Phenotypic properties of transmitted founder HIV-1

Nicholas F. Parrish,a,b Feng Gao,c,d Hui Li,e Elena E. Giorgi,g Hannah J. Barbian,a,b Erica H. Parrish,a Lara Zajic,c Chip S. Iyer,a,b Julie M. Decker,f Amit Kumar,g BhavnaHora,g Anna Berg,f Fangping Cay,g Jennifer Hopper,g Thomas N. Denny,j,h Hai Tao Ding,g Christina Ochensebauer,g John C. Kappes,g Rachel P. Galimidi,g Anthony P. West, Jr.,g Pamela J. Bjorkman,b,d Craig B. Wilen,a,b Robert W. Doms,a,b Meagan O’Brien,g Nina Bhnardjaj,b, Persephone Borrows,a Barton F. Haynes,c,d, Mark Muldoon,g James P. Theiler,e Bette Korber,g George M. Shaw,c,g and Beatrice H. Hahnd,b

Departments of aMedicine and bMicrobiology, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA 19104; Duke Human Vaccine Institute and Departments of aMedicine and bImmunology, Duke University School of Medicine, Durham, NC 27710; Theoretical Biology and Biophysics, Theoretical Division, Los Alamos National Laboratory, Los Alamos, NM 87545; bDepartment of Medicine, University of Alabama, Birmingham, AL 35294; bDivision of Biology and Howard Hughes Medical Institute, California Institute of Technology, Pasadena, CA 91125; Department of Pathology and Laboratory Medicine, Children’s Hospital of Philadelphia, Philadelphia, PA 19104; Departments of Medicine, Pathology, and Dermatology, New York University School of Medicine, New York, NY 10016; bNuffield Department of Clinical Medicine, University of Oxford, Weatherall Institute of Molecular Medicine, John Radcliffe Hospital, Headington, Oxford OX3 9DS, United Kingdom; and eSchool of Mathematics, University of Manchester, Manchester M13 9PL, United Kingdom

This contribution is part of the special series of Inaugural Articles by members of the National Academy of Sciences elected in 2012.

Contributed by Beatrice H. Hahn, March 5, 2013 (sent for review January 29, 2013)

Defining the virus-host interactions responsible for HIV-1 transmission, including the phenotypic requirements of viruses capable of establishing de novo infections, could be important for AIDS vaccine development. Previous analyses have failed to identify phenotypic properties other than chemokine receptor 5 (CCR5) and CD4+ T-cell tropism that are preferentially associated with viral transmission. However, most of these studies were limited to examining envelope (Env) function in the context of pseudoviruses. Here, we generated infectious molecular clones of transmitted founder (TF) and chronic control (CC) viruses. TF viruses were more efficiently captured by monocyte-derived dendritic cells 1.7-fold more efficiently compared with CC viruses. TF viruses were also captured by monocyte-derived dendritic cells 1.7-fold more efficiently compared with CC viruses. This difference was significant for subtype B (P = 0.00013) but not subtype C (P = 0.53) viruses, possibly reflecting demographic differences of the respective patient cohorts. Together, these data indicate that TF viruses are enriched for higher Env content, enhanced cell-free infectivity, improved dendritic cell interaction, and relative IFN-α resistance. These viral properties, which likely act in concert, should be considered in the development and testing of AIDS vaccines.

Understanding the host and viral factors that influence the ability of HIV-1 to cross mucosal barriers may be critical for the development of an effective AIDS vaccine (1). HIV-1 virions or infected cells are believed to cross the epithelium shortly after sexual exposure, although the mechanisms by which this transfer occurs remain largely unknown. Within the mucosa, viruses are believed to interact with dendritic cells (DCs), such as Langerhans cells, but do not productively infect these cells (2). In the simian model of HIV-1, the first cells to become productively infected are resting intraepithelial CD4+ T cells, which represent the most abundant target cell type in the lamina propria (3). Simultaneous with initial infection events, local innate immune responses are elicited, with plasmacytoid DCs accumulating rapidly at sites of virus entry. These cells secrete cytokines and chemokines, such as macrophage inflammatory protein (MIP)-1β and type 1 IFNs, and orchestrate an early local innate immune response (4). After infection in mucosal and submucosal tissues is established, HIV-1 spreads to regional and distant lymphoid tissues, including the gut-associated lymphoid tissue, where virus expands exponentially, triggering a systemic cytokine storm preceding peak viremia (5, 6).

Many host factors can influence whether virus exposure leads to productive infection, including the physical barrier of the mucosa and its associated mucous secretions (7, 8), target cell availability (9, 10), immune activation (11), genital inflammation (12), and altered mucosal microbiota (13). Although chronic HIV-1 infection is characterized by high genotypic and phenotypic diversity, the extent to which this variation influences the transmission process remains unclear (14, 15). Transmission across intact mucosal barriers is inherently inefficient and invariable associated with a viral population bottleneck (1, 16). Indeed, in 60–80% of mucosal infections, a single transmitted founder (TF) virus is responsible for productive clinical infection (16). This finding has raised the question of whether the transmission process represents a stochastic event, in which every replication-competent virus has an equal chance of establishing a new infection, or whether the bottleneck selects for viruses that exhibit particular biological properties that predetermine them to establish new infections more efficiently (14). In support of the latter, transmitted viruses have generally been found to exhibit chemokine receptor 5 (CCR5) tropism, infect CD4+ T cells but not macrophages, and share certain genetic features, including shorter variable loops, fewer potential N-linked glycosylation sites, and amino acid signatures, which may affect envelope glycoprotein (Env) surface expression and/or other viral properties (17–22). However, to date, no consistent phenotypic correlate of these genetic signatures has been identified.


The authors declare no conflict of interest.


See Profile on page 6613.

1To whom correspondence should be addressed. E-mail: bhhahn@upenn.edu.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1304288110/-/DCSupplemental.
Initial phenotypic studies of viruses obtained during acute and early infection were limited to analyses of Env functions, which were almost exclusively conducted in the context of pseudoviruses. These analyses led to a number of hypotheses as to how transmitted viruses might differ from their chronic counterparts, including that they engage their receptor and/or coreceptor more efficiently (23, 24), that they are more sensitive to neutralizing antibodies (17, 25), and that their Env glycoproteins interact preferentially with the integrin pair α4β7 (26). However, most of these studies used Env obtained weeks or months postinfection after substantial virus evolution and selection had already occurred. We, thus, used single-genome amplification and a model of random virus evolution to generate Env clones corresponding to actual TF genomes (16, 27). Although we confirmed that TF Env uses CCR5 as the coreceptor for entry, we failed to find evidence that TF Env used CD4 and CCR5 particularly efficiently (28, 29), interacted specifically with α4β7 (29), or exhibited a preferential tropism for particular CD4+ T-cell subsets (28, 29). We also failed to identify an enhanced overall sensitivity of TF Env to neutralization, although subtype B TF Env were slightly more sensitive to CD4 binding site neutralizing antibodies than subtype C TF Env (28). Most recently, we discovered that TF Env are more completely inhibited by small-molecule CCR5 antagonists than control Env constructs, suggesting differences in CCR5 interaction (30); however, the biological relevance of this finding remains to be determined (30). Thus, except for preferential CCR5 coreceptor use and CD4+ T-cell tropism, no other reproducible phenotypic difference between TF and chronic viruses has been identified.

Env pseudotypes are generated by cotransfection of an env expression cassette with an env−minus proviral backbone. Because this backbone represents a standard HIV-1 genome, the contribution of gene products other than Env to the TF phenotype cannot be assessed. Moreover, overexpression of Env, which is inherent in the cotransfection assay, precludes a meaningful assessment of virus-host cell interactions and particle composition. Reasoning that these shortcomings may have obscured important biological differences, we developed methods to clone full-length TF genomes and showed that these produced replication-competent viruses (31–33). However, a systematic evaluation of the TF phenotype also required a sufficient number of molecularly cloned chronic control (CC) viruses for comparison. Here, we describe the generation and biological characterization of a comprehensive panel of TF (n = 27) and CC (n = 14) infectious molecular clones (IMCs), representing two major group M subtypes. Using these reagents, we compared biological properties that would be expected to influence viral fitness during the earliest stages of the transmission process. Our results reveal that TF viruses share common traits that likely enhance their fitness in crossing mucosal surfaces and promoting the establishment of a productive initial infection.

Results

IMCs of TF and CC Viruses. Although a limited number of TF and CC IMCs has previously been reported (29, 31–33), available clones, especially from chronically infected individuals, were too few to conduct meaningful phenotypic comparisons (29). To create a more balanced panel, with respect to both the number of TF and CC IMCs and their subtype representation, we cloned additional viral genomes from individuals enrolled in acute and chronic HIV-1 infection cohorts. Using previously reported methods (31–33), we inferred 12 additional TF genomes (Fig. S1) representing mucosally transmitted viruses from single (n = 8) and multivariant (n = 4) infections. Together with existing constructs, these clones comprised a panel of 27 TF IMCs, with equal representation of subtypes B (n = 13) and C (n = 14) viruses (Table S1).

