of full-length genomes crippled by PCR-induced mutations from the polymerases used, and dependence on restriction enzyme sites within those sequences in order to generate clones (9,21,22). Due to the inherent genetic diversity and instability of RNA viruses such as HIV-1, manipulation of the genetic material from which it exists has proven challenging with commonly used methods. Here, we demonstrate, in the context of a HIV-1 subtype C infected heterosexual transmission pair, a novel method to amplify and clone full-length infectious molecular clones in a faster, cheaper, and more accurate manner. Using these methods, we are able to show that a majority of genomes found in a chronically-infected individual are in fact functional and display a wide range in particle infectivity. We also compare for the first time, the particle infectivity of a full-length TF virus with respect to full-length non-transmitted (NT) viruses from the chronically-infected epidemiologically linked donor.

Results

Rapid high-fidelity single genome amplification of HIV-1 subtype C near-full-length genomes

Using similar methods as described previously (21,23), but with a newly-released high fidelity enzyme, Q5 Hot-Start Polymerase (New England Biolabs), we have amplified near full-length (NFL) single genomes by limiting dilution PCR from infected patient plasma in an acutely-infected linked recipient (Fiebig Stage II, 22 Days post estimated day of infection) and the chronically-infected epidemiologically-linked donor partner near the time of the transmission event (Table 1). Using a slightly modified master mix and cycling conditions from those recommended by the manufacturer (see methods), we shortened PCR amplification times and decreased the cost of reagents from previously employed methods using the Roche Expand Long Template enzyme. The improved fidelity of the Q5 enzyme as compared to other polymerases is highly advantageous, as single mutations in the genome can have a significant impact on viral fitness. In limiting dilution PCRs from full-length viral cDNA, we have also observed that the efficiency of NFL PCR was increased, since samples at the same dilution of cDNA, yielded increased numbers of positive reactions with Q5 than with Expand Long Template. Using the included GC enhancer, which enhances polymerase read-through of high GC content templates, further increased DNA yield (data not shown). NFL single genome nested PCR followed by direct Sanger sequencing of 8 NFL genomes amplified from a p24+/Ab- plasma sample of the linked recipient of a heterosexual transmission pair allowed us to determine the TF

sequence based on the consensus of all reads after confirming star-like phylogeny (as described previously (21))(Figure 1A).

In conjunction, we amplified 17 single NFL genomes from the plasma of the chronically-infected donor near the time of transmission. A maximum likelihood tree of the combined 25 viral sequences showed that: 1) the pair was epidemiologically linked, 2) subtype was confirmed by phylogenetic clustering with other known subtype C variants, and 3) the recipient was infected with a single variant from the donor quasispecies (Figure 1B). Three of the 17 NFL genomes from the chronically-infected donor's quasispecies contained either a stop codon or frameshift mutation in one of its genes, indicating that these three non-functional genomes would fail to produce functional virus following infection. The 14 remaining sequences from the donor, as well as the TF of the recipient, showed intact coding regions with no stop codons, frameshift mutations or APOBEC3G hypermutation, suggesting that the these have the potential to produce functional genomes. The donor's viral population showed extensive genetic diversity across the whole genome with differences of up to 4.2% between viruses. In order to determine the functionality of viruses circulating in a chronically-infected patient, we chose to generate infectious molecular clones (IMCs) for the recipient TF virus genome and for 8 of the 14 intact amplicons that represented the diversity of viruses from the chronically-infected donor.

Generation of a TF specific LTR vector from the linked recipient

A hallmark of retroviruses is that the virion-packaged RNA genomes do not contain the entire long terminal repeat (LTR), as the two full LTRs that flank the genome are generated during the reverse transcription steps following cell entry. The repeated nature of the LTRs makes it difficult to generate amplicons that contain the equivalent of the integrated proviral DNA required for cloning a transcriptionally functional genome. Previous methods for generation of full-length TF genomes have required a separate RNA amplification of the 5' LTR, followed by complex engineering of virus LTRs (9). Since the NFL amplification primers cover a portion of the viral LTR sequence, we chose to derive the unknown TF LTR sequence from LTRs amplified from genomic DNA isolated from a white cell pellet from the infected patient at an early time point after infection (43 days after the estimated date of infection) (Table 1). Using patient-specific LTR primers designed from the available TF sequence of the 5' end of U3 and 3' end of U5 characterized in the NFL amplicon, we amplified the ~650 bp LTR from patient genomic DNA via nested PCR (Figure 2A). This LTR amplicon was then cloned into a pBluescript vector that had 15bp of complimentary sequence at the ends by a two-piece fusion reaction using the In-Fusion HD Cloning Kit (Clontech). We then selected multiple clones and directly Sanger sequenced the entire LTR in order to infer the LTR sequence for the TF virus, based on the fact that there is limited diversity at early time points following infection. The confirmed LTR sequence matched the RNA-derived NFL TF sequence in all overlapping regions. Only the R sequence, which is where the primers are located in the NFL sequence, was newly inferred. This LTR vector

was then used to generate the missing pieces needed to form a full-length IMC from the available NFL genome amplicons (Figure 2A).

In-Fusion HD cloning of near full-length genomes into TF LTR

In order to generate IMCs using the already amplified NFL genomes that were sequenced, we utilized a ligation-independent cloning strategy using the Clontech In-Fusion HD system to perform a three-piece DNA fusion reaction followed by transformation of high transformation efficient chemically competent bacteria. Using the previously generated LTR vector as a template, we performed two independent PCR reactions that generate ~1,700 bp segments that contain half of the pBluescript vector and a majority of either U3 or U5 (Figure 2A). We then generated 9kb NFL amplicons amenable to this cloning strategy by reamplifying from the selected positive first round amplicons, using cloning primers that matched15bp sequences at the ends of the two LTR vector pieces (Figure 2B). In a 15-minute In-Fusion HD cloning reaction, the three pieces: 1) 1.7kb, 2) 1.8kb & 3) 9kb, generated a full-length molecular clone of the TF virus. For the NT viruses, we obtained a full-length IMC that is chimeric in the transcribed genome only for the TF virus R region (Figure 2C). Thus after a single round of replication these viruses would contain all sequences derived from the single genome amplicon except for the R region of the LTRs.

This process allowed us in a single experiment to clone IMCs for the amplicon corresponding to the TF virus, as well as the 8 viruses from the chronically infected epidemiologically-linked donor. Although clones were generated from PCR reactions from single HIV genomes, previous methods have

led to many PCR-induced mutations that generated clones which were distinct from the original single genome derived amplicons and that then required repair by subsequent mutagenesis protocols (9). In contrast, when we sequenced a single representative clone of the 9 IMCs, we observed that each sequence was identical to that of its cognate amplicon. The increase in fidelity that was observed following cloning, where an individual DNA PCR molecule out of the SGA amplicons is selected, suggests that the initial first round amplicons better reflect the sequence of the virus in the plasma. After additional Q5 amplification and subsequent cloning of amplicons from chronically-infected individuals, we have detected only a single base change between the amplicon and the cloned genome in more than 25 IMCs, representing approximately 225,000 nucleotides. This estimate is consistent with experiments performed by NEB where they detected 2 mismatches in clones following PCR amplification with Q5 in approximately 440,000 nucleotides sequenced (New England Biolabs).

We hypothesized, based on higher likelihood of sequence identity to the *in vivo* virus and the lack of PCR-based recombination, that IMCs derived by this protocol would yield viruses that exhibit the true phenotypic characteristics of viruses present in the diverse quasispecies of a chronically-infected individual, as well as those that establish infection and ultimately lead to disease.

Particle infectivity of TF & NT Variants from a linked heterosexual transmission pair

All HIV-1 full-length genome TF virus IMCs studied to date have been functional and replication competent, but the range of functionality and

replication competence of viruses in the plasma of a chronically-infected individual has not been fully elucidated. In order to determine IMC functionality, we generated virus stocks by transfection of 293T cells with the IMCs. We then determined the infectious titer of these stocks by infecting Tzm-bl cells, a CD4+CCR5+ reporter cell line which expresses ß-gal following infection. We determined the relative quantity of virus particles present in each stock by performing a radio-labeled reverse transcriptase (RT) assay, assuming consistent RT incorporation into virions. Particle infectivity ratios were then calculated by dividing the infectious titer by the RT activity of the stock. The particle infectivity ratios of the 9 IMC derived viruses is shown in Figure 3, and we furthermore demonstrated in three replicate transfections that all IMCs produced infectious viruses. The range of infectivities of viruses derived from one individual is more than one order of magnitude. Moreover, we saw evidence that viruses, which appear genetically similar on the phylogenetic tree (Figure 1B), exhibit similar particle infectivities (Figure 3). Variants F5, F11, and the TF, which are derived from the lower branch of the donor tree, have higher particle infectivities than the viruses (variants F6, F9, F14, F15, F21) from the upper branch of the phylogenetic tree. Variant F_8 IMC might be predicted to also have high particle infectivity, but has a very low infectivity to particle ratio, which we attribute to a six amino acid deletion at the C-terminal end of the Gag protein. Thus, in this female to male transmission pair, the TF that established infection in the new host came from the branch of more highly infectious viruses, as has been suggested by others (18,24), although it did not have the highest infectivity per particle of the donor quasispecies.

Although the R region of the donor IMCs was derived from the TF virus LTR, we confirmed, through mutagenesis of the Tar sequence, by reverting the sequence to its authentic sequence, that the chimeric nature of the Tar-Tat interaction in these viruses had no impact on the particle infectivity (Figure 4). For two viral varaints, F6 & F9 from the chronically infected donor, we compared the R sequence of the clone to the known R sequence of the original single genome amplicon (Figure 4A). We found that there were two base pair mismatches in the clone between the R of the TF LTR used for cloning and the R of the original amplicons for the two variants. Using site directed mutagenesis we reverted the two base pairs in the clone, to generate the variants F6 & F9 with their original R sequence, called LTR Mut. We determined the particle infectivity of the two clones from each variant, differing only by two base pairs in the Tar sequence (Figure 4B). We found that the two base pair change in the R region did not have any signficiant impact on virus particle infectivity. Although a slight decrease in infectivity was observed this does not account for the more than one log difference between all the variant IMCs.

Discussion

We describe here a method for amplification and cloning of HIV-1 full-length genome IMCs that is more cost-effective, higher throughput, and more precise than previous methods. Using this approach, we have sequenced and cloned the TF virus from a linked heterosexual transmission pair, as well as multiple single viral variants from the chronically-infected donor of that same pair. We have shown that each viral variant, including both TF and NT variants, produce infectious viruses following transfection. These methods provide evidence that, at least in this individual, a majority of the genetic variants circulating in the chronically-infected patient's plasma contain the genetic information for a functional virus and that those viruses exhibit a large range of infectivities. Previous studies have encountered difficulty in cloning functional full-length IMCs for HIV-1 and have therefore suggested that a significant percentage of circulating virus genomes are defective (18,20). Based on the results presented here, we suggest that this result is due to PCR-induced mutations accrued using previous methods, and that full-length RNA genomes circulating in a chronically-infected patient are predominantly functional. Although research on more transmission pairs will be required to determine specific fitness requirements for transmission, we show that, in this transmission pair, the TF virus has a higher particle infectivity than a majority of the viruses in circulation in the donor, as suggested previously (18,24), although it is not the most infectious.

This study also demonstrates that non-clonally expanded viral variants from a chronically-infected individual's quasispecies are infectious and are

capable of being high replicators *in vitro*. Barring a gross genetic defect, the sampled RNA genomes found *in vivo* are functional *in vitro*. Studies of Enterovirus 71 have shown that cloning from half genomes and construction of recombinant genomes versus cloning from full-length PCR amplicons led to less functional clones from the same virus stock which again suggests that laboratory-induced recombination is detrimental to virus integrity (25). Although viruses such as HIV-1 can often generate recombinants *in vivo*, these likely emerge only after *in vivo* selection for replication competent viruses and that in the absence of this selection *in vitro,* constructed chimeras will more likely produce defective genomes. It will be important in the future to determine the fitness diversity of viral populations in HIV-1 infected individuals, since this may both predict transmissibility and the impact of the virus on pathogenesis in that host (12,24,26).

Hepatitis C virus (HCV) is another highly diverse RNA virus with a large global health burden that suffers from inadequate laboratory models. The difficulty to culture and grow HCV *in vitro* stems from a lack of an adequate cell culture system as well as difficulty in generating functional clones. Studies from 1997 showed that HCV clones were non-functional due to PCR-induced or cloning-induced mutations (27,28). The high fidelity system described here for generating IMCs could be used in the HCV field to screen a larger number of viral strains for *in vitro* growth. It would also allow a better estimate of viral fitness, which would not rely on restriction based cloning systems or necessitate extra sequencing costs. Similarly to HIV-1, the genomic instability of viruses such as Hepatitis C and Respiratory Syncytial virus create obstacles to the construction of

infectious clones and could benefit from accurate amplification used in this method (27,29). The combination of high fidelity DNA amplification and an enzymatic restriction site independent method of cloning provides a valuable tool for more detailed *in vitro* studies of full-length viral genomes.

Materials and Methods

Viral RNA Extraction & cDNA Synthesis

Viral RNA was extracted from plasma and converted to full-length cDNA as described previously(21). In brief, 140µl of patient plasma was used to extract viral RNA using the QIAamp Viral RNA Mini Kit (Qiagen). RNA was recovered and immediately used to synthesize cDNA using the SuperScript III Reverse Transcriptase (LifeTechnologies) enzyme with an anchor oligo-dT primer (Table 2).

Near Full Length Single Genome Amplification

cDNA was serially diluted in replicates of eight PCR wells and subjected to nested PCR amplification with HIV-specific primers that yield a 9-kb fragment beginning at the first nucleotide of the U5 region of the 5' long terminal repeat (LTR) and extending to the last nucleotide of the R region of the 3' LTR. cDNA dilutions were tested to identify a dilution where ~30% of wells were positive for amplification products (30). First-round PCR was performed in 1x Q5 Reaction Buffer, 1x Q5 High GC Enhancer, 0.35 mM of each dNTP, 0.5µM of primers 1.U5Cc and 1.3'3'PlCb (Table 2), and 0.02 U/µl of Q5 Hot Start High-Fidelity DNA Polymerase (NEB) in a total reaction volume of 25µl. PCR conditions for the first round are: 98^{o}C for 30s, followed by 30 cycles of 98^{o}C for 10s, 72^{o}C for 7.5 min, with a final extension of 72^{o}C for 10 min. 1µl of first round PCR product was then used as template for the second round PCR, with identical cycling conditions and PCR mix except for the primers. The second round primers are 2.U5Cd and

2.3'3'plCb (Table 2). PCR reactions were then run on a 1% agarose lithium acetate gel at 300V for 25 min in order to determine the presence of a 9kb band.

LTR Amplification

White cell pellet from acutely-infected patients were used to extract genomic DNA using the Qiagen DNeasy Blood and Tissue kit (Qiagen). Patient-specific primers determined from the sequence of their Transmitted/Founder U3 and U5 regions were generated to directly amplify the entire LTR from genomic DNA (~650 bp) in a nested PCR reaction (Figure 2A). First round primers, LTR_For1 and LTR_Rev1, are the first 19 bps of U3 and the reverse complement of the last 19bp of U5, respectively. First round PCR was performed in a total reaction volume of 50µl containing 1x Q5 Reaction Buffer, 1x GC Enhancer, 0.35 mM of each dNTP, 0.5 µM of each primer, 0.02 U/µl of Q5 enzyme, and 1 µl of DNA template. PCR conditions were: 98°C for 30s, followed by 30 cycles of 98°C for 10s, 59°C for 30s, and 72°C for 20s, with a final extension of 72°C for 2 min. 1µl of first round PCR product was then used as template for the second round PCR in a total volume of 50µl with similar components as first round, except for the primers. Second round forward primers, LTR_For2 and LTR_Rev2, are composed of the first 33 bps of patient U3, and the reverse complement of the last 34 bps of U5. PCR conditions for the second round were: 98°C for 30s, followed by 30 cycles of 98°C for 10s and 72°C for 40s, with a final extension of 72°C for 2 min. The amplicons were sized on a 1% agarose lithium acetate gel, and any positive bands were extracted using the Promega Wizard SV Gel Kit (Promega) and eluted into 40µl Nuclease-free water.

Cloning of Patient-Specific LTR Vector

In order to generate a patient-specific LTR vector, we generated a linear vector amplicon containing a bacterial origin of replication and ampicillin resistance cassette with 15 bp ends that match the 15 bp ends of the patient LTR by PCR amplification. Primers for our vector were LTRVec_For and LTRVec_Rev, where the underlined sequences represent sequences that will overlap with patient specific LTR (Table 2). We amplified the vector piece from the pBluescript plasmid using Q5, where PCR conditions were: 98^{o}C for 30s, followed by 35 cycles of 98^{o}C for 10s and 72^{o}C for 3 min, with a final extension of 72^{o}C for 2 min. The ~3,100 bp vector piece was gel purified as described previously. In order to combine the patient-specific LTR and the vector piece we used the Clontech In-Fusion HD cloning kit. Briefly, in a PCR tube, 100ng of patient-specific LTR, 100ng of Vector, 2µl of the In-Fusion HD 5x Enzyme Premix, were mixed with nuclease-free water to make 10µl final reaction volume. The reaction was incubated at 50^{o}C for 15 min, and then held at 4^{o}C until used for transformation. Stellar Competent cells were transformed as specified by the manufacturer's protocol (Clontech). Briefly, competent cells were thawed on ice, 50µl of competent cells was transferred to a cold 1.5 ml microcentrifuge tube, to which 2.5µl of In-Fusion HD cloning reaction was added followed by incubation on ice for 30 min. The cells were heat shocked in a heat block at 42^{o}C for 1 min, and then returned to ice for 2 minutes. A volume of 450µl of room temperature SOC medium was then added, and 100 µl of cells plated onto a Luria Broth/agar plate supplemented with 100 µg/ml Ampicillin. The plates were incubated at

30°C overnight, after which multiple colonies were picked and miniprep DNA prepared using the PureYield Plasmid Miniprep System (Promega). The LTR vector was directly Sanger sequenced and a consensus LTR that matched in sequence to the inferred TF was selected.

Full-Length Amplicon Cloning into TF LTR

The previously generated LTR vector is used to generate two PCR products used in the final cloning reaction that will encode the missing LTR regions from the NFL amplicons (Figure 2A). In two separate PCR reactions, we generate the two linear pieces from the LTR vector. The first piece was amplified by PCR in 1x Q5 Reaction Buffer, 1x Q5 High GC Enhancer, 0.35 mM of each dNTP, 0.5µM of primers Vec_1_For and Vec_1_R (Table 2), and 0.02 U/µl of Q5 Hot Start High-Fidelity DNA Polymerase (New England Biolabs) in a total reaction volume of 25µl. PCR conditions were: 98°C for 30s, followed by 30 cycles of 98°C for 10s, 66°C for 30s and 72°C for 1 min, with a final extension of 72°C for 2 min. The second vector piece was amplified using the same reaction mixture except with primers Vec_2_For and Vec_2_Rev, with the following cycling conditions: 98°C for 30s, followed by 30 cycles of 98°C for 10s, 67°C for 30s and 72°C for 1 min, with a final extension of 72°C for 2 min. The cloning NFL amplicon was derived by reamplification from the selected first round amplicon using cloning primers 3618_Clone_For and 3618_Clone_Rev (Figure 2B). PCR conditions were identical to the NFL PCR described previously except for cycling conditions which were: 98°C for 30s followed by 30 cycles of 98°C for 10s, 60°C for 30s, and 72°C for 7.5 min, with a final extension at 72°C for 10min. The three

PCR products were independently gel extracted as described previously. In order to generate the full-length clone, a final In-Fusion HD reaction with the 9kb NFL cloning amplicon was combined with the 2 LTR vector pieces, which complement the missing LTR regions, in a 10µl reaction with 150ng of each linear gel purified amplicon, 2µl of 5x In-Fusion HD Enzyme Mix, and nuclease free (Figure 2C). The reaction was placed at 50°C for 15 min, and then kept at 4°C until used for transformation. Stellar Competent cells (Clontech) were transformed, and DNA minipreps of colonies were performed as described above. We confirmed correct cloning by restriction digest and gel visualization, followed by growth of appropriate clones in 250ml Luria Broth plus ampicillin in order to purify with the PureYield Plasmid Maxiprep System (Promega). All clonal virus inserts were completely sequenced to confirm identity to the original NFL amplicon (see below).

Generation of Virus Stocks & Particle Infectivity

1.5µg of proviral plasmid clones were used to transfect 293T cells with Fugene-HD Transfection reagent (Promega). 293T cell supernatants were collected 72 hours after transfection and clarified by centrifugation. Virus stocks were then titered on the Tzm-bl reporter cell line as described previously(26,31). The virus stocks were also directly analyzed for RT activity using a radio-labeled RT assay as described previously(31,32).

Sequencing

All amplicons and clones were directly Sanger sequenced, aligned, and analyzed as has been described previously(21,30).

Acknowledgements

We respectfully recognize the Rwanda-Zambia HIV Research Group (RZHRG) and Zambia-Emory HIV Research Project (ZEHRP) staff and participants for their invaluable contributions to this study. The Emory University Institutional Review Board and the University of Zambia Research Ethics Committee approved these studies. This research was supported by grants from the National Institutes of Health (R37 AI-51235 (EH)) and the International AIDS Vaccine Initiative (SA). The Fogarty AITRP Grant FIC 2D43 TW001042 supported WK. This project was also funded in part by the National Center for Research Resources P51RR165 and the Office of Research Infrastructure Programs/ODP51OD11132 for the Yerkes National Primate Research Center.

Table 1 - Characteristics of the samples from the subtype C HIV-1 transmission pair.

ID	Status[a]	EDI[b]	Plasma date	DNA date[c]	VL[d]
Z3618M	LR	5/27/09	6/18/09	7/9/09	16,800,000
Z3618F	D	N/A	7/11/09		20,400

[a] LR – epidemiologically linked recipient partner, D – Donor partner
[b] EDI – Estimated date of infection
[c] DNA Date – Date of collection of PBMCs from which DNA was extracted
[d] VL – Viral Load – on 6/18/09 for LR, 7/15/09 for D

Table 2 - DNA primers used in the study.

