Published in final edited form as:

Curr Opin HIV AIDS. 2010 July ; 5(4): 327–334. doi:10.1097/COH.0b013e32833a0b9b.

Identifying and Characterizing Recently Transmitted Viruses

Brandon F. Keele

AIDS and Cancer Virus Program, SAIC-Frederick, Inc., National Cancer Institute–Frederick, Frederick, Maryland, USA

Abstract

Purpose of review—Improvements in sequencing approaches and robust mathematical modeling have dramatically increased information on viral genetics during acute infection with human immunodeficiency virus (HIV) and simian immunodeficiency virus (SIV) infection, providing unprecedented insight into viral transmission and viral/immune Interactions.

Recent findings—Overall viral genetic diversity is reduced significantly during mucosal transmission. Remarkably, in the vast majority of sexual transmissions this diversity is reduced to a single viral variant that establishes the initial productive clinical infection. By identifying and enumerating transmitted/founder viruses researchers can begin to define the characteristics that are necessary and sufficient for successful viral replication within a new host.

Summary—Acute HIV infection is a critical window of opportunity for vaccine and therapeutic intervention. New sequencing technologies and mathematical modeling of transmission and early evolution have provided a clearer understanding of the number of founder viruses that establish infection, the rapid generation of diversity in these viruses and the subsequent evasion of host immunity. The information gained by identifying transmitted viruses, monitoring the initial host responses to these viruses, and then identifying mechanisms of viral escape could provide better strategies for vaccine development, pre-exposure prophylaxis, microbicides or other therapeutic interventions.

Keywords

HIV; SIV; viral evolution; acute infection; transmitted virus

Introduction

Understanding HIV transmission and the possibility of blocking transmission and preventing infection has been the focus of intense research since HIV was discovered. Current global estimates of HIV infection reveal that ~90% of infections occur across a mucosal surface [1]. These epidemiological data indicate that different modes of exposure associated with different risk behaviors impact the probability of infection. For example, anal receptive intercourse is statistically more likely to result in infection than vaginal intercourse [2,3]. Furthermore, the high viral load associated with primary infection in a donor can contribute to an increased probability for transmission [4]. While there is a great deal of epidemiological information surrounding viral transmission, a molecular understanding of viral transmission is still emerging. During HIV transmission, viral diversity is often severely reduced, indicating a genetic bottleneck where a significant number of variants found in the donor are lost at transmission. Until recently, the extent of this bottleneck was

only vaguely quantified. A desire to understand this restriction in viral diversity provided the motivation for more quantitative and detailed analyses of the virus(es) responsible for transmission. Here I review current literature surrounding viral transmission and early evolution including both human and animal model data.

Methods of identifying and enumerating the number of transmitted/founder viruses

For many years sequencing of viral quasispecies has been performed by either bulk PCR amplification of plasma or infected cells. An important new sequencing method termed single genome amplification (SGA), has improved viral sequencing and allowed for an unprecedented view into HIV and SIV transmission and early viral evolution. SGA is based on a limiting dilution algorithm in which proviral DNA or synthesized complementary DNA is diluted to a single amplifiable molecule per reaction prior to standard polymerase chain reaction (PCR) and direct sequencing. One can either experimentally determine the dilution where only ~20% of reactions are positive and consequentially where most reactions contain only one amplifiable template [5–7] or one can use real time PCR to predict the end-point dilution [8]. The advantages of SGA over standard sequencing methodologies are three-fold: i) direct proportionality of sequences with circulating variants, ii) no *Taq* induced point mutations, and iii) no *in vitro* recombination. Firstly, direct proportionality of sequences is attained using universal primers to reduce bias, in which each viral genome is equally likely to be amplified and sequenced. Therefore, if one generates 40 sequences, and 10 have a shared polymorphism, then one can predict that at the time of sampling ~25% of circulating viruses share this polymorphism. Secondly, there is a lack of *Taq*-induced error. Since amplification is from a single genome, when *Taq*-errors occur during the first or second round of PCR, there will be a polymorphism within the sequencing chromatograph at that site resulting in exclusion of that sequence. Additional *Taq*-induced mutations at later PCR cycles will be undetectable in the final sequence analysis due to low frequency. Furthermore, polymorphisms within the sequencing chromatograph will also reveal reactions that contained two distinct templates which also are excluded. Lastly, with only a single template per reaction, *in vitro* recombination cannot occur. These theoretical benefits of SGA were directly tested and confirmed experimentally by a number of investigators [5–7] thereby ensuring that the sequences being analyzed are identical to the sequences that exist *in vivo*.