Chronically HIV-1–infected individuals harbor complex quasispecies of genetically diverse HIV-1 variants. Because it is impossible to predict based on sequence inspection alone which variants are biologically active and which variants are functionally impaired, we amplified between 20 and 40 env genes or 3′ half-genomes from chronic infection plasmas and searched for clusters of nearly identical sequences as indicators of recent clonal expansions (Fig. S2). We reasoned that the inferred common ancestors of these clusters must encode persistently replicating viruses and thus, represent relevant controls for biological comparisons with TF viruses. Consistent with this interpretation, we found that all chronic IMCs generated from such expansion races produced viruses that grew to high titers in CD4+ T cells. However, not all chronic infection plasmas were suitable for IMC construction. Analyzing specimens from over 60 individuals, we identified only 14, including 4 reported previously (29), that exhibited clonal expansion races in both 3′ and 5′ halves of their viral genomes (Fig. S2). These chronic plasma samples were used to construct CC IMCs representing both subtypes B (n = 5) and C (n = 9) infections (Table S1).

To determine the coreceptor use of 22 newly derived IMCs, we infected CCR5- and CXCR chemokine reporter type 4 (CXCR4+)-expressing reporter cells in both the presence and absence of their respective inhibitors (16). Consistent with previous analyses of Env pseudotypes (16, 28, 29, 34), the results indicated that all IMCs were CCR5 tropic, except for one TF and two CC viruses that were dual tropic for CCR5 and CXCR4 (Table S1).

Phenotypic Studies. All biological experiments were performed using viral stocks that were CD4+ T cell–derived, sucrose-purified, and depleted of CD45+ microvesicles. Virus was quantified by measuring reverse transcriptase (RT) activity, viral RNA copy number, and Gag p24 antigen content. Comparing these values, we noticed that subtype C stocks seemed to contain about fivefold less p24 antigen per unit of RT activity than subtype B stocks (Fig. S3A). This discrepancy was also seen when p24 antigen was normalized using RNA copy numbers (Fig. S3B). However, no such difference was observed when virion RT activity was compared with RNA copy number (Fig. S3 C and D). Similar results were obtained when viral stocks were tested in two additional p24 detection assays (Fig. S3 E and F). Thus, traditional ex vivo assays using CCR5 and CXCR4 signal proteins considerably less efficiently than subtype B core proteins, leading to a systematic underestimation of the number of viral particles in subtype C viral stocks. We, therefore, used RT activity to normalize virus input and measure virus replication in all subsequent experiments.

To determine whether TF and CC viruses differed in their phenotypic properties, we used two complementary statistical approaches. First, we used a conservative nonparametric permutation test (perm test) to address the central question of whether TF and CC viruses exhibited reproducible phenotypic differences. Second, we used a generalized linear model (GLM) to test for interactions between different parameters, such as subtype and cell donors in replicate experiments, and test for the independence of viruses from the same subjects. Because both virus status (TF or CC) and subtype (B or C) influenced the biological outcome, the GLM test allowed us to interpret the data with these variables taken into account.

TF Viruses Exhibit Enhanced Infectivity. Plasma virus collected during ramp-up stages of acute simian immunodeficiency virus (SIV) infection has been shown to be significantly more infectious than virus collected from later plasma samples (35). To examine whether this observation was also true for HIV-1, we used a single-round infection assay to determine whether TF virions were more infectious on a per-particle basis than CC virions. Serial dilutions of CD4+ T cell–derived viral stocks were used...
to infect TZM-bl cells (which express luciferase under the control of an HIV-1 promoter), and the resulting relative light units were expressed as a function of the input RT activity. The results showed that TF viruses as a group were 1.7-fold more infectious than CC viruses (Fig. 1A), although this difference was only marginally significant ($P = 0.049$ by perm test; $P = 0.062$ by GLM). Similar results were obtained when subtypes B and C viruses were considered separately (there was no subtype dependence of infectivity) (Fig. 1B). To enhance virus infectivity, we repeated these experiments in the presence of diethylaminoethyl (DEAE) dextran, which is known to increase virus attachment to target cells (36). As expected, DEAE dextran increased the per-particle infectivity of all viruses by up to three orders of magnitude (Fig. 1C), with a significantly greater effect on subtype C than subtype B viruses ($P = 0.015$ by GLM) (Fig. 1D). However, in the presence of DEAE dextran, the marginally significant difference between TF and CC viruses shifted to nonsignificance ($P = 0.068$ by perm test; $P = 0.083$ by GLM). Taken together, these data indicate that TF viruses are, on average, nearly two times as infectious as viruses that persist during chronic infection. However, when normal barriers to cell-free infection are mitigated by DEAE dextran, this difference is no longer significant (although a trend was still evident).

**TF Virions Have a Higher Env Content.** Previous computational analyses of TF and CC env gene sequences revealed genetic signatures in TF viruses that increased Env expression and particle incorporation when tested in the context of pseudoviruses (37). We, thus, examined whether TF virions packaged more Env per particle than CC virions. To detect both subtypes B and C Env proteins with comparable efficiency, we developed an ELISA that used antibodies previously shown to bind genetically highly diverse envelope (gp120) proteins. To capture Env, we used CD4-218.3-E51, a chimeric antibody that contains the first and second domains of human CD4 linked to the CD4-induced monomeric E51 (38). To detect bound Env, we selected affinity-purified anti-gp120-specific polyclonal antibodies that were isolated from human plasma (Advanced Bioscience Laboratories, Inc.). Using this ELISA, we found that TF viruses contained 1.9 times more Env per unit of RT activity than CC viruses ($P = 0.048$ by perm test; $P = 0.057$ by GLM) (Fig. 1E). Subtype B viruses appeared to package 2.4-fold more Env per particle than subtype C viruses (Fig. 1F), but this result was largely caused by a binding preference of the capture antibody, which recognized subtype B Env glycoproteins two times more efficiently than subtype C Env glycoproteins. Nonetheless, this subtype bias did not affect the differential between TF and CC viruses, which was seen for both subtypes B and C, and it was controlled for in the statistical analyses (Fig. 1F). As expected, there also was a significant correlation between Env content and particle infectivity ($r = 0.33; P = 0.036$).

**TF Viruses Bind DCs More Efficiently.** DCs have been proposed to play an important role in HIV-1 transmission, because they are located in the mucosa (39, 40), capture viruses using lectins (41, 42) and glycosphingolipid receptors (43), and efficiently transmit infectious particles to CD4+ T cells (44–46). To examine whether TF and CC viruses differ in their ability to bind DCs, we pulsed immature monocyte-derived DCs (moDCs) with equal amounts of virus (normalized by RT activity), washed the cells extensively to remove cell-free virions, and then lysed the cells to quantify the amount of cell-associated virus. Using cells from three different donors (Fig. 2A), we found that moDCs captured TF viruses 1.6 times more efficiently than CC viruses ($P = 0.040$ by perm test; $P = 0.060$ by GLM). This increase was observed for both subtypes B and C TF viruses (Fig. 2B), although subtype B viruses were captured 3.4 times more efficiently than subtype C viruses ($P = 4.6 \times 10^{-6}$ by GLM) (Fig. 2B).

Because in the above experiment, we had normalized virus input by RT activity but measured virion capture using the more sensitive p24 assay, we considered the possibility that the subtype-specific differences were an artifact of the p24 antigen capture assay. To explore this possibility, we repeated the binding experiment with cells from three additional donors, but this time, we normalized virus input by p24 content (Fig. 2C and D). Despite adding an estimated fivefold excess of subtype C virions, TF viruses were again captured 1.8 times more efficiently than CC viruses ($P = 0.030$ by perm test; $P = 0.014$ by GLM). Moreover, there was still a subtype-specific difference, with subtype B viruses being captured 1.9 times more efficiently than subtype C viruses ($P = 0.0029$ by GLM). As a control, we treated an aliquot of the pulsed moDCs with 0.25% trypsin-EDTA and showed that this treatment removed all detectable cell-associated virus, confirming that most of the DC-associated virus was surface-exposed (47). We also cultured virus-exposed moDCs and showed that these cells were not productively infected. Finally, we asked whether the percentage of captured virus for each strain correlated with different donors. This association

**Fig. 1.** Virion infectivity and Env content. (A–D) Infectivity values for TF and CC viruses (x axis) are expressed as relative light units (RLUs) per picogram of viral RT activity (y axis). (A) Bars indicate the median infectivity of TF (filled) and CC (open) viruses, with interquartile ranges indicated. TF viruses were 1.7-fold more infectious than CC viruses ($P = 0.049$). (B) Infectivity values are shown for each virus. Subtypes B and C viruses are shown in red and blue, respectively, with TF viruses in dark colors and CC viruses indicated in light colors, respectively. Values represent averages from four independent experiments. (C and D) Infectivity values are shown for TF and CC viruses as in A and B, except that infections were performed in the presence of DEAE dextran. Values represent averages from three independent experiments. (E and F) Env content of TF and CC viruses (y axis) is expressed as the mass ratio of Env and RT content (y axis). (E) Bars indicate the median values of Env content for TF (filled) and CC (open) viruses, with interquartile ranges indicated. TF viruses contained 1.9 times more Env per unit of RT activity than CC viruses ($P = 0.048$). (F) Env content is shown for each virus and color-coded as in B and D. Values represent averages from two independent experiments.
Fig. 2. Virus binding to moDCs. The percent of captured TF and CC virus is plotted (y axis) for moDC cell preparations from six different donors labeled A through F (x axis). (A and B) Virus input was normalized by RT activity. (C and D) Virus input was normalized by p24 content. (A and C) Bars indicate median values of moDC capture for TF (filled) and CC (open) viruses, with interquartile ranges indicated. TF viruses were captured 1.7-fold more efficiently than CC viruses (P = 0.035), (B and D) Values are plotted for each virus individually (color-coding for TF and CC viruses from subtypes B and C as in Fig. 1). Subtype B viruses were captured 3.4 times more efficiently than subtype C viruses (P = 4.6 × 10⁻³ by GLM).

was indeed observed regardless of whether virus input was normalized by RT activity or p24 content, thus validating the DC binding assay (Fig. S4A).