Primer name	Sequence (5′ - 3′)	Application	Color code (Fig. 2)
Oligo-dT	TTTTTTTTTTTTTTTTTVN	cDNA synthesis	Not shown
1U5Cc	CCTTGAGTGCTCTAAGTAGTGTGTGCCCGTCTGT	1st round NFL PCR	Not shown
2U5Cd	AGTAGTGTGTGCCCGTCTGTTGTGTGACTC	2nd round NFL PCR	Not shown
1.3′3′plCb	ACTACTTAGAGCACTCAAGGCAAGCTTTATTG	1st round NFL PCR	Not shown
2.3′3′plCb	TAGAGCACTCAAGGCAAGCTTTATTGAGGCTTA	2nd round NFL PCR	Not shown
LTRVec_For	GGAAAATCTCTAGCAGTGGCGCCCGAACAGGGA	LTR vector – pBluescript piece	Not shown
LTRVec_Rev	AAATCAACCCATCCACGAGCTCGCCCGCGGTG	LTR vector – pBluescript piece	Not shown
LTR_For1	TGGATGGGTTGATTTACTC	LTR PCR	Black
LTR_For2	TGGATGGGTTGATTTACTCCAAGAAAAGGCAAG	LTR PCR	Red
LTR_Rev1	TGCTAGAGATTTTCCACAC	LTR PCR	Black
LTR_Rev2	TGCTAGAGATTTTCCACACTACCAAAATGGTCTG	LTR PCR	Red
3618_Clone_For	GGTAACTAGAGATCCCTCAG	NFL cloning PCR	Cyan
3618_Clone_Rev	TGCTTATATGCAGCATCTG	NFL cloning PCR	Cyan
Vec_1_For	TGCTGCATATAAGCAGCTGC	NFL cloning vector piece PCR	Cyan
Vec_1_Rev	CTGTAGCAATGGCAACAACG	NFL cloning vector piece PCR	Blue
Vec_2_For	TTGCCATTGCTACAGGCATCG	NFL cloning vector piece PCR	Blue
Vec_2_Rev	GGATCTCTAGTTACCAGAGTCACAC	NFL cloning vector piece PCR	Cyan

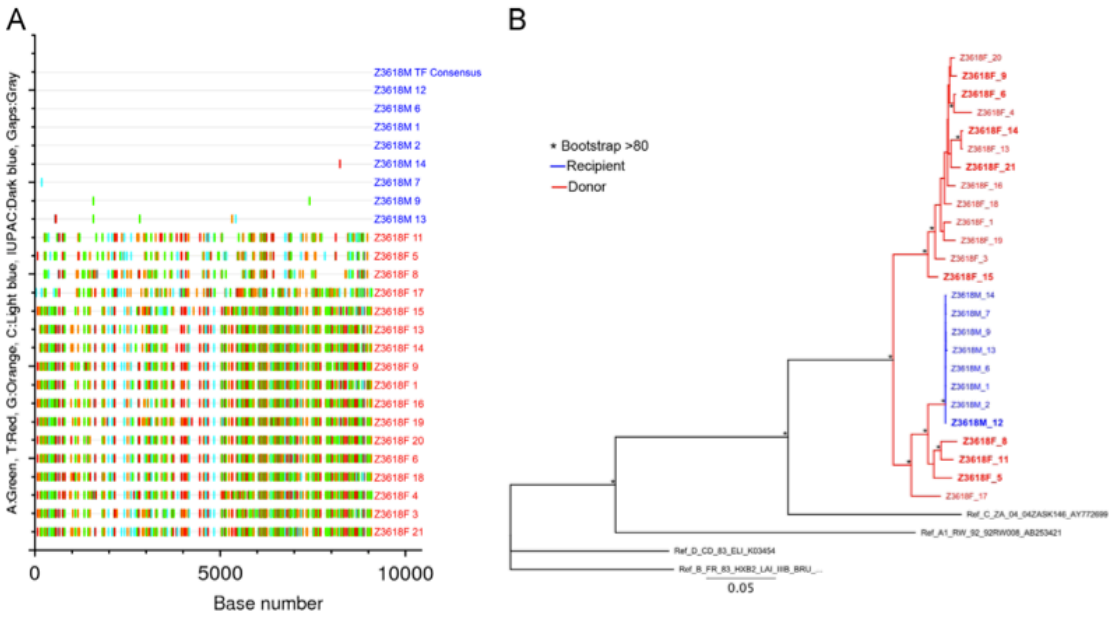


Figure 1 - Highlighter and phylogenetic analysis of HIV-1 near full-length single genome amplicons from a heterosexual epidemiologically-linked transmission pair.

(A) Highlighter analysis of the 25 single genome amplicons from both linked recipient and donor, where the master sequence is set as the TF consensus sequence. (B) Maximum likelihood tree of the near full-length genomes from a heterosexual epidemiologically-linked transmission pair. In blue are 8 amplicons from the linked recipient, after transmission. Near full-length amplicons from the donor are shown in red. HIV-1 reference sequences are denoted in black. Viral variants selected for cloning are indicated by bold lettering.

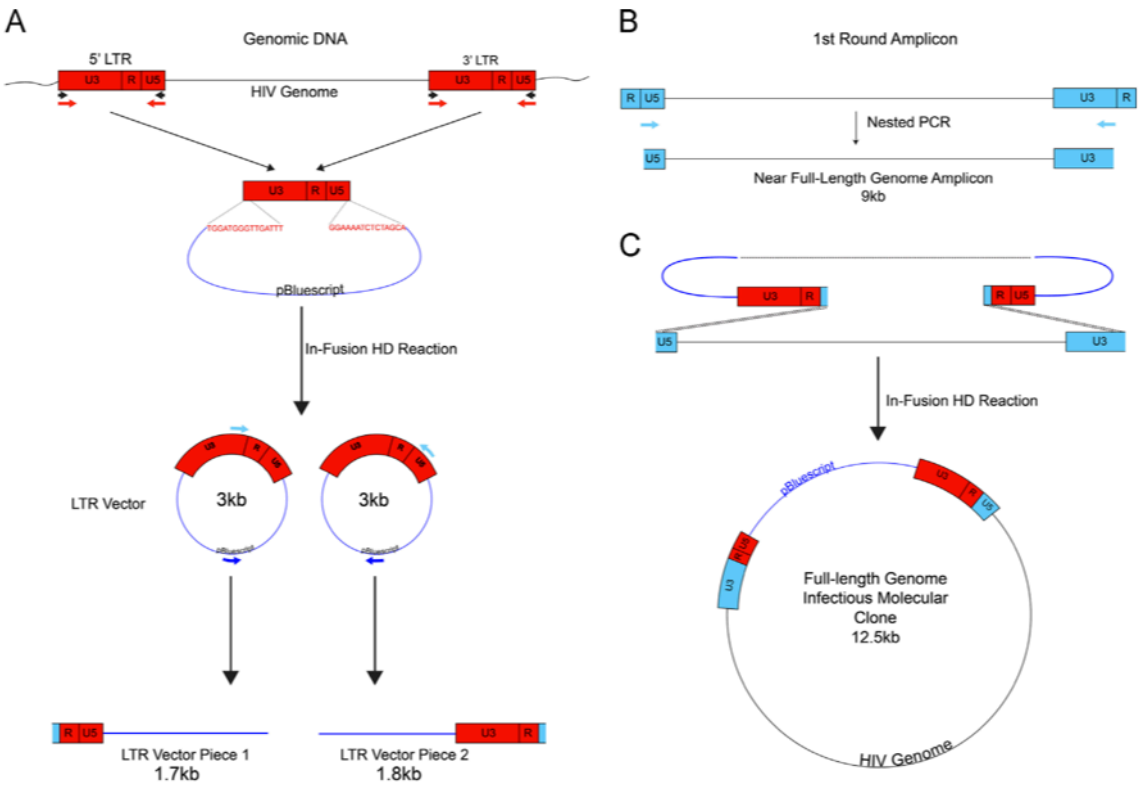


Figure 2 - Diagram for the full-length genome cloning strategy. (A) First, the LTR of the TF was amplified from genomic DNA and cloned using the In-Fusion system.

The LTR vector was then sequenced and a consensus LTR that matched in sequence to the inferred TF was selected. This LTR vector was used to generate two PCR products used in the final cloning reaction that will encode the missing LTR regions from the NFL amplicon. (B) Previously amplified single genomes from the recipient and donor partners were selected and reamplified using cloning primers with sequences that overlap with the ends of the two pieces amplified from the LTR vector. (C) In a final In-Fusion reaction the 9kb NFL amplicon was cloned with the 2 LTR vector pieces that complement the missing

LTR portions. This process yields a full-length molecular clone that contains all the sequence information from the NFL amplicon and has the TF R regions.

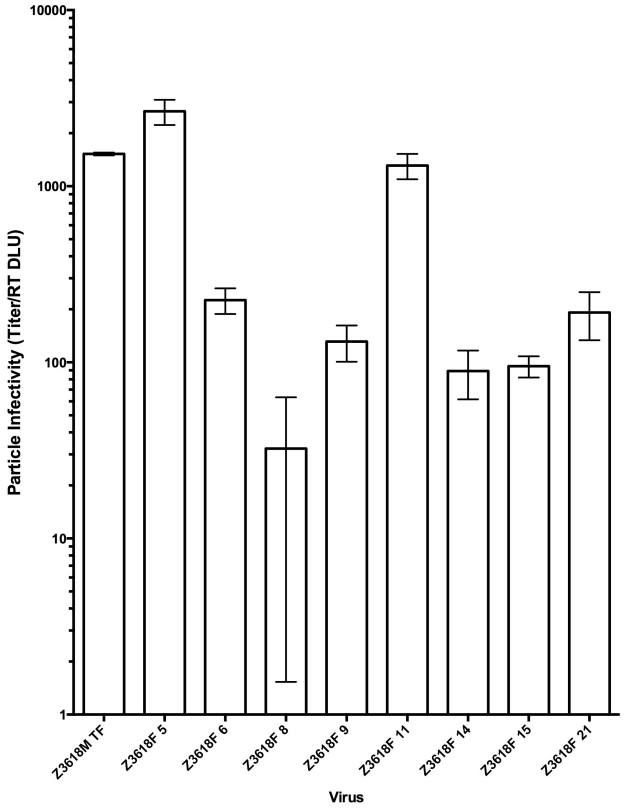


Figure 3 - Particle infectivities of HIV full-length genome infectious molecular clones from a linked heterosexual transmission pair.

Particle infectivity is measured by the infectious titer as determined on Tzm-bl cells, divided by the reverse transcriptase level of the stock measured via radio-labeled RT assays. The particle infectivity of each virus variant was measured for three separate virus stocks generated from independent 293T transfections, where error bars represent the standard error of the mean.

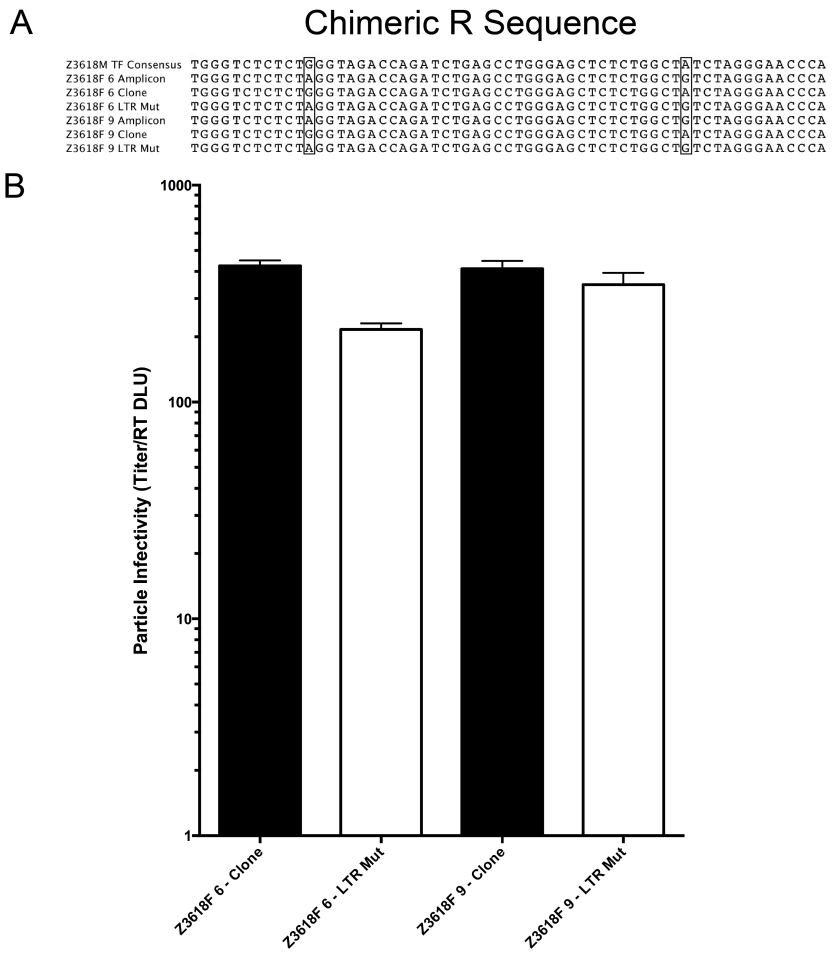


Figure 4 - Particle infectivities of two viral variants with the chimeric R and original amplicon R sequences.

(A) Chimeric R sequences of two single genome amplicons, from variants F6 and F9 from the donor partner are shown along with the two base pair mismatch found in the Tar region of the clone as compared to the amplicon. The mismatch was mutated back to the correct Tar sequence via site directed mutagenesis generating a LTR Mut clone for each virus. (B) The particle infectivity of the original clone versus the LTR Mut was compared for the two viruses and no significant difference in infectivity was seen.

References

1. UNAIDS. Global Report 2013. 2013Nov.5;:1–198.

2. Dunkle KL, Stephenson R, Karita E, Chomba E, Kayitenkore K, Vwalika C, et al. New heterosexually transmitted HIV infections in married or cohabiting couples in urban Zambia and Rwanda: an analysis of survey and clinical data. Lancet. 2008Jun.28;371(9631):2183–91.

3. Phillips RE, Rowland-Jones S, Nixon DF, Gotch FM, Edwards JP, Ogunlesi AO, et al. Human immunodeficiency virus genetic variation that can escape cytotoxic T cell recognition. Nature. 1991Dec.12;354(6353):453–9.

4. Richman DD, Wrin T, Little SJ, Petropoulos CJ. Rapid evolution of the neutralizing antibody response to HIV type 1 infection. P Natl Acad Sci Usa. 2003Apr.1;100(7):4144–9.

5. Coffin J, Swanstrom R. HIV Pathogenesis: Dynamics and Genetics of Viral Populations and Infected Cells. Cold Spring Harb Perspect Med. 2013Jan.2;3(1):a012526–6.

6. Arts EJ, Hazuda DJ. HIV-1 Antiretroviral Drug Therapy. Cold Spring Harb Perspect Med. 2012Apr.2;2(4):a007161–1.

7. Sanchez G, Xu X, Chermann JC, Hirsch I. Accumulation of defective viral genomes in peripheral blood mononuclear cells of human immunodeficiency virus type 1-infected individuals. J Virol.

1997Mar.;71(3):2233–40.

8. Eriksson S, Graf EH, Dahl V, Strain MC, Yukl SA, Lysenko ES, et al. Comparative Analysis of Measures of Viral Reservoirs in HIV-1 Eradication Studies. Douek DC, editor. PLoS Pathog. 2013Feb.14;9(2):e1003174.

9. Ochsenbauer C, Edmonds TG, Ding H, Keele BF, Decker J, Salazar MG, et al. Generation of Transmitted/Founder HIV-1 Infectious Molecular Clones and Characterization of Their Replication Capacity in CD4 T Lymphocytes and Monocyte-Derived Macrophages. J Virol. 2012Feb.11;86(5):2715–28.

10. Fenton-May AE, Dibben O, Emmerich T, Ding H, Pfafferott K, Aasa-Chapman MM, et al. Relative resistance of HIV-1 founder viruses to control by interferon-alpha. Retrovirology. 2013;10:146.

11. Meyerhans A, Cheynier R, Albert J, Seth M, Kwok S, Sninsky J, et al. Temporal fluctuations in HIV quasispecies in vivo are not reflected by sequential HIV isolations. Cell. Elsevier; 1989;58(5):901–10.

12. Fraser C, Lythgoe K, Leventhal GE, Shirreff G, Hollingsworth TD, Alizon S, et al. Virulence and Pathogenesis of HIV-1 Infection: An Evolutionary Perspective. Science. 2014Mar.20;343(6177):1243727–7.

13. Shaw GM, Hunter E. HIV Transmission. Cold Spring Harb Perspect Med. 2012;2(11).

14. Wolfs TF, Zwart G, Bakker M, Goudsmit J. HIV-1 genomic RNA diversification following sexual and parenteral virus transmission.

Virology. 1992Jul.;189(1):103–10.

15. Wolinsky SM, Wike CM, Korber BT, Hutto C, Parks WP, Rosenblum LL, et al. Selective transmission of human immunodeficiency virus type-1 variants from mothers to infants. Science. 1992;255(5048):1134.

16. Keele BF, Giorgi EE, Salazar-Gonzalez JF, Decker JM, Pham KT, Salazar MG, et al. Identification and characterisation of transmitted and early founder virus envelopes in primary HIV-1 infection. P Natl Acad Sci Usa. 2008;105(21):7552–7.

17. Derdeyn CA, Decker JM, Bibollet-Ruche F, Mokili JL, Muldoon M, Denham SA, et al. Envelope-constrained neutralization-sensitive HIV-1 after heterosexual transmission. Science. 2004Mar.26;303(5666):2019–22.

18. Parrish NF, Gao F, Li H, Giorgi EE, Barbian HJ, Parrish EH, et al. Phenotypic properties of transmitted founder HIV-1. Proceedings of the National Academy of Sciences. 2013Apr.23;110(17):6626–33.

19. Ndung'u T, Renjifo B, Essex M. Construction and analysis of an infectious human Immunodeficiency virus type 1 subtype C molecular clone. J Virol. 2001Jun.;75(11):4964–72.

20. Ndung'u T. Molecular Cloning and Biological Characterization of Full-Length HIV-1 Subtype C from Botswana. Virology. 2000Dec.20;278(2):390–9.

21. Salazar-Gonzalez JF, Salazar MG, Keele BF, Learn GH, Giorgi EE, Li H, et al. Genetic identity, biological phenotype, and evolutionary pathways of transmitted/founder viruses in acute and early HIV-1 infection. Journal of Experimental Medicine. 2009Jun.8;206(6):1273–89.

22. Baalwa J, Wang S, Parrish NF, Decker JM, Keele BF, Learn GH, et al. Molecular identification, cloning and characterization of transmitted/founder HIV-1 subtype A, D and A/D infectious molecular clones. Virology. 2012Nov..

23. Rousseau CM, Birditt BA, McKay AR, Stoddard JN, Lee TC, McLaughlin S, et al. Large-scale amplification, cloning and sequencing of near full-length HIV-1 subtype C genomes. J. Virol. Methods. 2006Sep.;136(1-2):118–25.

24. Carlson JM, Schaefer M, Monaco DC, Batorsky R, Claiborne DT, Prince J, et al. Selection bias at the heterosexual HIV-1 transmission bottleneck. Science. 2014Jul.10;345(6193):1254031–1.

25. Lazouskaya NV, Palombo EA, Poh C-L, Barton PA. Construction of an infectious cDNA clone of Enterovirus 71: Insights into the factors ensuring experimental success. J. Virol. Methods. Elsevier B.V; 2014Mar.1;197:67–76.

26. Prince JL, Claiborne DT, Carlson JM, Schaefer M, Yu T, Lahki S, et al. Role of Transmitted Gag CTL Polymorphisms in Defining Replicative Capacity and Early HIV-1 Pathogenesis. Lifson J, editor. PLoS Pathog. 2012Nov.29;8(11):e1003041.

27. Yanagi M, Purcell RH, Emerson SU, Bukh J. Transcripts from a single full-length cDNA clone of hepatitis C virus are infectious when directly transfected into the liver of a chimpanzee. P Natl Acad Sci Usa. 1997Aug.5;94(16):8738–43.

28. Lohmann V, Bartenschlager R. On the History of Hepatitis C Virus Cell Culture Systems. J. Med. Chem. 2014Mar.13;57(5):1627–42.

29. Hotard AL, Shaikh FY, Lee S, Yan D, Teng MN, Plemper RK, et al. A stabilized respiratory syncytial virus reverse genetics system amenable to recombination-mediated mutagenesis. Virology. Elsevier; 2012Dec.5;434(1):129–36.

30. Salazar-Gonzalez JF, Bailes E, Pham KT, Salazar MG, Guffey MB, Keele BF, et al. Deciphering Human Immunodeficiency Virus Type 1 Transmission and Early Envelope Diversification by Single-Genome Amplification and Sequencing. J Virol. 2008Mar.25;82(8):3952–70.

31. Claiborne DT, Prince JL, Hunter E. A restriction enzyme based cloning method to assess the in vitro replication capacity of HIV-1 subtype C Gag-MJ4 chimeric viruses. Journal of Visualized Experiments.

32. Current Protocols in Immunology. Hoboken, NJ, USA: John Wiley & Sons, Inc; 2001.

Chapter III : Heterosexual Transmission of Subtype C HIV-1 Selects Consensus-like Variants Without Increased Replicative Capacity or Interferon-α Resistance

Martin J. Deymier[1¶], Zachary Ende[1¶], Angharad E. Fenton-May[2], Dario A. Dilernia[1], William Kilembe[3], Susan A. Allen[4], Persephone Borrow[2], Eric Hunter[1,4*]

[1] Emory Vaccine Center at Yerkes National Primate Research Center, Emory University, Atlanta, Georgia, United States of America

[2] Nuffield Department of Medicine, University of Oxford, Oxford, United Kingdom

[3] Zambia-Emory HIV Research Project, Lusaka, Zambia

[4] Department of Pathology and Laboratory Medicine, Emory University, Atlanta, Georgia, United States of America

* Corresponding author

E-mail: ehunte4@emory.edu

¶ These authors contributed equally to this work.

Published in *Plos Pathogens (2015)*

Abstract

Heterosexual transmission of HIV-1 is characterized by a genetic bottleneck that selects a single viral variant, the transmitted/founder (TF), during most transmission events. To assess viral characteristics influencing HIV-1 transmission, we sequenced 167 near full-length viral genomes and generated 40 infectious molecular clones (IMC) including TF variants and multiple non-transmitted (NT) HIV-1 subtype C variants from six linked heterosexual transmission pairs near the time of transmission. Consensus-like genomes sensitive to donor antibodies were selected for during transmission in these six transmission pairs. However, TF variants did not demonstrate increased viral fitness in terms of particle infectivity or viral replicative capacity in activated peripheral blood mononuclear cells (PBMC) and monocyte-derived dendritic cells (MDDC). In addition, resistance of the TF variant to the antiviral effects of interferon-α (IFN-α) was not significantly different from that of non-transmitted variants from the same transmission pair. Thus neither *in vitro* viral replicative capacity nor IFN-α resistance discriminated the transmission potential of viruses in the quasispecies of these chronically infected individuals. However, our findings support the hypothesis that within-host evolution of HIV-1 in response to adaptive immune responses reduces viral transmission potential.