Mathematical model

Until recently, the amount of diversity lost in transmission was characterized only in general terms. However, sampling acutely infected individuals, performing SGA, and utilizing a mathematical model of early viral evolution, the number of founding variants and the actual nucleotide sequence of each variant can be unambiguously identified. During primary infection, in the absence of immune selection, viable offspring of each founder genome are assumed to accumulate random mutations at a constant rate. When a single virus is transmitted, the consensus sequence of all obtained sequences represents the actual founder virus (Fig. 1a). Importantly, when two or more viruses establish infection in a host, the data fails to conform to a Poisson distribution. However, when each lineage is analyzed independently, the Poisson distribution again is maintained and the consensus of each founding strain can be identified (Fig. 1b). To better understand early evolution in the context of an identified transmitted virus, Lee et al. utilized this mathematical model and a Monte Carlo simulation to assess the overall diversity during primary infection [9]. This model accurately identified single or multiple viral infections using a cohort of human subjects and SIV infected macaques [6,9,10]. Additionally, this model is incredibly sensitive and can discriminate between two unique variants differing by as few as 3 nucleotides

within the ~ 2,600 nucleotide sequence encoding envelope [6]. Several recently published studies utilized this approach to determine the number of founder viruses in subjects infected with HIV-1 subtypes A, B, and C [5–7,11–14] and an additional report from Sagar et al. recently reported a similar reduction in viral diversity during transmission of subtypes D and A [14]. Remarkably consistent data have revealed that ~80% of heterosexual infections occur due to a single virus after sexual transmission. These findings argue that a severe reduction in viral diversity is a universal feature of mucosal HIV-1 transmission and provides a precise assessment of the well-established genetic bottleneck previously reported [5,12,15–26].

Animal model

Experimental SIV infection of rhesus macaques (*Macacca mulatta*) is a key animal model, central to studies of AIDS virus pathogenesis, and evaluation of candidate prevention approaches, including vaccines. This non-human primate model also provides a unique opportunity to study early events in viral transmission and evolution, both systemically and within important tissue compartments. However, the relevance of this model for many applications hinges on whether it accurately recapitulates the features of human mucosal transmission identified through the new approaches described above. Using SGA analysis the challenge inocula used for such studies can be genetically characterized and quantified prior to exposure and inoculated animals can be sampled (plasma and tissues) at defined times relative to a known time of exposure, including times earlier than is typically possible with humans. Two seminal papers highlight the usefulness of SGA sequencing and phylogenetic/mathematical analysis after rectal challenge with or without vaccine intervention [27,28]. As reported by Wilson et al., the inoculum to be used for intrarectal challenge was first titrated on naïve animals to identify a dose where most animals would become infected with only one or few variants. Vaccinated and control animals were then challenged with this dose and sequence analysis confirmed the correct dilution with few infecting variants, accurately recapitulating the outcome of typical human mucosal transmissions of HIV [28]. Using a DNA prime, recombinant Ad5-based vaccine approach with constructs expressing all SIV proteins except envelope, this group found a significant reduction in peak viremia after a heterologous, low-dose intrarectal challenge, with many vaccinated animals effectively controlling infection into the chronic phase [28]. Interestingly, while there were lower numbers of transmitted viruses in vaccinated animals, this difference was not statistically significant [28]. Utilizing a challenge inoculums titrated to reproduce the outcomes of typical human HIV transmissions, and monitoring the number of transmitted viruses, different vaccine protocols and formulations can now be evaluated by comparing the number of viral variants transmitted when infection occurs.

Transmission signatures

Due to the dramatic genetic bottleneck during viral transmission, understanding how one or a few viruses evolve during the earliest moments after transmission is important for therapeutic intervention strategies. A number of recent reports have tried to identify key features of transmitted HIV variants including coreceptor use, length and charge of variable loops, and overall accumulation of mutations after transmission. In rapidly evolving viruses like HIV, maintaining function while evading host selection is essential to survival. Poon et al. have recently suggested that an abundance of epistatic interactions, where identical functions can be performed by highly divergent sequences, is a salient feature of HIV [29]. Utilizing the Nef gene from 686 patients, these authors identified two pairs of interactions that were linked together in every instance, indicating convergent evolution within various hosts [29]. Furthermore, it has been reported that most transmitted/founder viruses utilize the C-C chemokine receptor 5 (CCR5) as a coreceptor [6] but Huang et al. have identified 6

patients out of 150 individuals recently infected whose virus utilizes either the CXCR4 chemokine receptor 4 (CXCR4) only or both CCR5 and CXCR4 [30]. It will be informative to determine the frequency of coreceptor usage at the moment of transmission. Other features of transmitted viruses recently identified by Sagar et al. include accumulated mutations immediately following the variable-3 (V3) loop, an overall shorter envelope, and a lower V3 charge in 13 individuals infected with subtype C and D [14]. However, there was no significant difference in the number of potential N-linked glycosylation sites which contradicts previous reports [15]. Recently, a number of viral Env and full length clones have been generated to begin testing the phenotype of the transmitted virus [6,7]. Utilizing these clones may provide insight into the first cells infected, replication capacities, or potential weaknesses. If clear patterns of genetic or phenotypic signatures can be identified, vaccine development can hone in on these potentially important transmission motifs for more effective vaccine design.