Although the combined data from all six donors indicated that DCs captured TF viruses 1.7-fold more efficiently than CC viruses (P = 0.035 by perm test; P = 0.005 by GLM), there was considerable donor variability, with moDCs from two donors (A and F in Fig. 2) failing to yield significant binding differences (Fig. S4). To examine potential reasons, we compared surface expression levels of DC-specific intracellular adhesion molecule 3-grabbing nonintegrin (DC-SIGN) and macrophage mannose receptor (MMR) by flow cytometry. Interestingly, moDCs of donors A and F expressed lower levels of these molecules (Fig. S4 B and C), suggesting that the number of Env-specific binding sites on their cells was limited. For the remaining four donors, moDC virus capture tended to correlate with particle Env content (Fig. S4D). These data suggest that virion Env content represents one important determinant in moDC binding, although expression of lectins and other factors clearly also play a role (48).

To examine the efficiency of virus transfer from moDC to CD4+ T cells, we pulsed moDCs with equivalent amounts of TF and CC viruses, cocultured these cells with autologous CD4+ T cells, and then measured RT activity in culture supernatants as an indicator of virus replication. Analysis of cells from two different donors showed that TF viruses replicated to slightly higher titers than CC viruses each day over the course of 10 d (one representative donor is shown in Fig. 3). Although these differences were not significant in the perm test (P = 0.104), significant differences in TF and CC virus titers were observed in the GLM analysis (P = 0.025). To explore the reason for this apparent discordance, we ran the perm test on subtypes B and C viruses separately (Fig. 3B), and we found that subtype B TF viruses (P = 0.004), but not subtype C TF viruses (P = 0.23), grew to significantly higher titers than their respective chronic controls. Averaging data from both donors and across all time points, this difference was 11.2-fold for subtype B viruses but not significant for subtype C viruses. Because subtype B virions bound moDCs 3.5-fold more efficiently than subtype C viruses (Fig. 2 B and D), we interpret these findings to indicate that virions that bind moDCs more efficiently are also transferred to CD4+ T cells more efficiently.

Fig. 3. DC-mediated trans infection. (A) Virus replication expressed as picograms of RT activity per milliliter of culture supernatant (y axis) is shown after cocultivation of virus-pulsed moDC with CD4+ T cells for one representative donor (of two analyzed) over 10 d (x axis). Bars indicate median values of viral replication for TF (filled) and CC (open) viruses, with interquartile ranges indicated (there was no significant difference between TF and CC viruses). (B) Values are plotted for each virus individually (color-coding for TF and CC viruses from subtypes B and C as in Fig. 1). Averaging data from two different donors, subtype B TF viruses grew to 11.2-fold higher titers than subtype B CC viruses (P = 0.004), whereas no significant differences were observed for subtype C TF and CC viruses (P = 0.23).
activity in culture supernatants (Fig. 4A, C, and E, gray lines). Although TF viruses replicated to slightly higher titers, especially at the earliest time point postinfection, this difference was not statistically significant ($P = 0.16$ by perm test; $P = 0.186$ by GLM). Thus, in activated CD4+ T cells, TF and CC viruses replicated with comparable efficiency.

IFN-α is produced early in infection by plasmacytoid DCs (50), which are among the first cells to be recruited to the mucosal site of virus entry (4). This cytokine stimulates expression of hundreds of host genes [IFN-stimulated genes (ISGs)] (51), many of which have anti–HIV-1 activity (52). IFN-α has also been shown to limit Env incorporation into virus particles that are released from infected CD4+ T cells (53). To determine the sensitivity of TF and CC viruses to the antiviral activity of this cytokine, we analyzed their replication kinetics in CD4+ T cells in the presence of 500 U/mL IFN-α. Determining RT activity in culture supernatants as a measure of viral replication, we found significant differences in the growth rate of TF and CC viruses (Fig. 4A, C, and E, black lines). However, this effect was highly subtype-specific (Fig. S5). Averaging replication data from all donors and across all time points, we found that subtype B TF viruses grew to 62-fold higher titers than subtype B CC viruses ($P = 0.000013$ by GLM). In contrast, subtype C TF viruses grew only 1.7-fold more efficiently than their respective CC viruses, which was not significant ($P = 0.53$ by GLM). Importantly, this difference was not because of a lack of IFN-α resistance of the subtype C TF viruses but an increased IFN-α resistance of the respective chronic controls (Fig. 4B, D, and F). Inclusion of sex and risk group information into the statistical analyses indicated that neither was significantly associated with IFN-α resistance; however, the study was not sufficiently powered to examine the impact of sex and risk group, and a contribution of these factors cannot, therefore, be excluded (Fig. S5).

As a different way to analyze these data, we also calculated the ratio of virus production in the presence and absence of IFN-α for each virus strain (Fig. 4B, D, and F). Comparing across all donors, we again found that TF viruses were significantly more resistant to inhibition by IFN-α than chronic viruses ($P = 0.027$ by perm test). However, when considering the two clades separately, this finding was significant only for subtype B ($P = 0.0004$ by one-sided Wilcoxon test) and not for subtype C ($P = 0.84$). Viruses that were highly resistant in one donor were also highly resistant in the other donors (Fig. S6), indicating that their IFN-α phenotype was consistent between donors.

To determine whether the higher titers of TF viruses were because of increased particle release or differences in viral spread, we used flow cytometry to measure the percent of infected cells in IFN-α-treated cultures from two donors (Fig. S7). In the presence of IFN-α, TF viruses infected a larger number of cells than CC viruses, although the overall effect was only marginally significant ($P = 0.049$ by perm test). However, GLM analysis again detected a significant interaction between subtype and TF/CC status ($P = 0.041$), indicating that the differences were primarily caused by subtype B viruses. Because the ratio of virus production in the presence and absence of IFN-α correlated well with the percent of Gag-positive cells ($P < 0.0001$) (Fig. S7 B and C), it is likely that IFN-α resistance represents an enhanced ability to spread between CD4+ T cells. Interestingly, this phenotype correlated only weakly with particle infectivity and Env content, indicating that the higher levels of particle-associated Env in TF viruses are not the main drivers of their relative IFN-α resistance.

**Discussion**

A primary goal of AIDS vaccine development is to prevent acquisition of HIV-1 at mucosal surfaces. In this context, it is critical to determine whether newly transmitted HIV-1 strains share traits that provide novel targets for effective immunization. For such an analysis, the use of full-length IMCs is essential, because they contain the complete genetic information of viruses that successfully transmit infection. Moreover, genetic linkage is maintained to preserve potentially important regulatory and structural protein interactions (16, 31–33). Characterizing a large set of TF and CC IMCs from two major HIV-1 group M subtypes, we found that TF viruses were slightly more infectious than CC viruses on a per particle basis, packaged slightly more
Env, bound to moDCs more efficiently, and were relatively more resistant to the antiviral effects of IFN-α. These biological differences could act in concert to enhance cell-free infection and virus replication in the face of an early innate immune response.

Analyzing particle composition and infectivity, we found that TF virions were, on average, two times as infectious and contained two times as much Env as CC virions. These phenotypic differences are consistent with data from a recent genetic study (22) that identified sequence signatures in the signal peptide of TF Env sequences. When tested in the context of pseudoviruses, these signatures increased steady-state Env expression and particle incorporation (37). Assuming 35 enzymatically active RT dimers per virion, we estimate that the median TF virion contains 18 Env trimer spikes, whereas the median CC virion contains 7 spikes, a value essentially identical to previous estimates (54). Because only one or a few functional Env trimers are believed to be necessary for cell fusion and infection (55), it would seem that TF viruses carry an excess of envelope glycoprotein. Such excess is likely of advantage during the earliest stages of HIV-1 infection, because viruses with a higher Env content may also contain more functional trimers. Particle-associated Env may also play a role beyond the mere fusion and cell entry function. For example, excess Env could ensure more stable attachment of virions to target cells and/or be involved in receptor and/or coreceptor-mediated signaling (56). In addition, excess Env may compensate for shedding or inactivation by semen (57) and cervicovaginal mucus (58). Thus, it seems likely that the increased Env content and associated enhanced infectivity of TF viruses serve to increase their transmission fitness. Conversely, the lower Env content of CC viruses may be more advantageous after infection is established and may even be selected for by Env-specific neutralizing antibodies.