Author Summary

Despite the available HIV-1 diversity present in a chronically infected individual, single viral variants are transmitted in 80-90% of heterosexual transmission events. These breakthrough viruses may have unique properties that confer a higher capacity to transmit. Determining these properties could help inform the rational design of vaccines and enhance our understanding of viral transmission. We isolated the transmitted variant and a set of related non-transmitted variants from the transmitting partner near the estimated date of transmission from six epidemiologically linked transmission pairs to investigate viral correlates of transmission. The simplest explanation that transmitted variants are inherently more infectious or faster replicators *in vitro* did not hold true. In addition, transmitted variants did not replicate more efficiently than their non-transmitted counterparts in dendritic cells or in the presence of interferon-alpha *in vitro,* suggesting that they are not uniquely adapted to these components of the innate immune system. More ancestral genomes that were relatively sensitive to antibody neutralization tended to transmit, supporting previous reports that mutational escape away from the adaptive immune response likely reduces the ability to transmit. Our investigation into the traits of transmitted HIV-1 variants adds to the understanding of viral determinants of transmission.

Introduction

HIV-1 transmission is characterized by an extreme genetic bottleneck, the basis of which is unclear. Studies of both the highly diverse envelope glycoprotein (1-3) and full HIV-1 genomes (4) demonstrated that 80-90% of heterosexual transmissions are initiated by a single virus variant selected from the diverse viral quasispecies present in the chronically infected transmitting partner. These variants, which are different in each transmission event, have been named transmitted/founder (TF) viruses. Studying TF viruses could enhance our understanding of viral transmission and inform HIV prevention strategies.

The TF is rarely the dominant variant in the plasma or genital tract of the transmitting partner (5,6), which suggests that transmission is not entirely stochastic and may involve selection. A number of prior studies have identified distinctive properties of TF variants (4,7-19), particularly in analyses of the TF viral envelope (Env) glycoprotein. Reported characteristics of TF virus Envs include a selection for CCR5-tropism (2,20), a predominance of shorter and less glycosylated Env proteins (1,11,15,18,19), a preference for binding α4β7 (10,21) and a selection for more ancestral variants (8,22). Although these studies observed selection of viral traits, others found that acute and chronic variants had similar characteristics. By generating infectious molecular clones (IMC) with the *env* genes from linked recipients and transmitting partners in a common viral backbone, acute and chronic donor viruses displayed similar CD4 and CCR5 requirements for cell entry, low macrophage tropism, and no preferential usage of alternative coreceptors (23,24). Furthermore, studies of *env* only clones from acute infection compared with chronic control viruses have shown similar CD4 T

cell subset tropism, low macrophage tropism, and a lack of effect of blocking α4β7 on infection (25).

Selection of viral traits outside of the *env* gene has also been observed during heterosexual transmission. We recently described a selection bias during transmission for more consensus-like HIV-1 variants, in *gag*, *pol* and *nef* genes, from a cohort of 137 subtype C infected epidemiologically-linked transmission pairs (7). This study suggested that *in vivo* fitness of consensus-like HIV-1 variants increased their likelihood of transmission (7). Studies of full-length infectious molecular clones of TF viruses, in comparison to control viruses derived from chronic infection, have also demonstrated increased particle infectivity, as well as an enhanced resistance to interferon-α (IFN-α) in TF viruses (13,17).

While informative, conclusions of these previous studies are limited in that only individual genes were examined, or corresponding non-transmitted (NT) variants from the transmitting partner were unavailable as controls. HIV-1 IMC with the full complement of HIV-1 proteins have not been generated from both partners of transmission pairs nor evaluated for genetic and phenotypic signatures during transmission. Characterizing TF variants in comparison to NT variants from epidemiologically-linked partners could provide further insight into the viral requirements of HIV-1 transmission, potentially leading to new targets for intervention.

Here, we describe genetic and phenotypic comparisons of full-length genome TF and NT variants from six subtype C epidemiologically-linked heterosexual transmission pairs. We amplified and sequenced near full-length

HIV-1 genomes by single genome amplification (SGA) to assess genetic selection during transmission. In addition, we cloned the complete TF genome along with a representative panel of NT variants. These clones were used to assess the relative *in vitro* fitness of TF variants as measured by particle infectivity, neutralizing antibody resistance, replicative capacity in PBMC and dendritic cells, as well as IFN-α resistance. We found a strong selection bias toward consensus sites across the entire genome, at both the amino acid and nucleotide level, in all six pairs. The TF variants were also more sensitive to neutralization by donor antibodies than NT variants. However, no evidence was found for TF variants exhibiting increased particle infectivity, replicative capacity, or IFN-α resistance when compared to the transmitting partner's NT variants. Thus, in these six subtype C transmission pairs the transmission potential of TF variants is not discriminated by inherent *in vitro* replicative capacity or interferon resistance, and may be determined by alternate phenotypes difficult to dissect in these *in vitro* systems.

Results

Transmission pairs and amplification of near full-length genomes

Full-length genome HIV-1 variants derived from linked transmission pairs have yet to be evaluated for characteristics associated with transmission. To define whether TF variants exhibit distinct properties, we compared them to their NT counterparts in six heterosexual epidemiologically-linked transmission pairs. We selected five female-to-male and one male-to-female therapy-naïve subtype C epidemiologically-linked transmission pairs from the Zambia-Emory HIV Research Project (ZEHRP) based on the availability of plasma samples at the nearest time points following transmission (average 28 days post estimated date of infection) (Table 1). We PCR amplified, using a high-fidelity polymerase, and sequenced a total of 167 HIV-1 near full-length single genome amplicons as described previously (26). All six linked recipients were in Fiebig Stage II of infection, and were infected with a single variant from the donor quasispecies, as demonstrated by star-like phylogeny in a median of 8 near full-length genome amplicons per sample (2,4). This allowed us to infer an unambiguous consensus TF sequence from the genetically homogeneous population of sequences in each linked recipient.

For phylogenetic analyses, we aligned full-length nucleotide sequences as well as concatenated full proteome amino acid sequences of 115 HIV-1 single genomes (each TF virus represented by a single consensus sequence), with the HIV-1 consensus/ancestral alignment from the Los Alamos National Laboratory (LANL) HIV database. We generated maximum likelihood trees of the full-length

genome and proteome alignments for all six transmission pairs, and confirmed that all pairs were epidemiologically linked, since each TF variant fell clearly within the branches of the linked donor virus variants. Each transmission pair clustered independently on the phylogenetic tree with bootstrap values of 100 (Fig. 1). All six linked donor partners were chronically infected and demonstrated viral diversity in their plasma near the time of transmission (Fig. 1).

Consensus-like nature of TF and NT viruses

We previously demonstrated a consistent transmission bias for variants with consensus-like amino acid residues across the Gag, Pol and Nef proteins by population sequencing in a cohort of 137 epidemiologically-linked subtype C transmission pairs (7). Although this finding has been shown for the *gag* and *env* genes independently, it has not been confirmed by full-length genome SGA from the transmitting partner's quasispecies (7,8). We examined the selection bias for more consensus-like viruses by measuring the pairwise distance (branch length), of each viral variant to the LANL subtype C consensus node on the full-length nucleotide and amino acid phylogenetic trees (Fig. 1). TF variants had a significantly shorter pairwise distance to the subtype C consensus node than the median of their corresponding NT variants for both nucleotide (Fig. 2A; $p = 0.0156$) and amino acid (Fig. 2B; $p = 0.0469$) sequences. These transmission pairs confirm, as previously described, a selection bias for consensus-like amino acid and nucleotide sites across the viral genome during transmission.

Particle infectivity of TF and NT viruses

In a previous study, TF virions exhibited enhanced infectivity in comparison to chronic control viruses on TZM-bl cells (13). To test particle infectivity within transmission pairs, we generated full-length IMC for 40 viral variants, including the 6 TF variants and 3-8 NT variants from each chronically infected transmitting partner, as described previously (26). We selected variants to represent the genetic diversity present in the donor near the time of transmission (Fig. 1), and confirmed that the IMC and amplicon sequences were identical by whole genome sequencing. We also excluded the rare sequences that contained gross genetic defects, such as large deletions and frameshift mutations in gene coding regions. For each IMC, we generated virus stocks by transfection of 293T cells.

We defined particle infectivity as the ratio of infectious units, as measured by the virus titer on TZM-bl cells, a standard reporter cell line whose permissivity correlates with that of PBMC (27), over total amount of virions, measured by reverse transcriptase activity of the virus stock. We confirmed that the particle infectivity of a subset of virus stocks generated from 293T cells and harvested 48 hours after transfection (for consistency, as particle infectivity decreased over time post-transfection, S1A Fig.) correlated with the particle infectivity of virus stocks generated from PBMC 8 days following infection (S1B Fig., $p < 0.0001$, $r = 0.9455$). Analysis of the particle infectivity of virus stocks produced from all of the infectious molecular clones showed that the particle infectivities of all viruses tested ranged from $7x10^{-5}$ to $1x10^{-2}$, and that there was also a wide range of particle infectivities within each transmitting partner's quasispecies (Fig. 3). In

pair 3678, the TF variant was the most infectious virus compared to the rest of the transmitting partner's variants, while the TF from pair 3576 was the least infectious (Fig. 3). TF variants spanned the thousand-fold range of particle infectivities measured for all the viruses tested, as can be seen by the TF from pairs 3618 and 4473, which are found on extreme ends of the particle infectivity spectrum. Across all six transmission pairs, we observed no significant selection for infectivity when comparing the TF to the median of the transmitting partner's quasispecies (Fig. 3; $p = 0.6875$). In these subtype C transmission pairs particle infectivity did not constitute a dominant determinant of transmission fitness.

Sensitivity of TF and NT variants to neutralization by antibodies present in the transmitting partner

We previously reported that Env glycoproteins derived from early viruses in acutely infected linked recipients were on average more sensitive to neutralization by plasma from the transmitting partner, compared to autologous Envs directly derived from the transmitting partner (1). Antibody neutralization of SGA-derived genome length TF and NT variants, derived from the first month of infection, from heterosexual epidemiologically-linked transmission pairs, has not been examined to date. Using a previously described TZM-bl neutralization assay (1,28,29), we evaluated neutralization of full-length TF and autologous NT IMC by plasma from the transmitting partner near the time of transmission. Donor plasma (diluted 1:100) demonstrated relatively weak neutralization against the majority of viruses tested in each panel, with a median of 18% neutralization. The highest level of neutralization was seen in pair 4473 against

the TF (51%) (Fig. 4A). Overall, TF variants were more efficiently neutralized compared to the medians of the transmitting partner's NT variants (Fig. 4A; p = 0.031). Additionally, greater neutralization negatively correlated with distance to the amino acid subtype C consensus (Fig. 4B; p = 0.011, r= -0.4995), suggesting a link between these two measurements. Consistent with our previous findings, transmission did not select for TF variants with greater neutralizing antibody resistance to donor plasma.

In vitro replication of TF and NT viruses

Selection for consensus-like TF variants in these six transmission pairs may indicate a selection for viruses with greater *in vivo* fitness, as hypothesized from a study of 137 linked transmission pairs (7). To determine whether this translated into a similar fitness advantage in activated CD4 T cells, we measured the replicative capacity (RC) of viruses *in vitro*. TF and NT IMC were tested for *in vitro* replication by infection of stimulated peripheral blood mononuclear cells (PBMC), at equal multiplicities of infection (MOI). Since the number of infectious particles to total particles varied greatly between all virus stocks tested, we based the amount of virus used for each replication experiment on a consistent MOI (0.01), rather than equal amounts of virus particles, in order to normalize for initial infectivity. We measured virus growth by reverse transcriptase activity of cell culture supernatants every 48 hours for ten days (Fig. 5A). RC scores were generated for each virus based on the area under the curve of virus growth, as described in the methods. TF viruses exhibited a wide range of RC among all the viruses tested, and the relative RC of TF as compared to NT viruses from the

same donor also varied substantially (Fig. 5B). For instance, the TF from pair 3576 had the lowest RC when compared to the transmitting partner's quasispecies. Alternatively, pairs 3618 and 3678 had TF viruses with relatively high RC, although they were not the highest replicators from their transmitting partner's quasispecies (Fig. 5B). In total, we saw no significant selection for TF viruses having higher *in vitro* RC than the median RC of the NT viruses tested (Fig. 5C; $p = 0.219$). Similar to particle infectivity, which correlated with RC over all the viruses tested (S2 Fig.; $p = 0.0005$, $r = 0.5712$), there was no evidence for a distinct replicative capacity profile associated with transmission. In addition, viruses closer to consensus typically had lower *in vitro* replicative capacities, since the pairwise distance to subtype C consensus correlated with *in vitro* replicative capacity (Fig. 5D; $p = 0.0158$, $r = 0.4168$).

Since dendritic cells have also been implicated as an initial target cell for establishment of HIV-1 infection in the genital mucosa (30), we examined the ability of the 6 TF and a limited set of 6 NT variants with similar *in vitro* RC scores, to productively infect and replicate in immature monocyte derived dendritic cells (MDDC) *in vitro*. We cultured MDDC by isolation and differentiation of blood-derived CD14+ monocytes from healthy donors and infected them with virus at a high MOI of 1. We assayed virus production by measuring the reverse transcriptase activity present in cell culture supernatants every 48 hours for twelve days. We found that the TF and NT variants studied did not significantly differ in their ability to replicate in MDDC (S3 Fig.; $p = 0.87$). Of the twelve TF & NT variants, six had detectable replication in MDDC (3 TF & 3 NT), suggesting that productive infection of MDDC is limited, even at a high

MOI, and is not a requirement for transmission. Overall, these data suggest that HIV-1 transmission is permissive to TF variants with a wide range of *in vitro* replicative capacities relative to the transmitting partner's quasispecies.

Interferon-α resistance in HIV-1 subtype C transmission

By conducting *in vitro* replication assays in cells pre-treated with exogenous interferon-α (IFN-α), previous studies found that subtype B and subtype C TF variants were relatively resistant to IFN-α compared to a panel of chronic viruses (13) or later variants from the same individual (17). These studies suggested a selection during the HIV-1 transmission bottleneck for variants adept at escaping innate immunity, specifically the antiviral effects of IFN-α. However, these studies were not done in epidemiologically-linked transmission pairs, and thus were unable to directly compare TF viruses to related NT variants in the donor quasispecies near the time of transmission. To test whether the subtype C TF viruses investigated here exhibited relative resistance to IFN-α, as compared to NT variants derived from the transmitting partner's quasispecies, we assayed *in vitro* virus replication in PBMC in the presence and absence of IFN-α.

We assayed viral replication in activated CD8-depleted PBMC in the presence and absence of 5,000 U/ml of IFN-α, which was added 24 hours prior to infection in order to maximally inhibit viral replication, as described previously (17). Supernatant HIV-1 p24 antigen levels were measured every 48 hours for 10 days to assess the kinetics of viral replication. In the initial 21 variants tested, growth of virus in the presence of IFN-α was tightly correlated with *in vitro* RC scores in the absence of IFN-α (Fig. 6A; $p < 0.0001$, $r = 0.8844$), suggesting that

in vitro growth in the presence of IFN-α was largely determined by viral replicative capacity.

In light of this, we attempted to delineate subtle differences in IFN-α resistance by performing further experiments with selected NT variants that exhibited relatively similar replication kinetics to the TF in each pair (to minimize the impact of replication differences). The replication of these selected viruses was assessed in activated PBMC in the presence and absence of 1,000 U/ml of IFN-α (added 24 hours prior to infection), monitoring virus replication by reverse transcriptase activity in the supernatant. An example of such an assay for transmission pair 331 is shown in Fig. 6B. When compared to the tested NT variants, TF viruses did not differ significantly in resistance to IFN-α (assessed as the ratio of the RC score in the presence and absence of IFN-α) (Fig. 6C, $p = 0.219$). In pair 331 and 4473, the TF appeared to be more IFN resistant than the NT viruses from the same donor (Fig. 6C). In pair 3618 the TF was near the median of the NT variants, while in three pairs (3576, 3678, and 4248), the TF was the most sensitive to IFN-α. Overall, the IFN-α resistance of the TF viruses did not differ significantly from the median of the NT variants (Fig. 6C).

Because the TF viruses were not found to be more IFN-resistant than donor NT viruses, we validated the method used for analysis of IFN resistance with 3 subtype B TF and 6-month consensus virus pairs that had previously been demonstrated to differ in their IFN-α resistance (17). As shown in Figure 6C, the 3 TF viruses were each confirmed to be more IFN-resistant than the matched 6-month virus from the same subject, verifying the ability of the methods used here to detect previously documented differences in viral IFN resistance. In the six

subtype C epidemiologically-linked transmission pairs studied we also observed that IFN-a resistance correlated with the virus' ability to replicate (S4A Fig.). Although the RC and IFN-α resistance of the six subtype B TF and 6-month viruses was not statistically correlated, these subtype B TF viruses did have higher RC scores than their matched 6-month variants (S5 Fig.). Overall, these data suggest that a component of IFN-α resistance is the ability of TF and NT HIV-1 variants to replicate.

To confirm that this finding was independent of the amount of IFN used to inhibit viral growth, we measured replication at day 7 for four viruses with a representative range of RC scores using a range of IFN-α concentrations (0.5 U/ml - 10,000 U/ml). The relative sensitivity of these viruses was consistent across the range of IFN-α concentrations tested (S4B Fig.). Additionally, we tested a limited subset of TF and NT variants for their ability to induce IFN-a, which may have influenced IFN-a resistance measurements, and found that IFN-a levels above background were not detectable at day 8 in either PBMC or MDDC infected cultures (S4C Fig.). Hence HIV-1 transmission from these donors was not mediated by TF viruses that exhibited higher levels of interferon resistance than NT viruses, indicating that heterosexual HIV-1 transmission is permissive to viruses anywhere within the range of *in vitro* interferon resistance profiles observed in the donors studied here, and factors other than IFN-a resistance constituted the dominant determinants of transmission fitness in these pairs.

Discussion

The rapid within-host diversification of HIV-1 observed during chronic infection, which represents a primary obstacle to effective HIV prevention strategies, contrasts starkly with the viral homogeneity evident following transmission. The stringent genetic bottleneck is most pronounced in heterosexual transmission, where a vast majority of new infections are established by single viral variants. Correlates of transmission may become evident by studying the properties of these transmitted/founder (TF) variants, which, in turn, could help inform effective HIV-1 vaccine design. Studies of early and transmitted variants have found genetic and phenotypic signatures associated with transmission; however, none have examined full-length TF variants and corresponding non-transmitted (NT) variants present near the time of transmission from epidemiologically-linked transmission pairs.

In this study, we applied new molecular techniques to investigate the requirements of HIV-1 transmission in six subtype C transmission pairs. We amplified, sequenced, and generated infectious molecular clones (IMC) of matched full-length TF viruses very early after infection (Fiebig stage II) and near full-length NT variants from 22-45 days following the estimated date of transmission. Technical limitations associated with amplifying full-length virus from genital tract samples required us to amplify from patient plasma. Despite this limitation, we previously showed in eight epidemiologically-linked transmission pairs that the TF was most highly related to NT variants that were absent from the predominant genital tract subpopulations, and were found in both blood and genital tract of the donor partner (5). Consistent with this, in pair

331 we observed a NT variant in the plasma of the transmitting partner with only three amino acid differences from the TF across the entire proteome (Fig. 1B).

In this study, we generated IMCs from the diverse donor quasispecies with great sequence accuracy and selected variants in an unbiased fashion. Five of the six pairs were female-to-male, the route by which the most stringent bottleneck occurs (7,31). Since high donor viral load and the presence of genital ulcers and inflammation (GUI) in the recipient can, to a certain degree, mitigate selection bias in the bottleneck, it is important to note that these six pairs include three donors with viral loads >100,000 RNA copies/ml, as well as one recipient with a reported GUI in the twelve months prior to seroconversion (3,7). Despite these caveats, single variant transmission was observed in all six pairs.

Consistent with our previous findings (7), we observed selection during transmission for variants with more consensus-like amino acid and nucleotide DNA sequences from the available quasispecies present in the donor at the time of transmission, across the full viral proteome and genome, respectively (Fig. 2). It has been shown that HIV-1 within-host diversity during chronic infection is greater than between-host diversity, suggesting conservation of certain genetic elements during transmission (32). In conjunction, studies of subtype A and D heterosexual transmission pairs demonstrated transmission of more ancestral viral variants, by measuring distances of each variant to their most recent common ancestor (MRCA) on a phylogenetic tree of Env sequences (8,22). In the current study, the LANL subtype C consensus node falls near the subtype C MRCA highlighting the equivalence of these two measurements (Fig. 1). Thus, HIV-1 transmission consistently selects for variants that more closely resemble

ancestral and consensus-like viruses, indicating that evolution in the host decreases transmission potential.

The viral diversification observed during chronic infection due to adaptive immune pressure targeted specifically against HIV-1 is likely driving viral evolution away from consensus (1,7,33,34). We have previously shown that acquisition of resistance to antibody neutralization comes with a transmission fitness cost (1). We similarly found that TF viruses were more sensitive to neutralization by donor plasma acquired near the time of transmission when compared to the corresponding NT variants. It should be noted, however, that a limitation to this finding is that NT variants were cloned and tested with plasma from approximately four weeks after the estimated date of infection, although the two pairs with the largest time gap between transmission and sampling did not show the greatest neutralization of the TF. These data also reaffirm that TF variants are generally not resistant to antibody neutralization (1,35). As expected, donor plasma tested against contemporaneous viruses (TF or NT) demonstrated limited neutralization capacity. Moreover, neutralization sensitivity correlated with the distance to consensus over all the viruses tested. Considering these observations, it is reasonable to propose that selection of antibody sensitive variants during transmission is a side effect of the transmission cost associated with non-consensus adaptations in general, and not an underlying mechanism of transmission itself.