Early viral diversification

A number of mechanisms drive retroviral diversification including: i) reverse transcriptase errors, ii) host polymerase errors, iii) cytosine deamination, and iv) recombination. Apolipoprotein B mRNA-editing enzyme, catalytic polypeptide-like (APOBEC3) are innate viral restriction factors inducing deamination of cytosine to uracil on the negative strand of retroviral DNA during reverse transcription. This corresponds to a guanine to adenine (G-to-A) mutation in the newly generated plus strand proviral genome. The viral gene Vif counteracts APOBEC3 and prevents its incorporation into virions. Hypermutations are largely found within the context of GRD (where the G is mutated to A, the R is an A or G, and D is any nucleotide but cytosine)[31]. Recently a number of investigators have inquired if there is a relationship in the amount of hypermutation and disease progression. Pace et al. discovered that individuals with detectable hypermutation in peripheral blood mononuclear cells had a reduction in viral load [32]. Conversely, Ulena et al. recently found no association of hypermutation with viral load [33]. Land et al. also reported no differences in viral load but reported an increase in CD4 cells when hypermutation was detected [34]. Finally, Piantadosi et al. found no overall relationship between hypermutation and markers of disease progression [35]. Remarkably, one individual in this study retained a normal viral load despite infection with a virus with a partially defective Vif [35]. These studies all identified severely mutated sequences from integrated and putatively nonfunctional genomes. Recent studies suggest there is a moderate level of G-to-A mutation that does not render HIV completely non-functional but provides mutations that can often be utilized by the virus for rapid diversification. This low-level G-to-A mutation was first identified within acutely infected HIV-1 subjects [6,36]. By simply comparing the transmitted/founder sequence to all sequences during primary infection, minor G-to-A mutations within a GRD motif could be readily identified by their frequency within a single sequence which precluded a good fit to the Poisson model [6]. In addition to detecting low-level mutations within an individual, Wood et al. interrogated informative sites across multiple acutely infected hosts and found G-to-A substitutions despite having excluded heavily mutated sequences prior to analysis [36]. In several cases, a rapidly evolving site was embedded both in an APOBEC motif and in a CD8+ cytotoxic T lymphocyte (CTL) epitope, suggesting that APOBEC activity may actually facilitate early immune escape. Further evidence of low-level G-to-A mutations comes from Jern et al. who used both *in silico* modeling and cell culture experiments to identify non-lethal G-to-A mutations [37]. With putatively functional genomes the authors predict that these low-level G-to-A mutations are likely to survive selection pressure and accumulate overtime within the population. These predictions were validated by the detection of G-to-A imprints on current HIV compared to the ancestral genome [37]. Overall, hypermutation may not affect disease progression directly, but moderate mutations may be advantageous to the virus by rapidly accumulating genetic

diversity. Recombination can occur very early after infection but has been found typically after peak viremia when potential targets become more limited [6]. The effects of recombination are greatly increased in patients infected with more than one viral variant. This rapid increase in overall genetic diversity could lead to more efficient immune escape and increase progression to AIDS [38]. Onafuwa-Nuga and Telesnitsky recently reviewed recombination in detail [39], but it is notable that recombination is seen very early in primary infection in humans [6] and in non-human primates infected with SIV (Keele unpublished). Finally, studies comparing HIV transmission to other sexually transmitted diseases and to airborne infections show that there is a cost to high diversity and that cost is a lower transmission rate [40]. Understanding transmission of particular variants and early diversity is fundamental to eventually inhibiting these events.

Early host responses to infection

There are a number of host barriers to infection that in total can explain the genetic bottleneck seen in HIV-1 transmission. Each mechanism of host defense represents a battle line at which, the host and virus compete for life itself. In a typical case if the host cannot eliminate the invader, within days to perhaps weeks after exposure, the battle is over and the fate of the host is sealed. For mucosal infections, the mucus itself and an intact epithelial barrier most likely represent the greatest obstacle to infection. However, even after this barrier is breached, infecting virus must still find an appropriate target cell, mediate entry via CD4 and coreceptor, successfully reverse transcribe and finally integrate its genome. Generating progeny from an integrated genome, at a basic reproductive ratio large enough to overcome innate immune responses, becomes the next challenge for viral survival. Recently, Stacey et al. [41] measured cytokine and chemokine levels in plasma within the first days of HIV-1 infection. After synchronizing each individual based on first detectable vRNA (~10–21 days post infection), they found a rapid increase in many cytokines and chemokines, including alpha interferon, interleukin-15 (IL-15), inducible protein 10, tumor necrosis factor alpha and monocyte chemoattractant protein 1, with more slowly initiated increases in IL-6, IL-8, IL-18 and gamma interferon. The magnitude of this “cytokine storm” was not observed in acute hepatitis B or C infection and it may negatively impact the host by promoting viral replication and fueling viral dissemination [41,42]. In exposed and uninfected individuals, perhaps a strong innate response helps protect against initial infection, but once infection occurs, this immune activation can promote target cell recruitment and increase viral replication [42]. Importantly, this cytokine storm is quenched during primary infection of non-pathogenic natural hosts. Studies by Jaccquelin et al. and Bosinger et al. have shown a rapid control of innate responses in both African green monkeys [43] and sooty mangabeys [44] suggesting that natural hosts actively suppress immune activation thereby protecting T cells from infection and death. Characterizing the earliest host responses, including innate immunity in humans, pathogenic models and nonpathogenic hosts, will be critical to understanding and potentially impeding infection.