We also found that TF viruses bind more efficiently to moDCs than their chronic counterparts. Although moDCs are only a proxy for mucosal DC subsets, previous studies have shown that tissue- and in vitro-derived DCs have similar virus capture capabilities (59, 60). Thus, it is likely that the improved interaction of TF viruses with tissue culture-derived moDCs is also a reflection of their in vivo biological properties. The differential binding of TF viruses was statistically significant across different donors and across both clades, and thus, it seems to represent a general feature of TF viruses. There are at least two steps in the transmission process where this phenotype may confer an advantage. First, more efficient binding to tissue resident DCs may increase the likelihood of a successful transfer to CD4+ T cells. Second, more efficient binding to emigrating DCs may promote the seeding of regional lymph nodes (61). In this context, it is of interest that adaptive changes that conferred mucosal transmissibility to a commonly used simian-human immunodeficiency virus (SHIV-SF162P3) increased the binding capacity of its Env glycoprotein to DC-SIGN (62). Our data are, thus, consistent with the view that DCs are among the first cells to interact with HIV-1 after exposure and that this interaction is important for transmission, although DCs themselves are not productively infected.

One unexpected finding was that TF viruses replicated and spread in CD4+ T cells in the presence of IFN-α more efficiently than CC viruses. Although this effect was most pronounced for subtype B TF viruses, a trend was also observed for subtype C TF viruses. This finding raises the question of whether IFN-α responses in the mucosa are protective against HIV-1 infection. On first glance, a review of the literature would suggest that the answer to this question is no. Previous studies in macaques concluded that the initial innate immune response is ineffective at containing SIV infection and may even enhance early viral replication (4). Moreover, mucosal pretreatment with toll-like receptor agonists induced IFN-α but did not protect animals from SIV challenge (63). However, these studies were not only assessing the effects of IFN-α but also causing immune activation by inducing proinflammatory cytokines and chemokines. In fact, it has been shown more recently that blockade of type 1 IFN during acute SIV infection resulted in accelerated progression to AIDS and death (64). It, thus, seems that early IFN responses can be protective against HIV-1/SIV infection. The fact that IFN-α levels (65) as well as IFN regulatory factor polymorphisms (66) have been associated with protection in highly exposed but uninfected individuals is consistent with this hypothesis. Thus, the induction of an effective early antiviral state seems to contribute to the population bottleneck associated with mucosal HIV-1 transmission.

We hypothesized that virions containing more Env would be more resistant to IFN-α, because in previous studies, this cytokine was reported to decrease cell-free virus transmission (67), particle release (68), and particle infectivity (53). However, we failed to see a significant correlation between IFN-α resistance and Env content or virion infectivity. Thus, higher baseline Env incorporation cannot explain the IFN-α-resistant phenotype of TF viruses. Nonetheless, the ability of TF viruses to replicate in the presence of IFN-α indicates that resistance-conferring determinants must exist. There are several candidates, such as the accessory proteins Vpu and Vif, that are known to counteract the IFN-stimulated genes tetherin and APOBEC3G, respectively (69, 70), as well as the Tat protein, which has been reported to modulate the IFN-induced RNase L antiviral pathway (71) and down-regulate the ISG protein kinase R (72). Structural proteins, such as Gag, may also contribute to IFN-α resistance by exhibiting differential sensitivity to the tripartite motif family of proteins (73–75). Finally, IFN-α resistance may not only be protein-mediated, because variation in the number of transcription factor binding sites in the long terminal repeat (LTR) region could also influence IFN-α sensitivity (76, 77). Notably, there is a relationship between IFN-α resistance and efficiency of transmission in other viruses, including Venezuelan Equine Encephalitis virus (78) and Bovine Viral Diarrhea Virus (79), suggesting that increased IFN-α resistance may be a general property of transmitted founder viruses.

If IFN-α resistance was a prerequisite for efficient HIV-1 transmission, one would expect this viral property to be conserved across all clades. The lack of significance between subtype C TF and CC viruses seems to argue against this conclusion; however, Fig. 4 shows that the lack of significance is not caused by a loss of IFN-α resistance of subtype C TF viruses. Instead, the 62-fold cumulative replication differential observed for subtype B viruses reflects a much greater IFN-α sensitivity of the respective chronic controls. Although our panel of TF and CC IMCs is much larger than any previously tested, there are epidemiological differences between the patient cohorts from which they were derived. For example, all but one of the chronic subtype B IMCs were derived from men who had sex with men, whereas seven of eight chronic subtype C IMCs were derived from heterosexual women (Table S1). This sex bias may have contributed to the increased IFN-α resistance of chronic subtype C viruses, because plasmacytoid dendritic cells (pDCs) of women have been reported to generate more IFN-α in response to HIV-1 than pDCs of men (80). In addition, all subtype C viruses were derived from individuals living in Africa, whereas all subtype B viruses were derived from patients residing in the United States (Table S1). Thus, environmental factors, such as concurrent infections that increase general inflammation levels, may have selected for the much higher IFN-α resistance of the respective chronic viruses. Finally, if loss of IFN-α resistance is the result of pressure from host adaptive responses, then duration of infection may have been an important variable. We have recently found that subtype B viral isolates obtained from subjects later in infection...
are more sensitive to IFN-α than the viruses isolated shortly after transmission. Thus, it seems clear that, to decipher the reasons for the IFN-α resistance differences observed here, future studies will need to consider patient demographics, route of transmission, and duration of infection in addition to TF and CC virus status. Such studies may also shed light on factors that fuel the HIV-1 epidemic in Africa.

One of the motivations to characterize the phenotypic properties of TF viruses is the possibility that virus traits might be uncovered that could represent useful vaccine targets. In this context, the increased Env content of TF viruses may increase their sensitivity to neutralization by allowing antibodies to bind bivalently and potentially cross-link more densely packed Env spikes (51). Similarly, the identification of key ISGs that contribute significantly to viral control during the most vulnerable stages of HIV-1 infection may provide a new vaccination strategy, by either vaccine vector-mediated type 1 IFN production and/or induction of T cells that express the antiviral ISGs.

The molecular identification and biological analysis of TF viral genomes are powerful enabling tools for characterizing the transmission requirements and subsequent evolution of HIV-1, SIV, and other viruses (16, 31–33, 82). In the present study, we used TF analyses to characterize the biological properties of full-length HIV-1 genomes captured at the moment of transmission. Although we identified significant phenotypic differences, none of these differences completely differentiated TF from CC viruses, and the magnitude of the observed differences was generally modest. This finding is not surprising, because TF viruses, by necessity, represent a subset of a much larger and more diverse set of viruses that replicates persistently throughout chronic infection. We speculate that the viral population bottleneck associated with mucosal transmission is largely caused by inherent physical barriers but also specific virus–host cell interactions necessary for early virus infection and replication, each of which may pose hurdles to virus acquisition. Thus, even slight advantages of TF viruses at one or more of these biophysical checkpoints could significantly increase their transmission fitness. Defining the viral determinants that underlie this fitness will provide additional insights into the biology of mucosal HIV-1 transmission and may also inform vaccine design.

Materials and Methods

**IMC Construction.** The sequences of full-length TF and CC viruses were inferred and molecularly cloned as described (27, 29, 31–33). GenBank accession numbers are listed in Table S2.

**Viral Stocks.** CD4+ T cells pooled from nine donors were infected with 293T cell-derived transfection supernatants. Viral stocks were harvested after 11 d and depleted of CD4+ microvesicles. Viral p24 core protein content was measured using the PerkinElmer High-Sensitivity AlphaLISA, the PerkinElmer Alliance ELISA, and the method described by Bianchot et al. (83). RT activity was measured using the Roche Reverse Transcriptase Assay. Viral RNA was quantified using the Roche COBAS AmpliPrep/COBAS Taqman HIV-1 v. 2.0 Test.

**Phenotypic Analyses.** Coreceptor use was determined as described (33). Virion infectivity was measured using the TZM-bi reporter cell line. Particle-associated Env content was measured using a newly developed ELISA (SI Materials and Methods). To assess DC capture and CD4+ T-cell transfer, peripheral blood monocytes were differentiated to moDCs using 100 IU/mL granulocyte-macrophage colony-stimulating factor (GM-CSF) and 200 IU/mL IL-4. For moDC capture experiments, 1 × 106 cells were incubated with viral stocks containing either 30 pg RT activity or 600 pg p24 protein. For CD4+ T-cell transfer experiments, viral stocks containing 50 pg RT activity were added, and moDCs were cultured at a 1:5 ratio with autologous CD4+ T cells. For CD4+ T-cell infection studies, 2 × 106 negatively selected CD4+ T cells were infected overnight with viral stocks containing 30 pg RT activity. Cells were then propagated in media containing 30 U/mL IL-2 in the presence and absence of 500 U/mL IFN-α. Infected cells were identified by flow cytometry (SI Materials and Methods).