In order to address the role of viral fitness in transmission, we measured the *in vitro* fitness of a subset of viruses from six transmission pairs. Although a previous study found that TF viruses were more infectious than chronic control

viruses (13), we found no bias towards increased infectivity when comparing the TF to the corresponding NT variants. Particle infectivity in TZM-bl cells correlated with replicative capacity in PBMC, suggesting that entry into TZM-bl cells is representative of a component of viral replication in primary cells. Viral replicative capacity in activated PBMC, a fundamental measure of *in vitro* fitness, was also not higher for TF variants in comparison to the corresponding NT variants, and none of the TF variants exhibited the highest replicative capacity from among the tested NT variants. We found that more consensus-like variants, which are typically those that transmit, had lower *in vitro* replicative capacities over all the variants tested, indicating that higher *in vitro* replicative capacity is not linked to transmission. TF variants were also not observed to have enhanced replicative capacity in monocyte-derived dendritic cells, an *in vitro* model for dendritic cells, which may act as an initial target cell for establishment of HIV-1 infection. These findings argue against the original hypothesis that consensus-like variants would have higher *in vitro* replicative capacities. Thus, *in vitro* RC in activated PBMC or MDDC may not reflect *in vivo* transmission fitness, potentially because replication in stimulated PBMC may recapitulate the inflammatory environment that occurs some time after transmission and during chronic infection rather than conditions initially encountered at initial sites of virus replication. We cannot rule out the possibility that replication assays in cell types more representative of mucosal transmission, such as tissue resident CD4+ T cells or Langerhans cells, may yield different results. However, consistent with our observations, previous studies found a significant negative correlation between similarity to consensus and *in vitro* RC in a larger number of patients in

differing cohorts using gag-chimeras (36,37). Transmission of low *in vitro* fitness variants may seem counterintuitive; however, full-length TF IMC as well as over 200 transmitted Gag chimeras have been shown to exhibit a wide range of *in vitro* replicative capacities (36,38,39), as we found for our six TF viruses. A recent theoretical model of HIV transmission predicted that variants with lower replicative capacity via increased latency would exhibit a greater transmission potential *in vivo* (40), and it is therefore possible that modestly lower *in vitro* replicative capacity is an advantage during transmission.

A potential selection factor during mucosal transmission is the early innate immune response to HIV-1. Innate antiviral cytokines including IFN-a are induced at initial sites of HIV-1 replication in the mucosa and draining lymph nodes (41,42), hence HIV-1 variants that are more resistant to the antiviral effects of IFN-α may have an advantage during transmission. Indeed, cross-species transmission of Simian Immunodeficiency Virus (SIV) to humans required escape from the interferon stimulated APOBEC3 restriction factors by enhanced Vif antagonism (43). A recent in-depth study using the rhesus macaque model also found that IFN-α treatment prior to intrarectal SIV_{MAC251} inoculation reduced the number of transmitted variants and increased the number of challenges necessary to initiate infection (44). Consistent with the hypothesis that type 1 IFNs contribute to the transmission bottleneck, previous studies using HIV-1 found that TF variants are generally more resistant to IFN-α *in vitro* than viruses present during early chronic infection (13,17). Fenton-May et. al. (17) found that TF viruses from both subtype B and C infected subjects were more resistant to IFN-α when compared to matched variants generated from the same

individual six months post-infection or during early chronic infection. Parrish et. al. (13) found that TF viruses are more resistant to IFN-α than viruses from unmatched chronic controls, though this was true only for the subtype B and not for the subtype C variants they studied. In six subtype C transmission pairs studied here we did not observe that TF viruses exhibited enhanced resistance to IFN-α compared to NT viruses. TF variants did not replicate to higher levels in the presence of IFN-α, nor did they have higher ratios of replication in the presence versus the absence of IFN-α.

These differing results could be due to differences in experimental protocols, as well as difficulties in separating inherent replicative capacity from interferon resistance. We therefore tested the IFN-α resistance of previously studied TF and 6-month viruses and confirmed that these TF variants were more resistant to the effects of IFN-α, consistent with previous observations. In addition, we found that the TF variants had higher replicative capacities than the 6-month consensus variants, although for this group of viruses IFN-a resistance did not directly correlate with viral replicative capacity. The influence of viral replicative capacity on measures of interferon resistance is not fully understood. The impact of multiplicity of infection on measured interferon resistance has been noted previously (45), so in the current studies we utilized a low multiplicity to ensure adequate target cell availability even for the higher replicating viruses. We chose a MOI of 0.01 for our assays since it represented an input virus dose at which we were able to measure both replication differences between viruses, as well as IFN-α resistance differences (S4D Fig.). For the viruses tested from the six Zambian transmission pairs, we found that *in vitro* replication in the presence of

interferon correlated with replication in the absence of IFN-α, such that the rank order of virus replicative capacities from lowest to highest was similar in the presence or absence of interferon. Even when a subset of viruses with more closely-matched levels of replication were studied so that we were better able to observe IFN-α resistance differences, we found that TF variants were not IFN-α resistant compared to the matched NT variants.

The lack of difference in the IFN resistance of TF and NT viruses in these transmission pairs may be due to the length of time for which the chronically infected viral donors had been infected prior to viral transmission to their partners and derivation of the viruses studied. Fenton-May et. al. showed that while IFN-α resistance decreased over the first 6 months following infection, it subsequently increased in different subjects at timepoints from 2-7 years post-infection (17). Edlin et. al. and Kunzi et. al. further showed that viruses isolated from individuals who had progressed to AIDS were more IFN resistant than viruses from asymptomatic chronically-infected individuals (46,47). Likewise, Parrish et. al. proposed that their observation of differences in IFN resistance between TF IMCs and IMCs from unmatched chronically-infected subjects in a subtype B-infected cohort, but not in a subtype C-infected cohort, may have been due to the subtype C-infected donors being sampled at later timepoints in chronic infection (13). However, it should be noted that we did observe a range of interferon sensitivities across the six transmission pairs, with greater than a 100-fold difference being observed between TF viruses. In future, it would be of interest to determine whether chronically-infected donors in the Zambian discordant couples cohort who failed to transmit infection to their partners

harbor more IFN-sensitive viruses than those present in the virus-transmitting donors studied here. However on the basis of the current results it seems likely that IFN-a does not make a major contribution to the HIV-1 transmission bottleneck, or may do so only in some transmission scenarios.

Transmission selection for consensus-like and more neutralization-sensitive TF variants suggests that within-host evolution of HIV-1 in response to human adaptive immune responses may cause a loss of fitness required for the establishment of infection in a naive host following transmission. We show that relatively high *in vitro* replicative capacity and preferential IFN-α resistance were not selected for during transmission of subtype C HIV-1 in the six pairs studied here. Thus, the *in vitro* assays of HIV-1 replication employed here may not be measuring some of the key determinants of transmission fitness, and other models of HIV transmission, such as low dose intravaginal challenges of humanized mice, or human genital explant cultures, may be needed to determine the phenotypic requirements of HIV-1 transmission that genetic differences are pointing to.

Materials and Methods

Study subjects

The six HIV-1 subtype C transmission pairs investigated in this study were enrolled in the heterosexual discordant couple cohort at the Zambia-Emory HIV Research Project (ZEHRP) in Lusaka, Zambia. Human subjects protocols were approved by both the University of Zambia Research Ethics Committee and the Emory University Institutional Review Board. HIV-1 serodiscordant couples in this cohort were provided counseling and testing on a monthly basis prior to the negative partner becoming HIV-1 positive. The recipients were enrolled in the International AIDS Vaccine Initiative (IAVI) Protocol C early-infection cohort. Epidemiological linkage was defined by phylogenetic analyses of HIV-1 *gp41* sequences from both partners (48). All individuals in this study were ART naive during the time of sampling.

Viral RNA extraction and PCR amplification

Viral RNA extraction and near full-length genome single genome amplification were performed as described in Deymier et al. 2014 (26). Briefly, viral RNA was extracted from 140µl of plasma using the QIAamp Viral RNA mini kit (Qiagen) and was used for cDNA synthesis carried out with Superscript III (Life Technologies) and an anchored Oligo$(dT)_{18}$ primer. The cDNA was used immediately for PCR amplification. Near full-length single genome PCR amplification was performed by serially diluting cDNA, followed by two rounds of PCR amplification, so that ~30% of wells became positive. Both rounds of PCR

were performed in 1x Q5 Reaction Buffer, 1x Q5 High GC Enhancer, 0.35 mM of each dNTP, 0.5 μM of primers and 0.02 U/μl of Q5 Hot Star High-Fidelity DNA Polymerase (NEB) in a total reaction volume of 25 μl. First round primers were, 1U5Cc and 1.3'3'PlCb, and second round primers were 2U5Cd and 2.3'3'plCb (49). Cycling conditions for both reactions are 98°C for 30s, followed by 30 cycles of 98°C for 10s, 72°C for 7.5min, with a final extension at 72°C for 10min. PCR reactions were run on a 1% agarose lithium acetate gel at 300 V for 25 min in order to determine the presence of a 9 kb band.

Sequencing

Positive ~9kb single genome amplicons were gel-extracted using the Wizard SV Gel and PCR Clean-Up System (Promega). Purified ~9 kb PCR amplicons were sent for sequencing to the University of Alabama Birmingham (UAB) sequencing core for Sanger sequencing.

In conjunction, multiple amplicons from recipient 3576 were sequenced by single-molecule nucleic acid sequencing (Pacific Biosciences), to confirm the TF. Briefly, SMRTbell libraries were constructed according to the manufacturer's instructions for 10kb amplicons. PCR reactions of DNA amplicons were purified using Wizard SV Gel and PCR Clean-Up System (Promega) and mixed at equal concentrations to a total of 3ug DNA. Library preparation quality was assessed on a Bioanalyzer and SMRT sequencing on the PacBio RSII was performed following primer annealing and P4 polymerase binding to the library preparations. The consensus of the reads, aligned to the HXB2 reference sequence, were then taken to form a TF sequence, which matched the Sanger sequence.

Sequence Analysis

All 9kb viral sequences were aligned in Geneious bioinformatics software (Biomatters, Aukland, NZ) using MUSCLE (50), followed by hand aligning. The Los Alamos National Database HIV Consensus/Ancestral Sequence Alignments were used as reference sequences (http://www.hiv.lanl.gov/content/sequence/NEWALIGN/align.html). Phylogenetic trees were generated using the DIVEIN web server (http://www.indra.mullins.microbiol.washington.edu/DIVEIN/) (51). Phylogenetic analyses were performed by maximum likelihood parsimony under Phylogeny/Divergence/Diversity. For nucleotide sequence analysis a general time reversible model was used, with a fixed gamma distribution parameter of 1, and performed with 100 bootstraps. Amino acid phylogenetic analysis was performed using the HIVw model of evolution, with 100 bootstraps (52). Pairwise distances from each branch node to the subtype C consensus node were extracted from the distance matrices of the phylogenetic trees.

Generation of full-length IMC

HIV full-length genome infectious molecular clones were generated as described in Deymier et al. 2014 (26). Briefly, linked recipient specific primers were generated in order to amplify the full long terminal repeat (LTR) from the linked recipient white cell pellet DNA. This LTR was cloned into a pBluescript vector, and the TF sequence of the LTR sequence was inferred as the consensus sequence from multiple clones. Subsequently, a three-piece DNA HD In-Fusion

HD cloning (Clontech) ligation reaction using a reamplified clonable near-full length amplicon and two LTR pieces generated by PCR from the linked recipient LTR generated the full-length IMC. TF IMC were correct for the entire genome, whereas NT variants were chimeric for only for the R region of both 5' and 3' LTR, which was taken from the TF of that transmission pair. IMC were sequenced in order to confirm a match to the sequence of the single genome amplicon from which it was derived.

Generation of virus stocks and particle infectivity

293T (American Type Culture Collection) cells were transfected with 1.5µg of plasmid DNA, using the Fugene HD transfection reagent (Roche) according the manufacturer's protocol. Viral stocks were collected 48 hours post transfection and clarified by centrifugation. These virus stocks were then titered for infectivity on TZM-bl cell, as described previously (53). The virus stocks were also measured for reverse transcriptase (RT) activity using a radiolabeled reverse transcriptase assay (53). Particle infectivity of each virus was determined as the ratio of titer (infectious units/µl) over RT signal (RT/µl) for 3 independent experiments. Particle infectivity over time was measured by sampling 8ul (0.4%) per time point over a 3 day period.

Replication in PBMC and interferon resistance

Frozen peripheral blood mononuclear cells (PBMC) from buffy coats were thawed and stimulated with 20 U/ml of interleukin-2 (IL-2) and 3ug/ml of phytohemagglutinin (PHA) in R10 (Roswell Park Memorial Institute (RPMI)

1640 Medium supplemented with 10% defined fetal bovine serum (FBS), 1 U/ml penicillin, 1ug/ml streptomycin, 300ug/ml L-glutamine) for 72 hours at 37C. After 48 hours, 1,000 IU/ml of interferon-α2a (Sigma Aldrich, Product # SRP4594) was added to a portion of cells 24 hours prior to infection. $1x10^6$ cells were then infected in 15ml conical tubes by 2 hour spinoculation at 2,200 rpm with an MOI of 0.01 based on the TZM-bl titer in triplicate. Cells were then washed twice in 13ml RPMI, resuspended in 500ul of R10 media and plated in a 48 well plate in triplicate. 50ul of supernatant was then sampled every 48 hours starting with a day zero time point taken 2-3 hours after plating to get a baseline reverse transcriptase activity for each infection well using the radiolabeled reverse transcriptase assay.

Where noted in the text, an alternative strategy for another independent experiment with CD8-depleted PBMC was used with a few differences: anti-CD3 (R&D Systems clone UCHT1; 50ng/ml working concentration) and anti-CD28 (eBioscience clone CD28.2; 100ng/ml working concentration) antibodies were used to stimulate MACS microbead (from Miltenyi plus the MACS LD columns) CD8-depleted PBMC from three separate donors in a mixed lymphocyte reaction and then infected at an MOI of 0.1 based on TZM-bl titer. $2x10^5$ cells were then infected in the presence and absence of 5,000 IU/ml of interferon-α2a (Peprotech) and cells were washed three times with 10ml of RPMI and supernatant tested by a modified ELISA assay using the AlphaLISA HIV p24 (high sensitivity) kit (Product # AL291C PerkinElmer) per protocol instructions, using the same media for the standard as in the sample and loading 5ul per well. The replication score (RC score) for each variant was calculated using a

normalized area under the curve. The median of the replicates were background subtracted using the day 2 time point, adjusted for sampling by a measured exponential decay correction, and area under the curves (AUC) were divided by the AUC for a standard lab adapted subtype C virus, MJ4, to compare across transmission pairs analyzed on different days. Interferon-α2a resistance was measured in a similar fashion, followed by calculating the ratio of the RC score in the presence of interferon divided by the RC score in the absence of interferon.

Monocyte derived dendritic cell infections

Monocyte derived dendritic cells (MDDC) were isolated from two healthy blood donors by CD14 positive bead isolation (Miltenyi Biotec), followed by culture at 37°C in R10, supplemented with 40 ng/ml IL-4 (Peprotech) and 20ng/ml GM-CSF (Peprotech) for 7 days. MDDC differentiation was confirmed by flow cytometry using the following antibodies and stains: α-CD14 PB (clone M5E2), α-CD11c APC (clone S-HCL-3), α-HLA-DR V500 (clone G46-6) (BD Biosciences), and the LIVE/DEAD Fixable Near-IR Dead Cell Stain Kit (Life Technologies). The phenotype of MDDC after 7 days was CD14 low, CD11c high, and HLA-DR high, as expected. Cells were harvested, and $3x10^5$ MDDC were seeded in a flat-bottom 96-well plate. MDDC were infected in a volume of 250µl of R10 with virus at an MOI of 1 for 4 hours. Cells were then washed three times with RPMI, and cultured for 12 days in R10 supplemented with 40 ng/ml IL-4 and 20ng/ml GM-CSF. 50µl of culture supernatant was collected every two days and replaced with fresh media. The supernatant was then analyzed for virus production by the radiolabelled RT assay(53).

Interferon Elisa

IFN-a levels were measured by the VeriKine™ Human IFN Alpha ELISA Kit from supernatants 8 days after PBMC and MDDC infections with a subset of viruses from pairs 331 and 3678. The negative controls included media from PBMC and MDDC uninfected cultures. The positive controls included IFN-a spiked media equal to the initial amount of IFN-a utilized in these infections, along with supernatant from an infection carried out in the presence of IFN-a.

Neutralization assay

IMC derived virus and plasma taken from the same time point in the transmitting partner (donor) along with the TF from the recipient, were used to test antibody neutralization of variants circulating near the time of transmission. The TZM-bl neutralization assay was adapted for use with IMCs, in a similar fashion to what has been published previously for IMC (28) and pseudoviruses (1,29). Briefly, heat inactivated plasma was serially diluted 5-fold starting at 1:100, and each dilution was then mixed with 20 IU/ul of virus at a 1:1 ratio. After incubation at 37^{o}C for 1 hour, the plasma and virus mixtures were used to infect previously seeded TZM-bl cells (24 hours prior to infection at 6x10^{3} cells per well in a 96-well plate). After a 40 hour incubation, the Promega Reporter Buffer was used to lyse cells according to manufacturer instructions and, following two freeze-thaw cycles, luciferase was measured with the Luciferase Assay System from Promega (Catalog # E1501) in the supernatants on a luminometer using the Gen5 2.00 software. Maximal percent inhibition (compared to the no plasma

control) was calculated at a dilution of 1:00 after background subtraction and removal of variants with a signal less than three times background for cell only control wells. The data is averaged from each virus run in duplicate from two independent experiments.

Acknowledgements

We acknowledge all participating volunteers in Zambia as well as the staff at the Zambia Emory HIV Research Project who made this study possible. We thank Beatrice Hahn for providing subtype B TF & 6-month consensus plasmids. We thank Jon Allen, Sheng Luo, Katja Pfafferott, Katie Kilgore, Ling Yue and Paul Farmer for technical assistance and sample management. We thank Daniel T. Claiborne for experimental advice, Debby Basu, Jessica Prince, Evonne Woodson, and Kelsie Brooks for manuscript review and Tianwei Yu for expert statistical advice.

Table 1 – Transmission pair characteristics

Coded ID[a]	Partner Status[b]	EDI[c]	Sample Date[d]	Days after EDI	VL[e]	VL Sample Date	# Sequences Analyzed
Z331M	LR	17-Mar-09	18-Apr-09	32	2,864,668	18-Apr-09	9
Z331F	D		15-Apr-09	29	2,620	2-May-09	20
Z3576F	LR	6-Mar-09	28-Mar-09	22	6,460,200	28-Mar-09	18
Z3576M	D		18-Apr-09	43	54,100	23-Apr-09	15
Z3618M	LR	27-May-09	18-Jun-09	22	16,600,000	18-Jun-09	8
Z3618F	D		11-Jul-09	45	20,400	15-Jul-09	17
Z3678M	LR	26-Aug-09	17-Sep-09	22	3,017,616	30-Sep-09	8
Z3678F	D		23-Sep-09	28	269,240	23-Sep-09	18
Z4248M	LR	13-May-10	4-Jun-10	22	24,993,584	4-Jun-10	6
Z4248F	D		11-Jun-10	29	119,320	11-Jun-10	21
Z4473M	LR	16-May-11	7-Jun-11	22	16,891,328	7-Jun-11	9
Z4473F	D		7-Jun-11	22	104,427	18-Jun-11	18

[a] Z=Zambia, M=male, F=female
[b] LR – epidemiologically linked recipient partner, D – donor partner.
[c] EDI – estimated date of infection (22 days prior to Ag+Ab- sample) (3)
[d] Sample Date – date of collection of plasma from which viral RNA was extracted
[e] VL – Viral Load (RNA copies/mL)

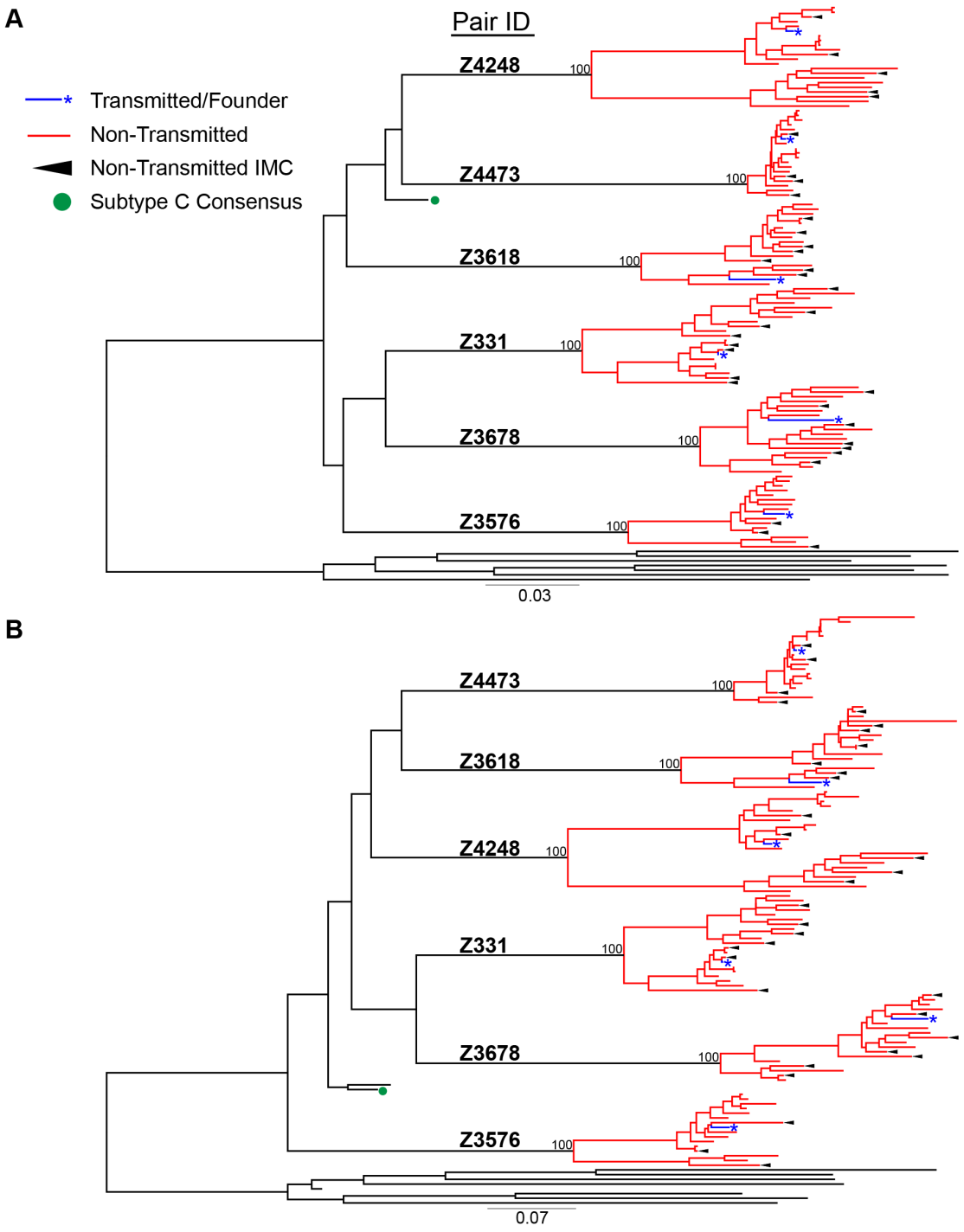

A
Pair ID
Transmitted/Founder
Non-Transmitted
Non-Transmitted IMC
Subtype C Consensus
Z4248
Z4473
Z3618
Z331
Z3678
Z3576
100
0.03
B
Z4473
Z3618
Z4248
Z331
Z3678
Z3576
100
0.07

Figure 1 - HIV-1 Full-length genome phylogenetic analysis of six epidemiologically-linked heterosexual transmission pairs.