Cellular immunity

The next significant challenge for transmitted viruses is the adaptive immune response and although recent evidence suggests that this activation occurs even within the first weeks of infection [45], it appears to be too little, and too late for successful viral elimination. From exposure to peak viremia, approximately 14–28 days later, the virus is expanding exponentially and generating millions of progeny virus. While it is true that in acute infection any given molecule of vRNA is limited in its overall diversity, which allows one to identify the transmitted virus, the viral population as a whole contains numerous mutations. It is this enormous potential for diversity that allows the virus to rapidly evade elimination by cellular immunity [45,46]. Utilizing the inferred transmitted/founder genome identified in

four acutely infected individuals [46]. Goonetilleke et al. identified and functionally confirmed a number of very early HIV specific CTL responses [45]. These responses supplied sufficiently strong selective pressure that within weeks after the onset of detectable viremia the entire viral population was replaced by escape mutants. Importantly, most mutations were not just a simple single amino acid replacement; often various mutations would initially occur and through progressive selection eventually a final solution could be identified [45]. This phenomenon of high variability within CTL epitopes was recently further elucidated by Bimber et al., who used pyrosequencing to identify and characterize early SIV specific mutations in rhesus macaques [47]. Importantly, mathematical modeling of these early CTL responses suggest that 15–35% of infected cells per day can be killed by specific T cells, which decreases viremia and would ultimately be successful in clearing virus if not for rapid escape mutations [45]. Conversely, two recent papers challenge the idea that CTLs directly alter the death rate of HIV infected cells [48,49]. Both groups used SIV infected rhesus macaques and CD8 depletion in combination with antiviral therapy to show that viral decay was indistinguishable in either group suggesting that CTLs did not increase the death rate of infected T cells [48,49]. Although additional experiments are required to validate and clarify the role of CTLs in retroviral infection, recent work by Louis Picker and colleagues suggests that a T cell based vaccine can provide effective protection in rhesus macaques [50]. Here immunization with a rhesus cytomegalovirus vector expressing SIV Gag, Rev-Tat-Nef and Env induced effector memory T cell responses at mucosal sites and provided increased protection against acquisition after repeat, low-dose intrarectal challenge [50]. Therefore it is clear that CTLs have enormous potential to effect viral kinetics and can even protect against infection if optimally activated.

Reversion

In addition to CTL escape, mutations can also accumulate due to reversion of escape mutations after transmission to a new host with different major histocompatibility complex I (MHC-I) alleles. Recently, Novitsky et al. utilized SGA from a cohort of 42 individuals recently infected with HIV-1 subtype C to identify gag mutations that occur during primary infection [13]. Although there were complex mutational pathways involving both reversion and immune escape, overall, the authors identified reversion to the wild-type occurred significantly earlier than immune induced selection [13]. In a second subtype C study, Treurnicht et al. bulk-sequenced full genomes of 20 individuals sampled during primary infection and 6 months later and reported a 2:1 ratio of reversion to wild type over immune escape [51]. Wang et al. also saw evidence of reversion, but examined only chronically infected subtype B subjects [52]. Interestingly, in two studies of acute subtype B infection, reversion to wild-type was reportedly very slow and infrequently occurred during primary infection [5,45,53]. Therefore, while it is certainly possible for reversion to occur rapidly after transmission to a new host with divergent MHC-1 alleles, the timing and frequency of reversion will largely depend on the relative fitness cost of each mutation. Two studies last year reported divergent fitness in various escape mutations compared to the consensus or wild-type sequence [54,55]. While Troyer et al. [55] reported no fitness costs for Env mutations they identified a large fitness cost within a p24 epitope which corroborates earlier studies involving Gag sequences [56–58]. Christie et al. however failed to identify a fitness cost in the Gag SL9 epitope [54]. Escape and reversion have recently been well reviewed by Davenport et al. [59]. Understanding the patterns of reversion and escape could prove useful for rational vaccine design and should be rigorously pursued.

Humoral immunity

Finally, after weeks to months of infection, antibodies and neutralizing antibodies (Nab) are detectable and can exert humoral immune selection pressure on established virus

populations. Typically, the first detectable antibodies are specific for gp41 and appear 2–3 weeks following infection [60]. After an additional 2 weeks, non-neutralizing gp160 reactive antibodies appear which can however mediate antibody-dependent cell-mediated cytotoxicity. But the earliest antibody responses that cause detectable genetic selection do not occur until ~12 weeks post infection. These first Nabs are typically narrow in scope often only reactive against a single epitope (e.g. variable loops). The virus has little difficulty in generating escape mutations to these antibodies due to their delay, the narrowness of the response and the already pre-existing diverse viral population [61–63]. However, since passive Nab can provide sterilizing protection in NHP model [64], preexisting Nab could prove valuable when the virus is most vulnerable during the first days after exposure when one or few variants with limited genetic diversity are struggling to survive. In fact, activation of the humoral immune response was most likely the successful element of the RV144 trial however, additional experiments are needed to understand the immunologic correlates of protection [65].

Clinical significance

Identifying and enumerating the number of infecting viruses can have important ramifications for clinical treatment regimens, pre-exposure prophylaxis (PrEP), and vaccine design. When deciding when a patient should be treated with anti-retroviral therapy, physicians should consider that very early after infection, viral diversity is limited and less likely to generate resistance than later in infection. Furthermore, if PrEP is successful and readily available, there is a risk that when infection does occur it will be with resistant viruses. One important target for early treatment or PrEP is CCR5. While there are clearly documented cases of CXCR4 usage, the vast majority of all identified transmitted viruses rely on CCR5 and successfully blocking this interaction could reduce the number of new infections. While resistance to both CCR5 inhibitors (or any other antiretroviral drug) is possible, the requirement for using CCR5 in most transmissions makes it an excellent candidate. If resistant strains are present but at a low proportion, it is possible that all minor, resistant strains are eliminated by the physiological barriers prior to infection. In other words, since most individuals are infected with a single variant and the overall likelihood of infection is low, blocking even a portion of potentially infecting variants should reduce overall infection rates.