**Statistical Analyses.** Data were analyzed using a nonparametric perm test that controlled for virus subtype, donor cells, and time of sampling as well as a GLM that tested for interactions between and the relative importance of these variables (SI Materials and Methods).

ACKNOWLEDGMENTS. We thank John Moore, Frank Kirchhoff, Paul Sharp, and Stuart Shapiro for helpful discussions; the University of Pennsylvania’s Center for AIDS Research (CFAR) Human Immunology, Flow Cytometry, and Viral and Molecular Core facilities for reagents and protocols; and Patricia Crystal for artwork and manuscript preparation. This work was supported by National Institutes of Health (NIIH) Grants R01 AI45378, R01 AI040080, P30 AI45008, and P30 AI27767, Center for HIV/AIDS Vaccine Immunology Grant U19 AI67854, Center for HIV/AIDS Vaccine Immunology and Immunogen Discovery Grant U11 AI00645, and Bill and Melinda Gates Foundation Grant 37874, N.F.P., S.S.I., and C.B.W. were supported by NIH Training Grant T32 AI07632.


Supporting Information

Parrish et al. 10.1073/pnas.1304288110

SI Materials and Methods

Study Subjects. Plasma samples were obtained from subjects enrolled in acute and established HIV-1 infection cohorts (CHAVI cohort 001). A summary of available epidemiological, clinical, and infection status data is shown in Table S1. Acutely infected individuals were staged as reported previously (1). Whole blood was collected in acid citrate dextrose, and plasma was separated and stored at ~70 °C. All subjects provided written informed consent for the collection of samples and subsequent analyses. The study was approved by the Institutional Review Boards of the University of Pennsylvania and Duke University.

Single-Genome Amplification. Single-genome amplification (SGA) was performed as described previously (2–5). Briefly, plasma viral RNA was extracted and reverse-transcribed using SuperScript III Reverse Transcriptase (Invitrogen), and the resulting cDNA was diluted in 96-well plates such that each well contained on average less than one cDNA template; 3′ and 5′ genome halves were amplified by nested PCR using primers and conditions as described (2–5). Amplicons were sequenced directly, and any sequence containing double peaks in the chromatogram was excluded from additional analysis. The resulting sequences are, thus, representative of viral genomes present in the plasma and devoid of PCR artifacts and cloning biases.

Infectious Molecular Clone Construction. Full-length transmitted founder (TF) and clonally expanded chronic control (CC) genome sequences were inferred as described (3–7). To clone TF genomes, overlapping half-genome and long terminal repeat (LTR) fragments were amplified from peripheral blood mononuclear cell (PBMC) DNA, cloned, and combined within a single plasmid vector (pBR322, pUC9, or pCR-XL-TOPO) as a complete proviral genome (3). To clone CC genomes, consensus sequences of viral expansion rakes were chemically synthesized as three to four fragments (Blue Heron Biotech or GenScript) and used to construct a complete provirus (5). This approach was also used for TF viruses derived from patients who were infected by multiple variants, except for ZM247Fv1 and ZM247Fv2, which have been described previously (3). All newly derived infectious molecular clones (IMCs) have been submitted to the National Institutes of Health AIDS Research and Reference Reagent Program.

Virus Stock Preparation. Virus stocks were generated by transfecting 30% confluent 293T cells (in a 10-cm dish) with 6 μg IMC DNA. Virus was harvested 72 h posttransfection and passed through a 0.45-μm polyvinylidene fluoride filter. Viral infectious units for each stock were determined using TZM-bl cells as described previously (8). To generate physiologically relevant stocks, CD4+ T cells were positively selected (Milltenyi Biotech) from buffy coats of nine healthy donors (Research Blood Components) and cryopreserved using CS10 freezing media (CryoStor). Cell aliquots were thawed quickly in a 37 °C water bath, resuspended at a density of 1 x 10^6 cells/mL, allowed to recover overnight in RPMI 1640 media containing 10% (vol/vol) FBS and 30 μU/mL IL-2 (CD4+ T-cell media) in a 37 °C incubator with 5% (vol/vol) CO_2, and then stimulated by adding Staphylococcal Enterotoxin B (SEB) at a concentration of 50 ng/mL for 48 h. Cells were then pooled, and 1 x 10^7 cells were inoculated with 1 x 10^7 infectious units each 293T-derived virus stock overnight in 1 mL CD4+ T-cell media. Cells were washed and resuspended at 1 x 10^6 cells/mL in CD4+ T-cell media, and additional 10 mL media were added at day 3 after infection. On day 5 after infection, all media were exchanged, and cells were resuspended in 30 mL CD4+ T-cell media. At day 11 post-infection, microvesicles were depleted as described elsewhere (9) with the following modifications: biotinylated anti-CD45 antibody (clone 2D1; R&D systems) was added to streptavidin-coated magnetic beads (M-270 beads; Invitrogen) for 1 h at room temperature at a concentration of 10 μg mAb/mg bead, washed three times in PBS, and then stored at 4 °C. Beads were added to clarified supernatant to a final concentration of 100 μg/mL for 1 h at room temperature and then magnetically captured using a 50-mL conical tube magnet (StemCell) at 4 °C for 20 min. Supernatant was passed through a 0.45-μm polyvinylidene fluoride filter, and virus was pelleted through a 20% (vol/vol) sucrose cushion at 100,000 g, resuspended in PBS, and stored in aliquots at ~80 °C.

The p24 antigen content in virus stocks was quantified using commercially available PerkinElmer High-Sensitivity AlphaLISA and Alliance ELISA p24 Assays as well as an assay developed by Biancotto et al. (10). Viral RNA was quantified using the Roche COBAS AmpliPrep/COBAS Taqman HIV-1 v.2.0 Test. Reverse transcriptase (RT) activity was determined using the Roche colorimetric RT assay.

Because the AlphaLISA p24 Assay was approximately two orders of magnitude more sensitive than the RT assay, we sought to convert p24 antigen concentrations to RT activity to use this assay when RT determinations were below the limits of detection. Thus, we calculated the ratio of p24 to RT activity for each virus in four independently prepared high-titer virus stocks. Because these ratios were highly reproducible, we used the average ratio for each virus strain to convert p24 to RT activity.

Coreceptor Tropism. Coreceptor tropism was assessed by pre-incubating TZM-bl cells with 1.2 μM AMD3100 and/or 10 μM TAK779 before infection with 1 x 10^6 infectious units 293T cell supernatant. Viruses were considered dual tropic if inhibited at least 5% by AMD3100 and no more than 95% by TAK779.

Analysis of Per-Particle Infectivity. TZM-bl cells (1 x 10^4) were seeded per well in a 96-well plate in 100 μL DMEM containing 10% (vol/vol) FBS. After 24 h, media were removed from each well, and cells were infected with 50 μL virus serially diluted in DMEM containing 1% FBS with or without 10 μg/mL diethylaminoethyl dextran; 2 h later, 50 μL DMEM containing 20% (vol/vol) FBS were added, and 36 h after infection, luciferase production was assayed using Brite-Glo luciferase substrate (Promega) and a Biotek H4 reader. As a control, a third-generation fusion inhibitor (T2635) was added at a saturating concentration (1 μg/mL) 12 h after virus addition to one of the duplicate wells (11). Relative light units (RLUs) generated per volume of each virus stock were calculated using all virus dilutions in the linear range of the assay (2 x 10^3 to 6 x 10^5 RLU). The infectivity per particle was then calculated as the RLU generated per picogram RT activity present in each virus stock.

Particle Envelope Glycoprotein Content. A sandwich ELISA was developed to measure virion envelope glycoprotein (Env) content; 200 pg monoclonal E51-218.3-CD4 were added in 100 mL 0.2 M sodium carbonate/bicarbonate (pH 10) solution to each well of 96-well ELISA plates (Nunc MaxiSorp) and incubated overnight. E51-218.3-CD4 is a chimeric antibody that contains the first and second domains of human CD4 linked to the CD4-
Dendritic Cell Capture and Trans Infection Assays. Monocyte-derived dendritic cells (moDCs) were generated from peripheral blood mononuclear cells (PBMCs) purified by CD14 immunomagnetic selection from buffy coats of healthy donors. Monocytes were further enriched by removing cells that did not adhere to poly-styrene after 1 h incubation at 37 °C in RPMI 1640 with 5% (vol/vol) normal human serum (NHS) and 1% Hepes buffer. Adherent cells were cultured in the same media supplemented with 100 IU/mL granulocyte-macrophage colony-stimulating factor (GM-CSF) and 200 IU/mL IL-4 (refreshed every 2 d). At day 6, 800 μL T-cell media were sampled, and calcium and magnesium to remove cell-free virus. Cells were then lysed in PBS containing 0.5% Triton X-100, and the viral p24 antigen of each virus was determined, and if an experiment was repeated, ranks were split, with one-half of the cells processed to isolate monocytes as above and the other one-half depleted of non-CD4+ T cells (RosetteSep; StemCell Technologies). moDCs were generated as above, and CD4+ T cells were stimulated for 48 h with 50 ng/mL SEB and then cultured at 1 × 10^6 cells per milliliter in CD4+ T-cell media. On day 7, 1 × 10^5 DCS were pulse for 3 h with 50 pg RT of each virus, washed as above, and then cocultured with 5 × 10^5 CD4+ T cells in 1 mL CD4+ T-cell media in 48-well plates. To measure virus replication, 200 μL media were sampled, with replacement every 2 d. At day 6, 800 μL media were exchanged.