(A) Nucleotide sequences for all 115 single genomes amplified from six linked transmission pairs were aligned to the curated LANL consensus/ancestral alignment and a maximum likelihood tree was generated. (B) Single genome nucleotide sequences for each viral gene (*gag, pol, vif, vpr, vpu, tat, rev, env,* & *nef*) were translated to their amino acids and then concatenated. These were aligned with LANL consensus/ancestral concatenated protein sequences and a maximum likelihood tree was generated. Transmitted/founder sequences from linked recipients are in blue, donor non-transmitted variants are in red, LANL database curated consensus/ancestral sequences are shown in black, where the LANL subtype C consensus is indicated by a green circle. Black arrows indicate virus variants from the donor quasispecies that were selected for generation of non-transmitted (NT) infectious molecular clones.

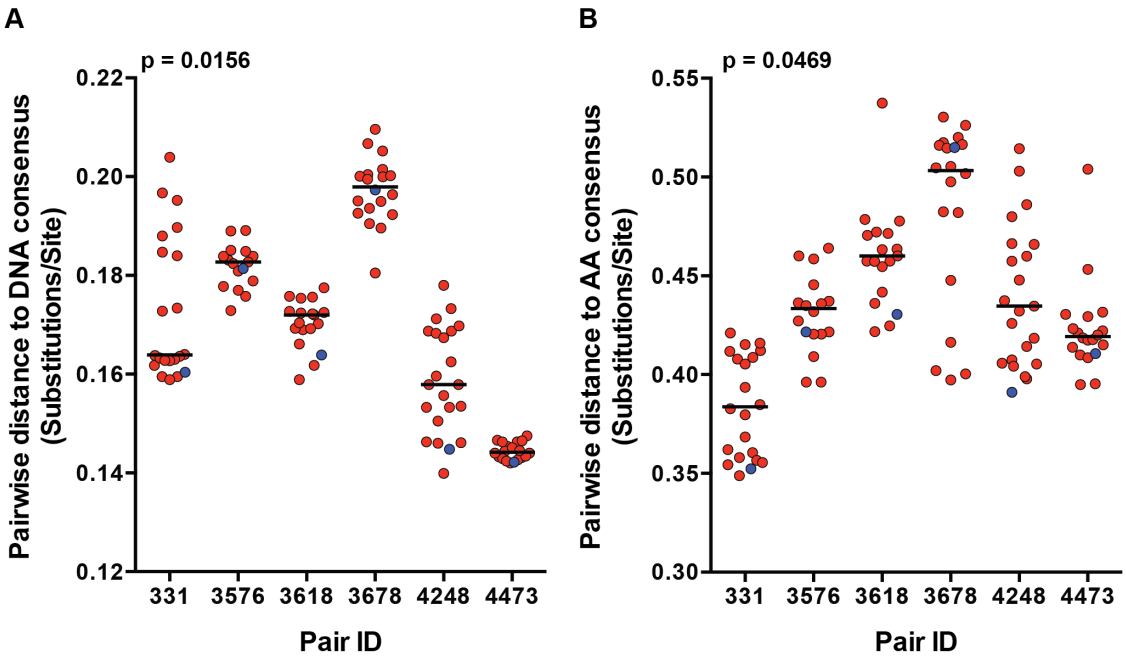


Figure 2 - Transmission selects for more consensus-like TF variants.

The pairwise distance of each viral variant on the (A) nucleotide and (B) amino acid phylogenetic trees to the LANL subtype C consensus node were measured and plotted for each transmission pair. Transmitted/founder variants are in blue, and non-transmitted variants are in red. The median of the non-transmitted variants is designated with a black line. The statistical significance of the difference between TF and NT donor median values was analyzed using a one-tailed Wilcoxon matched-pairs signed rank test.

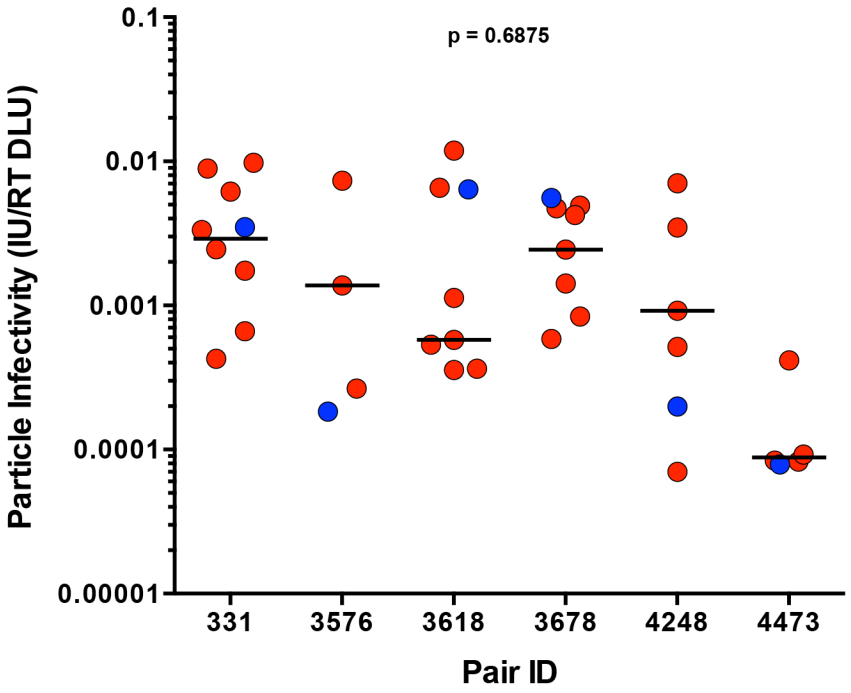


Figure 3 - Particle infectivity of TF and NT infectious molecular clones.

293T cells were transfected with TF (blue) and NT (red) infectious molecular clones. Supernatants collected 48 hours post-transfection were titered on TZM-bl cells to define the number infectious units per microliter (IU/ul), while reverse transcriptase activity was measured simultaneously for total viral particles per microliter (RT DLU, reverse transcriptase digital light units). The particle infectivity (IU/RT DLU) of each infectious molecular clone is plotted for each transmission pair. The median of the NT variants is designated with a black line. The statistical significance of the difference between TF and NT donor median values was determined using a two-tailed Wilcoxon matched-pairs signed rank test ($p = 0.6875$).

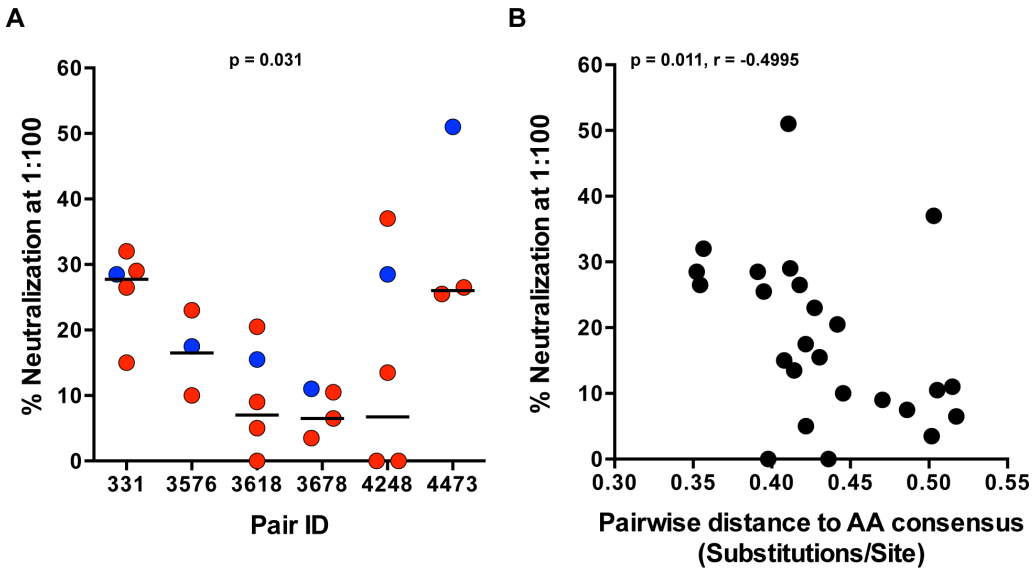


Figure 4 - TF variants are more sensitive to neutralization by donor plasma than NT viruses

(A) Neutralization of TF (blue) and NT (red) IMC by donor plasma was measured for each pair in a TZM-bl neutralization assay. Percent neutralization by donor plasma (diluted 1:100) is depicted on the y-axis, and representative TF and NT viruses tested for each transmission pair are depicted on the x-axis. The median of the NT variants is designated with a black line. The statistical significance of the difference between TF and NT donor median values was determined using a two-tailed Wilcoxon matched-pairs signed rank test ($p = 0.031$). (B) Spearman correlation of the pairwise distance to the amino acid subtype C consensus and donor plasma neutralization described in part A over all the variants tested from 6 transmission pairs ($p = 0.011$, $r = -0.4995$).

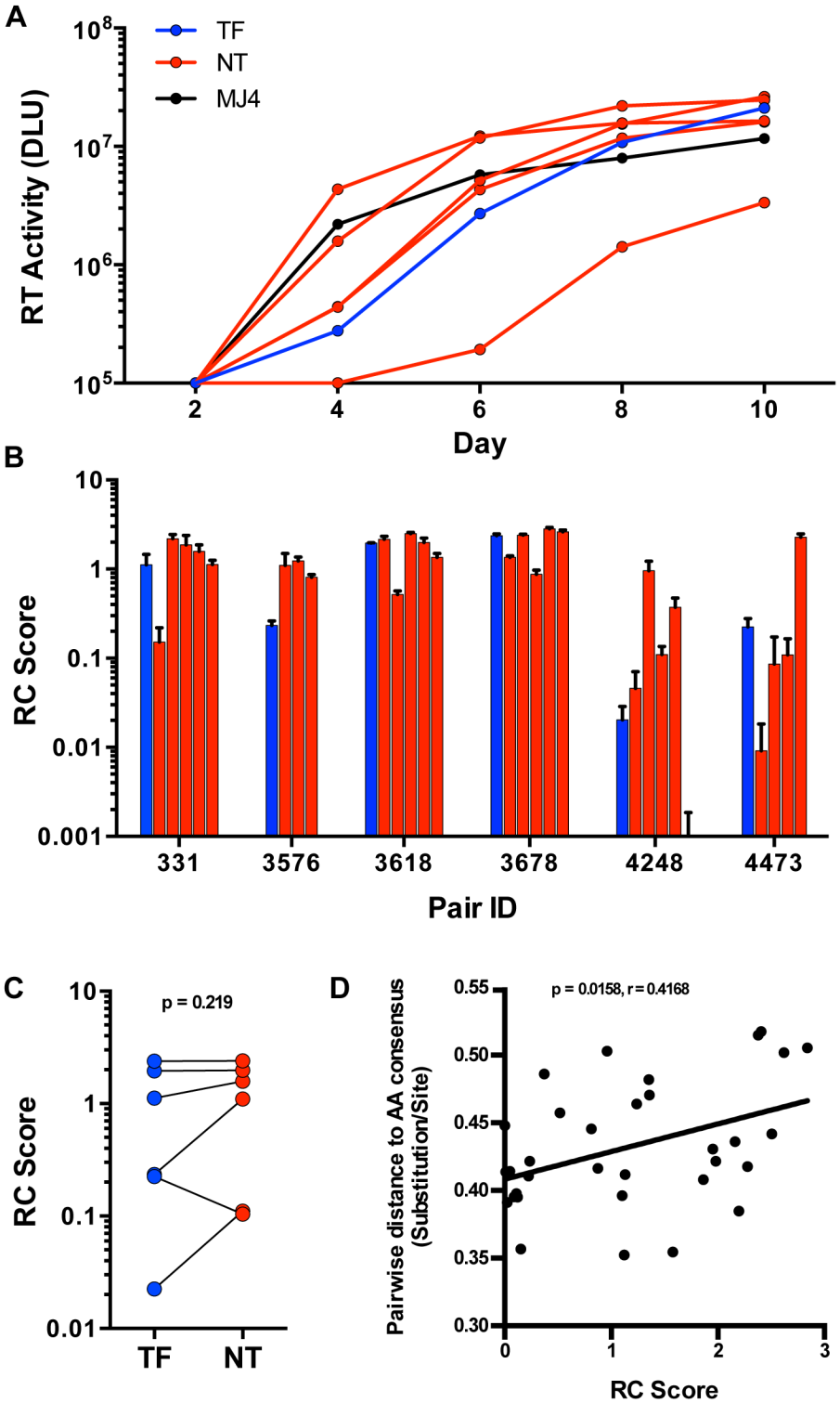

A
TF
NT
MJ4
RT Activity (DLU)
Day
B
RC Score
331
3576
3618
3678
4248
4473
Pair ID
C
p = 0.219
RC Score
TF
NT
D
p = 0.0158, r = 0.4168
Pairwise distance to AA consensus (Substitution/Site)
RC Score

Figure 5 - *In vitro* replication of TF and NT viruses in PBMC.

(A) Virus growth over 10 days in PBMC culture as measured by reverse transcriptase (RT) activity (DLU = digital light units) of the TF (blue), NT variants (red) and MJ4 standard (black) for one representative transmission pair, 331. (B) Replicative capacity (RC) scores, based on the area under the curve relative to MJ4, of all tested TF (blue) and NT (red) variants from six transmission pairs. (C) RC scores of the TF (blue) compared to the median of the corresponding NT variants (red) (Wilcoxon matched-pairs signed rank test, two-tailed p = 0.219). (D) Spearman correlation of the pairwise distance to the amino acid subtype C consensus described in Fig. 2 and RC scores over all variants tested (Non-parametric Spearman p = 0.0158, r = 0.4168.) The linear regression line is shown for visualization purposes.

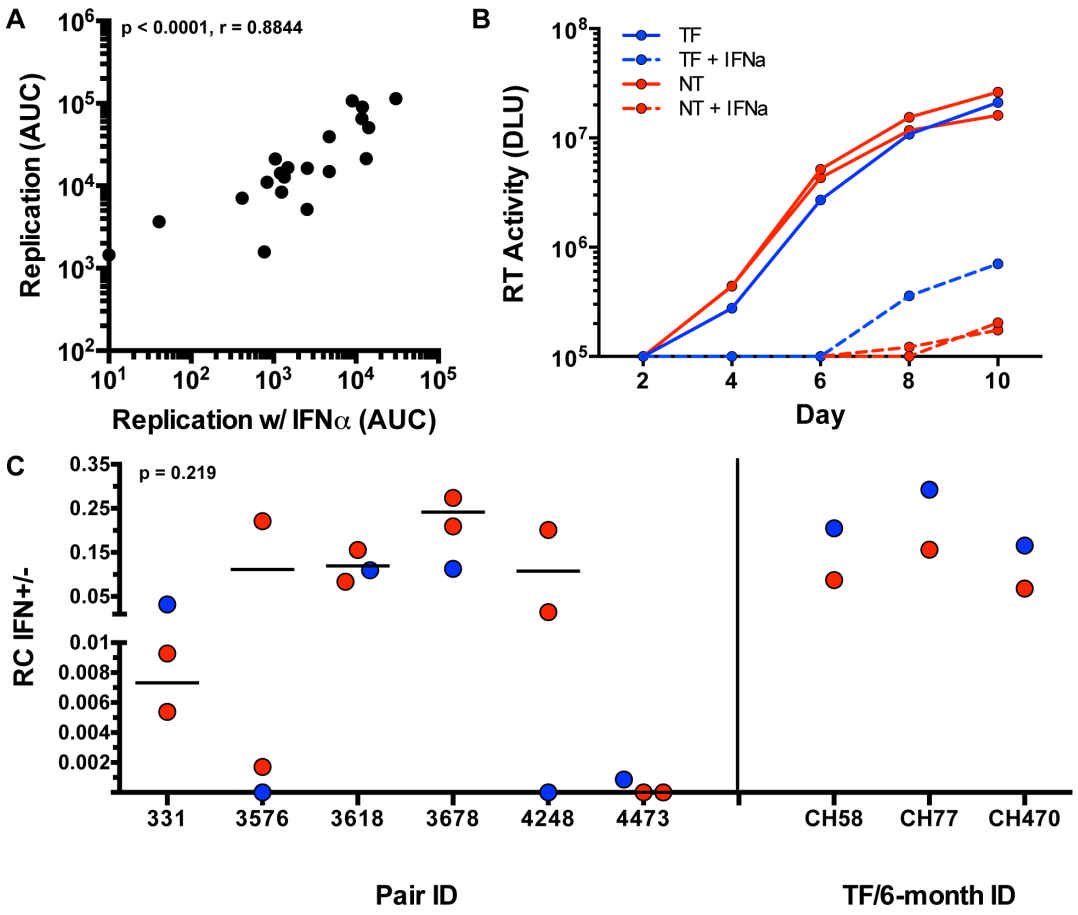
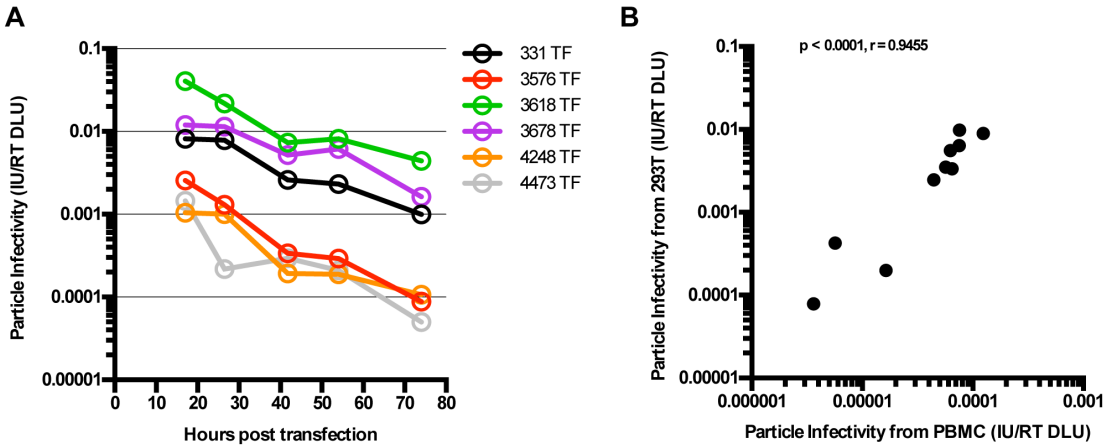


Figure 6 - Interferon- α resistance of TF and NT viruses.

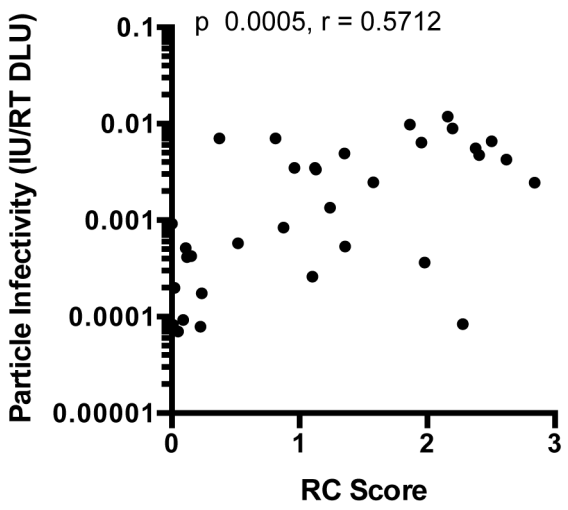
(A) Spearman correlation of the replication measured by area under the curve (AUC) (y-axis) of each tested variant (black dots) and the replication (AUC) in the presence of interferon alpha (x-axis) (Non-parametric Spearman $p < .0001$, $r = 0.8844$). (B) Virus growth over 10 days in culture as measured by reverse transcriptase (RT) activity (DLU = digital light units) of the TF (blue), and NT variants (red) in the presence of IFN-α (dotted) and absence of IFN-α (solid lines) in an example pair 331. (C) RC scores in the presence of IFN-α were divided by the RC score in the absence of IFN-α for TF (blue) and selected NT (red) viruses with similar replication kinetics (Wilcoxon matched-pairs signed rank test, two

tailed, p = 0.219). Subtype B TF (blue) and 6-month consensus (red) viruses are shown as controls on the right.



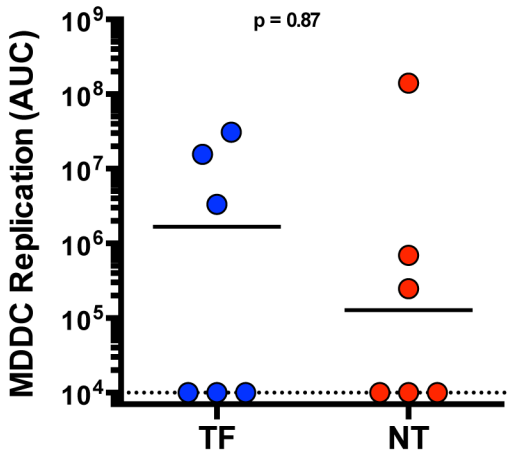
Supplemental Figure 1 - Particle infectivity from 293T and PBMC derived virus.