Conclusion

New and exciting data continue to help us understand viral transmission and early virus/host interactions. The field now has a quantitative foundation regarding the number of viruses responsible for productive clinical infection, the conditions that may increase this number, the phenotype of transmitted viruses, the mechanisms driving viral diversification, the host factors trying to impede viral progression, and the subsequent selection of viral escape mutants. Incorporating these emerging concepts into approaches for blocking HIV infection should help in developing new and more effective vaccine and other therapeutic strategies. In my opinion, a sterilizing vaccine that protects all individuals might not be achievable, however, the severe genetic bottleneck at transmission and limited early evolution offers hope that a prophylactic vaccine can be developed to significantly reduce new infections.

Acknowledgments

The author is grateful to Jacob Estes and Jeff Lifson for many valuable suggestions to this manuscript. The author was supported with federal funds from the National Cancer Institute, National Institutes of Health under contract HHSN266200400088C.

References and recommended reading

* Here the authors describe rapid viral escape to neutralizing antibodies. A number of escape mutations were ge

1. UNAIDS. Report on the global AIDS epidemic. Geneva: UNAIDS; 2008.
2. Powers KA, Poole C, Pettifor AE, et al. Rethinking the heterosexual infectivity of HIV-1: a systematic review and meta-analysis. *Lancet Infect Dis* 2008;8:553–563. [PubMed: 18684670]
3. Boily MC, Baggaley RF, Wang L, et al. Heterosexual risk of HIV-1 infection per sexual act: systematic review and meta-analysis of observational studies. *Lancet Infect Dis* 2009;9:118–129. [PubMed: 19179227]
4. Cohen MS. Preventing sexual transmission of HIV. *Clin Infect Dis* 2007;45 (Suppl 4):S287–292. [PubMed: 18190301]
5. Kearney M, Maldarelli F, Shao W, et al. Human immunodeficiency virus type 1 population genetics and adaptation in newly infected individuals. *J Virol* 2009;83:2715–2727. [PubMed: 19116249]
6. Keele BF, Giorgi EE, Salazar-Gonzalez JF, et al. Identification and characterization of transmitted and early founder virus envelopes in primary HIV-1 infection. *Proc Natl Acad Sci U S A* 2008;105:7552–7557. [PubMed: 18490657]
7. Salazar-Gonzalez JF, Bailes E, Pham KT, et al. Deciphering human immunodeficiency virus type 1 transmission and early envelope diversification by single-genome amplification and sequencing. *J Virol* 2008;82:3952–3970. [PubMed: 18256145]
8. Butler DM, Pacold ME, Jordan PS, et al. The efficiency of single genome amplification and sequencing is improved by quantitation and use of a bioinformatics tool. *J Virol Methods* 2009;162:280–283. [PubMed: 19698751]
- 9**. Lee HY, Giorgi EE, Keele BF, et al. Modeling sequence evolution in acute HIV-1 infection. *J Theor Biol* 2009;261:341–360. The authors use sophisticated mathematical modeling and computer simulation to better understand viral transmission and early diversification. Sequence diversity at each viral generation is modeled to accurately determine a most recent common ancestor. When diversity is greater than expected the model confirms multiple variant infections. [PubMed: 19660475]
10. Bimber BN, Chugh P, Giorgi EE, et al. Nef gene evolution from a single transmitted strain in acute SIV infection. *Retrovirology* 2009;6:57. [PubMed: 19505314]
11. Abrahams MR, Anderson JA, Giorgi EE, 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;83:3556–3567. [PubMed: 19193811]
12. Haaland RE, Hawkins PA, Salazar-Gonzalez J, et al. Inflammatory genital infections mitigate a severe genetic bottleneck in heterosexual transmission of subtype A and C HIV-1. *PLoS Pathog* 2009;5:e1000274. [PubMed: 19165325]
- 13*. Novitsky V, Wang R, Margolin L, et al. Timing constraints of in vivo gag mutations during primary HIV-1 subtype C infection. *PLoS One* 2009;4:e7727. Here the authors used SGA and direct sequencing of 42 HIV-1 infected subjects to identify the timing and robustness of viral reversion. Reversion occurred significantly earlier than CTL escape mutations. [PubMed: 19890401]
14. Sagar M, Laeyendecker O, Lee S, et al. Selection of HIV variants with signature genotypic characteristics during heterosexual transmission. *J Infect Dis* 2009;199:580–589. [PubMed: 19143562]
15. Derdeyn CA, Decker JM, Bibollet-Ruche F, et al. Envelope-constrained neutralization-sensitive HIV-1 after heterosexual transmission. *Science* 2004;303:2019–2022. [PubMed: 15044802]
16. Gottlieb GS, Heath L, Nickle DC, et al. HIV-1 variation before seroconversion in men who have sex with men: analysis of acute/early HIV infection in the multicenter AIDS cohort study. *J Infect Dis* 2008;197:1011–1015. [PubMed: 18419538]
17. Greenier JL, Miller CJ, Lu D, et al. Route of simian immunodeficiency virus inoculation determines the complexity but not the identity of viral variant populations that infect rhesus macaques. *J Virol* 2001;75:3753–3765. [PubMed: 11264364]