Virus Replication in the Presence and Absence of IFN-α. CD4+ T cells were isolated from buffy coats by depleting non-CD4+ T cells and then stimulated with 50 ng/mL SEB for 48 h in T-cell media at a concentration of 1 × 10^6 cells per milliliter. The mixture used for depletion contains anti-CD123 antibodies, and thus, it also depletes plasmacytoid DCs, which can produce INF-α. Cells were washed and resuspended at 1 × 10^6 cells per milliliter, with 2 × 10^6 cells incubated overnight with an equivalent amount of each virus (30 pg RT activity). Next, cells were washed two times with 1 mL PBS and divided, with one-half cultured in T-cell media alone and the other one-half supplemented with 500 U/mL recombining INF-α (A2; PBL Laboratories) in 1 mL in 48-well plates; 200 μL media were sampled, with replacement every 3 d. At day 6, 800 μL were exchanged with fresh media containing 500 U/mL INF-α.

On day 12, cells were analyzed by flow cytometry. Cells were washed one time in PBS and then incubated with 1 μL 1:1,000 dilution live/dead green stain (Invitrogen) in DMSO. Cells were incubated for 30 min at 4 °C. Cells were fixed in 1% paraformaldehyde and analyzed on an Amnis ImageStream instrument. Gag-positive cells were expressed as a percentage of all living cells. All samples were collected and analyzed under code, and four mock-infected cultures were used to define baseline anti-Gag staining. Cell viability at day 12 in mock-infected cultures was comparable in the presence and absence of INF-α.

Statistical Analyses. We used two statistical strategies to determine whether TF and CC viruses have distinct phenotypic characteristics. A nonparametric rank-based permutation test was used to determine whether values from TF viruses were significantly different from those values derived from CC viruses. This test was designed to account for clade and donor differences, and the P value reflected the likelihood of observing differences between TF and CC viruses by chance alone. The test took into account that (i) viruses were derived from two clades (B and C), (ii) repeat experiments were performed using different donor cells, and (iii) in some experiments, responses were evaluated at different times. Specifically, the TF ranking for each experiment was determined, and if an experiment was repeated, ranks were summed across replicate experiments. TF and CC labels were induced monoclonal ES1. Its architecture is similar to the architecture of other chimeric antibodies (12), except that the peptide that links the CD4 and ES1 moieties was modified for stability. Wells were washed two times with 200 μL PBS with magnesium and calcium containing 0.2% Tween 20 (PBS-T), blocked with 200 μL 5% (vol/vol) nonfat milk in PBS-T for 2 h at room temperature, and washed three times with 200 μL PBS-T. Plates were stored at 4 °C for up to 1 mo. Virus was diluted in PBS-T with 1% Triton X-100 detergent to disrupt virions, added to wells in three 10-fold serial dilutions (100 μL), and incubated at 37 °C for 2 h. Commercially available gp120 standards of known concentrations were added to wells in serial twofold dilutions starting at 100 ng/mL. Wells were washed five times with PBS-T and then incubated with 100 μL HRP-conjugated, affinity-purified human anti-gp120 (Advanced Bioscience Laboratories) for 1 h at 37 °C. Wells were washed five times with PBS-T and incubated with o-phenylenediamine dihydrochloride substrate followed by 2N sulfuric acid, and absorbance was read at 490 nm with a 630-nm reference. After 15 min of development, the assay had a dynamic range from ~100 pg gp120 per milliliter to 50 ng gp120 per milliliter. Multiple dilutions were used to determine the Env content for each virus stock, which was then normalized by RT activity.
then randomly reassigned 10,000 times, and the sum of the various ranks from the randomized data was tallied. The fraction of occurrences of a given rank-sum value observed in 10,000 randomized datasets provided an estimate for the probability of observing that rank-sum value by chance alone. Thus, the fraction of occurrences in the randomized data of rank-sum values equal to or less than that observed rank sum in the actual data provided a P value that the observed rank sum was statistically significant. Each IMC was randomized, including IMCs derived from the same subject, because they were genetically and phenotypically different. The independence of the IMCs was supported by subsequent general linearized model (GLM) analyses (see below). The code for this permutation test was written in-house using perl.

The GLM consisted of fitting a log-normal generalized mixed-effect model to the data; all analyses related to the model were performed using R (http://www.r-project.org/). The advantage of using this model was that it allowed us to estimate the impact of and assess the relationships between different parameters. Using this approach, we modeled the response variables in each experiment vs. subtype (B or C) and TF or CC status, considering the cell donor and the IMC as random effects to account for the different levels of donor and IMC responses. We compared nested models, simplifying from more complex models to simpler ones (with less variables) in a stepwise manner using the Akaike Information Criterion, and we only kept variables that reached a significance level of 0.05 or lower. This approach allowed us to test for possible interactions between variables (for example, when the response variable had a different TF vs. CC effect between clades). To test whether the various IMCs were independent when derived from the same subject, we also introduced a variable patient. The patient variable was never statistically supported in our models, suggesting the IMCs were behaving independently, even when derived from the same individual. The GLM was implemented using the glmer function (http://lme4.r-forge.r-project.org/), and the ANOVA function (13) was used to compare models using the Akaike Information Criterion. We used a Shapiro–Wilcoxon test (14) to show that the data were consistent with a log-normal distribution.