(A) Particle infectivity (TZM-bl titer divided by reverse transcriptase activity) of 293T cell derived TF virus stocks at different time points post transfection. (B) Correlation of particle infectivity assessed from day 8 of a PBMC infection and the particle infectivity from 293T derived stocks 48 hours after transfection of a subset of eleven viruses ($p < 0.0001$, $r = 0.9455$).



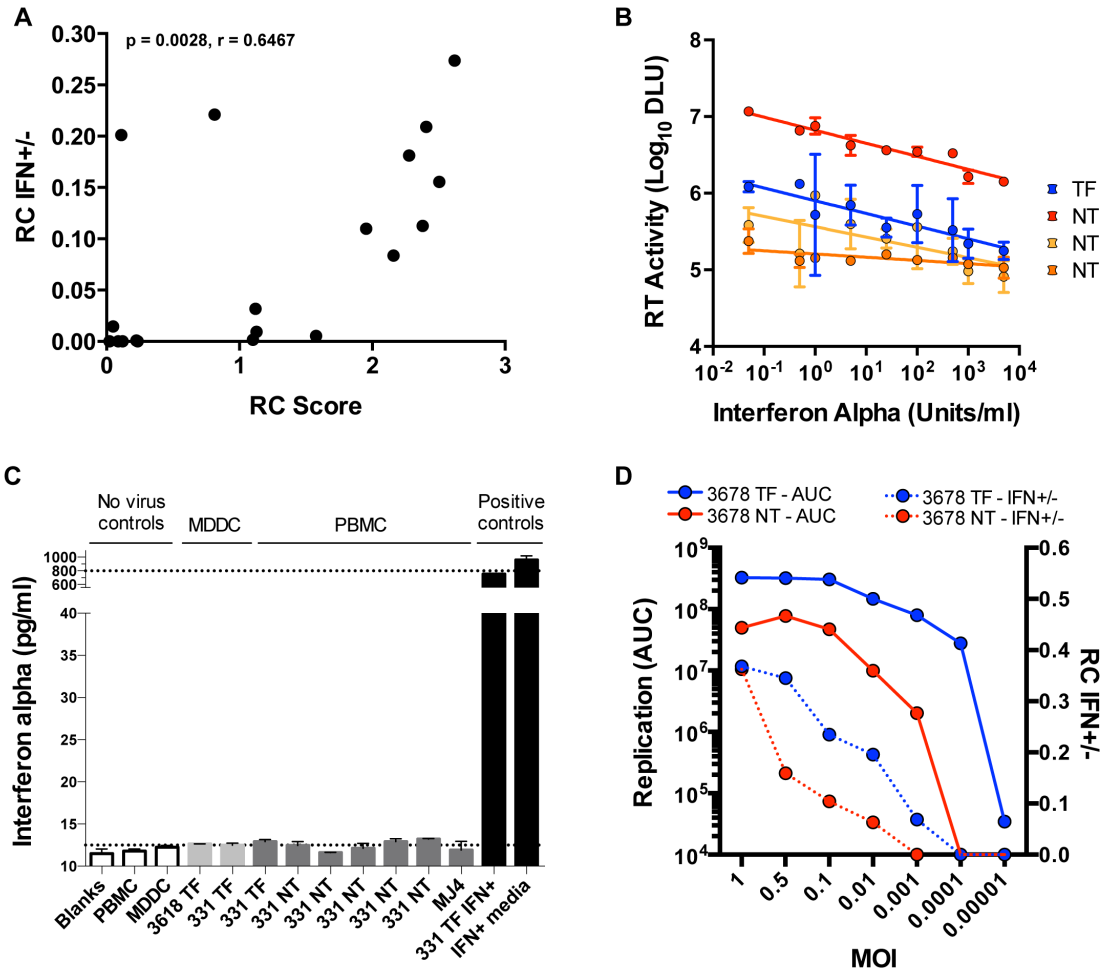
Supplemental Figure 2 - Particle infectivity correlates with replicative capacity.

Spearman correlation of particle infectivity and replicative capacity score of all TF & NT virus variants ($p = 0.0005$, $r = 0.5712$).



Supplemental Figure 3 - Replication of TF and NT viruses in monocyte derived dendritic cells.

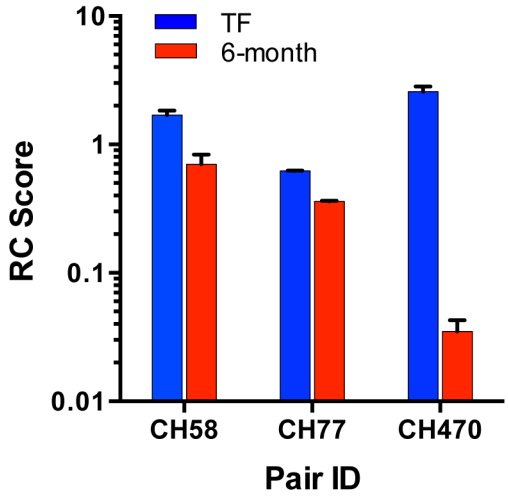
Virus growth in monocyte derived dendritic cells was measured by analyzing supernatant reverse transcriptase activity for 12 days following infection. Replication is depicted (y-axis) as the area under the curve for each virus variant. TF (blue) and NT (red) are presented with their group median. The difference between the groups was analyzed using a two-tailed Mann Whitney test (p = 0.87). Results are the average of replication in two healthy donors.



Supplemental Figure 4 - TF and NT resistance to IFN- α.

(A) Correlation of IFN-α resistance (RC IFN+/-) and RC Score of variants from figure 6C (p = 0.0028, r = 0.6467). (B) The TF (blue) and three NT (red, yellow, orange) variants from pair 331 with a representative range of RC scores were tested for replication in the presence of IFN-α concentrations from 0.5 U/ml - 10,000 U/ml. Supernatant reverse transcriptase (RT) activity at day 7 post-infection are shown. (C) Analysis of IFN-a levels in day 8 supernatants from PBMC and MDDC infected with a subset of viruses, to test for IFN-α induction *in vitro*. Negative controls are shown in white, MDDC infections in light gray,

PBMC infections in dark gray, and positive controls in black. (D) Area under the curve (solid lines) and IFN-α resistance ratios (dotted lines) from infections initiated at a range of MOI for the 3678 TF (blue) , along with an NT variant (red) with a different replicative capacity.



Supplemental Figure 5 - Replication of TF and 6-month consensus infectious molecular clones.

RC scores of three subtype B TF/6-month virus pairs (described in (17)) in activated PBMC.

References

1. Derdeyn CA, Decker JM, Bibollet-Ruche F, Mokili JL, Muldoon M, Denham SA, et al. Envelope-constrained neutralization-sensitive HIV-1 after heterosexual transmission. Science. 2004Mar.26;303(5666):2019–22.

2. Keele BF, Giorgi EE, Salazar-Gonzalez JF, Decker JM, Pham KT, Salazar MG, et al. Identification and characterisation of transmitted and early founder virus envelopes in primary HIV-1 infection. P Natl Acad Sci Usa. 2008;105(21):7552–7.

3. Haaland RE, Hawkins PA, Salazar-Gonzalez J, Johnson A, Tichacek A, Karita E, et al. Inflammatory Genital Infections Mitigate a Severe Genetic Bottleneck in Heterosexual Transmission of Subtype A and C HIV-1. Trkola A, editor. PLoS Pathog. 2009Jan.23;5(1):e1000274.

4. Salazar-Gonzalez JF, Salazar MG, Keele BF, Learn GH, Giorgi EE, Li H, et al. Genetic identity, biological phenotype, and evolutionary pathways of transmitted/founder viruses in acute and early HIV-1 infection. Journal of Experimental Medicine. 2009Jun.8;206(6):1273–89.

5. Boeras DI, Hraber PT, Hurlston M, Evans-Strickfaden T, Bhattacharya T, Giorgi EE, et al. Role of donor genital tract HIV-1 diversity in the transmission bottleneck. Proceedings of the National Academy of Sciences. 2011Nov.7.

6. Frange P, Meyer L, Jung M, Goujard C, Zucman D, Abel S, et al. Sexually-

Transmitted/Founder HIV-1 Cannot Be Directly Predicted from Plasma or PBMC-Derived Viral Quasispecies in the Transmitting Partner. Mammano F, editor. Plos One. 2013Jul.9;8(7):e69144.

7. Carlson JM, Schaefer M, Monaco DC, Batorsky R, Claiborne DT, Prince J, et al. Selection bias at the heterosexual HIV-1 transmission bottleneck. Science. 2014Jul.10;345(6193):1254031–1.

8. Sagar M, Laeyendecker O, Lee S, Gamiel J, Wawer MJ, Gray RH, et al. Selection of HIV Variants with Signature Genotypic Characteristics during Heterosexual Transmission. J Infect Dis. 2009Feb.15;199(4):580–9.

9. Pena-Cruz V, Etemad B, Chatziandreou N, Nyein PH, Stock S, Reynolds SJ, et al. HIV-1 envelope replication and α4β7 utilization among newly infected subjects and their corresponding heterosexual partners. Retrovirology. 2013;10:162.

10. Nawaz F, Cicala C, van Ryk D, Block KE, Jelicic K, McNally JP, et al. The Genotype of Early-Transmitting HIV gp120s Promotes α4β7 –Reactivity, Revealing α4β7+/CD4+ T cells As Key Targets in Mucosal Transmission. Trkola A, editor. PLoS Pathog. 2011Feb.24;7(2):e1001301.

11. Gnanakaran S, Bhattacharya T, Daniels M, Keele BF, Hraber PT, Lapedes AS, et al. Recurrent Signature Patterns in HIV-1 B Clade Envelope Glycoproteins Associated with either Early or Chronic Infections. Young JAT, editor. PLoS Pathog. 2011Sep.29;7(9):e1002209.

12. Cavrois M, Neidleman J, Santiago ML, Derdeyn CA, Hunter E, Greene WC. Enhanced Fusion and Virion Incorporation for HIV-1 Subtype C Envelope Glycoproteins with Compact V1/V2 Domains. J Virol. 2013Dec.11.

13. Parrish NF, Gao F, Li H, Giorgi EE, Barbian HJ, Parrish EH, et al. Phenotypic properties of transmitted founder HIV-1. Proceedings of the National Academy of Sciences. 2013Apr.23;110(17):6626–33.

14. Ochsenbauer C, Edmonds TG, Ding H, Keele BF, Decker J, Salazar MG, et al. Generation of Transmitted/Founder HIV-1 Infectious Molecular Clones and Characterization of Their Replication Capacity in CD4 T Lymphocytes and Monocyte-Derived Macrophages. J Virol. 2012Feb.11;86(5):2715–28.

15. Liu Y, Curlin ME, Diem K, Zhao H, Ghosh AK, Zhu H, et al. Env length and N-linked glycosylation following transmission of human immunodeficiency virus Type 1 subtype B viruses. Virology. 2008May;374(2):229–33.

16. Baalwa J, Wang S, Parrish NF, Decker JM, Keele BF, Learn GH, et al. Molecular identification, cloning and characterization of transmitted/founder HIV-1 subtype A, D and A/D infectious molecular clones. Virology. 2012Nov..

17. Fenton-May AE, Dibben O, Emmerich T, Ding H, Pfafferott K, Aasa-Chapman MM, et al. Relative resistance of HIV-1 founder viruses to control by interferon-alpha. Retrovirology. 2013;10:146.

18. Frost SDW, Liu Y, Pond SLK, Chappey C, Wrin T, Petropoulos CJ, et al.

Characterization of Human Immunodeficiency Virus Type 1 (HIV-1) Envelope Variation and Neutralizing Antibody Responses during Transmission of HIV-1 Subtype B. J Virol. 2005Apr.27;79(10):6523–7.

19. Chohan B, Lang D, Sagar M, Korber B, Lavreys L, Richardson B, et al. Selection for Human Immunodeficiency Virus Type 1 Envelope Glycosylation Variants with Shorter V1-V2 Loop Sequences Occurs during Transmission of Certain Genetic Subtypes and May Impact Viral RNA Levels. J Virol. 2005Apr.27;79(10):6528–31.

20. van't Wout AB, Kootstra NA, Mulder-Kampinga GA, Albrecht-van Lent N, Scherpbier HJ, Veenstra J, et al. Macrophage-tropic variants initiate human immunodeficiency virus type 1 infection after sexual, parenteral, and vertical transmission. Journal of Clinical Investigation. 1994Nov.;94(5):2060–7.

21. Richardson SI, Gray ES, Mkhize NN, Sheward DJ, Lambson BE, Wibmer CK, et al. South African HIV-1 subtype C transmitted variants with a specific V2 motif show higher dependence on α4β7 for replication. Retrovirology. BioMed Central; 2015Jun.23;:1–16.

22. Redd AD, Collinson-Streng AN, Chatziandreou N, Mullis CE, Laeyendecker O, Martens C, et al. Previously Transmitted HIV-1 Strains Are Preferentially Selected During Subsequent Sexual Transmissions. J Infect Dis. 2012Oct.8;206(9):1433–42.

23. Isaacman-Beck J, Hermann EA, Yi Y, Ratcliffe SJ, Mulenga J, Allen S, et al.

Heterosexual transmission of human immunodeficiency virus type 1 subtype C: Macrophage tropism, alternative coreceptor use, and the molecular anatomy of CCR5 utilization. J Virol. 2009Aug.;83(16):8208–20.

24. Alexander M, Lynch R, Mulenga J, Allen S, Derdeyn CA, Hunter E. Donor and recipient envs from heterosexual human immunodeficiency virus subtype C transmission pairs require high receptor levels for entry. J Virol. 2010Apr.;84(8):4100–4.

25. Parrish NF, Wilen CB, Banks LB, Iyer SS, Pfaff JM, Salazar-Gonzalez JF, et al. Transmitted/Founder and Chronic Subtype C HIV-1 Use CD4 and CCR5 Receptors with Equal Efficiency and Are Not Inhibited by Blocking the Integrin α4β7. Trkola A, editor. PLoS Pathog. 2012May31;8(5):e1002686.

26. Deymier MJ, Claiborne DT, Ende Z, Ratner HK, Kilembe W, Allen S, et al. Particle infectivity of HIV-1 full-length genome infectious molecular clones in a subtype C heterosexual transmission pair following high fidelity amplification and unbiased cloning. Virology. 2014Sep.19;468-470C:454–61.

27. Wei X, Decker JM, Liu H, Zhang Z, Arani RB, Kilby JM, et al. Emergence of Resistant Human Immunodeficiency Virus Type 1 in Patients Receiving Fusion Inhibitor (T-20) Monotherapy. Antimicrobial Agents and Chemotherapy. 2002Jun.1;46(6):1896–905.

28. Dugast A-S, Chan Y, Hoffner M, Licht A, Nkolola J, Li H, et al. Lack of

protection following passive transfer of polyclonal highly functional low-dose non-neutralizing antibodies. Plos One. 2014;9(5):e97229.

29. Wei X, Decker JM, Wang S, Hui H, Kappes JC, Wu X, et al. Antibody neutralization and escape by HIV-1. Nature. 2003Mar.20;422(6929):307–12.

30. Hu J, Gardner MB, Miller CJ. Simian immunodeficiency virus rapidly penetrates the cervicovaginal mucosa after intravaginal inoculation and infects intraepithelial dendritic cells. J Virol. 2000Jul.;74(13):6087–95.

31. Ping LH, Joseph SB, Anderson JA, Abrahams MR, Salazar-Gonzalez JF, Kincer LP, et al. Comparison of Viral Env Proteins from Acute and Chronic Infections with Subtype C Human Immunodeficiency Virus Type 1 Identifies Differences in Glycosylation and CCR5 Utilization and Suggests a New Strategy for Immunogen Design. J Virol. 2013Jun.10;87(13):7218–33.

32. Alizon S, Fraser C. Within-host and between-host evolutionary rates across the HIV-1 genome. Retrovirology. 2013;10:49.

33. Phillips RE, Rowland-Jones S, Nixon DF, Gotch FM, Edwards JP, Ogunlesi AO, et al. Human immunodeficiency virus genetic variation that can escape cytotoxic T cell recognition. Nature. 1991Dec.12;354(6353):453–9.

34. Richman DD, Wrin T, Little SJ, Petropoulos CJ. Rapid evolution of the neutralizing antibody response to HIV type 1 infection. P Natl Acad Sci Usa. 2003Apr.1;100(7):4144–9.

35. Wilen CB, Parrish NF, Pfaff JM, Decker JM, Henning EA, Haim H, et al. Phenotypic and Immunologic Comparison of Clade B Transmitted/Founder and Chronic HIV-1 Envelope Glycoproteins. J Virol. 2011Aug.8;85(17):8514–27.

36. Prince JL, Claiborne DT, Carlson JM, Schaefer M, Yu T, Lahki S, et al. Role of Transmitted Gag CTL Polymorphisms in Defining Replicative Capacity and Early HIV-1 Pathogenesis. Lifson J, editor. PLoS Pathog. 2012Nov.29;8(11):e1003041.

37. Wright JK, Brumme ZL, Carlson JM, Heckerman D, Kadie CM, Brumme CJ, et al. Gag-protease-mediated replication capacity in HIV-1 subtype C chronic infection: associations with HLA type and clinical parameters. J Virol. 2010Oct.;84(20):10820–31.

38. Yue L, Pfafferott KJ, Baalwa J, Conrod K, Dong CC, Chui C, et al. Transmitted Virus Fitness and Host T Cell Responses Collectively Define Divergent Infection Outcomes in Two HIV-1 Recipients. Douek DC, editor. PLoS Pathog. 2015Jan.8;11(1):e1004565.

39. Claiborne DT, Prince JL, Scully E, Macharia G, Micci L, Lawson B, et al. Replicative fitness of transmitted HIV-1 drives acute immune activation, proviral load in memory CD4 +T cells, and disease progression. Proceedings of the National Academy of Sciences. 2015Feb.17;:201421607.

40. Rouzine IM, Weinberger AD, Weinberger LS. An Evolutionary Role for HIV Latency in Enhancing Viral Transmission. Cell. Elsevier Inc;

2015Feb.26;160(5):1002–12.

41. Abel K, Rocke DM, Chohan B, Fritts L, Miller CJ. Temporal and anatomic relationship between virus replication and cytokine gene expression after vaginal simian immunodeficiency virus infection. J Virol. 2005Oct.;79(19):12164–72.

42. Li Q, Estes JD, Schlievert PM, Duan L, Brosnahan AJ, Southern PJ, et al. Glycerol monolaurate prevents mucosal SIV transmission. Nature. 2009Mar.4;458(7241):1034–8.

43. Etienne L, Hahn BH, Sharp PM, Matsen FA, Emerman M. Gene Loss and Adaptation to Hominids Underlie the Ancient Origin of HIV-1. Cell Host Microbe. Elsevier Inc; 2013Jul.17;14(1):85–92.

44. Sandler NG, Bosinger SE, Estes JD, Zhu RTR, Tharp GK, Boritz E, et al. Type I interferon responses in rhesus macaques prevent SIV infection and slow disease progression. Nature. Nature Publishing Group; 2014Jul.22;511(7511):601–5.

45. Shirazi Y, Pitha PM. Alpha interferon inhibits early stages of the human immunodeficiency virus type 1 replication cycle. J Virol. 1992Mar.;66(3):1321–8.

46. Künzi MS, Farzadegan H, Margolick JB, Vlahov D, Pitha PM. Identification of human immunodeficiency virus primary isolates resistant to interferon-alpha and correlation of prevalence to disease progression. J Infect Dis.

1995Apr.;171(4):822–8.

47. Edlin BR, St Clair MH, Pitha PM, Whaling SM, King DM, Bitran JD, et al. In-vitro resistance to zidovudine and alpha-interferon in HIV-1 isolates from patients: correlations with treatment duration and response. Ann. Intern. Med. 1992Sep.15;117(6):457–60.

48. Trask SA, Derdeyn CA, Fideli U, Chen Y, Meleth S, Kasolo F, et al. Molecular epidemiology of human immunodeficiency virus type 1 transmission in a heterosexual cohort of discordant couples in Zambia. J Virol. 2002Jan.;76(1):397–405.

49. Rousseau CM, Birditt BA, McKay AR, Stoddard JN, Lee TC, McLaughlin S, et al. Large-scale amplification, cloning and sequencing of near full-length HIV-1 subtype C genomes. J. Virol. Methods. 2006Sep.;136(1-2):118–25.

50. Edgar RC. MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res. 2004Mar.8;32(5):1792–7.

51. Deng W, Maust B, Nickle D, Learn G, Liu Y, Heath L, et al. DIVEIN: a web server to analyze phylogenies, sequence divergence, diversity, and informative sites. Biotech. 2010May;48(5):405–8.

52. Nickle DC, Heath L, Jensen MA, Gilbert PB, Mullins JI, Kosakovsky Pond SL. HIV-Specific Probabilistic Models of Protein Evolution. Pybus O, editor. Plos One. 2007Jun.6;2(6):e503.

53. Claiborne DT, Prince JL, Hunter E. A Restriction Enzyme Based Cloning

Method to Assess the In vitro Replication Capacity of HIV-1 Subtype C Gag-MJ4 Chimeric Viruses. JoVE (Journal of Visualized Experiments). 2014;(90):e51506–6.

Chapter IV : Discussion

Viruses, like HIV-1, are important human pathogens and cause global pandemics which require effective treatments and vaccines. A major challenge associated with developing effective vaccine strategies against HIV is due to the complexity of viral evolution, and the level of adaptation that these small RNA viruses go through. Understanding the mechanisms of transmission is important, not solely in the context of contributing to enhanced approaches to vaccine development, but can help describe the evolutionary bottlenecks these viruses undergo, providing insight into environmental pressures these parasites have been co-evolving with in their natural hosts. The genetic bottleneck, or founder effect, which occurs during HIV transmission offers a unique opportunity to target homogeneous viral population. In addition, if viral characteristics required for transmission can be uncovered, this already highly inefficient process may be manipulated to prevent all future infections.