18. Long EM, Martin HL Jr, Kreiss JK, et al. Gender differences in HIV-1 diversity at time of infection. *Nat Med* 2000;6:71–75. [PubMed: 10613827]
19. Margolis L, Shattock R. Selective transmission of CCR5-utilizing HIV-1: the 'gatekeeper' problem resolved? *Nat Rev Microbiol* 2006;4:312–317. [PubMed: 16541138]
20. Miller CJ, Li Q, Abel K, et al. Propagation and dissemination of infection after vaginal transmission of simian immunodeficiency virus. *J Virol* 2005;79:9217–9227. [PubMed: 15994816]
21. Poss M, Martin HL, Kreiss JK, et al. Diversity in virus populations from genital secretions and peripheral blood from women recently infected with human immunodeficiency virus type 1. *J Virol* 1995;69:8118–8122. [PubMed: 7494333]
22. Ritola K, Pilcher CD, Fiscus SA, et al. Multiple V1/V2 env variants are frequently present during primary infection with human immunodeficiency virus type 1. *J Virol* 2004;78:11208–11218. [PubMed: 15452240]
23. Sagar M, Kirkegaard E, Long EM, et al. Human immunodeficiency virus type 1 (HIV-1) diversity at time of infection is not restricted to certain risk groups or specific HIV-1 subtypes. *J Virol* 2004;78:7279–7283. [PubMed: 15194805]
24. Wolfs TF, Zwart G, Bakker M, et al. HIV-1 genomic RNA diversification following sexual and parenteral virus transmission. *Virology* 1992;189:103–110. [PubMed: 1376536]
25. Wolinsky SM, Wike CM, Korber BT, et al. Selective transmission of human immunodeficiency virus type-1 variants from mothers to infants. *Science* 1992;255:1134–1137. [PubMed: 1546316]
26. Zhu T, Mo H, Wang N, et al. Genotypic and phenotypic characterization of HIV-1 patients with primary infection. *Science* 1993;261:1179–1181. [PubMed: 8356453]
- 27*. Keele BF, Li H, Learn GH, et al. Low-dose rectal inoculation of rhesus macaques by SIVsmE660 or SIVmac251 recapitulates human mucosal infection by HIV-1. *J Exp Med* 2009;206:1117–1134. This article describes the first use of single genome amplification in SIV infected rhesus macaques. After low-dose rectal infection with SIVmac251 or SIVsmE660, infected animals showed one or few variants that established systemic infection. In some instances, an inoculum sequence was also identified as the transmitted/founder virus thus confirming the overall approach to identifying transmitted variants. [PubMed: 19414559]
- 28*. Wilson NA, Keele BF, Reed JS, et al. Vaccine-induced cellular responses control simian immunodeficiency virus replication after heterologous challenge. *J Virol* 2009;83:6508–6521. Here the authors utilize SGA to test a vaccine candidate in rhesus macaques. Titration of inocula was performed to identify a dose that most closely represents HIV-1 infection. After intrarectal infection, vaccinated animals had a significantly lower viral peak and set point compared to control animals. [PubMed: 19403685]
29. Poon AF, Swenson LC, Dong WW, et al. Phylogenetic Analysis of Population-Based and Deep Sequencing Data to Identify Coevolving Sites in the nef Gene of HIV-1. *Mol Biol Evol.* 2009
30. Huang W, Toma J, Stawiski E, et al. Characterization of human immunodeficiency virus type 1 populations containing CXCR4-using variants from recently infected individuals. *AIDS Res Hum Retroviruses* 2009;25:795–802. [PubMed: 19678765]
31. Rose PP, Korber BT. Detecting hypermutations in viral sequences with an emphasis on G → A hypermutation. *Bioinformatics* 2000;16:400–401. [PubMed: 10869039]
32. Pace C, Keller J, Nolan D, et al. Population level analysis of human immunodeficiency virus type 1 hypermutation and its relationship with APOBEC3G and vif genetic variation. *J Virol* 2006;80:9259–9269. [PubMed: 16940537]
33. Ullenga NK, Sarr AD, Hamel D, et al. The level of APOBEC3G (hA3G)-related G-to-A mutations does not correlate with viral load in HIV type 1-infected individuals. *AIDS Res Hum Retroviruses* 2008;24:1285–1290. [PubMed: 18851679]
34. Land AM, Ball TB, Luo M, et al. Human immunodeficiency virus (HIV) type 1 proviral hypermutation correlates with CD4 count in HIV-infected women from Kenya. *J Virol* 2008;82:8172–8182. [PubMed: 18550667]
35. Piantadosi A, Humes D, Chohan B, et al. Analysis of the percentage of human immunodeficiency virus type 1 sequences that are hypermutated and markers of disease progression in a longitudinal