### Table S1. Generation of IMCs of TF and CC viruses

<table>
<thead>
<tr>
<th>Subject</th>
<th>IMC</th>
<th>Subtype</th>
<th>Infection status</th>
<th>Sex</th>
<th>Risk factor</th>
<th>CD4 count</th>
<th>VL</th>
<th>Sample date</th>
<th>Stage</th>
<th>Country of origin</th>
<th>No. TF variants</th>
<th>Coreceptor tropism*</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>AD17</td>
<td>AD17</td>
<td>B</td>
<td>TF</td>
<td>M</td>
<td>MSM</td>
<td>n/a</td>
<td>47,600,000</td>
<td>6/14/99</td>
<td>II†</td>
<td>United States</td>
<td>1</td>
<td>R5</td>
<td>2</td>
</tr>
<tr>
<td>700010040</td>
<td>CH040</td>
<td>B</td>
<td>TF</td>
<td>M</td>
<td>MSM</td>
<td>n/a</td>
<td>2,197,248</td>
<td>7/25/06</td>
<td>II†</td>
<td>United States</td>
<td>1</td>
<td>R5</td>
<td>3</td>
</tr>
<tr>
<td>700010106</td>
<td>CH106</td>
<td>B</td>
<td>TF</td>
<td>M</td>
<td>MSMW</td>
<td>n/a</td>
<td>84,545,454</td>
<td>10/19/06</td>
<td>II†</td>
<td>United States</td>
<td>1</td>
<td>R5</td>
<td>3</td>
</tr>
<tr>
<td>700010470</td>
<td>CH470</td>
<td>B</td>
<td>TF</td>
<td>M</td>
<td>MSM</td>
<td>n/a</td>
<td>84,193</td>
<td>8/1/00</td>
<td>V†</td>
<td>United States</td>
<td>1</td>
<td>R5</td>
<td>3</td>
</tr>
<tr>
<td>700010607</td>
<td>CH607</td>
<td>B</td>
<td>TF</td>
<td>M</td>
<td>MSM</td>
<td>n/a</td>
<td>8,009</td>
<td>7/8/08</td>
<td>IV†</td>
<td>United States</td>
<td>2</td>
<td>R5</td>
<td>This study</td>
</tr>
<tr>
<td>REJO4541</td>
<td>REJO</td>
<td>B</td>
<td>TF</td>
<td>M</td>
<td>HSX</td>
<td>n/a</td>
<td>722,349</td>
<td>9/28/01</td>
<td>V†</td>
<td>United States</td>
<td>1</td>
<td>R5</td>
<td>3</td>
</tr>
<tr>
<td>RHPA4256</td>
<td>RHPA</td>
<td>B</td>
<td>TF</td>
<td>F</td>
<td>HSX</td>
<td>n/a</td>
<td>1,458,354</td>
<td>12/5/00</td>
<td>V†</td>
<td>United States</td>
<td>1</td>
<td>R5</td>
<td>3</td>
</tr>
<tr>
<td>SUMA0874</td>
<td>SUMA</td>
<td>B</td>
<td>TF</td>
<td>M</td>
<td>MSM</td>
<td>n/a</td>
<td>939,260</td>
<td>5/13/91</td>
<td>II†</td>
<td>United States</td>
<td>1</td>
<td>R5</td>
<td>3</td>
</tr>
<tr>
<td>THRO4156</td>
<td>THRO</td>
<td>B</td>
<td>TF</td>
<td>M</td>
<td>MSMW</td>
<td>n/a</td>
<td>5,413,140</td>
<td>8/1/00</td>
<td>V†</td>
<td>United States</td>
<td>1</td>
<td>R5</td>
<td>3</td>
</tr>
<tr>
<td>700010077</td>
<td>CH077</td>
<td>B</td>
<td>TF</td>
<td>M</td>
<td>MSM</td>
<td>n/a</td>
<td>394,649</td>
<td>8/31/06</td>
<td>II/III</td>
<td>United States</td>
<td>1</td>
<td>R5/X4</td>
<td>3</td>
</tr>
<tr>
<td>700010131</td>
<td>CH131</td>
<td>C</td>
<td>TF</td>
<td>M</td>
<td>MSMW</td>
<td>n/a</td>
<td>411,873</td>
<td>3/8/07</td>
<td>I/II</td>
<td>Malawi</td>
<td>1</td>
<td>R5</td>
<td>This study</td>
</tr>
<tr>
<td>700010162</td>
<td>CH162</td>
<td>C</td>
<td>TF</td>
<td>M</td>
<td>MSMW</td>
<td>n/a</td>
<td>16,758</td>
<td>1/14/03</td>
<td>I/II</td>
<td>Malawi</td>
<td>1</td>
<td>R5</td>
<td>This study</td>
</tr>
<tr>
<td>700010185</td>
<td>CH185</td>
<td>C</td>
<td>TF</td>
<td>M</td>
<td>HSX</td>
<td>n/a</td>
<td>13,100,000</td>
<td>6/22/07</td>
<td>III†</td>
<td>South Africa</td>
<td>1</td>
<td>R5</td>
<td>This study</td>
</tr>
<tr>
<td>700010215</td>
<td>CH215</td>
<td>C</td>
<td>TF</td>
<td>F</td>
<td>HSX</td>
<td>n/a</td>
<td>216</td>
<td>7/13/07</td>
<td>I/II</td>
<td>Malawi</td>
<td>1</td>
<td>R5</td>
<td>This study</td>
</tr>
<tr>
<td>700010215</td>
<td>CH215</td>
<td>C</td>
<td>TF</td>
<td>M</td>
<td>MSMW</td>
<td>n/a</td>
<td>165,501</td>
<td>5/11/07</td>
<td>III†</td>
<td>Malawi</td>
<td>3</td>
<td>R5</td>
<td>This study</td>
</tr>
<tr>
<td>700010228</td>
<td>CH228</td>
<td>C</td>
<td>TF</td>
<td>M</td>
<td>HSX</td>
<td>n/a</td>
<td>165,501</td>
<td>5/12/07</td>
<td>III†</td>
<td>Malawi</td>
<td>3</td>
<td>R5</td>
<td>This study</td>
</tr>
<tr>
<td>700010293</td>
<td>CH293</td>
<td>C</td>
<td>TF</td>
<td>M</td>
<td>HSX</td>
<td>n/a</td>
<td>165,501</td>
<td>4/1/08</td>
<td>n/a</td>
<td>Malawi</td>
<td>n/a</td>
<td>R5</td>
<td>This study</td>
</tr>
<tr>
<td>700010432</td>
<td>CH432</td>
<td>C</td>
<td>TF</td>
<td>F</td>
<td>HSX</td>
<td>n/a</td>
<td>47,549</td>
<td>6/1/08</td>
<td>n/a</td>
<td>Malawi</td>
<td>n/a</td>
<td>R5</td>
<td>This study</td>
</tr>
<tr>
<td>700010440</td>
<td>CH440</td>
<td>C</td>
<td>TF</td>
<td>F</td>
<td>HSX</td>
<td>n/a</td>
<td>30,434</td>
<td>4/9/08</td>
<td>n/a</td>
<td>Malawi</td>
<td>n/a</td>
<td>R5</td>
<td>This study</td>
</tr>
<tr>
<td>700010534</td>
<td>CH534</td>
<td>C</td>
<td>TF</td>
<td>F</td>
<td>HSX</td>
<td>n/a</td>
<td>63,300</td>
<td>8/2/08</td>
<td>n/a</td>
<td>Malawi</td>
<td>n/a</td>
<td>R5</td>
<td>This study</td>
</tr>
<tr>
<td>700010141</td>
<td>CH141</td>
<td>C</td>
<td>TF</td>
<td>F</td>
<td>HSX</td>
<td>326§</td>
<td>151,282</td>
<td>11/13/07</td>
<td>n/a</td>
<td>Malawi</td>
<td>n/a</td>
<td>R5</td>
<td>This study</td>
</tr>
<tr>
<td>700010167</td>
<td>CH167</td>
<td>C</td>
<td>TF</td>
<td>F</td>
<td>HSX</td>
<td>358§</td>
<td>73,505</td>
<td>5/24/07</td>
<td>n/a</td>
<td>Malawi</td>
<td>n/a</td>
<td>R5</td>
<td>This study</td>
</tr>
<tr>
<td>700010256</td>
<td>CH256</td>
<td>C</td>
<td>TF</td>
<td>F</td>
<td>HSX</td>
<td>531§</td>
<td>28,066</td>
<td>6/30/09</td>
<td>2‡</td>
<td>Malawi</td>
<td>n/a</td>
<td>R5</td>
<td>This study</td>
</tr>
<tr>
<td>700010475</td>
<td>CH475</td>
<td>C</td>
<td>TF</td>
<td>F</td>
<td>HSX</td>
<td>450§</td>
<td>234,671</td>
<td>6/1/08</td>
<td>n/a</td>
<td>Tanzania</td>
<td>n/a</td>
<td>R5</td>
<td>This study</td>
</tr>
</tbody>
</table>

F, female; HSX, heterosexual; M, male; MSM, men who have sex with men; MSMW, men who have sex with men and women; n/a, data not available; R5, chemokine receptor 5 (CCR5) tropic; R5/X4, CCR5/CXCR4 dual tropic; VL, viral load (RNA copies per milliliter of plasma).

*Coreceptor tropism as determined by AMD3100 and TAK779 inhibition.

†Defined by Fiebig et al. (1).

‡Duration of infection in years; n/a, data not available.

§Average CD4 count (cells per millimeter cubed of blood) from samples taken before and after IMC generation.

{A single CD4 count available for a sample collected 12 wk before IMC generation.
Table S2. GenBank accession numbers of SGA-derived sequences

<table>
<thead>
<tr>
<th>Subject</th>
<th>IMC</th>
<th>Subtype</th>
<th>Infection status</th>
<th>5' half-genome sequences</th>
<th>3' half-genome sequences</th>
</tr>
</thead>
<tbody>
<tr>
<td>700010470</td>
<td>CH470</td>
<td>B</td>
<td>TF</td>
<td>JX973388-JX973412</td>
<td>JX972238-JX972249</td>
</tr>
<tr>
<td>700010607</td>
<td>CH607</td>
<td>B</td>
<td>TF</td>
<td>JX973482-JX973502</td>
<td>JX972390-JX972410</td>
</tr>
<tr>
<td>704010042</td>
<td>CH042</td>
<td>C</td>
<td>TF</td>
<td>JX973846-JX973854</td>
<td>JX972838-JX972845</td>
</tr>
<tr>
<td>705010067</td>
<td>CH067</td>
<td>C</td>
<td>TF</td>
<td>JX973930-JX973938</td>
<td>JX972931-JX972964</td>
</tr>
<tr>
<td>703010131</td>
<td>CH131</td>
<td>C</td>
<td>TF</td>
<td>JX973549-JX973555</td>
<td>JX972455-JX972487</td>
</tr>
<tr>
<td>705010162</td>
<td>CH162</td>
<td>C</td>
<td>TF</td>
<td>JX973955-JX973969</td>
<td>JX972986-JX972998</td>
</tr>
<tr>
<td>706010164</td>
<td>CH164</td>
<td>C</td>
<td>TF</td>
<td>JX974138-JX974164</td>
<td>JX973234-JX973270</td>
</tr>
<tr>
<td>703010228</td>
<td>CH228v1</td>
<td>C</td>
<td>TF</td>
<td>KC149442-KC149450</td>
<td>KC149275-KC149299</td>
</tr>
<tr>
<td>705010185</td>
<td>CH185</td>
<td>C</td>
<td>TF</td>
<td>JX974014-JX974023</td>
<td>JX973075-JX973101</td>
</tr>
<tr>
<td>705010198</td>
<td>CH198</td>
<td>C</td>
<td>TF</td>
<td>JX974082-JX974094</td>
<td>JX973171-JX973180</td>
</tr>
<tr>
<td>703010200</td>
<td>CH200v1</td>
<td>C</td>
<td>TF</td>
<td>KC149260-KC149274</td>
<td>KC149138-KC149185</td>
</tr>
<tr>
<td>STC05453</td>
<td>STCo2</td>
<td>B</td>
<td>CC</td>
<td>KC312399-KC312433</td>
<td>KC312434-KC312466</td>
</tr>
<tr>
<td>STCOr1</td>
<td>STCOr2</td>
<td>B</td>
<td>CC</td>
<td>KC312330-KC312366</td>
<td>KC312367-KC312398</td>
</tr>
<tr>
<td>WARO5662</td>
<td>WARO</td>
<td>B</td>
<td>CC</td>
<td>KC312356-KC312357</td>
<td>KC312572-KC312600</td>
</tr>
<tr>
<td>MCST4474</td>
<td>MCST</td>
<td>B</td>
<td>CC</td>
<td>KC312510-KC312539</td>
<td>KC312467-KC312509</td>
</tr>
<tr>
<td>RHGA1581</td>
<td>RHGA</td>
<td>B</td>
<td>CC</td>
<td>KC156252-KC156268</td>
<td>KC156222-KC156251</td>
</tr>
<tr>
<td>703010269</td>
<td>CH269</td>
<td>C</td>
<td>CC</td>
<td>KC156344-KC156371</td>
<td>KC156318-KC156343</td>
</tr>
<tr>
<td>702010293</td>
<td>CH293</td>
<td>C</td>
<td>CC</td>
<td>KC156297-KC156317</td>
<td>KC156269-KC156296</td>
</tr>
<tr>
<td>702010440</td>
<td>CH440</td>
<td>C</td>
<td>CC</td>
<td>KC156472-KC156499</td>
<td>KC156431-KC156471</td>
</tr>
<tr>
<td>703010167</td>
<td>CH167</td>
<td>C</td>
<td>CC</td>
<td>KC156402-KC156430</td>
<td>KC156372-KC156401</td>
</tr>
</tbody>
</table>