Ernst Mayer, an evolutionary biologist, proposed, in 1942, a theory which stated that founder effects, where from a diverse genetic pool of individuals, a few or single individuals establish a new independent population, contribute to the generation of new species (107). During this process the founding population may contain, by chance, distinct characteristics from the original population which may drive the evolution of an entirely new species (107). This commonly occurs during population bottlenecks where new environmental pressures reduce a species' diversity to a single or few variants with increased fitness in the face of this new pressure. In Chapter 1 we discussed in depth the factors present during HIV-1 transmission, and found that this event can be described by such a founder

event which is not purely stochastic, but has a selection bias for variants with certain transmission fitness advantages (58,59,86). An essential point in researching HIV transmission bottlenecks was to determine whether this process is purely stochastic or whether it selects for variants with certain traits. The first evidence to suggest this was observed by sequencing the viral population in the chronically infected donors of linked transmission pairs where Transmitted/Founder (TF) variants were often found to be minor variants within the donor virus populations (61,85,108). Although it is essentially impossible to observe the viral population pool present in the donor genital tract at the time of transmission we have previously shown that the genital tract composition can remain stable for a period of a few months, within the timeframe of our sampling (85). In order to determine whether natural selection was occurring during transmission, we determined in a cohort of 137 heterosexual transmission pairs the amino acid sequences of *gag*, *pol* and *nef* genes from both partners and observed that consensus amino acid residues in the cohort were more favorably transmitted than non-consensus polymorphisms (86). This selection bias was mitigated by factors known previously to mitigate the bottleneck during transmission, thus strengthening the physiological relevance of the finding (60,86). This selection bias had also been previously seen in smaller cohorts by sequencing *gag* and *gp41* of the *env* gene, wherein selection of transmitted variants that more closely resembled the most recent common ancestor (MRCA) of the donor quasispecies were transmitted (74,87). In Chapter III, by sequencing full-length single genome amplicons we observed in six linked transmission pairs that the TF virus was genetically closer to the subtype C consensus sequence

across the entire viral genome than the majority of the donor variants (106). We therefore confirmed that selection for consensus-like variants is present across the entire viral genome, both for DNA and amino acid sequences (106). Together these findings all contribute to an evolutionary observation, that while the virus population evolves in the chronically infected host, it accumulates mutations, which decreases those variants' transmission potential. The accumulation of mutations driven by the host immune response drives a majority of the viral population away from a form most optimized for transmission. We observed by full-length sequencing multiple variants present in a chronically infected individual a complex population of viruses with varying genetic composition of up to approximately 5% diversity across the viral genome (106). This high intrahost viral diversity arises from the viruses' error-prone reverse transcriptase, high viral turnover and viral levels *in vivo*, as well the ability for the virus to remain latent (23,54-57). This viral diversity occurs randomly yet mutations may become fixed in the face of a an ever changing and persistent human immune response (52,53). Importantly viral latency allows for the quasispecies at a single time to be composed of viruses that have undergone differing amounts of rounds of replication. The concept of stepping back in time during transmission, can also be observed cross-sectional analysis, where others have shown that within-host diversity of HIV is more diverse than the between-host diversity (91). These findings reinforce the thought that during transmission bottlenecks the diverse viral quasispecies is reduced across the population to a smaller set of viruses with some common genetic structure. Consistent with these findings it had previously been suggested that adaptation and escape from antibody responses by extension

of variable loops in Env had a fitness cost in terms of transmission (58). We also observed in six linked transmission pairs that although neutralization of contemporaneous virus to autologous plasma is generally low, the TF variants were more sensitive than NT variants to neutralization (106). Therefore as the virus adapts within the host, it evolves away from a form, which was favored during the transmission or early establishment of infection of the previous host. This suggests that the increased rate of transmission observed from acutely infected individuals may not be solely based on viral load, but may include a pool of viruses with increased transmission potential, that has been established before acquiring too many cytotoxic T lymphocyte (CTL) and antibody escape mutations (29,109). HIV transmission clearly selects for variants that are more consensus-like or ancestral, but so far no phenotypic characteristic has been associated with this genotype that can explain the mechanistic advantage observed during the transmission process. During Subtype B transmission, viruses often switch during chronic infection from using CCR5 coreceptor to using CXCR4, this switch is an example of dead end evolution, since these CXCR4 tropic have been found to be much less able to transmit to a naïve host (59,110). Although this coreceptor switch does not occur as frequently during Subtype C transmission, some phenotypic change that is preventing subsequent transmission is occurring, and determining the mechanism required might help discover better preventative therapies for subtype C infection in sub-Saharan Africa, where the global health burden is greatest (70,111).

An obvious advantage for establishing a new population of a species in a naïve environment would be to reproduce or replicate well. Although no

consistent phenotypic property associated with TF variants, besides CCR5 tropism, has been observed, the replicative fitness of full-length variants from heterosexual transmission pairs had never been tested (92,93,101). We hypothesized that viruses with increased replicative capacity may have increased transmission potential and would therefore be selected for during transmission. In order to test this hypothesis we developed novel methods, described in Chapter II, with which to amplify and generate full-length infectious molecular clones of HIV-1 Transmitted/Founder and Non-Transmitted variants in an accurate and high-throughput way (105). Previous studies, without the advantage of samples from transmission pairs, described for the first time with full-length infectious molecular clones that TF viruses had increased infectivity and resistance to Interferon-α (IFNα) *in vitro* compared to chronic control viruses (103). Follow up studies of IFNα resistance also demonstrated that TF variants had increased resistance to IFNα compared to 6-month consensus viruses determined from the same individual (104). In all these studies the relative advantage in the presence of IFNα was measured as a ratio of their replication in the presence of IFNα to replication in the absence of IFNα (103,104). These studies did not report whether TF variants were able to replicate better in the presence of IFNα compared to the chronic control strains, although they did obverse no significant differences between TF and chronic control viruses for replication in the absence of IFNα (103). Similar studies also described potential phenotypic properties of TF virus Env gene that suggested increased expression of Env leading to an viral increase in particle infectivity (78). For the first time, as described in Chapter III, using viruses derived from heterosexual transmission

pairs we measured *in vitro* infectivity, replication, and IFNα resistance for subtype C HIV-1 full-length genome TF and NT variants, and found that TF viruses had no increased replicative capacity, particle infectivity, or resistance to IFNα compared with the NT viruses in six transmission pairs (106). Along with our confirming that consensus-like TF viruses were selected for during the transmission bottleneck, we observed a significant correlation between the replicative capacity of a virus and its distance from consensus, where more consensus-like viruses presented with decreased replicative capacity (106). We suggest that despite the expectation of TF having increased transmission fitness, this transmission fitness may not be associated with *in vitro* replicative fitness measured by the assays here, which may more likely represent the replication of viruses during chronic phase of infection in activated CD4+ T cells.

As an alternative to our proposition, it may be possible that variants with decreased replicative capacity are advantaged during the transmission process. We previously noted in two separate cohorts, that the replicative capacity of Gag-chimeric viruses correlated similarly to ours, with distance to consensus, where more consensus-like Gag proteins led to more poorly replicating viruses (112,113). Additionally a previous study of phenotypic properties of Env-chimeras from linked transmission pairs observed an increased *in vitro* replication advantage from donor virus Envs in comparison to Envs from the linked recipients of nine transmission pairs (114). A recent theoretical paper published by Rouzine I.M. *et al.* proposes, after having independently provided evidence for latency to be hardwired into the viral replication cycle as a stochastic event, and not necessarily entirely based on the activation status of the infected cells, that the

role of inherent viral latency may be to enhance viral transmission (115,116). This theory suggests that viruses have been optimized to stochastically enter latency as a method of establishing infection during transmission upon further reactivation, as opposed to undergoing rapid replication and viral extinction in the genital mucosa prior to dispersal (115). This model predicts that if strains with increased replicative capacity are less likely to go latent upon infection they will exhibit a lower transmission potential (115). Further measurements in determining the potential of viral strains to undergo latency would be required to test this characteristic with TF and NT viruses.

Quasispecies theory, based on the evolution of RNA virus populations, also predicts that viruses that are lower replicators will be favored if they give rise to progeny that are on average more fit in their relative environments (117-119). This theory was developed to explain the survival and fitness landscapes of populations of asexual species with great genetic diversity such as RNA virus populations, similar to HIV-1 (118,120). Within this framework a virus known to have a high mutation rate will accumulate errors and undergo extinction unless recombination occurs, which is a highly efficient process during HIV replication (54). As opposed to survival of the fittest, this process selects via 'survival of the flattest', where flattest represents a flat mutational landscape. A genetic variant with flatter mutational landscape is one whose spectrum of mutations are most likely to remain fit, as opposed to a sharp peak from a landscape where often a single mutation would kill the individual variant (118,121). Much of the theoretical framework mentioned here has yet to be tested in the lab, and reagents generated for HIV research may be useful to these scientific fields.

If our *in vitro* models simply do not represent the conditions of transmission properly, we may need to improve upon the assays from which to probe the mechanisms of transmission, in order to more accurately determine the transmission potential of TF viruses in comparison to NT viruses. The most common *in vivo* model of HIV infection is the Rhesus Macaque (*Macaca mulatta),* where many transmission studies have been performed, and results of intravaginal, intrarectal, and penile infection have provided important information about target cells for infection, and the effects of potential transmission prevention strategies (35,122-124). Although this model accurately reflects HIV pathogenesis, these animals can only be infected with Simian Immunodeficiency Virus (SIV), not HIV. Extensive studies of the role of genetic selection, and viral fitness in HIV transmission cannot be accomplished in this model, and although the model is susceptible to SIV/HIV (SHIV) chimeras, these studies are useful only in order to test antibodies against the native HIV Env and not an accurate measure of HIV viral phenotypes (125-127). SHIV replication in the Rhesus Macaques requires SIV genes, and the multiple passages *in vivo* required for establishing pathogenic clones, would potentially remove any key phenotypic properties outside of Envelope, that HIV variants contain across their entire genome that are favorable for *in vivo* transmission.

In order to assess the true transmission potential of full-length HIV variants, one must use human *in vivo* models where the full-length HIV can productively infect and replicate. Such models exist, in *ex vivo* models of human genital tissues, such as inner foreskins following adult elective circumcision or cervicovaginal tissues provided following hysterectomy of healthy donors (128-

130). These samples may provide an important medium to observe the virus behavior in an environment relevant to HIV transmission. Studies with these tissues have been done primarily in the context of human immunology, and host requirements for HIV infection, yet very few have asked whether certain virus variants infect these tissues disproportionately well and whether they contradict or confirm *in vitro* replication findings. The bone-marrow liver thymus (BLT) humanized mouse model of HIV infection, has been used and shown to be susceptible to intravaginal HIV challenge and can sustain replication, even undergoing latent infection following antiretroviral treatment (131-134). These mice develop cytotoxic T cell responses that lead to common escape mutations in the viral genome (135). Within the vaginal tract and rectal mucosa of these humanized mice the cells responsible for HIV transmission, such as human resident CD4+ T cells as well as myeloid cells such as dendritic cells and macrophages are present (131). Future experiments using reagents developed in our studies, will investigate whether low dose mixed inoculum infections of these humanized mice distinguish *in vivo* transmission fitness of TF and NT variants.

The importance of understanding the viral requirements for transmission is crucial to develop a vaccine against HIV-1, which so far has met with little success (reviewed in (136)). Immunological knowledge of the required protective antibody or CTL responses must come together with the virological transmission requirements via common genetic and phenotypic traits for transmission. Importantly the studies described here provide an important perspective into evolution on the level of viral populations, and can provide insight into the mechanisms by which, not only HIV but many other species of

viral pathogens co-evolve with their host, in order to reveal weaknesses we may be able to leverage into potential vaccine targets.

References

1. Gottlieb MS, Schroff R, Schanker HM, Weisman JD, Fan PT, Wolf RA, et al. Pneumocystis carinii pneumonia and mucosal candidiasis in previously healthy homosexual men: evidence of a new acquired cellular immunodeficiency. N Engl J Med. 1981Dec.10;305(24):1425–31.

2. Barré-Sinoussi F, Chermann JC, Rey F, Nugeyre MT, Chamaret S, Gruest J, et al. Isolation of a T-lymphotropic retrovirus from a patient at risk for acquired immune deficiency syndrome (AIDS). Science. 1983May20;220(4599):868–71.

3. UNAIDS. Global Report 2013. 2013Nov.5;:1–198.

4. Worobey M, Gemmel M, Teuwen DE, Haselkorn T, Kunstman K, Bunce M, et al. Direct evidence of extensive diversity of HIV-1 in Kinshasa by 1960. Nature. 2008Oct.2;455(7213):661–4.

5. Sharp PM, Hahn BH. Origins of HIV and the AIDS Pandemic. Cold Spring Harb Perspect Med. 2011Sep.1;1(1):a006841–1.

6. Huet T, Cheynier R, Meyerhans A, Roelants G, Wain-Hobson S. Genetic organization of a chimpanzee lentivirus related to HIV-1. Nature. 1990May24;345(6273):356–9.

7. Hirsch VM, Olmsted RA, Murphey-Corb M, Purcell RH, Johnson PR. An African primate lentivirus (SIVsm) closely related to HIV-2. Nature.

1989Jun.1;339(6223):389–92.

8. Ndung'u T, Weiss RA. On HIV diversity. Aids. 2012Jun.;26(10):1255–60.

9. Hemelaar J, Gouws E, Ghys PD, Osmanov S. Global and regional distribution of HIV-1 genetic subtypes and recombinants in 2004. Aids. 2006Oct.24;20(16):W13–23.

10. Korber B, Gaschen B, Yusim K, Thakallapally R, Kesmir C, Detours V. Evolutionary and immunological implications of contemporary HIV-1 variation. Br. Med. Bull. 2001;58:19–42.

11. Arts EJ, Hazuda DJ. HIV-1 Antiretroviral Drug Therapy. Cold Spring Harb Perspect Med. 2012Apr.2;2(4):a007161–1.

12. St-Louis M-C, Cojocariu M, Archambault D. The molecular biology of bovine immunodeficiency virus: a comparison with other lentiviruses. Anim. Health. Res. Rev. 2007Feb.28;5(02):125–43.

13. Luckheeram RV, Zhou R, Verma AD, Xia B. CD4+T Cells: Differentiation and Functions. Clinical and Developmental Immunology. 2012;2012(12):1–12.

14. Dalgleish AG, Beverley PC, Clapham PR, Crawford DH, Greaves MF, Weiss RA. The CD4 (T4) antigen is an essential component of the receptor for the AIDS retrovirus. Nature. 1984Jan.;312(5996):763–7.

15. Alkhatib G, Combadiere C, Broder CC, Feng Y, Kennedy PE, Murphy PM,

et al. CC CKR5: a RANTES, MIP-1alpha, MIP-1beta receptor as a fusion cofactor for macrophage-tropic HIV-1. Science. 1996Jun.28;272(5270):1955–8.

16. Feng Y, Broder CC, Kennedy PE, Berger EA. HIV-1 entry cofactor: functional cDNA cloning of a seven-transmembrane, G protein-coupled receptor. Science. 1996May10;272(5263):872–7.

17. Baltimore D. RNA-dependent DNA polymerase in virions of RNA tumour viruses. Nature. 1970Jun.27;226(5252):1209–11.

18. Hu WS, Hughes SH. HIV-1 Reverse Transcription. Cold Spring Harb Perspect Med. 2012Oct.1;2(10):a006882–2.

19. Craigie R, Bushman FD. HIV DNA Integration. Cold Spring Harb Perspect Med. 2012Jul.2;2(7):a006890–0.

20. Wei X, Ghosh SK, Taylor ME, Johnson VA, Emini EA, Deutsch P, et al. Viral dynamics in human immunodeficiency virus type 1 infection. Nature. 1995Jan.12;373(6510):117–22.

21. Murray JM, Kelleher AD, Cooper DA. Timing of the Components of the HIV Life Cycle in Productively Infected CD4+ T Cells in a Population of HIV-Infected Individuals. J Virol. 2011Sep.21;85(20):10798–805.

22. Yamashita M, Emerman M. Retroviral infection of non-dividing cells: Old and new perspectives. Virology. 2006Jan.;344(1):88–93.

23. Siliciano RF, Greene WC. HIV Latency. Cold Spring Harb Perspect Med. 2011Sep.1;1(1):a007096–6.

24. Tyagi M, Pearson RJ, Karn J. Establishment of HIV Latency in Primary CD4+ Cells Is due to Epigenetic Transcriptional Silencing and P-TEFb Restriction. J Virol. 2010Jun.8;84(13):6425–37.

25. Davey RT, Bhat N, Yoder C, Chun TW. HIV-1 and T cell dynamics after interruption of highly active antiretroviral therapy (HAART) in patients with a history of sustained viral suppression. 1999.

26. Shaw GM, Hunter E. HIV Transmission. Cold Spring Harb Perspect Med. 2012;2(11).

27. Dunkle KL, Stephenson R, Karita E, Chomba E, Kayitenkore K, Vwalika C, et al. New heterosexually transmitted HIV infections in married or cohabiting couples in urban Zambia and Rwanda: an analysis of survey and clinical data. Lancet. 2008Jun.28;371(9631):2183–91.

28. Galvin SR, Cohen MS. The role of sexually transmitted diseases in HIV transmission. Nature Publishing Group. 2004Jan.;2(1):33–42.

29. Quinn TC, Wawer MJ, Sewankambo N, Serwadda D, Li C, Wabwire-Mangen F, et al. Viral load and heterosexual transmission of human immunodeficiency virus type 1. Rakai Project Study Group. N Engl J Med. 2000Mar.30;342(13):921–9.

30. Keele BF, Estes JD. Barriers to mucosal transmission of

immunodeficiency viruses. Blood. 2011Jul.28;118(4):839–46.

31. Lai SK, Hida K, Shukair S, Wang YY, Figueiredo A, Cone R, et al. Human Immunodeficiency Virus Type 1 Is Trapped by Acidic but Not by Neutralized Human Cervicovaginal Mucus. J Virol. 2009Oct.8;83(21):11196–200.

32. Fanibunda SE, Modi DN, Gokral JS, Bandivdekar AH. HIV gp120 binds to mannose receptor on vaginal epithelial cells and induces production of matrix metalloproteinases. Plos One. 2011;6(11):e28014.

33. Miller CJ, Li Q, Abel K, Kim E-Y, Ma Z-M, Wietgrefe S, et al. Propagation and dissemination of infection after vaginal transmission of simian immunodeficiency virus. J Virol. 2005Jul.;79(14):9217–27.

34. Stieh DJ, Maric D, Kelley ZL, Anderson MR, Hattaway HZ, Beilfuss BA, et al. Vaginal Challenge with an SIV-Based Dual Reporter System Reveals That Infection Can Occur throughout the Upper and Lower Female Reproductive Tract. Silvestri G, editor. PLoS Pathog. 2014Oct.9;10(10):e1004440.

35. Spira AI, Marx PA, Patterson BK, Mahoney J, Koup RA, Wolinsky SM, et al. Cellular targets of infection and route of viral dissemination after an intravaginal inoculation of simian immunodeficiency virus into rhesus macaques. J Exp Med. 1996Jan.1;183(1):215–25.

36. Dinh MH, Anderson MR, McRaven MD, Cianci GC, McCoombe SG, Kelley

ZL, et al. Visualization of HIV-1 Interactions with Penile and Foreskin Epithelia: Clues for Female-to-Male HIV Transmission. Silvestri G, editor. PLoS Pathog. 2015Mar.6;11(3):e1004729.

37. Auvert B, Taljaard D, Lagarde E, Sobngwi-Tambekou J, Sitta R, Puren A. Randomized, Controlled Intervention Trial of Male Circumcision for Reduction of HIV Infection Risk: The ANRS 1265 Trial. Plos Med. 2005;2(11):e298.

38. Cameron PU, Freudenthal PS, Barker JM, Gezelter S, Inaba K, Steinman RM. Dendritic cells exposed to human immunodeficiency virus type-1 transmit a vigorous cytopathic infection to CD4+ T cells. Science. 1992Jul.17;257(5068):383–7.

39. Pope M, Betjes MG, Romani N, Hirmand H, Cameron PU, Hoffman L, et al. Conjugates of dendritic cells and memory T lymphocytes from skin facilitate productive infection with HIV-1. Cell. 1994Aug.12;78(3):389–98.

40. Cameron PU, Lowe MG, Crowe SM, O'Doherty U, Pope M, Gezelter S, et al. Susceptibility of dendritic cells to HIV-1 infection in vitro. J. Leukoc. Biol. 1994Sep.;56(3):257–65.

41. Zhou Z, Barry de Longchamps N, Schmitt A, Zerbib M, Vacher-Lavenu M-C, Bomsel M, et al. HIV-1 Efficient Entry in Inner Foreskin Is Mediated by Elevated CCL5/RANTES that Recruits T Cells and Fuels Conjugate Formation with Langerhans Cells. Silvestri G, editor. PLoS Pathog.

2011Jun.30;7(6):e1002100.

42. Turville SG, Santos JJ, Frank I, Cameron PU, Wilkinson J, Miranda-Saksena M, et al. Immunodeficiency virus uptake, turnover, and 2-phase transfer in human dendritic cells. Blood. 2004Mar.15;103(6):2170–9.

43. Brenchley JM. CD4+ T Cell Depletion during all Stages of HIV Disease Occurs Predominantly in the Gastrointestinal Tract. Journal of Experimental Medicine. 2004Sep.20;200(6):749–59.

44. Alter G, Heckerman D, Schneidewind A, Fadda L, Kadie CM, Carlson JM, et al. HIV-1 adaptation to NK-cell-mediated immune pressure. Nature. 2011Aug.3;476(7358):96–100.

45. Haase AT. Perils at mucosal front lines for HIV and SIV and their hosts. Nat Rev Immunol. 2005Oct.;5(10):783–92.

46. Haase AT. Targeting early infection to prevent HIV-1 mucosal transmission. Nature. 2010Mar.11;464(7286):217–23.

47. Shaw GM, Hahn BH, Arya SK, Groopman JE, Gallo RC, Wong-Staal F. Molecular characterization of human T-cell leukemia (lymphotropic) virus type III in the acquired immune deficiency syndrome. Science. 1984Dec.7;226(4679):1165–71.

48. Byrareddy SN, Kallam B, Arthos J, Cicala C, Nawaz F, Hiatt J, et al. Targeting α4β7 integrin reduces mucosal transmission of simian immunodeficiency virus and protects gut-associated lymphoid

tissue from infection. Nat Med. Nature Publishing Group; 2014Nov.24;20(12):1397–400.

49. Nawaz F, Cicala C, van Ryk D, Block KE, Jelicic K, McNally JP, et al. The Genotype of Early-Transmitting HIV gp120s Promotes α4β7 –Reactivity, Revealing α4β7+/CD4+ T cells As Key Targets in Mucosal Transmission. Trkola A, editor. PLoS Pathog. 2011Feb.24;7(2):e1001301.

50. Lazarovits AI, Tidswell M. Expression and function of the MAdCAM-1 receptor, integrin alpha 4 beta 7, on human leukocytes. The Journal of …. 1994.