- cohort, including one individual with a partially defective Vif. *J Virol* 2009;83:7805–7814. [PubMed: 19494014]
36. Wood N, Bhattacharya T, Keele BF, et al. HIV evolution in early infection: selection pressures, patterns of insertion and deletion, and the impact of APOBEC. *PLoS Pathog* 2009;5:e1000414. [PubMed: 19424423]
 - 37*. Jern P, Russell RA, Pathak VK, et al. Likely role of APOBEC3G-mediated G-to-A mutations in HIV-1 evolution and drug resistance. *PLoS Pathog* 2009;5:e1000367. This article describes the effects of low-level G-to-A mutations on HIV-1 evolution overtime. Using in vitro experimentation, simulations and phylogenetic inferences, the authors conclude that non-lethal, moderate G-to-A mutation is beneficial to HIV and the genetic imprint of such mutations can be found by comparing ancestral and contemporaneous sequences. [PubMed: 19343218]
 38. Sagar M, Lavreys L, Baeten JM, et al. Infection with multiple human immunodeficiency virus type 1 variants is associated with faster disease progression. *J Virol* 2003;77:12921–12926. [PubMed: 14610215]
 39. Onafuwa-Nuga A, Telesnitsky A. The remarkable frequency of human immunodeficiency virus type 1 genetic recombination. *Microbiol Mol Biol Rev* 2009;73:451–480. Table of Contents. [PubMed: 19721086]
 40. Lange A, Ferguson NM. Antigenic diversity, transmission mechanisms, and the evolution of pathogens. *PLoS Comput Biol* 2009;5:e1000536. [PubMed: 19847288]
 - 41*. Stacey AR, Norris PJ, Qin L, et al. Induction of a striking systemic cytokine cascade prior to peak viremia in acute human immunodeficiency virus type 1 infection, in contrast to more modest and delayed responses in acute hepatitis B and C virus infections. *J Virol* 2009;83:3719–3733. This article reports on the cytokine profiles during acute HIV-1 infection. Patients were serially sampled just prior to and following the eclipse phase of infection. The authors report an abundance of cytokines and immune activation markers. [PubMed: 19176632]
 - 42*. Li Q, Estes JD, Schlievert PM, et al. Glycerol monolaurate prevents mucosal SIV transmission. *Nature* 2009;458:1034–1038. Here the earliest events surrounding SIV vaginal infection are elucidated using immunohistochemistry and in situ hybridization. Early viral replication occurs locally and only after recruitment of activated T cells does the virus spread to lymph nodes and systemically. The authors also identified glycerol monolaurate as an inhibitor of vaginal infection. [PubMed: 19262509]
 - 43*. Jacquelin B, Mayau V, Targat B, et al. Nonpathogenic SIV infection of African green monkeys induces a strong but rapidly controlled type I IFN response. *J Clin Invest* 2009;119:3544–3555. One of two back-to-back publications elucidating early innate responses in nonpathogenic hosts. Here the authors utilize African green monkeys to show that robust innate responses are rapidly dampened in this model suggesting a reason for increased pathogenesis in humans. [PubMed: 19959873]
 - 44*. Bosinger SE, Li Q, Gordon SN, et al. Global genomic analysis reveals rapid control of a robust innate response in SIV-infected sooty mangabeys. *J Clin Invest* 2009;119:3556–3572. The second of two back-to-back publications elucidating early innate responses in nonpathogenic hosts. Here the authors utilize sooty mangabeys to show that robust innate responses are also rapidly dampened in this model. [PubMed: 19959874]
 - 45*. Goonetilleke N, Liu MK, Salazar-Gonzalez JF, et al. The first T cell response to transmitted/founder virus contributes to the control of acute viremia in HIV-1 infection. *J Exp Med* 2009;206:1253–1272. This article utilizes SGA derived full genome sequencing to identify and quantify the first T cell responses to the transmitted virus. Patient specific peptides were utilized to generate profiles of CTL reactivity. Primary CTL responses were concurrent with reduction in viral loads and earlier than previously recognized. After the initial immune escape, these initial immune responses rapidly waned but were replaced by CTL responses to epitopes where escape was slower and at a higher fitness cost to the virus. [PubMed: 19487423]
 - 46*. Salazar-Gonzalez JF, Salazar MG, Keele BF, et al. Genetic identity, biological phenotype, and evolutionary pathways of transmitted/founder viruses in acute and early HIV-1 infection. *J Exp Med* 2009;206:1273–1289. Here the author utilize SGA for the first time to generate full length genome sequences. After unambiguously identifying the entire transmitted/founder genome,