Fig. 51. Inference of TF viral sequences from acutely infected individuals. Highlighter plots of SGA-derived 5' (Left) and 3' (Right) half-genome sequences are shown for subjects infected with subtype B (A–D) and subtype C (E–V) viruses, respectively (Highlighter v2.1.1; hiv.lanl.gov); tick marks indicate differences compared with the sequence on the top, which represents the inferred TF sequence (red, T; green, A; blue, C; orange, G; gray, indel; A and B, CH607; C and D, CH470; E and F, CH200v1; G and H, CH228v1; I and J, CH131; K and L, CH164; M and N, CH185; O and P, CH198; Q and R, CH042; S and T, CH067; U and V, CH162). Two TF variants were inferred from subject 703010200, with the second variant shown in bold below the sequences used to infer the first variant.

Fig. 52. Generation of IMCs of CC viruses. The phylogenetic relationships of SGA-derived 5' and 3' half-genome sequences from chronically infected subjects are shown. Trees composed of S' (A, C, E, G, I, K, M, O, and Q) and 3' (B, D, F, H, J, L, N, P, and R) half-genome sequences were constructed using maximum likelihood methods. Sequences highlighted in blue (subtype C) and red (subtype B) represent clonal viral expansions that were used to generate half-genome consensus sequences (A and B, 702010141; C and D, 703010167; E and F, 703010269; H and G, 702010293; I and J, 702010440; K and L, MSCT4474; M and N, RHGA1581; O and P, STCOS453; Q and R, WAROS662). These consensus sequences, which were identical in a 1-kb region of overlap, were chemically synthesized and used to construct full-length IMCs. Asterisks indicate bootstrap support of greater than 70%. (Scale bar: 0.001 or 0.01 substitutions per site as indicated.)

Fig. 53. Quantiﬁcation of viral stocks. For each virus stock, p24 antigen content (measured in picograms), RT activity (measured in picograms), and viral RNA (vRNA) copy number were determined, and their relationships were compared for subtypes B (solid circles) and C (open circles) viruses, respectively. (A–C) p24 antigen content was measured by the PerkinElmer AlphaLISA (averaged from four experiments). vRNA copies were measured by the Roche COBAS AmpliPrep/COBAS Taqman HIV-1 v. 2.0 Test at different viral dilutions. RT activity was measured using the colorimetric Roche Reverse Transcriptase Assay (averaged from three experiments). Values for subtypes B and C viruses were compared using a Mann–Whitney test (P values are indicated). (D) The relationship between vRNA copy number (x axis) and RT activity (y axis) is shown for each virus stock, with Spearman correlation coefficient and P value indicated; (E) p24 antigen content was measured by the PerkinElmer Alliance ELISA (averaged from two experiments), and (F) p24 antigen content was measured using a method developed by Biancotto et al. (1) (averaged from two experiments).


Fig. 54. Virus capture by moDCs. (A) The percent virion captured for each virus strain (y axis) by cells from a given donor (indicated by different symbols) is shown in relation to the average percent virion captured of that virus strain by cells from all donors (x axis). Virion capture by cells from donors A and F were the least well-correlated (Spearman correlation coefficient R_s = 0.72 and 0.40, respectively; other donors R_s > 0.8). (B) The mean ﬂuorescent intensity of DC-specific intracellular adhesion molecule 3-grabbing nonintegrin (DC-SIGN) and macrophage mannose receptor (MMR) staining (yellow and green bars, respectively) is shown for different donors (A–F) and virus input (RT activity or p24 content). Cells from donors A and F expressed less DC-SIGN and MMR than the other moDC donors. (C) Flow cytometry images of representative moDCs expressing DC-SIGN (yellow) and MMR (green) for each donor. (D) The percent moDC capture of a given virus (y axis) is shown in relation to its Env content (x axis). This relationship was signiﬁcant for cells from donor B, with a trend observed for the other donors (R_s and P values are shown for each donor). Subtype B TF and CC viruses are shown in dark and light red, respectively; subtype C TF and CC viruses are shown in dark and light blue, respectively. Symbols indicate different moDC donors.

Fig. 55. Virus replication in CD4+ T cells in the absence and presence of IFN-α. (A and B) Activated CD4+ T cells from three healthy donors (rows) were infected with an equivalent amount of virus and cultured in the absence (Left) and presence (Right) of 500 U IFN-α. Virus replication was monitored by RT activity in culture supernatants (expressed as picograms per milliliter) at four different time points (days) postinfection (x axis). Values are shown for each virus. Subtypes B and C viruses are shown in red and blue, respectively, with TF viruses indicated in dark colors and CC viruses indicated in light colors, respectively. Values indicate average RT activity from duplicate measurements. (A) Open circles denote IMCs from female patients, and filled circles denote IMCs from male patients. (B) Open circles denote IMCs from heterosexual transmissions (HSX), and filled circles denote IMCs from homosexual transmissions (MSM). The broken line indicates the limits of virus detection by the RT assay. Circles beneath the broken line indicate the absence of detectable virus. There was no significant interaction between subtype and sex (P = 0.162 by GLM); sex was not a significant fixed effect when included as an independent variable (P = 0.8) or when subtype was excluded from the model (P = 0.4). All MSM samples were from the B clade, and therefore, subtype-risk interactions could not be explored.

Fig. 56. IFN-α resistance across donors and time points. The ratio of virus production in the presence and absence of IFN-α for each virus strain (y axis) is shown in relation to the average ratio of virus production in the presence and absence of IFN-α for the same virus strain (x axis). Although there was donor-to-donor variability in the magnitude of the IFN-α effect, the observed levels of sensitivity or resistance were consistent across all donors and time points (R values ranged from 0.75 to 0.90; P < 0.0001 for all donors and time points).

Parrish et al. www.pnas.org/cgi/content/short/1304288110
Fig. S7. Cell-to-cell virus spread in the presence of IFN-α. (A) CD4+ T cells from two donors (donors A and B as in Fig. 4) were infected with TF (filled bars) and CC (open bars) viruses in the presence of 500 U IFN-α and analyzed for Gag-positive cells by flow cytometry 12 d postinfection. Bars indicate the median percent of Gag-positive cells (y axis), with interquartile ranges indicated. TF viruses infected a higher percentage of cells by day 12 than CC viruses (P = 0.049). (B and C) The ratio of virus production in the presence and absence of IFN-α is shown for each virus at day 12 (x axis) in relation to the percent of Gag-positive cells infected by this virus (y axis). Subtypes B and C viruses are shown in red and blue, respectively, with TF viruses indicated in dark colors and CC viruses indicated in light colors, respectively. Viruses that were resistant to IFN-α infected a higher percentage of cells than viruses that were sensitive to IFN-α (R_β = 0.80 and 0.63 for B and C, respectively; both P < 0.0001).
**5’ Half Genome**

A

**3’ Half Genome**

B

C

D

Figure S1
Figure S1
Figure S1
Figure S1
Figure S2
Figure S2
Figure S3

A. pg p24 per pg RT

B. pg p24 per vRNA

C. vRNA per pg RT

D. pg RT per vRNA copies per ml

E. pg p24 per pg RT

F. pg p24 per copy vRNA

Figure S3
Figure S4
Figure S5
Figure S6
Figure S7

A

% Gag+ cells

B

% Gag+ cells

C

% Gag+ cells

+ IFN / - IFN

+ IFN / - IFN

Figure S7