51. Santangelo PJ, Rogers KA, Zurla C, Blanchard EL, Gumber S, Strait K, et al. Whole-body immunoPET reveals active SIV dynamics in viremic and antiretroviral therapy–treated macaques. Nat Methods. Nature Publishing Group; 2015Mar.9;:1–9.

52. Richman DD, Wrin T, Little SJ, Petropoulos CJ. Rapid evolution of the neutralizing antibody response to HIV type 1 infection. P Natl Acad Sci Usa. 2003Apr.1;100(7):4144–9.

53. Phillips RE, Rowland-Jones S, Nixon DF, Gotch FM, Edwards JP, Ogunlesi AO, et al. Human immunodeficiency virus genetic variation that can escape cytotoxic T cell recognition. Nature. 1991Dec.12;354(6353):453–9.

54. Burke DS. Recombination in HIV: an important viral evolutionary

strategy. Emerging Infect. Dis. 1997Jul.;3(3):253–9.

55. Roberts JD, Bebenek K, Kunkel TA. The accuracy of reverse transcriptase from HIV-1. Science. 1988Nov.25;242(4882):1171–3.

56. Cuevas JM, Geller R, Garijo R, López-Aldeguer J, Sanjuán R. Extremely High Mutation Rate of HIV-1 In Vivo. Rowland-Jones SL, editor. PLoS Biol. 2015Sep.16;13(9):e1002251.

57. Ho DD, Neumann AU, Perelson AS, Chen W, Leonard JM, Markowitz M. Rapid turnover of plasma virions and CD4 lymphocytes in HIV-1 infection. Nature. 1995Jan.12;373(6510):123–6.

58. Derdeyn CA, Decker JM, Bibollet-Ruche F, Mokili JL, Muldoon M, Denham SA, et al. Envelope-constrained neutralization-sensitive HIV-1 after heterosexual transmission. Science. 2004Mar.26;303(5666):2019–22.

59. Keele BF, Giorgi EE, Salazar-Gonzalez JF, Decker JM, Pham KT, Salazar MG, et al. Identification and characterisation of transmitted and early founder virus envelopes in primary HIV-1 infection. P Natl Acad Sci Usa. 2008;105(21):7552–7.

60. Haaland RE, Hawkins PA, Salazar-Gonzalez J, Johnson A, Tichacek A, Karita E, et al. Inflammatory Genital Infections Mitigate a Severe Genetic Bottleneck in Heterosexual Transmission of Subtype A and C HIV-1. Trkola A, editor. PLoS Pathog. 2009Jan.23;5(1):e1000274.

61. Wolinsky SM, Wike CM, Korber BT, Hutto C, Parks WP, Rosenblum LL, et al. Selective transmission of human immunodeficiency virus type-1 variants from mothers to infants. Science. 1992;255(5048):1134.

62. Zhu T, Mo H, Wang N, Nam DS, Cao Y, Koup RA, et al. Genotypic and phenotypic characterization of HIV-1 patients with primary infection. Science. 1993;261(5125):1179.

63. Zhang LQ, MacKenzie P, Cleland A, Holmes EC, Brown AJ, Simmonds P. Selection for specific sequences in the external envelope protein of human immunodeficiency virus type 1 upon primary infection. J Virol. 1993Jun.;67(6):3345–56.

64. Wolfs TF, Zwart G, Bakker M, Goudsmit J. HIV-1 genomic RNA diversification following sexual and parenteral virus transmission. Virology. 1992Jul.;189(1):103–10.

65. Roos MT, Lange JM, de Goede RE, Coutinho RA, Schellekens PT, Miedema F, et al. Viral phenotype and immune response in primary human immunodeficiency virus type 1 infection. J Infect Dis. 1992Mar.;165(3):427–32.

66. van't Wout AB, Kootstra NA, Mulder-Kampinga GA, Albrecht-van Lent N, Scherpbier HJ, Veenstra J, et al. Macrophage-tropic variants initiate human immunodeficiency virus type 1 infection after sexual, parenteral, and vertical transmission. Journal of Clinical Investigation. 1994Nov.;94(5):2060–7.

67. Taylor BS, Hammer SM. The challenge of HIV-1 subtype diversity. N Engl J Med. 2008Oct.30;359(18):1965–6.

68. Berger EA, Murphy PM, Farber JM. Chemokine receptors as HIV-1 coreceptors: roles in viral entry, tropism, and disease. Annu. Rev. Immunol. 1999;17:657–700.

69. Connor RI, Sheridan KE, Ceradini D, Choe S, Landau NR. Change in coreceptor use correlates with disease progression in HIV-1--infected individuals. J Exp Med. 1997Feb.17;185(4):621–8.

70. Ping LH, Nelson JA, Hoffman IF, Schock J, Lamers SL, Goodman M, et al. Characterization of V3 sequence heterogeneity in subtype C human immunodeficiency virus type 1 isolates from Malawi: underrepresentation of X4 variants. J Virol. 1999Aug.;73(8):6271–81.

71. Liu R, Paxton WA, Choe S, Ceradini D, Martin SR, Horuk R, et al. Homozygous defect in HIV-1 coreceptor accounts for resistance of some multiply-exposed individuals to HIV-1 infection. Cell. 1996Aug.9;86(3):367–77.

72. Li M, Salazar-Gonzalez JF, Derdeyn CA, Morris L, Williamson C, Robinson JE, et al. Genetic and Neutralization Properties of Subtype C Human Immunodeficiency Virus Type 1 Molecular env Clones from Acute and Early Heterosexually Acquired Infections in Southern Africa. J Virol. 2006Nov.10;80(23):11776–90.

73. Chohan B, Lang D, Sagar M, Korber B, Lavreys L, Richardson B, et al. Selection for Human Immunodeficiency Virus Type 1 Envelope Glycosylation Variants with Shorter V1-V2 Loop Sequences Occurs during Transmission of Certain Genetic Subtypes and May Impact Viral RNA Levels. J Virol. 2005Apr.27;79(10):6528–31.

74. Sagar M, Laeyendecker O, Lee S, Gamiel J, Wawer MJ, Gray RH, et al. Selection of HIV Variants with Signature Genotypic Characteristics during Heterosexual Transmission. J Infect Dis. 2009Feb.15;199(4):580–9.

75. Frost SDW, Liu Y, Pond SLK, Chappey C, Wrin T, Petropoulos CJ, et al. Characterization of Human Immunodeficiency Virus Type 1 (HIV-1) Envelope Variation and Neutralizing Antibody Responses during Transmission of HIV-1 Subtype B. J Virol. 2005Apr.27;79(10):6523–7.

76. Salazar-Gonzalez JF, Bailes E, Pham KT, Salazar MG, Guffey MB, Keele BF, et al. Deciphering Human Immunodeficiency Virus Type 1 Transmission and Early Envelope Diversification by Single-Genome Amplification and Sequencing. J Virol. 2008Mar.25;82(8):3952–70.

77. Gnanakaran S, Bhattacharya T, Daniels M, Keele BF, Hraber PT, Lapedes AS, et al. Recurrent Signature Patterns in HIV-1 B Clade Envelope Glycoproteins Associated with either Early or Chronic Infections. Young JAT, editor. PLoS Pathog. 2011Sep.29;7(9):e1002209.

78. Asmal M, Hellmann I, Liu W, Keele BF, Perelson AS, Bhattacharya T, et

al. A Signature in HIV-1 Envelope Leader Peptide Associated with Transition from Acute to Chronic Infection Impacts Envelope Processing and Infectivity. Vartanian J-P, editor. Plos One. 2011Aug.18;6(8):e23673.

79. Abrahams MR, Anderson JA, Giorgi EE, Seoighe C, Mlisana K, Ping LH, et al. Quantitating the Multiplicity of Infection with Human Immunodeficiency Virus Type 1 Subtype C Reveals a Non-Poisson Distribution of Transmitted Variants. J Virol. 2009;83(8):3556–67.

80. Li H, Bar KJ, Wang S, Decker JM, Chen Y, Sun C, et al. High Multiplicity Infection by HIV-1 in Men Who Have Sex with Men. Douek DC, editor. PLoS Pathog. 2010May13;6(5):e1000890.

81. Masharsky AE, Dukhovlinova EN, Verevochkin SV, Toussova OV, Skochilov RV, Anderson JA, et al. A Substantial Transmission Bottleneck among Newly and Recently HIV-1–Infected Injection Drug Users in St Petersburg, Russia. J Infect Dis. 2010Jun.;201(11):1697–702.

82. Bar KJ, Li H, Chamberland A, Tremblay C, Routy J-P, Grayson T, et al. Wide variation in the multiplicity of HIV-1 infection among injection drug users. J Virol. 2010Jun.;84(12):6241–7.

83. Joseph SB, Swanstrom R, Kashuba ADM, Cohen MS. Bottlenecks in HIV-1 transmission: insights from the study of founder viruses. Nature Publishing Group. Nature Publishing Group; 2015Jun.8;13(7):414–25.

84. Zhu T, Wang N, Carr A, Nam DS, Moor-Jankowski R, Cooper DA, et al.

Genetic characterization of human immunodeficiency virus type 1 in blood and genital secretions: evidence for viral compartmentalization and selection during sexual transmission. J Virol. 1996May;70(5):3098–107.

85. Boeras DI, Hraber PT, Hurlston M, Evans-Strickfaden T, Bhattacharya T, Giorgi EE, et al. Role of donor genital tract HIV-1 diversity in the transmission bottleneck. Proceedings of the National Academy of Sciences. 2011Nov.7.

86. Carlson JM, Schaefer M, Monaco DC, Batorsky R, Claiborne DT, Prince J, et al. Selection bias at the heterosexual HIV-1 transmission bottleneck. Science. 2014Jul.10;345(6193):1254031–1.

87. Redd AD, Collinson-Streng AN, Chatziandreou N, Mullis CE, Laeyendecker O, Martens C, et al. Previously Transmitted HIV-1 Strains Are Preferentially Selected During Subsequent Sexual Transmissions. J Infect Dis. 2012Oct.8;206(9):1433–42.

88. Li B, Decker JM, Johnson RW, Bibollet-Ruche F, Wei X, Mulenga J, et al. Evidence for potent autologous neutralizing antibody titers and compact envelopes in early infection with subtype C human immunodeficiency virus type 1. J Virol. 2006Jun.;80(11):5211–8.

89. West AP Jr, Scharf L, Scheid JF, Klein F, Bjorkman PJ, Nussenzweig MC. Structural Insights on the Role of Antibodies in HIV-1 Vaccine and Therapy. Cell. Elsevier Inc; 2014Feb.13;156(4):633–48.

90. Overbaugh J, Morris L. The Antibody Response against HIV-1. Cold Spring Harb Perspect Med. 2012Jan.3;2(1):a007039–9.

91. Alizon S, Fraser C. Within-host and between-host evolutionary rates across the HIV-1 genome. Retrovirology. 2013;10:49.

92. Alexander M, Lynch R, Mulenga J, Allen S, Derdeyn CA, Hunter E. Donor and recipient envs from heterosexual human immunodeficiency virus subtype C transmission pairs require high receptor levels for entry. J Virol. 2010Apr.;84(8):4100–4.

93. Isaacman-Beck J, Hermann EA, Yi Y, Ratcliffe SJ, Mulenga J, Allen S, et al. Heterosexual transmission of human immunodeficiency virus type 1 subtype C: Macrophage tropism, alternative coreceptor use, and the molecular anatomy of CCR5 utilization. J Virol. 2009Aug.;83(16):8208–20.

94. Wilen CB, Parrish NF, Pfaff JM, Decker JM, Henning EA, Haim H, et al. Phenotypic and Immunologic Comparison of Clade B Transmitted/Founder and Chronic HIV-1 Envelope Glycoproteins. J Virol. 2011Aug.8;85(17):8514–27.

95. Parker ZF, Iyer SS, Wilen CB, Parrish NF, Chikere KC, Lee FH, et al. Transmitted/Founder and Chronic HIV-1 Envelope Proteins are Distinguished by Differential Utilization of CCR5. J Virol. 2012Dec.26.

96. Cicala C, Martinelli E, McNally JP, Goode DJ, Gopaul R, Hiatt J, et al. The

integrin alpha4beta7 forms a complex with cell-surface CD4 and defines a T-cell subset that is highly susceptible to infection by HIV-1. P Natl Acad Sci Usa. 2009Dec.8;106(49):20877–82.

97. Arthos J, Cicala C, Martinelli E, Macleod K, van Ryk D, Wei D, et al. HIV-1 envelope protein binds to and signals through integrin α4β7, the gut mucosal homing receptor for peripheral T cells. Nat Immunol. 2008Feb.10;9(3):301–9.

98. Richardson SI, Gray ES, Mkhize NN, Sheward DJ, Lambson BE, Wibmer CK, et al. South African HIV-1 subtype C transmitted variants with a specific V2 motif show higher dependence on α4β7 for replication. Retrovirology. BioMed Central; 2015Jun.23;:1–16.

99. Salazar-Gonzalez JF, Salazar MG, Keele BF, Learn GH, Giorgi EE, Li H, et al. Genetic identity, biological phenotype, and evolutionary pathways of transmitted/founder viruses in acute and early HIV-1 infection. Journal of Experimental Medicine. 2009Jun.8;206(6):1273–89.

100. Ochsenbauer C, Edmonds TG, Ding H, Keele BF, Decker J, Salazar MG, et al. Generation of Transmitted/Founder HIV-1 Infectious Molecular Clones and Characterization of Their Replication Capacity in CD4 T Lymphocytes and Monocyte-Derived Macrophages. J Virol. 2012Feb.11;86(5):2715–28.

101. Parrish NF, Wilen CB, Banks LB, Iyer SS, Pfaff JM, Salazar-Gonzalez JF, et al. Transmitted/Founder and Chronic Subtype C HIV-1 Use CD4 and

CCR5 Receptors with Equal Efficiency and Are Not Inhibited by Blocking the Integrin α4β7. Trkola A, editor. PLoS Pathog. 2012May31;8(5):e1002686.

102. Baalwa J, Wang S, Parrish NF, Decker JM, Keele BF, Learn GH, et al. Molecular identification, cloning and characterization of transmitted/founder HIV-1 subtype A, D and A/D infectious molecular clones. Virology. 2012Nov..

103. Parrish NF, Gao F, Li H, Giorgi EE, Barbian HJ, Parrish EH, et al. Phenotypic properties of transmitted founder HIV-1. Proceedings of the National Academy of Sciences. 2013Apr.23;110(17):6626–33.

104. Fenton-May AE, Dibben O, Emmerich T, Ding H, Pfafferott K, Aasa-Chapman MM, et al. Relative resistance of HIV-1 founder viruses to control by interferon-alpha. Retrovirology. 2013;10:146.

105. Deymier MJ, Claiborne DT, Ende Z, Ratner HK, Kilembe W, Allen S, et al. Particle infectivity of HIV-1 full-length genome infectious molecular clones in a subtype C heterosexual transmission pair following high fidelity amplification and unbiased cloning. Virology. 2014Sep.19;468-470C:454–61.

106. Deymier MJ, Ende Z, Fenton-May AE, Dilernia DA, Kilembe W, Allen SA, et al. Heterosexual Transmission of Subtype C HIV-1 Selects Consensus-Like Variants without Increased Replicative Capacity or Interferon-α Resistance. Malim MH, editor. PLoS Pathog. 2015Sep.17;11(9):e1005154.

107. Mayr E. Systematics and the Origin of Species. Columbia University Press; 1942.

108. Frange P, Meyer L, Jung M, Goujard C, Zucman D, Abel S, et al. Sexually-Transmitted/Founder HIV-1 Cannot Be Directly Predicted from Plasma or PBMC-Derived Viral Quasispecies in the Transmitting Partner. Mammano F, editor. Plos One. 2013Jul.9;8(7):e69144.

109. Ma ZM, Stone M, Piatak M, Schweighardt B, Haigwood NL, Montefiori D, et al. High Specific Infectivity of Plasma Virus from the Pre-Ramp-Up and Ramp-Up Stages of Acute Simian Immunodeficiency Virus Infection. J Virol. 2009Mar.10;83(7):3288–97.

110. Margolis L, Shattock R. Selective transmission of CCR5-utilizing HIV-1: the “gatekeeper” problem resolved? Nat Rev Microbiol. 2006Apr.;4(4):312–7.

111. Peeters M, Vincent R, Perret JL, Lasky M. Evidence for differences in MT2 cell tropism according to genetic subtypes of HIV-1: syncytium-inducing variants seem rare among subtype C HIV-1 viruses. JAIDS Journal of …. 1999.

112. Prince JL, Claiborne DT, Carlson JM, Schaefer M, Yu T, Lahki S, et al. Role of Transmitted Gag CTL Polymorphisms in Defining Replicative Capacity and Early HIV-1 Pathogenesis. Lifson J, editor. PLoS Pathog. 2012Nov.29;8(11):e1003041.

113. Wright JK, Brumme ZL, Carlson JM, Heckerman D, Kadie CM, Brumme CJ, et al. Gag-protease-mediated replication capacity in HIV-1 subtype C chronic infection: associations with HLA type and clinical parameters. J Virol. 2010Oct.;84(20):10820–31.

114. Pena-Cruz V, Etemad B, Chatziandreou N, Nyein PH, Stock S, Reynolds SJ, et al. HIV-1 envelope replication and α4β7 utilization among newly infected subjects and their corresponding heterosexual partners. Retrovirology. 2013;10:162.

115. Rouzine IM, Weinberger AD, Weinberger LS. An Evolutionary Role for HIV Latency in Enhancing Viral Transmission. Cell. Elsevier Inc; 2015Feb.26;160(5):1002–12.

116. Razooky BS, Pai A, Aull K, Rouzine IM, Weinberger LS. A Hardwired HIV Latency Program. Cell. Elsevier Inc; 2015Feb.26;160(5):990–1001.

117. Lauring AS, Andino R. Quasispecies Theory and the Behavior of RNA Viruses. Manchester M, editor. PLoS Pathog. 2010Jul.22;6(7):e1001005.

118. Domingo E, Sheldon J, Perales C. Viral Quasispecies Evolution. Microbiology and Molecular Biology Reviews. 2012Jun.11;76(2):159–216.

119. Wilke CO. Quasispecies theory in the context of population genetics. BMC Evol. Biol. 2005;5:44.

120. Eigen M, McCaskill J, Schuster P. Molecular quasi-species. The Journal of Physical …. 1988.

121. Codoñer FM, Darós J-A, Solé RV, Elena SF. The Fittest versus the Flattest: Experimental Confirmation of the Quasispecies Effect with Subviral Pathogens. PLoS Pathog. 2006;2(12):e136.

122. Enose Y, Ibuki K, Tamaru K, Ui M, Kuwata T, Shimada T, et al. Replication capacity of simian immunodeficiency virus in cultured macaque macrophages and dendritic cells is not prerequisite for intravaginal transmission of the virus in macaques. J Gen Virol. 1999Apr.;80 (Pt 4):847–55.

123. Roederer M, Keele BF, Schmidt SD, Mason RD, Welles HC, Fischer W, et al. Immunological and virological mechanisms of vaccine-mediated protection against SIV and HIV. Nature. 2013Dec.18.

124. Barouch DH, Klasse PJ, Dufour J, Veazey RS, Moore JP. Macaque studies of vaccine and microbicide combinations for preventing HIV-1 sexual transmission. Proceedings of the National Academy of Sciences. 2012May29;109(22):8694–8.

125. Hessell AJ, Rakasz EG, Poignard P, Hangartner L, Landucci G, Forthal DN, et al. Broadly Neutralizing Human Anti-HIV Antibody 2G12 Is Effective in Protection against Mucosal SHIV Challenge Even at Low Serum Neutralizing Titers. Farzan M, editor. PLoS Pathog. 2009May15;5(5):e1000433.

126. Hessell AJ, Rakasz EG, Tehrani DM, Huber M, Weisgrau KL, Landucci G, et al. Broadly Neutralizing Monoclonal Antibodies 2F5 and 4E10 Directed

against the Human Immunodeficiency Virus Type 1 gp41 Membrane-Proximal External Region Protect against Mucosal Challenge by Simian-Human Immunodeficiency Virus SHIVBa-L. J Virol. 2010Jan.11;84(3):1302–13.

127. Hessell AJ, Poignard P, Hunter M, Hangartner L, Tehrani DM, Bleeker WK, et al. Effective, low-titer antibody protection against low-dose repeated mucosal SHIV challenge in macaques. Nat Med. Nature Publishing Group; 2009Jun.7;:1–5.

128. Patterson BK, Landay A, Siegel JN, Flener Z, Pessis D, Chaviano A, et al. Susceptibility to human immunodeficiency virus-1 infection of human foreskin and cervical tissue grown in explant culture. Am. J. Pathol. 2002Sep.;161(3):867–73.

129. Maher D, Wu X, Schacker T, Horbul J, Southern P. HIV binding, penetration, and primary infection in human cervicovaginal tissue. P Natl Acad Sci Usa. 2005Aug.9;102(32):11504–9.

130. Hladik F, Sakchalathorn P, Ballweber L, Lentz G, Fialkow M, Eschenbach D, et al. Initial events in establishing vaginal entry and infection by human immunodeficiency virus type-1. Immunity. 2007Feb.1;26(2):257–70.

131. Berges BK, Akkina SR, Folkvord JM, Connick E, Akkina R. Mucosal transmission of R5 and X4 tropic HIV-1 via vaginal and rectal routes in humanized Rag2-/- gammac -/- (RAG-hu) mice. Virology.

2008Apr.10;373(2):342–51.

132. Denton PW, Garcia JV. Humanized mouse models of HIV infection. AIDS Rev. 2011Jul.;13(3):135–48.

133. Denton PW, Estes JD, Sun Z, Othieno FA, Wei BL, Wege AK, et al. Antiretroviral pre-exposure prophylaxis prevents vaginal transmission of HIV-1 in humanized BLT mice. Plos Med. 2008Jan.15;5(1):e16.

134. Denton PW, Olesen R, Choudhary SK, Archin NM, Wahl A, Swanson MD, et al. Generation of HIV Latency in Humanized BLT Mice. J Virol. 2011Dec.9;86(1):630–4.

135. Dudek TE, No DC, Seung E, Vrbanac VD, Fadda L, Bhoumik P, et al. Rapid evolution of HIV-1 to functional $CD8^+$ T cell responses in humanized BLT mice. Sci Transl Med. 2012Jul.18;4(143):143ra98.

136. O'Connell RJ, Kim JH, Corey L, Michael NL. Human Immunodeficiency Virus Vaccine Trials. Cold Spring Harb Perspect Med. 2012Dec.3;2(12):a007351–1.