- longitudinal sequencing revealed numerous sites of selection. These selected mutations resulted directly from early immune responses and possibly reversion. [PubMed: 19487424]
- 47*. Bimber BN, Burwitz BJ, O'Connor S, et al. Ultradeep pyrosequencing detects complex patterns of CD8+ T-lymphocyte escape in simian immunodeficiency virus-infected macaques. *J Virol* 2009 83:8247–8253. This paper utilizes the 454 pyrosequencing technique to identify patterns of immune escape during acute SIV infection of rhesus macaques. The authors found that using bulk sequencing techniques did not provide sufficient depth to understand the very complex patterns of some CTL escape mutations.
 48. Klatt NR, Shudo E, Ortiz AM, et al. CD8+ lymphocytes control viral replication in SIVmac239-infected rhesus macaques without decreasing the lifespan of productively infected cells. *PLoS Pathog* 2010;6:e1000747. [PubMed: 20126441]
 49. Wong JK, Strain MC, Porra R, et al. In vivo CD8+ T-cell suppression of siv viremia is not mediated by CTL clearance of productively infected cells. *PLoS Pathog* 2010;6:e1000748. [PubMed: 20126442]
 50. Hansen SG, Vieville C, Whizin N, et al. Effector memory T cell responses are associated with protection of rhesus monkeys from mucosal simian immunodeficiency virus challenge. *Nat Med* 2009;15:293–299. [PubMed: 19219024]
 51. Treurnicht FK, Seoighe C, Martin DP, et al. Adaptive changes in HIV-1 subtype C proteins during early infection are driven by changes in HLA-associated immune pressure. *Virology* 2010;396:213–225. [PubMed: 19913270]
 52. Wang YE, Li B, Carlson JM, et al. Protective HLA class I alleles that restrict acute-phase CD8+ T-cell responses are associated with viral escape mutations located in highly conserved regions of human immunodeficiency virus type 1. *J Virol* 2009;83:1845–1855. [PubMed: 19036810]
 53. Brumme ZL, John M, Carlson JM, et al. HLA-associated immune escape pathways in HIV-1 subtype B Gag, Pol and Nef proteins. *PLoS One* 2009;4:e6687. [PubMed: 19690614]
 54. Christie NM, Willer DO, Lobritz MA, et al. Viral fitness implications of variation within an immunodominant CD8+ T-cell epitope of HIV-1. *Virology* 2009;388:137–146. [PubMed: 19368950]
 55. Troyer RM, McNevin J, Liu Y, et al. Variable fitness impact of HIV-1 escape mutations to cytotoxic T lymphocyte (CTL) response. *PLoS Pathog* 2009;5:e1000365. [PubMed: 19343217]
 56. Brockman MA, Schneidewind A, Lahaie M, et al. Escape and compensation from early HLA-B57-mediated cytotoxic T-lymphocyte pressure on human immunodeficiency virus type 1 Gag alter capsid interactions with cyclophilin A. *J Virol* 2007;81:12608–12618. [PubMed: 17728232]
 57. Crawford H, Prado JG, Leslie A, et al. Compensatory mutation partially restores fitness and delays reversion of escape mutation within the immunodominant HLA-B*5703-restricted Gag epitope in chronic human immunodeficiency virus type 1 infection. *J Virol* 2007;81:8346–8351. [PubMed: 17507468]
 58. Schneidewind A, Brockman MA, Yang R, et al. Escape from the dominant HLA-B27-restricted cytotoxic T-lymphocyte response in Gag is associated with a dramatic reduction in human immunodeficiency virus type 1 replication. *J Virol* 2007;81:12382–12393. [PubMed: 17804494]
 59. Davenport MP, Loh L, Petravic J, et al. Rates of HIV immune escape and reversion: implications for vaccination. *Trends Microbiol* 2008;16:561–566. [PubMed: 18964018]
 60. Tomaras GD, Yates NL, Liu P, et al. Initial B-cell responses to transmitted human immunodeficiency virus type 1: virion-binding immunoglobulin M (IgM) and IgG antibodies followed by plasma anti-gp41 antibodies with ineffective control of initial viremia. *J Virol* 2008;82:12449–12463. [PubMed: 18842730]
 61. Davis KL, Gray ES, Moore PL, et al. High titer HIV-1 V3-specific antibodies with broad reactivity but low neutralizing potency in acute infection and following vaccination. *Virology* 2009;387:414–426. [PubMed: 19298995]
 62. Moore PL, Gray ES, Morris L. Specificity of the autologous neutralizing antibody response. *Curr Opin HIV AIDS* 2009;4:358–363. [PubMed: 20048698]
 - 63*. Rong R, Li B, Lynch RM, et al. Escape from autologous neutralizing antibodies in acute/early subtype C HIV-1 infection requires multiple pathways. *PLoS Pathog* 2009;5:e1000594. This article describes rapid viral escape to neutralizing antibodies in acute subtype C infected

individuals. Utilizing a number of envelope clones, the authors identified mutations in variable regions as the principle cause of neutralizing antibody escape. [PubMed: 19763269]

64. Xu W, Hofmann-Lehmann R, McClure HM, et al. Passive immunization with human neutralizing monoclonal antibodies: correlates of protective immunity against HIV. *Vaccine* 2002;20:1956–1960. [PubMed: 11983253]
65. Rerks-Ngarm S, Pitisuttithum P, Nitayaphan S, et al. Vaccination with ALVAC and AIDSVAX to prevent HIV-1 infection in Thailand. *N Engl J Med* 2009;361:2209–2220. [PubMed: 19843557]

Figure 1. Identifying the transmitted virus using single genome amplification and mathematical modeling

Infection with a single variant is determined by generating sequences during acute infection and comparing all sequences to themselves. In panel A, of the 21 sequences shown in the tree (left) and in the highlighter alignment (center) over half (62%) are identical to themselves and identical to the consensus over the entire env gene. The remaining sequences differ from this consensus sequence by one (33%) or two (5%) mutations. The differences between all sequences are compared to a Poisson distribution and deviations are identified by a goodness of fit test. When two sequences are transmitted (B), variant 1 (black) can be distinguished from variant 2 (grey) and the consensus of each lineage represents a unique transmitted/founder viral sequences. When all sequences from both lineages are compared, the Poisson distribution does not fit, but is restored when each lineage is analyzed independently.